Architecture and the Built environment

#21
2016

From access to re-use: a user's perspective on public sector information availability

Frederika Welle Donker

From access to re-use: a user's perspective on public sector information availability

Frederika Welle Donker
Delft University of Technology, Faculty of Architecture and the Built Environment, Department of OTB - Research for the Built Environment

Architecture and the Built environment

abe.tudelft.nl

Design: Sirene Ontwerpers, Rotterdam

ISBN 978-94-92516-27-5
ISSN 2212-3202

From access to re-use: a user's perspective on public sector information availability

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 6 december 2016 om 15:00 uur
door Frederika Meinarda WELLE DONKER
Master of Science in Systems Engineering and Policy Analysis
geboren te Rotterdam

Dit proefschrift is goedgekeurd door de

promotor: Prof. dr. W.K. Korthals Altes
copromotor: Dr. ir. B. van Loenen

Samenstelling promotiecommissie bestaat uit

Rector Magnificus,	voorzitter
Prof. dr. W.K. Korthals Altes,	promotor
Dr. ir. B. van Loenen,	copromotor

Onafhankelijke leden

Prof. dr. ir. M.F.W.H.A. Janssen, Faculty of Technology, Policy and Management, Delft University of Technology
Prof. dr. ir. P.J.M. van Oosterom, Faculty of Architecture and the Built Environment, Delft University of Technology
Prof. dr. ir. J. Crompvoets, Public Governance Institute, University of Leuven
Prof. C. de Terwagne, Faculty of Law, University of Namur
Prof. mr. dr. ir. J.A. Zevenbergen, Faculty of Geo-Information Science and Earth Observation, University of Twente

"There's no point in acting surprised about it. All the planning charts and demolition orders have been on display at your local planning department in Alpha Centauri for 50 of your Earth years, so you've had plenty of time to lodge any formal complaint and it's far too late to start making a fuss about it now. ... What do you mean you've never been to Alpha Centauri? Oh, for heaven's sake, mankind, it's only four light years away, you know. I'm sorry, but if you can't be bothered to take an interest in local affairs, that's your own lookout. Energize the demolition beams."

Douglas Adams (1979). The Hitchhiker's Guide to the Galaxy (p.26).

Preface and acknowledgements

Life is what happens when you're busy making other plans. This dissertation, somehow, illustrates that. Life is what happened to me and always seemed to lead me to into a different direction. When I attended high school, I always intended to go to university. However, by the time I left school, a university degree was no longer an option. Instead, I enrolled in a teachers training college because I wanted to teach. There, I learned one important lesson: never combine a full-time study with a part-time cleaning job. After that failed experiment, I went to Sydney, Australia, for an extended holiday. I fell in love with the country, and one of its citizens, so I decided to stay and enrol in an electronics engineering course at a TAFE College. The only place where I could get a traineeship was the National Building Technology Centre at North Ryde. The NBTC was a small, fairly run-down institute, in the middle of a wild park reserve. More than once did we have to remove snakes from our furniture before we could start work in the field laboratories. On my first day, I met an older man walking out of one of the most desolate field laboratories. He wore big glasses, had tufts of long grey hair standing upright and he muttered some incantations (or maybe formulae). I felt at home straight away and decided at that moment I wanted to be researcher as well. After I finished my traineeship, I worked as a (Senior) Technical Officer at Sydney University. During those 10 years, I enjoyed working with students, assisting during practical classes. However, as a non-academic, I could never become a lecturer nor a researcher, alas.

When I re-migrated to the Netherlands, I was told that my working permit application would take about three months. In the meantime, I was not allowed to work, a strange experience for someone who had worked all her life. A friend of mine had just enrolled in the part time degree course at the Faculty of Technology, Policy and Management of Delft University of Technology. Although I had not considered a university degree for a long time, I decided to give it a go for a couple of months, maybe a year. My friend dropped out after one year, I graduated three years later. Life is what happens. ...

After a three year stint as a research assistant at Erasmus MC in Rotterdam, I knew for certain I wanted to be a researcher. So, when a Ph.D. position at the OTB was advertised, I jumped at the chance to apply. I did not get the job but I did get a one-year contract to carry out a part of a longer research project. After one year, the contract was extended, and then more projects and contracts followed. In the end, this dissertation has taken nearly ten years to complete and during that time, many contract research projects were carried out.

Writing a dissertation based on peer-reviewed articles resulting from a variety of contract research projects has advantages and disadvantages. One has to research

many aspects, including some that may not have been originally considered. This may be an enrichment to the dissertation or a distraction and yet another delay in the writing process. It was enriching, however, to stay in touch with the "real world". Writing and submitting peer-reviewed articles as you go, is one way of keeping the research sharp. It is also a time-consuming exercise that can be distracting when the next project has already started. It also means that some articles are outdated by the time they are published as a dissertation chapter. If I could have started all over again, would I have done it differently? Probably not. The fact that the dissertation took so long, also had an advantage. During the course of this research, the data policy goal posts moved towards Open Data. This meant that there was an excellent opportunity to test some hypotheses. How many policy researchers get such a chance? Life is what happens ...

No dissertation is complete without thanking your colleagues, project partners, family and friends. In my case, there are too many to name each one. I will name a few though. First of all, I would like to thank Bastiaan van Loenen, who has dragged me through the hard times when I lost faith and focus. We have cooperated in many of the research projects and we have written many publications complementing each other. Next, I would like to thank Prof. Korthals Altes who took over as my promotor after my first promotor, Prof. de Jong retired. Jitske de Jong had helped me enormously in the early stages of my Ph.D. research and we had already co-written a number of publications. Willem Korthals Altes took over during a period that my research had slowed down, and he enabled me to regain focus. I would also like to thank the two roommates I have had over the years. Firstly, Henk Koerten, who started his Ph.D. (yes, the one I applied for) at the same time I started my time at the OTB. We participated in the same project and he always managed to put complex issues into perspective with his typical Frisian down-to-earthness. Secondly, I would like to thank my current roommate Hendrik Ploeger, whose views on 20th Century dictatorships are refreshing, to say the least. These views, somehow, manage to put the concept of Open Data into a totally different perspective. I would also like to thank Dirk Dubbeling, who has been a great help with the publication of this dissertation, and who I must have driven mad with yet another change... Finally, I would like to thank all the other colleagues of OTB, with whom I have had long lunches and interesting discussions.

I would also like to thank the many project partners I have worked with over the last 11 years. Some of them I have known for years as we have participated in more than one project; others I have known only for a short time as they were more short-lived project partners. They all have added to my knowledge base of data access policies and have helped me keep touch with the world outside the University.

Of course, I have to thank my partner and my family who had to put up with me working long hours, and not being very sociable, especially in the last 12 months. I faithfully promise I will do my fair share of household chores again after my defence.

Finally, I would like to thank my bicycle mates who have participated in the annual OTB Cycle Tours. They have had to put up with my "little surprises" along the routes I had planned. However, what do you expect when you let a non-geodesist in charge of navigating … On the positive side, we have seen parts of Great Britain, Belgium, France and Germany not often seen by other cyclists. Life is indeed what happens when you're making other plans.

Contents

Summary

If data are the building blocks to generate information needed to acquire knowledge and understanding, then geodata, *i.e.* data with a geographic component (geodata), are the building blocks for information vital for decision-making at all levels of government, for companies and for citizens. Governments collect geodata and create, develop and use geo-information - also referred to as spatial information - to carry out public tasks as almost all decision-making involves a geographic component, such as a location or demographic information. Geo-information is often considered "special" for technical, economic reasons and legal reasons. Geo-information is considered special for technical reasons because geo-information is multi-dimensional, voluminous and often dynamic, and can be represented at multiple scales. Because of this complexity, geodata require specialised hardware, software, analysis tools and skills to collect, to process into information and to use geo-information for analyses.

Geo-information is considered special for economic reasons because of the economic aspects, which sets it apart from other products. The fixed production costs to create geo-information are high, especially for large-scale geo-information, such as topographic data, whereas the variable costs of reproduction are low which do not increase with the number of copies produced. In addition, there are substantial sunk costs, which cannot be recovered from the market. As such, geo-information shows characteristics of a public good, *i.e.* a good that is non-rivalrous and non-excludable. However, to protect the high investments costs, re-use of geo-information may be limited by legal and/or technological means such as intellectual property rights and digital rights management. Thus, by making geo-information excludable, it becomes a club good, *i.e.* a non-rivalrous but excludable good. By claiming intellectual property rights, such as copyright and/or database rights, and restricting (re-)use through licences and licence fees, geo-information can be commercially exploited and used to recover some of the investment costs.

Geo-information is considered special for a number of legal reasons. First, as geo-information has a geographic component, *e.g.* a reference to a location, geo-information may contain personal data, sensitive company data, environmentally sensitive data, or data that may pose a threat to the national security. Therefore, the dataset may have to be adapted, aggregated or anonymised before it can be made public. Secondly, geo-information may be subject to intellectual property rights. There may be a copyright on cartographic images or database rights on digital information. Such intellectual property rights may be claimed by third parties involved in the information chain, *e.g.* a private company supplying aerial photography to the National Mapping Authority. The data holder may also claim intellectual property rights to

commercially exploit the dataset and recoup some of the vast investment costs made to produce the dataset. Lastly, there may be other (international) legislation or agreements that may either impede or promote publishing public sector information, whereby in some cases, these policies may contradict each other.

It has been recognised that to deal with national, regional and global challenges, it is essential that geo-information collected by one level of government or government organisation be shared between all levels of government via a so-called Spatial Data Infrastructure (SDI). The main principles governing SDIs are that data are collected once and (re-)used many times; that data should be easy to discover, access and use; and that data are harmonised so that it is possible to combine spatial data from different sources seamlessly. In line with the SDI governing principles, this dissertation considers accessibility of information to include all these aspects. Accessibility concerns not only access to data, *i.e.* to be able to view the data without being able to alter the contents but also re-use of data, *i.e.* to be able to download and/or invoke the data and to share data, including to be able to provide feedback and/or to provide input for co-generated information.

Accessibility to public sector geo-information is not only essential for effective and efficient government policy-making but is also associated with realising other ambitions. Examples of these ambitions are a more transparent and accountable government, more citizens' participation in democratic processes, (co-)generation of solutions to societal problems, and to increase economic value due to companies creating innovative products and services with public sector information as a resource. Especially the latter ambition has been the subject of many international publications stressing the enormous potential economic value of re-use of public sector (geo-) information by companies. Previous research indicated that re-users of public sector information in Europe encountered barriers related to technical, organisational, legal and financial aspects, which was deemed to be the main reason why in Europe the number of value added products and services based on public service information were lagging compared to the United States. Especially the latter two barriers (restrictive licence conditions and high licence fees) were often cited to be the main barriers for re-users in Europe. However, in spite of considerable resources invested by governments to establish spatial data infrastructures, to facilitate data portals and to release public sector information as open data, *i.e.* without legal and financial restrictions, the expected surge of value added products based on public sector information has not quite eventuated to date and the expected benefits still appear to lag expectations.

When this research started a decade ago, the debate around accessibility of public sector information focussed on access policies. Access policies ranged from open access (data available with a minimum of legal restrictions and for no more than marginal dissemination costs) to full cost recovery, whereby all costs incurred in collection, creation, processing, maintenance and dissemination costs to be recovered

from the re-users. Most of the public sector bodies in the European Union adhered to a cost recovery policy for allowing re-use of public sector information. In 2003, the European Commission adopted two directives to ensure better accessibility of public sector information Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC, the so-called Access Directive, provided citizens the right of access to environmental information. Citizens should be able to access documents related to the environment via a register, preferably in an electronic form and if a copy of a document was requested, the charges must not exceed marginal dissemination costs. Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, the co-called PSI Directive, intended to create conditions for a level playing field for all re-users of public sector information. However, the PSI Directive of 2003 left room for public sector organisations to maintain a cost recovery regime with restrictive licence conditions. In spite of these directives, access policies for geographic data were slow to change in most European nations.

At the end of the last decade, accessibility of public sector information received two major impulses. The first major impulse was the implementation of Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), the co-called INSPIRE Directive, established a framework of standardisation rules for the data and publishing via web services, which significantly contributed to the accessibility of public sector geo-information. The second major impulse was the development of open data policies following the Digital Agenda for Europe adopted in 2010 and the USA Open Government Directive of 2009 and the Digital Agenda for Europe of 2010. These two impulses were the main drivers in Europe to start a careful move from cost recovery policies to open access or open data policies and for more public sector information to be made available as open data. Thus, of the four barriers to re-use of public sector information data cited in Chapter 1 (legal, financial, technical and organisational barriers), two barriers should have been lifted to a large degree due to open data. This shift to open data provided an excellent opportunity to test the hypothesis that the main barriers for re-users of public sector information were indeed restrictive licences and high fees as suggested by earlier research.

Chapter 2 showed that by 2008, most European Union Member States had transposed and implemented the 2003/98/EC PSI Directive, however, in various ways and with considerable delay. By 2008, the effects of the PSI Directive were only slowly starting to emerge. A number of Member States reviewed their access policies and more public sector information became available for re-use. Some Member States made the information available free-of-charge or reduced their fees significantly. In many cases, where re-use fees were reduced the number of regular re-users increased significantly and total revenue even increased in spite of lower fees. Although the 2007/2/EC

INSPIRE Directive paved the way for technical interoperability by providing guidelines for web services and catalogues, neither the INSPIRE Directive nor the PSI Directive had tackled the issue of legal interoperability. Chapter 2 also demonstrated that a major barrier to creating a level playing field for the private sector was the fact that some public sector bodies acted as value added resellers by developing and selling products and services based on their own data. Thus, the level playing field envisioned by the European Commission had not been realised.

Chapter 3 researched the aspect of harmonised licences as a first step towards legal interoperability. Earlier research had indicated that one of the biggest barriers for re-users were complex, intransparent and inconsistent licence conditions, especially for re-users wanting to combine data from multiple sources. A survey of licences used by public sector data providers in the Netherlands demonstrated that although there were differences in length and language, there were also many similarities. The conclusion was that the introduction of a licence suite inspired by the Creative Commons concept would be a step towards increased transparency and consistency of geo-information license agreements. This chapter introduced a conceptual model for such a geo-information licence suite, the so-called Geo Shared licences. Both Creative Commons and Geo Shared licence suites enable harmonisation of licence conditions and promote transparency and legal interoperability, especially when re-users combine data from different sources. The Geo Shared licence suite became a serious option for inclusion into the draft version of the INSPIRE Directive as an annex. Unfortunately, the concept of one licence suite for the entire European Union came too early in 2006. The Geo Shared licences were further developed and implemented into the Dutch National Geo Register.

In 2009, the European Commission recognised that PSI was the single largest source of information in Europe and the potential for re-use of PSI needed to be highlighted in the digital age. As part of a review of the 2003/98/EC PSI Directive, the European Commission carried out a round of consultations with stakeholders to seek their views on specific issues to be addressed in the future in 2010. In addition, the Commission commissioned a number of studies. These studies included a review of studies on public sector information re-use and related market studies, an assessment of the different models of supply and charging for public sector information and a study on public sector re-user in the cultural sector. The first study, carried out by Graham Vickery in 2011, showed that the overall economic gain from opening up public sector information as a resource for new products and services could be in the order of €40 billion per annum in the European Union. Both the Vickery Report and the second study, the so-called POPSIS Study, showed that for most public sector data providers their revenues from licence fees were relatively low in comparison to their total budget. After the evaluation, Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information was adopted and came into force on 17 July 2013. Chapter 4

described the main changes of the 2013/37/EU Amended PSI Directive, including the recommendation to employ open data licences. This chapter continued with a review of the various open data licences in use in Europe and analysed their interoperability. Although adoption of open data licences for public sector information should have addressed legal interoperability barriers for re-users, in practice, the different types of open data licences might not be so interoperable after all. Effectively, only a public domain declaration, such as a Creative Commons Zero (CC0) declaration, is suitable for open data re-users requiring with cross-border data sets and that such a public domain declaration is published in a prominent place to remove uncertainty for re-users. Without a public domain declaration, re-use of open data is still impeded as re-users are loathe to invest time into the development of value added products or services when it is uncertain if and which restrictions may be applicable and what the impact may be on their product or service.

This dissertation also researched the financial and economic aspects of public sector information accessibility. Chapters 1 and 2 indicated that a cost recovery regime for dissemination of public sector information provided a financial barrier for private sector re-users because the fees charged were perceived to be too high. However, in 2008, there were still many advocates for maintaining a cost recovery regime. Especially public sector bodies that are not funded by the national Treasury, the so-called self-funding agencies, needed revenue from data sales to cover a substantial part of their operational costs. A sustainable source of revenue was viewed as essential to maintain the data at an adequate level, and to ensure actuality and continuity. Chapter 5 explored the potential business models and pricing mechanisms for public sector INSPIRE web services. Although, depending on the type of web service, and type of re-user, there might have been an argument for employing a subscription model as a pricing mechanism, business models based on generating revenue from public sector information would not be viable in the long run and were not in the spirit of the INSPIRE Directive. This research concluded that public sector information web services employing different pricing regimes were counterproductive to achieving financial interoperability.

In Chapter 6, business models for public sector data providers were revisited, this time from an open data perspective. Government agencies, including self-funding government agencies are under increasing pressure to implement open data policies. This chapter analysed the business models of self-funding agencies either already providing open data or under pressure to provide (some) open data in the near future. The analysis showed which adaptions might be necessary to ensure the long-term availability of high quality open data and the long-term financial sustainability of self-funding agencies. The case studies confirmed that providing (raw) open data does not necessarily lead to losses in revenue in the long term as long as the organisation has enough flexibility to adapt its role in the information value chain, especially when revenue from licence fees represents only a relative small part of their

total budget. The case studies indicated that switching to open data has resulted in internal efficiency gains. In practice, it is difficult to isolate and quantify the internal efficiency gains that are solely attributable to open data as the researched organisations continuously implement efficiency measures. However, the reported decreases in internal and external transaction costs due to open data are in line with the case study carried out in Chapter 7.

Open data also provided an excellent opportunity to assess the effects of open data ex ante as baseline measurements could be carried out. To develop both quantitative and qualitative indicators to assess the success of a policy change is a challenge for open data initiatives. In Chapter 7, a model to assess the effects on the organisation of an open data provider was developed. Liander, a private energy network administrator mandated with a public task, planned to publish some of their datasets as open data in the autumn of 2013. This offered an excellent opportunity to apply the developed assessment model to provide an insight into internal, external, and relational effects on Liander. A benchmark was carried out prior to release of open data and a follow-up measurement one year later. The benchmark provided an insight into the then work processes and into the preparations required to implement open data. The follow-up monitor indicated that Liander open data are used by a wide range of users and have had a positive effect on the development of apps to aid energy savings. However, it remains a challenge to quantify the societal effects of such apps. The follow-up monitor also indicated that regular re-users of Liander data used the open data to improve existing applications and work processes rather than to create new products. The case study demonstrated that private energy companies could successfully release open data. The case study also showed that Liander served as a best-practice case for open data and had a flywheel effect on companies within the same sector. By 2015, nearly all energy network administrators had published similar open data. The monitoring model developed in this project was assessed to be suitable to monitor the open data effects on the organisation of the data provider.

The assessment model developed and tested in Chapter 7 proved to be suitable to monitor the effects of open data on organisational level. However, to provide a more complete picture of the effects of open data and to assess if there are other barriers for re-users, a more holistic approach was required to assess the maturity of open data. Therefore, a holistic open data assessment framework addressing the supplier side, the governance side, and the user side of the open data was developed and applied to the Dutch open data infrastructure in Chapter 8. This Holistic Open Data Maturity Assessment Framework was used to evaluate the State of the Open Data Nation in the Netherlands and to provide valuable information on (potential) bottlenecks. The framework showed that geographic data scored significantly better than other types of government data. The standardisation and implementation rules laid down by INSPIRE Directive framework appear to have been a catalyst for moving geographic data to a higher level of maturity. The maturity assessment framework provided Dutch

policy makers with useful inputs for further development of the open data ecosystem and development of well-founded strategies that will ensure the full potential of open data will be reached. Since the publication of the State of the Open Data Nation in 2014, a number of the recommendations have already been implemented.

This dissertation demonstrated that many aspects that should facilitate accessibility, such as standardised metadata, have already been addressed for geodata. This research also showed that for other types of data, there is still a long way to go. There is a growing demand for other types of data, such as financial data and healthcare data. Public sector organisations holding such types of data need hands-on guidelines to enable publication of their datasets, preferably as open data. However, data published as open data are forever and cannot be recalled. Therefore, the decision to publish public sector data as open data is complex: datasets are often of a heterogeneous nature and may contain microdata (data that quantify observations or facts, such as data collected during surveys) Although microdata may not necessarily contain personal data, the datasets will probably have to be processed before publication to address confidentiality and data quality issues. In addition, there is a tension between open data and protection of personal data. The big question remains to which level the data need to be aggregated and/or anonymised to ensure protection of personal data now and in the future, and at the same time keeping sufficient significance to be re-usable. Another issue that needs further research is data-ownership of sensor data and co-created data. Increasingly, sensor data generated by *e.g.* smart phones, smart energy meters and traffic sensors are collected by the public sector and the private sector and become part of a big data ecosystem. In addition, public sector organisations cooperate with other public sector organisations and the private sector to create information from their data, so-called co-created information. Citizens also collect data or complement information on a voluntary basis, *e.g.* bird counts data. Co-created information will become more commonplace in the coming decades, as will the contribution of sensor data to a big data ecosystem. However, the aspect of who owns the data in which part of the information value chain has not been researched. Uncertainty related to third party rights will pose a barrier to publishing open data. Therefore, the aspect of data-ownership for sensor data and for co-created data should be further researched.

Samenvatting

Als data de bouwstenen zijn om de informatie te genereren die nodig is om kennis en begrip te vergaren, dan zijn geodata, d.w.z. data met een geografische component, de bouwstenen voor de informatie die noodzakelijk is voor besluitvorming bij de overheid, het bedrijfsleven en burgers. Geo-informatie wordt gecreëerd, verzameld, ontwikkeld en gebruikt om publieke taken uit te voeren, aangezien vrijwel alle besluitvorming een ruimtelijk component nodig heeft, zoals een locatie of demografische gegevens.

Geo-informatie of ruimtelijke informatie wordt vaak al "speciaal" beschouwd vanuit een technisch, economisch en juridisch oogpunt. Geo-informatie is vanuit een technisch oogpunt special omdat geo-informatie multidimensionaal is, vaak een groot bestand (terabytes) betreft die een dynamische aard kan hebben (bijvoorbeeld meteorologische informatie), en op meerdere schalen kan worden weergegeven. Vanwege die complexiteit is er gespecialiseerde hardware, software en kennis nodig om geodata in te winnen, te verwerken tot informatie en te gebruiken voor analyses.

Vanuit een economisch oogpunt geldt dat geo-informatie net als andere vormen van informatie, specifieke eigenschappen heeft waardoor het zich van andere producten onderscheidt. Vaak geldt voor informatie dat de vaste productiekosten hoog zijn, net als de verzonken kosten die niet kunnen worden terugverdiend uit de markt. De variabele reproductiekosten zijn laag en nemen niet toe als er extra kopieën worden geproduceerd. Bovendien bestaan er geen natuurlijke capaciteitsgrenzen aan het aantal kopieën. Informatie wordt daarom als een collectief goed beschouwd, d.w.z. het gebruik van een goed gaat niet ten koste van het gebruik door een ander, en het is vaak onmogelijk om mensen uit te sluiten van het gebruik van dat goed ook als die niet betalen voor gebruik (de zogenaamde free riders). Vanwege de hoge investeringskosten voor vooral grootschalige geo-informatie, wordt de meeste geo-informatie door overheden ingewonnen omdat de overheid schaalvoordelen hebben die de private sector niet heeft. Om de hoge investeringskosten van (grootschalige) geo-informatie te beschermen, kunnen er wettelijke en technische middelen worden ingezet om het free rider probleem deels op te vangen.

Vanuit een juridisch oogpunt kan geo-informatie als "speciaal" kan worden beschouwd, voor verschillende redenen. Ten eerste bevat geo-informatie een verwijzing naar een locatie, en kan daardoor vaak gevoelige gegevens bevatten, zoals persoonsgegevens, gevoelige bedrijfsgegevens, gegevens die het milieu kunnen schaden, of gegevens die een bedreiging voor de Staat zouden kunnen vormen. Daarom moet geo-informatie vaak worden aangepast, zoals aggregeren of anonimiseren, voordat de bestanden beschikbaar kunnen worden gesteld. Ten tweede kan geo-informatie intellectuele eigendomsrechten bevatten, zoals auteursrecht

op kaartafbeeldingen of databankrechten. Die intellectuele eigendomsrechten kunnen worden voorbehouden door (private) partijen die data hebben aangeleverd om de informatie te genereren (bijvoorbeeld luchtfoto's) of de intellectuele eigendomsrechten kunnen worden geclaimd door de overheid om hun economische belangen te beschermen. Door intellectuele eigendomsrechten te claimen kan geo-informatie commercieel worden geëxploiteerd door het bestand onder licentie en tegen een vergoeding beschikbaar te stellen. De licentie bevat vaak beperkende voorwaarden voor hergebruiker, zoals een verbod op het doorleveren van onbewerkte data. Op die manier kunnen overheidsorganisaties hun investeringskosten (deels) terug verdienen en kan de gevoelige aard van de informatie contractueel worden beschermd. Ten slotte kunnen er andere (internationale) wetgeving of afspraken zijn die het publiceren van overheidsinformatie belemmeren of juist promoten, waarbij het niet altijd duidelijk is welk belang zwaarder weegt.

Het wordt erkend dat om nationale, internationale en globale problemen aan te kunnen pakken, het van groot belang is dat geo-informatie die door één overheid is verzameld met andere overheden wordt gedeeld via een infrastructuur voor ruimtelijke informatie. De hoofdprincipes voor een ruimtelijke informatie infrastructuur is dat gegevens eenmalig worden ingewonnen meervoudig worden gebruikt. Dus diezelfde gegevens worden door andere overheden gebruikt en herbruikt zodat die gegevens niet dubbel hoeven worden verzameld. De gegevens moeten laagdrempelig te vinden zijn via dataportalen, en laagdrempelig te gebruiken zijn. Niet alleen kunnen andere overheidsorganisaties informatie die door één overheidsorganisatie is gegenereerd, gebruiken voor het uitvoeren van hun publieke taken, ook door partijen buiten de overheid zouden de informatie kunnen hergebruiken, d.w.z. gebruiken voor een ander doel dan waarvoor de informatie oorspronkelijk was gegenereerd. Daarom is het voor effectief hergebruik van belang dat de gegevens geharmoniseerd zijn zodat de gegevens eenvoudig gecombineerd kunnen worden met gegevens van andere bronnen.

In dit proefschrift wordt met toegankelijkheid van overheidsinformatie alle facetten van toegang tot informatie bedoeld. Dat houdt in dat informatie kan worden ingezien zonder dat er aanpassingen mogelijk zijn (bijv. toegang tot beleidsdocumenten), kan worden gedownload en/of aangeroepen, en worden gedeeld (bijv. binnen een bepaalde doelgroep). Toegankelijkheid van overheidsinformatie is niet alleen essentieel voor effectieve en efficiënte beleidsvoering, maar wordt ook geassocieerd met het bereiken van andere doelen zoals het nastreven van een transparantere en controleerbare overheid, en het faciliteren van een participatiemaatschappij door, bijvoorbeeld burgers te betrekken in democratische en bestuurlijke processen en het (gezamenlijk) aanpakken van maatschappelijke problemen. Daarnaast kan overheidsinformatie worden hergebruikt door het bedrijfsleven als grondstof voor het creëren van toegevoegde waardeproducten en –diensten. Vooral in de internationale literatuur wordt dit laatste aspect vaak benadrukt en worden er hoge verwachtingen gewekt voor de potentiële economische waarde van overheidsinformatie. Onderzoek

uit het verleden toonde aan dat in Europa hergebruikers van overheidsinformatie technische, organisatorische, juridische en financiële barrières ondervonden waardoor de economische waarde van toegevoegde waardeproducten- en diensten achter bleef op bijvoorbeeld de Verenigde Staten. Vooral de laatste twee barrières, veroorzaakt door restrictieve gebruiksvoorwaarden en hoge licentiekosten, werden vaak als de meest belangrijke barrières benoemd. Open data, d.w.z. bestanden beschikbaar voor hergebruik zonder juridische en financiële beperkingen, zou daar een kentering in brengen. Het concept van open overheidsdata kwam aan het eind van het vorig decennium op, en werd als het "gouden ei" door de overheden omarmd als de beleidslijn voor de toegankelijkheid van (een deel van) overheidsinformatie. Echter, hoewel in de afgelopen jaren overheden middelen hebben geïnvesteerd in het opzetten van informatie-infrastructuren en dataportalen om de toegankelijkheid van overheidsinformatie als open data te faciliteren, laat de voorspelde stroom van toegevoegde waardeproducten nog op zich wachten, en lopen de gerealiseerde baten nog steeds achter op de verwachtingen.

Toen dit onderzoek tien jaar geleden startte, werd het debat over toegankelijkheid van overheidsinformatie gedomineerd door de discussie over kostenregimes. Beleid voor toegankelijkheid van overheidsinformatie varieerde van open toegankelijkheid (overheidsinformatie beschikbaar met minimale verstrekkingsvoorwaarden en voor kosten die niet hoger waren dan de marginale verstrekkingskosten) tot volledig kostendekkend, waarbij alle kosten gemoeid met inwinning, verwerking, onderhoud en verstrekking van de informatie doorberekend worden aan de gebruikers. De meeste lidstaten van de Europese Unie hanteerden een kostendekkend regime voor hergebruik van geo-informatie om ten minste een deel van de hoge inwinningskosten en beheerkosten terug te kunnen verdienen. In 2003 werden er in de Europese Unie een aantal richtlijnen aangenomen om betere toegankelijkheid van overheidsinformatie te faciliteren. Richtlijn 2003/4/EG van het Europees Parlement en de Raad van 28 januari 2003 inzake de toegang van het publiek tot milieu-informatie en tot intrekking van Richtlijn 90/313/EEG van de Raad voorzag in een raamwerk voor burgers om toegang te krijgen tot milieu-informatie. Documenten die betrekking hebben op het milieu moeten via een register geraadpleegd te kunnen worden, bij voorkeur in een elektronische vorm en indien a kopie van een document werd opgevraagd, dan mochten de verstrekkingskosten niet hoger zijn dan de marginale verstrekkingskosten. Richtlijn 2003/98/EG van het Europees Parlement en de Raad van 17 november 2003 inzake het hergebruik van overheidsinformatie, de zogenaamde PSI Richtlijn, voorzag in het scheppen van voorwaarden voor een gelijk speelveld voor alle hergebruikers van overheidsinformatie. De PSI Richtlijn liet echter de lidstaten toe om een kostendekkend regime en restrictieve licentievoorwaarden voor hergebruik te handhaven. Daardoor duurde het nog lang voordat lidstaten hun toegankelijkheidsregime aan pasten.

Aan het einde van het vorig decennium waren er twee belangrijke stimulansen om de toegankelijkheid van overheidsinformatie te verbeteren. De eerste stimulans

TOC

was Richtlijn 2007/2/EG van het Europees Parlement en de Raad van 14 maart 2007 tot oprichting van een infrastructuur voor ruimtelijke informatie in de Gemeenschap (INSPIRE), de zogenaamde INSPIRE Richtlijn, die een raamwerk schepte met standaardisatieregels voor de data en ontsluiting via webdiensten, waardoor de toegankelijkheid van overheidsgeo-informatie sterk verbeterde. De tweede stimulans was de ontwikkeling van open data beleid in navolging van de Open Overheid Richtlijn uit de Verenigde Staten in 2009, en de Digitale Agenda voor Europa in 2010. De INSPIRE Richtlijn en de Digitale Agenda waren de grootste drijfveren in Europa om tot een voorzichtige koerswijziging, van een kostendekkend regime naar een meer open toegankelijksregime te komen. Sinds 2011 wordt steeds meer overheidsinformatie beschikbaar als open data. Hoofdstuk 1 identificeerde de vier grootste barrières van toegankelijkheid van overheidsinformatie als zijnde van technische, organisatorische, juridische en financiële aard. Open data zouden de laatste twee barrières voor een groot deel moeten slechten. De invoering van een open databeleid verschafte een uitgelezen kans om de hypothese te testen dat de grootste barrières voor hergebruikers inderdaad de restrictieve gebruiksvoorwaarden en –kosten waren.

Hoofdstuk 2 toonde aan dat tegen 2008 bijna alle Europese lidstaten de PSI Richtlijn 2003/98/EG hadden omgezet in nationale wetgeving, maar allemaal met verschillende variaties en met veel vertraging. Een van de eerste effecten van de PSI Richtlijn was dat de meeste lidstaten hun toegankelijkheidsbeleid hadden geëvalueerd. Meer overheidsinformatie werd beschikbaar gesteld voor hergebruik en een aantal lidstaten hadden licentiekosten afgeschaft of significant verlaagd. In die lidstaten waar licentiekosten werden verlaagd, nam het aantal hergebruikers toe, waardoor de totale omzet van een aantal overheidsorganisaties zelfs toenam. De INSPIRE Richtlijn 2007/2/EG effende de weg voor technische interoperabiliteit met richtlijnen voor harmonisatie van standaarden voor webdiensten en (metadata)catalogi. Maar geen van beide richtlijnen droegen bij aan het verbeteren van juridische interoperabiliteit. Hoofdstuk 2 toonde ook aan dat er nog een barrière voor hergebruikers werd opgeworpen doordat sommige aanbieders van overheidsinformatie zelf toegevoegde waardeproducten en –diensten ontwikkelde en verhandelden in competitie met de private sector. Daardoor was het gelijke speelveld dat de Europese Commissie voor ogen had, aan het einde van het vorig decennium nog niet gerealiseerd.

In Hoofdstuk 3 werd het aspect van het harmoniseren van licenties onderzocht als een eerste stap naar juridische interoperabiliteit. Eerder onderzoek had uitgewezen dat één van de grootste barrières voor hergebruikers van overheidsinformatie werd veroorzaakt door non-transparante en inconsistente licentievoorwaarden. Een inventarisatie van de licentievoorwaarden gehanteerd door Nederlandse overheidsorganisaties voor de populairdere geo-informatie toonde aan dat, hoewel er veel verschillen in lengte en taalgebruik waren, er ook veel overeenkomsten waren. In dit hoofdstuk werd een conceptueel model voor zogenaamde GeoGedeelde licenties geïntroduceerd voor het harmoniseren van geo-informatielicentie. Het concept van GeoGedeeld was gebaseerd

op het concept van Creative Commons licenties. Het toepassen van GeoGedeelde was een eerste stap in de richting van het harmoniseren van licentievoorwaarden en van juridische interoperabiliteit, speciaal voor hergebruikers die verschillende bestanden van verschillende bronhouders willen combineren. De GeoGedeelde licenties zijn als een serieuze optie beschouwd als een bijlage van de toenmalige conceptversie van de INSPIRE Richtlijn. Helaas kwam het concept van één licentiesysteem voor de gehele Europese Unie te vroeg in 2006. Het GeoGedeeld concept is verder uitgewerkt en geïmplementeerd in het Nationale GeoRegister in Nederland.

De Europese Commissie erkende in 2009 dat overheidsinformatie de grootste bron van informatie in Europa was, en dat het potentieel van hergebruik van overheidsinformatie beter benadrukt moest worden in het digitale tijdperk. Als onderdeel van de evaluatie van de 2003/98/EG PSI Richtlijn, voerde de Europese Commissie een consultatieronde uit met stakeholders in 2010. In die consultatieronde werden de meningen van de stakeholders gevraagd over specifieke aandachtspunten die in de toekomst aangepast zouden moeten worden. De Europese Commissie heeft ook een aantal aanvullende onderzoeken laten uitvoeren. Deze onderzoeken behelsden een literatuuronderzoek van overheidsinformatie hergebruik en gerelateerde marktonderzoeken, een beoordeling van de verschillende financiële modellen voor aanbod en tarifering van overheidsinformatie, en een onderzoek naar hergebruik van overheidsinformatie in de culturele sector. Het eerste onderzoek, uitgevoerd door Graham Vickery in 2011, toonde aan dat de totale potentiële economische baten van het openstellen van overheidsinformatie als grondstof voor het creëren van nieuwe producten en diensten, in de orde van grootte van €40 miljard per jaar voor de Europese Unie zou kunnen zijn. Zowel het onderzoek van Vickery als het tweede onderzoek uitgevoerd i.o.v. de Europese Commissie, het zogenaamde POPSIS rapport, toonden ook aan dat voor de meeste aanbieders van overheidsinformatie inkomsten uit het verstrekken van informatieproducten onder licentie relatief laag was in verhouding met hun totale begroting. De PSI Richtlijn van 2003 aangepast, en in 2013 trad Richtlijn 2013/37/EU van het Europees Parlement en de Raad van 26 juni 2013 tot wijziging van Richtlijn 2003/98/EG inzake het hergebruik van overheidsinformatie die op 17 juli 2013 in werking.

Hoofdstuk 4 beschrijft de voornaamste veranderingen van de Herziende PSI Richtlijn 2013/37/EU, inclusief de aanbeveling voor het gebruik van open data licenties. Verder geeft dit hoofdstuk een overzicht van de verschillende open data licenties die in Europa in gebruik zijn, en analyseert in hoeverre deze open data licenties de juridische interoperabiliteit bevorderen. Hoewel het gebruik van open data licenties de barrière van juridische non-interoperabiliteit had moeten slechten, de praktijk leert dat de verschillende types open data licenties toch niet zo uitwisselbaar zijn. De conclusie van dit hoofdstuk is dat alleen een publieke domeinverklaring zoals een Creative Commons Zero (CC0) verklaring geschikt is voor hergebruikers die open data van verschillende bronhouders in verschillende landen willen combineren. Bovendien is het belangrijk

dat een dergelijke verklaring in een prominente plek wordt gepubliceerd zodat elke onzekerheid voor hergebruikers wordt weggenomen. Zonder een dergelijke verklaring zal het hergebruik van open data nog steeds belemmerd worden omdat hergebruikers minder geneigd zullen zijn om tijd te investeren in het ontwikkelen van toegevoegde waardeproducten en –diensten wanneer het onbekend is wat de mogelijke juridische beperkingen kunnen zijn.

In dit proefschrift worden ook de financiële en economische aspecten van de toegankelijkheid van overheidsinformatie in beschouwing genomen. Hoofdstuk 1 en 2 toonden aan dat een kostendekkend regime voor het verstrekken van overheidsinformatie een financiële barrière op zou werpen voor hergebruikers in de private sector omdat de licentiekosten te hoog zouden zijn. Rond 2008 waren er nog steeds veel voorstanders van het behouden van een kostendekkend regime voor het verstrekken van overheidsinformatie. Vooral overheidsinstellingen die niet uit algemene middelen worden gefinancierd en voldoende inkomsten moeten genereren om een groot deel van hun operationele kosten te dekken, de zogenaamde zelf-financierende instellingen, zagen geen heil in het opgeven van een vaste inkomstenbron. Bovendien moeten er voldoende middelen beschikbaar blijven om te zorgen dat de informatie die door die instellingen ingewonnen, bewerkt, en beheerd worden, van een voldoende kwaliteitsniveau zijn en om de actualiteit en continuïteit te garanderen. In hoofdstuk 5 worden de verschillende businessmodellen en prijsmechanismen onderzocht die toegepast kunnen worden op overheidsinformatie toegankelijk gemaakt via INSPIRE webdiensten. Dit hoofdstuk laat zien dat, afhankelijk van het type webdienst, type informatie en type hergebruiker, het te verdedigen is om een abonnementsmodel als prijsmechanisme in te zetten. Echter, businessmodellen die gebaseerd zijn op het genereren van inkomsten uit het verstrekken van overheidsinformatie zullen op de lange duur niet levensvatbaar zijn, en zijn zeker niet in de geest van de INSPIRE richtlijn. Dit hoofdstuk concludeert dat door verschillende prijsmechanismen te hanteren voor overheidsinformatie webdiensten, er geen financiële interoperabiliteit kan worden bereikt.

In Hoofdstuk 6 worden businessmodellen voor aanbieders van overheidsinformatie opnieuw onderzocht, deze keer vanuit een open data perspectief. Sinds de introductie van open data leid in Nederlands komen zelf-financierende overheidsinstellingen steeds meer onder druk te staan om dit beleid binnen de eigen organisatie te implementeren en informatie als open data aan te bieden. Dit hoofdstuk beschrijft de businessmodellen van overheidsinstellingen die nu al, of binnen afzienbare tijd, open data aanbieden. Het hoofdstuk toont de verschuivingen van de rol die de overheidsorganisatie in de informatiewaardeketen in neemt nadat open data beleid was ingevoerd. De case studies die voor dit onderzoek zijn uitgevoerd, toonden aan dat het aanbieden van (onbewerkte) open data niet noodzakelijk leidt tot omzetverlies op de lange termijn, zolang de betreffende overheidsinstelling voldoende flexibiliteit heeft om haar rol in de informatiewaardeketen aan te passen. Dit geldt vooral voor

organisaties waar de inkomsten uit dataverstrekking slechts een klein percentage van de totale begroting betreft. De case studies toonden ook aan dat de switch naar open data aanbieden waarschijnlijk heeft geleid tot interne efficiëntieslagen bij de data-aanbieders. In de praktijk blijkt het moeilijk te zijn om interne efficiëntieslagen die geheel ten gevolge zijn van open data toe, te isoleren en kwantificeren, aangezien de onderzochte instellingen continue werken aan efficiëntie verbeteringen. Echter, de in de case studies gerapporteerde dalingen van interne en externe transactiekosten ten gevolge van open data zijn in lijn met de resultaten van de case studie die in hoofdstuk 7 is uitgevoerd.

De switch naar open data beleid in Nederland bood ook een uitgelezen gelegenheid om de effecten van open data ex ante te bepalen omdat er nulmetingen konden worden uitgevoerd. Voor open data-initiatieven is het een uitdaging om zowel kwantitatieve als kwalitatieve indicatoren te ontwikkelen waarmee het succes van een beleidswijziging gemeten kan worden. In hoofdstuk 7 is een beoordelingsmodel ontwikkeld om de effecten van open data op de organisatie van een open data-aanbieder te kunnen bepalen. Liander is een private energienetwerkbeheerder met een publieke taak, en was voornemens om in het najaar van 2013 enkele datasets als open data beschikbaar te stellen. Deze gelegenheid bood een unieke kans om het ontwikkelde beoordelingsmodel toe te passen om een inzicht te krijgen van de interne, externe en relationele effecten van open data bij Liander. Een nulmeting werd uitgevoerd vlak voor het openstellen van de datasets, en een vervolgmeting één jaar later. De nulmeting gaf inzicht in de toenmalige werkprocessen en voorbereidingswerk voor het implementeren van open data. De vervolgmeting toonde aan dat de open data van Liander gebruikt werden door een uiteenlopende groep gebruikers, en dat er een positief effect was geweest op de ontwikkeling van een aantal apps ten behoeve van energiebesparingen. Het blijft echter een uitdaging om dergelijke maatschappelijke baten te kwantificeren. De vervolgmeting toonde ook aan dat de reguliere gebruikers van Liander data de open data eerder gebruikten om bestaande toepassingen en werkprocessen te verbeteren dan om nieuwe producten en diensten te ontwikkelen. De case studie toonde aan dat private netwerkbeheerders open data met succes beschikbaar kunnen stellen. De case studie toonde ook aan dat Liander als open data boegbeeld de weg heeft geëffend voor de andere private netwerkbeheerders in Nederland om ook open data beschikbaar te stellen. Het voor dit onderzoek ontwikkelde beoordelingsraamwerk, werd geschikt bevonden om de effecten van open data op de organisatie van de data-aanbieder te monitoren.

Het beoordelingsraamwerk dat in Hoofdstuk 7 was ontwikkeld en getoetst, werd geschikt geacht om de effecten van open data op organisatieniveau te monitoren. Echter, om een completer beeld te kunnen krijgen van de effecten van open data en om te bepalen of er andere barrières zijn voor hergebruikers, is een meer holistische aanpak nodig. Daarom werd een holistisch open data beoordelingsraamwerk ontwikkeld dat het hele open data ecosysteem in beschouwing nam vanuit

TOC

verschillende perspectieven: aanbod, governance en de hergebruikers. In Hoofdstuk 8 werd het holistisch beoordelingsraamwerk op de Nederlandse data infrastructuur toegepast in 2014 om de volwassenheid van open data in Nederland, de zogenaamde "Staat van Open Dataland" te bepalen en om eventuele knelpunten in beeld te krijgen. Het holistisch beoordelingsraamwerk toonde aan dat geodata significant beter scoorden dan andere soorten overheidsdata. De standaardisatie en implementatieregels die in het INSPIRE raamwerk zijn vastgelegd, hebben waarschijnlijk als een katalysator gewerkt om geodata naar een hoger niveau van volwassenheid te tillen. Het holistische beoordelingsraamwerk bood Nederlandse beleidmakers bruikbare aanknopingspunten aan om het open data ecosysteem verder te ontwikkelen en om onderbouwde strategieën te ontwikkelen zodat het potentieel van open data naar een hoger niveau te tillen. Naar aanleiding van de Staat van Open Dataland in 2014, zijn een aantal van de aanbevelingen reeds uitgevoerd.

Dit proefschrift heeft aangetoond dat veel aspecten die de toegankelijkheid van overheidsdata moeten faciliteren, zoals gestandaardiseerde metadata, voor geodata reeds zijn opgepakt. Maar dit proefschrift heeft ook aangetoond dat voor andere overheidsdata er nog een lange weg te gaan is. Er is een groeiende vraag naar overheidsdata zoals financiële data en gezondheidszorgdata. Overheidsinstellingen die dergelijke niet-geodata beheren, hebben behoefte aan praktische richtlijnen om een onderbouwde beslissing te kunnen maken om data als open data toegankelijk te maken. Dergelijke beslissingen zijn niet eenvoudig: de data zijn vaak van een heterogene aard en kunnen microdata (gegevens die observaties of feiten kwantificeren, zoals gegevens verzameld in enquêtes) bevatten. Hoewel microdata niet altijd persoonsgebonden gegevens bevatten, zullen dergelijke gegevens waarschijnlijk toch bewerkt moeten worden. Er is nog veel onzekerheid over kwesties zoals hoe om te gaan met potentieel vertrouwelijke informatie of hoe de data voldoende kan beschreven worden om misinterpretaties te voorkomen.

Verder blijven open data en bescherming van persoonsgegevens op gespannen voet met elkaar staan. De grote vraag blijft tot welk niveau de gegevens geaggregeerd en/of geanonimiseerd moeten worden om de bescherming van persoonsgegevens nu en in de toekomst te waarborgen, en tegelijkertijd ook nog een bruikbare dataset over te houden, d.w.z. met voldoende betekenis om voor hergebruikers nog interessant te zijn. Een ander punt dat verder onderzocht zou moeten worden, is het aspect van data-eigendom. Steeds meer sensordata van bijvoorbeeld telefoons, slimme energiemeters en meetlussen in de weg, worden verzameld door de overheid en het bedrijfsleven, en worden een onderdeel van een steeds groter worden "big data" ecosysteem. Daarnaast worden er steeds meer gegevens verzameld door overheidsinstellingen in samenwerking met andere overheidsinstellingen en het bedrijfsleven. Burgers leveren ook gegevens aan de overheid op een vrijwillige basis, bijvoorbeeld vogeltellingen. Dergelijke informatie, zogenaamde geco-creëerde informatie zal steeds meer gemeengoed worden in de komende jaren, net als sensor data als onderdeel van een

big data ecosysteem. Echter, het aspect van wie de data-eigenaren zijn, welke rechten zij hebben in welk deel van een informatieketen is nog niet onderzocht. Onzekerheid over rechten van derden, zoals intellectuele eigendomsrechten, creëert een barrière voor het publiceren van open data. Voordat geco-creëerde data en sensordata aangemerkt kunnen worden als open data, zal het aspect van data-eigendom verder onderzocht moeten worden.

TOC

1 General introduction

§ 1.1 Geographic data

Data can be considered to be the building blocks for information needed to acquire knowledge and understanding. Data have no meaning on itself but by selecting, combining and/or processing data, a meaning or interpretation can be given to transform data into Information. Knowledge can be considered as information to which value has been added and understood based on context, experience and purpose (*cf.* Ackhoff, 1989; Longley *et al.,* 2001). Geographic data are the main resource for information vital for making sound decisions at the local, regional, and global levels (GSDI, 2009). Geographic data - also known as spatial data – are data that refer to a location on Earth, *e.g.* maps, aerial photography, and satellite images but also registers such as cadastral registers and address data. The location component 'where', complemented with 'who', 'what' and 'when', is an important component of government policy, business decisions and personal choices (GSDI, 2009). Geographic data are often considered "special" for technical, economic reasons and legal reasons. Geographic data are special for technical reasons because geographic data are multi-dimensional, voluminous, can be represented at multiple scales and turn a 3D world into 2D projection. Geographic data require special analysis tools and skills, and are, therefore, complex, time-consuming and expensive to process and maintain (Longley *et al.,* 2001). Geographic data, and especially large-scale geographic data, are special for economic reasons because the fixed production costs to collect the data and transform into geo-information are high. As the variable costs of reproduction are low in a digital environment, the sunk costs of geo-information are substantial. To protect the high investments costs, (re-)use of geo-information may be limited by legal and/or technological means such as intellectual property rights and digital rights management. By claiming intellectual property rights, such as copyright and/or database rights and restricting (re-)use through licences and licence fees, geo-information can be commercially exploited and used to recover some of the investment costs. In addition, because of the geographic component, geographic data may contain sensitive personal data, sensitive company data or environmentally sensitive data and, therefore, may require adaptation before the datasets are published. Thus, geographic data are special for legal reasons.

Governments create, collect, develop and use geographic data for carrying out public tasks. On a regional level, geographic data are essential for tackling cross-

border challenges, such as water management, emergency response and disaster management and protection of the environment. On a global level, spatial information is needed to address issues such as mitigating the effects of climate change, protection of the oceans or mapping epidemic outbreaks. It has been recognised that to deal with national, regional and global challenges, it is essential that geographic data collected by one level of government be shared between all levels of government via a so-called Spatial Data Infrastructure (SDI).[1] The main principles governing SDIs are that data are collected once and (re-)used many times; that data should be easy to discover, access and use; and that data are harmonised so that it is possible to seamlessly combine spatial data from different sources.[2]

Geographic data are only a part of the vast amount of data collected by governments. The umbrella term 'public sector information (PSI)' is often used to denote all documents, datasets and content produced or held by public sector bodies. In the European Union, the term 'document' is defined as "any content whatever its medium (written on paper or stored in electronic form or as a sound, visual or audiovisual recording); and any part of such content.[3] In other words, PSI is "any kind of information that is produced and/or collected by a public body and is part of the institution's mandated role" (OECD 2006, p.7). PSI covers a vast array: from static documents, *e.g.* Hansard record, to highly dynamic and voluminous data, *e.g.* meteorological data.

Geographic data are only a part of the vast amount of data collected by governments. The umbrella term 'public sector information (PSI)' is often used to denote all documents, datasets and content produced or held by public sector bodies. In the European Union, the term 'document' is defined as "any content whatever its medium (written on paper or stored in electronic form or as a sound, visual or audiovisual recording); and any part of such content.[4] In other words, PSI is "any kind of information that is produced and/or collected by a public body and is part of the institution's mandated role" (OECD, 2006, p.7). PSI covers a vast array: from static documents, *e.g.* Hansard record, to highly dynamic and voluminous data, *e.g.* meteorological data.

1 See, for instance, http://www.anzlic.gov.au/ for the Australian and New Zealand SDI strategy; http://cgcrt.ca/ for the Canadian SDI strategy; http://www.fgdc.gov/nsdi-plan for the United States SDI strategy and http://inspire.ec.europa.eu/ for the European Union SDI strategy.

2 http://inspire.ec.europa.eu/inspire-principles/9

3 Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information, art. 2(3)a. and b.

4 Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information, art. 2(3)a. and b.

Not only are government data essential for effective and efficient government policy-making, accessibility, *i.e.* access to and (re-)use of all government data are associated with realizing more ambitions. One driver for accessibility to government data is to foster political accountability and democratic participation, such as a more transparent and accountable government (*e.g.* Huijboom and van den Broek, 2011; Algemene Rekenkamer, 2014), reducing corruption (*e.g.* Granickas, 2014; David-Barrett *et al.,* 2015); improving citizens' participation (Jetzek, 2013), and solving societal problems (*e.g.* Uhlir, 2009; Attard *et al.,* 2015). Another driver is to foster efficiency and effectiveness, both within the government (*e.g.* Huijboom and van den Broek, 2011; van Eechoud, 2014) but also efficiency improvements for organisations (*e.g.* McKinsey Global Institute, 2013; WISE Institute, 2014). A third driver is to foster innovation, economic growth and job creation as companies can re-use government data as a resource for creating innovative and value added products and services, resulting in increasing economic value (*e.g.* Vickery 2011; Omidyar Network, 2014) and societal benefits for the end-user (*e.g.* McKinsey Global Institute, 2013; Jeztek, 2013). Especially the latter ambition, economic growth, has been the subject of many publications stressing the enormous potential economic value of public sector spatial data re-use (see *e.g.* Omidyar Network, 2014; McKinsey Global Institute, 2013; Vickery, 2011; Houghton, 2011; de Vries *et al.,* 2011; Pollock, 2008; Dekkers *et al.,* 2006). However, in spite of considerable resources invested by governments to establish spatial data infrastructures (see, *e.g.* Rhind, 2000), to facilitate data portals (see *e.g.* Martin *et al.,* 2013) and to release government geographical data without restrictions (*e.g.* Danish Government, 2012; Kamerstukken, 2011; National Audit Office, 2012), the expected surge of value added products based on government data has not quite eventuated to date (Algemene Rekenkamer, 2014; European Commission, 2013; Rothenberg, 2012; du Preez, 2012). Previous research indicated that re-users of government data encountered barriers related to technical, organisational, legal and financial aspects. Especially the latter two barriers (restrictive licence conditions and high licence fees) are often cited to be the main barriers for re-users (*cf.* Groot *et al.,* 2007; Fornefeld *et al.,* 2008; Pollock, 2011). However, with more government data becoming available as open data, *i.e.* available for re-use without legal and financial restrictions, the expected benefits still seem to lag expectations. Maybe other already identified barriers have been underestimated. Maybe there are barriers that have not been identified yet. Maybe the benefits of open government data are not purely economic benefits and not measurable with standard economic indicators. Maybe the benefits of open government data are internal benefits (increased efficiency and effectiveness) or external benefits (to the society at large). The main challenge for open data will be how to measure the actual impact of open government data. To do so, an open data assessment framework is required that evaluates open government data not only from a data supplier's perspective, but also from a (re-)user's perspective.

This dissertation aims to bridge the gap between current government data re-use practices and its full potential. User barriers are identified, theoretical concepts

TOC

are developed and designed, and practical bridges are provided to enable re-use to the max. An open data assessment model to determine the effects of open data was developed and tested on open data supplied by Liander, an energy network administrator. This assessment model was refined and extended to assess the maturity of the State of Open Data in the Netherlands. Although most of the research was primarily performed in the Netherlands and aimed at the Dutch geo-sector, the results are equally applicable to a broader perspective.

§ 1.2 Background motivation

This research started after a student asked our research group for advice. The question related to a problem he had encountered with a government data holder. His problem, and other problems encountered by other start-ups, is sketched as a narrative below.

§ 1.2.1 Data requests for an innovative application

Once upon a time, there was a student with a brilliant idea for a smartphone "killer" app. To develop this app, the student needed a national topographic map, municipal real-time public transport information, road maintenance data and crime statistics related to bicycle theft hotspots. As he had only limited funding, our student decided to use a public sector large-scale topographic dataset as layer for the app. He had already used this dataset at the university for his assignments, so to re-use the dataset was a logical choice. Our student now needed to find the public transport data. An internet search showed that the required public transport data covered an area of about 40 municipalities in two provinces. The public transport concessions in the area were fulfilled by three private transport companies and three municipal public sector organisations. As there appeared to be no central portal for re-using such data, the student e-mailed the concession holders with the data request. The student also sent a request to the Department of Public Works, an agency resorting under the Ministry of Infrastructure and the Environment, for a road dataset including actual road maintenance data, and a request to the Ministry of Justice for actual crime statistics related to bicycle theft on detail level by street.

§ 1.2.2 Outcome data requests

The Department of Justice was quick to respond. Within a day, the reply came back that theft data on street detail level will not be made publicly available as such data contain personal data. The student was advised to obtain the aggregated data through the National Statistics Bureau. However, the aggregated data were not timely and the location references were too coarse to be usable for assessing bicycle theft hotspot locations for the intended "donotleaveyourbikehere" app.

The Department of Public Works responded about four weeks later: the data request was under consideration but more time was needed to assess the request. Another four weeks later, the data request was refused on the grounds that releasing the national road dataset for free might disadvantage other companies that had already invested into developing similar information services. The requested dataset could, however, be acquired for a fee.

The responses of the public transport concession holders were varying. One district private sector concession holder was quick to respond and provided the required real-time information to the student. Another district private sector concession holder refused the request outright, claiming that such information was confidential business information and pointed out that the timetables were already published on their website as a document in pdf format. One municipal concession holder responded after some weeks, claiming that real-time information was not provided because it would encourage fare-dodging. Another municipal concession holder would not provide real-time information because such information was already provided to another quasi-public organisation delegated to provide access to public transport timetables. The other two concession holders did not respond at all. The student sent a request to the quasi-public organisation but, alas, to no avail. The organisation claimed they were not in a position to redistribute such data as they were merely a data holder and not the owner of the data and referred the student back to the concession holders.

To cap it all off, the mapping authority, which provided the large-scale base map to the university, sent a letter to the student demanding that the student must immediately cease to use the base map outside the university. The mapping authority had an exclusive contract with the University for re-use of the data for educational purposes only, and not with individual students. The student could obtain the base map from the mapping authority after the student provided a business plan including details of the intended service to be provided, as the mapping authority wanted to prevent misuse of the data. After some hesitation, the student provided the mapping organisation with the requested information. After some time, he received login details to access the required dataset accompanied by a lengthy licence agreement containing terms written in legal language the student did not understand. He was after all, an engineering student and not a legal eagle.

After all these months, our intrepid student started to doubt the viability of his killer app. After the arrival of a hefty invoice for the large-scale base map, he decided to call the whole thing off. When a couple of months later, the mapping authority introduced an app, which looked incredibly similar to the killer app described in the business plan submitted, the student was devastated.

§ 1.2.3 An isolated case or not

A sad story or harsh reality? Unfortunately, for our student, the latter seems to be the case. The above example is based on real-life experiences from case studies into discovering and obtaining public sector data. In April 2006, the Dutch Department of Public Works (Rijkswaterstaat) received two applications under the Public Information Access Act (Wet openbaarheid van bestuur) for their National Roads Dataset. After obtaining legal advice, Rijkswaterstaat provided the dataset to the two applicants and announced in 2007 that the National Roads Dataset would be downloadable for all without licence restrictions for re-use and for free. However, a few days later, the dataset was withdrawn from public access after a mapping company complained that publishing the dataset for free constituted unfair trading practices as the mapping company had already invested in creating a similar product. Although the National Roads Dataset was published without restrictions in December 2011, the momentum for re-using this dataset appeared to have passed as most potential re-users had already found alternative data sources (Welle Donker and van Loenen, 2013).

There are ample examples of public sector bodies developing similar services in competition with the private sector. In the Netherlands, for instance, the Ministry of Education, Culture and Science developed a web service with information for parents to compare primary schools and secondary schools, even though a private sector company had already developed a similar free web service (www.10000scholen.nl) (van Loenen and Welle Donker, 2014). In 2007, Arcadis BV, a private sector provider of geographical information products and consultancy services launched their "Landinkaart" product. The product provided information on plots of land to allow potential buyers to assess the immediate environment of the plot of land. Arcadis BV obtained the necessary data from various public sector bodies, such as municipalities and the Cadastre, Land Registry and Mapping Agency (Kadaster). To obtain the data, Arcadis BV lodged nearly one hundred applications under the Public Information Access Act. Even though the Public Information Access Act required a response within four weeks, some of the municipalities took more than one year to respond to the application. In 2008, the Kadaster launched a similar product, which was not only considerately cheaper but also contained more timely information. As Arcadis BV could not compete under these circumstances, their product was discontinued

in 2010 (de Vries *et al.,* 2011). Landmark BV developed a similar product in 2007 and also lodged applications under the Public Information Access Act to obtain environmental data from all (then) 443 Dutch municipalities. Only one municipality could actually provide Landmark BV unlimited access to their environmental data register.[5] 85 Per cent of the municipalities either supplied reports rather than raw data and often only one report per request. The other 15 per cent of municipalities either did not have the data available in a digital format or argued about the height of the fees charged for data supply (Leenaers, 2008). Although Landmark BV won a legal battle with the Municipality of Amsterdam related to the height of the fees charged for a soil pollution dataset (de Vries *et al.,* 2011), Landmark discontinued their operations in the Netherlands. In the United Kingdom, Intelligent Addressing, a company which operated a gazetteer compiled and run by local councils, encountered prolonged problems with obtaining data from the Ordnance Survey; during which time the Ordnance Survey had developed a similar service based on the same data (OPSI, 2006). In Germany, a private meteorological information provider set up a parallel infrastructure of weather stations after negotiations between the company and the German National Meteorological Services failed (Fornefeld, 2009).

§ 1.3 The legal framework for government geo-information accessibility

§ 1.3.1 Explanation of terms used

In literature related to government data in general, the terms "government data" and "public sector information" are often used. Although "data" and "information" are not synonyms, in this dissertation, they will be used (almost) synonymous. Although (raw) data are the building blocks to generate information, *i.e.* to give meaning to raw data, in nearly all cases, some transformation will have taken place to make raw data suitable

5 In 2003, the UN Aarhus Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters (the 'Aarhus Convention') was transposed into Dutch legislation. The Aarhus Convention provides for a right of access to environmental information as part of every citizen's right to an adequate environment and duty to safeguard the environment for future generations. One of the requirements of the transposed Aarhus Convention was that public sector bodies had to establish and publish a register of all their environmental data collected after 2003. The intention of the register was to provide citizens a list of which environmental data was available and where to find the data. By 2006, only a small percentage of Dutch municipalities had actually published such a register (Leenaers, 2008).

to execute a public task or to support informed decision-making. Before the data are fit for publication, more transformations may have to be done, such as to adapt the data format from a propriety format to an open format, or to aggregate or anonymise the data to avoid that a natural person may identifiable or that the data contain sensitive corporate data. Thus, there is no clear distinction between "data" and "information" in the context of government data.

Furthermore, the terms "government" and "public sector" are also often used as synonyms in literature. According to Article 2(1) of Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information (the so-called PSI Directive), the term public sector bodies "means the State, regional or local authorities, bodies governed by public law and associations formed by one or several such authorities or one or several such bodies governed by public law". The term "government data" refers to data produced or commissioned by government or government controlled entities."[6]

In this dissertation, both terms are used as synonyms.

The terms "geographic" and "spatial" are used interchangeably in this dissertation, as are the terms "geo-information," "geo-data" and "spatial data." In literature, both "geographic" and "spatial" are used interchangeably and both refer to data or information that have a direct or indirect reference to a location or geographical area (INSPIRE Directive (2007/2/EC), art. 3(2)).

In this dissertation, the term "accessibility" is used to denote all aspects of access to and (re-)use of public sector information. Access to public sector information is often linked to a growing demand for governments to become more transparent and accountable. Government bodies may provide passive access (a copy is provided after a request) or active access (pro-actively published online) to information. Use of public sector information is often linked with a growing demand for governments to become more efficient and effective, and usually refers to sharing of (authentic) information to carry out public tasks. Re-use of government information is often linked to economic motivations, and usually refers to third parties combining and transforming information to create a new product. To enable re-use, the information should be published in a way that the information can be downloaded or invoked. Although in a number of publications the terms "access", "use" and "re-use" are used interchangeably, there is a difference between providing free access (*e.g.* allowing views of a register per record) and providing services and tools for re-users to download

6 http://opendefinition.org/government/

or invoke all the records. However, for reasons of linguistic simplicity, the term "accessibility" is used as a collective noun.

Legal interoperability for data relates to a compatible legal environment of laws, policies and agreements needed to allow seamless exchange, combination and re-use of data between different organisations and countries (*cf.* van Loenen, Janssen and Welle Donker, 2012). To achieve legal operability, not only agreement on a legal regime is needed but also agreement on the interpretation of the regime. According to the GEO Data Sharing Working Group legal interoperability occurs when use conditions are clearly and readily determinable for each dataset, the use conditions allow creation and use of combined or derivative products, and data may be legally accessed and re-used without seeking authorisation from data creators on a case-by-case basis (GEO DSWG, 2014).

§ 1.3.2 Public sector information accessibility prior to 2010

At the end of the last decade it was recognised that information generated by public bodies was a major, but so far under-exploited asset, which could be exploited to help maximise the value of this public sector information to governments, citizens and businesses alike. Commercial exploitation of public sector information in the European Union was still limited compared to the United States of America and to Canada.(Pira *et al.*, 2000; Dekkers *et al.*, 2006; Uhlir, 2009). The European Commission recognised that a framework for re-use of public sector information was needed to ensure that companies that there were barriers to re-use of public sector data. These barriers related to legal, financial, organisational and technical aspects. To address these barriers, and to realise the full potential of government geographic data, the European Commission adopted a number of directives, the so-called Access Directive 2003/4/EC[7], PSI Directive 2003/98/EC,[8] and INSPIRE Directive.[9] These directives each deal

7 Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC, 2003, Official Journal of the European Union L 41/26-32,. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:041:0026:0032:EN:PDF

8 Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, 2003,(Official Journal of the European Communities L345/90-96. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02003L0098-20130717&qid=1395944373135&-from=EN

9 Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), 2007, Official Jour-

TOC

with one specific aspect of accessibility, namely, to access data; to create conditions for re-use of data; and to provide a framework to share data. Especially the PSI Directive strived to create a level playing field for the private sector by setting a framework for re-use of public sector information.

Although the EU legal framework related to accessibility of government geographic data attempts to address the problems encountered in the example above, the impact the EU framework had on the geo-sector appeared to be minimal. To create, collect and maintain geographic datasets, especially large-scale datasets require vast investments (van Loenen, 2006; BCR, 2016). The public sector relied on claiming copyright and/or database rights to protect and monetise these investments. Thus, many government spatial datasets were available for re-use with unique licence agreements and licence conditions. The private sector felt that these legal and financial restrictions formed big barriers to the re-use of government spatial data for value added products and services (*e.g.* Fornefeld *et al.,* 2008; Groot *et al.,* 2007). The public sector bodies each used unique licences, which created complex and intransparent licencing structures. In addition, each public sector body used different pricing mechanisms for their data, such as a fee per km^2, per kB or for access time. When combining multiple datasets, the legal interoperability of the different licences became very complicated, especially for a layperson. The financial interoperability was hampered by the different pricing regimes, which made it difficult to determine the total fee payable. It seems that the EU's main objective – to promote a genuine cross-border market without legal limitations – did not eventuate. To promote legal interoperability, a uniform and harmonised licence framework was required. To promote financial interoperability, public sector bodies needed to reconsider their business model and whether with providing access to data via web services; alternative ways may be available to finance their data operations.

§ 1.3.3 Government open data policies

In 2010, the European Union adopted the Digital Agenda for Europe.[10] Within the Digital Agenda, government data published as open data are considered to be an engine for innovation, growth and transparent governance (European Commission, 2011). According to the European Commission, open data refers to the idea that

nal of the European Union L108/ 1-14. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF

10 https://ec.europa.eu/digital-agenda/en

certain data should be freely available for use and re-use[11], *i.e.* there should be no legal or financial barriers to access and re-use government data. The Digital Agenda for Europe intends to overcome the current barriers to re-use of government data by adopting legislative and non-legislative measures and to facilitate open data portals. The European Commission has taken a number of actions, such as to develop an Open Data Portal for its own documents and to establish a pan-European digital service infrastructure aggregating content of existing open data portals inside the EU. The revised Directive on the re-use of public sector information 2013/37/EU amended the framework for publishing government data for re-use by encouraging the use of open licences.

§ 1.3.4 Data infrastructures

There have been other initiatives related to developing data portals and infrastructures, with varying levels of user-friendliness. The INSPIRE Directive 2007/2/EC requires Member States to take measures that address exchange, sharing, access and use of interoperable spatial datasets and spatial data services across the various levels of public authority and across different sectors (recital 3 INSPIRE Directive). On national level, data services and portals have been developed to address the technical barriers related to re-use of government data. However, many of these initiatives are data-driven and appear to be unable to engage the end-user in developing such infrastructures/portals. Although the observation of McLaughlin & Nichols (1994, p. 72) that "users, however, will probably be the most mentioned group and yet actually the least considered" is nigh 30 years old, it seems to still be true today. On national level, platforms for providing access to environmental and health information, *e.g.* the Atlas Leefomgeving, have been developed; however, such platforms only provide access to government information and do not facilitate re-use. Re-users must be able to find the data, *i.e.* via search engines and discovery services. Having located the data, re-users must be able to re-use the data, *i.e.* be able to invoke and/or download the data. The data interoperable, *i.e.* there should be the possibility to combine and transform the datasets without repetitive manual intervention, in such a way that the result is coherent and the added value of the data sets and services is enhanced (INSPIRE Directive, art.3(7)). It appears that there are still legal, financial, technical and organisational obstacles for (re-)users of public sector geo-information. In addition, as shown in the previous section, some public sector bodies produce value added services considered part of their public task. Sometimes these public sector value added services are provided in direct competition with the private sector, thus,

11 https://ec.europa.eu/digital-agenda/en/open-data-0

providing uncertainty for the private sector. The level playing field for the private sector as envisioned by the EU may not be apparent in the geo-sector.

§ 1.3.5 Users and re-users of government information

In Section 1.3.1, the difference between providing access to public sector information and facilitating (re-)use of public sector information was explained. There is a difference between 'users' and 're-users'. The 2003 PSI Directive, art. 4 defines 're-use' as "the use by persons or legal entities of documents held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the documents were produced. Exchange of documents between public sector bodies purely in pursuit of their public tasks does not constitute re-use". Therefore, one could argue that 'users' are any persons or legal entities that use government information for similar purposes for which the information was produced, *i.e.* other government bodies and citizens. The difference between users and re-users is not very distinct as a large percentage of users of public sector information are other public sector bodies (*e.g.* OFT, 2006; DECA, 2010). In an open data ecosystem, the distinction between user and re-users will blur altogether.

§ 1.4 Research aim and scope

With more public sector information becoming available as open data and with rapid technological developments, a trickle of web services and apps based on public sector information can be witnessed. However, the predicted free flow of information products and services based on public sector information has not eventuated yet. The question begs why this is the case and what can be done. Open data policies appear to tackle the often-cited legal and financial barriers, so what is impeding the predicted tsunami of value-added products and services? This dissertation aims to assess which barriers remain for public sector data re-users and to assess the maturity level of open data. The main research question will be:

How can the accessibility of public sector information (PSI) for re-users be improved?

To be able to answer this question a number of sub-questions must be examined.

1 What are the legal barriers to public sector information (PSI) accessibility and how can legal interoperability be achieved?
2 What are the financial barriers to public sector information (PSI) accessibility and how can financial interoperability be achieved and still maintain sustainable PSI provision?
3 How can the effects of open data be assessed and how can such an assessment be used to overcome organisational barriers?
4 a) What are the technical barriers to public sector information (PSI) accessibility and how can the governance be improved?
b) Are there other, not yet identified, barriers for re-users of PSI and how may these barriers be addressed?

This dissertation is based on peer-reviewed articles and book chapters written as part of contract research projects. The order of the research questions, and subsequent chapters, is largely determined by the chronological order of the research projects.

§ 1.5 Research methodology

In this dissertation, different research methods are used. In all cases, literature reviews were combined with case studies. To address the first research sub-question, a review of EU legislation, policy documents, and literature was carried out and analysed to gain a better understanding of the complexity related to the EU legal framework governing re-use of public sector geo-information. A model for standardised geo-licences was developed to harmonise and streamline re-use of public sector geo-information framework (Geo Shared licences). This framework for Geo Shared licences was a major step towards solving a major barrier to legal interoperability. In 2014, an update of open licences used by public sector geo-information was carried out to assess if the variety of open licences contribute to legal interoperability in the European Union.

To address the second sub-question, interviews were held with public sector data providers to analyse existing web services technology and standards. This analysis was combined with a theoretical framework for business models, including financial models and pricing strategies. This has resulted in a framework indicating which business model may be best suited and robust for specific types of geo-information web services. Since the Digital Agenda, public sector bodies are under increasing pressure to publish data as open data. However, public sector bodies required to generate sufficient revenue to cover a substantial part of their operating costs – so-called self-funding agencies – face a challenge in this open data ecosystem. Self-funding agencies need to adapt their business model to cope with losses in revenue due to open data and still guarantee sustainable supply of high quality data in the future. Business models for

self-funding open data providers were researched in combination with their position in the information value chain.

To address the third sub-question, an open data assessment model was developed to measure and monitor the internal, external, and relational effects of open data on the organisation of a data supplier. The assessment model was applied to Liander, an energy network administrator and the results were used to improve organisational aspects of open data supply.

To address the last sub-question, the open data assessment model was fine-tuned and extended to enable open data assessment from multiple perspectives. The holistic open data assessment framework measures the maturity of open data from a data supply perspective, a user perspective and a governance perspective. The open data maturity assessment framework was used to measure the State of Open Data in the Netherlands in 2014. In addition, the framework was used to assess which other barriers may exist for re-users of PSI.

§ 1.6 Research limitations

This dissertation analyses economic and financial aspects of public sector information re-use. However, this dissertation does not provide an in-depth analysis of the socio-economic impact of public sector information re-use. Such research requires a longitudinal case study of the PSI-sector, which was beyond the scope of this dissertation.

This dissertation focuses on public sector geographic data as a resource for value adding in Europe. To understand the complexity of this topic, the concept of public task is an important part of this research. The way a government interprets the scope of a public task, often determines the policies related to publishing data generated as part of that public task. However, this dissertation will not attempt to provide an exact definition of a public task. There has been ample research into the concept of public tasks and public interest (*e.g.* Damme and Schinkel, 2009; Janssen, 2009). What is deemed to be a public task is a political decision and, thus, a dynamic concept depending on political, economic and cultural circumstances. Therefore, this dissertation will concentrate on which actors may carry out a public task in which role, and will not provide a definition of a public task.

Furthermore, although technical aspects will be taken into account, this dissertation will not attempt to go into technical details of these types of barriers to re-use of government

data. This dissertation will be limited to describing and analysing technical aspects of accessibility in relation to governance of public sector information.

There is a growing demand for statistical data and scientific information as open data.[12] The Organisation for Economic Co-operation and Development (OECD) recognised that fostering broader, open access to and wide use of research data would enhance the quality and productivity of science systems worldwide and adopted a Declaration on Access to Research Data from Public Funding in 2004.[13] Open access to research data not only refers to access to reports and peer-reviewed articles but also to the (raw) underlying data. Public trust and confidence may suffer if the data that underpin scientific information and statistics are not freely available along with the methods used (Jackson, 2012, p.3). However, many statistical datasets and research data contain microdata, *i.e.* data that quantify observations or facts, such as data collected during surveys. Confidentiality issues are often cited as the main reason for not publishing microdata. Not all microdata contain personal data, *e.g.* data related to government expenditures could be identified as microdata, which may be classified as confidential information. In addition to a potential breach of confidentiality, publishing microdata may pose other threats, such as revealing poor data quality or misinterpretation of the data due to a lack of knowledge (Jackson, 2012). Although National Statistical Bureaus are in the process of developing standards for publishing statistical data as open data, there is still a long way to go for research microdata. During the latter part of this research, the issue of microdata has been raised as a concern by a number of government organisations. However, this is a complex topic and requires further research beyond the scope of this dissertation.

Finally, this dissertation will not address the tension between open data and the protection of personal data. Government data may contain data on a personal level, for instance, a trade register containing the names and addresses of all companies, including details of one-man companies, which operate from a home address. Such data are personal data and must be processed before publication, *e.g.* aggregation and/or anonymisation or removal of the sensitive files from the open data version. However, the dataset may lose meaning and become less valuable for re-users. In the case of the trade register, if all records of one-man companies were removed from the open data version of the register, the remaining dataset would only contain just over half of all registered companies (van Loenen *et al.,* 2016).Even when data are processed to be deemed suitable for publication as open data, there are other issues. Information Technology, including cloud computing, is developing fast and data processing power

12 See, *e.g.* http://ec.europa.eu/isa/documents/publications/report-on-high-value-datasets-from-eu-institutions_en.pdf.

13 http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=157

and storage capacity have increased exponentially. In 2016, the average person's smartphone has more computing power than home computers in 2006. As more data become available as open data, the advances in technology make it easier to combine datasets. Although open data may not seem to be personal data on first glance especially when anonymised or aggregated, it may become personal data by combining with other data or when de-anonymised (Kulk and van Loenen, 2012). In addition, with apps and tools based on open government data, there is nothing to prevent the use of open data for profiling, data mining and other activities, which have privacy implications for individuals (Scassa, 2014, p.407). The EU Data Protection Directive and the new EU General Data Protection Regulation framework may have a serious impact on publishing open data. However, this topic requires a legal approach and is beyond the scope of this dissertation.

§ 1.7 Outline of this dissertation

This dissertation is based on scientific peer-reviewed articles and book chapters, written for research carried out between 2005 and 2016. Although the papers were written as contract part of a variety of research projects, there is a common theme, namely the various aspects of accessibility to public sector data (see Figure 1.1 for the dissertation outline).

The research builds on work carried out between 2005 and 2008 for the GeoPortal Networks project (RGI-006) and Geodata project (RGI-117), part of the Dutch Space for Geo-Information (RGI) Programme (Chapters 2, 3 and 5), and on work carried out between 2011 and 2015 for the Open Data and Beyond projects, part of the Dutch Next Generation Infrastructures (NGI) programme (Chapter 7). The RGI programme and the NGI programme were co-funded by the Dutch Bsik Programme, which provided subsidies to stimulate research and investments into a knowledge infrastructure.[14]

14 Bsik (Besluit subsidies investeringen kennisinfrastructuur) was an incentive programme of the Ministry of Economic Affairs to stimulate research in order to strengthen the foundations of a knowledge infrastructure. The Bsik programme subsidised 4-6 year research projects carried out by consortia consisting of knowledge institutes, the public sector, and the private sector (http://www.rvo.nl/subsidies-regelingen/besluit-subsidies-investeringen-kennisinfrastructuur-bsik).

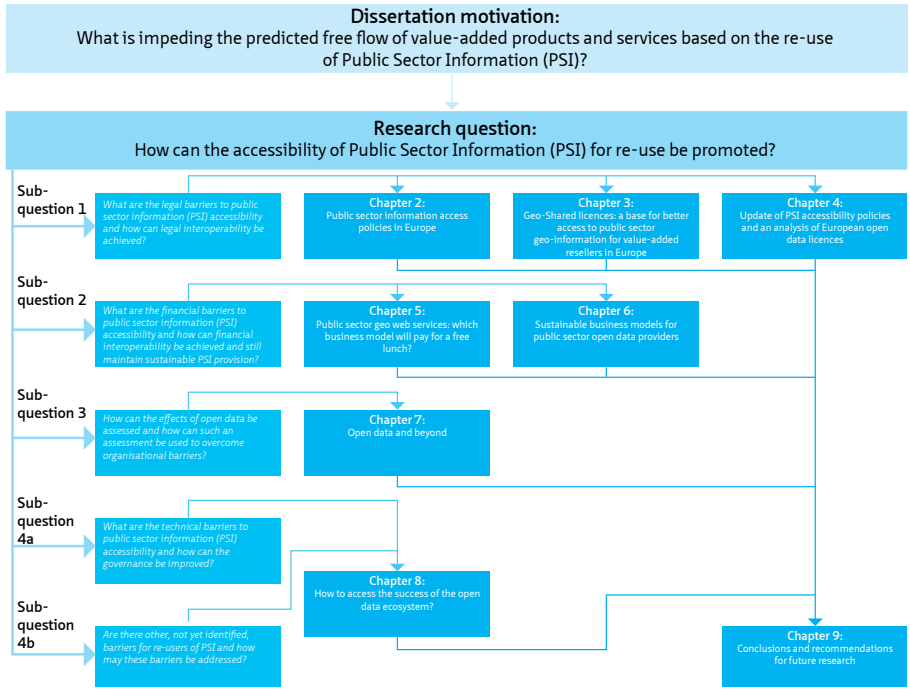


FIGURE 1.1 Dissertation outline

Chapter 4 builds on work carried out between 2014 and 2015 for the European Location Framework (ELF) project, co-funded by the European Commission's Competitive and Innovation Framework Programme (CIP)[15]. Chapter 6 builds on research carried out in 2014 and 2015 and commissioned by the Dutch Kadaster and TNO. Chapter 8 builds on research carried out in 2014 as part of the ICT Breakthrough Project "Open Geo-data as a component of growth and innovation." The consortium members of the ICT Breakthrough Project Open Geodata were representatives of the public sector (Ministry of Economic Affairs, Ministry of Infrastructure and the Environment, and the Ministry of the Interior and Kingdom Relations), representatives of the private sector, and representatives of knowledge institutes (Delft University of Technology and TNO).

Chapters 2 to 4 address sub-question 1: what are the legal barriers to accessibility of public sector information and how can legal interoperability be achieved. Chapter 2, 'Public Sector Information Access Policies in Europe' provides an overview and analysis of public sector geo-information access policies in Europe that were in place

15 The Competitiveness and Innovation Framework Programme (CIP) ran from 2007 to 2013 and supported innovation activities (including eco-innovation), encouraged a better take-up and use of information and communication technologies (ICT), and helped to develop the information society (http://ec.europa.eu/cip/).

TOC

around 2008. The paper describes the legal framework related to providing access to environmental data for citizens, facilitating re-use of public sector information for the private sector, and facilitating the exchange of geographical information between governments in Europe. The paper also describes how the access policies were interpreted in five European countries. This paper provides an analysis of access policies around 2008 and is, therefore, somewhat dated in 2016. However, the main aims of the described legal framework are still valid today.

Chapter 3, 'Geo-Shared licences: a base for better access to public sector geo-information for value-added resellers in Europe' addresses one of the most-cited barriers to re-use of public sector information, namely complex and intransparent licencing structures. The paper provides an overview of the various licences for re-use of public sector information in place in 2007 in a number of European countries and proposes a mechanism for achieving legal interoperability. The developed Geo-Shared licencing structure – based on the Creative Commons concept – provides a framework for transparent and harmonised licences that should be a serious option for a pan-European geographic data infrastructure to enable legal interoperability. The Geo-Shared Licences concept proposed in the paper was further developed and implemented in the Netherlands to provide a standard set of licences for the Dutch National GeoRegister. The Geo-Shared licences became an important lever for adopting Creative Commons licences for public sector information in the Netherlands.

Chapter 4, 'Update of PSI accessibility policies and an analysis of European open data licences' provides an update of the legal framework for re-use of public sector information described in Chapter 2. In 2013, the so-called PSI Directive (2003/98/EC Directive on the re-use of public sector information) was amended by Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information. Chapter 4 describes the changes made by the 2013/37/EU PSI Amended Re-use Directive. In addition, this chapter analyses the various open data licences used by topographic and cadastral data suppliers, which participate in the European Location Framework project. This chapter concludes that although open data licences are a step into the right direction to remove legal barriers, to facilitate true cross-border re-use of public sector information, more needs to be done to achieve true legal interoperability.

Chapters 5 and 6 address sub-question 2: what are the financial barriers to public sector information (PSI) accessibility and how can financial interoperability be achieved and still maintain sustainable PSI provision. Chapter 2 had identified the financial barriers encountered by re-users to not only constitute the level of licence fees charged but also the different pricing mechanisms employed by data holder. Chapter 5, 'Public sector geo-web services: which business model will pay for a free lunch?' provides an overview of different pricing mechanisms suitable for public sector information web services. The paper describes the different types of web services available for publishing

public sector information and introduces the theoretical framework for business models. The components of business models and various pricing mechanisms are explained. The paper provides a number of business models suitable for public sector bodies that supply data web services and concludes that, although there may be a case for public sector bodies to charge for their data, a full cost recovery regime may not be viable in the long-term. The research for this paper was carried out in 2007 for the web services that had to be developed as part of the INSPIRE Directive.

Chapter 6, 'Sustainable business models for public sector open data providers' revisits public sector business models. Since 2011, public sector bodies are under increasing pressure to adopt open data policies. Financial interoperability for re-users would be achieved with open data, *i.e.* available for free. However, between proposal of an open data policy and successful implementation are practicable obstacles, especially for public sector agencies required to generate sufficient revenue to cover their operating costs, so-called self-funding agencies. With revenue losses due to open data, there is a real risk that the update frequency and the quality of data may suffer or that the open data policy may even have to be reversed. This chapter researches the financial effects of open data policies for self-funding agencies on their business model. The paper provides some hands-on proposals for self-funding agencies having to implement an open data policy whilst ensuring their long-term sustainability.

Sub-question 3: how can the effects of open data be assessed and how can such an assessment be used to overcome organisational barriers is addressed by Chapter 7, 'Open Data and Beyond'. This chapter describes the open data experiences of Liander, a Dutch private energy network administrator. Not only governments worldwide see the potential benefits of publishing their data as open data, such as more transparency, increased governmental efficiency and effectiveness, and external benefits, including societal and economic benefits. A number of private sector companies also recognise the potential benefits of making their some of their datasets available as open data (*e.g.* Omidyar Network, 2014). One such company is Liander, an energy network administrator in the Netherlands. Liander views open data as a contributing factor to energy conservation. This chapter identifies the organisational aspects of open data supply. In addition, the chapter describes a monitoring framework that was developed to assess the effects of open data, and applies the framework to Liander's small-scale energy consumption dataset as a first step towards assessing the effects of open data.

Chapter 8, 'How to assess the success of the open data ecosystem?' addresses sub-questions 4a and 4b, and explores why the actual (re-)use of open public sector data fails to live up to its expectations. In spite of more public sector datasets made available, user requirements appear to be ignored. To develop a successful and sustainable open data strategy, it is essential to consider the technical aspects of open data supply and open data governance aspects, and user requirements holistically. The paper describes an open data assessment framework developed by

the Knowledge Centre GI Governance, which will provide the basis for developing well-founded strategies that will ensure the full potential of open data will be reached. The assessment framework is a useful tool to assess the level of open data maturity in a nation and to indicate which aspects of the open data ecosystem are successful and which aspects require attention. The open data assessment framework uncovered other, not yet identified barriers for re-users of public sector data.

Chapter 9 concludes this dissertation with the main conclusions and recommendations for future research. In this chapter, the research carried out in Chapters 2, 3 and 5 are updated where needed and analysed in the context of open data. This chapter describes the lessons learned from the case studies and provides an overview of what has been achieved to date. The chapter concludes with recommendations for more research into issues that need to be addressed for open data to become successfully embedded into society.

References

Ackhoff, R.L. 1989. "From data to wisdom." Journal of Applied Systems Analysis 16, pp. 3-9.

Algemene Rekenkamer. 2014. Trendrapport open data [Trend Report Open Data 2014]. http://www.rekenkamer.nl/Publicaties/Onderzoeksrapporten/Introducties/2014/03/Trendrapport_Open_data, Accessed 28 March 2014.

Attard, J., F. Orlandi, S. Scerri and S. Auer. 2015. "A systematic review of open government data initiatives." Government Information Quarterly http://dx.doi.org/10.1016/j.giq.2015.07.006:20.

BCR [Bureau of Communications Research]. 2016. Open government data and why it matters. A critical review of studies on the economic impact of open government data. Forrest ACT 2603, Australia https://www.communications.gov.au/departmental-news/open-government-data-and-why-it-matters-now, Accessed June 10, 2016.

Damme, E. v. and M.P. Schinkel. 2009. Marktwerking en Publieke Belangen. Koninklijke Vereniging voor de Staatshuishoudkunde Preadviezen 2009. Amsterdam: Koninklijke Vereniging voor de Staatshuishoudkunde.

Danish Government / Local Government Denmark. 2012. Basic Data for improved efficiency and growth. The eGovernment strategy 2011-2015. Albertslund: Rosendahls - Schultz Distribution. http://www.digst.dk/Home/Servicemenu/English/Digitisation/~/media/Files/English/BasicDataUKweb20121008.ashx, Accessed October 30, 2012.

David-Barrett, E., P. M. Heywood and N. Theodorakis. 2015. TACOD Towards a European Strategy to reduce corruption by enhancing the use of open data. National research United Kingdom: RiSSC - Research Centre on Security and Crime - IT. http://www.tacod.eu/wordpress/wp-content/uploads/2015/04/National_Research_UK_def.pdf, Accessed January 28, 2016.

de Vries, M., L. Kapff, M. Negreiro Achiaga, P. Wauters, D. Osimo, P. Foley, K. Szkuta, J. O'Connor and D. Whitehouse. 2011. Pricing of Public Sector Information Study. Models of Supply and Charging for Public Sector Information (ABC) Final Report. Brussels: Deloitte Consulting,. https://ec.europa.eu/digital-single-market/en/news/pricing-public-sector-information-study-popsis-models-supply-and-charging-public-sector, Accessed January 1, 2012.

DECA (Danish Enterprise and Construction Authority). 2010. The value of Danish address data: Social benefits from the 2002 agreement on procuring address data etc. free of charge: 8. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf, Accessed July 7, 2010.

Dekkers, M., F. Polman, R. te Velde and M. de Vries. 2006. Measuring European Public Sector Information Resources (MEPSIR) Study. Final report of study on exploitation of public sector information - benchmarking of EU framework conditions: HELM Group of Companies of Moira, Northern Ireland & ZENC, the Netherlands. http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=1198, Accessed 30 June 2013.

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). 25 April 2007.

Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information. 26 June 2013.

du Preez, D. 2012. "Prime minister's special envoy 'disappointed' with open data use Computerworld UK, October 18, 2012.

European Commission. 2011. Open Data: An engine for innovation, growth and transparent governance. Brussels. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF, Accessed June 10, 2013.

European Commission. 2011. Open Data: An engine for innovation, growth and transparent governance. Brussels. COM(2011) 882 final: 13, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF, Accessed June 10, 2013.

European Commission. 2013. European Union Location Framework Strategic Vision. : Joint Research Centre. http://ec.europa.eu/isa/actions/documents/isa-2.13_eulf-strategic-vision-lite-v0-3_final_en.pdf, Accessed October 14, 2014.

Fornefeld, M. 2009. "The value to Industry of PSI: the business sector perspective The Socioeconomic Effects of Public Sector Information on Digital Networks: Towards a Better Understanding of Different Access and Reuse Policies. Workshop Summary, Paul Uhlir, Ed., pp. 10-16. Washington: The National Academies Press.

Fornefeld, M., G. Boele-Keimer, S. Recher and M. Fanning. 2008. Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report. Study commissioned by EC in 2007. Dusseldorf: MICUS. http://ec.europa.eu/newsroom/document.cfm?doc_id=1258., Accessed January 28, 2009.

GEO DSWG [GEO Data Sharing Working Group]. 2014. White Paper: Mechanisms to Share Data as Part of GEOSS Data-CORE. https://www.earthobservations.org/documents/dswg/Annex%20VI%20-%20%20Mechanisms%20to%20share%20data%20as%20part%20of%20GEOSS%20Data_CORE.pdf, Accessed June 29, 2016.

Granickas, K. 2014. Open Data as a tool to fight corruption: European Public Sector Information Platform Topic Report no. 2014/04. https://ofti.org/wp-content/uploads/2014/05/221171136-Open-Data-as-a-Tool-to-Fight-Corruption.pdf, Accessed January 28, 2016.

Groot, N., A. Bregt, J. Crompvoets and Y. Pluijmers. 2007. "Toegankelijkheid publieke geodata vooral gehinderd door juridische barrières [Accessibility public geodata mostly hampered by legal barriers]." Vi Matrix, dec. 2007-jan. 2008, pp. 6-8.

GSDI [Global Spatial Data Infrastructure]. 2009. The SDI Cookbook. http://www.gsdi.org/gsdicookbookindex, Accessed July 1, 2010.

Houghton, J. 2011. Costs and benefits of data provision. Report to the Australian National Data Service. Melbourne: Centre for Strategic Economic Studies, Victoria University. http://ands.org.au/resource/houghton-cost-benefit-study.pdf, Accessed August 27, 2012.

Huijboom, N. and T. van den Broek. 2011. "Open Data: an international comparison of strategies." European Journal of ePractice no. 12 (March/April 2011):13. http://share-psi.eu/papers/TNO-epractice.pdf.

Jackson, P. 2012. Microdata Exchange and the Challenges of Open Data and Transparency, Expert Group for International Collaboration on Microdata Access: 11. Paris: OECD.

Janssen, K. 2009. "The EC legal framework for the availability of public sector spatial data. An examination of the criteria for applying the Access Directive, the PSI Directive and the INSPIRE Directive Ph.D. Dissertation, Katholieke Universiteit Leuven.

Jetzek, T. 2013. "The value of Open Government Data." Geoforum Perspectiv 12 (23 (2013)), pp. 47-56. http://journals.aau.dk/index.php/gfp/article/view/489/400, Accessed February 4, 2014.

Kamerstukken. 2011. Hergebruik en Open Data: naar betere vindbaarheid en hergebruik van overheidsinformatie

Reuse and Open Data: for better visibility and reuse of government information. Letter to the Chairman of the Lower House of Parliament]. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Den Haag, doi:kst-32802-2, Accessed 14-06-2011.

Kulk, S. and B. van Loenen. 2012. "Brave New Open Data World?" International Journal of Spatial Data Infrastructures Research 7:196-206. doi: 10.2902/1725-0463.2012.07.art10.
Leenaers, H. 2008. "Digitale deur milieu op een kier [Digital door environment ajar]." Binnenlands Bestuur, 25 januari 2008, pp. 30-32.
Longley, P.A., M.F. Goodchild, D.J. Maguire and D.W. Rhind. 2001. Geographic information systems and science. Chichester, UK: John Wiley & Sons.
Martin, S., M. Foulonneau, S. Turki and M. Ihadjadene. 2013. "Risk Analysis to Overcome Barriers to Open Data." e-Journal of e-Government 11 (2):348-359. www.ejeg.com/issue/download.html?idArticle=296, Accessed June 22, 2015.
McKinsey Global Institute. 2013. Open Data: Unlocking innovation and performance with liquid information: McKinsey Global Institute. http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information, Accessed October 31, 2013.
McLaughlin, J. and S. Nichols. 1994. "Developing a National Spatial Data Infrastructure." Journal of Surveying Engineering 120 (2):62-76. doi: http://dx.doi.org/10.1061/(ASCE)0733-9453(1994)120:2(62).
National Audit Office. 2012. Cross-government review: Implementing transparency. London: National Audit Office. http://www.nao.org.uk/publications/1012/implementing_transparency.aspx, Accessed July 2, 2012.
OECD (Organisation for Economic Co-operation and Development). 2006. Digital Broadband Content: Public Sector Information and Content: 82, http://www.oecd.org/internet/ieconomy/36481524.pdf, Accessed 30 March 2006.
OFT (Office of Fair Trading). 2006. The commercial use of public information (CUPI): 183, http://www.oft.gov.uk/shared_oft/reports/consumer_protection/oft861.pdf, Accessed May 1, 2007.
Omidyar Network. 2014. Open for Business: How open data can help achieve the G20 growth target. A Lateral Economics report commissioned by Omidyar Network. https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf, Accessed September 15, 2015.
OPSI [Office of Public Sector Information]. 2006. Office of Public Sector Information report on its investigation of a complaint (SO 42/8/4): Intelligent Addressing and Ordnance Survey. http://www.nationalarchives.gov.uk/documents/information-management/so-42-8-4.pdf, Accessed February 8, 2016.
PIRA International Ltd, University of East Anglia and KnowledgeView Ltd. 2000. Commercial exploitation of Europe's public sector information - Final report edited by Pira International Ltd: European Commission Directorate General for the Information Society. ftp://ftp.cordis.lu/pub/econtent/docs/commercial_final_report.pdf, Accessed November 20, 2006.
Pollock, R. 2008. The Economics of Public Sector Information: University of Cambridge. Original edition, November 2008. http://rufuspollock.org/economics/papers/economics_of_psi.pdf, Accessed February 1, 2012.
Pollock, R. 2011. Welfare gains from opening up Public Sector Information in the UK: University of Cambridge. http://rufuspollock.org/papers/psi_openness_gains.pdf, Accessed November 1, 2013.
Rhind, D. 2000. "Funding an NGSI Geospatial Data Infrastructure Concepts, Cases and Good Practices, R. Groot and J. McLaughlin (eds), 39-55. Oxford University Press.
Rothenberg, J. 2012. Case study International Benchmark. Open data and use of standards. Towards a better supply and distribution process for open data. https://www.forumstandaardisatie.nl/fileadmin/os/documenten/Internationale_benchmark_v1_03_final.pdf, Accessed October 14, 2014.
Scassa T. 2014. "Privacy and Open Government". Future Internet 6(2), pp. 397-413. doi: 10.3390/fi6020397.
Uhlir, P. F., ed. 2009. The socioeconomic effects of public sector information on digital networks. Towards a better understanding of different access and reuse policies - Workshop Summary. http://www.nap.edu/openbook.php?record_id=12687. Washington: National Academies Press.
van Damme, E. and M.P. Schinkel. 2009. Marktwerking en Publieke Belangen. Koninklijke Vereniging voor de Staatshuishoudkunde Preadviezen 2009. Amsterdam, Koninklijke Vereniging voor de Staatshuishoudkunde, ISBN 978-90-807422-8-4.
van Eechoud M.M.M. 2014. "De lokroep van open data" [The call of open data], Amsterdam: Universiteit van Amsterdam, Amsterdam, p.21. http://www.ivir.nl/publicaties/download/1407.pdf, Accessed: October 28, 2016.
van Loenen, B. 2006. "Developing geographic information infrastructures: The role of information policies. PhD Dissertation, Delft University of Technology.
van Loenen, B., K. Janssen and F. Welle Donker. 2012. "Quest for a global standard for geo-data licences." refereed book chapter GSDI 13 World Conference Global Geospatial Conference 2012: Spatially enabling government, industry and citizens, Québec City, 14-17 May 2012. doi: http://repository.tudelft.nl/view/ir/uuid%3Aca7682be-9bca-47d9-8d29-0db047051cb5/.

van Loenen, B. and F. Welle Donker. 2014. De stand in opendataland [The State of the Open Data Nation]. Delft: OTB - Research for the built environment, doi:uuid:7b425db6-2f9a-4be6-b5f0-3492055268d8
van Loenen, B., W. Korthals Altes, D. Groetelaers and F. Welle Donker. 2016. Ontsluiten handelsregister met open data nader belicht [Unlocking trade register as open data further discussed]. Delft: OTB - Research for the built environment, doi: uuid:2c053185-fb19-49cb-a6db-55efdd567647.
Vickery, G. 2011. Review of recent studies on PSI re-use and related market developments. Paris: Information Economics. http://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments, Accessed November 1, 2011.
Welle Donker, F. and B. van Loenen. 2013. Monitoren van de effecten van Open Data voor Rijkswaterstaat; De (mogelijke) effecten van het NWB en het AHN als open data. Delft: Onderzoeksinstituut OTB.
WISE Institute. 2014. Big and open data in Europe. A growth engine or a missed opportunity? Warsaw: WISE Institute [Warsaw Institute for Economic Studies], http://www.bigopendata.eu/, Accessed November 29, 2014.

2 Public sector information access policies in Europe

Frederika Welle Donker

In: Fitzgerald, B. (Ed) 2010. *Access to Public Sector Information: Law, Technology & Policy. Volume 1*. Sydney (Sydney University Press), pp. 253-279, ISBN 978-1-920899-39-4.

§ 2.1 Introduction

In the digital age geo-information has become embedded in our daily lives, such as navigation systems, community platforms, real estate information and weather forecasts. Everybody uses geo-information for their day-to-day decision making. Therefore, access to geo-information is of vital importance to the economic and social development of the nation. Most geo-information, especially the more valuable large scale geo-information is owned by governments all over the world. Government bodies create, collect, develop and disseminate geo-datasets and geo-information to support their public tasks. Although this information is primarily created and collected for internal use, it forms a rich resource for other public sector bodies, citizens and the private sector.

There have been a number of initiatives within the European Union (EU) to provide access to and reuse of this public sector information in order to create a free flow of information and services within the EU. Initially aimed at paper documents, these initiatives had little effect on geo-information. Geo-information existed as paper maps or geo-information systems requiring specialised software. But in the last decade improved computer processing capabilities, broadband internet and interoperability of systems have led to mass digitalisation and thus better availability of information in general. EU initiatives to improve access to information, especially the 2003 Directive on reuse of public sector information, the so-called PSI Directive (2003/98/EC), should have had a flow-on effect on geo-information. But five years after adoption, its impact has not quite lead to the expected surge of value added geo-information products and services as predicted by some (*e.g.* PIRA, 2000; RAVI, 2000). The private sector still faces legal, financial and organisational obstacles when trying to access public sector information (*e.g.* MICUS, 2003 and 2008; Groot *et al.*, 2007).

So, maybe access to public sector geo-information is still not as simple as EU legislation intended it to be. The level playing field as envisioned by EU legislation may not be apparent in the geo-sector. What impact has the EU framework had on access to public sector geo-information to date? This paper will provide a description of the current EU framework. A brief history of public sector geo-information availability will be presented, and a description of the current situation in a number of European countries. The paper will finish with some conclusions and recommendations.

§ 2.2 Geo-information

§ 2.2.1 Geo-information use and users

What is geo-information exactly and why is it so different from other products? To start with, there are many different descriptions of geo-information, depending on the country and the application. Also, the terms "geo-information", "geo-data", "spatial information" and "spatial data" are interchangeably used as synonyms. For the purpose of this paper only the term geo-information (GI) will be used. There are many definitions for the concept of GI. MICUS (2008) defines GI fairly narrowly as "topographical data in all scales, cadastral information (including address coordinates and aerial photography" because these are the categories with the highest reuse rates. In the EU GI is defined as "any data with a direct or indirect reference to a specific location or geographic area" (EU, 2007). After a literature study, Longhorn and Blakemore (2008, p.5) came up with possibly the broadest definition:

"Geo-information is a composite of spatial data and attribute data describing the location and attributes of things (objects, features, events, physical or legal boundaries, volumes, etc.), including the shapes and representations of such things in suitable two-dimensional, three-dimensional or four-dimensional (x, y, z, time) reference systems (e.g. a grid reference, coordinate system reference, address, postcode, etc.) in such a way as to permit spatial (place-based) analysis of the relationship between and among thing so described, including their different attributes".

GI may exist as static information such as aerial images, topographic maps, statistical data, land administration data or census data, but also as dynamic information such as meteorological radar data. In short, GI is more than just digital maps or cadastral information, it also includes administrative information such as address codes,

environmental data, government spatial planning and legal system information. Because of its broad scope GI has become a valuable resource in current society.

One of the most efficient ways of making GI available is through an infrastructure. In the EU it will be mandatory for Member States to set up geo-information infrastructures (GIIs) in order to share public sector geo-information (PSGI) between governments. It is envisaged that such infrastructures will also be used by other users. Van Loenen (2006) distinguishes four types of users of a GII, namely primary users (the collector and major users); secondary users (incidental users for similar purposes as the primary user); tertiary users (users that use the dataset for other purposes than the purposes for which the information was collected and the dataset created); and end-users. Van Loenen (2006) asserts that the tertiary users will be the main drivers of the development of a GII. The private geo-sector, including firms that add value to existing GI and resell those products and services, the so-called value added resellers (VARs), form a large proportion of this tertiary users group. But also the end-users are becoming more influential in the development of GIIs. By exploring the viewing possibilities of GIIs they provide essential feedback. This is why consistent access policies are vital for the development of GIIs.

§ 2.2.2 Limitations

Geo-information – like all other forms of information – has economic aspects which sets it apart from other products. In the case of large scale GI, the fixed production costs of creating information are high and there are also substantial sunk costs. Sunk costs are costs which must be incurred to compete in a market but are not recoverable on exiting the market. The variable costs of reproducing information are low and do not increase if additional copies are produced, i.e. the marginal costs are low. There are also no natural capacity limits to the number of copies produced (Shapiro and Varian, 1999). As such, information shows characteristics of a public good, i.e. a good that is non-rivalrous and non-excludable. Consumption of information does not reduce its availability for consumption by others, and in principle no-one can be excluded from consuming the good. However, because of the high investments costs consumption of GI may be limited by legal and/or technological means such as copyright and digital rights management. Thus, by making GI excludable, GI becomes a club good, i.e. a non-rivalrous but excludable good. By claiming intellectual property rights (IPRs) such as copyright – and in the EU also database rights – (re)use of GI can be controlled and commercially exploited through licences. Restricting use with licence conditions and charging a fee allows for recouping some of the investments made. If the public sector makes GI available, fees may vary from marginal cost recovery, e.g. the costs of burning a DVD and postage, to full cost recovery including all investment costs and personnel costs. Especially large scale GI may end up costing millions of Euros for land-covering datasets.

GI may consist of many base datasets to make a total package. Integrating and analysing the many varied types of data may be time-consuming, and the process of updating is complex (Longley et al., 2001). Also, these individual base datasets are often from different sources and owned by different parties. These parties may or may not claim IPRs. Therefore, even if only one party supplying only a small part to the total information limits use by IPR, then the entire information will be limited as well. For example, a government agency produces a file containing information related to roads. The information includes datasets such as type of road surface, maintenance schedules, topographical layers, address coding, et cetera. The topographical layers are created by another public sector agency and are derived from aerial photographs. The aerial photographs are supplied by a private firm, specialised in such products. The firm claims copyright as a way to commercially exploit their images. The firm may stipulate that for each government agency a separate contract has to be negotiated. The firm may also stipulate that the derived products may not be made available to third parties because the same firm also sells the same aerial images to these third parties.

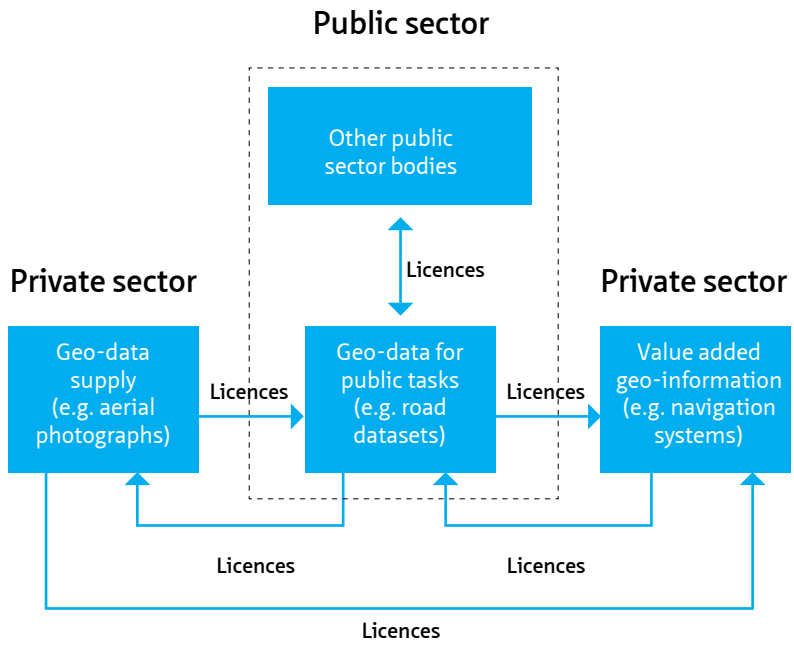


FIGURE 2.1 Flow of geo-information between public and private sector (Welle Donker, 2009)

Another reason why reuse of GI may be limited is that GI may contain data that are subject to privacy protection legislation, *e.g.* data linked to a natural person. Data may also be limited because of security issues, *e.g.* satellite images showing army bases or GI may be linked to sensitive information such as breeding sites of endangered

animals. As such, GI may have to be adapted before it is made available for (re)use, or may even be withheld or withdrawn from publication altogether.

§ 2.2.3 Public sector geo-information

GI, and especially large scale GI, is primarily used by the public sector for public tasks such as policy making, spatial planning, flood prediction and relief, emergency services, environmental assessments and many other applications. Large-scale GI generally refers to geographic datasets (to a scale of approximately 1:1,000) in densely populated areas. The scale of a dataset, its technical characteristics, and type are among the factors that determine the cost of data collection, which can vary significantly. A 1:1,000 dataset with comprehensive content for a complete jurisdiction is expensive compared to a 1:1,000,000 dataset that covers only one type of data for a sub-jurisdiction (Van Loenen, 2006). Also, large scale GI needs to be updated frequently to be useful. Due to the high investment costs, there are only a few private sector enterprises that are able to produce large scale GI. Therefore, producing large scale GI is most often done by the public sector because of the economies of scale. The public sector may also create large scale GI for historic reasons (*e.g.* producing topographical maps traditionally for military purposes).

Large scale GI is usually produced for a specific purpose. Sometimes the public sector body acquires base data from the private sector to produce large scale GI, *e.g.* aerial photographs. These private sector enterprises usually make the data available to the public sector under a licence agreement. After the original purpose has been fulfilled, the public sector geo-information (PSGI) can be (re)used by others, either with or without licence conditions. The largest group of PSGI (re)users consists of other public sector organisations. These organisations will adapt the PSGI again to suit their own purposes. Depending on the original licence conditions, they may or may not make this PSGI available for reuse by *e.g.* the private sector. The private sector can use this PSGI for their own business purposes (*e.g.* soil data for engineering firms) or they can enrich and add value to the existing PSGI for commercial purposes. This last category of companies is known as the so-called value added resellers (VARs) as they create differentiated products and services, both for the public sector and the market. However, VARs will not be able to produce value added products if the purchase price is too high or the licence conditions too strict. Thus a vicious circle can arise: the public sector starts to develop value added products themselves because the private sector is not doing so to a satisfactory extent (Groot *et al.*, 2007).

§ 2.3 Access regimes for public sector geo-information

§ 2.3.1 Open access

There are two funding regimes for financing public sector bodies that produce PSGI. The first model is the so-called marginal costs regime. With this regime PSGI is funded out of general revenue, and then made available for reuse for no more than the costs of dissemination and with a minimum of restrictions. Disseminating information for free with no user restrictions is called an open access model. The philosophy behind this model is that once taxpayers have paid for producing PSGI, the information belongs to the taxpayers and they should not have to pay again to reuse this information. This regime is applied to *e.g.* geo-information of United States (US) federal agencies. The expectations are that with an open access model the knowledge economy will be stimulated, more value-added products will be produced and thus revenue will flow back to the government in the form of taxes such as value added taxes and company taxes (Van Loenen, 2006). With the marginal costs regime the costs are shared by all the taxpayers. However, this funding regime is sensitive to political decisions. If funding for a public sector body out of the general budget is reduced, the update frequency and quality of the datasets may be reduced. Also, there is no guarantee that revenue raised from taxation will be returned to the appropriate public sector body (Longhorn and Blakemore, 2008).

There is another possible hitch with the open access model, especially when a public sector agency decides to switch to an open access model. Making PSGI available may be deemed to be an economic activity, even if it is for free. As such, it may be in breach with national Fair Trade Legislation in some countries as it may constitute an act of unfair trading practices if the private sector already has made vast investments to create similar datasets. The Dutch Department of Public Works ran into a dispute with some geo-companies after the Department made their National Roads Dataset available for free, in line with existing policy. The geo-companies had produced similar datasets for car navigation producers and for emergency services. The Department of Public Works withdrew the dataset after the geo-companies threatened to sue for unfair trading practices because the free National Roads dataset was competing with the fee-based datasets.

§ 2.3.2 Cost recovery

The other regime for funding PSGI is by recovering all costs incurred in production and dissemination of the PSGI from the actual users, i.e. a user-pay system. The fees may include a return on investments. The information is only made available for (re) use under, often restrictive, licence conditions. The pricing model may be a fee per area, subscription fees, fixed access fees, royalties or a combination of these models (Welle Donker, 2009). Providing fee-based access to information is called a cost recovery access model. This model is applied to e.g. data from United Kingdom Trading Funds16 such as the Ordnance Survey (the British Mapping Authority). The advantage of this regime is that all costs incurred in producing the information, are shared by the actual users. Also, the appropriate public sector body can use the revenue raised for updating and improving the information thus guaranteeing continuous high-quality information. However, when the number of likely (re)users is not known in advance, it may be difficult to set reasonable fees based on cost-recovery (Welle Donker, 2009). There is no natural ceiling for prices as the public sector body often enjoys monopolistic advantages. Also, setting fees is complicated because the value of GI depends on many factors and assumptions (Longhorn and Blakemore, 2008). Another risk with this regime is the boundary between public and private tasks is becoming blurred as the public sector body is also a market party.

The funding regimes described above are two extremes on a sliding scale. In the EU most governments employ a form of cost recovery regime for GI. In some countries a mixture of open access and cost recovery regimes is employed, sometimes even within the same level of government.

§ 2.4 European Union legal framework

Until the 1990's, there was no formal framework for marketing PSI. With each country setting their own policies, there was a variety of different policies with a variety of

16 A Trading Fund is an operation of a government department that has been established by a Trading Fund Order in accordance with the Government Trading Fund Act 1973 (as amended by the Government Trading Act 1990). A Trading Fund may be established where a Minister of the Crown judges that the revenue of an operation could "consist principally of receipts in respect of goods or services provided in the course of the operations in question", and that setting one up would lead to "improved efficiency and effectiveness of the management of those operations". Trading Funds are required by statue to principally recover their costs (*i.e.* to recover a majority of their costs) through income derived from operations within the trading fund (Cambridge University, 2008).

fees and user conditions. From about the mid 1990's a general rethink occurred in a number of EU countries. Studies carried out in Europe and the US indicated that PSI would be a rich resource for creating value added products and services produced by the private sector (*e.g.* PIRA, 2000). As such, PSI has a potential economic value worth thousands of million Euros. However, due to restrictions in availability, exploitation of PSI in Europe is lagging in comparison to the US. The potential economic value of PSI in general was estimated to be between 28 and 134 thousand million Euros in 1999 (PIRA, 2000). Similar national studies came up with comparable figures (*e.g.* RAVI, 2000; MICUS, 2001), although other studies came up with more conservative estimates (MEPSIR, 2006; OFT, 2006). Even with more conservative estimates, the potential value ranges from 10 to 48 billion Euros (MEPSIR, 2006).

§ 2.4.1 Creating a level playing field

The current legal framework related to PS(G)I is not so straight forward in Europe. Countries that are members of the EU have to abide to EU Directives and Treaties, national legislation and policies. A number of older EU Member States such as Germany and the United Kingdom already have established national legislation such as a Freedom of Information Act, Fair Trade legislation and Copyright Act, as well as specific statutes such as Cadastre Acts or Anti-Terrorism legislation. Other EU countries, especially the newer Member States from Eastern Europe, may not have such an advanced legislative framework yet. However, by adopting and implementing the EU directives a general EU-wide framework is slowly emerging.

There are a number of Treaties and Directives which attempt to create a level playing field for businesses and to provide access to information within the EU. The Treaty establishing the European Community (EC Treaty), the Aarhus Convention, the PSI Directive, the INSPIRE Directive and the framework for the protection of intellectual property probably contribute most to setting a general framework. A brief description will follow below. There are additional EU Directives and Guidelines which are in some way relevant to PSI access models. This includes, inter alia, legislation relating to the protection of information and of personal data; broadband Internet access; the need for transparency within financial transactions and supervision by government agencies and the establishment of a regulatory framework for electronic communications networks and services. However, these will not be dealt with in this paper.

§ 2.4.2 The Treaty establishing the European Community and the Treaty of Maastricht

The Treaty establishing the European Economic Community of 1957 (in 1992 the name was changed to the Treaty establishing the European Community) provided two fundamental freedoms, namely the freedom of establishment and the freedom to provide services. After incorporation of the EC Treaty into the Treaty of Maastricht in1993[17], the number of fundamental freedoms were extended to four, namely (1) free movement of goods; (2) free movement of persons, including free movement of workers and freedom of establishment; (3) free movement of services; and (4) free movement of capital. Both treaties seek to establish a level playing field for a European internal market. These fundamental freedoms are further specified in various directives and guidelines. The Treaties also deal with aspects such as State Aid in order to set a rough framework for governments and agencies when competing with the private sector.

§ 2.4.3 The PSI Directive

The 2003 Directive on the reuse of Public Sector Information (2003/98/EC), the so-called PSI Directive, was established in order to set a general framework for governing the reuse of public sector information and to ensure fair, proportionate and non-discriminatory conditions for reuse. The objectives of the PSI Directive are twofold: 1) to provide access to and use of public sector information as an important ingredient for EU-residents to be well-informed and to participate in the democratic process; and 2) to facilitate the creation of Community-wide information products and services based on public sector information and to enhance the effective cross-border use of public sector information by the private sector in order to create value-added information products and services. The PSI Directive cannot enforce publication or reuse of information. The decision to authorise reuse remains with the Member State or the public sector body concerned. The PSI Directive does stipulate that information should be made available in electronic formats as much as possible. The PSI Directive leaves IPRs unaffected. A public sector body may continue to use licences and/or charge fees for reuse of PSI if they were already doing so in the past. Where charges are made, the

17 The Treaty of Maastricht consolidated a number of older treaties related to various European Communities that were forerunners of the European Union. Since then, the Treaty of Maastricht was amended to some extent by the Treaty of Amsterdam (1999) and the Treaty of Nice (2003). The Treaty will most likely be amended again in the near future when the Treaty of Lisbon (signed in December 2007) will be ratified, although the target date of January 1, 2009 was not met.

total income should not exceed the total costs of collecting, producing, reproducing and disseminating documents, together with a reasonable return on investment. Unfortunately, what exactly is deemed to be a reasonable rate on investment is not specified in the Directive. Any conditions applicable to reuse and charges must be pre-established and published through electronic means where possible. Upon request, a public sector information holder (PSIH) has to give an account of how the charges were calculated and which costs were taken into account. The PSI Directive does not deal with redress issues, leaving that to individual Member States.

§ 2.4.4 The INSPIRE Directive

Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), adopted in 2007, intends to establish a common framework for annotating and sharing geographic data between Member States, thus setting a framework for a geo-information infrastructure (GII). The Directive emphasises the environmental reasons to share data between official agencies in different EU countries, rather than focusing on access to that data as a way of promoting wider cross-border usage of geo-information. This INfrastructure for SPatial InfoRmation in Europe (INSPIRE) will be based on (N)GIIs created by Member States that are made interoperable with common implementing rules. The Directive applies to all PSGI used for carrying out public tasks. The INSPIRE Directive leaves IPR claims and the PSI Directive unaffected as far as access regimes and charges are concerned. However, it should be possible to at least view information without incurring fees. As far as INSPIRE is concerned, it will be necessary to facilitate access to PSGI that extend over national or administrative borders, in order to stimulate the development of value-added services by third parties. This should be achieved by developing technical standards to improve cross-border interoperability. Although INSPIRE describes all environmental information to be included in a NGII, it foresees a limited number of policy domains in which specific risks can occur when disclosing certain information, *e.g.* bird breeding grounds on military sites. The INSPIRE Directive has yet to be transposed into national legislation with the first step due in May 2009.

§ 2.4.5 Copyright framework

Intellectual property is divided into two categories, namely industrial property (trademarks, patents, trade secrets) and creative works (copyright and related rights, database rights). Copyright was originally conceived as a way to restrict printing

by granting exclusive rights to make copies. Nowadays copyright should provide an incentive for the creation of, and investment in, works such as music, films, print media, software, and their economic exploitation. There is no EU Directive establishing copyright as such as Member States already had established national Copyright Acts. The EU Directive on the harmonising of certain aspects of copyright and related rights in the information society (2001/29/EC), the so-called Copyright Harmonisation Directive, merely harmonises terms of copyright protection within the EU. The Copyright Harmonisation Directive specifies the exceptions and limitations to the rights. The Directive also adapts the existing framework to reflect technological developments and allows digital rights management to control access to works. The Copyright Harmonisation Directive implements the framework of the World Intellectual Property Organisation (WIPO) Treaties of 1996. However, the Copyright Harmonisation Directive leaves Member States national legislation unaffected[18] (see item Copyright Changes).

Copyright Changes

The European Commission announced in July 2008 that some more changes will be made to copyright legislation, mainly to bring performers' protection more in line with that already given to authors. The European Commission also released a Green Paper on Copyright in the Knowledge Economy. In this Green Paper the Commission has highlighted the need to promote free movement of knowledge and innovation in the EU single market. According to the Green Paper, the free movement of knowledge and innovation should be considered to be the fifth fundamental freedom in the EU. The Green Paper will now focus on how research, science and educational materials are disseminated to the public and whether knowledge is circulating freely in the internal market. The consultation document will also look at the issue of whether the current copyright framework is sufficiently robust to protect knowledge products and whether authors and publishers are sufficiently encouraged to create and disseminate electronic versions of these products (Commission EC, 2008)

18 The Copyright Harmonisation Directive harmonised a number of aspects of existing national copyright acts but did not alter the spirit of the existing national copyright acts with respect to data protection, conditional access and access to public documents, which may affect the protection of copyright or related rights (recital 60). (Footnote added October 28, 2016).

§ 2.4.6 Database directive

Europe, unlike the US, has recognised that creating databases requires vast investments. But databases are not subject to copyright protection as databases fail to comply with the creativity requirement. Some EU countries already had incorporated a "sweat of the brow" doctrine in their Copyright Acts, *i.e.* having invested a substantial amount of resources to produce a work like a database, the creator could claim copyright. The 1996 Directive on the legal protection of databases (96/6/EC) established a sui generis[19] right granting a 15 year protection period from date of publication or completion. Any change which could be considered to be a substantial new investment will lead to a new 15 year term. A database is defined as "a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means". A database may contain all sorts of works or materials. The contents are described as "information" in the widest sense of that term (EU, 1996). Database rights prevent the unauthorised extraction and reuse of the entire or substantial part of the contents of the database. Since most GI is stored in some form of database and these databases are continually updated, the protection period is almost perpetual.

The objective of the Database Directive was to encourage investment in the information industry by providing protection from copying. However, the protection provided by the Database Directive has had an anticompetitive effect on the information market (Hugenholtz, 2005). In effect, all databases are prevented from (re)use because of the ambiguity of terms like "substantial". Even government bodies claim database rights so licence restrictions and fees for reusing PSI can be imposed. In recent years, the EU national Courts, by adopting the Spin-Off Doctrine, have given some clarity as to when a database may be protected. The Spin-Off Doctrine questions if the requirement of "substantial investment" is fulfilled when the database is generated as a by-product of other activities (spin-off), *i.e.* a database can only invoke rights if all investments are made solely to produce that specific database. The mere fact that substantial costs were made to collect the data is not enough to invoke protection under the Database Directive (see item Spin-off doctrine).

Spin-off doctrine

Public sector bodies regularly claim database right to recoup investments made for producing public sector databases. Some national courts in the EU have interpreted the substantial investment test in such a way that it rules out investment in "spun-off"

19 Sui generis means "of its own kind" in Latin.

databases (i.e. databases that are created to support its own operations or that are created as a result of these operations but not created as a core activity), the so-called spin-off doctrine. On November 9, 2004 the European Court of Justice (ECJ) had to rule in four closely related cases brought before it by a number of national courts. The ECJ confirmed the spin-off doctrine and thereby denied protection to producers of single-source databases. Only if the database in question was produced with the sole purpose of commercial exploitation, can database right be invoked, see, e.g. British Horseracing Board v William Hill (EC) joint cases C-46/02, C-338/02 and C-442/02).

The ECJ ruled in cases against private sector and semi-public sector operators but the spin-off doctrine is also applicable to public sector organisations. In the Netherlands, the spin-off doctrine was confirmed by the District Court of Amsterdam on February 11, 2008 in the case of the Municipality of Amsterdam v Landmark Ltd. Landmark Ltd, a private company, had requested a file pertaining to soil pollution under the Freedom of Information Act. Initially the Municipality of Amsterdam refused to make the file available, claiming it was not public information. After Landmark Ltd lodged a formal complaint about breaching the Freedom of Information Act, the Municipality of Amsterdam decided to make the file available after all but charged a hefty fee by invoking database rights. Landmark Ltd sued the Municipality of Amsterdam claiming that database rights were not applicable. The District Court of Amsterdam ruled that a government or public sector body could not invoke database rights because the investments made to produce the database had not carried a substantial risk as such, even though the Municipality of Amsterdam had made a considerable investment to create the file. The soil database had been produced with public money for a specific public task, and not for commercial purposes (Amsterdam District Court, reg. no. LJN BG1554). The Municipality of Amsterdam lodged an unsuccessful appeal as the Council of State, the highest Dutch Court of Appeal for Administrative Law, upheld the District Court's decision on April 29, 2009 (Raad van State, case nr. 200801985/1).

§ 2.4.7 The Aarhus Convention

The Convention on Access to Information, Public Participation in Decision making and Access to Justice in Environmental Matters, was adopted in Aarhus, Denmark, on 25 June 1998. The Aarhus Convention is a United Nations Economic Commission for Europe (UNECE) environmental agreement and links environmental rights to human rights. It links government accountability and environmental protection. The Aarhus Convention specifies that governments should not only grant passive access to environmental information (giving access to information after an application has been lodged) but also active access (publishing reports, environmental registries, et cetera). The INSPIRE Directive recognises these principles and have adopted similar

terms. Although most European countries have ratified the Aarhus Convention, they have adopted different interpretations. Some countries are setting up websites or web services showing environmental information. Some governments are using the Aarhus Convention as a lever to chance existing access policies for environmental information. The Norwegian Government passed legislation making all environmental thematic information available for free. The Dutch government is in the process of setting up a web service which will allow viewing and combining information related to one's direct environment for free. This web service will include PSGI that is currently fee-based.

§ 2.5 Obstacles to accessibility

In spite of the EU framework there are still obstacles to accessibility of PSGI. PSGI is difficult to find as it is scattered throughout different public sector organisations. Often public sector organisations claim IPRs to maintain control over (re)use of PSGI. Each organisation applies its own licence conditions and pricing regime. A survey of PSGI licences in the Netherlands in 2006 revealed that most PSGIHs employ a wide variety of licences, all vastly different in length and phrasing. The licences varied from a couple of paragraphs in plain language to dozens of pages in legalese. The restrictions varied from only having to attribute the source, to having to supply a fully developed business plan showing what the user intends to use the data for. The fees also varied from free to hundreds of thousands of Euros for large scale land covering datasets (Welle Donker and van Loenen, 2006). It is this inconsistency and intransparency in user conditions that forms one of the biggest obstacles for VARs in their decision to (re)use public sector geo-information for their activities (see Groot et al., 2007; STIA, 2001; RAVI, 2000). Other obstacles frequently mentioned by VARs are unfavourable pricing and restrictive licence conditions (see e.g. MICUS, 2008). As a consequence, value-added use remains limited.

Another obstacle to reuse of PSGI is that some public sector organisations will act as a VAR themselves by combining and enriching their datasets, and promoting these in the market. After the privatisation and unbundling wave of the last decade or so, a number of public sector organisations have become (semi-)private enterprises that are required to recover their operating costs. These organisations are also often PSGIHs such as the British Ordnance Survey. In some cases the geo-datasets were part of a privatisation "dowry". Thus the original costs of collection and creation are reduced to zero, leaving only ongoing costs for maintenance, development and dissemination. Because of the cost recovery requirements, their GI is traded as a commodity with user restrictions. So, not only does the private sector find it hard to obtain GI from the public sector, they may also have to compete with the same public sector that may enjoy advantages private sector enterprises do not have. This may constitute distortion of the internal European market.

§ 2.6 PSGI availability in Europe

Although all EU Member States have to abide by the PSI Directive, there are still quite some differences with respect to access and licence conditions. Information regarding Nord Rhein Westfalen (Germany), Norway, France and the United Kingdom was collected as part of a study (Van Loenen *et al.*, 2007). Information regarding the Netherlands was collected as part of earlier research by the author. In this chapter a brief summary of access policies of these countries will be provided.

§ 2.6.1 North Rhine Westphalia (Germany)

§ 2.6.1.1 Background

Germany is a federal republic with 16 States that have a high level of autonomy. The German federal government acknowledges the economic, political and societal importance of the availability of GI. The federal programme Deutschland on-line has incorporated the GII, the so-called GDI-DE. Implementation of GDI-DE at the federal level is coordinated by the Inter-Ministerial Committee for Geo Information (IMAGI). IMAGI is supported by the GDI-DE Steering Committee and set about developing collaborations with the private sector and academia. IMAGI is now responsible for developing and operating a meta-information system as part of a federal geo-portal. Each German federal authority or agency currently defines its own data policy on a case-by-case basis under the direction of the appropriate Minister. The GDI-DE Steering Committee and IMAGI are – directly or indirectly – working towards the development of a harmonised and simplified licensing framework and a comparable pricing regime for GI (SADL, 2008).

Each of the 16 states in Germany is responsible for its own topographic service, land and property register, environmental and statistical information collection, and in general for information policies. Information collection is largely decentralised and carried out mostly on the regional and local level. The different states have issued laws ('Surveying and Cadastral Acts') that regulate the work and the mandate of the surveying and mapping authorities, including defining the production of cartographic material as a public task. With regard to GII development, the developments of the GDI-NRW is closely watched by other states and IMAGI, as it may be an example for other state GIIs and GDI-DE.

North Rhine Westphalia (NRW) is one of the 16 states in the west of Germany and borders the Netherlands, Belgium and France. It covers about 34,600 km^2 and has a population of over 18 million. Since March 2005 there is an Act stipulating that all PSI must be available for sharing between all levels of government and agencies. The government structure has three distinct levels of public authority: national, regional and local, all of which generate and hold PSGI. The levels are organised as follows: at the national level a State government; at the regional level 5 Regierungsbezirke (larger districts) and 54 Kreis government (small districts); at the local level Gemeinden (municipalities). In NRW small-scale topographical information (*e.g.* 1:10,000) is the responsibility of the State Topographical Service. The Kreisen are responsible for large-scale geo-information (*e.g.* 1:1,000). Municipalities are users and the Regierungsbezirke will oversee that the Cadastre Reform Act is adhered to and will assist the Kreisen on a technical level. A Kreis cannot collect its own taxes and is financially dependant on the State (income and property taxes) and Gemeinden (company tax).

§ 2.6.1.2 Access to PSGI

Access to PSGI is largely controlled by the Cadastre Reform Act and corresponding legislation. GI not covered by the Cadastre Reform Act, the so-called non-geo base data, *e.g.* aerial photography of the districts, is covered by local policies. All local governments claim copyright and database rights in their information and only grant a "limited use" licence for reuse. Use of geo-base data is free within the public sector. Other users pay a fee based on cost recovery regime. There are different tariffs depending on the format, category of the layers, size of the area required and information density. Different types of users also pay different fees. The pricing structure as set down in the Tariff Regulation is complicated and difficult to understand. Also, prices can be quite steep: a copy of the ALK (Automated Property Map) covering entire NRW amounted to about €3,400,000 in 2006. The private sector has indicated that the Tariff Regulation's complexity is one of the main obstacles to reusing PSGI. Also, the Tariff Regulation is too inflexible to be of use for web service applications (MICUS, 2003).

Because of the barriers reuse of PSGI for developing value added products and services by the private sector remains limited. Some of the Gemeinden, like the City of Aachen, have developed value added services to fill the gap. The Cadastre Reform Act does have a clause which allows experimental use of geo-data. This allows the State government to provide private companies with free access to explore the possibilities of PSGI. If a product appears successful then the free supply of PSGI will be stopped and a contract will be negotiated. An example of one experiment was *e.g.* www.mySDI.com by Con Terra and Vodafone. However, PSGI is mostly used by other public sector organisations and semi-public sector organisations such as utilities. Another problem

for VARs in NRW is access to thematic data. Socio-economic data are not available from one single access point and are therefore harder to obtain. In addition, as production of topographical information is defined as a public task, the State Surveying Authority considers creating spin-off services such as leisure maps also to be a public task (MICUS, 2008).

Some Gemeinden and Kreisen provide on-line access to PSGI via Web Mapping Services (WMSs) but they are not obliged to do so. The State government provides online access to its topographic and cadastral information via a web service called TIM-online (www.tim-online.nrw.de). Private use of the web service is free but downloading the reference information is illegal. A user can view information via a WMS. The user can also merge further geodata via a Web Feature Service (WFS).[20] Due to the popularity of TIM online and feedback provided by users, the update frequency of TIM online has increased from annually to fortnightly. In addition, the popularity of TIM online has raised awareness of the value of GI at the decision making levels, although this has not resulted (yet) in major policy changes or additional finances.

§ 2.6.2 Norway

§ 2.6.2.1 Background

Norway is a mountainous long stretched country with an extensive coastline of over 2,000 km and an area of 307,000 km². Norway is part of Scandinavia and is located in the north-west of Europe. Norway is a monarchy with a State government, 19 counties (both as regional units of the state government and as a local government) and 431 kommunes (municipalities). Most of its population of 4.6 million reside in the southern part and is otherwise less populated. Norway is not a member of the EU but has strong ties with the EU. Therefore Norway adheres to general EU policy and implements most European Directives, probably even faster than most Member States. However, implementation of the PSI Directive took longer because it was tied to a renewal of the Norwegian FoI Act. The PSI Directive is now implemented in the

20 There are many technical differences between a WMS and a WFS. The main difference is that with a WMS an image is generated on screen from raster data but no actual data transferred to the user, whereas with a WFS actual data is transferred to the user. WMSs are often used for free web services because the image generated is of a low resolution. WFSs are used for vector data so that the data can be manipulated and analysed. Because features of the data are transferred to the user, WFSs are most often used for fee-based services.

Act on the right to access to objects in the public sector (public law), which came into effect on 1 January 2009. The new Act sets an upper limit for pricing of public sector information by stipulating that that the right to take a profit can only be used in special cases (http://www.epsiplus.net/news/psi_re_use_innovation). The Norwegian Ministry of Trade and Industry released a White Paper on 5 December 2008, in which it re-stated its commitment to establish favourable conditions for wealth creation based on sound solutions in the public sector and the increased use of public data as a driver for innovation (Norwegian Ministry of Trade & Industry, 2008). In Norway, it is generally accepted that thematic GI is freely available. For environmental information, this has been enshrined in domestic Norwegian law since 1993. Both the State and local government have such data available on-line. Often this data is only on-line in raster formats but upon request it is possible to obtain the vector version as well. This principle seems to precede the Aarhus Convention (Van Loenen *et al.*, 2007).

§ 2.6.2.2 Access to PSGI

Within the public sector several organisations handle GI. The Norwegian Mapping and Cadastre Authority (Statents Kartverk SK), residing under the Ministry of Environment, is responsible for the coordination of the Norwegian GII. In 2003, a White Paper authorised GI sharing within the public sector by setting up a GII. This program, called Norge Digitalt (Digital Norway, www.GeoNorge.no), provides not only a portal but also a framework for cooperation within the public sector. Nearly all state departments and agencies, as well as local governments, have joined or are in the process of joining Norge Digitalt (ND). After paying a contribution, the government organisation then makes its GI available free of charge to other participating organisations. The contribution paid is related to the importance of base geo-data and the size of the organisation. Within ND all participants can use free GI for its own internal business processes. More than 30 state and almost all local government organisations are a member of ND. For historic reasons, some private sector organisations are allowed to join ND (see Figure 2.2).

If the private sector wants to use PSGI, it can buy datasets from a government-owned intermediary, the Norsk Eiendomsinformasjon (NE). The NE acts as a one-stop shop for VARs to get the data and resell it to end-users. A contract is drafted with the NE and NE pays royalties to ND. NE uses the same (restrictive) licence conditions for all information it resells. However, there are some unresolved issues with this system. As part of the decision to let SK coordinate the ND, the marketing activities of SK were sold off. A private firm, Ugland IT, now has an exclusive right to produce certain map series. SK is not allowed to sell its own GI to the private sector, as this was handed over to NE. However, other members of ND are still allowed to market their own GI. Several public sector organisations provide this GI for free through WMSs.

Until 1 January 2007, all SK services were freely available on the web. To be in line with the access policy from the 2003 white paper, SK had to limit free access to ND partners only. NE does not have a publicly known pricing policy. In order for ND to operate more transparently, GI should be made available to outsiders under clear and equal conditions. NE was set up as a one-stop shop for VARs and distributors but is increasingly selling to end-users as well. By doing so NE acts more and more as a market party, thus blurring the separation between public and private sector. Because there is no legal framework for ND as such (only a white paper) there are no clear boundaries.

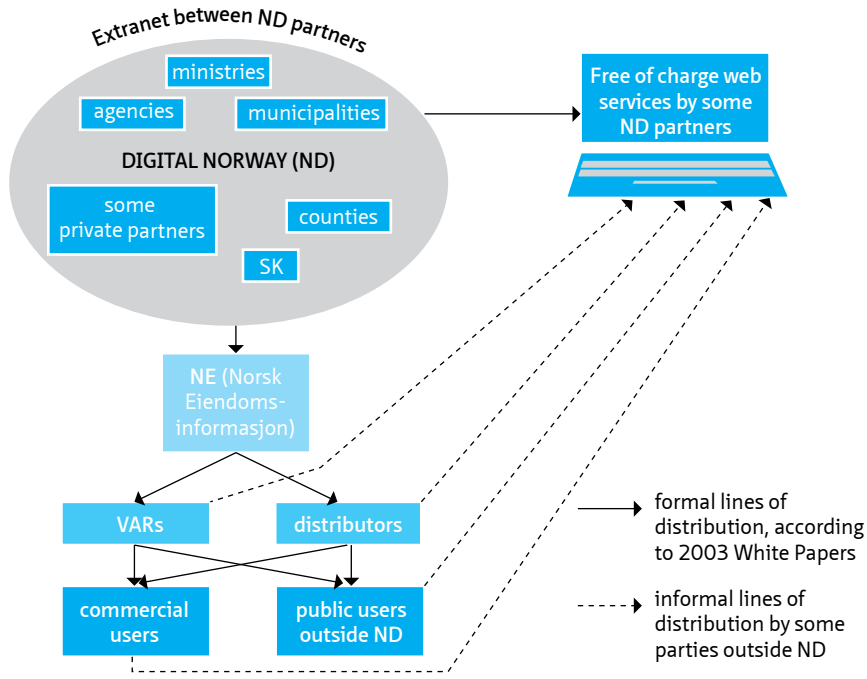


FIGURE 2.2 Norway Digital access model, formal and informal lines of distribution (Welle Donker, 2009)

TOC

§ 2.6.3 France

§ 2.6.3.1 Background

The Republic of France is the largest country in Western Europe. Mainland France (excluding overseas territories of the French Republic) has an area of approximately 543,965 km² and a population of circa 65 million. France is governed by a centralised government, presiding over 22 Regions that are further subdivided into 96 Departments. These Departments are then further divided into Arrondissements and Communes. Most of PSGI is collected and used by these administrative divisions. Designing a common access policy in France is not so simple. The administrative divisions, especially the Communes have a high level of autonomy. Thus, a top-down approach has to be carefully implemented as the Communes cannot be compelled to adopt a Central Government policy, they can only be asked to participate in the interests of the Republic. A number of initiatives have commenced in order to modernise the French government's approach to access to (national) PSI and services to citizens. One of those initiatives is the Direction Générale pour la Modernisation de l'Etat (DGME) initiative which was launched in January 2006. The Ministry of Public Works, Infrastructure and Land Planning is now working on an intranet geo-catalogue/ geo-portal system for internal Ministry usage with a view to making this service available to other ministries in the future.

§ 2.6.3.2 Access to PSGI

Within the DGME initiative, Geoportail has been set up as the main PSGI portal (www.geoportail.fr). There are three organisations responsible for the implementation and maintenance of Geoportail. The overarching organisation is the DGME, since Geoportail is a part of the DGME initiative. The DGME is responsible for coordinating the policies necessary to ensure that public sector bodies (and where possible local governments and the private sector) make their data available to Geoportail. The Ministry of Geology (BRGM) is the second organisation responsible for the implementation of Geoportail. BRGM's role is to design, implement and maintain the catalogue component (Le Geocatalogue) of Geoportail. With the catalogue function, datasets can be located. The third organisation involved in Geoportail is the Institut Geographique National (IGN). IGN's function is to implement the other main component of Geoportail, the visualisation component (the Visualiser). With the Visualiser, datasets can be viewed and downloaded. Viewing is free of charge but only custodians of the datasets can download data for free. Other parties like the private

sector can download data on a subscription basis. With an API, Geoportail is available for the private sector to upload their own information. Geoportail is envisaged to become a community-oriented and development platform (IGN, 2008).

Since its inception in July 2007, Geoportail has attracted millions of viewers with numbers now hovering around 1.2 million users per month (IGN, 2008). Most of the datasets accessible through Geoportail belong to BRGM, IGN and some partners and contains topographical, cadastral, hydrographic and thematic information, and historical maps. The Visualiser allows 2D and 3D viewing, rivalling private sector platforms such as Google Earth in speed and performance. Thus, Geoportail far exceeds the requirements of INSPIRE. To increase the performance, images are stored as tiles on the server(s) in advance, requiring Terabytes of storage capacity. Geoportail requires 3 Gbps broadband capacity, two 50 Tb caches and a 100 Tb storage capacity (IGN, 2008). Although Geoportail is set up to make PSGI accessible for reuse by both the public and the private sector, it is unclear to what extent revenue through downloads will help to recover the costs of development (circa 6 million Euros) and the annual operating costs (circa 1.5 million Euros). Also, as the lower governments cannot be compelled to participate, the success of Geoportail will depend on their willingness to make their datasets available. Funding will have to be made available to the lower governments to make their data compatible to Geoportail. Already a number of the local authorities have their own web services to provide access to local PSGI. Linking their websites to Geoportail may produce volumes of traffic that these sites were not designed to handle (Van Loenen, 2007).

§ 2.6.4 England and Wales (United Kingdom)

§ 2.6.4.1 Background

The United Kingdom (UK) is an island nation in north-western Europe located between the Atlantic Ocean and the North Sea, to the west of France, Belgium and the Netherlands. The total area of the UK is circa 245,000 km² and its population is nearly 61 million. The UK is a constitutional monarchy and is centrally governed by a national government. Furthermore, there are three Executives (the governments of Northern Ireland, Scotland and Wales), and a complex system of local government. England, the largest country of the UK, has no devolved executive and is administered directly by the UK government on all issues. There are nine Government office regions, each further divided into boroughs, counties, district councils and unitary authorities, about 500 in total. Policy decisions are made by the central government and their agencies.

Local governments are mainly responsible for local planning and everyday operations of their areas. The larger local authorities, such as the City of London, have a greater autonomy. The Executives of Scotland and of Northern Ireland have strong levels of independence. The Welsh Executive has more limited powers. For this paper England and Wales are combined as their access policies are very similar.

In the UK, there are different copyright regimes applicable to GI. The main copyright law affecting PSGI is the Crown Copyright. Crown Copyright applies to PSGI produced by central government agencies referred to as Crown Bodies. However, it is not always easy to distinguish which public sector organisations are Crown Bodies and thus affected by Crown Copyright because of technical legal reasons (APPSI, 2004). Therefore different central government agencies will have different copyright regimes regulating their information, resulting in different rules for reuse.

§ 2.6.4.2 Access to PSGI

Because of the centralised structure, the central government and its agencies require access to detailed information at both local and national level. The public sector is therefore the biggest producer of information. To support the service-orientated market, the UK government has implemented a number of initiatives to encourage the use and reuse of PSI. These are:

- the promotion by the Cabinet Office of the reuse of PSI to enhance the knowledge economy and the quality of government in the UK;
- the initiatives of HM Treasury to leverage PSI to generate revenue and reduce the cost of government;
- the Efforts by the DCA to promote transparent government through the Freedom of Information Act, and
- the DTI efforts to enhance the competitiveness of the UK information sector and the join-up government policy (APPSI, 2004).

However, some of these initiatives show conflicts of interest with each other (APPSI, 2004). In 2006, as part of a general review, the Advisory Panel on Public Sector Information (APPSI) had its mandate changed to a non-departmental public body of the Ministry of Justice to – among other things – review and consider complaints related to reuse of PSI.

Most PSGI is generated by the Ordnance Survey (OS), although other parties like the United Kingdom Hydrographic Office (UKHO), Her Majesty Land Registry (HMLR) and the Royal Mail Group are also active. OS, UKHO and HMLR are all classified as Trading Funds and are required to generate a surplus. Therefore, these agencies all use

restrictive licence conditions and fees to make their datasets available for reuse. There is no single access policy for PSI in the UK. UKHO use a network of VARs which reuse hydrographic information on a royalty basis. OS also have licence agreements with various VARs on a royalty basis.

As far as reuse within the public sector is concerned, OS uses a system of Collective Licensing Agreements (CLAs) to make their PSGI available to other public sector organisations. A CLA is a contract between OS and a group of public bodies whereby access is given to OS information for a set fee. There are at least four distinct CLAs between OS and the public sector. These are:

1. the Pan-Government Agreement (PGA). This is a contractual arrangement between the OS and Central Government Agencies;
2. Mapping Services Agreement (MSA). This is the contractual arrangement between OS and Local Government Agencies for the provision of GI;
3. London Government Agreement (LGA). The contractual agreement between the Local Government Authority of London and OS for the provision of GI; and
4. National Health Services Agreement (NHSA). This a blanket agreement amongst the different health sectors of England and the OS for the provision of GI.

The advantage of a CLA is that participants collectively only have to negotiate once with OS to get quick access to high quality information. However, the information may only be used for internal purposes. The public body concerned is not even allowed to place the information on its website. Within a CLA there may be sublicences for large scale and small scale GI. Central government agencies with different sublicences are not allowed to share OS information.

In the UK there is no central portal for PSGI but the major suppliers of PSGI offer GI web services with – where applicable – click-through licences. On-line access can be obtained to OS and UKHO datasets via their websites but the access is not open to the general public, only to business partners. There are GI web services that are freely accessible to the general public for viewing such as GI Gateway (www.gigateway.org.uk). GI Gateway is a free web service aimed at increasing awareness of and access to GI in the UK.

§ 2.6.4.3 Implementation of the PSI Directive

The PSI Directive was implemented in the UK in the form of the Re-use of Public Sector Information Regulations 2005 (the Re-use Regulations), dealing with reuse of government documents. Although the term "document" is broadly defined and explicitly includes "any part" of any content (art. 2), the Re-use Regulations do not

apply to a document where supply of the document is not part of a public task (art.5(1) a) or if a third party owns relevant IPR in the document (art.5(1)b). The concept of 'public task' is not defined in the Regulations. The Re-use Regulations were quickly tested when in 2006 a private firm called Intelligent Addressing complained about the way in which OS licenced its address database called AddressPoint (see item Intelligent Addressing v Ordnance Survey).

Intelligent Addressing v Ordnance Survey

Intelligent Addressing (IA), as partner of a joint venture with Local Government Information House Ltd, needed a database called AddressPoint to produce the National Land and Property Gazetteer (NLPG). Local governments can obtain data for the NLPG through the Mapping Services Agreement (MSA) with Ordnance Survey (OS) but IA is not a party to the MSA. IA claimed that OS offered licence terms which unnecessarily restricted competition. OS claimed the database was not a document as defined in the Re-use Regulations because the file contained third party (Royal Mail) proprietary postal coding address file. Therefore OS did not have to abide by the Re-use Regulations. In February 2006, IA lodged a complaint to the Office of Public Sector Information (OPSI), the regulatory body for PSI regulations and Fair Trade schemes, about breaches of the Re-use Regulations. In their defence OS claimed that as Royal Mail held third party IPR, the database was not a document as such. Oddly enough, OS's claim that commercialisation of the information held by OS to be "a core part of its task" was not contested by IA. If commercially marketing of PSI is a public task then the Re-use Regulations should have applied. OPSI ruled in July 2006 that OS had breached the Re-use Regulations. It was then mutually agreed that APPSI would review the findings of OPSI. APPSI ruled in April 2007 that the Regulations did not apply to AddressPoint because Royal Mail held third party IPR. APPSI also ruled that producing value added products was not a public task. Because the Re-use Regulations did not apply, the case was referred to the Office of Fair Trade (OFT).

From about 2007 there has been a marked increase across central government in the level of interest and debate in the reuse of PSI, including a debate about the position of the Trading Funds (APPSI, 2007). Reports like the so-called Cambridge Report (2008) concluded that in most cases a marginal cost recovery regime would be welfare improving and would not have a detrimental effect on the quality of the data. Although OS, UKHO and the Met Office would have to receive additional funding from central government, the benefits would be commensurably bigger (Cambridge Report, 2008). In its 2008 pre-Budget Report, the UK government stated that the Treasury will publish some key principles for the reuse of PSI, consider how these currently apply in each of the trading funds and how they might apply in the future, and the role of the OPSI in ensuring that government policy is fully reflected in practice. For OS, this will involve consideration of its underlying business model (http://www.hm-treasury.gov.uk/prebud_pbr08_index.htm).

§ 2.6.5 Netherlands

§ 2.6.5.1 Background

The Netherlands, located in north-western Europe, is a low-lying densely populated country of about 41,500 km² and circa 16.4 million inhabitants. The Netherlands is a constitutional monarchy with a national government, 12 Provincial Councils, 26 Waterschappen (democratically elected water boards) and 441 Gemeenten (municipalities) as per 1 January 2009. The lower governments have a fairly high level of autonomy enshrined in legislation. Politics and governance in the Netherlands are characterised by an effort to achieve broad consensus on major issues. Therefore, the process of policy forming and governance may appear slow but generally, final outcomes are broadly supported by all parties involved. The Ministry of Housing, Spatial Planning and the Environment (Volkshuisvesting, Ruimtelijke Ordening & Milieubeheer (VROM)) is responsible for coordinating GI and the establishment of a NGII. Most of the PSGI is collected and used by lower levels of government although VROM, some other Ministries and their related agencies hold large scale base datasets. Some of these PSGI agencies, such as Kadaster (Netherlands Cadastre, Land Registry & National Mapping Agency) and National Co-operation Large-Scale Base Map of the Netherlands[21] (LSV GBKN), are public sector enterprises, *i.e.* they are self-funded public bodies that generate revenue from sales of their products and services. Other PSGI agencies such as the Department of Public Works are funded out of consolidated revenue. Lower levels of government are self-funded through levies and rates, and receive subsidies from the national government for delegated tasks.

§ 2.6.5.2 Access to PSGI

Until the 1990 there was no overriding policy for access to PSI or government bodies engaging in market activities. After many complaints from the private sector about unfair trading practices by enterprising public sector organisations, an inquiry was held in 1995. This inquiry resulted in a policy document in 1998, the so-called Guidelines for Economic Activities by National Public Sector Bodies (Guidelines),

21 Members of the National Co-operation are the Federation of Energy Providers; Kadaster; KPN (former public and still largest telecom provider in the Netherlands); Union of Waterschappen; the Association of Water Providers; and the Association of Municipalities. In association with The Department of Public Works the LSV GBKN produces and maintains the most detailed large scale base map of the Netherlands.

pending formulation of overarching legislation. The Guidelines state that a national public sector body may only engage in economic activities if the private sector will not or cannot (due to *e.g.* security reasons). If a public sector agency engages in economic activities, then all costs incurred in collecting, processing and disseminating must be passed on to the customer and the agency must pay all due taxes (VAT, etc.). The Guidelines only apply to national public sector bodies not covered by specific legislation. Lower levels of government do not have to abide by the Guidelines. Some national agencies are governed by specific legislation with varying mandates. For instance, Kadaster – as a self-funded public sector enterprise – is allowed to employ a cost recovery regime and may produce value-added products from its own data as enshrined in the Cadastre Act. This means that the PSIHs of the more desirable datasets such those of Kadaster and the municipalities are not covered by the Guidelines. Also, the Guidelines only have the status of pseudo-legislation. In the few (lower) court cases where breach of the Guidelines was contested, the courts have set the Guidelines aside. The overarching legislation, although rewritten a number of times, has not proceeded beyond the draft stage to date.

Access to PSI in the Netherlands is covered since 1991 by the Freedom of Information Act (FoIA). The FoIA provides for access to public information, *i.e.* all information within government except information relating to national security, the security of the Crown, trade secrets, and information covered by privacy legislation. The general pricing regime is dissemination costs only. PSI covered by specific legislation, such as by the Cadastre Act, is subject to its own pricing regime. The dissemination costs regime also does not apply to data for which the policy line would result in financial problems for the supplier of the information. The FoIA was amended in 2006 when the PSI Directive was implemented as a separate Chapter, 5A, in the FoIA. Chapter 5A stipulates that for reuse of PSI subject to IPR the total income out of supply of information should not exceed the costs of collection, production, reproduction and distribution, increased by a reasonable return on investments. With the ever-decreasing blur between access to PSI and reuse of PSI in a web-based environment, the duality of pricing regimes in the FoIA[22] is confusing to both the public and the private sector. For national public sector bodies there is an additional clash between the policy line of no more than dissemination costs and the earlier mentioned Guidelines, which state that all costs made must be passed to customers. Provincial Councils and Waterschappen adopted the dissemination costs regime around 2006. Municipalities, however, use a variety of cost regimes. The larger municipalities, such as Amsterdam and Rotterdam, use full

22 The FoIA is currently under review again and it is expected that all information covered by the current Act will be made available for dissemination costs only, unless it is a threat to the direct revenue of a public sector organisation. Although the amendment will not affect the pricing regime of most national public sector enterprises, the amendment will affect the pricing regime of the municipalities. The amendment was adopted by the Lower Chamber on 24 March 2009 but still has to be passed by the Upper Chamber.

cost recovery regime for making their GI available because they have to finance their surveying departments. Most PSGIHs with a cost recovery regime basis, market their GI for area-based pricing or on a subscription basis. The only exception is the Dutch Hydrographic Service which markets its GI to a set number of VARs on a royalty basis.

In the Netherlands there is a portal for all government information, but only for administrative documents such as copies of legislation (www.overheid.nl). There is no NGII as such, although serious efforts have been undertaken in the past to establish one. Currently – as part of INSPIRE requirements – Geonovum, the Dutch NGII Executive Committee is in the process of setting up a geo-catalogue service as precursor to an NGII. At the moment if one wants to find specific PSGI one still has to muddle through search engines. Most PSGIHs have their own web services, usually offering (samples of) PSGI free for viewing. Downloading is usually only possible after a paper contract has been signed.

§ 2.6.5.3 Base Registers

The Dutch national government is in the process of establishing a system of base registers. The idea is that authentic public information is only collected once and reused many times. For instance, municipalities will be responsible for maintaining a single register for residents and addresses in its district. These 441 municipal registers are then combined into one national register. Other governmental bodies at all levels must reuse data from that register so that citizens do not have to resubmit name and address details every time they deal with a public sector body. Municipalities will be responsible for the quality of the data, and other government bodies must report back any mistakes to the municipality. The Dutch government has designated ten base registers so far, another three are nominated and will most likely follow suit. The base registers will include GI datasets such as the 1:10,000 Topographic Map of the Netherlands (TOP10NL), Cadastral Register, Cadastral Map, DINO (data pertaining to the subsoil) and the Large Scale Base Map. The base registers are interrelated, *i.e.* information out of one register will form an essential part of another register. For example, property ownership information from the municipal Buildings & Addresses Register will be combined with the definition of property objects from the national Cadastral Register and type of usage, *e.g.* commercial usage, to form the basis of a Register for Property Values (see Figure 2.3).

As far as financing the roll-out of the base registers is concerned, the national government has made funding available. Future funding for maintenance and quality control of all the base registers is not guaranteed yet. Kadaster, the agency responsible for the TOP10NL, Cadastral Map and Cadastral Register, may continue charging

other public sector bodies for their information[23] even though reuse is compulsory. The base registries are primarily aimed at sharing authentic information between the different public sector bodies. Once fully established, reuse by the private sector may be considered for the public datasets. The base registries will have to be adapted before making them available to the non-public sector so that only aggregated information will be provided. A survey completed in 2007 indicated that the private sector regards base register information as the most valuable resource for creating value added products (Groot *et al.*, 2007).

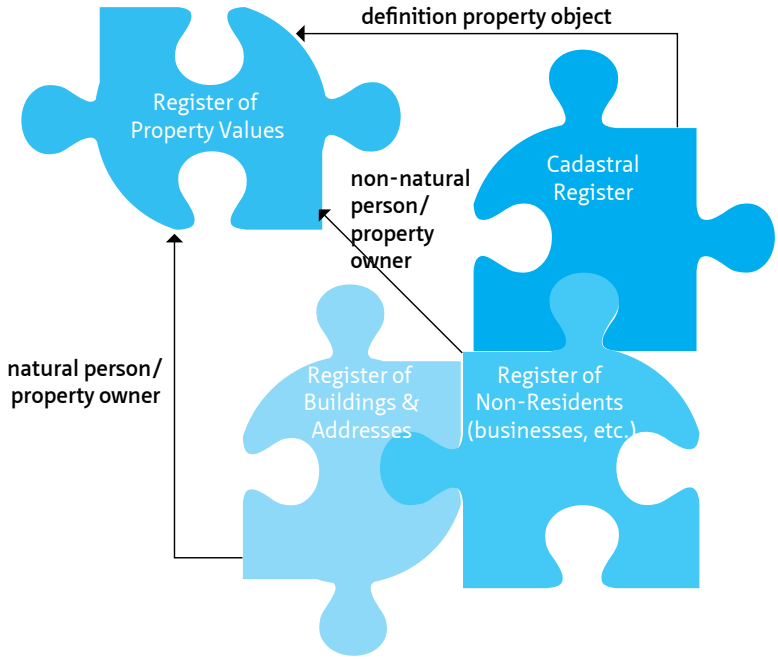


FIGURE 2.3 Interrelationship between Dutch Base Registers (Welle Donker, 2009)

23 In 2008 the Ministry of VROM and Kadaster started negotiation about future funding of their base registers. Although formal agreement still has to be reached, the Ministry will most likely allow Kadaster to charge only dissemination costs and the Ministry will foot the bill for maintenance, etc. so that fees will not be an impediment to other public sector organisations for compulsory reuse. LSV GBKN will receive an additional 7 million Euros annually to allow reuse within the public sector for dissemination costs.

§ 2.7 Conclusion

The EU has tried to promote a level playing field for the private sector by setting conditions for the free flow of information and services. This legal framework includes a number of Treaties and Directives such as the Aarhus Convention, the PSI Directive and the INSPIRE Directive. Different Member States have implemented this legal framework in different ways. Some countries such as Norway and the Netherlands have used the legal framework, including the Aarhus Convention, to make thematic geo-information available for free, at least for viewing purposes. France has taken the requirement of the PSI Directive to make PSI available in electronic format, one step further by setting up a geo-portal rivalling Google Earth. Most Member States use the cost recovery clause of the PSI and INSPIRE Directives to use raised revenue to maintain a continuous level of quality. In most comparisons between the EU and the US, the US marginal cost regime is often lauded as a best-practice example. However, the US marginal cost regime only applies to federal PSGI. It is debatable to what extent the quality of PSGI can be guaranteed if funding is dependable on political decisions. In the US some federal PSGI has not been updated for years. The Dutch Kadaster nearly went bankrupt at the end of the last century. Only by changing its organisational structure to that of an independent administrative agency with a cost-recovery regime could Kadaster guarantee the continuation of services and quality.

The PSI Directive has been in force in the EU since 2003, but transposition into a national framework has taken longer with some Member States only having finished implementation in 2008. The effects of the PSI Directive are slowly starting to emerge, in spite of the fact that awareness of the existence of the PSI Directive among reusers is very low (MICUS, 2008). But the PSI Directive and its evaluation in 2008 show that Member States are now reviewing their pricing regimes and policies. Some Member States are making more PSGI available for dissemination costs only or have reduced their fees significantly. For example, the Austrian National Mapping and Cadastral Agency (Bundesamt für Vermessungswesen BEV) has decreased its prices for digital orthophotos by 97%. Due to the fact that sales volume has increased by up to 7,000%, the total turnover of the BEV has remained more or less stable. New users from small to medium sizes enterprises are now purchasing data from BEV (MICUS, 2008).
The Dutch New Map of the Netherlands (a GIS file containing planning information from all levels of government) had its access regime changed from cost recovery to open access and was made available for free in April 2006. Since then the number of regular users has significantly increased (Welle Donker and Van Loenen, 2006). Thus, by decreasing prices total revenue will in most cases be offset by increases in the number of new users. Especially when the additional revenue to the government in the form of value added taxes, company, income taxes, is taken into account, the total revenue will actually increase in the long term (Van Loenen, 2006).

The PSI and INSPIRE Directives are have been instrumental in improving access to PSGI. In the past users of PSGI have indicated that the biggest obstacles to reusing PSGI was poor accessibility - both in terms of access rights and physical access - inconsistent and non-transparent access policies, differences in pricing, liability regimes and user conditions (*e.g.* KPMG, 2001; RAVI, 2000; PIRA, 2000). Thanks to the PSI and INSPIRE Directives and technological advances, physical access to PSGI is improving. PSGIHs are setting up portals and WMS/WFSs that allow information from different sources to be combined. If those web services are also used to sell downloadable information, care should be taken to ensure that the pricing mechanism does not become too complex to calculate (MICUS, 2003). Setting up geo-catalogues as part of NGIIs is a big step towards being able to find appropriate PSGI.

But there are still some more obstacles for (re)users. The biggest obstacle still appears to be restrictive and intransparent licence conditions. PSGI has little value to users if the information cannot be reused to create new products, either because the licence conditions are unclear or because the user is not allowed to reuse the PSGI. This is not just a problem for VARs which will have to obtain the necessary information from other sources. End-users wanting to reuse PSGI for their personal websites or community platforms may encounter the same problems. Already, community-driven initiatives to develop parallel GI are emerging. One such initiative is Open StreetMap which was originally set up in the UK in 2004 because OS did not allow their data to be reused on community websites. Open StreetMap is a project whereby volunteers go out with GPS units to produce open source street maps for free usage. Open StreetMap now operates in many countries on six continents. Some private geo-companies have donated cartographic information or money to the project as well in return for their data or as a platform for innovative applications (http://www.opengeodata.org/?p=223). Open StreetMap is a prime example of an alternative GI platform purely developed because local PSGI just is not accessible for end-users.

Complicated and inconsistent licence conditions are a particular problem when combining different datasets. The INSPIRE Data and Services Sharing Drafting Team (2008) has come up with a guideline for licence implementing rules, including types of licences and a model for specific licences. Unfortunately this is only a guideline as the implementing rules are not compulsory. The model is a step forward because it addresses issues such as reuse by third parties. The model also contains an Emergency Use clause and a Transparency clause, similar to the transparency clause in the PSI Directive. The Creative Commons system of licensing can also be applied to free PSGI since the Creative Commons does not allow financial gain to be made. Creative Commons also provides a useful template to adapt the licensing framework to fee-based PSGI (Welle Donker and Van Loenen, 2006).

Finally, there is a conflict of interest when public sector agencies act as VARs themselves, especially when in direct competition with the private sector. In the UK,

Trading Funds act as VARs because they are required to recoup their costs. In Germany, production of topographical information is defined as a public task. Therefore, creating spin-off services such as cycling maps are also deemed to be a public task, thus effectively locking the private sector out. In Norway when ND was set up, the SK was forced to sell its marketing activities. But other ND-participants can still sell their own data, making it more confusing for the private sector because of varying pricing and licensing regimes. In the Netherlands, Kadaster is legally mandated to produce value added products and services but only from their own data. Because of its monopoly position Kadaster takes part in many co-operative organisations. Within those co-operations Kadaster produces value added services using non-Kadaster data as well, and then sells those services to third parties. Just as OS does in the UK, Kadaster is pushing the boundaries of its legal mandate.

If there is to be a true free flow of geo-information and geo-services in the EU, there is still a long way to go. The legal framework is paving the way but the devil is in the interpretation into national legislation. Every Member State has its own legacy of PSGI access policies. Concepts like "public task" are interpreted in different ways. What is deemed to be a public task in one Member State is deemed to be a task for the private sector in another. All the EU Member States have different legally mandated PSGI bodies with different cost regimes and different existing policies and legislation. Changing access policies will require extra funding and may also run into unforeseen problems. If a public sector body changes its access policy to unrestricted reuse for free, it may be in breach of national Fair Trade legislation if the supply of PSGI is deemed to be an economic activity. So, even if the Directives are transposed in their most liberal sense, they may still be in breach of existing national legislation. Whilst developing a functioning framework in the EU is a long term goal, legacy systems may slow down the required changes. Although it will take a long time before a level playing field is truly developed, at least the PSI Directive has had the effect that Member States are now seriously looking at and harmonising access policies in the EU. INSPIRE will probably give an additional impetus when it becomes operational.

ABBREVIATIONS	
ALK	Automatisierten Liegenschaftkarte (Computerised Property Map)
APPSI	Advisory Panel on Public Sector Information
BEV	Bundesamt für Vermessungswesen (Austrian National Mapping & Cadastral Agency)
BRGM	The Ministry of Geology
CLA	Collective Licence Agreement
DGME	Direction Générale pour la Modernisation de l'Etat
E(E)C	European (Economic) Commission
ECJ	European Court of Justice
EU	European Union
FoIA	Freedom of Information Act
GI(I)	Geo Information (Infrastructure)
HMLR	Her Majesty Land Registry
IA	Intelligent Addressing
IGN	Institut Geographique National (National Cadastral & Mapping Agency)
IMAGI	Inter-Ministerial Committee for Geo Information
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
IPR	Intellectual Property Rights
LSV GBKN	Landelijk Samenwerkingsverband Grootschalige Basiskaart Nederland (National Co-operation Large Scale Base Map of the Netherland)
ND	Norge Digitalt (Digital Norway)
NE	Norsk Eiendominformasion
(N)GII	(National) Geo Information Infrastructure
NLPG	National Land and Property Gazetteer
NRW	Nord Rhein Westfalen (North Rhine Westphalia)
OFT	Office of Fair Trading
OPSI	Office of Public Sector Information
OS	Ordnance Survey
PS	Public Sector
PS(G)I	Public Sector (Geo) Information
PSGIH	Public Sector Geo Information Holder
SK	Statents Kartverk (Norwegian Mapping & Cadastre Authority)
TOP10NL	Topographic Map 1:10,000 of the Netherlands
UK	United Kingdom
UKHO	United Kingdom Hydrographic Office
UNECE	United Nations Economic Commission for Europe
US	United States
VAR	Value Added Reseller
VROM	Volkshuisvesting, Ruimtelijke Ordening en Milieubeheer (Dutch Ministry of Housing, Spatial Planning and the Environment)
WFS	Web Feature Service
WIPO	World Intellectual Property Organisation
WMS	Web Map Service

References

APPSI (Advisory Panel on Public Sector Information) (2008). Realising the value of Public Sector Information. Annual Report 2007. APPSI (Advisory Panel on Public Sector Information), ISBN 978 0 108 50744 1.

APPSI [Advisory Panel on Public Sector Information] (2004) "1st Annual Report." Her Majesty Stationery Office, London, ISBN 0711504598.

Cambridge Report [Newbery, D., L. Bently & R. Pollock] (2008). Models of Public Sector Information Provision via Trading Funds. Study commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury in July 2007, Cambridge University, Cambridge, United Kingdom.

Commission of the European Communities (2008). Copyright in the Knowledge Economy, Green Paper by Commission of the European Communities. COM(2008) 466/3. http://ec.europa.eu/internal_market/copyright/docs/copyright-infso/greenpaper_en.pdf.

EU (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Official Journal of the European Union. L 108, vol. 50: p.1-14.

EU (2003). Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information. Official Journal of the European Communities, L345, pp.90-96.

EU (2001). Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society. Official Journal of the European Communities. L167, pp. 10-19.

EU (1996). Directive 96/9/EC of the European parliament and of the Council of 11 March 1996 on the legal protection of databases. Official Journal of the European Communities. L077, pp.20-28.

Groot, N., A. Bregt, J. Crompvoets & Y. Pluijmers (2007). Toegankelijkheid publieke geodata vooral gehinderd door juridische barrières [Accessibility public geodata mostly hampered by legal barriers]. Vi Matrix. 116: pp.6-8.

Hugenholtz, P.B. (2005). Abuse of database right. Sole-source information banks under the EU Database Directive. in F. Lévêvre & H. Shelanski (eds.), Antitrust, patents and copyright: EU and US perspectives, Cheltenham: Edward Elgar, pp. 203-219.

IGN [Richard, D.] (2008). Géoportail: The French geoportal. Access online to public datasets by means of geographic interfaces. IGN Institut Geographique National. Presentation INSPIRE drafting team network services, Geonovum Oct. 15, 2008, Amersfoort.

INSPIRE Data and Services Sharing Drafting Team (2008). Implementing rules for governing access and rights of use of spatial data sets and services for Community institutions and bodies – Description of the rules with accompanying Commentary and Guidance. Draft INSPIRE Document D4.9, 12 December 2008, http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataandServiceSharing/D4.9_Draft_IR_Data_and_Service_sharing_v2.0.pdf.

KPMG Consulting Inc. [Garry Sears] (2001). Geospatial Data Policy Study (executive summary). GeoConnections Policy Advisory Node. Ottawa, Canada, GeoConnections.

Loenen, B. van (2006). Developing geographic information infrastructures: The role of information policies. OTB Research Institute for Housing, Urban & Mobility Studies, Delft University of Technology, The Netherlands. PhD dissertation.

Loenen, B. van, J. Zevenbergen, G. Giff & J. Crompvoets (2007). Open toegankelijkheidsbeleid voor geo-informatie vergeleken: het gras leek groener dan het was [Open access policy compared for geo-information: the grass seemed greener than it was]. Report for the Dutch Ministry of the Interior. OTB Research Institute for Housing, Urban & Mobility Studies, Delft University of Technology, The Netherlands.

Longhorn, R.A. & M. Blakemore (2008). Geographic information: value, pricing, production, and consumption. Boca Raton, CRC Press, Taylor & Francis Group.

Longley, P.A., M.F. Goodchild, D.J. Maguire, & D.W. Rhind (2001). Geographic Information Systems and Science (Chichester, UK: Wiley).

MEPSIR [Dekkers, M., F. Polman, R. te Velde & M. de Vries] (2006). Measuring European Public Sector Information Resources. Final report of study on exploitation of public sector information – benchmarking of EU framework conditions. HELM Group of Companies of Moira, Northern Ireland & ZENC, the Netherlands.

MICUS Management Consulting GmbH (2008). Assessment of the Re-use of Public Sector Information (PSI) in the Geographical Information, Meteorological Information and Legal Information Sectors. Final Report. Study commissioned by the EC in 2007. http://ec.europa.eu/information_society/policy/psi/docs/pdfs/micus_report_december2008.pdf.

MICUS Management Consulting GmbH, 2003, Der markt fuer Geoinformationen: Potenziale fuer Beschaeftigung, Innovation und Wertschoepfung [The market for geo-information: potential for employment, innovation and value adding], (http://www.micus.de/downloads.html#).

MICUS Management Consulting GmbH (2001). Market survey: Boosting of the geospatial market in North Rhine Westphalia, (Media NRW), http://www.micus.de/pdf/micus_marktstudie_nrw_en.pdf.

Minister for the Cabinet Office (2005). The Re-use of Public Sector Information Regulations 2005. no. 1515, OPSI (Office of Public Sector Information), http://www.opsi.gov.uk/si/si2005/20051515.

Norwegian Ministry of Trade and Industry (2008). An Innovative and Sustainable Norway. Short version of the White Paper, report no. 7 to the Storting (2008-2009). http://www.regjeringen.no/upload/NHD/Vedlegg/brosjyrer_2008/innomeld_kortv_eng.pdf.

OFT (Office of Fair Trading) (2006). The commercial use of public information (CUPI). http://www.oft.gov.uk/shared_oft/reports/consumer_protection/oft861.pdf.

PIRA International Ltd, University of East Anglia & KnowledgeView Ltd (2000). Commercial exploitation of Europe's public sector information. Final report for the European Commission Directorate General for the Information Society (ftp://ftp.cordis.lu/pub/econtent/docs/commercial_final_report.pdf).

RAVI Bedrijvenplatform [Nijkamp, P., R. Vreeker, M.J.J. Koetse & F.R. Bruinsma] (2000). Economische effecten van laagdrempelige beschikkingstelling van overheidsinformatie [Economic effects of making public sector information easily accessible]. RAVI publication 00-02.

SADL (Spatial Applications Division K.U. Leuven Research & Development) (2008). Spatial data infrastructures in Germany, State of play 2007. Country report on SDIs elaborated in the context of a study commissioned by the EC (EUROSTAT) in the framework of the INSPIRE initiative, http://inspire.jrc.it/state_of_play.cfm.

Shapiro, C. and H. R. Varian (1999). Information Rules: a strategic guide to the network economy. Boston, Massachusetts, Harvard Business School Press.

STIA (Spatial Technologies Industry Association) (2001). Phase 1 report: Increase private sector awareness of, and enthusiastic participation in the National Spatial Data Infrastructure (NSDI). http://www.fgdc.gov/library/whitepapers-reports/sponsored-reports/stia/.

UN (1998). The Aarhus Convention. UN/ECE Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters. UNECE (United Nations Economic Commission for Europe).

Welle Donker, F. (2009). Public Sector Geo Web Services: which business model will pay for a free lunch? in: B. van Loenen, J.W.J. Besemer, J.A. Zevenbergen (eds), Spatial Data Infrastructure Convergence: Research, Emerging Trends, and Critical Assessment, Delft, Netherlands Geodetic Commission (NCG) (to be published in June 2009).

Welle Donker, F. & B. van Loenen (2006). Transparency of Accessibility to Government-Owned Geo-Information. 12th EC-GI&GIS Workshop, Innsbruck, Austria, JRC-IES.

3 Geo Shared Licenses: A base for better access to Public Sector Geo-Information for value added resellers in Europe

Frederika Welle Donker, Bastiaan van Loenen & Jaap Zevenbergen

Environment and Planning B: Planning and Design Journal, 2010, Vol.37 (2), pp.326-343, doi 10.1068/b35043. Submitted 2 April 2008; revised 2 April 2009; published online 8 January 2010.

Abstract

In a digital age public sector geoinformation (PSGI) is potentially a vital link in the added-value chain. Yet private sector value-added resellers (VARs) still face a number of barriers to using PSGI. Price is only one impediment. The complexity of licences and restrictive licence conditions of PSGI may be an even bigger obstacle. Especially when combining different datasets, VARs can face a quagmire of conflicting licence conditions. Batty (2006 Environment and Planning B: Planning and Design 33 163 – 164) called for research that would stimulate value-added use of PSGI. However, inconsistent and intransparent licence conditions for PSGI are among the biggest obstacles of PSGI for VARs. This paper explores the current PSGI licences to assess the actual restrictions and how current obstacles can be levelled. The Creative Commons licensing concept was explored and adapted to make it suitable for licensing PSGI. The resulting concept of Geo Shared licences is a means to harmonise licence conditions for PSGI. Our research shows that the Geo Shared concept can be a valuable contribution to further harmonisation of PSGI licences and thus development of value-added chains. Furthermore, development of geographic information infrastructures will also be stimulated. Similarly, the concept can be considered as a serious option within the Infrastructure of Spatial Information for Europe (INSPIRE), as a way towards transparent harmonised licences in Europe and beyond.

Keywords: Creative Commons; licences; geographic information infrastructure; INSPIRE; reuse of public sector information: Geo Shared

§ 3.1 Geographic information infrastructure development

§ 3.1.1 Framework

The terms 'geographic information', 'geographic data', 'spatial information' and 'spatial data' are interchangeably used as synonyms. For the purpose of this article only the term geographic information (GI) will be used. Access to GI is of vital importance to the economic and social development of the nation. Nations around the world are developing geographic information infrastructures (GIIs), also referred to as spatial data infrastructures (SDIs), with access to GI at the core. For more advanced GIIs (re) use is considered to be the driver of a GII. In this respect special reference is made to value added use of available basic or framework GI. Most GI belongs to public sector bodies with access and use governed by specific access policies. In Europe many public sector bodies use licence fees to finance their operations and to guarantee certain levels of GI quality. However, each body applies different licence conditions and pricing structures. It is this inconsistency and intransparency that forms one of the biggest obstacles for value-added reusers (VARs) in their decision to (re)use public sector geographic information (PSGI) for their activities (see Groot *et al.,* 2007; RAVI, 2000; STIA, 2001; van Loenen *et al.,* 2007). As a consequence, value-added use, the driver for advanced GIIs, remains limited.

A GII or SDI may be defined as the framework to facilitate the management of information assets, with a focus on better communication channels for the community for sharing and using data assets, instead of aiming toward the linkage of available databases (Rajabifard *et al.,* 2002). Governments have an important role in the development of GIIs. They are often both providers and users of GI, and most often government agencies lead GII development. This is especially true when the government is the main provider of GI. They can decide what information is collected and maintained and, through its access policies, they also determine the extent to which a dataset can be used. Pricing of PSGI is an important factor for users in their decision to use a data set for value-adding. However, surveys held in 2007 and 2008 among VARs in Europe suggests that the most prominent barriers for value-added (re) use are the complexity, inconsistency, intransparency and restrictive use conditions (Groot *et al.,* 2007; MICUS, 2008a);. The European Directive on the re-use of public sector information 2003/98/EC, the so-called PSI Directive, is explicitly directed at promoting value-added use of PSI (EC, 2003). However, it only prescribes a minimum of harmonisation for licences, keeping the hindering status quo alive. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community –

the so-called INSPIRE Directive – requires Member States to exchange, share, access and use interoperable spatial data and spatial data services across the various levels of public authority and across different sectors. INSPIRE should assist policy-making in relation to policies and activities that may have a direct or indirect impact on the environment. However, while INSPIRE requires data to be shared between public sector bodies, it also allows public sector bodies to charge for (re)using the data by leaving the regime of the PSI Directive unaffected. So far, these Directives have not resulted in a harmonisation of licence conditions leaving value-added use of PSGI hindered as before. INSPIRE must be transposed into national legislation by 15 May, 2009. A harmonised licensing framework has been developed by the INSPIRE Data and Service Sharing Drafting Team. However, this framework will be voluntary.

§ 3.1.2 Users and their needs

Users of the GII "will probably be the most mentioned group and yet actually the least considered" (McLaughlin and Nichols, 1994, p.72). Van Loenen (2006) distinguishes four user groups:

1. primary users (the collector and major users);
2. secondary users (incidental users for similar purposes as the primary user);
3. tertiary users (users that use the dataset for purposes other than those for which the information was collected and the dataset created), and
4. end-users.

Primary users are those that use the dataset for the initial purpose of information collection on a continuous basis. They typically belong to the organisation that has collected and processed the information. Secondary users use the information incidentally for similar purposes, and tertiary users are those that add value to the framework dataset by using the data set for other purposes then the collection purpose. Finally, the end users are the fourth group of users. This group consists of citizens, decision makers, and others that use the end product of geographic information – for example, a map or an answer to a query – mostly through services provided by the tertiary users. Although secondary, tertiary and end users all may reuse PSGI, it is the tertiary user that by definition reuses PSGI for value adding. Therefore, this paper's primary focus is on the tertiary user.

Users require transparency of the information policies (*e.g.* Groot *et al.*, 2007; RAVI, 2000) and require consistency in the access policies throughout government (KPMG, 2001; PIRA International, 2000; QSIO, 2006; RAVI, 2000; STIA, 2001). Differences in pricing, use restrictions, and liability regimes may result in confusion and ultimately

limited use of the dataset (Meixner and Frank, 1997). The user is, for example, uncertain about the cost he or she should calculate for complete jurisdiction coverage. A consistent or harmonised access policy throughout government may promote the use of framework information. In this paper we assess the extent to which the concept of the Creative Commons (CC) can be used as a tool to develop a model that will harmonise current PSGI licences.

§ 3.1.3 Reading guide

First, in Section 3.2, we will consider attempts to standardise licences in general, with a focus on CC. In Section 3.3 we describe the present situation with regard to the use of licences in the Netherlands, Norway, Germany, and England. In Section 3.4 we look at the pros and cons of applying CC for geographic information and look at the issues that remain when aiming at the extended use of CC. In Section 3.5 we introduce the Geo Shared concept as an alternative framework. We conclude with an analysis. Further, we will discuss the issues that CC can and cannot solve with regard to access to PSGI for VARs.

§ 3.2 Standardising lLicences

§ 3.2.1 Information licences

Access to and (re)use of geographic information is often regulated by licences to allow the information holder to economically or otherwise exploit the information. A licence is a contract imposing express limits on the use of the data (Dreyfuss, 1999). One can generally redistribute a licenced copy only if especially contracted for the right to do this (Samuelson, 1998). Other legislation such as privacy legislation may impose restrictions as well. Intellectual property rights (IPR) can be considered a prerequisite for successfully exploiting information.

§ 3.2.1.1 Copyright and database rights

Intellectual property rights such as copyright, and in Europe also database rights, may be had in many types of geographic information, such as topographic information. Copyright gives the creator of an original work exclusive rights to it (*e.g.* right to publish, distribute, and adapt), most often for a limited time, usually in the order of seventy years after the death of the author. The primary objective of copyright is to promote creativity and innovation.[24] It assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work (Onsrud and Lopez, 1998). Copyright protection extends to expressions and not to ideas, procedures, methods of operation, or mathematical concepts as such (WIPO, 1996). Differences among the copyright laws of various nations have resulted from a wide range of interpretations that nations have developed for the concept of originality (Onsrud and Lopez, 1998).

The EU Directive on the legal protection of databases (96/9/EC), the so-called Database Directive, made a significant change to intellectual property rights in Europe. This directive created a 'new' sui generis[25] right for the creators of databases that do not qualify for copyright as such. While copyright protects the creativity of an author, database rights protect the substantial investments in obtaining, verification, or presentation made by the producers. Under the terms of the Database Directive the rightholder may prohibit the extraction and/or reuse of the whole or a substantial part of the database. Database rights last for fifteen years from the end of the year that the database was made available to the public. Any substantial changes, which could be considered to be a substantial new investment, will extend the protection for another fifteen years. Therefore, databases that are regularly updated could effectively have a perpetual protection. Database rights may be reserved only if the investments in obtaining, verifying or presenting the data are made as a main commercial activity of the database producer. If a database is created without substantial investments or as a by-product of another activity, the so-called spin-off doctrine applies (Hugenholtz, 2005). The European Court handed down a number of rulings in 2004 confirming the spin-off doctrine.[26] On April 2009 the Council of State, the highest Dutch Court

24 See for instance, Feist Publications Inc. v. Rural Telephone Service Co., 1991, 499 US 340, 349.

25 The literal meaning of sui generis is of its own kind. In law it is a term used to identify a legal classification that exists independently of other categorisations because of its uniqueness or owing to the specific creation of an entitlement or obligation.

26 See for instance, The British Horseracing Board Ltd and Others v. William Hill Organization Ltd. ECJ, joint cases C-46/02, C-338/02 and C-442/02, 9 November 2004.

of Appeal for Administrative Law, upheld[27] the District Court of Amsterdam's ruling [28] that a public sector may not claim database rights for public sector databases, if the database was created as part of its public task and was funded by taxpayers' money. Thus, the spin-off doctrine has a significant bearing on public sector bodies claiming database rights as there may be no legal basis if they are publicly funded.

§ 3.2.1.2 Some Rights Reserved Licenses

In the 1990s changes were made to United States Copyright Act in order to offer better protection of works in a digital environment. These changes included retroactively extended copyright terms, thereby threatening to prevent the so-called orphaned works from being published on the Internet.[29] As a reaction to these copyright changes, several organisations were founded to provide alternative licensing systems, on the basis of a 'some rights reserved' approach. The terms 'some rights reserved' is used to denote a concept somewhere in between the 'all rights reserved' approach of the Copyright Act and the 'no rights reserved' approach of the Public Domain.

There are now over 60 'some rights reserved' type licences currently recognised by the Open Source Initiative. The most popular types currently in use for small group or individual users for non-software works are Creative Commons (CC) licences and variations of the GNU Free Documentation Licence (FDL). Although the latter was designed originally to apply to software manual documentation, it has been applied far more widely – for example, for projects of the Wikimedia Foundation (Onsrud, 2006). The fact that there are so many different 'some rights reserved'-type licences is a fair illustration that attempts to standardise these have not succeeded, as illustrated by the attempts of the Science Commons to develop a licence framework since 2005. Even to enable just one transaction, namely the transfer of biological materials, Science Commons have developed four different Material Transfer Agreements (http://www.sciencecommons.org/projects/licensing/).

27 Raad van State case nr. 200801985/1. The Council of State reiterated in its ruling that databases funded by public money and produced for a public task rather than specifically for commercial purposes, cannot be protected by database rights as the investments made to produce the database–even though the investments were vast–had not carried a substantial risk.

28 Landmark Nederland BV v. Municipality of Amsterdam, Amsterdam District Court reg. nr. LJN BG1554, 11 February 2008.

29 As contested in Eldred v. Ashcroft, 2003, 537 US 186.

In the US the National Research Council (NRC) suggests that, in order to facilitate finding and (re)using geoinformation, a national GI marketplace should be set up. The would-be customer could search for GI and buy the suitable data after 'clicking-through' to the appropriate server. In more advanced implementations, the seller or licensor might define for each dataset or group of datasets a pricing formula that varies with differing standard licence or sale conditions (National Research Council, 2004).

§ 3.2.2 Creative Commons

CC was founded in 2001 as a non-profit organisation to offer flexible copyright licences for creative works such as text articles, music, and graphics. They advocate a system whereby works can be made available through the Internet without forfeiting their intellectual property rights. To facilitate this, they have developed a licensing system, the co-called CC licences. Thus, works can be made easily accessible for dissemination or for reuse. As at February 2009, fifty countries around the world have set up national CC organisations and have transposed the US version of CC licences into national legislation. CC licences are becoming very popular; at the end of 2003 there were worldwide about 1 million CC licences in use, and at the end of 2008 this number has exploded to 130 million and at the time of writing is still growing exponentially (www.creativecommons.org). Within Open Geospatial Consortium efforts to arrive at a geospatial rights management standard, variations of CC licences are also considered (Vowles *et al.*, 2007).

CC licences try to find a balance between the 'all rights reserved' concept of traditional IPR and the 'no rights reserved' concept of the public domain, by employing a 'some rights reserved' approach. Through their website (http://www.creativecommons.org) they offer six standard licences for anyone wanting to publicise their work. Each CC licence contains the following standard clauses:

1. The licence applies worldwide.
2. The licence is irrevocable.
3. The licence is granted for the term of the appropriate IPR legislation.
4. Licensors do not forfeit their IPR.
5. Acknowledgement of the source is compulsory (attribution the way the author requests).
6. Licensees must seek permission for actions that are not allowed by that specific licence.
7. Each copy of the work must contain a link to the licence.
8. Licencees may not alter the terms of the licence agreement.
9. Licencees may not employ technology or other means to limit access to the work in a way that is contradictory with the terms of the licence agreement.

10 Works are offered on an 'as-is' basis without any guarantees and the licensor does not accept any liability claims.

Apart from each of these standard clauses, the six CC licences offer one or more of the following terms:

1 You let others to distribute derivative works only under a licence identical to the licence that governs your work (share alike).
2 You let others copy, distribute, display, and perform only verbatim copies of your work, not derivative works based upon it (no derivative works).
3 Others may copy, distribute, display and perform your work – and derivative products based upon it – but for non-commercial purposes only (noncommercial).
The six main licences are described in Table 3.1.

LICENCE TYPE	ICONS	LICENCE CONDITIONS
Attribution (by)		This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered, in terms of what others can do with your works licensed under Attribution.
Attribution Share Alike (by-sa)		This license lets others distribute, remix, tweak, and build upon your work, even for commercial reasons, as long as they credit you and license their new creation under the identical terms. This license is often compared with open source software. All new works based on yours will carry the same licence, so any derivatives will also allow commercial use.
Attribution No Derivatives (by-nd)		This license allows for redistribution, commercial and noncommercial, as long as it is passed along unchanged and in whole, with credit to you.
Attribution Non-Commercial (by-nc)		This license lets others remix, tweak, and build upon your work noncommercially, and although their new works must also acknowledge you and be noncommercial, they don't have to license their derivative works on the same terms.
Attribution Non-Commercial Share Alike (by-nc-sa)		This license lets others remix, tweak, and build upon your work noncommercially, as long as they credit you and license their new creations under the identical terms. Others can download and redistribute your work just like the by-nc-nd licence, but they can also translate, make remixes, and produce new stories based on your work. All new work based on yours will carry the same licence, so any derivatives will also be noncommercial in nature.
Attribution Non-Commercial No Derivatives (by-nc-nd)		This licence is the most restrictive of the six main licenses, allowing redistribution. This licence is often called the 'free advertising' licence because it allows others to download your works and share them with others as long as they mention you and link back to you, but they can't change them in any way or use them commercially.

TABLE 3.1 Creative Commons licenses
http://creativecommons.org/about/licenses/, symbols Trademark by Creative Commons, http://creativecommons.org)

Each of the CC licences generates three versions of the same licence agreement. The first version – a commons deed in plain language suitable for laymen – is a summary of the licence complete with the relevant symbols as displayed in Table 3.1. The second version – a legal code – is the actual licence and is legally binding. The legal code is suitable for lawyers and consists of a number of pages in legal terminology. The third version – a digital code – is a machine-readable translation of the licence that helps computer programs such as search engines to identify the work by its terms of use.

CC aims to promote access to IPR protected works as an open content organisation. Open access works, while copyrighted, allow use without obtaining prior permission since a general licence is granted ahead of any specific use. A basic condition of a CC licence is that user rights are supplied without royalties, although the right to receive a reward is not forfeited under a CC licence. The licences were designed to suit creators who want to distribute their work independently to gain publicity or to build up a reputation, or to suit creators or organisations that act out of ideological or non-profit objectives. The CC-licences are also applied to digital works to stimulate sales of the printed version of the same work, or to promote the use of paid support services (Boyle, 2007; National Research Council, 2004). CC-licences appear to be suitable for those that do provide their data for free such as non-profit organisations, academia, and government organisations, but also suitable for VARs that may use the data as the trigger to generate revenue from the sales of related products or services.

§ 3.3 PSGI licences in Europe

Although all EU Member States have to abide by the PSI Directive, there are still quite some differences with respect to access and user licences. Information regarding the Netherlands, Norway, North Rhine Westphalia (Germany), and England and Wales (United Kingdom) was collected as part of a study (van Loenen *et al.*, 2007). In this chapter, we will give a brief summary of access policies of these countries.

§ 3.3.1 The Netherlands

In the Netherlands access to PSI and reuse of PSI are both regulated by the Freedom of Information Act (Wet openbaarheid van bestuur, known as the Wob). The Wob states that, with respect to access, fees should not exceed dissemination costs as far as

possible. However, for reuse of PSI subject to IPR, charges should not exceed the total costs of collecting, producing, reproducing and disseminating documents, together with a reasonable return on investment. Some public sector organisations have their own specific legislation setting their own framework for disseminating information – for example, the Cadastre Act and the Meteorological Institute Act. At ministerial level there is a framework of policies and guidelines recommending that information should be made available to other national public sector organisations for dissemination costs. However, this framework does not apply to regional and municipal organisations (yet). The Wob is currently under review and the differences in pricing regimes will probably be amended.

In 2006, current licence agreements for PSGI were reviewed (Welle Donker, 2006). Licence terms and conditions appeared to be very diversely formulated, ranging from a few paragraphs written in plain language to countless pages written in legal language that is hard to understand for a layperson. The licence fees also vary significantly, ranging from free to hundreds of thousands of Euros for a complete dataset. Sometimes a differentiation is made between different types of users – that is, libraries, schools, universities and research institutes will pay lower fees than the private sector. Almost all of the licence agreements usually specify that the data are to be used only for internal purposes and if the dataset is to be used for any other purposes a separate licence agreement will have to be negotiated. In some cases one has to indicate what the data will be used for before access or permission for reuse is granted. Sometimes the dataset has to be returned after a (predetermined) goal has been attained. Sometimes one has to purchase an entire dataset and sometimes one gets access through a web service. None of the licence agreements contain provisions for the combined use of data from more than one source (Welle Donker, 2006). Formally no differentiation is made between public sector users and non-public sector users. In practice, some public sector organisations have data-for-data agreements, in which they share data to create and maintain large-scale datasets. Some public sector organisations charge fees to other public sector organisations.

In spite of all these differences, all these licence agreements also show a lot of similarities as far as the main provisions are concerned. These similar provisions are:

- A non-exclusive user right is granted.
- Intellectual property remains with the supplier.
- The data may not be transferred to a third party without prior consent of the rightholder.
- Derivative products obtained by adaptation of the data (if allowed) must be clearly credited with the original source (name of supplier and year of acquisition).
- The supplier of the data indemnifies himself or herself against any claims to the comprehensiveness and accuracy of the data or any damage resulting from use of the data.

– General (nonspecific) financial provisions related to terms of payment.

§ 3.3.2 Norway

Within the public sector several organisations handle geographic information. The Norwegian Mapping and Cadastre Authority (SK) – which falls under the Ministry of Environment – is responsible for the coordination of the Norwegian GII. In 2003, a white paper authorised GI sharing within the public sector by setting up a GII. This program, called Norge Digitalt (Digital Norway), provides not only a portal but also a framework for cooperation within the public sector. Nearly all state departments and agencies as well as local governments and some private partners have joined or are in the process of joining Norge Digitalt (ND). After paying a contribution, the government organisation then makes its GI available free of charge to other participating organisations. This way all participants can use free GI for its own internal processes. More than thirty state and almost all local government organisations are a member of ND. In Norway thematic GI should be available – often online – free of charge for everybody to view. For environmental information this has been the case by law since 1993.

If the private sector wants to use PSGI, they can buy datasets from a government-owned intermediary, the Norsk Eiendomsinformasjon (NE). NE acts as a one-stop shop for VARs to get the data and resell them to end users. A contract is drafted with the NE and NE pays royalties to ND. NE uses the same (restrictive) licence conditions for all information it resells. However, there are some unresolved issues with this system. SK is not allowed to sell information directly to third parties but other members of ND are. Several public sector organisations provide this information for free through web mapping services. Until 1 January 2007, all SK services were available freely on the web. To be in line with the access policy from the 2003 white paper, SK had to limit free access to ND partners only. NE does not have a publicly known pricing policy. NE is supposed to operate as a wholesale distributor but NE is also selling PSGI to end users thus blurring the boundaries between public and private tasks (Welle Donker and Zevenbergen, 2007).

§ 3.3.3 North Rhine Westphalia (Germany)

North Rhine Westphalia (NRW) is one of the sixteen states of the federal republic of Germany. Each German state is responsible for its own topographic service and land register, environmental and statistical information collection, and in general for

information policies. Information collection is largely decentralised and carried out mostly on the regional and local level. The different states have issued laws ('surveying and cadastral acts') that regulate both the work, and the authorities of the surveying and mapping agencies.

All local governments in NRW claim copyright and database right in their information. In NRW users of PSGI are granted a 'limited use right' as described in the Copyright Act and further in the Cadastre Act. Only with permission of the concerned organisation can information from local government be multiplied, made public, or provided to third parties. Making copies and processing the (digital) information for internal use are permitted.

The Cadastre Act rules that access to PSGI within government is without cost. The free access provision does not apply to access for VARs. One has to pay a fee according to the fee ordinance if cadastre information is used for commercial purposes. The fee for the information depends on the category of the layers, the information density, the size of the area requested, and the format requested (analogue, vector, raster). Further, there are different fees for different users. Although the fee ordinance provides the legal framework for the price setting of PSGI, it is generally regarded as complex and difficult to understand, and too inflexible to be of use for Internet applications. As in the Netherlands, VARs in NRW find the current restrictive licence conditions a major obstacle to reusing PSGI (MICUS, 2008b).

§ 3.3.4 England and Wales (United Kingdom)

Within the UK, Scotland and Northern Ireland have devolved responsibilities. In England and Wales policy is set by the UK government. Therefore, we will limit ourselves to England and Wales. Local governments are responsible mainly for local planning and everyday operations of their areas. The UK has different copyright regimes that apply to GI. The main copyright law affecting PSGI is the Crown Copyright. Crown Copyright applies to PSGI produced by central government agencies referred to as Crown Bodies. However, it is not always easy to distinguish which public sector organisations are Crown Bodies and thus affected by Crown Copyright because of technical legal reasons (APPSI, 2004). Therefore, different central government agencies will have different copyright regimes regulating their information, resulting in different rules for reuse.

Most PSGI is generated by the Ordnance Survey (OS). PSGI is also provided by central government parties like the United Kingdom Hydrographic Office (UKHO), Her Majesty Land Registry (HMLR) and the Royal Mail Group. The local authorities of the UK

(approximately 500, excluding the local authorities of London) have an agreement with public and private GI producers for the provision of GI products and services they require for performing their activities. This agreement is known as the Mapping Services Agreement (MSA). This competitive procurement results in the responsibility for the provision of GI to local authorities falling into the hands of three GI suppliers. In the MSA the OS is still the main provider of GI datasets with supporting datasets being provided by Intermap and Intelligent Addressing. However, the majority of the more widely used GI in the UK is derived from or is actually OS datasets. OS, UKHO and HMLR are all classified as trading funds and are required to generate a surplus. Therefore, these agencies all use restrictive licence conditions and fees to make their datasets available for reuse. Hence, access to these datasets will be governed by the underlining policies of these trading funds.

Like the Netherland, the UK has no single access policy for PSI. As far as reuse within the public sector is concerned, OS uses a system of Collective Licensing Agreements (CLAs). A CLA is an agreement between the OS and a group of public sector organisations which allow the public sector organisations access to OS information for internal processes. As far as reuse by the private sector is concerned, UKHO uses a network of VARs which reuse hydrographic information on a royalty basis. OS also have licence agreements with various VARs on a royalty basis and/or upfront fees.

From the above examples we can see that there are vastly different approaches to PSGI licensing in Europe.

§ 3.4 Applying Creative Commons to PSGI licences in Europe

Although CC licences appear vastly different from the PSGI licences currently in use, the general terms of most licence agreements do not differ that much from the CC licences. Thus, CC offers a promising access model. However, not all the available CC-licences can be applied to geographic information as such, especially if our aim is to make datasets available as input for commercial value-added products and services. Tables 3.2a and b show that there are some inherent problems when applying CC licences to PSGI for VARs. In this section we will discuss some of these concerns.

§ 3.4.1 Matches and differences between PSGI licences and Creative Commons licences

As we have shown, there are matches and differences between current PSGI-licences and CC-licences. These are listed in Tables 3.2(a) and 3.2(b) respectively. No colour indicates a match, medium blue a near-match and dark blue a substantial difference.

The table shows that there are discrepancies in several locations. These discrepancies and shortcomings of Creative Commons are addressed in the following sections.

CC	NL	NORWAY	NRW	ENGLAND
Adaptation of the information is in some cases allowed. Derivatives must be clearly attributed to the creator(s) of the original source	Yes	Yes	Yes	Yes
Information is accessible on-line after the licence terms have been agreed to	Sometimes	Yes	Yes	Sometimes
The intellectual property rights remain with the right holder	Yes	Yes	Yes	Yes
The user obtains a non-exclusive user right	Yes	Yes	Yes	Yes

TABLE 3 2(a): Matches in licence conditions CC and European case studies

Background medium blue: Near Match
Background no colour: Match

CC	NL	NORWAY	NRW	ENGLAND
On-line acceptance of licence is available (no paper application or signature required)	Sometimes	Need formal agreement	Need formal agreement	Need formal agreement
The user may transfer the information and/or derivatives to a third party without prior consent of the right holder	No	Only to Norge Digitalt (Digital Norway) participants	Only to public sector parties	No
All Information is available for (re)use at no upfront charges and free of royalties	Some information	Only thematic (environmental) information	Only for other public sector bodies	Very little information
No differentiation between types of users	Sometimes	Differentiation between public sector and other users	Differentiation between public sector and other users	Differentiation between public sector and other users
Licence is valid for the duration of copyright/database right	Sometimes valid for fixed period	Sometimes valid for fixed period	Sometimes valid for fixed period	Only valid for fixed period

TABLE 3 2(b): Differences in licence conditions CC and European case studies

Background medium blue: Near match
Background dark blue: Substantial difference

§ 3.4.2 Commercial use

One of the cornerstones of CC is sharing information, usually for noncommercial purposes. However, what exactly constitutes 'commercial use'? In its legal code CC defines noncommercial in article 4b as:

"You may not exercise any of the rights granted to You [the licensee] ... in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital file-sharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works."

This definition is clear with regard to a private sector organisation that wants to use the dataset to produce a product or service with the intention to sell this product or service for a profit. But what about use by nonprofit organisations, are they entitled to use data made available under a 'non-commercial' condition when they do not intend to make a profit? Should there be a differentiation between public and private schools since private schools are institutes that ultimately intend to make a financial profit? And what about a company representative visiting a client using a car navigation system, does this constitute commercial or internal use? The courts will not only look for a legalese interpretation of the word 'commercial' but also look at the contract situation as a whole, when interpreting the situation (Pawlo, 2004). On a national level, some consensus may be reached what the meaning of 'commercial' will be, but on an international level this may not be the case. In the Netherlands, the District Court ruled in favour of a CC licensor. A well-known DJ had published photographs of his family on flickr.com under a CC-nc licence. A magazine used some of these photographs without permission. The DJ successfully sued the magazine for breach of the CC licence, although no damages were awarded.[30]

Therefore, the CC question 'Allow commercial uses of your work?' would always have to be answered with 'yes', or else the private sector would not be able to use the datasets. Even if they were only to use the datasets for internal use rather than to produce directly value-added services, this may still constitute commercial use, given the uncertainty of the concept 'non-commercial' in various jurisdictions. To avoid a potential quagmire, it would be best if only by-nd, by-sa or by licences are used for reusing PSGI.

30 Curry v Audax Publishing, District Court Amsterdam, 9 March 2006, ECLI: NL: RBAMS: 2006: AV4204.

In most European jurisdictions public sector organisations make PSGI available for producing value-added products and services only after a formal agreement has been negotiated. This allows the public sector organisation to customise licence agreements depending on the type and quantity of data. This is one of the reasons why the current licensing system is not transparent. It might be more practical to replace the current CC noncommercial use symbol with an 'advance permission' symbol. The licence condition as it is currently in use by a number of public sector organisations would thus be better represented. It would also avoid a philosophical discussion concerning commercial use. However, it would be better to abolish the distinction between non-commercial (internal use only) and commercial (external use) entirely. Especially as a non-commercial CC licence will not prevent the user from reproducing the data using web services or posting the data on websites. As long as there is no financial gain for the licencee, the licencee is allowed to do so as long as the right attribution has been made.

§ 3.4.3 Derivatives and Share Alike

In the older CC-versions there was a mismatch between different 'some rights reserved' licences such as CC and FDL. If you wanted to remix works issued under different 'some rights reserved' licences you could not make the derivative product available if the derivative has to be licenced under exactly the same licence as the original. By selecting one 'some rights reserved' licence over the other, you were in breach of the original licence and therefore neither could be selected. Version 3.0 of CC, released in the spring of 2007, has rectified these incompatibility problems. Products may now be made available under other types of open content licences, as long as they have the same properties.

The CC licence concepts of 'no derivatives' and 'share alike' also may pose a problem if the aim is to make datasets available for value-added products. If PSGI is only to be used without being able to produce derivatives, then it will only be suitable for internal business processes or for end users. Whilst this makes the licence suitable for GI reuse by secondary users and end users, it will not stimulate value adding by tertiary users. The same applies to the share alike option. In a creative environment the concept of sharing works, adapting them and making the derivatives available under similar conditions can be very important. Institutes like Wikipedia could not exist without share alike licences. But when PSGI is made available to tertiary users for value adding, the concept of making the value-added services and products available under the same conditions would be counterproductive. The concepts are therefore only suitable to make PSGI available to secondary users and end users, provided the GI was supplied for no more than marginal costs of dissemination. This constitutes discrimination

for different types of users which is in conflict with the non-discriminatory provision in the PSI Directive.

PSGI licences found typically are non-transferable licences without so-called viral use conditions (licences conditions requiring derived works should be made available under the same some-rights-reserved conditions). Therefore, the share-alike condition of CC cannot be applied in these instances.

§ 3.4.4 Fees and royalties

CC aim to protect some rights of the author, which should also include the right to receive fair compensation. But CC also stated in their earlier licence conditions that the licencee is under no obligation to pay "any royalties, compulsory licence fees, residuals or any other payments". However, in a number of jurisdictions collective music rights systems are in place. With version 3.0 CC addresses this problem of compulsory contributions to collecting societies. In the older licensing versions the right to collect royalties had to be waived. CC has now acknowledged that this is not possible to do so in those jurisdictions where there are statutory or compulsory licensing schemes. Whilst this amendment addresses the problem of musicians having to compulsory join a collection society and still wanting to publish their work under a CC licence, it does not directly address the problem of a licensor intending to charge licence fees and/or royalties. CC licences as such therefore seem to be effectively only suitable for organisations that intend to make the datasets available free of charge. However, if PSGI is made available for dissemination costs, then one does not pay for the actual dataset. Rather, one pays a compensation for setting up and maintaining a web service, cost of DVD, or postal charges. In that case we hold the opinion that a CC licence can be used for PSGI as long as it is clear that the data itself is free and one only pays for the costs of dissemination. However, much European PSGI is available at a price exceeding the marginal cost of dissemination. In these instances, CC cannot be applied.

§ 3.4.5 Liability

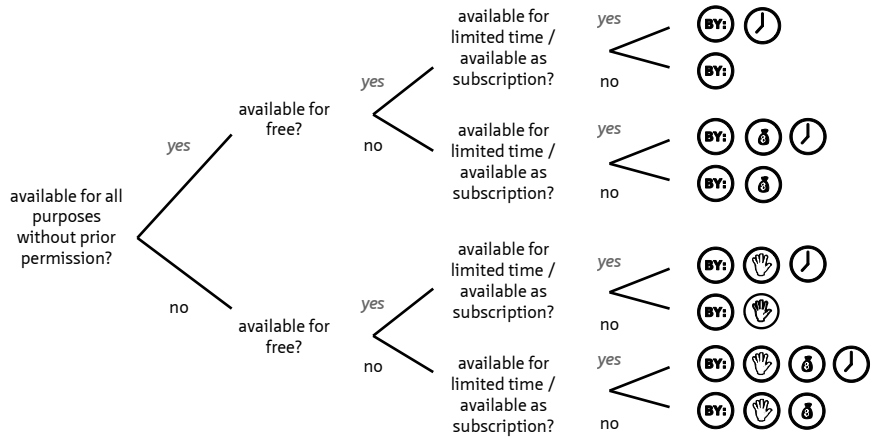
Geo datasets incur a different liability regime than most other data. Suppose a company is commissioned by a municipality to produce a road system for a new housing development. Afterwards it turns out there is a mistake in the dataset because two street names were switched. The municipality suffers losses because they have used the dataset to produce a new street plan and have already distributed 10,000

copies. Others may suffer losses as well because of this mistake. What if one of the residents suffers a heart attack and dies because the ambulance was delayed due to the street name mix-up? Can his relatives claim damages? (van Loenen *et al.*, 2006).

We will not go into the legal details of liability here as liability regimes differ in Europe. In general though, in the Netherlands, if a public sector organisation makes (geographic) data available for reuse by third parties, the datasets should be accurate and exhaustive enough to carry out the original public task (van Loenen *et al.*, 2006). The metadata should display the original use of the dataset. Potential users of the dataset can determine if the dataset is suitable for the intended (re)use by inspecting the supplied metadata. However, although INSPIRE will prescribe metadata standards, in practice metadata is poorly maintained, especially for older GI and non-GI PSI. In the CC licences v3.0, works are offered 'as-is' unless mutually agreed by parties in writing. So, if the metadata is incomplete, liability will remain a problem as the licencee does not have enough information to determine the suitability of the data. Furthermore, consumer protection legislation might prevent the use of a total disclaimer. In the Netherlands, for example, disclaiming liability for gross negligence is not allowed in general conditions between companies and consumers (it is on the so-called black list). The Dutch CC licence allows for such legal provisions at the end of article 6.

§ 3.4.6 In summary

This means that – apart from a public domain licence – in effect only one out of the six CC licences can be considered for supplying PSGI, namely the CC-BY licence. This conclusion corresponds with the conclusions of research carried out about the suitability of CC-licences for public sector information in general (van Eechoud and van der Wal, 2008). Nevertheless, with the additional symbols as shown in Figure 3.1, most of the current PSGI policies would be covered. However, when changes are to be made to the original CC model, the name 'Creative Commons' can no longer be used. The name 'Geo Shared' is more applicable to an adapted model. Geo Commons seems more obvious as a moniker, but is not so suitable. The name Commons implies communal use – that is, prior permission for use does not have to be sought before the GI is used. This may be misleading, hence the name Geo Shared.



= **Attribution**. Others may copy, distribute, display, and perform the copyrighted work - and derivate works based upon it - but only if they give credit the way rightholder request.

= **Permission in advance**. Data and/or derivative products may only be made available to third parties after obtaining permission from right holder in advance.

= **Costs**. The user is required to pay licence fees and/or royalties for the use of the data/information.

= **Limited period**. The Data and/or derivative products is available for a limited period, see full licence for exact period

FIGURE 3.1 Geo Shared licences

§ 3.5 Geo Shared licenses

Although CC licences are considered for PSGI in Queensland (Australia) (QSIC, 2007; QSIO, 2006) and are successfully used in the Netherlands for the New Map of the Netherlands, available at no cost, the analysis of licences currently available for PSGI in Europe shows that a one-to-one translation into CC licences is not possible (see Table 3.3). The first difference – formal licences – can be solved by online registration and password-controlled entry procedures. Many organisations which supply GI already use online registration forms and password-controlled entry procedures. The second difference does not pose a problem either as it can be included in the legal code. To make this condition clearer on the common deed, the non-commercial use symbol could be replaced with another symbol. The third difference could be overcome by including an extra symbol to indicate the difference between free or fee-based data. To indicate the last difference, another symbol could be included on the common deed.

However, it is debatable if this is necessary. A lot of GI dates quickly, having most of its value in the degree to which it is up-to-date.

By adapting the existing schema for CC with additional symbols, we can no longer use the name Creative Commons. Therefore, we will refer to the new schema as Geo Shared licences. This would be in line with the recommendations of the National Research Council (2004).

The licence conditions are reduced to the following terms:

1. Others may use your data as long as they credit you for the original creation the way you request it.
2. Others may use, copy, display and distribute your data – and derivative products based upon it – either for commercial or non-commercial purposes, but only after they have contacted you in advance the way you request it (prior permission).
3. The data are available for an upfront fee and/or attracts royalties payable (fee-based).
4. The data are only available for a limited period, either on a subscription basis or data to be returned after a specified period (time limit).

LICENCE TYPE	ICONS	LICENCE CONDITIONS
Attribution		This license lets others copy, build upon and distribute the data without prior permission, as long as they credit the rightholder for the original creation.
Attribution Time Limit		This license lets others copy, build upon and distribute the data without prior permission, as long as they credit the rightholder for the original creation. The data are only available for a limited period.
Attribution Fee-based		This license lets others copy, build upon and distribute the data without prior permission, as long as they credit the rightholder for the original creation. The data incur upfront fees and/or royalties payable.
Attribution Time Limit Fee-based		This license lets others copy, build upon and distribute the data without prior permission, as long as they credit the rightholder for the original creation. The data incur upfront fees and/or royalties payable. The data are only available for a limited period.
Attribution Prior Permission		This license lets others copy, build upon and distribute the data only after prior permission. The rightholder must be credited for the original creation.
Attribution Prior Permission Time Limit		This license lets others copy, build upon and distribute the data only after prior permission. The rightholder must be credited for the original creation. The work is only available for a limited period.
Attribution Prior Permission Fee-based		This license lets others copy, build upon and distribute the data only after prior permission. The rightholder must be credited for the original creation. The work incurs upfront fees and/or royalties payable.

Attribution Prior Permission Fee-based Time Limit		This license lets others copy, build upon and distribute the data only after prior permission. The rightholder must be credited for the original creation. The work is only available for a limited period. The work incurs upfront fees and/or royalties payable.

Table 3.3 Geo shared licence framework

§ 3.6 Conclusion

Only if the restrictive reuse conditions and financial issues have been resolved is value-added use expected to thrive. Until that very moment, the introduction of a CC inspired concept such as the Geo Shared concept in the world of GI may help to increase the transparency and consistency of licence agreements, especially when combining data from different sources. Although CC licences are not suitable for all types of GI licences, they do provide a tool to review the current PSGI licences. Both CC and Geo Shared licence categories provide a way to review and categorise current licences. The Geo Shared licensing concept also enables the harmonisation of fee-based datasets. Using symbols in a layman's version of licence agreements makes it easier for users to identify datasets suited to specific purposes. Uniform and legible licence agreements would certainly help to make the whole process more transparent, especially when combining datasets from different suppliers. In this way, the Geo Shared concept is a valuable contribution to the development of many geographic information infrastructures around the world, including INSPIRE. Therefore Geo Shared licences should also be considered as a serious option within INSPIRE as one concept of transparent harmonised licences for geographic information as a key for the utilisation of the geographic information infrastructure in Europe. To the same end, other nations across the globe may take advantage of the Geo Shared concept by harmonising existing licence conditions of PSGI. Ultimately, this may result in a standard set of licences for PSGI providing the consistency and transparency required by value-added resellers.

Acknowledgements This article has been written as part of the activities of the project 'GeoPortal Network – Liberty United' (Geoloketten – Vrijheid in Verbondenheid) in the framework of the Bsik program 'Space for Geo-information'.

ABBREVIATIONS	
CC	Creative Commons
CLA	Collective Licence Agreement
FDL	Free Documentation Licence
GI(I)	Geographic Information (Infrastructure)
HMLR	Her Majesty Land Registry
INSPIRE	Infrastructure for Spatial Information in Europe
IPR	Intellectual Property Rights
MSA	Mapping Services Agreement
ND	Norge Digitalt (Norwegian Geographic Information Infrastructure)
NE	Norsk Eiendomsinformasjon (Norwegian GI One-Stop Shop)
NRC	National Research Council
NRW	North Rhine Westphalia (German State)
OS	Ordnance Survey (United Kingdom)
PS(G)I	Public Sector (Geographic) Information
SDI	Spatial Data Infrastructure
SK	Statents Kartverk (Norwegian Mapping and Cadastre Authority)
UKHO	United Kingdom Hydrographic Office
VAR	Value Added Reseller
Wob	Wet openbaarheid van bestuur (Dutch Freedom of Information Act)

References

APPSI (Advisory Panel on Public Sector Information), 2004, 1st Annual Report, http://www.appsi.gov.uk/reports/annual-report-2004.pdf.

Batty, M., 2006, Public sector information: chains of added value, Environment and Planning B: Planning and Design 33(2). pp. 163-164.

Boyle, J., 2007, Text is free, we make money on volume(s), Financial Times, 22 January.

Dreyfuss, R. C., 1999, Do you want to know a trade secret? How article 2B will make licensing trade secrets easier (but innovation more difficult) Californian Law Review 87, pp. 193-267.

EC, 2003. "Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the reuse of public sector information Official Journal of the European Communities L345, pp. 90-96, 31 December 2003.

van Eechoud, M., van der Wal, B., 2008, Creative Commons licensing for public sector information. Opportunities and pitfalls, Institute for Information Law (University of Amsterdam).

Groot, N., Bregt, A., Crompvoets, J., Pluijmers, Y., 2007, Toegankelijkheid publieke geodata vooral gehinderd door juridische barrières [Accessibility of public sector geodata mostly hampered by legal barriers] VI Matrix 116 15)8), pp. 6-8.

Hugenholtz, P.B., 2005, Abuse of database right. Sole-source information banks under the EU Database Directive , in Antitrust, patent and copyright: EU and US perspectives, Eds. F. Lévêvre & H. Shelanski Edward Elgar, (Cheltenham), pp. 203-219.

KPMG Consulting Inc. [Garry Sears], 2001, Geospatial Data Policy Study (executive summary), GeoConnections Policy Advisory Node, GeoConnections (Ottawa, Canada).

van Loenen, B., 2006, Developing geographic information infrastructures: The role of information policies, PhD Dissertation, OTB Research Institute for Housing, Urban and Mobility Studies, Delft University of Technology (Delft, The Netherlands).

van Loenen, B., Welle Donker, F., Ploeger, H., Zevenbergen, J., Besemer, J., 2006, Beschikbaar stellen van geo-informatie bij Rijkswaterstaat. Analyse van de (on)mogelijkheden van het op korte termijn beschikbaar stellen van vier geo-data sets. Making geo information available by Dept. of Public Works: Analysis of the (im) possibilities of making four geo-datasets available in the short term, (OTB Research Institute for Housing, Urban and Mobility Studies, Delft (Delft University of Technology).

van Loenen, B., Zevenbergen, J., Giff, G., Crompvoets, J., 2007, Open toegankelijkheidsbeleid voor geo-informatie vergeleken: het gras leek groener dan het was [Open access policy of geo information compared: the grass seemed greener than it was], (OTB Research Institute for Housing, Urban and Mobility Studies, Delft (Delft University of Technology).
McLaughlin, J., Nichols, S., 1994, Developing a National Spatial Data Infrastructure, Journal of Surveying Engineering 120, pp. 62-76.
Meixner, H., Frank, A.U., 1997, Study on policy issues relating to geographic information in Europe, IMPACT program (European Commission).
MICUS Management Consulting GmbH, 2008a, Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report, Study commissioned by EC in 2007, http://ec.europa.eu/information_society/policy/psi/docs/pdfs/micus_report_december2008.pdf.
MICUS Management Consulting GmbH, 2008b, The value to industry of PSI: the business sector perspective, in The socioeconomic effects of Public Sector Information on digital networks: Toward a better understanding of different access and reuse policies, Ed. P. Uhlir, OECD Headquarters 4 February 2008, Paris, France.
NRC (National Research Council), 2004, Licensing geographic data and services, Washington DC (The National Academies Press,).
Onsrud, H., 2006, Management of Intellectual Property Use Licensing in a Commons of Geographic Data, CGD IP White Paper v4 (Dept. of Spatial Information Science & Engineering, University of Maine, USA), http://geodatacommons.umaine.edu/wpapers/CGD%20IP%20White%20Paper.pdf.
Onsrud, H., Lopez, X., 1998, Intellectual property rights in disseminating digital geographic data, products and services: Conflicts and commonalities among European Union and United States approaches, in: European Geographic Information Infrastructures: Opportunities and pitfalls, Eds. I. Masser & F. Salgé London (Taylor and Francis), pp. 153-167.
Pawlo, M., 2004, What is the meaning of non-commercial? in International Commons at the Digital Age. La création en partage, Eds. D. Bourcier & M. Dulong de Rosnay, Paris (Romillat),, pp. 69-81.
PIRA International Ltd, University of East Anglia, KnowledgeView Ltd, 2000, Commercial exploitation of Europe's public sector information - Final report, European Commission Directorate General for the Information Society, ftp://ftp.cordis.lu/pub/econtent/docs/commercial_final_report.pdf.
QSIC (Queensland Spatial Information Council) and Queensland Government Natural Resources Mines and Water, 2007, Queensland Spatial Imagery Acquisition Program 2007-2011 (Qld Govt Natural Resources, Mines and Water), www.qsic.qld.gov.au.
QSIO (Queensland Spatial Information Office), OESR (Office of Economic and Statistical Research), Queensland Treasury, 2006, Government Information and Open Content Licensing: An Access and Use Strategy, QSIC (Queensland Spatial Information Council) (Queensland Government), www.qsic.qld.gov.au.
Rajabifard, A., Feeney, M.-E.F., Williamson, I.P., 2002, Directions for the future of SDI development, International Journal of Applied Earth Observation and Geoinformation 4, pp. 11-22.
RAVI Bedrijvenplatform, 2000, Economische effecten van laagdrempelige beschikkingstelling van overheidsinformatie [Economic effects of making public sector information easily available], RAVI Bedrijvenplatform 00-02 (Amersfoort).
Samuelson, P., 1998, Legally speaking: Does information have to be licensed?, Communications of the ACM 41(9), pp. 15-20.
STIA (Spatial Technologies Industry Association), 2001, Phase 1 report: Increase private sector awareness of, and enthusiastic participation in the National Spatial Data Infrastructure (NSDI), FGDC, http://www.fgdc.gov/library/whitepapers-reports/sponsored-reports/stia/.
Vowles, G., Onsrud, H., Wagner, R., Wilson, C., 2007, Models for rights management, in Workshop Open Geospatial Consortium GeoRM Working Group Conference, 15-20 July 2007, Cambridge, United Kingdom.
Welle Donker, F., 2006, Belemmeringen in de toegankelijkheid van (overheids)geoinformatie [Restrictions in accessibility of (public sector) geo information], in Geoloketten Publicaties, Ed. M. Hoogerwerf Wageningen (Geoloketten Consortium).
Welle Donker, F., Zevenbergen, J., 2007, Value Added GI Services: Respective roles of private and public sector players, in Proceedings of 13th EC-GI&GIS Workshop. INSPIRE Time: ESDI for the Environment 4-6 July 2007, Eds. K. Fullerton and E. Paukenerova, IES Institute for Environment and Sustainability & JRC European Commission, Porto, Portugal, pp. 53-56.
WIPO (World Intellectual Property Organisation), 1996, International protection of copyright and related rights, http://www.wipo.int/export/sites/www/copyright/en/activities/pdf/international_protection.pdf.

4 Update public sector information accessibility policies and open data licences in Europa

Chapters 2, 3 and 5 of this dissertation were written between 2007 and 2009. Since then, there have been a number of developments in the field of public sector information re-use. One of these developments is the emergence of open data. In the previous chapters, legal and financial aspects of public sector information access regimes were discussed. This chapter starts with a description of the revised PSI Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information. This Directive and Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) provide the basis for implementing open data policies, as described in Chapters 6, 7 and 8. This chapter continues with an update of the licence framework discussed in Chapter 3. This update was written for the European Location Framework project. The chapter concludes with an analysis of the various open data licences currently in use in Europe.

§ 4.1 Emergence of open data

The idea of open data, *i.e.* data that are freely available to everybody to (re-)use without restrictions, is not a new concept. In 1942, the sociologist Robert King Merton explained the importance of research results to be freely accessible to all. All researchers should contribute to a "common pot" and give up intellectual property rights to allow knowledge to move forward (Chignard, 2013). The concept of open access to scientific data was also adopted by International Council for Science when the World Data Center System was established in 1958.[31]

In the digital age where information can be accessed and shared easily, science and technology ministers of all nations of the Organisation for Economic Co-operation

31 Several World Data Centres were established around the world to minimize the risk of data loss and to maximise data accessibility, see http://www.icsu-wds.org/organization

and Development (OECD) recognised that fostering broader, open access to and wide use of research data would enhance the quality and productivity of science systems worldwide. In 2004, the ministers adopted a Declaration on Access to Research Data from Public Funding, and asked the OECD to take further steps towards proposing Principles and Guidelines on Access to Research Data from Public Funding. The OECD published these Principles and Guidelines for Access to Research Data from Public Funding in 2007.[32]

In 2007, thirty open data pioneers met in Sebastopol, California to write eight open data principles, intended for adoption by US presidential candidates (Chignard, 2013). Among these pioneers was Lawrence Lessing, the founder of Creative Commons licences. The objective was to find a simple way to express values about how the government should make their data available in a way that enables a wider range of people to help make the government function better.[33]

The eight principles – that data should be complete, primary, timely, accessible, machine-processable, non-discriminatory, non-propriety and licence-free – formed the foundation for the open data movement since then. In 2010, these eight principles were updated to ten by the Sunlight Foundation and included the principles of data permanence and (no) usage costs (Sunlight Foundation, 2010). In 2014, the open data principles were extended to fourteen by including principles on publishing data with trust and provenance and principles on the openness process (public input, public review and coordination (Tauberer, 2014).

§ 4.1.1 Open data principles

In 2007, thirty open data pioneers met in Sebastopol, California, to write eight open data principles, intended for adoption by US presidential candidates (Chignard, 2013). Among the pioneers of the Open Government Working Group was Lawrence Lessing, the founder of Creative Commons licence suite described in Chapter 3. The objective was to find a simple way to express values about how the government should make their data available in a way that enables a wider range of people to help make the government function better.[34] The Open Government Working Group considered

32 http://www.oecd.org/science/sci-tech/oecdprinciplesandguidelinesforaccesstoresearchdatafrompublicfunding.htm

33 Larry Lessing on Open Government Data Principles, https://www.youtube.com/watch?v=AmlzW980i5A

34 Larry Lessing on Open Government Data Principles, https://www.youtube.com/watch?v=AmlzW980i5A

government data to be open if it was made public in a way that it complied with the eight principles that data should be complete, primary, timely, accessible, machine-processable, non-discriminatory, non-propriety and licence-free. These eight principles formed the foundation for the open data movement since then. In 2010, The Sunlight Foundation updated these eight principles to ten and included the principles of data permanence and (no) usage costs (Sunlight Foundation, 2010).[35] Tauberer proposed in 2014 to extend the open data principles to fourteen by including principles on publishing data with trust and provenance and principles on the openness process (public input, public review and coordination (Tauberer, 2014).

The fourteen principles according to Tauberer are:

1. Information is not meaningfully public if it is not available on the Internet for free.
2. Primary: Primary data is data as collected at the source, with the finest possible level of granularity, not in aggregate or modified forms, including audio-visual content.
3. Timely: Data are made available as quickly as necessary to preserve the value of the data. Data is not open if it is only shared after it is too late for it to be useful to the public.
4. Accessible: Data are available to the widest range of users for the widest range of purposes. Data should be made available in formats that support both intended and unintended uses of the data by being published with current industry standard protocols and formats, preferably open, non-proprietary protocols and formats. Data should be discoverable and be provided with sufficient metadata and documentation so that the user understands the structure of the data.
5. Analysable: Data should be published in a format that is machine-processable, so that users can perform their own analyses without having to rely on government analyses.
6. Non-discriminatory: Data are available to anyone, with no requirement of registration, including access via APIs.
7. Non-proprietary: Data are available in a format over which no entity has exclusive control, *i.e.* in a recommended (open) format that can be processed with non-propriety software.
8. Licence-free. Dissemination of the data is not limited by intellectual property law such as copyright, patents, or trademarks, contractual terms, or other arbitrary restrictions. This includes a requirement to attribute the original source.
9. Permanent: Data should be made available at a stable Internet location indefinitely, *e.g.* through the use of persistent URLs (PURLs) or URIs. When data changes over time, copies of all published versions of the data should be retained and stability of format from version to version should be maintained.

35 In Chapter 7, the ten principles proposed by the Sunlight Foundation are described.

10 Safe file formats: Government bodies publishing data online should always seek to publish using data formats that do not include executable content. Executable content within documents poses a security risk to users as it may be malware. Therefore, documents containing macros should be avoided.

11 Provenance and trust: Published content should be digitally signed or include attestation of publication/creation date, authenticity, and integrity. Digital signatures help data users validate the source of the data they find so that they can trust that the data has not been modified since it was published.

12 Public input: The public is in the best position to determine what information technologies will be best suited for the applications the public intends to create for itself.

13 Public review: Not only the data should be public but the process of data creation should also be transparent.

14 Interagency coordination: interoperability makes data more valuable by making it easier to derive new uses from combinations of data. Public data from different departments should be published in the same standard formats with the same definitions.

§ 4.1.2 Some issues with the 14 open data principles

Even within these principles, there is some tension. For instance, Principle (5) prescribes that data should be analysable and Principle (7) states that data should be published in an open format. However, not all users are familiar with open standards and open software to analyse the data. Therefore, if data are published according to Principles (2) and (4) it may be advisable to publish data both in the original (propriety but a de facto) format.

Another concern is Principle 6 (data available to everybody without prior registration). Data may be available via an Application Programming Interface (API), which allow re-users to acquire a small part of the data without downloading the entire dataset. APIs are a suitable interface for applications that require re-use of dynamic and/or voluminous datasets, *e.g.* real-time traffic information. Government data providers develop APIs to facilitate re-users. The government body may then require that re-users register prior to use and agree with the API service conditions. Such agreement conditions may be used to terminate or deny access by users that are suspected of misusing the data. Although prior registration is in violation of Principle 6, this is not always recognised by government data providers. In addition, APIs can limit the amount of data queried each time (rate limiting) to prevent the server from being overtaxed or to prevent misuse of the data. However, rate limiting also violates Principle 4 (access in bulk) (Tauberer, 2014).

Finally, there is the issue of the effort and resources governments should invest in publishing data according to all principles. To make data accessible and interoperable requires resources and time. Data documentation has to be written, metadata have to be filled according to metadata standards and data formats have to be adapted to an open format. It may be that a shared data standard has to be developed or an existing standard has to be adapted. This shared data format has to be adopted within government through coordination across departments, agencies and other government organisations as part of open data governance. This aspect of open data governance may lead to delays in publishing the data, which is contrary to Principle 3 (timely published).

As will be demonstrated in the following sections, most nations have adopted open data policies that include most of the original eight principles proposed by the Open Government Working Group in 2007. However, the eight principle, licence-free, is still a potential issue.

§ 4.1.3 Adoption of open data policies

The concept of open data gained momentum when on his first day in office in January 2009, President Obama issued a memorandum on transparency and open data, which declared that "openness will strengthen our democracy and promote efficiency and effectiveness in government" (Obama, 2009, p.1). The Executive Order of 2013 ordered that "making open and machine readable the new default for government information", *i.e.* all government agencies were to publish their data in a machine-readable form for free public re-use (Obama, 2013). In 2010, the European Commission published the Digital Agenda for Europe as one of the seven pillars of the Europe 2020 Strategy, which sets objectives for the growth of the European Union by 2020. The Digital Agenda's main objective is to develop a digital single market in order to generate smart, sustainable and inclusive growth in Europe.[36]

The European Commission views opening public data as a way to untap the potential for re-use in new products and services and for efficiency gains in administrations (European Commission, 2011). Other countries, such as Australia, India and Kenya have adopted open data policies with transparency, accountability, public participation and economic potential as the main drivers, although each country has its own specific motivation for opening their data. In the United States, transparency and

36 https://ec.europa.eu/digital-single-market/en/europe-2020-strategy

accountability are the main drivers, whereas in European countries there is more emphasis on innovation and growth, and Australia did not want to fall behind Open Government leadership of the United States (Huijboom and van den Broek, 2011).

§ 4.2 The Amended Public Sector Information Re-use Directive 2013/37/EU

Directive 2003/98/EC on the re-use of public sector information – the so-called PSI Directive – aimed to remove major barriers to re-use of public sector information (PSI) in the European internal market, such as discriminatory practices, monopoly markets and a lack of transparency. Chapter 2 showed that the PSI Directive had a number of shortcomings, such as a lack of clear definitions that allowed room for public sector bodies to offer commercial services in competition with the private sector. The intended ceiling on charges left sufficient room for public sector bodies to charge fees above cost recovery ("cost recovery plus a reasonable rate on return"). Although progress had been made to remove barriers to re-use of PSI since the adoption of the PSI Directive, Member States needed to take further steps to unlock the full potential of PSI for the EU economy (European Commission, 2009).

In 2009, the European Commission recognised that public sector information (PSI) was the single largest source of information in Europe and the potential for re-use of PSI needed to be highlighted in the digital age. (European Commission, 2009). As one of the key actions of the Digital Agenda for Europe[37] was a review of the PSI Directive, the European Commission carried out a round of consultations with stakeholders to seek their views on specific issues to be addressed in the future Commission guidelines in 2010. In addition, the Commission commissioned a number of studies. These studies included a review of studies on PSI re-use and related market studies by Graham Vickery[38], an assessment of the different models of supply and charging for PSI

37 https://ec.europa.eu/digital-single-market/en/our-targets#Our Actions

38 Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Paris, Information Economics: 44, http://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments.

(the POPSIS study)[39] and a study on PSI re-use in the cultural sector. [40] In addition, the European Commission carried out an impact assessment of the proposed revisions of the PSI Directive.[41]

The review highlighted the different ways in which PSI rules were being applied by Member States (European Commission, 2011). In addition, Vickery's review of PSI-re-use studies showed that the overall economic gain from opening up public sector data as a resource for new products and services could be in the order of €40 billion per annum. The Pricing of PSI Study (POPSIS) assessed different models of supply and charging for PSI and their effects through the analysis of 21 case studies, covering a wide range of public sector bodies and different PSI sectors. The case studies showed that for public sector bodies that charged for PSI re-use, the revenue was relatively to extremely low in comparison to the total budget of the public sector body. The study concluded that lowered charges could lead to more economic activity, market dynamism, innovation and employment, and might also entail efficiency gains for the public sector body (de Vries *et al.*, 2011). The study on PSI re-use in the cultural; sector concluded that overall, the revenue resulting from PSI-re-use for cultural institutions was relatively limited and very few cultural institutions are dependent on revenue from PSI re-use. However, the current revenue was important to enable future re-use and future development of services. The institutions also indicated that digitising content was the limiting factor, in terms of costs and effort, to enable re-use. The institutions expressed concerns about becoming entirely dependent on public money (Clapton *et al.*, 2011). After the review, the 2003 PSI Directive was amended in 2013 by Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information and came into force on 17 July 2013.

The main changes of the 2013/37/EU Amended Public Sector Information Re-use Directive were that the Directive made it a general rule that all documents made accessible by public sector bodies can be re-used for any purpose, commercial or non-commercial, unless protected by third-party copyright. The scope was extended to libraries, museums and archives. Charges are limited to the marginal costs of distribution of the data, unless duly justified. Data are to be published in machine-

39 de Vries, M., L. Kapff, M. Negreiro Achiaga, P. Wauters, D. Osimo, P. Foley, K. Szkuta, J. O'Connor and D. Whitehouse (2011). Pricing of Public Sector Information Study. Models of Supply and Charging for Public Sector Information (ABC) Final Report. Brussels, Deloitte Consulting, 403, https://ec.europa.eu/digital-single-market/en/news/pricing-public-sector-information-study-popsis-models-supply-and-charging-public-sector.

40 Clapton, G., M. Hammond and N. Poole (2011). PSI re-use in the cultural sector - final report. Curtis+Cartwright Consulting Ltd. Guildford, European Commission: 43, http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/cc462d011_1_1final_report.pdf

41 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SEC:2011:1551:FIN:EN:PDF

readable formats, and are to be accompanied by metadata and cross-lingual search facilities to enable effective re-use. The Directive also requires Member States to establish independent regulatory authorities to deal with complaints.

§ 4.2.1 Open data principles for re-use of PSI

The Amended PSI Re-use Directive encourages implementation of open data policies. Recital 3 states that "open data policies: which encourage the wide availability and re-use of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints, and which promote the circulation of information not only for economic operators but also for the public, can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulate economic growth and promote social engagement …". However, the Amended PSI Re-use Directive does not address all open data principles; it merely sets recommendations for publishing documents as primary data, the use of open and machine-readable formats, and open licences. Other open data principles, such as timely publication and permanent (data available at a stable internet location indefinitely), are not addressed in the directive.

The 2013 Amended PSI Re-use Directive recommends that to facilitate re-use, public sector bodies should, where possible and appropriate, make documents available through open and machine-readable formats and together with their metadata, at the best level of precision and granularity, in a format that ensures interoperability and recommends consistency with the principles governing the compatibility and usability requirements for spatial data under Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) (recital 20).

The 2013 Amended PSI Re-use Directive promotes the use of open licences available online (recital 26) but does not mandate the use of open licences. In the Implementation Guidelines, the European Commission recommends the use of Creative Commons licences (European Commission, 2014). Section 4.3.1 describes the recommendations of the European Commission.

§ 4.2.2 Still room for charges for public sector information

The 2013 Amended PSI Re-use Directive applies the principle that where charges are made by public sector bodies for the re-use of documents, those charges should in principle be limited to the marginal costs. However the necessity of not hindering the normal running of public sector bodies that are required to generate revenue to cover a substantial part of their costs relating to the performance of their public tasks or of the costs relating to the collection, production, reproduction and dissemination of certain documents made available for re-use should be taken into consideration. In such cases, public sector bodies should be able to charge above marginal costs. Those charges should be set according to objective, transparent and verifiable criteria and the total income from supplying and allowing re-use of documents should not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment (recital 22). Libraries, museums and archives are allowed to charge above marginal costs in order not to hinder their normal running. When calculating the charges, the cultural institutions could consider the prices charged by the private sector for the re-use of identical or similar documents when calculating a reasonable return on investment (recital 23).

In the 2003 PSI Directive, the decision whether or not to authorise re-use remained with the Member States or the public sector body concerned. Under the 2013 Amended PSI Re-use Directive, a clear obligation for Member States to make all documents re-usable unless access is restricted or excluded under national rules on access to documents and subject to the other exceptions laid down in this Directive. The amendments made by this Directive do not seek to define or to change access regimes in Member States, which remain their responsibility (recital 8). Thus, the 2013 Amended PSI Re-use Directive does not provide a right to information.

§ 4.2.3 Compliance with protection of personal data principles

The Amended Directive should be implemented and applied in full compliance with the principles relating to the protection of personal data in accordance with Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.[42] In particular, it is worth noting that, according to

42 OJ L 281, 23.11.1995, p. 31.

that Directive, the Member States should determine the conditions under which the processing of personal data is lawful. Furthermore, one of the principles of that Directive is that personal data must not be processed further to collection in a way incompatible with the specified, explicit and legitimate purposes for which those data were collected (recital 11).

As described in Section 1.6, there is a tension between open data and protection of personal data. As more data become available as open data, the risk of misuse of personal data increases although open data may not seem to be personal data on first glance, especially when it is anonymised or aggregated. However, the data may become personal data by combining it with other data or when de-anonymised (Kulk and van Loenen, 2012). In addition, with apps and tools based on open government data, there is nothing to prevent the use of open data for profiling, data mining and other activities, which have privacy implications for individuals (Scassa, 2014, p.407). During the review process of the PSI Directive, the European Data Protection Supervisor (EDPS) was not consulted. In 2012, EDPS issued an Opinion on the "Open-Data Package" (of which the Amended PSI Re-use Directive was a part). In this Opinion, the EDPS recommended that the Amended PSI Re-Use Directive should address data protection more specifically (EDPS, 2012, p.5). The EDPS made some specific recommendations, including, inter alia, that public sector bodies should carry out a data protection assessment prior to publishing open data (EDPS, 2012, p.7). However, the Amended PSI Re-use Directive did not adopt the EDPS's recommendations.

§ 4.2.4 Level playing field

The Amended PSI Re-use Directive recognises that a level playing field at Union level is required in terms of whether or not the re-use of documents is authorised, as this cannot be achieved by leaving it subject to the different rules and practices of the Member States or the public sector bodies concerned. To prevent different rules in different Member States acting as a barrier to the cross-border offer of products and services, and to enable comparable public data sets to be re-usable for pan-European applications based on them, a minimum harmonisation is required to determine what public data are available for re-use in the internal information market, consistent with the relevant access regime (recital 6). Recital 13 states that where any document is made available for re-use, the public sector body concerned should retain the right to exploit the document. The Amended PSI Re-use Directive allows room for exceptions to the charges ceiling of marginal costs for public sector bodies that are required to generate revenue and for specifically excepted documents. Given the fact that the concept of "public task" is still not defined in the amended directive, there is room for such public sector bodies to define publication of (semi-)commercial products as a public task.

§ 4.2.5 Redress by an impartial body

The means of redress should include the possibility of review by an impartial review body. That body could be an already existing national authority, such as the national competition authority, the national access to documents authority or a national judicial authority. That body should be organised in accordance with the constitutional and legal systems of Member States and should not prejudge any means of redress otherwise available to applicants for re-use. It should however be distinct from the Member State mechanism laying down the criteria for charging above marginal costs. The means of redress should include the possibility of review of negative decisions but also of decisions, which, although permitting re-use, could still affect applicants on other grounds, notably by the charging rules applied. The review process should be swift, in accordance with the needs of a rapidly changing market (recital 28). Although the amended directive is more specific on redress procedures, there are no time limits set to deal with complaints, thus appeal procedures described in Chapter 1 of this dissertation, could still take a long time.

§ 4.3 Open licences in Europe[43]

Since the Digital Agenda for Europe, many of the EU Member States have adopted open data licences for publishing open data. This section provides an overview of the various open data licences employed by National Mapping and Cadastre Authorities, and an analysis to which extent these open licence contribute to legal interoperability in a pan-European project.

§ 4.3.1 European Commission recommendations for open licences

The European Commission recommends the use of open standard licences for publishing public sector data, e.g. Creative Commons licences. Open standard licences could allow

43 The following sections build on a study carried out for the European Location Framework platform to research the interoperability of the various open licences employed by the participating National Mapping and Cadastre Authorities (see van Loenen, B. and F. Welle Donker, 2015. Open licences for ELF. (p. 17). Delft: Knowledge Centre Geo-Information Governance).

the re-use of PSI without the need to develop and update custom licences at national or sub-national level. Especially the CC0 public domain dedication is interesting as a legal tool as it "allows waiving copyright and database rights on PSI, it ensures full flexibility for re-users and reduces the complications associated with handling numerous licences, with possibly conflicting provisions" (European Commission, 2014, p.2).

Further recommendations of the Commission include that the open standard licence should:

- Provide a reference to the conditions under which re-use is allowed should appear prominently at the point of display of, or accompanying, the information.
- Define the temporal and geographical scope of the rights covered by the licensing agreement.
- Define the types of rights granted and the range of re-use allowed.
- Grant a worldwide (to the extent allowed under national law), perpetual, royalty-free, irrevocable (to the extent allowed under national law) and non-exclusive rights to use the information covered by the licence.
- Explicitly set out the rights not covered by the licence.
- Define the types of right granted (copyright, database right, and related rights) broadly.
- Use the broadest possible wording to refer to what can be done with the data covered by the licence (terms, such as: use, re-use and "share" can be further described by an indicative list of examples).

The Commission continues "where licences are required by law and cannot be replaced by simple notices, it is advisable that they cover attribution requirements only, as any other obligations may limit licensees' creativity or economic activity, thereby affecting the re-use potential of the documents in question." (European Commission, 2014, p.3). However, the use of licences that require source attribution is in violation of Principle 6 listed in Section 4.1.1.

Several licences comply with the principles of 'openness'. They have been translated into many languages, centrally updated, and already used extensively worldwide. Open standard licences, for example the most recent Creative Commons (CC) licences (version 4.0), could allow the re-use of PSI without the need to develop and update custom-made licences at national or sub-national level.

In addition, the LAPSI[44] 2.0 thematic network discourages organisations to use their own open government licence since it raises all kinds of interoperability and licence

44 LAPSI stands for Legal Aspects of Public Sector Information.

management issues. If countries still prefer to do so, the LAPSI 2.0 thematic network advises to create only an Attribution-only licence (Tsiavos, 2012).

§ 4.3.2 Recommended open licences

The European Commission as well as the LAPSI 2.0 thematic network recommends for open government data the use of the CC0 declaration or, if CC0 appears not feasible or possible, a CC-BY 4.0 licence. In Tables 4.1 and 4.2, the main characteristics of CC0 and CC-BY are described.

§ 4.3.2.1 CC0

The Creative Commons Zero declaration (CC0) allows one to waive all copyrights and related or neighbouring rights in one's work, such as moral rights (to the extent that these can be waived), publicity or privacy rights, rights protecting against unfair competition, and database rights and rights protecting the extraction, dissemination and re-use of data.[45]

- Affirmer overtly, fully, permanently, irrevocably and unconditionally waives Copyright and Related Rights and associated claims and causes of action in the Work in all territories worldwide for the maximum duration provided by applicable law or treaty, in any current or future medium and for any number of copies, and for any purpose whatsoever, including without limitation commercial, advertising or promotional purposes.
- Work is provided "as-is".
- No trademark or patent rights held by Affirmer are waived.

TABLE 4.1 CC0

45 See Loenen, B. van, Janssen, K. and Welle Donker, F.M. (2012). Towards true interoperable geographic data: developing a global standard for geo-data licences. In K. Janssen and J Crompvoets (Eds.), Geographic Data and the Law. Defining New Challenges (pp. 19-36). Leuven: Leuven University Press; see also http://creativecommons.org/publicdomain/zero/1.0/legalcode

§ 4.3.2.2 CC-BY 4.0

 **Attribution**	• You let others Share (copy and redistribute) the material in any medium or format and/or Adapt (remix, transform, and build upon the material) for any purpose, even commercially - but only if they give appropriate credit, provide a link to the licence, and indicate if changes were made. • Non-sub licensable licence grant. • No endorsement (no use in any way that suggests the licensor endorses the use or the user). • Work is provided "as-is".

TABLE 4.2 CC-BY 4.0

In November 2013, the CC-BY 4.0 licence replaced the CC-BY-3.0 version of 2007. There are a number of differences between the two versions. The main differences are the licence scope beyond copyright (*e.g.* database rights) and the manner of attribution. Below, we list the relevant differences, as listed by https://wiki.creativecommons.org/License_Versions.

§ 4.3.2.3 Sui generis database rights

The CC 4.0 international suite licences database rights along with copyright. When the CC 4.0 licence is used for a database, sui generis database rights are implicated, whether or not copyright is implicated. The 3.0 version does not mention sui generis rights. In the ported 3.0 licences for jurisdictions where those rights exist, these rights are addressed according to CC's 3.0 database rights policy. Under this policy, sui generis rights must be licenced but licence restrictions for uses triggering database rights must also be waived. With the switch from ported licences to international licences, version 4.0 explicitly addresses licence conditions applicable to sui generis rights. Version 2.0 does not address sui generis database rights at all.

§ 4.3.2.4 Moral rights and trademark rights

There are other differences in the licence scope beyond copyright, such as the treatment of moral rights and trademark rights. Versions 1 to 2.5 did not address moral rights and version 3.0 did not include a waiver of moral rights. Version 4 harmonised the treatment of moral rights and limited the role of moral rights where the exercise of those rights by licensors would prevent uses the CC licences are designed to

permit, but only to the extent those rights are held by the licensor and may be waived or not asserted.[46]

Creative Commons licences do not cover trademark and patent right. In version 4.0, this was made explicit to avoid confusion.

§ 4.3.2.5 Attribution and marking

In version 4.0, a licensor may request removal of attribution by users whether the work is modified or not. In earlier versions of CC, the title of the work was required in the attribution. In version 4.0, this is no longer a requirement to increase flexibility and ease of compliance.

In version 4.0, (URI) is required for proper attribution if it is reasonably practicable to include. In previous versions, a URI is only required if it contains copyright notices of licensing information.

Version 4.0 includes a "no endorsement" clause, *i.e.* the licence is clear that the user is not granted permission to suggest the licensor endorses their use. In earlier versions, this is also the case but it was never explicitly mentioned. In version 4.0, this clause is expressed as a limitation on the rights granted by the licensor.

In version 4.0, licencees are required to indicate if they have made modifications to the licenced material. In version 3.0, this obligation only applies if they result in the creation of an adaptation.

§ 4.3.3 European Location Framework Project

The European Location Framework (ELF) project was established to provide a practical implementation of INSPIRE and to complement the activities of European national mapping, cadastral and land registry authorities. The intention of the ELF platform is to provide a single point of access to harmonised pan-European maps, geographic and land information from official sources to facilitate the wider use of geo-information and enable the creation of innovative value-added services (EuroGeographics, 2016). One of the

46 https://wiki.creativecommons.org/wiki/Version_4#Moral_rights.3B_similar_rights

objectives of the ELF project is to create a policy for sustainable interoperability to ensure that the data from the ELF platform will remain available for use and re-use after the end of the pilot. This policy should be consistent with the INSPIRE Directive rules for data and service sharing and network services, and the 2013 Amended PSI Re-use Directive. A number of the participating National Mapping and Cadastre Authorities, which provide data via the ELF platform, are self-funding authorities, i.e. they are required to generate sufficient revenue to cover a substantial part of their operating costs. ELF aims to establish a financially viable operational framework through agreements, which encourages open licences and minimum to no charge licence fees.

§ 4.3.4 Open licences in the ELF network

§ 4.3.4.1 Open licences currently in use in the ELF network

Table 4.3 provides an overview of the in the ELF network existing open licences.[47]

COUNTRY	LICENCE	KEY CHARACTERISTICS
Czech Republic	Unknown	· Data can be downloaded without reference to a licence
Denmark	Conditions for use of open public geographic data	· Register before access · Right to copy, distribute and publish, adapt and combine with other material, exploit commercially and non-commercially · Attribution required + link + note on whether the data were retrieved from the Licensor or through a data service · Copy of conditions available to third parties · Same conditions apply if forwarding data to a third party · No guarantee for the continued availability of the data · Licensor may change the licence and licence conditions at all times

47 Czech Republic CUZK Geoportal at http://geoportal.cuzk.cz/; Danish Geodata Agency at http://eng.gst.dk/media/gst/2364686/Conditionsforuseofopenpublicgeographicdata.pdf; Eurogeographics at http://www.eurogeographics.org/form/topographic-data-eurogeographics; National Land Survey of Finland at http://www.maanmittauslaitos.fi/en/opendata; France at https://www.etalab.gouv.fr/licence-ouverte-open-licence; the Netherlands Kadaster at http://www.kadaster.nl/web/artikel/Alle-producten-1/TOPvector.htm and the Dutch PDOK geoportal https://www.pdok.nl/en/products/downloading-data-pdok; Norwegian Mapping Authority at http://kartverket.no/en/Kart/Gratis-kartdata/Open-and-Free-geospatial-data-from-Norway/; the Surveying and Mapping Authority of the Slovenian Republic at http://www.gu.gov.si/en/services/free_access_database/; the Spanish Instituto Geográfico Nacional at http://www.ign.es/ign/main/index.do?locale=en and the Spanish Catastro at http://www.sedecatastro.gob.es/; UK Ordnance Survey at http://www.ordnancesurvey.co.uk/oswebsite/products/os-opendata.html.

Eurogeo-graphics	EuroGlobalmap licence	· Right to reproduce, distribute, adapt, extract, re-utilize and communicate to the public for any legal purpose including commercial exploitation · Attribution required + link · Sublicensing allowed · No endorsement · As is provided · No right to use trademark
Finland	CC-BY 4.0	· *See Table 4.2*
France	Licence ouverte	· Right to reproduce, copy, publish, transmit, disseminate, redistribute the information, to adapt, modify, transform and extract from the information, to exploit the information commercially and non-commercially · Attribution required (name + date last updated or URL link) · No endorsement · As is provided · No misleading third parties · Licence is compatible with CC-BY 2.0, OGL (UK) and ODC-BY (Open Knowledge Foundation)
Nether-lands	CC-BY version 3.0 and CC-BY 4.0	· *See Table 4.2*
Nether-lands	CC0	· -
Norway	CC-BY 4.0	· Register before access.
Slovenia	Open data licence Slovenia ("CC-like")	· Only available in Slovenian language. It is very similar with Danish license "Conditions for use of open public geographic data" · Data can be copied, distributed, published, re-used, and adapted in new products for com-mercial or non-commercial use · Attribution required (name source + year) · As is provided; SI NMCA does not take any liabilities regarding data/service quality and continued availability
Spain	"CC-BY like"	· Request for attribution as "© IGN. National Geographic Institute of Spain"
Spain	Resolution of 23 March 2011	· Register before access · Data must be transformed when re-using data · Authorisation for re-use and transform is granted for a period of 10 years · Attribution required (name source + access date) · As is provided · No guarantee for the continuous availability of the Service
United Kingdom	Ordnance Survey Open Data Licence (based on OGL Version 2.0)	· Right to copy, publish, distribute, transmit, adapt, combine and exploit the information commercially and non-commercially · Attribution required (Name source + year) · Attribution passed on in any sub-licences · No endorsement · As is provided

TABLE 4.3 Overview open data licences used within ELF network

We see that most countries build in one way or another on the framework of Creative Commons. Finland, Norway and the Netherlands are using CC-BY 3.0/4.0 and/or CC0; the other open licences are similar in the rights granted, the licence conditions, the rights not licenced, and the disclaimer and limitation of liability. Sometimes, there are differences in the wording of the use rights and sometimes, issues are addressed

that may not need to be addressed in an open data licence (e.g. no guarantee on the data availability).

§ 4.3.5 Differences in open licences

Although many of the open licences build on the Creative Commons suite framework, and have many similarities, there are also a number of differences. These differences may pose barriers for some ELF network participants.

§ 4.3.5.1 Denmark

The Danish Open Data licence differs from CC-BY 4.0. In Denmark, users must register first. The attribution is very specific (name of Agency + name of dataset + retrieval date + data retrieved from Licensor or through a data service). If the data are made available to third parties, the original attribution licence terms must be available to these third parties, *e.g.* by using a link. In addition, there is an explicit clause that the Authority does not guarantee the continued availability of the data and that the Authority may at any time modify the right to use the data and under what circumstances. This last clause means that the Danish Open Licence for the data is revocable at any time.

Prior registration before access should not be considered as a barrier to re-use via the ELF platform. Although CC prohibits the use of technical protection measures to prevent others from exercising the licenced rights, prior registration as such does not prevent the usage of the data. However, prior registration may be viewed as a barrier by re-users outside the ELF platform.

The main differences between the Danish licence and CC-BY 4.0 are the specific attribution requirement and revocable data licence. As far as specific attribution is concerned, CC licences have a flexible attribution requirement. The proper method for giving credit will depend on the medium, means, and context in which a licencee is redistributing licenced material. The user may satisfy the attribution requirement if a link is provided to a place where the attribution information may be found.

As far as revocable licences are concerned, CC-BY licences for data are irrevocable by definition. However, with every updated version of the data, a new licence could be reapplied. The old licence would still apply to all data obtained under the older licence terms. However, as the value of ELF data lies in the actuality of the data, older downloaded versions would probably cease to be in use within foreseeable time.

The specific attribution requirements and the revocable licence may pose a barrier to international use of ELF data. The first barrier can only be overcome by making the attribution requirement more flexible in line with CC-BY 4.0. The latter barrier could be overcome by notifying potential re-users with a disclaimer on the ELF platform.

§ 4.3.5.2 France

The French licence ouverte declares to be compatible with CC-BY 2.0. The differences between version CC-BY 2.0 and CC-BY 4.0 lie in the application to:

- sui generis database rights;
- the treatment of moral right;
- an explicit waiver of rights to enforce, and grant permission to circumvent technological protection measures;
- automatic reinstatement after termination if violations occur;
- attribution-specific elements.

In Section 4.3.2, the main differences between versions 2.0 and 4.0 were described. In CC-BY 2.0 the title of the work is required as part of the attribution; in CC-BY 4.0, this requirement was eliminated to increase flexibility and ease of compliance. However, this point should not be an issue as the French licence ouverte requires attribution by acknowledging its source "(at least the name of the « Producer ») and the date on which it was last updated. The « Re-user » may fulfil this condition by providing one or more hypertext links (URL) referring to the « Information » and effectively acknowledging its source."[48]

In the CC-BY 4.0 version, licencees are required to indicate if they made modifications to the licenced material. This obligation applies whether or not the modifications produced adapted material. In 3.0 and earlier licence versions, the indication of changes is only required if a derivative is created. This clause does not specifically appear in the French licence ouverte.

The differences to attribution between CC-BY 4.0 and the French licence ouverte should not pose a legal barrier as such for international use of ELF data.

48 http://ddata.over-blog.com/xxxyyy/4/37/99/26/licence/Licence-Ouverte-Open-Licence-ENG.pdf

§ 4.3.5.3 Spain

The Spanish Cadastre uses an open data licence, which is not compatible with CC-BY licence. The main differences between the Catastro licence and CC-BY are the requirement that the data must be transformed and the licence term of 10 years. The latter should not pose a barrier, as most re-users will update the data within the 10-year period. However, the former requirement may pose an enforcement problem when the data is re-used by users outside Spain. The licence condition implies that data cannot be hosted by ELF and can only be invoked from the Spanish web service. The ELF platform should notify potential re-users of the transformation requirement and the licence term limitation. However, it is expected that ELF users will transform the data anyway by combining with other data.

§ 4.3.5.4 United Kingdom

The standard UK Open Government Licence 2.0 is similar to a CC-BY licence, and is compatible with CC-BY 4.0. However, the UK Ordnance Survey has added a clause to the standard OGL licence making the Ordnance Survey Open Data licence incompatible with CC-BY 4.0. In the Ordnance Survey OpenData licence, a user has to include the same acknowledgement requirement (name source + year) in any sub-licences of the data and a requirement that any further sub-licences do the same.[49] CC-BY 4.0 prohibits such a restriction. If the standard OGL 2.0 licence were applied to the UK contribution to ELF data, there would be no impediment to using CC-BY 4.0 for international use. However, the Ordnance Survey OpenData licence is currently incompatible with CC-BY 4.0. The ELF platform could provide a notification alerting potential re-users of the sublicensing requirement.

§ 4.3.6 Summary open data licences currently in use

Most of the countries build in one way or another on the framework of Creative Commons with CC-BY the most common licence. It would, therefore, be the most obvious to recommend that ELF would adopt the CC-BY 4.0 licence for open data. Although some of the identified differences, such as user registration prior to access or the CC-BY 2.0 compatibility of the French licence ouverte, will not pose a barrier to CC-

49 http://www.ordnancesurvey.co.uk/business-and-government/licensing/using-creating-data-with-os-products/os-opendata.html

BY 4.0 compatibility, there are, however, some differences that may pose a barrier to adopting a CC-BY 4.0 licence. The current open data licences for Spanish Cadastral Data and for the UK Ordnance Survey data are incompatible with CC-BY 4.0. If ELF decides to invoke the national services rather than hosting data, a disclaimer and a link to the specific licence conditions should suffice. The national authority will be responsible for enforcement of the specific licence conditions.

§ 4.3.6.1 Remaining questions

To comply with the Spanish open data licence for Cadastral data, ELF should only use invoking services. In this case, national open data licence will be applicable to the data. This raises the first question if the ELF platform needs to use a separate ELF licence for the web service, and if so, what type of licence.

If a CC-BY-type licence were chosen for the invoked ELF open data, the second question that needs to be answered, is how attribution should be given, given the fact that CC-BY does not allows specific attribution in a specific place. The seemingly most logical way of recommending attribution would be to allow for multiple source attribution, *e.g.* "Contains ELF data + year". Similar to the CC-BY licence, ELF should allow for flexibility in attribution for compliance reasons.

Another remaining issue is whether the ELF platform should refer to a single ELF licence (for invoked data) or to link to the individual licences per data holder? If the latter is the case, will ELF provide a link to the information provided by the national authority? This may pose problems with missing information (*e.g.* Czech Republic), mismatch in information (*e.g.* the Netherlands) or language issues (*e.g.* Slovenia). It may be preferable to refer to an ELF page with specific information about the licences of the individual data holders in multiple languages. This page should also describe the main differences between the different open licences. However, licence changes in the individual countries need to be monitored regularly.

The fourth question that needs to be addressed is how to deal with the specific licence differences as these differences cannot be addressed in a single licence. For instance, the UK requirement of users having to pass on attribution requirement to all subsequent licences is incompatible with the CC-BY 4.0 licence and/or CC0 declaration used in Finland, Norway and the Netherlands, as none of the Creative Commons licences grants permission to sublicence the licenced material. It would not be practical to include such a clause in an ELF open licence, as this would cause a problem of having a more restrictive licence for a product obtained through ELF rather than obtained through a national service. Especially for data licenced under a CC-BY 4.0 or a CC0 licence.

The last question that needs to be addressed is how to deal with data obtained via services when the national licence is revoked or amended.

References

Chignard, S. (2013). A brief history of open data. ParisTech Review, http://www.paristechreview.com/2013/03/29/brief-history-open-data/, accessed April 29, 2016.

Clapton, G., M. Hammond and N. Poole (2011). PSI re-use in the cultural sector - final report. Curtis+Cartwright Consulting Ltd. Guildford, European Commission, http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/cc462d011_1_1final_report.pdf: 43, accessed 10 May 2011.

de Vries, M., L. Kapff, M. Negreiro Achiaga, P. Wauters, D. Osimo, P. Foley, K. Szkuta, J. O'Connor and D. Whitehouse (2011). Pricing of Public Sector Information Study (POPSIS) - Summary Report. Brussels, Deloitte Consulting CVBA,, https://ec.europa.eu/digital-single-market/en/news/pricing-public-sector-information-study-popsis-summary-report, accessed October 2011.

EDPS (European Data Protection Supervisor) (2012). Opinion of the European Data Protection Supervisor on the 'Open-Data Package' of the European Commission including a Proposal for a Directive amending Directive 2003/98/EC on re-use of public sector information (PSI), a Communication on Open Data and Commission Decision 2011/833/EU on the reuse of Commission documents. Brussels: 13. https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2012/12-04-18_Open_data_EN.pdf, Accessed October 28, 2016.

EuroGraphics (2016). What is the European Location Framework?, http://elfproject.eu/sites/default/files/ELF2_leaflet_2016_FINAL.pdf: 2, accessed 30 March 2016.

European Commission (2009). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Re-use of Public Sector Information- Review of Directive 2003/98/EC – [SEC(2009) 597]. Brussels, European Commission, http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52009DC0212, accessed May 26, 2015.

European Commission (2011). Commission Staff Working Paper. Impact Assessment Accompanying the document Proposal for a Directive of the European Parliament and the Council amending European Parliament and Council Directive 2003/98/EC on the re-use of public sector information {COM(2011) 877 final} {SEC(2011) 1551 final}. Brussels, https://ec.europa.eu/digital-agenda/en/news/commission-staff-working-paper-impact-assessment-accompanying-document-proposal-directive: 84, accessed December 12, 2011.

European Commission (2011). Open Data: An engine for innovation, growth and transparent governance. Brussels. COM(2011) 882 final: 13, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF, accessed June 10, 2013.

European Commission (2014). Guidelines on recommended standard licences, datasets and charging for the reuse of documents (2014/C 240/01). Brussels (Official Journal of the European Union). 57: 10, http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2014.240.01.0001.01.ENG&toc=OJ:C:2014:240:TOC, accessed September 11, 2015.

Huijboom, N. and T. van den Broek (2011). Open Data: an international comparison of strategies, European Journal of ePractice no. 12 (March/April 2011): 4-16, http://share-psi.eu/papers/TNO-epractice.pdf, accessed April 14, 2015.

Obama, B. (2009). Transparency and Open Government. Memorandum for the Heads of Executive Departments and Agencies. Washington, The White House, https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment, accessed March 29, 2016.

Obama, B. (2013). Executive Order -- Making Open and Machine Readable the New Default for Government Information. Washington DC, https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-, accessed May 10, 2013.

Scassa, T. (2014). "Privacy and Open Government". Future Internet 6(2), pp.397-413. doi: 10.3390/fi6020397.

Sunlight Foundation. (2010). Ten Principles for Opening Up Government Information, accessed August 31, 2010, from https://sunlightfoundation.com/policy/documents/ten-open-data-principles/.

Tauberer, J. (2014). Open Government Data: The Book. The principles, practices, and a history of the open government data movement, https://opengovdata.io/: 196, accessed October 6, 2015.

Tsiavos, P. (2012). Licence Interoperability Report, LAPSI 2.0, http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=3899, Accessed April 19, 2016.

5 Public sector geo web services: which business model will pay for a free lunch?

Frederika Welle Donker

In: B. van Loenen, J.W.J. Besemer, J.A. Zevenbergen (eds.), 2009. *SDI Convergence: Research, Emerging Trends, and Critical Assessment*, Delft (Netherlands Geodetic Commission - (NCG), Vol. 48, pp.35-51, ISBN 978-90-6132-310-5.

Abstract

Geo-information (GI) is increasingly having a bigger impact on socio-economic benefits. Over the last decade, use of GI has shifted from a specialised GIS niche market to serving as a direct input to planning and decision-making, public policy, environmental management, readiness to deal with emergencies, creation of value added products, citizen mobility and participation, and community platforms.
The emergence of Google Earth and Google Maps has created a geo-awareness and has catalysed a thirst for custom-made geo-information. Governments possess, often high-quality large-scale GI, primarily created, collected, developed and maintained to support their public tasks. This rich source of GI begs to be used and reused both within the public sector and by society. Both the INSPIRE Directive (2007/02/EC) and the Directive on re-use on Public Sector Information - the so-called PSI Directive - (2003/98/EC) underwrite the philosophy of "collect once, reuse many times".
Web services are an effective way to make public sector geo-information available. They allow information to be accessed directly at the source and to be combined from different sources. However, the costs of web services are high and revenues do not always cover the costs. Assuming that there is no such thing as a free lunch related to public sector GI (Longhorn and Blakemore, 2008), which business models and which financial models form the basis for public sector geo web services? This article explores the different models currently in use and illustrates them with examples.

Keywords: geographic information; public sector web services; business models; financial models; revenue models

§ 5.1 Introduction

The terms "geographic information", "geographic data", "spatial information" and "spatial data" are interchangeably used as synonyms. For the purpose of this article, only the term geographic information (GI) will be used. Access to GI is of vital importance to the economic and social development of the nation. Nations around the world are developing geographic information infrastructures (GIIs), also referred to as spatial data infrastructures (SDIs), with access to GI at the core. For more advanced GIIs (re)use is considered the driver of a GII (van Loenen, 2006). One way to facilitate reuse of GI is through web services. The INSPIRE Directive even requires that as part of developing geo-information infrastructures network services should be used. National GIIs are now evolving from first to second generation GIIs. The existence of web services are regarded as the main technological drivers of second generation GIIs because they can fulfil the needs of users and improve the use of data (Crompvoets *et al.*, 2004; Rajabifard *et al.*, 2003). This article will give an inventory of the different models currently in use and illustrate them with examples. In Section 5.2 a description of various types of web services will be provided, including a case study illustrating costs involved setting up a commercial Web 2.0 platform and the potential revenue web services can generate. Section 5.3 will supply a theoretic framework for business models with a breakdown of the four essential parts of a successful business model. Section 5.4 will build on the business model framework with a framework for financial models, including various cost and revenue models and price strategies. In Section 5.5, the summary will show which business model and which financial model will be most suited and robust in a given situation. It will also show some current pricing trends for public sector geographic information (PSGI) in Europe. Section 5.6 will finish with some conclusions and offer some recommendations for public sector web services.

§ 5.2 Web services

§ 5.2.1 Different web services

A web service is a platform that is accessible with open standard protocols such as SOAP and XML. A web services sends a request from the client-computer to a server. The server sends queries to the appropriate source servers and transmits a reply back to the client-computer. The advantage is that data is queried at the original source so

it is as current as possible. There are a number of different types of GI web services, which roughly fall into two categories: web services using Open Geo Consortium (OGC) standards and web services using ICT standards.

§ 5.2.1.1 Open Geo Consortium web services

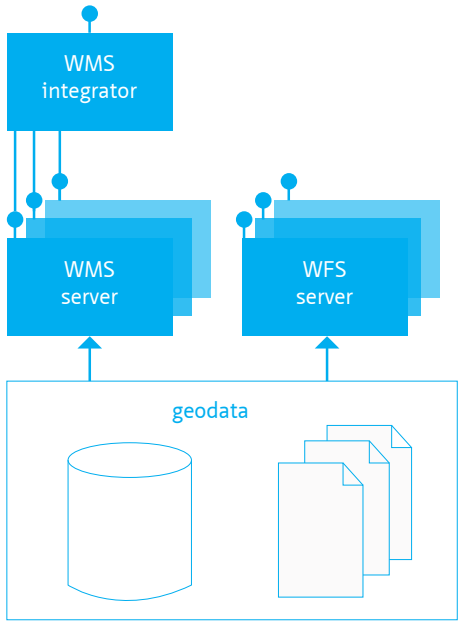


FIGURE 5.1 Serving geo-information using WMS, WFS and WIS (source: http://www.geoloketten.nl/wms_integrator_services.html)

The main OGC standards used for web services are Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS) and Web Integrator Service (WIS). WMS only produces a static image on screen from raster files. Because no actual data is transferred, no information can be downloaded. Therefore, it is easier to comply with protection of intellectual property rights (IPR). With WFS, discrete features (points, lines, polygons) are downloaded in XML to the client-computer. The same applies to a WCS whereby entire coverages (sets of features) are also downloaded in XML. Data from WFS and WCS are suitable for interpretation, extrapolation and other forms of analysis. Because the data itself is transferred from the server, measures to protect data subject to IPR are harder to implement for WFS and WCS than for WMS whereby no data is transferred. Therefore, WFS and WCS are probably more suitable for fee-

based web services. A WIS is a service that can horizontally integrate various WMSs. Horizontal integration of WMS means that different WMSs of different organisations are bundled into one new WMS. A WIS allows for instance to integrate all regional WMSs containing planning information to be bundled into one national WMS for planning information. To the end-user, the WIS will appear as one web service (see Figure 5.1).WMSs are very popular for "free" web services as they only produce a static image in a low-resolution format (*e.g.* jpg, pdf) that allows little to no editing. Often images generated from WMS are embedded into other services such as online route planners or community platforms. However, the images contain an attribution label as part of copyright requirements. If a map is generated from more than one WMS or from a WIS, multiple attribution labels will appear on the image, which may hamper legibility of the image (see Figure 5.2). In the Netherlands WMSs are the most popular web services used by both the public sector and the private sector. From interviews held for this research, it appeared that to date there is little demand yet for WFSs and WCSs. There are a few WFSs available, which are mainly used within the public sector and by specialised private sector companies such as engineering firms. However, the lack of demand for WFSs/WCSs in the Netherlands may be explained by the fact that potential users of these geo web services may be unaware these web services exist.

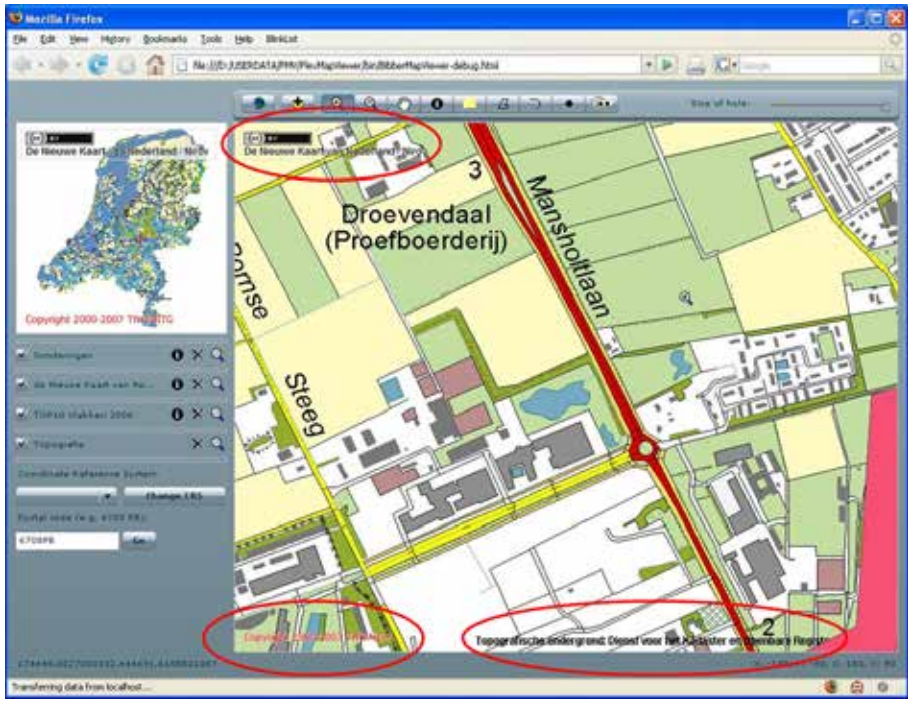


FIGURE 5.2 Several source attributions per map image (source: Bibber, GeoPortal Networks Working Party, https://portal.wur.nl/sites/geoloketten/default.aspx)

TOC

§ 5.2.1.2 ICT standards web services

For geo web services ICT standards such as SOAP, are actually used more often than OGC standards. The most popular type of web service is a Data Service (DS). The private sector uses DSs because custom-made information is delivered to the client. Furthermore, a DS can combine geo-information with data from other databases. Query tools can then be used to perform analyses on the metadata. Licenced information can be protected with firewalls, although the same firewalls can make it harder to set up query tools. Apart from DSs, there are also Sensor Web Services and Simulation Models. Sensor Web Service is a type of sensor network consisting of spatially distributed sensor platforms that wirelessly communicate with each other. They are most often deployed for environmental monitoring and control. For this research, all ICT standard web services will be bundled into Data Services.

Although the technical specifications and standards used for the various types of web services are different, the economic aspects of them are not so dissimilar. In this article, no distinction will be made between the different types of web services when describing the economic aspects.

§ 5.2.2 Costs of web services

The costs of setting up and keeping a web service operational are high. To develop a web service one has to invest in hardware, software, legal, technical, sociological and economic expertise, building up know-how, market and target group research, implementation costs, advertising and promotional costs, administrative and project management costs. Then there are the operational expenses such as servers, broadband capacity, licence fees for software and/or (geo) datasets, acquisition costs and personnel costs. During the operational phase of a web service reservations have to be made for future costs such as R&D, equipment depreciation and extra capacity.

The costs of an operational web service are very variable, depending on the type of service. Stieglitz et al. (2008) made a financial analysis of a virtual community as part of a case study. Virtual communities are a group of people sharing a common interest by using internet applications. Web 2.0 platforms are technologies, which enable formation of virtual communities. An increasing number of private sector organisations are using virtual communities to bridge the gap between users and the organisation by including users in the value chain. The financial analysis undertaken by Stieglitz et al. (2008) was conducted for a virtual community of retail investors at the Berlin Stock Market with memberships sold on a subscription base. Stieglitz et al. (2008) distinguish four separate phases in the life of a web service. These four phases are:

1. the development phase (analysis, design and implementation);
2. the operational phase;
3. the adaptation phase (evaluation and evolution); and
4. the disintegration phase.

Even in the disintegration phase, the web service still incurs costs such as migration costs to another platform, running contract costs and replacement of technology. Only in the operational phase is revenue raised through savings, advertisements and memberships / subscriptions. In their analysis, Stieglitz et al. (2008) noted that the total costs per month were relatively stable during the first year of the operational phase. Only after a critical mass of users and contributions is reached, growth can accelerate. Later in the operational phase, the costs will continue to increase but so will the revenue. With an increasing number of members, the cost per member will decrease until it approaches zero. However, when the number of active members reaches a certain level, the operational costs will step up because of the required extra capacity (servers, broadband, personnel). In addition, this specific virtual community is still in the operational phase. In later phases (adaptation and disintegration), the cost per member will probably increase again.

Although this case study applied to a commercial virtual community, the same principles apply to geo web services. From the various interviews held for this research, the biggest cost item mentioned is sufficient broadband capacity to keep the service operational at all times. Especially for WMS the required server and broadband capacity can be huge if there are many simultaneous users. In addition, it can take some time for an image to build up on the screen of the client-computer. If the build-up time is too slow, the user will abandon the web service. To save building-up time, images can be stored as tiles on the server(s) in advance. However, for large-scale information sets Terabytes of storage capacity is required. Geoportail, the French NGII web service requires 3 Gbps broadband capacity, two 50 Tb caches and a 100 Tb storage capacity (Richard, 2008).

§ 5.2.3 Web service revenue

Web services are set up by the public sector for several reasons: to share information with other public sector organisations, to inform citizens and the private sector (with or without a legal obligation to do so), or as a way to market public sector information (PSI) for reuse. PS(G)I forms a rich resource for value added resellers (VARs) to create value added products and services. Because the public sector enjoys scale of economies and scales of scope, the costs are relatively low. The benefits may be financial for fee-based services or increased taxation revenue from VARs; or the benefits may

be intangible such as a better-informed citizen or increased policy effectiveness. As intangible benefits are harder to measure, cost-benefit analyses tend to be negative. However, end-users of information also incur costs if information needed is scattered all around. These lost productivity costs can be significant when someone has to spend hours searching the Internet for useful information (Bates and Andersen, 2002). The savings made in search costs should be included in cost-benefit analyses when setting up web services for internal use.

§ 5.3 Business models

There are many definitions for the concept of business models. Rappa (2003) offers perhaps one of the simplest definitions, that a business model the method is of doing business by which a company can sustain itself - that is, generate revenue. A business model describes the strategies implemented to achieve a goal. A financial model is an essential part of a business model. The financial model describes the cost framework and how revenue will be generated. The simplest business model is producing and selling a good to customers with revenue higher than all costs incurred. Poorly worked out business models and financial models were one of the main causes of the demise of the dot-com companies at the end of the last century (see *e.g.* Razi *et al.*, 2004).

§ 5.3.1 Components of a business model

After a comparison of different business model definitions, Bouwman et al. (2008) distinguish four components of a successful business model, namely Service, Technology, Organisation and Finance. Together these components form the so-called STOF-model (see Figure 5.3). The four components should be addressed in balance with each other. The starting point is the service domain which addresses aspects such as type of service, intended user group and the value of a service for meeting customer demands. The service domain serves as a guide to the technical design. Some of the aspects addressed in the technical design are architecture, infrastructure, accessibility and payment mechanisms. To develop and market a successful service often requires organisations to collaborate. Collaboration can be as simple as one organisation wanting to launch a web service and needing financial backing from a bank or it can be different organisations bundling information into one web service. The organisation domain describes the value chain required to realise a specific service. A value chain consists of actors with specific resources and capabilities that interact and work

together to create value for customers and to realise their own strategies and goals (Faber *et al.,* 2008). The organisation domain has to address the network and actor aspects as well. The last component to be addressed is the finance domain, which is the bottom line of any business model with revenues on one side and investments, costs and risks on the other side.

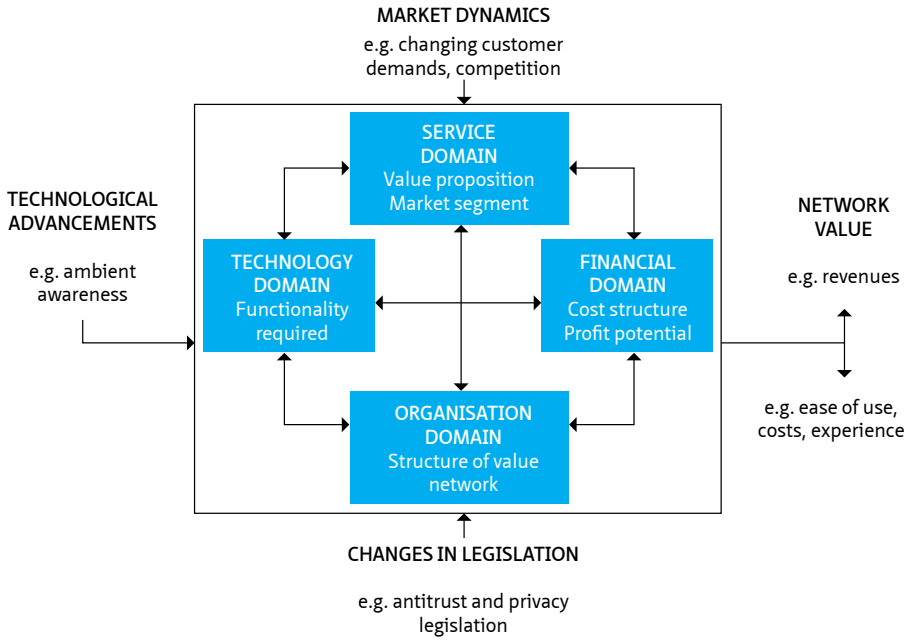


FIGURE 5.3 STOF model (source (Bouwman *et al.,* 2005)

Osterwalder and Pigneur (2002) note that a business model can only be successful if it includes the following three elements: (1) revenue and product aspects; (2) business actor and network aspects; and (3) marketing specific aspects. In their view, a business model should be based on the following columns:

- Product innovation;
- Customer relations;
- Infrastructure management, and
- Finances.

§ 5.3.2 Types of business models

Malone et al. (2006) designed a simple diagram of 16 types of business models based on two dimensions. The first dimension looks at the type of asset right being sold. These are:

1. Creator buys raw materials or components from suppliers and then transforms or assembles them to create a product sold to buyers;
2. Distributor buys a product and resells essentially the same product to someone else;
3. Landlord sells the right to use, but not own, an asset for a specified period of time;
4. Broker facilitates sales by matching potential buyers and sellers. Unlike a typical Distributor, a Broker does not take ownership of the product being sold, rather only receives a fee from the buyer, the seller, or both.

The second dimension takes into account the type of asset for which rights are being sold. These types are physical (durable goods), financial (*e.g.* cash, insurance), intangible (*e.g.* copyrights, knowledge, goodwill), and human (people's time, effort). Combining these dimensions offers the following 16 business models, although effectively there are only 14 as two (human creation and human trade *i.e.* slavery) will be illegal in most countries (Table 5.1 provides an overview of Malone's business models).

	CREATOR	DISTRIBUTOR	LANDLORD	BROKER
Physical	Manufacturer	Wholesaler/retailer	Leaser (*e.g.* real estate)	Auctioneer (*e.g.* eBay)
Financial	Entrepreneur	Bank, investment firm	Lender/insurer	Insurance broker
Intangible	Inventor	Intellectual property trader	Publisher/brand manager/ attractor (*e.g.* Google)	Intellectual property broker
Human	*Human creation*	*Slavery*	Contractor	Human resources broker

TABLE 5.1 Schema of 16 types of business models (after Malone *et al.*, 2006)

Since information is a physical good, only the business models on the top row are applicable to GI suppliers. GI suppliers are often both "Manufacturer" as well as "Leaser" because apart from producing GI, they often only sell the right to use the product rather than transfer ownership. There are some public business organisations trading as "Broker", such as DataLand brokering municipal GI in the Netherlands. However, most of these organisations also trade as "Leaser" and the brokerage is often only a secondary business activity. Hence, in this article they are included in the "Leaser" category. The schema of viable business models can be adapted now for GI suppliers illustrated in Table 5.2.

	CREATOR	DISTRIBUTOR	LANDLORD	BROKER
Physical	Manufacturer	Wholesaler/retailer	Leaser	Auctioneer
Financial	Entrepreneur	Bank, investment firm	Lender/insurer	Insurance broker
Intangible	Inventor	Intellectual property trader	Publisher/brand manager/ attractor	Intellectual property broker
Human	*Human creation*	*Slavery*	Contractor	Human resources broker

TABLE 5.2 Schema of viable business models for GI-suppliers (light blue) (after Malone *et al.*, 2006)

§ 5.4 Financial models

§ 5.4.1 Cost models

Financial models consist of two components: cost models and revenue models. The cost model describes which costs an organisation incurs to run a business. The revenue model describes how an organisation expects to generate income. For public sector organisations supplying PSGI there are two cost model regimes: marginal costs regime and cost recovery regime. With the marginal costs regime only costs of dissemination are taken into account, *e.g.* cost of a DVD or actual time taken to produce a copy. For web services, the marginal costs are zero if the operational costs of the web service are deemed part of supplying a public service. With the cost recovery regime, all costs that are made by the organisation to create, collect, process and maintain the information are included in calculating the dissemination costs. The PSI Directive even allows a reasonable return on investment.

§ 5.4.2 Revenue models

All organisations, including public sector organisations, will have to employ a Revenue Model for PSGI web services. In the literature, many revenue models are described. Rappa (2003) distinguishes nine different categories of revenue models. These categories are listed in Table 5.3.

Not all revenue models described by Rappa are suitable to PSGI web services, such as the Brokerage, Advertisement, Infomediairy and Merchant Model. In addition, the term 'Usage Model' may be a better description of the model than the term 'Utility Model'. Public sector organisations with a Marginal Costs regime will not need to charge for their web services at all. Therefore, some extra models are added to the list, including some revenue models out of the creative sector. As most public sector organisations are holders of (semi-)monopolistic data, they employ the Manufacturer Model by definition; therefore, this model is further omitted.

REVENUE MODEL	DESCRIPTION
Brokerage model	Brokers bring buyers and sellers together and facilitate transactions, usually for a fee or commission.
Advertising model	The web site provider provides content (usually, but not necessarily, for free) and services (such as e-mail or blogs) mixed with advertising messages in the form of banner ads.
Infomediary model	Infomediaries collect information, *e.g.* information about consumers and their consumption habits, or information about producers and their products useful to consumers when considering a purchase. The infomediary then acts as an information intermediary.
Merchant model	Wholesalers and retailers of goods and services. Sales may be made based on list prices or through auction.
Manufacturer (direct) model	The manufacturer or "direct model" allows a manufacturer to reach buyers directly and thereby compress the distribution channel.
Affiliate model	The affiliate model offers financial incentives (in the form of a percentage of revenue) to affiliated partner sites. The affiliates provide purchase-point click-through to the merchant. It is a pay-for-performance model: if an affiliate does not generate sales, it represents no cost to the merchant.
Community model	The viability of the community model is based on user loyalty. Users have a high investment in both time and emotion. Revenue can be based on the sale of ancillary products and services or voluntary contributions; or revenue may be tied to contextual advertising and subscriptions for premium services.
Subscription model	Users are charged a periodic fee to subscribe to a service. It is not uncommon for sites to combine free content with "premium" (*i.e.*, subscriber- or member-only) content. Subscription fees are incurred irrespective of actual usage rates.
Utility model	The utility or "on-demand" model is based on metering usage, or a "pay as you go" approach. Unlike subscriber services, metered services are based on actual usage rates.

TABLE 5.3 Categories of revenue models (after Rappa, 2003)

When the viable business models for PSGI suppliers (see Table 5.2) of Malone et al. (2006) are combined with the adapted revenue models of Rappa, the following revenue models appear:

1 Subscription model: Revenue is raised through periodic fees. This is a popular model for supplying access to a service that is frequently used, *e.g.* iTunes. The advantage for the web service provider is that revenue is raised in advance and thus providing more certainty of regular income. The advantage for the user is that costs of accessing

information are known in advance and access is unlimited within the subscription limit. A disadvantage is that both research and practice show that consumers are reluctant to pay for online services (Schiff, 2003), unless there is a direct relation with their private lives (Reitsma, 2007). Sometimes a basic subscription is offered for free and versions with more features attract a fee (*e.g.* Google Earth for free, Google Earth Plus $20/year and Google Earth Pro $400/year). Subscription models are best suited to specialist information, or (semi-)monopolistic information, *e.g.* large-scale base maps.

2 Usage Model: Revenue is raised through actual usage of a service. Usage may be measured in time, per bytes, per area or per session. The web service provider has to be able to cope with small amounts of money. The usage model is best suited to ad hoc users whereby access to services is more important than possession. In addition, the usage model is only suited to web services with geo-data from only a few suppliers, as the pricing structure will become very complicated and intransparent (MICUS, 2003; 2008b). Another disadvantage for geo web services is that charging per hectare or bytes will render large-scale area coverage very expensive.

3 Royalty model: Revenue is raised through royalties paid after a value added product has been successfully produced. The price of a service is dependent on the results of the user. The price, the royalty, is usually a fixed percentage of the turnover or the revenue of the value added product of the user. The advantage of this model is that a firm only has to pay for the GI after a value added product is successful so there is room for experimenting. The disadvantage of this model is that contracts have to be signed in advance making this model less suitable to click-through licences. Users of the supplied information have to be monitored. In addition, there is no short-term certainty of income.

4 Free Model: There is no direct revenue raised through this model, although there will be indirect benefits. Public sector organisations employ this model, either as a legal obligation or for efficiency reasons (no sales staff). The immediate benefits are intangible, *e.g.* a better-informed citizen or better policy effectiveness, or the benefits may be financial in the long term, *e.g.* extra taxes when value added products are created. However, making GI available free of charge may be in breach with national Fair Trade Legislation in some countries as it may be deemed an act of unfair trading practices if the private sector already has made vast investments to create similar services. The creative sector also uses the Free Model to achieve name recognition or for altruistic reasons.

5 Hybrid models: These are models showing some of the characteristics of the models described above. Below some of the more common varieties are described.

a Enticement model: A part of the content is provided free of charge as a lure to entice the user. Revenue is raised from sale of premium content or other related services. This is one of the oldest revenue models first introduced by King Gilette to create a market for his disposable razor blades (Anderson, 2008). Often cross-subsidising is employed, *i.e.* content is offered for free and revenue is raised from sale of related products, such as merchandising (*e.g.* free mobile phones with revenue from phone calls/text messages; songs downloadable for free and revenue is raised from sale of concert tickets and/or merchandising).

b Community model: The viability of the community model is based on user loyalty. Users invest both time and emotions to produce a communal service. Revenue can be raised by sale of ancillary products and services or by donations; or revenue may be tied to contextual advertising and subscriptions for premium services. The best-known example of a Community is Wikipedia. An example of a GI-community is OpenStreetMap (OSM), a project whereby volunteers go out with GPS units to produce open source street maps for distribution free of charge. OSM operates in many countries on six continents. Some private geo companies have donated cartographic information or money to OSM as well in return for their data or as a platform for innovative applications (http://www.opengeodata.org/?p=223). In Germany, the open geodata of OSM were used for the development of a 3D Geodata Infrastructure as part of the research project 'Geodata Infrastructure 3D' (http://www.gim-international.com/news/id3688-OpenStreetMap_D_Prototype_for_Entire_Germany.html). Virtual communities are frequently used by the private sector to involve users in the developmental and evaluation phases of services as the users provide useful feedback and ensure quality control.

c Street performer protocol: A protocol popular in the creative domain and with software developers. Under this protocol, a producer will release a work (*e.g.* a book or software application) into the public domain after a certain amount of money has been received in a trust fund. Interested parties pay their donations to this trust fund, which is managed by a publisher. When the work is released on time, both the producer and the publisher are paid from the trust fund. If the work is not released on time, the publisher repays the donors. In some variations, the product is commercially released on the market rather than into the public domain. The producer will repay a return on investment to the donors when the product makes a profit. This protocol is very dependent on the reputation of the producer. This protocol would also be suitable to screened-off web services whereby the users are known in advance. Once the participants

have donated their share of development costs and expected operational costs, the service would then be available to the participants to use.

d Combination model: Combinations of the above models are quite often employed, *e.g.* combining the Royalty Model with a start-up fee. The UK Ordnance Survey uses this model for VARs. Another possible combination would the Enticement Model combined with the Subscription Model, *e.g.* giving away a small sample of the Cadastral Map to consumers. The Dutch Large Scale Base Map combines the Subscription Model with the Utility Model as well as offering a user right for the entire dataset for a one-off fee. Another Model is the Data-For-Data model whereby different public sector organisations participate in a joint program, with or without paying an upfront contribution. They donate their data into this program to produce large-scale geo-information. In return, the organisations receive user rights for this large-scale geo-information, Norge Digitalt in Norway uses this model to finance large-scale datasets. The Data-For-Data Model can be combined with the Street Performer Model if a participant donates money instead of data.

§ 5.4.3 Summary of revenue models

Table 5.4 provides a summary of the various revenue models, their advantages and disadvantages and their suitability to various web services.

MODEL	ADVANTAGES	DISADVANTAGES	SUITABLE FOR
Subscription model	· Certainty of regular revenue · Adaptable to users · Lock-in of users · Suitable for click-through licences	· Not popular with consumers · Only suitable for specialised data that is required frequently	· WMS · WFS/WCS · DS
Usage model	· User-pay system, only pay for actual usage · Suitable for ad hoc users · Suitable for click-through licences	· Only suitable when access is more important than possession · Need mechanisms to deal with small payments · Pricing may be prohibitive for large quantities · Pricing mechanism complex when combined with other web services	· WMS · WFS/WCS · DS
Royalty model	· Suitable for experimentation/ innovation platform · Low accessibility · May generate long term indirect revenue for VA products	· Uncertainty of revenue (amount, time) · Must monitor progress of experimenters · No revenue from consumers · Nor suitable for click-through licence	· WMS · WFS/WCS · DS

Free model	· Low accessibility · Indirect revenue (better informed citizen, more effective policy) · May generate long term indirect revenue for VA products · Suitable for click-through licences (if still required)	· No direct or immediate revenue · May be in breach with national Fair Trade Legislation	· WMS · WFS/WCS · DS
Hybrid models			
a) Community model	· User is closely involved (feedback, quality control) · Improvement of service/user friendliness · Encourages experimentation/ innovation platform	· No direct or immediate revenue (unless combined with another model)	· WMS · WFS/WCS · DS
b) Enticement model	· Lures potential users · Lock-in of users	· No direct or immediate revenue (unless combined with another model)	· WMS · DS
c) Street performer model	· Financing service is done upfront · Unlimited use for donors/ participants	· Donors/participants must be known and willing to donate in advance · Dependant on good reputation of producer	· WMS · WFS/WCS · DS

TABLE 5.4 Revenue models with pros, cons, and their suitability to web services

§ 5.4.4 Price strategies

Apart from the Revenue Models described above, price discrimination can be applied as well. The British welfare economist A.C. Pigou described as early as in 1920 a pricing theory, which included price discrimination (Pigou, 1920). Price discrimination can only be applied in a limited fashion by the public sector, as the PSI Directive does not allow that a public sector body distinguishes between different groups of users using the data for similar purposes. It may be possible to offer rural GI cheaper than urban GI because the latter is more dynamic and needs to be updated more frequently (Longhorn and Blakemore, 2008). In addition, there may be more need for urban information, *i.e.* a larger market segment. Another form of price discrimination that may be applied, is offering volume discounts but the volume price is the same for everybody. An example would be to decrease the unit price per hectare when a larger area is selected, *e.g.* as applied to the Automatisierten Liegenschaftkarte (ALK) in North Rhine Westphalia, Germany. Alternatively, a time-based approach could be employed, *e.g.* charging a higher fee for more timely weather information products, or charging a lower fee for usage outside normal business hours.

In the last couple of years there appears to be a trend that large scale PSGI is coming down in price, because either it was too expensive for the private sector or the prices

created barriers to effective reuse within the public sector. With prices being lowered, the number of (re)users is increasing, so the actual total revenue may even go up. Recent examples are found in Austria, Netherlands and Spain. The Austrian Federal Office of Meteorology and Surveying (BEV) have significantly reduced their fees for their PSGI. For instance, the fee for the cartographic model was reduced by 93% and usage went up by 200-1500%, and the digital cadastral map went down by 97% and usage up by 250%. The majority of new users are small to medium enterprises (Schennach, 2008). In the Netherlands, the so-called New Map of the Netherlands (NMN) has been available online with a Creative Commons licence since January 2006 (see www.nieuwekaart.nl). The NMN offers a complete overview of planned spatial developments and functional changes in the Netherlands. Before the NMN became available free of charge, about 20 datasets were sold. Since then, the number of discrete reusers - both from the public and the private sector - downloading the NMN on a regular basis have stabilised to around 200 (Nirov, 2007). The Spanish Cadastre made the complete cadastral map of Spain available on the internet in March 2003. An analysis of the impact of free access to spatial data in Catalonia demonstrated that such initiative is highly profitable to public institutions, by saving a lot of time, simplifying processes and making optimal use of the available information. The impact on private companies is also positive (MICUS, 2008a).

§ 5.5 Summary of business models

Since the development and operational costs of web services are in general high and the distribution costs low, the underlying business model and financial model must be carefully considered. For public sector bodies the costs of web services will be relatively lower due to their economies of scale. Data often is already available as they are often the holder of such data, and personnel often can be drawn from ICT departments. However, some major aspects still have to be addressed.

The web service should be designed with a clear vision. The STOF Model offers a useful framework to address key components. Firstly, the service component must be addressed. Aspects such as intended users (other public sector bodies, private sector), which functionalities the web service should have, should be considered. Once a type of web service (WMS, WFS/WCS, WIS, DS) has been selected, technical adaptations may have to be made to cope with data protection and, if needed, payment facilities. Server and broadband capacity should match the expected number of simultaneous users, bearing in mind that new web services often attract many visitors in the first months before the number settles. Web services such as TIM-online in North Rhine Westphalia (Germany), GeoNorge in Norway and Geoportail in France attract millions of visitors per

year and their number still increase progressively. It is advisable to design a feedback mechanism for users for quality control.

Developing web services often requires collaboration with other departments or organisations. Therefore, attention must be paid to the actors and networks involved. However, networks are dynamic; changes in policy and legislation will cause actors and their roles to change during the period of collaboration. So, it is important to establish formal and informal agreements on the respective roles and responsibilities within the network. If information is used from third parties, *e.g.* aerial photography from the private sector, care has to be taken that licence restrictions are complied with. It is vital that when licence agreements with third parties are drawn up, it is made clear in advance that the information will be made available through web services to avoid legal problems afterwards.

Lastly, the financial aspects have to be considered. These aspects include selecting the most suitable revenue model for the type of information made available and which tariff scale, if applicable, will be employed. If fees are to be charged, it is important to set the fees appropriately, as the fee structure is the most visible part of a web service. If the fees are too high, they will form a bar for potential users and insufficient revenue will be raised to cover the costs. Fees that may appear too low to recover costs in the short term may turn out to attract more users that are new and thus actually increase revenue.

The Subscription Model is best suited to web services that offer frequently used information. The user has a clear indication of ongoing fees in return for unlimited use of data within the subscription limit. The supplier has a clear indication of revenue received upfront. The Usage Model is best suited to ad hoc users whereby access to services is more important than possession. However, the Usage Model is only suitable when data is only available from only one or a few sources as the pricing mechanism can become complicated. The Royalty Model is most suited to VARs who need some time to experiment to develop a viable product or service. For the supplier the short-term revenue is uncertain but the long-term revenue may compensate the initial losses. This model is therefore very suitable to public sector bodies that either have an additional source of funding or already have established a steady flow of income out of earlier royalties. The Free Model is best suited to information supplied by public sector bodies funded out of general revenue. It is an open access model, which should remove the current barriers to reuse of PSGI. However, supplying certain PSGI data may be in breach with Fair Trade Legislation if the private sector has already developed similar datasets. The Hybrid Models, either combining aspects of the above models or borrowing elements of revenue models from the creative domain, offer interesting possibilities. The Community Model involves the end-user and thus, provides essential feedback for a successful web service. The Enticement Model can be

used in combination with fee-based web services to attract new customers. The Street Performer Model can be adapted for establishing GIIs for the public sector.

§ 5.6 Conclusions and recommendations

In the last decade, the way GI is used has shifted from only being used in niche applications to becoming embedded everywhere in society. Technological and societal changes have made unlocking PSGI easier. As GIIs are evolving from first generation to second generation GIIs, more and more PSGI web services are set up. However, as technology has changed to make PSGI available, so should the underlying business models and financial models; especially in light of the upcoming INSPIRE implementation. If the only users of a PSGI web service are other public sector bodies, especially when the web service is part of a NGII, then the only viable revenue model is the Free Model or the Data-for-Data Model as variant of the Street Performer Model. Not only is it counterproductive for public sector organisations to invoice each other every time a web service is used, there is also a real risk that public sector organisations will prefer to use (a combination of) alternative "free" sources such as Google Earth and OpenStreetMap rather than their "own" public sector geographic information. This contradicts with the spirit of the INSPIRE Directive (see Giff *et al.*, 2008).

If PSGI web services are made available outside the public sector to society, then the only viable revenue model for viewing services such as WMS is the Free Model. The Royalty Model could also be used, as this is effectively a "free" model since no value added products will be created by just viewing. The private sector, which may need PSGI for their own business processes or to produce value added products, will be prepared to pay for good quality PSGI provided the fees are not too prohibitive. Therefore, for reusers of WFS, WCS and Data Services the Subscription Model, the Royalty Model or Hybrid Models would be suitable. Although the Usage Model is commonly applied, in the long term it is not be viable even for high-quality Large Scale Base Maps. The fees, even with price discrimination discounts, will become too steep for larger areas and the fee structure will become complicated when combined with other data.

To ensure that PSGI is truly shared through web services as envisaged by INSPIRE, national governments will have to provide sufficient funding to guarantee continuous quality. This means that the current cost recovery regime has to be reconsidered. Recent reports in 2008 such as the Cambridge Report (Cambridge University, 2008) and the MICUS Report on Assessment of Re-use of PSI (MICUS, 2008a) support this point of view. While the Cost Recovery model ensures that a public sector organisation can guarantee that PSGI is maintained at a sufficient level of quality of PSGI (van

Loenen, 2009), the model is no longer suited to using web services for PSGI. This is because the specific PSGI data is no longer just accessible from that public sector body but from multiple web service avenues. In the long term, the benefits of making PSGI available free of charge or for lower fees will pay off in the form of intangible benefits and extra revenue raised in the form of taxes when more value added products will be created.

Acknowledgements This article has been written as part of the activities of the RGI-006 project 'GeoPortal Network – Liberty United' (Geoloketten – Vrijheid in Verbondenheid) and the RGI-117 project 'Geodata: from supply to access' (Geodata: van verstrekking naar toegang) in the framework of the Bsik program 'Space for Geo-information'.

ABBREVIATIONS	
ALK	Automatisierten Liegenschaftkarte (Computerised Property Map)
BEV	Austrian Federal Office of Meteorology and Surveying
DS	Data Service
GI	Geo Information
(N)GII	(National) Geo Information Infrastructure
IPR	Intellectual Property Rights
OGC	Open Geo Consortium
OSM	Open Street Map
PS	Public Sector
PS(G)I	Public Sector (Geo) Information
NMN	New Map of the Netherlands
STOF	Service – Technology – Organisation - Finances
VAR	Value Added Reseller
WCS	Web Coverage Service
WFS	Web Feature Service
WIS	Web Integrator Service
WMS	Web Map Service

References

Anderson, C. (2008). Free! Why $0.00 is the future of business, Wired Magazine: 16.03 Tech Biz: IT (February 25, 2008, at http://www.wired.com/techbiz/it/magazine/16-03/ff_free [accessed October 10 2008].

Bates, M.E., and D. Andersen (2002). Free, fee-based and value-added information services, The Faciva 2002 White Paper Series (Dow Jones Reuters Business Interactive LLC), at http://www.factiva.com/index.asp?node=menuElem0 [accessed November 21 2008].

Bouwman, H., Faber, E., and T. Haaker (2005). Designing business models for mobile services: balancing customer and network value, Proceedings 2nd IEEE International Workshop on Mobile Commerce and Services (WMCS'05), July 19 2005, München, Germany.

Bouwman, H., Haaker, T. and H. de Vos (2008). Mobile service innovation and business models, Heidelberg: Springer Verlag.

Cambridge University [Newbery, D., Bently, L. and R. Pollock] (2008). Models of Public Sector Information Provision via Trading Funds. Study commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury in July 2007, Cambridge: Cambridge University.

Crompvoets, J., Bregt, A., Rajabifard, A. and I. Williamson (2004). Assessing the worldwide status of national spatial data clearinghouses, International Journal of Geographical Information Sciences, 18(7), pp. 665-689.

Faber, E., de Vos, H.,Haaker, T. and H. Bouwman (2008), About STOF, in Faber, E. and H. de Vos (Eds.) Creating successful ICT-services. Practical guidelines based on the STOF method, Enscheda: Telematica Institute, pp. 17-31.

Giff, G.A., van Loenen, B. and J.A. Zevenbergen (2008). PSGI policies in Norway and England: are they within the spirit of recent EU directives? International Journal of Spatial Data Infrastructures Research, 2008(3), pp. 118-145.

Longhorn, R.A. and M. Blakemore (2008). Geographic information: value, pricing, production, and consumption, Boca Raton: CRC Press, Taylor & Francis Group.

Malone, T.W., Weill, P., Lai, R. K., D'Urso, V.T., Herman, G., Apel, T.G. and S. Woerner (2006). Do Some Business Models Perform Better than Others?, MIT Sloan working paper 4615-06, (May 2006), at http://ssrn.com/paper=920667 [accessed March 12 2008].

MICUS Management Consulting GmbH (2003). The market for geospatial information: potentials for employment, innovation and value added - Short version, Study commissioned by Federal Ministry of Economy and Labour (Bundesministerium fuer Wirtschaft und Arbeit), Dusseldorf: MICUS, www.micus.de.

MICUS Management Consulting GmbH (2008a). Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report, Study commissioned by EC in 2007, Dusseldorf: MICUS, at: http://ec.europa.eu/information_society/policy/psi/docs/pdfs/micus_report_december2008.pdf.

MICUS Management Consulting GmbH, 2008b, The value to industry of PSI: the business sector perspective, Proceedings The socioeconomic effects of Public Sector Information on digital networks: Toward a better understanding of different access and reuse policies, OECD Workshop, February 4 2008, Paris, France, OECD, at: http://www.oecd.org/dataoecd/27/9/40047551.pdf.

Nirov (2007). Analysis GIS Download 2006-April 2007, Memo from Jan Kadijk, Project Leader De Nieuwe Kaart van Nederland).

Osterwalder, A. and Y. Pigneur (2002). An eBusiness Model Ontology for Modeling eBusiness, Proceedings 15th Bled Electronic Commerce Conference eReality: Constructing the eEconomy, June 17-19 2002, Bled, Slovenia, pp. 75-91, Bled: University of Maribor.

Pigou, A.C. (1920). The economics of welfare, London: Macmillan.

Rajabifard, A., Feeney, M.-E.F. and I.P. Williamson (2003). National SDI initiatives, London: Taylor & Francis.

Rappa, M. (2003). Business models on the Web, Managing the digital enterprise. Open educational resource, North Carolina State University at http://digitalenterprise.org/about.html [accessed May 13 2008].

Razi, M.A., Tarn, J.M. and F.A. Siddiqui (2004). Exploring the failure and success of DotComs, Journal of Information Management & Computer Security, 12(3): pp. 228-244.

Reitsma, D. (2007). De consument en zijn behoefte aan geografische data.
[The consumer and his need for geographic information], Proceedings Werkconferentie RGI-117 Geodata, van verstrekking naar toegang, February 13 2007, Den Haag, pp. 35-47. Rotterdam: Bedrijvenplatform Geo-Informatie.

Richard, D. (2008). Géoportail: The French Geoportal. Access online to public datasets by means of geographic interfaces, INSPIRE Drafting Team Network Services Meeting October 15 2008, Amersfoort, Netherlands. Geonovum.

Schennach, G. (2008). Reducing charges for national geographic reference data in Austria, ePSIplus Conference: PSI Re-use - Who takes action next?, June 13, 2008, Brussels, Belgium, at: http://www.epsiplus.net/events/epsiplus_conference_psi_re_use_who_takes_action_next/schennach.

Schiff, F. (2003). Business models of news web sites: a survey of empirical trends and expert opinions, First Monday [Online] 8(6) June 2 2003 at: http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1061/981 [accessed May 13 2008].

Stieglitz, S., Latteman, C., vom Brocke, J. and C. Sonnenberg (2008). Economics of virtual communities - a financial analysis of a case study at the Berlin Stock Exchange, 21st Bled eConference eCollaboration: Overcoming boundaries through multi-channel interaction, Bled, Slovenia, pp. 248-262, Bled: University of Maribor.

van Loenen, B. (2009). Developing geographic information infrastructures: the role of access policies, International Journal of Geographical Information Science, 23(2): pp. 195-212.

van Loenen, B. (2006). Developing geographic information infrastructures: The role of information policies, PhD Dissertation, Delft: OTB Research Institute for Housing, Urban and Mobility Studies, Delft University of Technology.

Interviews

Jellema, M., DataLand, Reeuwijk, Netherlands, November 20, 2007.
Kooijman, J., TNO Research Centre for Built Environment and Geo Sciences, Utrecht, Netherlands, October 21, 2008.
Murre, L., former director National Cooperation Large Scale Base Map Netherlands (LSV GBKN), Delft, Netherlands, May 21, 2008.
Rijkaart van Capellen, W., director www.buienradar.nl (private sector WMS), Capelle aan den IJssel, Netherlands, April 29, 2008.
Rijnders, J., senior consultant KPN Telecommunications, Utrecht, November 27, 2007.
Schreuder, P., former adjunct director Kadata (former commercial arm of the Dutch Cadastre), Utrecht,Netherlands, May 8, 2008.
Wormgoor, E. & J. Braaf, Hydrography Service Royal Navy (DHKM), Den Haag, Netherlands, June 25, 2008.

6 Sustainable business models for public sector open data providers

Frederika Welle Donker and Bastiaan van Loenen

JeDEM Journal of eDemocracy & Open Government 2016, vol. 8(1), pp.28-61, http://www.jedem.org/index.php/jedem/article/view/390. Submitted 18 September 2015; revised 22 March 2016; accepted 23 May 2016; published 1 July 2016.

Abstract

Since 2009, Open Government Data initiatives have been launched worldwide and the concept of open data is gaining momentum. Open data are often associated with realizing ambitions, such as a more transparent and efficient government, solving societal problems and increased economic value. However, between proposing an open data policy and successful implementation are some practicable obstacles, especially for government agencies required to generate sufficient revenue to cover their operating costs, so-called self-funding agencies. With lost revenue due to open data, there is a real risk that the update frequency and the quality of data may suffer or that the open data policy may even have to be reversed. This article has researched the financial effects of open data policies for self-funding agencies on their business model. The article provides some hands-on proposals for self-funding agencies having to implement an open data policy whilst ensuring their long-term sustainability.

Keywords: open data; implementation of open data policy; business models; sustainability; self-funding agencies

§ 6.1 Introduction

§ 6.1.1 Open data of self-funding agencies

Open government data are often associated with realizing ambitions such as a more transparent and efficient government (*e.g.* Huijboom and van den Broek 2011), improved citizens' participation (Jetzek, 2013), solving societal problems (*e.g.* Uhlir, 2009) and increasing economic value by companies creating innovative products and services using open data as a resource (*e.g.* Omidyar Network, 2014). Governments worldwide are adopting open data policies for their data, for instance, the G8 Open Data Charter of 2013. However, between proposing an open data policy and successful implementation are some practicable obstacles, especially for government agencies required to generate sufficient revenue to cover a substantial part of their operating costs – so-called self-funding agencies, *e.g.* the UK Trading Funds. Large-scale geographical data are expensive to collect and maintain. In many countries, self-funding agencies are often monopolist providers of such high-value data and rely on income generated by licence fees for supplying data as part of their business model (Onsrud, 1992).

Self-funding agencies are increasingly under political and societal pressure to release their high-value data as open data. However, without a sustainable open data business model for self-funding agencies there is a real risk that the update frequency and the quality of data may suffer or that the open data policy may have to be reversed (van Loenen, 2009). To offset revenue losses due to open data, self-funding agencies either have to be compensated by the central government or have to increase revenues raised by other activities, which may give rise to competition concerns (DotEcon, 2015). Another concern is that self-funding agencies may only publish open datasets with the least commercial value, while retaining the more valuable data to minimise the risk to commercial revenues (Rhind, 2014). There may also be a fear that 'open data could be immediately "swallowed up" [...] by big global companies" and not benefit the national economy' (Michael Fallon, UK Minister for Business and Enterprise, cited by PASC, 2014).

Much of the past research focussed on the potential benefits of open data without taking into account the financial effects on the data providers, even though financial aspects are a core part of an organisation's business model. One of the challenges in this tension field between lost revenue due to open data and still maintain adequate data service quality, is to develop a sustainable business model for open government data providers which ensures the availability of quality open data in the long term.

This article has researched different business models that may be suitable for self-funding government agencies considering adopting open data policies. We have researched alternative and innovative business models to find out if these business models can be adopted or adapted. Our research questions are which business models are currently employed by government open data providers and how can such a business model be made sustainable to ensure the availability of high quality open data in the long term? As part of our research, we have studied the effects of open data on the business model of three Dutch government agencies, the Kadaster, the National Transport Agency (RDW) and the Dutch Chamber of Commerce, and one research institute TNO. These cases were selected because the first three organisations are self-funding agencies allowed by law to charge fees for data, and the fourth case concerns a non-governmental organisation with a public mandate. All four organisations are key register data administrators and although these organisations rely on generating income from fee-based services to cover a substantial part of their operating costs, they are under pressure to supply key register data as open data. In addition, we have researched two international government agencies that can be considered 'best practice' open data providers, the Australian Bureau of Statistics and the United Kingdom Companies House. The Australian Bureau of Statistics was selected because it was one of the first organisations to switch to open data in 2005. Thus, this case demonstrates the longer-term effects of open data. The Companies House was selected because their public tasks and datasets are comparable to the Dutch Chamber of Commerce and have recently switched to providing open data. These case studies may offer inspiration for other government agencies having to make the switch from fee-based data and services to open data.

§ 6.1.2 Dutch legal framework

The legal framework for accessibility of government data in the Netherlands follows the European Union's framework laid down in the so-called Amended Public Sector Information (PSI) Re-use Directive (2013/37/EU) the so-called INSPIRE Directive (2007/2/EC) establishing an Infrastructure for Spatial Information in the European Community, and the system of key registers.

§ 6.1.2.1 Public Sector Information Re-use Act

The transposed Amended PSI Re-use Directive came into force on 18 July 2015 and modified the regime of the Public Information Access Act. The aim of the PSI Re-use Act is to make all suitable public government data available for re-use with

as few legal restrictions as possible, through open and machine-readable formats and together with their metadata. In principle, charges should not exceed marginal dissemination costs. The PSI Re-use Act lists only three self-funding public sector bodies (Kadaster, National Transport Agency (RDW) and Chamber of Commerce) that are allowed to charge higher fees, provided these fees do not exceed the total costs of production and dissemination. The PSI Re-use Act confirms the general open government data policy already in place in the Netherlands since the adoption of the EU Digital Agenda in 2011 that all government data should be "open, unless", *i.e.* available as open data unless there are limitations, *e.g.* privacy-sensitive data, commercially-sensitive data and data related to state security.

§ 6.1.2.2 Geographical information infrastructure implementation Act

The transposed INSPIRE Directive came into force on 1 September 2009. The INSPIRE Implementation Act provides a framework of general rules, implementing rules and measures aimed at the establishment of a geographical information infrastructure to facilitate policy-making and measures that may have an impact on the environment. The framework addresses accessibility and exchange of geographical data across various levels of public authority and across different sectors through harmonisation and standardisation. The INSPIRE Implementation Act requires that INSPIRE datasets are described through metadata, have interoperable formats and are accessible through network services with minimum performance criteria for those services. These criteria relate to performance, capacity and availability of network services. INSPIRE requires that public access to discovery and viewing services are without costs but download services may be subject to licences and charges. To meet the INSPIRE requirements; the Dutch government has developed PDOK, a portal with web services for discovering, viewing and downloading government geographical data. Many PDOK web services are based on open data and thus, available to everyone.

§ 6.1.2.3 Key register system

To improve the information infrastructure in the Netherlands, the government has developed a system of key registers of core reference data. The ultimate goal of this system that core reference data are collected once and re-used many times. To achieve this, key registers are regulated by law: one government organisation is responsible as key register data administrator and provider of a key register, other government bodies are obliged to re-use key register data to prevent multiple collection of similar data, and government users are obliged to report any errors to the data manager. Together, key registers form part of a System of Key Registers, whereby authentic data from one

register is re-used in another register, *e.g.* the addresses of the Key Register Addresses are re-used in the Key Register Cadastre. Of the 12 allocated key registers, six have a geographical component and six are administrative registers. The ultimate aim is that key registers will be available for re-use within the government without financial barriers. Key registers that are not under financial constraints and do not contain personal data, are made available as open data.

Although the Key Register System predates INSPIRE, together they form a framework for a national information infrastructure.

§ 6.1.3 Reading guide

In the next section, we describe business model theory and business model components. After an overview of the spectrum of current data policies, we elaborate further on two components of business models, the service component and the financial component and specifically potential revenue models. Section 6.3 introduces information value chains and describes the different activities that are part of an information value chain. We explain which roles a public sector data provider may adopt in this information value chain and illustrate the relation between business model components and roles within information value chains. Section 6.4 describes and analyses the six case studies. In addition, we address financial implications of open data for self-funding agencies. In Section 6.5, we use our analyses to provide conclusions and recommendations.

§ 6.2 Business models

A literature review has shown that there are many definitions for the concept of business models. For instance, Ovans (2015) proposes that a business model is more about strategies and assumptions than about money. Osterwalder and Pigneur (2010, p.14) state that "A business model describes the rationale of how an organization creates, delivers, and captures value". Rappa (2010) adds to this definition the goal of a company: namely to generate revenue. For this article, we adhere to the definition that a business model describes the strategies an organisation can employ to reach a certain goal. For a company producing physical goods the goal may be simple: sell products in order to generate a profit. For government agencies whose goal it is to provide a specific

public service, the business model becomes more complex as there may be specific legal constrains and obligations.

§ 6.2.1 Components of a business model

De Reuver et al. (2008) distinguish four common components of a business model, a service component, a technical component, an organisational component and a financial component (*cf.* Johnson, Christensen and Kagermann, 2008; Osterwalder and Pigneur, 2010). These components are not isolated but are mutually interconnected.

- The service component describes the type of service on offer, the target audience and market segment (government, businesses, consumers), and the added value of the service to the end-user. The service component determines the design of the other components.
- The technical component describes the architecture, infrastructure, data security, and other technical aspects needed to develop a service, and the way in which this service is delivered to the end-user.
- The organisational component describes the value network required to realise a certain service. A value network consists of actors with certain resources and capabilities, which interact and together perform value activities, to create value for customers and to realize their own strategies and goals (Bouwman, Haaker and de Vos, 2008, p.56). Legislation and policies provide the framework within an organisation can operate.
- The financial component describes the financial resources required to develop and deliver a service (cost model) and in which way revenue is generated (revenue model). In the financial component attention must be paid to incoming revenue on one side and all production costs, transaction costs, investments and risk mitigation on the other side. It should be noted that the costs for online services are characterised by a high ratio of fixed to variable costs (Shapiro and Varian, 1999).

We focus on the service component as it forms the starting point of any business model, and on the financial component as this component determines the sustainability of all other components, *i.e.* the finances determine the level of service, the technical and organisational aspects.

§ 6.2.2 Service component

The service model describes the type of service on offer (what), the intended target group and/or market segment (to whom) and the added value to the user (why). The service component has a direct relation with the position or role of an organisation in the (information) value chain (in which capacity). For government open data organisations the services on offer would typically be web services for viewing and/or downloading open data, possibly in combination with fee-based services. The target audience for open data is society at large, *i.e.* other government bodies, companies and citizens. As each user has its own preference for file formats, attributes, et cetera, it will be, therefore, impossible to set up services tailor-made to all types of users. However, a public sector organisation can elect to set up services for a specific market segment, such as other public sector bodies, and to outsource data supply to other segments, such as the private sector and citizens.

§ 6.2.3 Financial models and data policies

In the previous section, we have described which activities (what) an organisation can undertake and which role the organisation can adopt (in which capacity). In this section, we will describe the financial component of a business model. The financial component breaks down into two models, the cost model (what will it cost) and the revenue model (how will costs be recovered). After we describe the spectrum of data policies, we will explore existing revenue models.

Roughly speaking, there are two data policies for government agencies providing data to users: a cost recovery policy and an open data policy. With a cost recovery model, all costs incurred in production, maintenance and dissemination of data are recovered from the users of the data by charging licence fees (see *e.g.* de Vries *et al.*, 2011; Fornefeld *et al.*, 2008; Longhorn and Blakemore, 2008; van Loenen, 2006; Onsrud, 2004). Data are made available under a licence agreement, usually with restrictive licence conditions. The pricing models used to set the licence fees, may be (one-off) fees per area or per kBs, subscription fees, fixed access fees, royalties or a combination of these models (Welle Donker, 2009). The revenue raised by licence fees flows directly back to the data provider. Thus, the data provider has certainty of covering its (partial) operational costs. Figure 6.1 shows the data flows and financial flows of the cost recovery policy.

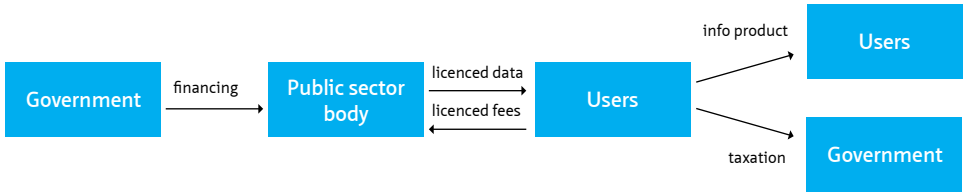


FIGURE 6.1 Cost recovery policy for data (P. Boers (personal communication April 17, 2015) after van Loenen, 2006, p.106)

On the other end of the spectrum is an open data policy. Data are provided free of charge and without licence restrictions. There is no direct revenue for the data provider; instead, the data provider depends on the Treasury to finance its operational costs. There is an expectation that when open data are used for value added products and services, revenue in the form of company taxes and value added taxes will flow back to the Treasury (*e.g.* Vickery, 2011; PIRA *et al.*, 2000; Koski, 2011; de Vries *et al.*, 2011). However, the data provider will largely depend on political decisions to maintain adequate funding levels (Onsrud, 1992). If funding for a data provider is reduced, the update frequency and quality of the datasets may have to be reduced. In addition, there are no guarantees that revenue raised from taxation will be returned to the appropriate public sector body (Longhorn and Blakemore, 2008). Figure 6.2 shows the data flows and financial flows of the open data policy.

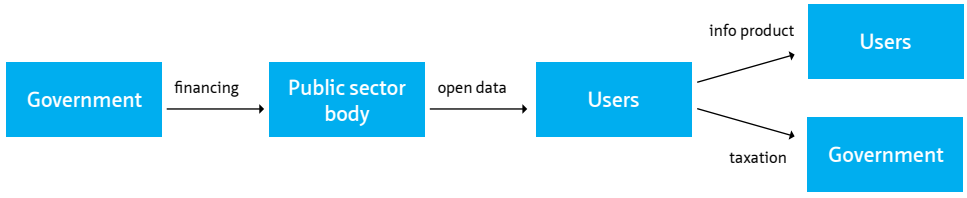


FIGURE 6.2 Open data policy (P. Boers (personal communication April 17, 2015) after van Loenen, 2006, p.106)

There are many data policies between these two ends of the spectrum. In the following section, we will explain these in more detail.

TOC

§ 6.2.4 Revenue models for public sector organisations

From a literature review, we identified eight possible revenue models for public sector bodies. Of these eight, the first five are single revenue models, the latter three are combinations of revenue models.

§ 6.2.4.1 Open data model: budget financing

Open data are supplied free of charge. The costs are financed out of general revenue/annual budget of the organisation. There are no direct benefits to the data provider. However, the indirect benefits may be efficiency gains (Houghton, 2011), improved decision-making (Janssen, Charalabidis and Zuiderwijk, 2012), strengthen law enforcement (Huijboom and van den Broek, 2011), a better informed citizen, lower levels of citizen's perception of corruption (Granickas, 2013), or more innovative applications based on the open data and thereby higher societal benefits (*e.g.* Newbery, Bently and Pollock, 2008; Huijboom and van den Broek, 2011).

§ 6.2.4.2 Legal instruments

Governments have legal instruments available to generate revenue. Next to general taxation, *e.g.* income taxes, (local) governments may levy taxes for a specific purpose, *e.g.* road taxes, council rates, compulsory company registration charges and/or adjust existing tariffs to balance the budget. Another legal instrument is to make authenticated data products mandatory for certain transactions, *e.g.* a certified copy of a cadastral registration is compulsory when buying and selling a property, and only the Land Registry Office can issue such certified copies.

§ 6.2.4.3 Subscription model

A product can be used after an upfront fee for a specific period is paid. Subscription models are best suited to services that are used frequently (Rappa, 2010). There may be subscription differentiation between different types of users. The advantage of this model is that the data provider has the assurance of a predictable and constant revenue flow known in advance. The advantage for the user is that fees are known in advance; irrespective of how many times the service is used.

Research has shown that users are in general only willing to pay for online services if this has a direct relation with their private life (Reitsma, 2007), the service fits in with their lifestyle (MRI, 2012) or if users have a positive perception of convenience, essentiality, added value and service quality (Wang *et al.,* 2005). Users are willing to pay for subscription services that are deemed essential, such as e-banking, but are less willing to pay for information services, such as online news sites, if the same information may be available free of charge via alternative channels. Therefore, a subscription-based service needs to differentiate itself by offering higher added value than alternative free services.

§ 6.2.4.4 Utility model

The utility model is based on the principle of a user-pay system. Users pay each time they use a service, often per unit. Units may be time-based, volume-based or per session. This model has to be able to handle micro-payments per unit. This model is best suited to services for ad-hoc users when access to content is more important than possession of content (Welle Donker, 2009). In addition, this model is only suitable to platforms where the pricing mechanisms are relatively simple, *e.g.* per km^2, as it will become complicated to calculate the fees when combining multiple datasets with different pricing mechanisms and, thus, less transparent for the user (Fornefeld *et al.,* 2008).

§ 6.2.4.5 Royalty model

In the royalty model when a product is used as a resource to develop and market value added products, the usage-based fee depends on the success of the final product. The fee, the royalty, is typically a percentage of gross or net revenue generated by the re-user (van Loenen, Zevenbergen and de Jong, 2006). The advantage of this model for the re-user is that the final fee is only payable after a value-added product is successfully marketed for a profit. The disadvantage of this model is that contracts must be exchanged in advance. In addition, if a company markets a value-added product without a profit objective, *e.g.* a free-of-charge app, royalties may not be payable to the data provider.

§ 6.2.4.6 Enticement models

These are models whereby some part is given away to lure potential customers, and the rest of the product/service is available for a fee.

Razor & Blades model

With the "razor and blades" model, a product is supplied for free to entice the user to pay a fee for associated services. The razor is provided free of charge, the blades are provided at a cost. The free product has little value in itself and only becomes valuable with the associated services. This model may be used for creating vendor lock-in, by *e.g.* supplying free data in a proprietary software format (Rappa, 2010). In an open data setting Ferro and Osella (2013) call this 'Infrastructural Razor & Blades': a product/service is provided for free via APIs ("razor") while re-users are charged only for computing power that they employ on-demand in as-a-service mode ("blades"). Application of this model is limited to contexts and domains in which the computational costs are significant (Ferro and Osella, 2013).

Open Source Like model

In the Open Source Like model a product/service is provided on top of a basic service that is provided in an open format, *e.g.* Linux software supplied by Red Hat. The costs of the basic service are covered by revenues generated by supplementary products/services that are based on open data (Ferro and Osella, 2013).

Unlike the Razor & Blades model, the part that is given away – the data – is usable in its own right and does not require the fee-based service. This model distinguishes itself from the infrastructural razor & blades model by not being limited to a specific platform/infrastructure.

Freemium/premium

The freemium/premium model, also known as versioning, uses a pricing strategy whereby a product/service – usually with a high intrinsic value – is provided in different versions (Ferro and Osella, 2013). A sample or a basic version of a dataset is provided free of charge (freemium), whereas a version with a higher level of quality is provided for a fee (premium). The difference between the versions may the available size, number of available attributes, update frequency or available level of service (*cf.* Tennison, 2012; de la Iglesia and Gayo, 2009). For instance, a highly dynamic dataset may be available for a fee guaranteeing access to (near) real-time data 24/7, whereas the freemium version the data may be 24 hours old without guaranteed 24/7 access.

§ 6.2.4.7 Community model

The Community model depends on the loyalty of the users (Rappa, 2010). Users invest their time and soul in developing and maintaining a specific service or platform, *e.g.* OpenStreetMap. The Community model may also be used to facilitate user feedback to improve a data service or to improve the quality of published data (Tennison, 2012) as a valuable addition to other forms of collecting data (DotEcon, 2015). Revenue is raised through voluntary contributions, sales of related products and/or contextual advertisements. Companies, such as Garmin and Nike, use community platforms to obtain an insight into the usage of their products, to receive feedback and to use the sensor data (mobility patterns) generated by the platform users. This business model can be very useful for organisations that have an obligation to publish information but lack the resources to do it well (Tennison, 2012).

Street Performer protocol

In this model, a producer will release a product into the public domain after a certain level of donations has been received by a trust fund. If the producer releases the product within an allocated time, the producer is paid by the trust fund. If the product is not published within the allocated time, or the product is commercially released, the donators are repaid by the trust fund (Kelsey and Schneider, 1999). In an open data context, it may be conceivable that this model is deployed for co-financing the added effort to release data as open data, *e.g.* adapting data to anonymise personal information or corporate sensitive information.

Crowd Funding model

Crowd funding is characterised by small donations from many individuals, mostly acquaintances (friends, family, colleagues) and, increasingly, from unknown parties. The investor may receive a small reward in return (acknowledgement on website, shares in start-up company), interest payment or the investor may donate for altruistic reasons (Zhang, Collins and Baeck, 2014). With crowd funding money is raised bypassing traditional lending organisations and often a direct relation between the donators and the funded project is created (Louzada, 2013). Although crowd-funding has been used by smaller public organisations for *e.g.* research projects in the United States, it is not suitable to large government agencies.

§ 6.2.4.8 Advertising models

Data and/or services are provided (usually, but not necessarily, free of charge) mixed with advertising messages. Revenue is generated through banner ads, intromercials and/or storing, combining and selling user registration data (Rappa, 2010).

Rappa (2010) describes more advertising revenue models, such as the contextual/targeted advertising and the affiliate model (providing purchase opportunities to internet surfers). Ferro and Osella (2013) describe a "Free as Branded Advertising" model (aimed at persuading an audience towards a brand or company) and a "White Label Development" model (outsourcing required expertise to specialised firms). Although advertising models may not be suitable to all public sector organisations depending on national policies and/or cultural attitudes, other countries have formulated policies for advertising on government websites.

§ 6.2.5 Summary revenue models

Table 6.1 provides a summary of the various revenue models. All models are suitable for public sector organisations, except that the advertising model may be limited by a national legal/cultural framework. Legal instruments are only suitable to private sector organisations if they are mandated with executing a public task.

REVENUE MODEL	SHORT DESCRIPTION/OPTIONS	SUITABLE TO PUBLIC SECTOR
Budget financing	No direct revenue raised; indirect benefits	Yes
Legal instruments	Specific levies/taxation	Yes
	Mandatory usage of certified data products	Yes
Subscription model	Fees in advance for a specific period independent of actual usage	Yes
Utility model	Pay-per-use/view	Yes
	Premium	Yes
	Work to order	Yes
Royalty model	Revenue once a derived product is profitable	Yes
Enticement model	(Infrastructural) razor & blades	Yes
	Open Source Like	Yes
	Freemium/premium	Yes
Community model	Street performer	Yes
	Crowd funding	Not always
Advertising model	Web/banner advertising	Not always

	Affiliation model	Not always
	Free as Branded Advertisement	Not always
	White Label Development	Not always

TABLE 6.1 Revenue models and their suitability for public sector organisations

Organisations can employ multiple revenue models to generate income, for instance a utility model for infrequent users in combination with a subscription model for frequent users. Infrastructural razor & blades (providing an open data platform with fee-based cloud-computing facilities) can be facilitated in combination with the open source like model (expert services to add value to open data).

§ 6.3 The information value chain

Although there is ample literature related to business models of organisations providing (open data) services, it appears that there is an tendency to use the term business model to actually describe the different activities an organisation may undertake to provide data services (*e.g.* Rappa, 2010; Janssen and Zuiderwijk, 2014), to describe the different roles an organisation may occupy in the information value chain (*e.g.* Malone *et al.*, 2006; Deloitte LLP, 2012), or to describe pricing mechanisms for information products (*e.g.* Rappa, 2010; Ferro and Osella, 2013). The first two aspects are part of the Service Component of a business model; the latter is part of the Financial Component. In the previous chapter, we have described the financial component of a business model. In this chapter, we will describe the service component, and specifically the roles an organisation may occupy in the information value chain.

§ 6.3.1 Roles in the public sector information value chain

Public sector organisations collect and process vast quantities of raw data to information as part of their public tasks. The data are combined and transformed in multiple steps to produce the information needed to carry out these public tasks. In the information value chain, most of the costs occur in the initial stages whereas most of the value is created in the latter stages (*cf.* Krek and Frank, 2000). We consider four activities in the information value chain: (1) collecting data, (2) aggregating and storing data, (3) processing data, including quality control and transforming into user-friendly

formats, and (4) using the information product (*cf.* van Loenen and Zevenbergen, 2010). Each step in the chain builds on the previous step and in each step, value is added to the information product. Value can be added by *e.g.* adding new attributes to the data or by combining with other data, *e.g.* data from companies, sensor data and/or crowd-sourced data. In addition, value may be added by developing tools or by supplying specialised knowledge to assist the users. The information product can be an end-product but may also be the first step of the next value chain. Figure 6.3 shows the information value chain for public sector data.

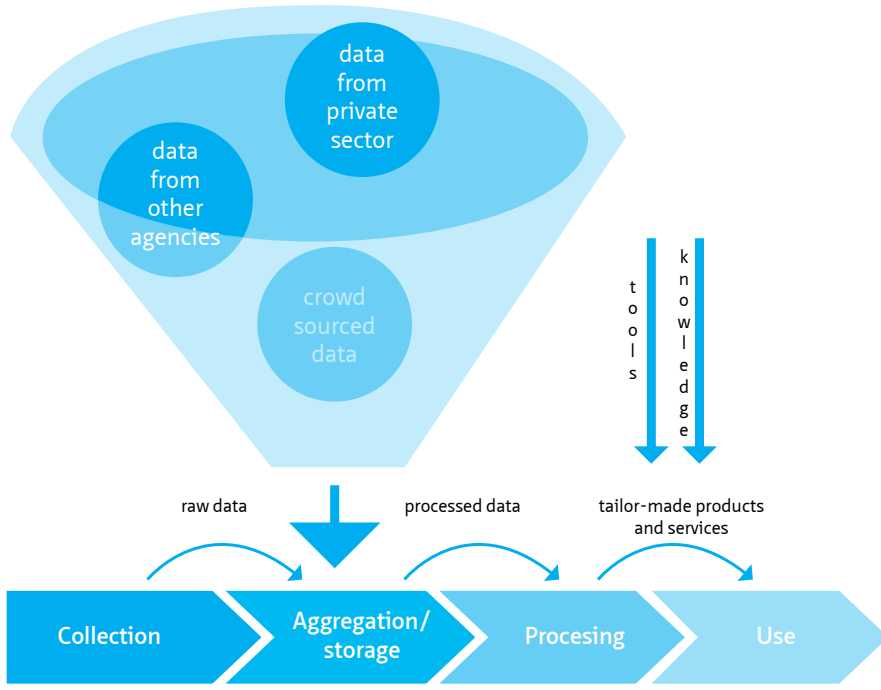


FIGURE 6.3 Basic information value chain with steps for adding value

In this section, we describe the different roles in the information value chain, using the five "archetype" roles for organisations within the public sector information value chain distilled by Deloitte LLP (2012) as a basis for the Service Component. These archetype roles are:

1 Suppliers of (open) data: may be public sector organisations but also companies. Data are supplied for re-use by third parties. There are no direct monetary benefits from supplying the data; instead, broader benefits may be found in greater transparency, enhanced reputation and/or societal benefits.

Users of open data may further be categorised as:

2 Aggregators: organisations that collect and aggregate open data, sometimes combined with proprietary data. Such aggregation often occurs on sectorial or geographical level. The aggregated data may be used to present the data more efficiently, or to perform analyses.

3 Enablers: organisations that provide a platform, tools and technology for third parties to use open data. The enablers do not use the data as such but act as an intermediary between data holders and users by providing cost-effective and efficient solutions, and/or by coordinating feedback. Winning platforms are those that are more convenient, more user-friendly, better organized, and more visually appealing (Schiff, 2003). In addition to providing an open data platform, Enablers can offer additional services, such as consultancy. The Enabler's open data platform can be a demand-oriented platform where users pay a charge for user-friendly and reliable access to data or a supply-oriented platform where the data holders pay a fee to use the Enabler's resources and expertise instead of developing their own open data platform.

4 Developers: organisations and individuals that design, develop and sell applications for end-users. Such applications, such as multi-modal route planners, typically use highly dynamic open data. Developers may sell their applications directly to end-users or build custom-made applications for other organisations.

5 Enrichers: organisations (typically larger companies) that use open data to enhance their existing portfolio through better insight, efficiency gains or as a tool to sell other products.

End-users are at the end of the information value chain and could be the start of a new value chain. In Figure 6.4, we have combined the different roles identified by Deloitte LLP (2012) and activities as part of the public sector information (PSI) value chain.

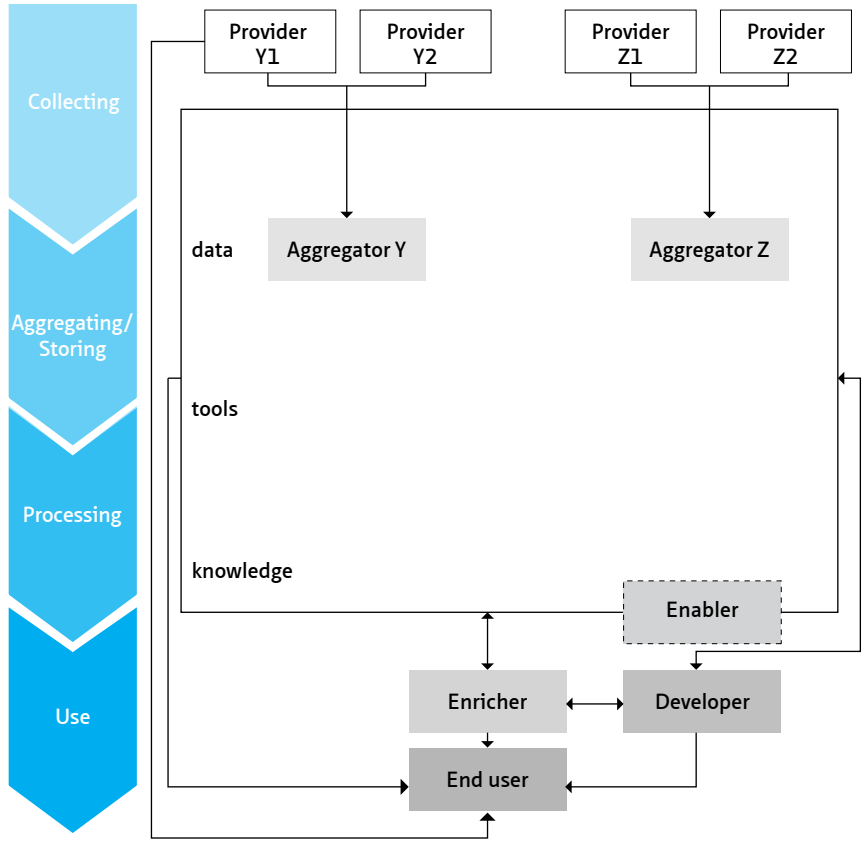


FIGURE 6.4 The roles and activities identified by Deloitte LLP (2012) and their relation to the PSI value chain

§ 6.3.2 Service component linked to financial component

In the previous chapter, we have described which revenue models are appropriate for public sector organisations. In this section, we link the five roles for public sector bodies in data processing to the financial component.

1 **Supplier:** activities are financed by budget financing, legal instrument charges or in combination with additional fee-based services, *e.g.* to use freemium/premium model to generate revenue from high value information.

2 **Aggregator:** can adopt different revenue models, *e.g.* access via APIs free of charge and fee-based services for combining data from other sources. Aggregators can also be financed by budget financing and/or legal instrument charges.

3 **Developer:** generate revenue by working to order or from sales of applications, using the subscription model, the utility model and/or advertising models. It is conceivable that a developer may use crowd funding.

4 **Enabler:** generate revenue directly from the use of the platform and/or supplementary information services. Combination of almost all revenue models can be used by the Enabler, depending on the status of the organisation (public, private) and the types of supplementary services. In addition, operational expenses may be reduced due to efficiency gains.

5 **Enricher:** often use open data to improve existing products/service. Revenue is generated by efficiency gains and/or added value for existing customers by facilitating effective use of existing products/services. The revenue models that are already in place will probably not be adapted.

Table 6.2 shows the relation between the different roles and revenue models. A plus indicates that the revenue model is suitable to both the public sector and the private sector. A hyphen indicates that the revenue model is only suitable to public sector organisations (and private organisations mandated with carrying out a public task). A cross indicates that the revenue model is suitable to the private sector but not always to the public sector, depending on the legal and cultural framework.

ROLE/MODEL	BUDGET FINANCING	LEGAL INSTRUMENTS	SUBSCRIPTION MODEL	UTILITY MODEL	ROYALTY MODEL	ENTICEMENT MODEL	COMMUNITY MODEL	ADVERTISING MODEL
Supplier	+	-	+	+	+	+	+	x
Aggregator	+	-	+	+	+	+	+	x
Developer	+	-	+	+	+	+	+	x
Enabler	+	-	+	+	+	+	+	x
Enricher	+	-	+	+	+	+	+	x

TABLE 6.2 Relation between roles and revenue models, and suitability to public sector and private sector

+ = suitable to public sector and private sector
- = only suitable to public sector but not always to private sector
x = suitable to private sector but not always to public sector

§ 6.4 Case studies

§ 6.4.1 Methodology

Having to implement an open data policy, may create a financial burden for government agencies that rely on generating revenue to cover a substantial part of their operating costs - so-called self-funding agencies. There may be a risk that revenue losses due to open data may endanger the agency's ability to maintain operations at a sustainable level and with an adequate level of quality. To test this hypothesis, we have researched the effects of open data on the business model of a number of self-funding agencies in the Netherlands and the United Kingdom, and a budget-financed agency in Australia.

To examine which open data business models would be suitable for (long-term) sustainability for self-funding government organisations, we analysed the current business model of four Dutch organisations by desk research and by face-to-face interviews. The cases were selected because these organisations are by law allowed to charge fees for re-use of their data. All four organisations are key register data administrators and are, therefore, monopolists for these authentic datasets. Although these self-funding organisations rely on income from fee-based services to cover a substantial part of their operating costs, they are also under pressure to supply key register data as open data. For comparison, we analysed the business model of two open data government agencies outside the Netherlands. These cases were selected from a shortlist of best practice cases after a literature review. The research for the international cases was carried out by desk research and interviews by phone. All interviews were structured, using open questions sent to the interviewees before the interview. Additional desk research was carried out by analysing annual reports and information available from the websites of the interviewed organisations. Reports of the interviews were sent to the interviewees for validation.

§ 6.4.2 Dutch National Transport Agency (RDW)

The Dutch National Transport Agency (RDW) is a self-funding government agency of the Ministry of Infrastructure and the Environment and is responsible for registering the life cycle of motorised vehicles, including a number plate register. In addition, the RDW is responsible for registering driver's licences and annual vehicle tests. The RDW is the managing authority for the Key Register Vehicles.

§ 6.4.2.1 RDW service component

Since 2013, the RDW supplies part of its registers as open data. These data are:

1. General vehicle data, such as make and model, fuel consumption and registration history. Reports of missing and/or stolen vehicles are not available as open data. The update frequency is daily. A request may contain 1 to 100 number plate registrations.
2. Parking data: static data pertaining to parking areas and corresponding parking fees, and near real-time data pertaining to availability of municipal and commercial parking areas.

The RDW also supplies similar data as a fee-based web service (OVI-business) with a higher service level: available as near real-time data, no data limit, and additional historical information available. OVI-business guarantees 24/7 availability through a Service Level Agreement.

§ 6.4.2.2 RDW financial component

The RDW receives no budget financing from the national government or compensation for open key register data. Operational costs are covered by revenue received from vehicle registration charges and annual vehicle tests. Open data supply is financed by internal budget allocations.

The RDW receives revenue from the private sector for their fee-based services. The total revenue in 2014 was €193.4M (RDW 2015, p.87) with revenue from information dissemination accounting for just over 2% of the total revenue. Compared to 2012, revenue from information products, especially from the vehicle traders, dropped in 2014 from €4.86 to €4.11M (RDW, 2015, p.87). According to the RDW interviews, the loss of revenue is offset by internal efficiency gains and fewer data requests.

§ 6.4.2.3 Summary RDW business model

The RDW employs legal instruments (registration fees and annual vehicle test charges) to cover their operational costs. The revenue model is a combination of freemium/premium: (downgraded) open data with a lower service level and fee-based data available as a subscription-based added-value service. The RDW has not developed additional value added services to compensate for revenue losses due to open data. Open data have not had a significant impact on the revenue of fee-based services: after

an initial drop, RDW actually received more applications for fee-based data services as more users became aware of the data potentials. Since the introduction of open data, the prime role of the RDW in the information chain has moved from Aggregator towards Enabler by offering a limited number of tools for open data. The RDW did not develop these tools themselves; instead, they opted to use an existing platform (MS Azure) with a proven track record.

§ 6.4.3 Dutch Cadastre, Land Registry and Mapping Agency (Kadaster)

The Dutch Cadastre, Land Registry and Mapping Agency (Kadaster) is a self-funding non-departmental public body operating under the responsibility of the Ministry of Infrastructure and the Environment. One of its statutory tasks is to manage four key registers (Key Register Topography, Key Register Cadastre, Key Register Real Estate Values and Key Register Large Scale Topography), as well as to manage datasets on behalf of the Ministry. In addition, the Kadaster manages the national geographic information platform PDOK. This case study is limited to the Kadaster's open data: the Key Register Topography.

§ 6.4.3.1 Kadaster service component

The Key Register Topography (BRT) consists of digital topographic data in various formats and scales, including 1:10,000 and has been available as open data since 1 January 2012. The BRT is available via web viewing services and download services from the PDOK platform. It is mandatory for other government bodies to use BRT data for their public tasks unless the data do not meet their specific needs. The introduction of BRT open data has led to a rapid increase in re-use and an increased demand for better actuality. The Kadaster has started a project to provide a user-friendly feedback system open to parties outside the public sector to improve the data quality (Grus and te Winkel, 2015).

§ 6.4.3.2 Kadaster financial component

The Kadaster nett turnover was €257.6M in 2014, of which 5% came from topography (non-standard products) (Kadaster, 2015, p.73).

Before open data, revenue from topographic data accounted for 8% (€21M) of the total turnover (Kadaster, 2012, p.96). To offset the loss in revenue, the national government pay compensation to the Kadaster, which covers about 50% of the BRT operating costs. The other 50% are expected to be covered by efficiency gains, such as staff reduction and increased data quality (de Vries, 2014). The annual compensation has decreased since 2013 due to government budget cuts (Kadaster, 2015). In addition to other forms of revenue, Kadaster manages PDOK, set up as a Supply-Oriented platform where other public organisations pay per dataset to host data.

§ 6.4.3.3 Summary of Kadaster business model

Before open topographic data, the Kadaster's prime role in the information chain was Aggregator using a combination of legal instruments, subscription model and utility model. For non-open data, the Kadaster still uses the same revenue models. The Kadaster receives some budget financing to compensate re-use within the public sector. The Kadaster has started to experiment with the community model (wisdom of the crowd) to provide feedback on the data quality. The Kadaster's secondary role in the information chain is Enabler by developing tools and offering specialist knowledge for its own consultancy and support services, and as developer and manager of PDOK. Although its revenue models have changed due to open topographic data, the Kadaster's role in the information chain was already shifting from Aggregator to Enabler, due to its role in managing PDOK.

§ 6.4.4 TNO Research Institute

TNO, the Netherlands Organisation for Applied Scientific Research is a self-funding independent research institute mandated with carrying out a number of public tasks. One of its public tasks as the Netherlands Geological Survey is to maintain the Key Register Soil and Subsurface Data.

§ 6.4.4.1 TNO service components

Under the Mining Act, it is compulsory for public and private organisations to supply TNO with subsurface data collected by these organisations. TNO has developed subsurface data models for its data platform as a way of presenting the data in a more user-friendly way. The data and data models are available as open data. Drilling and

probing data are used for large-scale building projects whereas re-use for small-scale building projects and for non-archaeological purposes has been almost zero (Twynstra Gudde, CE Delft and Tauw, 2011). TNO employs a cost recovery regime to provide specialised research and services based on geological data.

§ 6.4.4.2 TNO financial component

There are few figures available related to data dissemination. The data management costs for TNO are estimated at €4.8M for the period 2011-2025 + €0.3M per annum for non-personnel related expenditure (Twynstra Gudde *et al.*, 2011), which constitutes less than 1% of the total operational costs of €533.6M in 2014 (TNO, 2015, p.42).

§ 6.4.4.3 Summary TNO business models

TNO's prime role in the information chain is Aggregator using legal instruments (Mining Act). There is no need to generate revenue from subsurface data as the operational data costs are low compared to TNO's total operational costs. TNO's secondary role in the information chain is Enabler by developing tools and data models to present the subsurface data in a more user-friendly way. TNO employs an Open Source Like model by offering its knowledge to provide specialised geological services and research.

§ 6.4.5 Dutch Chamber of Commerce

The Dutch Chamber of Commerce (Kamer van Koophandel (KvK)) is a self-funding agency of the Ministry of Economic Affairs. Its core tasks are to register, inform and advise companies and to maintain the Key Trade Register. The Key Trade Register contains data related to companies, *e.g.* business profile, history and deposited documents. Incorporation is mandatory for all companies and legal entities. Companies require authenticated copies of registration for a number of business transactions, *e.g.* for bank loan applications. It is mandatory for other government bodies to re-use data from the Key Trade Register, *e.g.* to register a company in the Key Register Cadastre, the Kadaster must refer to the Key Trade Register.

§ 6.4.5.1 Kamer van Koophandel service component

The KvK provides a free viewing service for accessing the public part of Key Trade Register data only, *i.e.* name, street address and website address of a company per single entry. The KvK deems other Key Trade Register data to be non-public, *i.e.* to be privacy-sensitive data and commercially-sensitive data. The KvK provides a fee-based service for companies to obtain an authenticated copy of incorporation, and provides various services for additional (non-public) information, such as lodged annual reports, company address data, mutation subscriptions, and data integration services. Bulk products are available for one-off fees and/or subscription fees. The KvK offers advice and information services for companies. The KvK is in the process of improving their fee-based online services by developing apps, APIs and value added services combining Trade Register data with data from other sources.

§ 6.4.5.2 Kamer van Koophandel financial component

In 2014, the total income of the KvK was €218.5M with a net loss of €63.1M, mainly due to high reorganisation costs and decreased revenue from information products (Kamer van Koophandel, 2015, p.52). In 2012, the national government abolished annual company registration fees to reduce the administrative burden of companies. In addition, government bodies no longer pay each time they re-use Key Trade Register information. Instead, the national government pay an annual compensation to the KvK, which amounts to about sixty percent of the KvK's revenue. This annual contribution will be reduced from €134.9M in 2014 to €114.8M in 2015 (Kamer van Koophandel, 2015, p.51). KvK activities generated net €35.8M in 2014, of which 8.7% (€3.1M) was generated by selling information products.

Should (part of) the Key Trade Register become available as open data (depending on various levels of data protection), it is expected that usage of the services will increase but revenue will drop between €7.9M to €38.9M, depending on which data will become available as open data. It is expected that the estimated net loss in revenue will be limited to circa €0.75M per annum due to increased internal efficiencies (Verdonck Klooster & Associates, Ecorys and Berenschot, 2014, p.26). Open data could help to increase the level of data quality, as more people will be able to report data errors.

§ 6.4.5.3 Summary of Kamer van Koophandel business model

The KvK's prime role in the information chain is Aggregator using a combination of legal instruments (incorporation fees), budget financing and data services using

subscription model and utility model to generate revenue. In addition, the KvK offers specialist knowledge for consultancy and support.

The KvK recognises that its current business model is under threat (Kamer van Koophandel, 2015, p.55). It has to cope with revenue reductions due to the abolition of annual registration fees and budget cuts by the central government. After the Trade Register became part of the key register system, companies are no longer required to obtain a paper-based authenticated proof of registration from the KvK for many other government transactions, such as registering a company car, thus, further eroding the revenue base of the KvK. As the KvK is also under pressure to release at least some of its data for re-use as open data, the KvK must find additional sources of revenue.

§ 6.4.6 Australian Bureau of Statistics (ABS)

The Australian Bureau of Statistics (ABS) is Australia's official statistical organisation. In 2005, the ABS made its statistical data available as open data.

§ 6.4.6.1 ABS service component

The ABS provides Australian national statistics, key economic indicators, census data, consumer price index, labour force, national accounts, regional statistics, and measures of Australia's progress. The ABS provides the following statistical products/services:

- ABS information in the basic set, available on the ABS website free of charge;
- Self-service tools for generating statistical tables from a variety of data sources;
- Self-service statistical literacy resources via the ABS website free of charge;
- National Information and Referral Service (NIRS) providing free information over the phone for straightforward enquiries.

The ABS supplies printed publications, information consultancy services, selected Census products and Confidentialised Unit Record Files (CURFs) for a fee.

The ABS has developed value added services based on its statistical data, as there is a demand for such products and services, especially from the education sector. The ABS would prefer the private sector to take up the development of value added services; however, the private sector appears reluctant to fill the gap.

Since the introduction of open data, there has been a significant increase in web service traffic and in the types of users. Open data require a higher level of data quality and metadata quality as errors become visible for everybody. There has been a decrease in the number of e-mail inquiries and sales of hard copy publications.

Consultancy services have grown slightly every year instead of dropping after the introduction of open data, especially for complex work and specialised services. By offering free data, people have become more aware of the benefits and opportunities of statistics but do not always have the skills to analyse the statistical data.

§ 6.4.6.2 ABS financial component

Data customised to meet specific user needs are available for a fee. The ABS must recover costs associated with preparing customised requests for individual clients/ organisations in accordance with government cost-recovery guidelines. Costs are also recovered for training, consultancies, providing paper copies of ABS publications, information on CD-ROM and access to data via specialised products such as Census DataPacks, TableBuilder and CURFS. In all cases where costs are recovered, the ABS will charge "efficient costs" *i.e.* the minimum costs necessary to deliver products and services that are fit for purpose.

The ABS is wholly funded by the national government. In 2005, revenue from data was relatively low to the annual costs (6.9%). In 2013/14, sales of product and services accounted for 10.9% of the total budget (Australian Bureau of Statistics, 2014, p.133).

§ 6.4.6.3 Summary ABS business model

Prior to 2005, the ABS was primarily an Aggregator employing a combination of budget financing and premium products as revenue models. After 2005, the revenue models of the ABS have shifted to a combination of budget financing and Open Source Like. The Open Source Like revenue model was not a deliberate choice but rather a consequence of open data. By providing open data, consultancy activities have increased. The role of the ABS has shifted to the Enabler role as the ABS provides tools to facilitate use of their data services in a more effective way. Open data is used as a way of applying the ABS' knowledge to provide specialised statistical services.

§ 6.4.7 UK Companies House

The UK Companies House (CH-UK) is an Executive Agency of the Department for Business, Innovation and Skills. CH-UK has Trading Fund status, *i.e.* it is a self-funding agency directly managing its own finances and having to pay an annual dividend to the national government. The core tasks of CH-UK are to incorporate and dissolve limited companies, examine and store company information delivered under the Companies Act and related legislation; and make this information available to the public. CH-UK also carries out delegated functions for the Secretary of State, for example, enforcing the delivery of annual accounts and annual returns; directing companies to change their names that are similar to those currently on the public record, et cetera.

§ 6.4.7.1 UK-CH service component

UK-CH currently offers a number of data services. Two of these data services are free of charge; the other data services are available for a subscription fee and/or pay-per-use fee. These services of UK-CH are:

- WebCHeck service providing a searchable Company Names and Address Index (free of charge). An app providing access to basic company details, register statistics and warnings when certain reports are due to UK-CH (free of charge).
- WebFiling service allowing companies to submit information and reports electronically, and download Incorporation Certificates.
- WebCHeck to view a company's filing history and purchase copies of document images and a selection of company reports.
- Companies House Direct (CHD), a subscription service allowing access to all UK-CH records, including analogue records.

CH-UK also produces DVD and bulk data products of their directory on demand.

As part of improving internal efficiencies, the UK-CH is developing new services and/or improving existing services to replace the WebFiling, WebCheck and CH Direct services. Since 25 June 2015, CH-UK operates a beta service providing open data access and download facility to digital records on companies and directors per single entry.

CH-UK does not intend to produce value added products and services using its data as a resource, as this should be done by the private sector or anybody in the market.

§ 6.4.7.2 CH-UK financial component

As a trading fund, CH-UK is required to generate a surplus at the end of the financial year and pay a dividend to the Ministry, amounting to £4.4M in 2015. The main sources of CH-UK's income are incorporation fees and annual contributions. In addition, CH-UK receives revenue from dissemination activities. CH-UK uses a combination of a subscription model and utility model for dissemination services. The switch to open data services is financed out of its own budget.

CH-UK employs a full cost recovery regime for its operations. The total income out of fees and charges was £67.5M in 2014/15. Revenue from dissemination activities was £10.4M. Dissemination activities account for 15% of operating income, registration activities for 83% and other services account for 2% (Companies House, 2015, p.45). At this stage, it is too early to assess the full impact of open data.

§ 6.4.7.3 Summary CH-UK business model

CH-UK's prime role is Aggregator of data provided by companies and generates revenue with legal instruments (compulsory incorporation fees and annual contributions) and premium data products.

In the last two years, the CH-UK is moving towards an Enabler role by developing tools to facilitate use of (open) data in a more effective way as part of an internal efficiency drive. CH-UK does not use its raw data to provide advisory services or value added products.

§ 6.4.8 Summary case studies

The ABS is wholly funded by the government and made the switch to open data a decade ago. The self-funding agencies use a combination of legal instruments and fee-based data services to cover their operational costs. The Kadaster and the Dutch Chamber of Commerce receive additional budget financing to partially compensate free use of key register data within the public sector. Most researched agencies use hybrid revenue models combining open data with fee-based subscription and utility services. In addition, most of the agencies have developed tools to facilitate users and/or are offering their expert knowledge as (fee-based) value added services. The Kadaster has an additional revenue base as the manager of PDOK, the supplier-driven platform for geographic data.

To date, the effects of open data have been a decrease in revenue from information products and an increase in data traffic. However, as the revenue constituted a relative small part of the total revenue, the losses appear to be offset by internal efficiency gains and higher data quality. The datasets that are not (yet) available as open data often represent a larger part of the total turnover. Table 6.3 provides a summary of the financial component of the researched business models and the reported effects of open data.

	TOTAL REVENUE IN 2014 (€ MILLIONS)	DATA REVENUE BEFORE OPEN DATA (% OF TOTAL)	DATA REVENUE AFTER OPEN DATA (% OF TOTAL)	REVENUE MODEL	EFFECTS OF OPEN DATA
RDW	193.4	2.6%	2.1%	· Legal instruments · Freemium/ Premium services	· Increase in requests · Increase of requests for additional services · Increase in fee-based premium services · Increase in feedback/ reporting errors · Decrease in internal transaction costs · New applications developed by third parties · Operational costs reduced by using existing infrastructure
Kadaster	257.6	8%	5%	· Legal instruments · Budget financing · Supply-driven platform (PDOK)	· Increase in use, especially by citizens · Decrease of requests via Kadaster directly · Increase in requests via PDOK · Increase in feedback/ error reports · Increased demand for better actuality of data · Decrease in transaction costs internally and externally · New applications developed by third parties · Operational costs reduced by using existing infrastructure · Decrease in revenue · Increased dependency on budget funding

TNO	533.6	<1%	Not available	· Legal instruments · Open Source Like	· Increase in use for construction works · Increase in use for agriculture (groundwater data) · Development of tools and models · Operational costs reduced by using existing infrastructure
Dutch Chamber of Commerce	218.5	3.7%	Not applicable	· Budget financing · Legal instruments · Subscription model · Utility model	***Free viewing service:*** · Increase in company searches for basic data · Decrease in revenue · Increased dependency on budget funding
Australian Bureau of Statistics	353.1	6.9%	10.9%	· Budget financing · Open Source Like	· Increase in requests/downloads, etc. · Increase of requests for consultancy services · Increase in fee-based premium services · Increase in feedback/reporting errors · Decrease in internal transaction costs · New tools developed by ABS
Companies House - UK	96.8*	20%	15.4%	· Legal instruments · Subscription model · Utility model	· Increase in company searches and downloads · Decrease in requests via other channels · Decrease in revenue · Decrease in internal and external transaction costs

TABLE 6.3 Summary financial components of case studies
* 1 GBP = € 1.43 (31-07-2015).

Figure 6.5 shows the shift in roles for the researched agencies. The Chamber of Commerce does not provide open data. Its role is very slowly shifting down from Aggregator. The other agencies have shifted further from Aggregator towards Enabler since providing open data. The Kadaster has made the biggest shift as manager of PDOK.

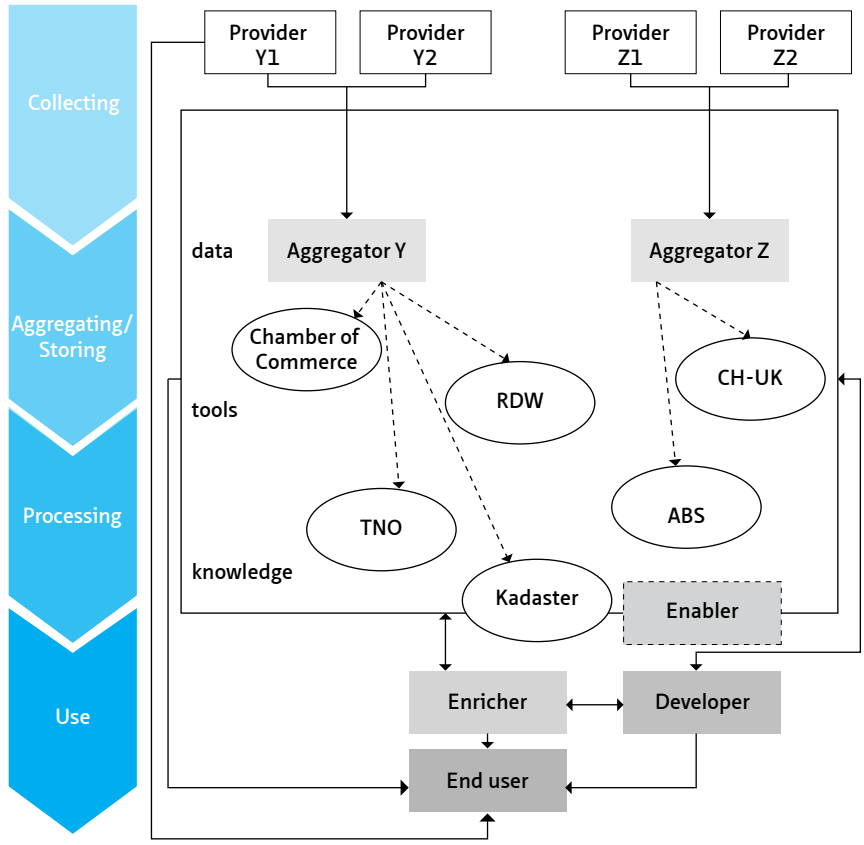


FIGURE 6.5 Shift in roles in the PSI chain for the researched agencies after introduction of open data

§ 6.5 Analysis and conclusions

§ 6.5.1 Case study findings

This article has researched business models that may be suitable for self-funding government agencies that are considering or are in the process of implementing open data policies. We have researched alternative and innovative business models

to find out if these business models can be adopted or adapted. We have analysed the business model of a number of government agencies that have already made the transition to open data, and one agency that is under pressure to provide (some) open data in the near future. Our case studies demonstrate that business models have adapted to ensure the long-term availability of high quality open data, thus addressing long-term financial sustainability of the agencies.

The case studies have shown that providing (raw) open data will not necessarily lead to losses in revenue in the long term. Where organisations have implemented open data in addition to fee-based services, there have been no negative effects on the fee-based services. It should be stressed though, that in all open data cases, revenue generated by fee-based services are relatively low when compared to the main source of income (budget-financed and/or legal instrument model) and the revenue is independent from the main source of revenue. In a number of cases, revenue from fee-based services has even increased. From interviews and from business analytics it emerged that open data has led to internal efficiency gains. In practice, it is difficult to quantify internal efficiency gains solely due to open data in isolation as the researched organisations continuously implement measures to increase efficiency. However, the reported decreases in internal and external transaction costs due to open data are in line with other research (*cf.* de Vries *et al.*, 2011; Houghton, 2011; Koski, 2015). An open data case study of a private sector energy network administrator in the Netherlands resulted in similar findings (Welle Donker, van Loenen and Bregt, 2016).

Our research suggests that open data results in new roles in the information value chain. Before open data, organisations were primarily Aggregators of (raw) data; now they are moving towards the Enabler role by developing tools and/or platforms to facilitate users. Often, these tools/platforms are developed to provide data in a more user-friendly way. In some cases, expert knowledge is offered as a fee-based value added service.

Where open data are not available, the business models are more uncertain as demonstrated by the Dutch Chamber of Commerce. Its revenue from information products and budget financing has decreased whilst operating costs have increased due to a reorganisation. Its business model becomes even more uncertain as on the one hand, a Bill is drafted that will allow stricter data control to protect the Chamber of Commerce's revenue base in the future, whilst on the other hand, there is strong pressure from society and from members of parliament to release Key Register data as open data. The UK Companies House demonstrates that open data may actually be more beneficial in the longer run; however, its revenue base from legal instruments is higher and more stable than that of their Dutch counterpart.

§ 6.5.2 Theory versus reality

Based on a literature study, we have identified a number of potential revenue models. In practice, the case studies have shown that government organisations use the following revenue models:

- Budget financing (Kadaster, KvK, ABS);
- Legal instruments/compulsory contributions (RDW, Kadaster, TNO, KvK, CH-UK);
- Freemium/premium (RDW);
- Open Source Like (TNO, ABS).

The researched cases have demonstrated that it is vital for a sustainable open data business model that there must be a guaranteed main source of revenue, whether it be budget financing or access to legal instruments for levying charges. The researched cases have also demonstrated that offering open data without additional tools are not sufficient to ensure effective use of the data. Thus, all researched organisations have moved from Aggregators towards Enablers in varying degrees. Of the researched organisations, only RDW re-used an existing platform and tools for open data supply. All other organisations have developed their own open data platform and tools.

Government agencies that are in the process of implementing open data should analyse their current role in the information value chain, and which role the organisation can would like/is allowed to move towards. The role in the information value chain determines which level of services (data services, tools, knowledge) are to be provided and whether to develop these services and tools themselves or to use external developers. Once the role in the value chain has been determined, the most optimal revenue model(s) that are appropriate can be adopted, allowing flexibility to cope with a dynamic economic and political landscape. Such choices will depend on the legal framework and the available resources. However, there is no one-size-fits-all sustainable open data business model.

We have found sustainable business models for open data, but in each case, these were unique. There are still a number of issues to be considered. Firstly, open data do not necessarily imply additional open/free services. Secondly, there are many different users and different user needs. Freemium/premium services can address the variety in user needs, such as 24/7 access to near real-time data. Our case studies indicate that some users actually prefer paying for a guaranteed service level. Thirdly, to implement open data in a user-friendly way requires investments. Such investments may provide a return on investment, such as efficiency gains or more feedback on data quality. However, in practice it proves difficult to quantify the direct benefits of open data.

§ 6.5.3 Further research

Open data is still in its infancy and the actual benefits of open data still have to materialise. However, measuring the benefits of open data is no mean feat. Our research indicated that initially such benefits may be found in internal efficiencies for organisations. However, more case studies are needed to quantify the (longer-term) effects on self-funding agencies and to verify our model.

Our research indicates that for sustainable business models, it is a key to be able to move in the information value chain. However, more research is needed to confirm our initial findings. This research has demonstrated that each agency employs its own strategies and develops its own platforms/tools to publish open data. To ensure the long-term viability of open data platforms, a governance framework should be developed which addresses not only aspects, such as long-term finances and commitments to manage these platforms, but also technical aspects, such as data interoperability, to ensure efficient re-use by third parties requiring data from multiple platforms. More research is needed to formulate good governance of open data especially to guarantee the long-term availability.

Acknowledgements: We would like to thank Martine van Heijnsbergen and Hans Lambeck of RDW; Dick Eertink of Kadaster; Rob van der Krogt and Tirza van Daalen of TNO; Freddy Keij of Kamer van Koophandel; Cassy Short, Kate Chaloner, Merry Branson, Stuart Kevin and Tim Montgomery of the Australian Bureau of Statistics; and Jo Jones and Stephen Lewis of Companies House, UK for their kind cooperation.

References

Australian Bureau of Statistics (2014). Annual Report 2013-14. Canberra. Retrieved September 23, 2014, from http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/0BCB69DE782667DACA257D-7200114E5D/$File/abs_annual_report_2013_14_web.pdf.

Bouwman, H., T. Haaker and H. de Vos (2008). Mobile service innovation and business models. Heidelberg: Springer Verlag.

Companies House (2015). Companies House Annual Report and Accounts 2014/15, Retrieved July 27, 2015, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/446159/Annual_Report_201415.pdf.

de la Iglesia, J.L.M. and J.E.L. Gayo (2009). Doing business by selling free services. In M.D. Lytras, E. Damiani and P. Ordóñez de Pablos (eds), Web 2.0. The business model, 89-102. New York: Springer.

de Reuver, M., H. Bouwman and T. Haaker (2008). Capturing value from mobile business models: Design issues that matter. 21st Bled eConference eCollaboration: Overcoming boundaries through multi-channel-interaction, Bled, Slovenia. Retrieved October 7, 2010, from http://domino.fov.uni-mb.si/proceedings.nsf/Proceedings/9D94EC6E7C9E47D2C1257481003D8ECF/$File/09Reuver.pdf.

de Vries, M. (2014). Wah kos'dah dan? Onderzoek naar de incrementele kosten van aan Open Data doen [What's it going to cost? Research on the incremental cost of doing Open Data]. The Green Land. Retrieved March 4, 2014, from http://open-overheid.nl/open-overheid/open-data-wah-kos-dah-dan/.

de Vries, M., L. Kapff, M. Negreiro Achiaga, P. Wauters, D. Osimo, P. Foley, K. Szkuta, J. O'Connor and D. Whitehouse (2011). Pricing of Public Sector Information Study (POPSIS) - Summary Report. Brussels: Deloitte Consulting CVBA,, Retrieved February 29, 2016, from https://ec.europa.eu/digital-agenda/en/news/pricing-public-sector-information-study-popsis-summary-report.

Deloitte LLP (2012). Open Growth: Stimulating demand for open data in the UK. A briefing note from Deloitte Analytics. London: Deloitte Touche Tohmatsu Limited. Retrieved November 27, 2013, from http://www2.deloitte.com/uk/en/pages/deloitte-analytics/articles/stimulating-demand-for-open-data-in-the-uk.html.

DotEcon (2015). Independent evaluation of the OFT's 2006 market study into the Commercial Use of Public Information (CUPI). A Report by DotEcon for the CMA. Retrieved March 26, 2015, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/418097/Evaluation_of_CUPI_study.pdf.

Ferro, E. and M. Osella (2013). Eight business model archetypes for PSI re-use. Open Data on the Web Workshop Shoreditch, London, 23rd-24th April 2013. Retrieved November 23, 2013, from http://www.w3.org/2013/04/odw/odw13_submission_27.pdf.

Fornefeld, M., G. Boele-Keimer, S. Recher and M. Fanning (2008). Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report. Study commissioned by EC in 2007. Dusseldorf: MICUS, Retrieved December 2, 2008, from: http://ec.europa.eu/newsroom/document.cfm?doc_id=1258.

Granickas, K (2013). Understanding the impact of releasing and re-using open government data. European Public Sector Information Platform Topic Report No. 2013/08. Retrieved November 28, 2014, from http://epsiplatform.eu/sites/default/files/2013-08-Open_Data_Impact.pdf.

Grus, M. and D. te Winkel (2015). Assessment of the volunteered geographic information feedback system for the Dutch topographical key register. ISPRS Geospatial Week 2015, La Grande Motte, 28 Sep – 03 Oct 2015, doi:10.5194/isprsarchives-XL-3-W3-51-2015.

Houghton, J. (2011). Costs and benefits of data provision. Report to the Australian National Data Service. Melbourne: Centre for Strategic Economic Studies, Victoria University, Retrieved September 30, 2012, from http://ands.org.au/resource/houghton-cost-benefit-study.pdf.

Huijboom, N. and T. van den Broek (2011). Open Data: an international comparison of strategies. European Journal of ePractice no. 12 (March/April 2011):4-16. Retrieved April 14, 2015, from http://share-psi.eu/papers/TNO-epractice.pdf.

Janssen, M., Y. Charalabidis and A. Zuiderwijk (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management 29 (4):258-268. doi: 10.1080/10580530.2012.716740.

Janssen, M. and A. Zuiderwijk (2014). Infomediary Business Models for Connecting Open Data Providers and Users. Social Science Computer Review 32 (5):563-576. doi: 10.1177/0894439314525902.

Jetzek, T. (2013). The value of Open Government Data. Geoforum Perspectiv 12 (23 (2013)):47-56. Retrieved February 4, 2014, from http://journals.aau.dk/index.php/gfp/article/view/489/400.

Johnson, M.W., C.M. Christensen and H. Kagermann (2008). Reinventing your business model. Harvard Business Review. doi: https://hbr.org/2008/12/reinventing-your-business-model.

Kadaster (2012). Voor onbegrensde kansen. Jaarverslag 2011 [For unlimited chances. Annual report 2011], Apeldoorn. Retrieved June 30, 2014, from http://www1.kadaster.nl/1/jaarverslag/2011/pdf/jaarverslag_2011.pdf.

Kadaster (2015). In beweging. Jaarverslag 2014 [On the move. Annual report 2014], Apeldoorn, Retrieved May 1, 2015, from http://www.kadaster.nl/web/Jaarverslag-2014/2014-in-het-kort/Financiele-resultaten.htm.

Kamer van Koophandel (2015). Transitie. Jaarverslag 2014 [Transition. Annual report 2014]. Retrieved July 31, 2015, from http://www.kvk.nl/download/KvK_jaarverslag_2014%20_tcm109-406118.pdf.

Kelsey, J. and B. Schneider (1999). The street performer protocol and digital copyrights. First Monday 4 (6). doi: 10.5210/fm.v4i6.673.

Koski, H. (2011). Does Marginal Cost Pricing of Public Sector Information Spur Firm Growth? Keskusteluaiheita Discussion Papers, 28 September 2011 no. 1260. Helsinki: ETLA - The Research Institute of the Finnish Economy, Retrieved November 28, 2011, from http://www.etla.fi/wp-content/uploads/2012/09/dp1260.pdf.

Koski, H. (2015). The Impact of open data – a preliminary study. Helsinki: ETLA, Retrieved March 1, 2016, from https://www.w3.org/2013/share-psi/wiki/images/6/67/Impact_of_Open_Data_in_the_Public_Sector_Koski_2015.pdf.

Krek, A. and A.U. Frank (2000). The Production of Geographic Information - The Value Tree. Journal for Spatial Information and Decision Making 13(3):10-12. Retrieved November 28, 2014, from https://www.researchgate.net/publication/2803544_The_Production_of_Geographic_Information_-_The_Value_Tree.

Longhorn, R.A. and M. Blakemore (2008). Geographic information: value, pricing, production, and consumption. Boca Raton: CRC Press, Taylor & Francis Group.

Louzada, K. (2013). Crowdfunding, nu ook voor wetenschappers en academici. [Crowd funding, now also for scientists and academics]. Agentschap NL, Ministerie van Economische Zaken Retrieved September 17, 2014, from http://www.rvo.nl/sites/default/files/Crowdfunding%20v2.pdf.
Malone, T.W., P. Weill, R.K. Lai, V.T. D'Urso, G. Herman, T.G. Apel and S. Woerner (2006). Do Some Business Models Perform Better than Others? MIT Sloan working paper 4615-06, May 2006. doi: http://ssrn.com/paper=920667.
MRI [Market Rates Insight] (2012). First Integrated Study on Service Fees. Retrieved September 17, 2014, from http://www.marketratesinsight.com/news/NewsRelease-MRI_IntegratedStudy_FirstResults_052112.pdf.
Newbery, D., L. Bently and R. Pollock (2008). Models of Public Sector Information Provision via Trading Funds. Study commissioned jointly by the Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury in July 2007. Cambridge: Cambridge University. Retrieved February 26, 2009, from http://webarchive.nationalarchives.gov.uk/20090609003228/http://www.berr.gov.uk/files/file45136.pdf.
Omidyar Network (2014). Open for Business: How open data can help achieve the G20 growth target. A Lateral Economics report commissioned by Omidyar Network. Retrieved July 28, 2015, from https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf.
Onsrud, H.J. (1992). In support of cost recovery for publicly held geographic information. GIS Law 1(2):1-7. Retrieved August 1, 2008 from http://www.spatial.maine.edu/~onsrud/pubs/Cost_recovery_for_GIS.html.
Onsrud, H.J. (2004). The US national Spatial Data Infrastructure. Legal & economic issues and developments. In B. van Loenen and B.C. Kok (eds), Spatial data infrastructure and policy development in Europe and the United States, pp. 87-100. Delft: OTB Research Centre for Housing, Urban and Mobility Studies.
Osterwalder, A. and Y. Pigneur (2010). Business Model Generation. Hoboken, NJ: John Wiley & Sons.
Ovans, A. (2015). What is a business model? Harvard Business Review. Retrieved June 30, 2015, from https://hbr.org/2015/01/what-is-a-business-model&cm_sp=Article-_-Links-_-Top%20of%20Page%20Recirculation#.
PASC [Public Administration Select Committee] (2014). Public Administration Committee - Tenth Report. Statistics and Open Data: Harvesting unused knowledge, empowering citizens and improving public services. London. Retrieved March 17, 2015, from http://www.publications.parliament.uk/pa/cm201314/cmselect/cmpubadm/564/56402.htm.
PIRA International Ltd, University of East Anglia and KnowledgeView Ltd. (2000). Commercial exploitation of Europe's public sector information - Final report. European Commission Directorate General for the Information Society. Retrieved September 20, 2006, from ftp://ftp.cordis.lu/pub/econtent/docs/commercial_final_report.pdf.
Rappa, M. (2010). Business models on the Web. (open educational resource). North Carolina State University Retrieved July 13, 2014, from. http://digitalenterprise.org/about.html.
RDW (2015). Jaarverslag 2014 [Annual report 2014]. Retrieved July 30, 2015, from http://jaarverslag.rdw.nl/Documents/Wettelijk%20Jaarverslag%202014.pdf.
Reitsma, D. (2007). De consument en zijn behoefte aan geografische data. [The consumer and his requirements for geographical data]. In Y. Pluijmers (ed), Werkconferentie RGI-117 Geodata, van verstrekking naar toegang, Den Haag, 13 February 2007. Bedrijvenplatform Geo-Informatie, Woerden, pp. 35-47.
Rhind, D. ed. (2014). What is the value of open data? What is the Value of Open Data? Proceedings of an APPSI Seminar on 28 January 2014, London, April 2014. Retrieved October 20, 2015, from https://www.nationalarchives.gov.uk/documents/meetings/20140128-appsi-what-is-the-value-of-open-data.pdf.
Schiff, F. (2003). Business models of news web sites: a survey of empirical trends and expert opinions. First Monday [Online] 8, nr. 6 - June 2 2003:27. Retrieved May 13, 2008, from http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1061/981.
Shapiro, C. and H.R. Varian (1999). Information Rules: a strategic guide to the network economy. 1 vols. Boston, Massachusetts: Harvard Business School Press.
Tennison, J. (2012). Open Data Business Models Jeni's Musings. Retrieved August 20, 2014, from http://www.jenitennison.com/blog/node/172.
TNO (2015). Annual Report 2014. Retrieved March 31, 2015, from https://www.tno.nl/media/5574/tno_annual_report_2014.pdf.
Twynstra Gudde, CE Delft and Tauw (2011). Maatschappelijke Kosten Baten Analyse van de Basisregistratie Ondergrond, fase 1 [Social cost-benefit analysis of the Key Register Subsurface, phase 1]. Amersfoort, Retrieved July 27, 2015 from: https://bro.pleio.nl/file/download/32139762.

Uhlir, P.F., ed. (2009). The socioeconomic effects of public sector information on digital networks. Towards a better understanding of different access and reuse policies - Workshop Summary. Washington: National Academies Press. Retrieved March 1, 2010 from http://www.nap.edu/openbook.php?record_id=12687.

van Loenen, B. (2006). Developing geographic information infrastructures: The role of information policies. PhD Dissertation, OTB Research Institute, Delft University of Technology.

van Loenen, B. (2009). Developing geographic information infrastructures: the role of access policies. International Journal of Geographical Information Science 23 (2) :195-212. doi: 10.1080/13658810701851412.

van Loenen, B. and M. Grothe (2014). INSPIRE Empowers Re-Use of Public Sector Information. International Journal of Spatial Data Infrastructures Research 9 (2014):86-106. doi: 10.2902/1725-0463.2014.09.art4.

van Loenen, B. and J. Zevenbergen (2010). Assessing geographic information enhancement. International Journal of Spatial Data Infrastructures Research 5 (2010):244-266. doi: 10.2902/1725-0463.2010.05.art10.

van Loenen, B., J. Zevenbergen and J. de Jong (2006). Access to government information unnecessary remains a case of professional elite. Shaping the Change, XXIII FIG Congress Munich, Germany, 8-13 October 2006.

Verdonck Klooster & Associates, Ecorys and Berenschot (2014). Businesscase SGO-3. Maatregel 4: Open data [Business case SGO-3. Measure 4: Open data]. Retrieved May 22, 2014, from http://www.rijksoverheid.nl/documenten-en-publicaties/publicaties/2014/05/22/business-cases-5.html.

Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Paris: Information Economics. Retrieved November 1, 2011, from http://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments.

Wang, C.L., Y. Zhang, L.R. Ye and D.-D. Nguyen (2005). Subscription to fee-based online services: what makes consumer pay for online content? Journal of Electronic Commerce Research 6 (4):304-311. Retrieved June 30, 2015 from http://www.isy.vcu.edu/~jsutherl/Info658/FFSINFO.pdf.

Welle Donker, F. (2009). Public sector geo web services: which business model will pay for a free lunch? In B. van Loenen, J.W.J. Besemer and J.A. Zevenbergen (eds), SDI Convergence: Research, Emerging Trends, and Critical Assessment, pp. 35-51. Delft: Nederlandse Commissie voor Geodesie / Netherlands Geodetic Commission (NCG).

Welle Donker, F., B. van Loenen and A.K. Bregt (2016). Open Data and Beyond. ISPRS International Journal of Geo-Information 5 (4). doi: 10.3390/ijgi5040048.

Zhang, B., L. Collins and P. Baeck (2014). Understanding alternative finance. The UK Alternative Finance Industry Report 2014. London: Nesta, Retrieved November 30, 2015, from https://www.jbs.cam.ac.uk/fileadmin/user_upload/research/centres/alternative-finance/downloads/2014-uk-alternative-finance-benchmarking-report.pdf.

7 Open data and beyond

Frederika Welle Donker, Bastiaan van Loenen and Arnold K. Bregt

ISPRS International Journal of Geo-Information 2016, Vol. 5 (4), 48, doi: 10.3390/ijgi5040048. Submitted 22 January 2016; revised: 11 March 2016; accepted: 22 March 2016; published: 7 April 2016.

Abstract

In recent years, there has been an increasing trend of releasing public sector information as open data. Governments worldwide see the potential benefits of opening up their data. The potential benefits are more transparency, increased governmental efficiency and effectiveness, and external benefits, including societal and economic benefits. The private sector also recognises potential benefits of making their datasets available as open data. One such company is Liander, an energy network administrator in the Netherlands. Liander views open data as a contributing factor to energy conservation. However, to date there has been little research done into the actual effects of open data. This research has developed a monitoring framework to assess the effects of open data, and has applied the framework to Liander's small-scale energy consumption dataset.

Keywords: open data; effects of open data; monitoring framework

§ 7.1 Introduction

§ 7.1.1 Open data expectations

Since the 2009 Obama Executive Order (Obama, 2009), the 2010 Digital Agenda of the European Commission (2010), the 2011 Open Government Partnership Initiative (2016), and the 2013 G8 Open Data Charter (2015) there has been an increasing trend of government datasets published as open data. Open government data are associated with realising ambitions, such as a more transparent and efficient

government (Huijboom and van den Broek, 2011), reducing corruption (Granickas, 2014; David-Barrett *et al.,* 2015); increasing citizen trust in government decision-making (Grimmelikhuijsen, 2012); improving citizens' participation (Jetzek, 2013) and increasing democratic control (Bregt *et al.,* 2013); solving societal problems; increasing economic value due to companies creating innovative products and services with open data as a resource (Vickery, 2011; Omidyar Network, 2014); and efficiency improvements (WISE Institute, 2014; McKinsey Global Institute, 2013). For this article, we consider open data to be data available for re-use without any cost and without any restrictions in use.

By re-using and sharing data between government organisations, it is expected that internal efficiencies will be realised as transaction costs will decrease (Wise Institute, 2014; Houghton, 2011), as there will be no longer need for contract negotiations and policing between government organisations. In addition, by employing the principle of "collect once, re-use many times" governments can work more efficient and be more effective in decision-making. Open data may also contribute to higher data quality by fostering user feedback (incomplete data, errors, etc.) (de Vries *et al.,* 2011).

External benefits include solving societal problems (*e.g.* Uhlir, 2009; Attard *et al.,* 2015), as well as economic benefits (Vickery, 2011; McKinsey Global Institute, 2013; Pollock, 2011). Companies can use open data to develop innovative products and services, which may not only contribute to their turnover, but also contribute to society in general. For instance, actual roads information may lead to more efficient route planning and, thus, to less CO_2 emissions and shorter travel time (Helbig *et al.,* 2012).

§ 7.1.2 Open data principles

To facilitate effective re-use, open data has to comply with a number of principles. In 2007, 30 open data advocates defined a list of criteria for open government data. In 2010, the Sunlight Foundation (2010) updated these original principles[50], which have become the basis for many open data policies:

1 Completeness, including release of descriptive metadata, with the highest possible level of granularity, which will not lead to personally identifiable information;
2 Primacy, collected at the source, including information on how and where data were collected to allow verification by users;

50 In Chapter 4, the number of open data principles are extended to 14. However, at the time of writing this article, there were only 10 open data principles.

3 Timeliness, *i.e.* data should be released as quickly as possible;
4 Ease of physical and electronic access;
5 Machine-readable, in formats that allow machine-processing;
6 Non-discrimination, available to anyone with no requirement of identification or justification;
7 Use of commonly-owned or open formats;
8 Licensing, *i.e.* no imposition of attribution requirements and preferably labeled as part of the public domain;
9 Permanence, *i.e.* data should remain online with appropriate version-tracking and archiving;
10 Usage costs, *i.e.* data available preferably free of charge.

Although most of these principles were agreed upon for public sector data (Kulk and van Loenen, 2012), they should equally apply to private sector organisations with a mandated public task. There are ample private organisations mandated to perform a certain public task and generate data in the process, yet, these organisations are often exempted from open government data policies. However, it could be argued that data generated in the process of performing a public task are a public good (Attard *et al.*, 2015) and should, therefore, adhere to the same open data principles. Geographical data, such as topographical maps and the underlying Earth observation data, energy data, and health data are top-listed by the European Commission for release as open government data due to the high demand from re-users (European Commission, 2014). However, as data holders of such data are often private companies or (semi) privatised government organisations, these organisations often fall outside the scope of the legal framework related to accessibility of public sector information.[51] There are a number of open access initiatives for non-government data, such as publicly-funded research data (OECD, 2007), earth observation data (CODATA, 2015), and health data U.S. Department of Health & Human Services, 2016). In addition, some private companies already adopting open data practices, often do so not for altruistic reasons but to market their products more effectively (Herzberg, 2016) or to enable data-as-a-service business models (Deloitte, 2012).

51 In article 2.2. of the Access Directive 2003/4/EC and in article 3.9 of the INSPIRE Directive 2007/2/EC an exception is made for environmental information. In these articles, 'public authority' is defined as "government or other public administration, including public advisory bodies, at national, regional or local level; any natural or legal person performing public administrative functions under national law, including specific duties, activities or services in relation to the environment; and any natural or legal person having public responsibilities or functions, or providing public services, relating to the environment under the control of a body or person falling within (a) or (b). (Footnote added October 28, 2016.)

§ 7.1.3 Assessment of the effects of open data

Although there is ample anecdotal evidence (Vickery, 2011; McKinsey Global Institute, 2013) and case studies (Hogge, 2015) to demonstrate the benefits of open data, to date little is known about the underlying organisational mechanisms and implications of open data as most open data impact assessments are *ex ante* Koski, 2015). The impact of open data has been only recently addressed in macro-economic research projects. Detailed studies on the costs of the implementation are scant, results on the impact of open data on an organisation's workflow processes and from a user perspective are very limited, and a monitoring framework that supports the assessment of the impact in a scientifically solid way is lacking. This research aims to provide an *ex post* assessment of open data effects for Liander, a private company responsible for energy network distribution and administration.

We have developed a monitoring framework to assess open data effects and have applied the framework to Liander's small-scale energy consumption and generation dataset. This research is performed through a literature study on open data and the performance of open data. The developed assessment framework is applied to the open data of Liander. The benchmark was performed through structured interviews with selected Liander staff and users of Liander data, and analysing web statistics. The follow-up was performed through structured interviews with Liander staff, questionnaires to users and analysing web statistics.

In Section 7.2, we explain the potential of open energy data and describe structure of the Dutch energy sector. We provide a description Liander data flows and data services prior to open data implementation in 2013. We also briefly describe the Open Data and Beyond projects. Section 7.3 provides the theoretical basis of impact assessment frameworks, and present indicators required for such a framework. We present transactions costs theory, and the organisational effects of implementing open data. In Section 7.4, we provide the assessment outcomes related to internal effects, external effects and relational effects. We describe the results of applying the framework to monitor the effects of opening a geographic dataset. We conclude in Section 7.5 with our recommendations.

§ 7.2 Liander data

§ 7.2.1 Importance of open energy data

Many countries and organisations are implementing or thinking of implementing open data policies for their data in line with open data agendas, such as the G8 Open Data Charter. In the European Union, the Digital Agenda for Europe of DG Information Society European Commission, 2010) provides an extra stimulus to start open data projects. In the Netherlands, the Dutch Digital Agenda of the Ministry of Economic Affairs (2011), the National Open Data Agenda (NODA) of the Ministry of the Interior and Kingdom Relations (2015) and the Open Data Roadmap of the Ministry of Infrastructure and Environment (2012) are examples of national open data agendas. In addition to the public sector, the private sector recognises potential benefits of making their datasets available as open data (*cf.* Deloitte, 2012).

It is widely recognised that open energy data can make a valuable contribution to inform consumers better about energy reduction and improved energy efficiencies (Vickery, 2011; Omidyar Network, 2014; McKinsey Global Institute, 2013; DECA, 2010). Energy efficiencies are necessary, as fossil fuels become a limited resource. In addition, to meet the targets set in the Paris Agreement, a significant reduction in emissions as part of the method for reducing greenhouse gas is required (UN, 2015).

Liander, an energy network administrator in the Netherlands wants to take the lead in the open energy data field as a trailblazer. Liander views open data as a contributing factor to energy conservation and the reduction of carbon emissions. Liander expects that it will benefit from open data in more than one way. Open data will not only lead to societal benefits and a more efficient operational management but also to a better image as a transparent monopolist and an open data trailblazer leading the way for other network companies.

§ 7.2.2 Dutch energy sector

The Dutch energy sector was liberalised in 1998, *i.e.* the formerly public sector utilities became private organisations mandated to execute certain public tasks. The sector was unbundled into four pillars: production, transmission (high voltage), distribution (low and medium voltage), and supply. Transmission and distribution are regulated public

tasks, whereas production and supply are liberalised to allow for more competition. Although it could be argued that liberalisation of the Dutch energy sector has its drawbacks (van Damme, 2006), this discussion is not within the scope of this article. Currently, the national high voltage grid is administered by TenneT, a government-owned company. The regional low to medium voltage networks are managed by eight privatised distribution network companies, of which Liander is one of the largest network administrators.

§ 7.2.3 Liander data and services

Liander is the largest Dutch energy network administrator covering five out of 12 provinces in the Netherlands, transporting electricity and natural gas to 5.2 million connections (households and businesses). The coverage of Liander's network is about 37% of the Netherlands (see Figure 7.1). Although Liander is a private company in the legal sense of the way, it has a regulated public task to administer and maintain energy networks, which significantly influences the way Liander is doing business. According to the Energy Act 1998, network administrators have an obligation to share and exchange their data between network administrators and TenneT. However, there is no legal requirement to disseminate energy data to the general public, other than general information related to, *e.g.* tariffs.

Within Liander, core data chains are established to streamline work processes. Within a core data chain, only the departments concerned have access to the specific datasets; departments outside the chain have no access to the data.

Liander has a legal obligation to supply large-scale (businesses) and small-scale (households) energy consumption and generation data to some government agencies, such as Statistics Netherlands (CBS). CBS receives monthly and quarterly reports on large-scale and small-scale energy consumption and generation. These reports are based on raw data (connection level), *i.e.* they contain addresses. CBS receives the data under strict conditions that the raw data must not be made available as open data.

Liander supplies companies (most often engineering firms) with customised data (at the connection level) under contract conditions and for a fee. The contract terms prohibit further distribution of the data. Liander also supplies energy consumption and generation data to local governments (municipalities and provinces) and building corporations. Local governments can obtain data in two ways: either by lodging a request directly to Liander or by using EnergieInBeeld.nl (Energy in Focus), a web service developed by Liander in cooperation with the other energy network administrators, Local governments can visualise and download aggregated data via

EnergieInBeeld.nl free of charge. If the raw data at the connection level are required, local governments can send an automated request via EnergieInBeeld.nl to Liander, as such data contain personal data.

Legend

1. RENDO Networks
2. Cogas Infra and Management
3. Liander
5. Stedin
6. Westland Infra
7. Enduris
10. Endinet Group
11. Enexis

FIGURE 7.1 Regional network administrator coverage in the Netherlands (source: Energieleveranciers.nl, 2016)

Until December 2013, the subscription fee for EnergieInBeeld.nl was around €6,000 per annum for local governments. Since 2014, EnergieInBeeld.nl can be used free of charge, although local governments must still register in advance to receive login details required to download data. Since October 2015, everybody can use EnergieInBeeld.nl with a public login code. To ensure the free downloadable data do not contain personal data (data at the connection level), the aggregation level for the general public is set to a five-digit postcode, and for local governments to a six-digit postcode; see Figure 7.2. A standard Dutch postcodes consists of four numbers and two letters and covers about 10–20 households. To publish data at a more aggregated level, postcodes may be limited to the first five-digits (about 1–2 blocks) or to the first four-digits (about a neighbourhood).

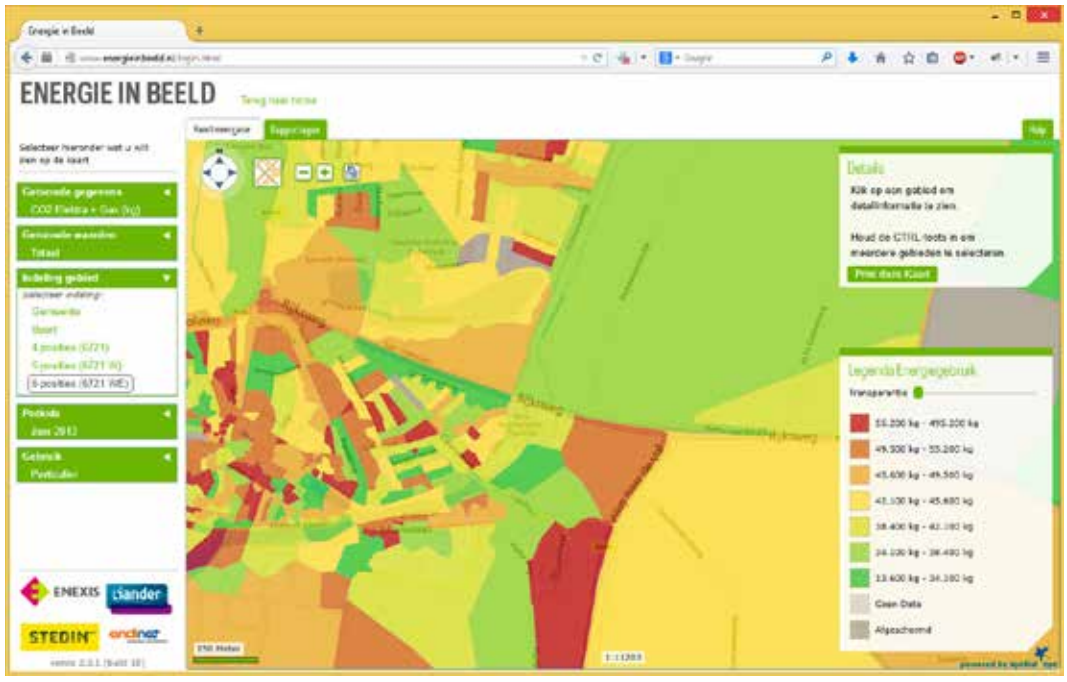


FIGURE 7.2 Sample of "Energie In Beeld" (Energy in Focus) web viewing service on the highest aggregation level (six-digit postcode level)(2016)

§ 7.2.4 Open Data and Beyond projects

As part of the Next Generation Infrastructures Program in the Netherlands, a consortium consisting of Liander, Delft University of Technology, and Wageningen University carried out a number of projects. Liander intended to release a number of datasets as open data. The first project "Open Data and Beyond I" was carried out in 2012 and focused on legal, technical, and organisational preconditions for implementing open data.

The "Open Data and Beyond II" project (July 2012 until March 2015) was established to develop a framework to monitor the effects of open data. Before opening (aggregated) small-scale energy consumption data and other datasets at a later date as open data, Liander wanted to assess the impact of open data on the organisation ex ante.

The assumption is that by implementing open data transaction costs will be reduced, as there are no more transaction costs incurred during contract negotiations and policing licence conditions. However, implementing open data will incur implementation costs, such as extra server capacity and human resources.

Most of the expected effects of releasing data as open data will only be visible in the long term. To be able to measure the effects, a benchmark measurement was carried out before releasing the dataset as open data.

§ 7.3 Monitoring framework to measure impact of open data

§ 7.3.1 Theoretic framework

Monitoring the effects of a policy change can be carried out several ways. One such way for monitoring is using performance indicators. For performance indicators to provide precise and accurate performance information, they should be designed and implemented within a performance management system (Giff and Crompvoets, 2008). Developing a framework for such a monitor framework consists of seven steps.

1. Develop a performance framework to describe what the program is about, description of the organization's mission and strategic goals;
2. Identify the most important elements, or key performance areas which are most critical to understanding and assessing your program's success;
3. Select the most appropriate performance measures;
4. Determine the "gaps between what information you need and what is available;
5. Develop and implement a measurement strategy to address the gaps;
6. Develop a performance report which highlights what you have accomplished and what you have learned;
7. Learn from your experiences and refine your approach as required.

These steps are iterative.

§ 7.3.2 Indicators

Monitoring is usually carried out using indicators. Performance indicators will provide a view of the impact of an activity, and should be (Giff and Crompvoets, 2008):

- Specific: clearly defined and easily understood;
- Measurable: quantifiable to facilitate comparison;
- Attainable/feasible: practical and cost-effective to implement;
- Relevant: true representation;
- Timely and free of bias;
- Verifiable and statistically valid;
- Unambiguous: a change in an indicator should result in clear and unambiguous interpretation
- Comparable and time-bound.

In addition, an indicator should be communicable. For this project, we selected five core requirements: specific, measurable, practical, relevant, free of bias, and communicable.

It is important to distinguish between output, outcome, and impact. Output concerns the products and services provided. Ultimately, the outcome of a program should relate to the mission and the mandate of the program provider (Environment Canada, 2000). Outcome is the result of an activity. Impact relates to the way the outcomes contribute to the (strategic) goals of an organisation or effects in society. Table 7.1 shows an example of the relation between activity, output, outcome, and impact for Liander.

	ACTIVITY	OUTPUT	OUTCOME	IMPACT
Liander	Releasing data as open data	Open small-scale consumption dataset	Energy apps based on open data	Contribution to energy conservation

TABLE 7.1 Example of output-activity-outcome-impact for Liander

§ 7.3.3 Potential effects of open data

Implementation of open data incurs costs for the data supplier, such as extra server capacity to facilitate higher download traffic and to host open data on a separate server. In addition, the data has to be made suitable for release as open data. For instance, the data must be aggregated from the household level to the postcode area level to comply

with requirements for protection of personal data. Investments must also be made to optimise open data usage, such as setting up systems to make use of user feedback, a help-desk, user-friendly interfaces, etc. These costs are counted as internal data supply effects. Most of the costs are incurred in the start-up phase. Once preparing open data for publication has become part of the workflow processes, the additional transaction costs of the data supplier are expected to go down. The expectations are that once data are released as open data, the transaction costs for internal users within Liander but outside the core data chains will go down. The monitor will measure internal user transaction costs to test these expectations.

Costs are only one aspect of monitoring the effects of open data. The expectation is that data policy change will also cause effects that are intangible. Examples of such effects are more effective decision-making related to energy saving measures, development of innovative applications, and more transparency, thus increasing the public image of Liander. These effects are hard to monetise, however, they are a real part of open data effects. In addition, implementation of open data will require a change in the organisation's culture, as not only work processes have to be amended, which require costs and efforts to maintain data suitable for publication. There may also be a reluctance to open data for fear of inadequate interpretation of data (Martin *et al.*, 2013), embarrassment over content or quality, and worries about privacy and liability (Deloitte, 2012).

§ 7.3.4 Transaction costs

One way to measure the impact of a policy change, is to compare transaction costs before the policy change with transaction costs after the policy change. Transaction costs theory deals with the cost of transacting. Every exchange of the product entails costs that result from both parties attempting to determine the valued characteristics of the good (North, 1990). It takes resources to measure these characteristics and additional resources to define and to measure the rights that are transferred to the user with the exchange of the goods. The costs associated with these efforts are considered part of the transaction costs (North, 1990; Williamson, 1985; Williamson and Masten, 1995; Sholtz, 2001).

According to van Loenen et al. (2010), information trade is a transaction which involves data and service providers, on one hand, and data users, on the other hand. In the process of exchanging data, the potential users and providers have to agree on the characteristics of the data, and on the conditions of exchange. In this process of communication, costs occur on both sides. A user (both internal and external) will typically undertake the following activities that incur transaction costs:

- Activity 1: searching for the data supplier;
- Activity 2: inquiring about the general conditions of exchange;
- Activity 3: inquiring about specific conditions related to price and availability;
- Activity 4: defining the exact characteristics of the data;
- Activity 5: acquiring and testing (a sample of) the product for fitness of use;
- Activity 6: reading and understanding the documentation related to the licence and fees;
- Activity 7: obtaining the actual dataset;
- Activity 8: uploading the data into the software, harmonising, adapting the format.

The data supplier incurs transaction costs related to making data and metadata available; setting up a portal and a contact point/help desk; negotiating, drafting, and exchanging contracts; and collecting fees and enforcing conditions (Poplin, 2010).

§ 7.3.5 Effects of releasing data

According to its formulated mission, Liander wants "to strive for proving a service that gives everybody access on equal conditions to reliable, affordable and sustainable energy" (Liander, 2016). Although this is a lofty and abstract mission, we translated this ambition into the following strategic goals:

1. Continuously optimise performance on services, security of supply, and costs;
2. Improve management of energy flows and insight into energy consumption;
3. Help customers save energy and switch over to renewable energy sources.

We expect to distinguish three different effects related to the release of open data. Figure 7.3 shows the relation between the different effects.

In order to be able to measure and monitor the effects of open data, an indicator framework was developed. The proposed indicator framework was fine-tuned by Liander stakeholders during a workshop on 20 September 2012. In addition, Liander actively participated in brainstorm sessions and by making the dataset available as a pilot during hackathons. Feedback received from these events contributed to assessing what the user needs were and how best to present the data.

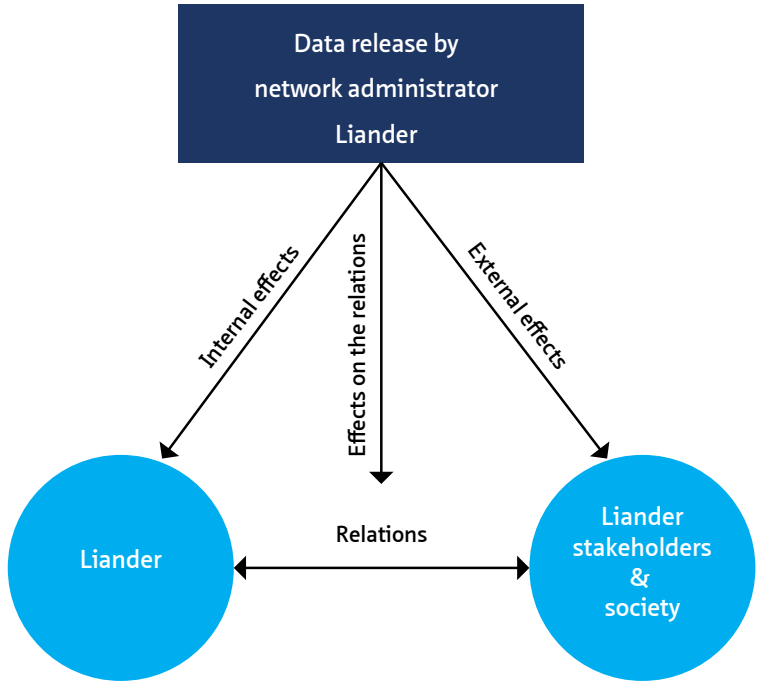


FIGURE 7.3 Effects of open data for Liander (after Bregt *et al.*, 2013)

After the Open Data and Beyond pilot project, it was decided that — according to the open data principles — Liander would not require prior registration of re-users. This means that it will not be possible to track re-users that download the data. Instead, the developed framework focuses on measuring, inter alia, transaction costs, both for the supply of open data and for the users. We have distinguished between different types of costs. To ensure that transaction costs will be comparable between organisations and between years, we have expressed transaction costs in full-time equivalents (FTE) or man-hours (MH). For this research, the following effects were measured:

1. The external effects of open data for Liander by measuring:
   a. usage of Liander open data in society (numbers and types of users);
   b. nature and intensity of use per type of user;
   c. effects on transaction costs/hours of external user.
2. The internal effects of open data for Liander by measuring:

TOC

a effects on supply transaction costs/hours within Liander (preparation and operational use of data);
b effects on data quality;
c effects on customer service of Liander;
d usage of Liander open data internally (numbers and types of users);
e nature and intensity of use per type of user;
f saved transaction costs/hours by internal users.

3 The relational effects of open data for Liander and society by measuring:
a effects on communication from society to Liander;
b effects on communication from Liander to society;
c effects on the image of Liander as a transparent energy network administrator.

Table 7.2 presents the monitor framework consisting of 10 indicators. We applied the framework to the small-scale energy consumption dataset, which was released as open data in September 2013. The indicator framework monitors the outcomes once a year to assess the impacts. To provide a basis for comparison a baseline measurement to set a benchmark was carried out prior to releasing the dataset as open data in September 2013. The text in italics denotes the results of the follow-up assessment carried out in December 2014. The benchmark and the follow-up were carried out by the authors and by Liander. The transaction costs measured are expressed in man-hours (MH) and full-time equivalent (FTE) for human resources.

GOAL MONITOR	INDICATOR	METHODOLOGY	RESULTS
Determining external effects on Liander	1a. Number and type of external users	Distinguish between different types of users (public sector/private sector) and frequency Derivable from current contracts, number of downloads; number of unique IP numbers	Public sector: • CBS (annually and quarterly); • municipalities (quarterly); • building corporations (quarterly). Private sector: • engineering firms (quarterly); • energy suppliers: (quarterly). *Increase in use by other private sector companies and citizens (from interviews and surveys)*
	1b. Nature and intensity of data usage per type of user	Distinguish between which data formats/services are requested/downloaded, and how often Derivable from (web) statistics for web services	• EnergieInBeeld.nl: ca. 40 page views/month • open data web service (pilot): ca. 48 page views/month. *Significant increase in page views/month, stabilizing after initial period*

TOC

	1c. Transaction costs per type of user	Interviews with two selected key users (CBS and City of Amsterdam) before implementation open data	• CBS: 5 MH to request (non-open) data; delivery time: max. one month; 40–56 MH to adapt data for re-use. *No changes* • Amsterdam: 8 MH to locate (non-open) data; 4 MH to assess suitability of sample; 8 MH to set up contract; 1 FTE/annum to adapt data for re-use. *Less time to locate data as open data and to adapt data*
Determining internal effects on Liander	2a. Transaction costs Liander to prepare release open data and keep operational open data	Derivable from project administration (personnel; IT-investments)	• Preparation and pilots: 1.6 FTE • Preparatory research (legal, technical and organisational preconditions): 1.3 FTE • Development monitor: 0.4 FTE *After initial phase, operational costs are negligible*
	2b. Numbers and type internal users	Distinguish between departments, unique IP-numbers, contact details	• This indicator could not be measured for benchmark as there were no such records. • For the follow-up Liander will set up a register. *No changes*
	2c. Nature and intensity of data usage per type of internal user	Distinguish between which data formats/services are requested/downloaded, and how often	• This indicator could not be measured for benchmark as there were no such records. • For the follow-up Liander will set up a register. *No changes*
	2d. Transaction costs of internal users outside core data chain	Derivable from Service Point estimates	• Requests from users without SAP authorisation: 2 MH/week. *No changes*
Determining relation effects between Liander and society	3a. Communication from society to Liander	Monitoring and analysing reaction (call centers, energy failure desk, e-mail). Monitoring social media (Facebook Likes, Twitter, LinkedIn)	• Measurements indicated that open data have no effect on communication, as most communication concerns individual connections. *No changes*
	3b. (Pro)active communication from Liander to society	Monitoring and analysing own social media (Facebook Likes, #Liander, #OpenData, LinkedIn, Open Data groups	• Twitter: energy failures are tweeted; therefore, number of tweets depends on energy failures. Max. was 30 tweets in July 2013. • Facebook: about 2 posts/week, mostly about sustainable energy generation. *No changes*
	3c. Image/transparency Liander	Not part of regular customer satisfaction survey but will be assessed through stakeholder surveys	• Roundtable meetings with stakeholders in 2013 indicated that stakeholders highly value transparency but customers to a lesser extent. *Liander has become a sought-after party for Smart City and other energy projects*

TABLE 7.2 Indicators for effect assessment.

TOC

§ 7.4 Benchmark outcomes and analysis

In 2012 and 2013, Liander participated in several open data hackathons (province of Flevoland, City of Amsterdam), whereby area samples of small-scale energy consumption data were made available as open data to the hackathon participants. In September 2013, Liander published the complete small-scale energy consumption dataset as open data, including historical data, and in several formats. Since then, Liander has released other datasets, such as energy failure data, as open data.

Prior to the release of open data, a benchmark was carried out using the monitor framework. In September 2014, the monitor was repeated. The repeat monitor only showed some short-term effects, which are analysed below.

§ 7.4.1 External effects

Prior to releasing the dataset, external users consisted mainly of municipalities, building corporations, Statistics Netherlands (CBS), engineering firms, and a few organizations producing value added services for public sector bodies. CBS collects small-scale and large-scale energy consumption and generation data as part of their legal mandate. Open data did not have an effect on the transaction costs of CBS, as the data required are on household level and, therefore, not available as open data.

Municipalities and building corporations can now access small-scale energy consumption and generation aggregated data either via Liander's open data portal or via existing web services, and large-scale data via channels in place prior to open data. Prior to open data in 2013, to obtain energy consumption and generation data on connection level, the average transaction costs for a municipality were about 32 man-hours to locate required data, contact the data holder, and to negotiate and exchange contracts, see Figure 7.4. Municipalities could also opt to pay a subscription fee for EnergieInBeeld.nl to obtain aggregated data. It is expected that more municipalities will use open datasets in addition to EnergieInBeeld.nl. Building Corporations require data at the connection level and, therefore, it is not expected that their transaction costs will decrease due to open data.

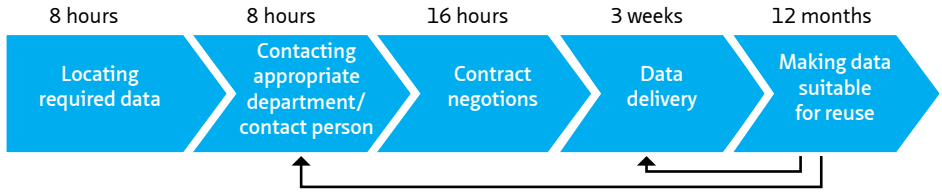


FIGURE 7.4 Example of a data supply chain for external users of Liander data prior to open data, expressed in man-hours, including feedback loops and data delivery time.

One year after opening small-scale energy consumption and generation data, the transaction costs for external users have decreased somewhat. However, this is probably due to Liander streamlining its data processes and the removal of the annual fee for EnergieInBeeld.nl rather than to open data as such.

Prior to open data there were no regular re-users of the datasets apart from municipalities and building corporations. Within weeks of release, one company has re-used the data for energy usage apps and web services. Although the Liander open data portal contains a page showing samples of open data applications (Liander, 2016), to date, only a few commercial companies have linked their products to this webpage. As Liander supplies datasets via the open data portal without prior registration, it is difficult to measure if and for what the datasets are used, a common problem with monitoring open data (*cf.*) (Bregt *et al.*, 2014). During an Apps4Energy hackathon in 2015, (potential) re-users of Liander open data were surveyed as part of the follow-up. Figure 7.5 shows the types of surveyed users as a percentage.

From the follow-up, it emerged that Liander open data are mainly used by external organizations for energy planning and policy advice. It appears that Liander data are used to improve the quality of existing applications, such as Energy Atlases, rather than for new applications. In general, Liander open data users appreciate the availability and quality of the open data, however, they would prefer the data to be timelier (near real-time). In addition, the users would like large-scale energy consumption open data to be available as well. However, the release of this dataset is not foreseeable in the near future as there are barriers related to data protection.

TOC

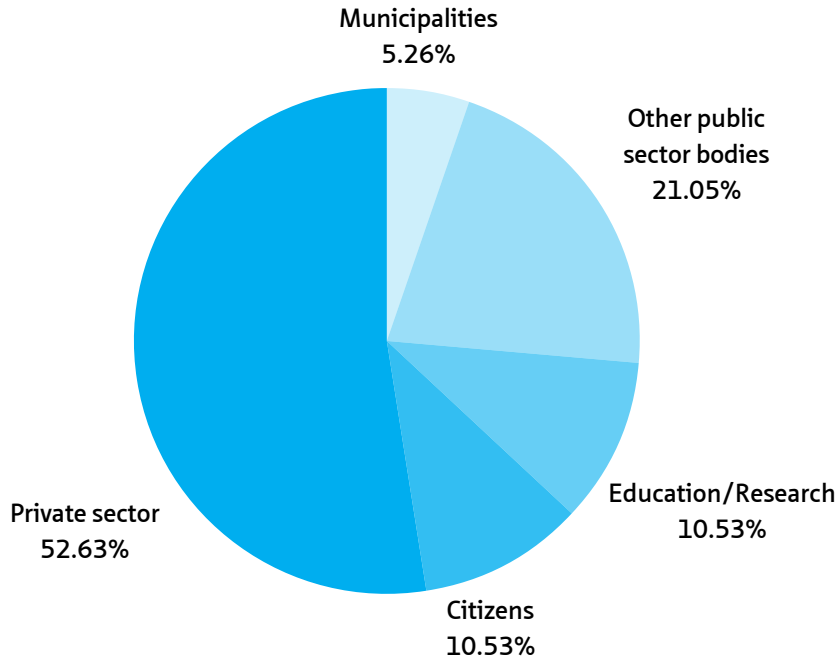


FIGURE 7.5 Types of Apps4Energy surveyed users as a percentage

§ 7.4.2 Internal effects

For the benchmark, the transaction costs both for data supply and for internal users of Liander outside the core data chain were set to zero, in order to be able to show the effects in the future. Re-use of the datasets within the core data chain and outside the core data chain ran via existing procedures. Within the core data chain, the transaction costs are zero because users have a SAP authorisation to access the data. Users without SAP-authorisation must send a request to the Service Point. As the IT department does not distinguish between different types of requests and different datasets, these transaction costs are set to zero as well. The costs of releasing open data are primarily personnel costs related to the research projects. An extra server was purchased to host the dataset to prevent hackers from entering the main server. Liander opted to develop their own data platform rather than re-use an existing platform to maintain control, thus alleviating some of the concerns of the IT department.

In the short term, the main effect of releasing one open dataset was a significant increase in the number of page views and time spent online. However, it could not

be assessed how many of these page views/downloads were due to internal or to external users. Since 2014, the number of visits has more or less stabilised with peaks during energy hackathons.

Another effect is that releasing one dataset has paved the way for releasing more datasets. The technical and operational lessons learned from the pilot phase have been taken onboard in the processes for releasing more open datasets. There have been no discernible internal effects on the transaction costs of data supply. Liander is streamlining its data supply services by having added a Datashare Service, in which data requests (both for open and for closed data) be automated. Open data has become embedded in Liander's primary processes and the Open Data team has been disbanded without consequences for the sustained availability of open data.

§ 7.4.3 Relational effects

The benchmark measured the relation effects prior to the release of open data. Liander makes only limited use of social media. Communication to society is mainly related to innovations, end usage, and sustainable energy generation, and is limited to Facebook. Energy failure inquiries and reports are mainly communicated via Twitter. There was no communication related to small-scale energy consumption prior to the release of open data.

Since the release of the dataset as open data in September 2013, Liander has communicated via social media and via the Dutch LinkedIn Open Data group. Initially, the reactions were positive until it emerged that the dataset did not comply to the open data criterion of being available without a licence, or with an "open" licence, such as a Creative Commons BY (For more information on Creative Commons licences, see https://creativecommons.org/licenses/) license, requiring only source attribution. Instead, the dataset was released with a licence prohibiting derived products to be used for commercial purposes, or to alter the dataset. After a small storm of protests, the licence was changed to a Creative Commons BY license early October. In spite of the somewhat heated discussion, Liander received 17 Likes out of 21 reactions.

After the release of open data, Liander has become better known in the open data community and has become an active participant in various Smart City projects.

§ 7.4.4 Benchmark outcomes summary

Table 7.2 shows a summary of the selected indicators and outcomes. Not all indicators will be measurable in the follow-up, as once open data is published, it is almost impossible to track the actual numbers and types of external users. Even if registration prior to downloading would be employed, users can still set up shadow datasets. It is expected that the number of internal users will remain stable. It is also expected that, once established, the operational costs of open data will be negligible.

The repeat monitor in December 2014 showed that providing open data has had significant relational effects. The open data tools and interfaces initially provided by Liander were not viewed as very user-friendly. Liander used the feedback to make improvements to its open data platform. In addition, Liander has made the EnergieInBeeld.nl web service more user-friendly in close cooperation with the City of Amsterdam and private sector organisations, and made the service available to the general public. As Liander has become part of the Dutch open data community, they are a sought-after party to participate in various Smart City projects and other energy projects. Liander has become an open data champion paving the way for other energy administrators in the Netherlands.

§ 7.5 Conclusions and recommendations

To develop both quantitative and qualitative indicators to assess the success of a policy change is a challenge for open data initiatives. In the orientation phase of this research project, indicators were proposed but not tested. Releasing multiple datasets as open data permanently by Liander offers an excellent opportunity to fine-tune the proposed framework, apply it, and provide insight into the effects. As the dataset had not yet been released as open data, we were able to carry out a baseline measurement.

Most of the effects of releasing data as open data will only be noticeable in the long term. Similar research in the Netherlands assessing the effects of releasing the large-scale topographic dataset (Bregt *et al.*, 2013; Bregt *et al.*, 2014; Grus *et al.*, 2015) and in the United Kingdom assessing the value of Ordnance Survey datasets to the economy (Carpenter and Watts, 2013) show that the short-term effects are mainly more downloads and page views, and more communication between data suppliers and users, and between users. Liander has experienced similar short-term-effects. Within a couple of weeks after the open data was launched, the first app based on Liander data was launched.

The expected long-term effects were initially that more apps and web services would be developed, new user groups would be accessed, transaction costs would lower for existing customers and for the organisation, and hopefully fewer questions about Liander's activities (*e.g.* about energy failures). To date, only lower transaction costs have been realised. From the follow-up, it emerged that Liander open data are used by a wide range of users and have had a positive effect on energy consumption visualisation applications. However, it remains a challenge to quantify such effects. It also remains a challenge to monitor re-users, as there is no registration.

The open data expectations of external users vary. Although the original small-scale energy consumption dataset is the most widely used open dataset of Liander, it has also created a demand for other datasets, such as large-scale energy consumption data and small-scale energy generation data (windmills, solar panels). The latter data were indeed added to the small-scale energy consumption data. More municipalities download the open datasets to supplement EnergieInBeeld.nl data. Both the private sector and the municipalities use Liander open data to improve existing applications and work processes rather than to create new products.

Liander has successfully demonstrated that private energy companies can release open data, and has successfully championed the other Dutch network administrators to follow suit. In 2015, the other network administrators in the Netherlands have also published their small-scale energy consumption data and Enexis, the second largest Dutch network administrator, has also published asset data as open data.

The monitoring framework developed in this project monitors the societal effects. Liander assessed the project to be very helpful and a key component in the process towards the successful implementation of the open data strategy at Liander. So far, the developed framework appears to be suitable to measure the open data effects.

Acknowledgments: We would like to thank the Next Generation Infrastructures Programme for making this research possible. In addition, we would like to thank Liander for their participation in this project.

Author contributions: As part of the Open Data and Beyond Consortium, Bastiaan van Loenen carried out a substantial part of the work for the Open Data and Beyond project to establish the legal, technical and organisational preconditions for implementing open data. Arnold Bregt provided an input for establishment of the open data organisational preconditions. The established preconditions were used as main criteria for the Open Data and Beyond II project. Frederika Welle Donker, Bastiaan van Loenen and Arnold Bregt, jointly developed the assessment framework based on a conceptual model developed by Arnold Bregt. Frederika Welle Donker carried out the benchmark and analysed the results. Frederika Welle Donker and Arnold Bregt carried out the follow-up and all three authors jointly analysed the results. Frederika Welle Donker as lead author and Bastiaan van Loenen as second author wrote this paper. Arnold Bregt made useful suggestions to the final version of the paper.

Conflicts of interest: The authors declare no conflict of interest.

References

Attard, J.; Orlandi, F.; Scerri, S.; Auer, S. A systematic review of open government data initiatives. Gov. Inf. Q. 2015, 32, pp. 399–418.

Bregt, A.; Grus, L.; Eertink, D. Wat Zijn de Effecten van een Open Basisregistratie Topografie na Twee Jaar? Wageningen University: Wageningen, The Netherlands, 2014; p. 49.

Bregt, A.K.; Castelein, W.T.; Grus, L.; Eertink, D. De effecten van een open basisregistratie topografie (BRT); Wageningen University: Wageningen, The Netherlands, 2013; p. 40.

Carpenter, J.; Watts, P. Assessing the Value of OS Open DataTM to the Economy of Great Britain—Synopsis; UK Ordnance Survey: England, UK, 2013; p. 32.

Danish Enterprise and Construction Authority (DECA). The Value of Danish Address Data: Social Benefits from the 2002 Agreement on Procuring Address Data Etc. Free of Charge; DECA: Copenhagen, Denmark, 2010; p. 8.

David-Barrett, E.; Heywood, P.M.; Theodorakis, N. TACOD: Towards a European Strategy to Reduce Corruption by Enhancing the Use of Open Data; Research Centre on Security and Crime (RiSSC): Turin, Italy, 2015; p. 102.

Deloitte. Open Data. Driving Growth, Ingenuity and Innovation; Deloitte: London, UK, 2012; p. 36.

de Vries, M.; Kapff, L.; Negreiro Achiaga, M.; Wauters, P.; Osimo, D.; Foley, P.; Szkuta, K.; O'Connor, J.; Whitehouse, D. Pricing of Public Sector Information Study (Popsis)—Summary Report; Deloitte Consulting CVBA: Brussels, Belgium, 2011.

EnergieLeveranciers.nl. Overzicht netbeheerders elektriciteit; Available online: http://www.energieleveranciers.nl/netbeheerders/elektriciteit (accessed on 29 March 2016).

Environment Canada. Manager's Guide to Implementing Performance-Based Management; Environment and Climate Change Canada: Gatineau, QC, Canada, 2000; p. 55.

European Commission. Communication from the Commission of 19 May 2010 to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions—A Digital Agenda for Europe; European Commission: Brussels, Belgium, 2010.

European Commission. Guidelines on Recommended Standard Licences, Datasets and Charging for the Reuse of Documents; European Commission: Brussels, Belgium, 2014; p. 10.

G8 Open Data Charter and Technical Annex. Available online: https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex (accessed on 15 September 2015).

Giff, G.A.; Crompvoets, J. Performance indicators a tool to support spatial data infrastructure assessment. Comput. Environ. Urban Syst. 2008, 32, pp. 365–376.

Granickas, K. Open Data as a Tool to Fight Corruption; ePSI Platform Topic Report No. 2014/04; ePSI Platform: Amsterdam, The Netherlands, 2014; p. 20.

Grimmelikhuijsen, S.G. Transparency and Trust: An Experimental Study of Online Disclosure and Trust in Government. Ph.D. Dissertation, Utrecht University, Utrecht, The Netherlands, 2012.

Grus, L.; Bregt, A.; Eertink, D. De Effecten van Open Data BRT na 3 Jaar; Wageningen University and the Kadaster: Wageningen, The Netherlands, 2015; p. 7.

Helbig, N.; Cresswell, A.M.; Burke, G.B.; Luna-Reyes, L. The Dynamics of Opening Government Data. A White Paper; Center for Technology in Government, University of Albany: Albany, NY, USA, 2012; p. 34.

Herzberg, B. The next frontier for open data: An open private sector, Voices. Perspectives on Development. Available online: http://blogs.worldbank.org/voices/next-frontier-open-data-open-private-sector (accessed on 7 March 2016).

Hogge, B. Open Data: Six Stories about Impact in the UK; Omidyar Network: Redwood City, CA, USA, 2015; p. 61.

Houghton, J. Costs and Benefits of Data Provision. Report to the Australian National Data Service; Centre for Strategic Economic Studies, Victoria University: Melbourne, VI, Australia, 2011; p. 66.

Huijboom, N.; van den Broek, T. Open data: An international comparison of strategies. Eur. J. ePract. 2011, 12, pp. 1–13.

Jetzek, T. The value of open government data. Geoforum Perspect. 2013, 12, pp. 47–56.

Koski, H. The Impact of Open Data—A Preliminary Study; The Research Institute of the Finnish Economy (ETLA): Helsinki, Finland, 2015; p. 27.

Kulk, S.; van Loenen, B. Brave new open data world? Int. J. Spat. Data Infrastruct. Res. 2012, 7, pp. 196–206.

Liander. Missie en Strategie; Available online: https://www.liander.nl/over-liander/wat-doet-liander/missie-en-strategie (accessed on 29 March 2016).

Liander. Voorbeelden Open Data. Available online: https://www.liander.nl/over-liander/innovatie/open-data/apps (accessed on 29 March 2016).
Martin, S.; Foulonneau, M.; Turki, S.; Ihadjadene, M. Risk analysis to overcome barriers to open data. e-J. e-Gov. 2013, 11, pp. 348–359.
McKinsey Global Institute. Open Data: Unlocking Innovation and Performance with Liquid Information; McKinsey Global Institute: New York, NY, USA, 2013.
Ministerie van Binnenlandse Zaken en Koninkrijkrelaties. Nationale Open Data Agenda 2016 (NODA). Ministerie van Binnenlandse Zaken en Koninkrijkrelaties: Den Haag, The Netherlands, 2015; p. 7.
Ministerie van Economische Zaken; Landbouw en Innovatie. Digitale Agenda.nl—ICT voor Innovatie en Economische Groei; Ministerie van Economische Zaken: Den Haag, The Netherlands, 2011; p. 44.
Ministerie van Infrastructuur en Milieu. Beslisnota Open Data: Roadmap; Ministerie van Infrastructuur en Milieu: Den Haag, The Netherlands, 2012; p. 13.
North, D.C. Institutions, Institutional Change and Economic Performance; Cambridge University Press: Cambridge, UK, 1990; p. 159.
Obama, B. Transparency and Open Government: Memorandum for the Heads of Executive Departments and Agencies; The White House: Washington, WC, USA, 2009.
Omidyar Network. Open for Business: How Open Data Can Help Achieve the G20 Growth Target. A Lateral Economics Report Commissioned by Omidyar Network; Omidyar Network: Redwood City, CA, USA, 2014; p. 84.
Open Government Partnership. Open Government Declaration. Available online: http://www.opengovpartnership.org/about/open-government-declaration (accessed on 29 March 2016).
Organisation for Economic Co-Operation and Development (OECD). OECD Principles and Guidelines for Access to Research Data from Public Funding; OECD: Paris, France, 2007.
Pollock, R. Welfare Gains from Opening up Public Sector Information in the UK; University of Cambridge: Cambridge, UK, 2011; p. 4.
Poplin, A. Methodology for measuring the demand geo-information transaction costs: Based on experiments in Berlin, Vienna and Zurich. Int. J. Spat. Data Infrastruct. Res. 2010, 5, pp. 168–193.
Sholtz, P. Transaction costs and the social costs of online privacy. First Monday 2001, 6, doi:10.5210/fm.v6i5.859.
Sunlight Foundation. Ten Principles for Opening Up Government Information. Available online: https://sunlightfoundation.com/policy/documents/ten-open-data-principles/ (accessed on 31 August 2010).
The Committee on Data for Science and Technology (CODATA). The Value of Open Data Sharing: A CODATA Report for the Group on Earth Observations; CODATA: Paris, France, 2015; p. 43.
Uhlir, P.F. The Socioeconomic Effects of Public Sector Information on Digital Networks. Towards a Better Understanding of Different Access and Reuse Policies—Workshop Summary; National Academies Press: Washington, DC, USA, 2009; p. 105.
United Nations. Adoption of the Paris Agreement; United Nations: New York, NY, USA, 2015; p. 12.
U.S. Department of Health & Human Services. Unleashing the Power of Data and Innovation To Improve Health. Available online: http://www.healthdata.gov/content/about (accessed on 29 March 2016).
van Damme, E. Pragmatic privatization: The Netherlands 1982–2002. In Privatization Experiences in the European Union; Köthenbürger, M., Sinn, H.-W., Whalley, J., Eds.; CES-IFO with MIT Press: Cambridge, MA, USA, 2006; pp. 289–337.
van Loenen, B.; Crompvoets, J.; Poplin, A. Assessing geoportals from a user perspective. In Proceedings of the 2nd Workshop on Value of Geoinformation, Hamburg, Germany, 30 September–2 October 2010; pp. 29–38.
Vickery, G. Review of Recent Studies on Psi Re-use and Related Market Developments; Information Economics: Paris, France, 2011; p. 44.
Williamson, O.E. The Economic Institutions of Capitalism. Firms, Markets, Relational Contracting; The Free Press: New York, NY, USA, 1985; p. 468.
Williamson, O.E.; Masten, S.E. Transaction Cost Economics, Theory and Concepts; Edward Elgar Publishing Limited: Cheltenham, UK, 1995.
WISE Institute. Big and Open Data in Europe. A Growth Engine or a Missed Opportunity? Warsaw Institute for Economic Studies: Warsaw, Poland, 2014; p. 116.

8 How to assess the success of the open data ecosystem?

Frederika Welle Donker and Bastiaan van Loenen

International Journal of Digital Earth 2016, Vol. 9, http://dx.doi.org/10.1080/17538947.2016.1224938. Submitted 14 March 2016; accepted 11 August 2016; published online 9 September 2016.

Abstract

Open data are currently a hot topic and are associated with realising ambitions such as a more transparent and efficient government, solving societal problems and increasing economic value. To describe and monitor the state of open data in countries and organisations, several open data assessment frameworks were developed. Despite high scores in these assessment frameworks, the actual (re)use of open government data fails to live up to its expectations. Our review of existing open data assessment frameworks reveals that these only cover parts of the open data ecosystem. We have developed a framework, which assesses open data supply, open data governance and open data user characteristics holistically. This holistic open data framework assesses the maturity of the open data ecosystem and proves to be a useful tool to indicate which aspects of the open data ecosystem are successful and which aspects require attention. Our initial assessment in the Netherlands indicates that the traditional geographical data perform significantly better than non-geographical data, such as healthcare data. Therefore, open geographical data policies in the Netherlands may provide useful cues for other open government data strategies.

Keywords: open data; geodata; assessment framework; open data governance; open data maturity

§ 8.1 Introduction

Open data are currently a hot topic. Around 2009, open government data initiatives started to emerge with *e.g.* the 2009 Obama Executive Order[52], the 2010 Digital Agenda of the European Commission[53], the 2011 Open Government Partnership Initiative[54], and the 2013 G8 Open Data Charter[55]. Open government data (OGD) are associated with realizing ambitions, such as a more transparent and efficient government (*e.g.* Huijboom and van den Broek, 2011), reducing corruption (*e.g.* Granickas, 2014; David-Barrett, Heywood and Theodorakis, 2015); improving citizens' participation (Jetzek, 2013), solving societal problems (*e.g.* Uhlir, 2009; Attard *et al.*, 2015), increasing economic value due to companies creating innovative products and services with open data as a resource (*e.g.* Vickery, 2011; Omidyar Network, 2014) and efficiency improvements (*e.g.* WISE Institute, 2014; McKinsey Global Institute, 2013).

Open data should comply with ten principles formulated in 2010. Government data shall be considered open if they are complete, primary, timely, accessible, machine processable, non-discriminatory, non-proprietary, permanent, licence-free and preferably free of charge (Sunlight Foundation, 2010). Open data are not limited to government data as the private sector also recognises the potential benefits of making their datasets available as open data (Welle Donker, van Loenen and Bregt, 2016).

For this article, we consider open data to be all data that can be reused without financial and legal restrictions, including data available with a licence requiring attribution, *e.g.* a Creative Commons Attribution (CC-BY)[56] licence.

52 http://www.whitehouse.gov/open/documents/open-government-directive

53 http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52010DC0245R(01)&from=EN

54 http://www.opengovpartnership.org/

55 https://www.gov.uk/government/publications/open-data-charter

56 http://creativecommons.org/licenses/

§ 8.1.1 Open data benefits yet to materialise

Open data initiatives in Europe were initially driven by the potential transparency and economic benefits (European Commission, 2011). However, in spite of more OGD made available, the predicted effects appear not to have materialised to date. Although a literature review by the authors demonstrated that there is ample anecdotal evidence of a positive impact of open data (*e.g.* Vickery, 2011; de Vries *et al.,* 2011), in practice it is difficult to measure the actual socio-economic impact of open data (Koski, 2015).

Research indicates that more OGD does not automatically lead to more transparency and increased trust in government (*e.g.* Gurstein, 2011; Grimmelikhuijsen, 2012; dos Santos Brito *et al.,* 2015) or to a surge of value added products and services based on OGD (*e.g.* Rhind, 2014). This may be due to a number of reasons, *e.g.* a mismatch between the datasets supplied and the actual dataset demand (IRM, 2015), a lack of cooperation by government agencies (Peled, 2011) or not enough care is taken when publishing datasets (Janssen, Charalabidis and Zuiderwijk, 2012). If governments cannot see a positive impact of open data, high-level political commitment may reduce and open data programmes may stall or even go backwards (World Wide Web Foundation, 2015). Therefore, it is vital that a positive impact of open data is demonstrated. However, before we can assess the success of the impact of open data, we need to assess the current state of open data as a benchmark.

§ 8.1.2 Assessment of open data initiatives

To improve the uptake of open data and successful embedding in society, an assessment and evaluation of the maturity of open data is a useful tool. Assessment frameworks are used for different reasons, such as benchmarking and comparing between different sectors and between countries, or to provide tools to improve the quality and governance of open data. Although already a number of open data assessment frameworks have been developed around the world, these models tend to focus on only one perspective of the open data ecosystem. To determine and assess the success factors of open data requires a holistic approach. Therefore, we have developed a holistic assessment framework to assess and to evaluate open data initiatives from multiple perspectives.

Our research methodology consisted of a literature study and interviews with users, providers, and open data policy makers. We used a literature review on open data assessment frameworks to draft the first model. During the interviews, users and providers were asked to reflect on the first draft and to assess the applicability to

their situation. The resulting framework was fine-tuned and applied to the open data ecosystem in the Netherlands.

In this article, we apply our holistic open data assessment framework to provide a snapshot of the Dutch open data 'State of the Nation'. In Section 8.2, we describe assessment framework theory and provide an overview of six open data assessment frameworks. Section 8.3 describes the holistic framework components of data accessibility, data governance and user characteristics. In Section 8.4, we apply our framework to assess the data supplier's perspective of the 'State in Open Data Land' of the Netherlands. The maturity of open data governance is assessed in Section 8.5. Section 8.6 describes the user characteristics required to develop and successfully market value added products and services based on open data. Section 8.7 concludes with our analysis and recommendations for open data assessment.

§ 8.2 Open data ecosystem assessment

The key to a well-functioning open data ecosystem is accessibility from a technical, legal, and organisational perspective. Therefore, it is important that there are policies in place that define the legal context, standards to facilitate data interoperability, and a stable and sustainable network for users of the data, as illustrated in Figure 8.1. Such data ecosystems are often created by governments to facilitate access to, sharing and (re)using of government data.

To facilitate open data accessibility, governments worldwide are developing open data platforms in varying forms and functionality (*cf.* Zuiderwijk, 2015). In recent years, there has been a number of international open data assessment frameworks developed, most of which focus on implementation of open data strategies. Below, six of these assessment models are summarised and reflected upon.

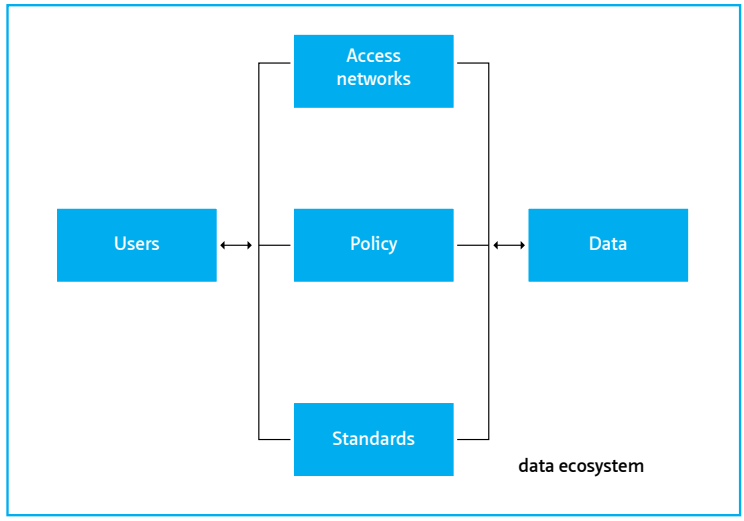


FIGURE 8.1 Key components of the open data ecosystem (after Rajabifard, Feeney and Williamson, 2002)

§ 8.2.1 Existing open data assessment frameworks

Open Knowledge International (OKI) developed a Global Open Data Index[57] to track the state of government open data, *i.e.* which countries are publishing data in the right way and in a timely way. In 2014, 97 countries were included in the index of 10 key datasets[58], with only 11% of the dataset entries were open according to their Open Definition.[59] The Index does not provide an insight into the quality of the data, however.

57 http://index.okfn.org/

58 These datasets are: Election Results; Company Register; National Map (1:250,000 or better); Government Spending; Government Budget; Legislation; National Statistical Office Data; National Postcode Data; Public Transport Timetables; and Pollutant Emissions.

59 "Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)." (http://opendefinition.org/)

The World Wide Web Foundation developed the Open Data Barometer[60] to provide a snapshot of Open Government Data practices. The Barometer focusses on open data readiness, implementation, and emerging impacts. The second edition of the Open Data Barometer assessed these aspects for a sample of 86 countries. The Open Data Barometer found that countries having open data initiatives that receive both senior-level government backing and sustained resources are much more likely to achieve impact. Only a low percentage of the countries included in the Barometer publish open data related to government transparency and accountability. Just over 10% of the surveyed datasets conformed with their open data criteria (published in bulk, machine-readable formats and under an open licence) (World Wide Web Foundation, 2015). The Barometer provides an insight into the maturity of open data governance from a data provider's perspective only.

Independent Reporting Mechanism (IRM) developed a tagging framework to assess the extent to which the Open Government Partnership (OGP) commitment addresses both supply and demand for open data in their action plans. Their framework used 26 tags grouped into three clusters: Data Supply/Infrastructure, Environment/Context (legal and institutional conditions), and Data Use. Of the 92 OGP countries assessed, IRM (2015) found that 53% of OGP commitments relate to Data Supply, 21% to Context (legal and institutional conditions) and 26% to Data Use. IRM concluded that there appears to be a misalignment between providers publishing "low hanging fruit" and users wanting high-value data (see also Davies, 2014). IRM assessed the governance of open data initiatives are being carried out but IRM did not include the user's perspective.

The Public Sector Information (PSI) Scoreboard is a 'crowd-sourced' tool to measure the status of open data and PSI reuse throughout the EU. The PSI Scoreboard measures seven aspects of PSI reuse, based on 25 indicators. The PSI Scoreboard focusses on the EU PSI Directive[61] implementation and other aspects, such as availability of local government data and events organised to promote open data. The Scoreboard does not include other governance aspects or the user's perspective.

The United Kingdom Open Data Institute (ODI) developed a Maturity Framework to assess how well an organisation publishes and consumes open data. The model consists of 15 organisational activities grouped into five themes, and five progress levels to assess and monitor organisational performance (Dodds and Newman, 2015).

60 http://opendatabarometer.org/

61 Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information.

The model focusses on organisational processes and data governance from a data-provider perspective.

The Capgemini Consulting Open Data Benchmark researched 23 open data portals in the EU and found that only 22% of countries shared data that can be classified as comprehensive. Almost all (96%) countries share data which are not regularly updated; over 60% of the countries lacked enhanced search capabilities; and 87% of the countries are not utilising user participation capabilities (Tinholt, 2013). Most countries (87%) have an open data portal but only 33% of the portals support feedback mechanisms for users to give their opinion and only 11% offer a contribution feature (Capgemini Consulting, 2015). The Benchmark assessed data availability, political leadership and data portal usability from the data provider's perspective.

§ 8.2.2 Summary existing open data assessment frameworks

These frameworks assess open data from a specific perspective, such as releasing data conform an open data definition, the type of data released, adhering to open data policy commitments or open data portal performance. Some of these frameworks assess a specific outcome, such as government transparency. Although these assessment frameworks provide interesting insights, they all focus on data supply and data environment, see Figure 8-2. Even IRM (2015) only considered what data providers had done to facilitate users but had not actually consulted users. Thus, the user's perspective appears to be missing in all these frameworks.

To include the user's perspective, we have developed a new multi-dimensional holistic assessment framework that builds on a variety of existing frameworks. Our framework not only adds the user characteristics to the existing frameworks, but also provides a holistic comprehensive approach to open data assessment building on existing frameworks, which only deal with single components of the open data ecosystem. Our holistic approach reuses elements of the existing frameworks. For example, our framework includes access through a portal (part of CapGemini's framework) as part of the indicator "recognisable", and the openness aspect of a dataset of the OKI framework, and some of the parts of the maturity framework of ODI. We do this, however, from a user perspective. A user needs to know that a dataset exists and where it can be accessed. This knowledge can be provided through a data portal, but also through a general search engine. Therefore, we do no limit ourselves to portal assessment but also include other relevant aspect for this specific indicator "recognisable". In the next section, we will describe the new framework.

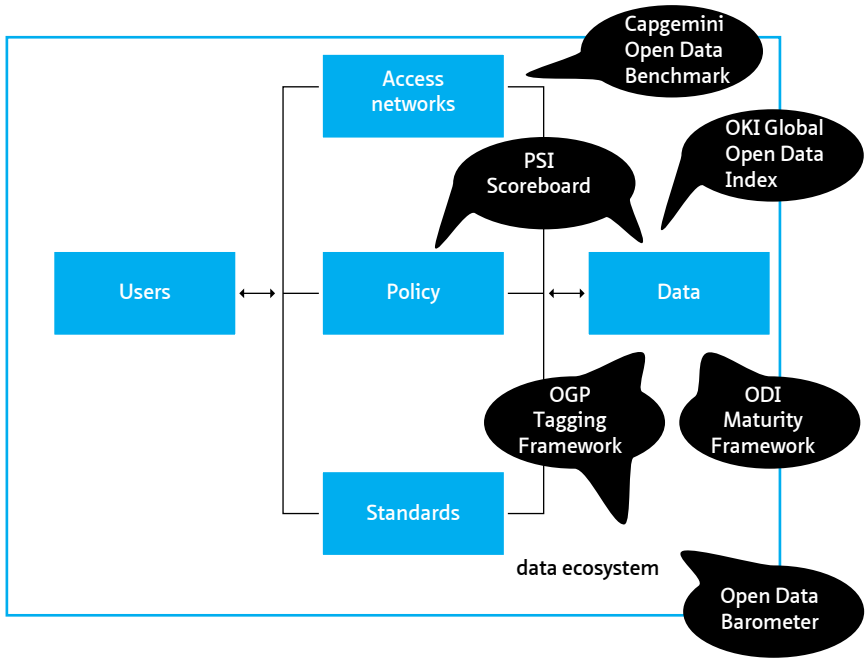


FIGURE 8.2 Focus of existing open data assessment frameworks

§ 8.3 Open data maturity assessment framework

There are several ways to assess the effect of a policy regulation. A commonly used method is to develop an assessment framework using indicators, whereby it is important that the indicators reflect the organisation's mission and core activities. We distinguish four elements: activity (action of an organisation), output (products/services of an organisation), outcome (results of an action), and impact (the way in which an outcome contributes to a strategic goal of the organisation) (Environment Canada, 2000). For instance, a government agency releases the national roads dataset as open data (activity), which results in open road data (output). A company uses the dataset to improve a car navigation system (outcome) thus, enabling drivers to avoid roads under repair and make more effective use of the roads infrastructure (impacts) (see also Helbig *et al.*, 2012).

For our research, we use three output indicators as conditions for a successful open data ecosystem, namely:

1 Data supply: The way in which data are provided as open data;
2 Data governance: The way in which governance aspects are organised;
3 User characteristics: The way in which the user characteristics enable the user to innovate with open data.

In this section, we will describe these three components.

§ 8.3.1 Open data supply indicators

The concentric shell model of Backx (2003) illustrates the open data supply from a user perspective, (see Figure 8.3). This model provides a good insight into the steps a user has to follow to assess if data may be suitable for his requirements (van Loenen and Grothe, 2014). The data should be:

1 known to the user (are the data identifiable and where can data be obtained?);
2 attainable by the user (can the user obtain the data, and under what conditions?);
3 usable for the intended purpose of the user (can the user assess the quality of the data)?

For a user to be able to reuse data, these three conditions must be satisfied.

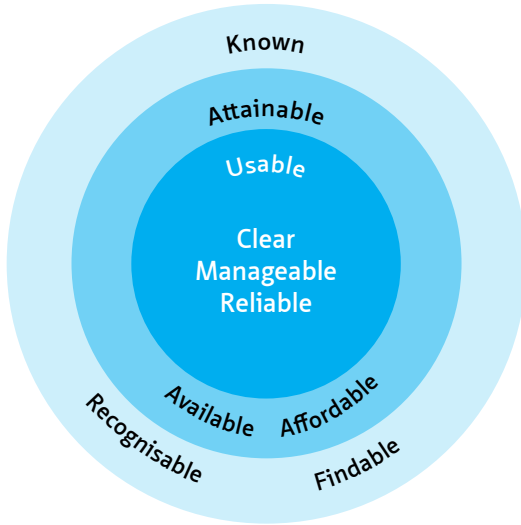


FIGURE 8.3 Concentric shell model. (Source: Backx, 2003)

Appendix A details the data supply indicators. Below, we provide the main characteristics of each shell.

§ 8.3.1.1 Known

The user has to know that a certain dataset exists: the user has to be able to recognise, *i.e.* identify the dataset. This can be achieved through resource metadata, *e.g.* resource titles or abstracts, through tags, *e.g.* internet bookmarks or textual keywords, or, for linked data, Resource Description Framework (RDF) resources.

In addition, a user has to know where to find the dataset. A user may either use a search engine or go to a data portal. If an open dataset is published but this is not clear to the public and cannot be found through a simple search, then the data can easily be overlooked and not put to good use (Open Knowledge International, 2014). The chance that data are discovered may increase if the data are published in a well-known and accessible portal. Government information portals have been around for several decades, however, these are often poorly stocked, obsolete, and particularly user-unfriendly (*cf.* van Loenen, Crompvoets and Poplin, 2010).

§ 8.3.1.2 Attainable

Once a dataset is found, it has to be attainable, *i.e.* a user has to be able to physically access the dataset (to view and/or to download it via services, or on request), to be allowed to (re)use the data (licences), and to be able to afford the data (fees). Unclear licence conditions, especially when combining multiple datasets, and high up-front fees may form a barrier for potential users (*cf.* Fornefeld *et al.*, 2008).

§ 8.3.1.3 Usable

A user will only be able to assess if the data are suitable to his/her needs after the data can be (physically) inspected. Aspects within this shell relate to data quality, *e.g.* available data formats, available documentation/metadata, level of coverage, timeliness and update frequency. Other key aspects are the presence of a helpdesk or forum for questions related to the data and guarantees for continuous availability of the data. Without such guarantees, a user may be hesitant to invest precious resources into developing a derivative product.

§ 8.3.2 Data governance

In addition, open data governance is relevant for facilitating open data use. We consider governance to be the interaction between public sector entities and/or private sector entities with the ultimate goal to realise common goals (Termeer *et al.*, 2011). Governance is a framework of policies, processes, and instruments that structure this interaction in order to enable parties to reach their common goals. Governance of open data not only provides a framework to facilitate the shells of Backx's model but also establishes who will assist the user when he/she stumbles over one of the shells. For the governance part of our open data assessment framework we use the five elements for assessing the governance of geographical information infrastructures identified by Kok and van Loenen (2005). Although this model was developed to assess the maturity of geographical information infrastructures, it can equally be applied to open data ecosystems. The aspects that help to determine the functionality of a data infrastructure are Vision, Leadership, Communication, Self-organising ability, and Long-term financing, (see Figure 8.4). In Section 8.5, we explain these aspects in detail.

FIGURE 8.4 Aspects of government data infrastructure governance

In addition, there are other important aspects, such as legal and policy frameworks (*e.g.* a right to (re)use public sector information, a right of redress to reinforce good governance values (Brewer, 2007), and a clear demarcation between public tasks and private sector activities (Janssen, Crompvoets and Dumortier, 2011)).

§ 8.3.3 User characteristics

Having data supply and governance ticked off, does not automatically mean that data will be re-used. Our assessment framework distinguishes itself by not only assessing open data readiness but also including the user's perspective, as "users will probably be the most mentioned group and yet actually the least considered" (McLaughlin and Nichols, 1994, p.72).

Next to data accessibility and governance, there are other factors that will enable the re-use of open data, such as technical connectivity, user capabilities and resources (*e.g.* Jetzek, Avital and Bjørn-Andersen, 2014; OECD, 2011). However, it may be that the user cannot or will not use the data, does not have enough technical knowhow and/or creative skills to transform the data, does not have access to sufficient capital or other resources, may not want to invest in a risky open data product, or be unfamiliar with the opportunities (*e.g.* Janssen *et al.*, 2012; Gurstein, 2011; McClean, 2011). These aspects, directly related to the user characteristics and his environment, are categorised as user characteristics.

People use open data for a number of reasons: maybe for personal reasons to address a certain (societal) issue or to fill a specific niche, or to experiment with data. However, to mature from hobbyist to developing a sustainable business model, requires more than just a good idea. Apart from being in touch with societal issues, one has to have knowledge of the supply market and of the needs of the end-market (*cf.* Osterwalder and Pigneur, 2010). To develop a marketable product or service, a right marketing mix of the right product sold at the right price at the right place using suitable promotion is required (Business Case Studies, 2016). As open data are available to everybody, everybody could theoretically create similar derivative products. The challenge is to develop a product or service that stands out from the crowd.

The characteristics someone should have to be innovative vary. You need a question or a problem that needs solving. This may stem from one's own motivation (*what if I*) or from a broader societal aspect (*what if we*). Therefore, a user should be in close touch with societal issues, as well as having good domain knowledge. As one of the interviewed users stated: "It is pointless to develop a multi-modal route planner

without intrinsic knowledge of the local infrastructure and bottlenecks if there is already a well-functioning multimodal route planner on the market."

Figure 8.5 shows the links between the elements of our assessment framework. The outcomes of the Governance model (data governance) and the Data Accessibility model (data supply) become inputs – next to other user characteristics – required for successful reuse of government open data. The impact of open data reuse could be measured using "traditional" outcome indicators, such as company turnover.

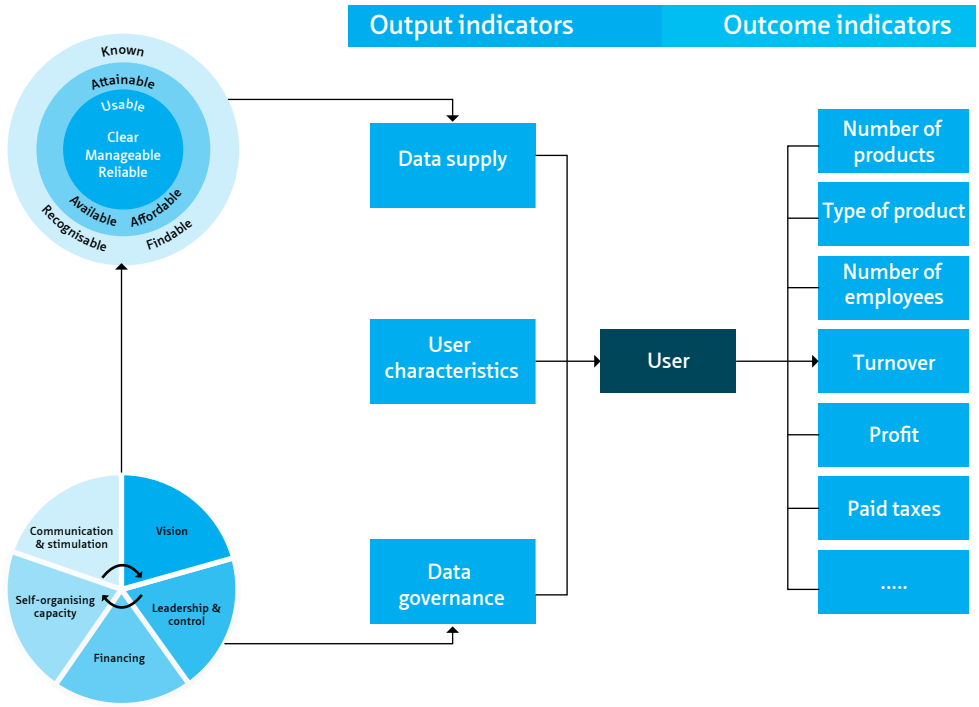


FIGURE 8.5 Output and outcome indicators of the holistic open data assessment framework

§ 8.4 Application of the framework to open data in The Netherlands: Supply

Using the indicators identified in Section 8.3, we assessed the maturity level of data supply by using a scale of 1 to 5, 1 being the lowest score.

In this section, we will apply the developed framework to the Dutch open data ecosystem. We will do this for each part of the ecosystem: the data supply, the data governance and the user characteristics. For each indicator we provide how the assessment was performed.

§ 8.4.1 Indicators for "Known"

To assess the first sub-indicator "Known", we used a profile-free search engine (https://ixquick.com/) to avoid the search engine adapting its behaviour to the used search terms.

§ 8.4.1.1 Recognisable

To assess if a dataset is recognisable, *i.e.* identifiable, we used a generic search term, *e.g.* "elevation data". If that did not resulted in a hit, we subsequently used the official name of the dataset, *e.g.* "Actual Heights Model Netherlands" and finally the official acronym, *e.g.* "AHN". A score of 1 indicates that the dataset was either not published or not identifiable; a score of 5 indicates that using a general search term provided the dataset as first hit.

§ 8.4.1.2 Findable

To assess if the dataset could actually be found, we used the official open government data portal data.overheid.nl (data.gov.nl) as well as the National Geodata Register (NGR) and, if applicable, the data provider's website, again using varying search terms. A score of 1 indicates that the dataset could not be located; a score of 5 indicated that the dataset could be located via ixquick.com (2), the data provider (3), as well as via NGR (4) and data.overheid.nl (5).

§ 8.4.2 Indicators for "Attainable"

To assess if a dataset is attainable from a financial, legal and practical aspect, we have used sub-indicators for finances (are tariffs, if applicable, published online?), licences (online, standardised licence) and service level (active/passive publication, type of data service, *e.g.* viewing/download/Application Programming Interface (API)) and delivery time if a dataset is not available online.

§ 8.4.3 Indicators for "Usable"

There are many sub-indicators to assess the usability of a dataset. Below, we describe the selected sub-indicators.

§ 8.4.3.1 Reliable

To assess the reliability of a dataset, a user should be able to check the quality of the data. As sub-indicators, we have checked the presence of metadata, their comprehensiveness and standardisation and if metadata are available in more than one language. A score of 1 indicates no metadata or documentation; a 5 indicates complete and standardised metadata.

In addition, we considered if the dataset is published in a reliable way, *i.e.* the data should not produce dead links, be available in the long term, and not be removed without a warning in advance to the users.

§ 8.4.3.2 Clear

Not all users have sufficient expertise to (re)use data (*cf.* Janssen *et al.*, 2012; Gurstein, 2011). To assess if it is clear to the user how to use the dataset, we have researched if additional documentation, such as (multi-lingual) manuals and a FAQ platform are available.

§ 8.4.3.3 Manageable

A user should be able to use the data with available resources and for the goal the user has in mind. As there is a large variation in different user needs (Bemelmans 1994, p.186), see Table 8.1, we could not develop an exact indicator to assess the manageability of the dataset. Instead, we quantified the manageability with a score of 1 for datasets published without options (only one version and format), and 5 to indicate more than three options.

USER NEEDS	RANGE
1 Coverage of required area	local <--> global
2 Actuality of data	historical <--> real-time
3 Data	thematic <--> Geodata Top 20
4 Aggregation level	1:1,000 <--> 1:10,000,000
5 Data formats	choice between propriety <--> open
6 Type of data	static <--> dynamic
7 Data service level	viewing – download – API
8 Dataset size	kilobytes <--> terabytes
9 Completeness of data	only most recent version <--> time series
10 Data consistency	consistent formats, location (URLs <--> URIs), etc.
11 Metadata	standardised and complete
12 Language and semantics	only in Dutch <--> multi-lingual

TABLE 8.1 Variation in the user needs

§ 8.4.3.4 Communication

For this indicator, we only researched if there is a helpdesk facility available with the data provider. The scope of our desk research did not extend to checking the response time and the level of knowledge of helpdesk staff, therefore, this indicator is only included in the user's perception part of our framework.

§ 8.4.3.5 Up-to-date data

We researched the actuality and the update frequency of the dataset. A score of 1 indicates that a version was published once off and never updated; a sore of 5 indicates that the most recent version (near real-time) is timely published.

§ 8.4.3.6 Long-term availability of data

We researched if a legal or policy commitment is published guaranteeing continuous availability of the dataset, for updates as well as historical versions. In addition, we assessed the technical sustainability of data availability, *i.e.* at which level are the data published. We used the five-star model of Tim Berners-Lee (TBL)[62] whereby a score of 1 indicates that a dataset is published with an open licence, but not in a structured or open format, a score of 2 means that the data is available as machine readable structured data, a score of 3 implies that the dataset has also a non-proprietary format. A score of 4 stands for dataset using open standards from W3C and a score of 5 indicates that the dataset is published as linked open data.

§ 8.4.4 The assessment framework applied to Top 20 most wanted datasets

To assess the open data supply part of our framework, we researched twenty datasets in the Netherlands.[63]

The datasets were selected by using a "Top 20 Most Wanted datasets" originally compiled by GeoBusiness Netherlands, an umbrella organisation for geographic information companies, in 2007 (ref. Groot *et al.*, 2007) and updated in 2014 by GeoBusiness Netherlands and by the interviewed users. Whereas the 2007 Top 20 contained mostly geographical data (geodata), the 2014 version reflected a desire for other data, including healthcare data and energy data: a trend also reported by the OECD (Ubaldi, 2013).[64] Our desk research of the Top 20 Most wanted resulted in 27 assessed datasets, of which 19 are managed by national government bodies, three by municipalities, and five by non-government organisations (NGOs). Seventeen datasets were publishes as open data, six as non-open data and four datasets were not accessible at all.

62 See http://5stardata.info/en/ for an explanation.

63 The 2014 "Top 20 Most Wanted" datasets were: Key Registration Topography (1:10,000), Company Register, Statistical information related to local areas, Key Registration Large Scale Base Map, municipal information, aerial photography, Key Registration Addresses & Buildings, cadastral information, energy usage data, energy labels of dwellings, soil information, national railway data, national roads data, real-time traffic information, spatial planning, digital elevation map, national waterways data, water levels (real-time), health risk areas, and healthcare information.

64 See also the European Commission which ranked these datasets as the highest priority for being made available for reuse due to the high demand from reusers across the EU (see European Commission, 2014; see also The Cabinet Office, 2013).

In addition to desk research, we used the experiences of open data users to assess the data supply. We selected our interviewees from a diverse group of users with diverse backgrounds. The interviewees represented companies of varying sizes (from one-person start-ups to large companies; from fulltime professionals to active amateurs; from geographical information / IT specialists to non-experts) and requiring open data for various purposes (value added services, information intermediary, consultancy, civil activism). In total, we interviewed nine users using structured interviews with semi-closed questions. We asked the users through open questions to reflect on the draft assessment framework and to apply the framework to their specific situation. In Table 8.2, we provide the resulting scores per category.

	NATIONAL GEODATA	NATIONAL NON-GEODATA	MUNICIPAL DATA	OPEN DATA*	FEE-BASED DATA*	NGO DATA
Recognisable	3.85	2.75	1.67	3.29	4.13	3.00
Findable	4.45	2.75	2.33	4.53	3.38	3.00
Known	**4.15**	**2.75**	**2.00**	**3.68**	**3.75**	**3.00**
Affordable	4.20	1.25	2.00	4.65	2.00	1.60
Licences	3.80	1.00	2.33	4.47	1.38	1.20
Service level	3.80	1.00	2.33	3.94	2.50	1.40
Delivery time	4.35	1.25	2.00	4.65	2.38	1.80
Attainable	**4.04**	**2.25**	**3.25**	**4.18**	**2.75**	**2.50**
Reliability	2.55	0.75	1.33	2.71	1.50	1.00
Clear	2.80	0.75	1.33	2.82	1.88	1.20
Manageable	3.40	0.50	1.00	3.18	2.38	1.00
Up-to-date	3.30	0.50	1.00	3.12	2.25	0.80
Continuity	3.20	0.25	1.67	2.94	2.50	0.60
TBL score	1.95	0.75	0.33	2.53	0.00	0.60
Usable	**3.05**	**1.10**	**1.90**	**2.79**	**2.80**	**1.53**
Average score for Known, Attainable and Usable	**3.75**	**2.31**	**2.03**	**3.46**	**3.13**	**2.47**

TABLE 8.2 Data supply scores (scale: 1 (low) - 5 (high))

* Two datasets were excluded from the categories "open data" and "fee-based data" as their status was unknown.

The table demonstrates that geodata score higher than non-geo data, that open data score better than fee-based data, and that national data score higher than municipal data and NGO data.

§ 8.4.4.1 Known

The researched datasets scored 3.8 overall for "Known". As seen in Table 8.2, national datasets already available in the traditional geographical information (geodata) domains scored better than non-geodata (healthcare and energy) for being recognisable and findable. Most of the researched datasets were recognisable but not findable unless the correct acronym was used or the data holder was known. Data that had only recently been available as OGP provided mainly search engine hits for private sector information services rather than links to the public data source. Municipal data scored only on average 2.0 as not all researched datasets could be found.

We found that using general search terms in the data portals data.overheid.nl and NGR resulted in non-related data and/or the search facility took a long time. Moreover, the researched local government open datasets were not registered in data.overheid.nl, and only one out of eight energy network administrators published their energy usage data as open data. The researched municipal websites offered even poorer search facilities than the national data portals. Our desk research findings were confirmed by the user interviews.

From interviews, it emerged that users use various strategies to find data: general search engines, data catalogues, social media, and professional networks were all named as strategies. Users will contact the data holder directly if the data holder is known rather than using a data portal link. However, users indicated that it is hard to find out which government organisation holds which datasets. Especially municipal data are difficult to locate. Often, data holders cite protection of personal data as a reason for not publishing data. However, most users perceive this to be a fallacy because any personal data can be aggregated, anonymised, or removed in the end-product.

§ 8.4.4.2 Attainable

The attainability of the researched datasets scored 3.9 overall, however, there were some points of concern.

Licences

Seventeen datasets were available as open data, however, two of which were published without a licence and three with a licence limiting re-use. Only one energy dataset was published with an open licence, the other two energy datasets were not public, as were the locations of healthcare providers. Not knowing which conditions apply creates uncertainty as not all open licences are equal (Van Loenen, Janssen and Welle Donkekr,

2012). The interviewed users confirmed they were hesitant to combine open datasets because of the uncertainty what can and cannot be done with the end-product. For instance, some health data may not be reused for commercial products but the intended end-product will be a free app. In one case, the licence conditions were hidden in a disclaimer.

Fees

We found that for one open dataset administration fees were applicable. Although these fees are only marginal, some interviewed users, mainly start-ups and activists, indicated that any charges pose a barrier to their use. Others, often professional users, indicated that paying a fee was not a barrier as long as the fee was in proportional to the business case of the end-product. For fee-based data, often tariffs (fees per unit, object, km^2) are published online but no tariff for the entire (land-covering) dataset.

Services

For open geodata, we found that in most cases viewing services and download services were available and some data available via APIs, resulting in a score of 3.8. However, we found that for many geo-datasets clicking the download button of data.overheid.nl could result in an error message if one did not have appropriate software, as many datasets are linked to geo-web services of the PDOK Portal[65]. Similar research found that 14 per cent of all datasets released via the Dutch open data portal are not accessible because of broken links (Open State Foundation 2015). Non-geodata and municipal data scored respectively 1.0 and 2.3 with often no download services or APIs available at all.

Delivery time

Open data are often downloadable directly. Fee-based data scored 2.4 as the time to respond to a data request varied from 1 to 5 working days. For data that are not published, a user has to make a formal request according to the Public Information Act procedures. This procedure can legally take up to 8 weeks, although appeal cases have been known to take years. Most users indicated that any delivery time of over 5 days is too long.

65 PDOK (Public Services on the Map) was established as national geographical data portal for viewing, invoking and downloading services, part of the INSPIRE Directive 2007/2/EC requirements. Although primarily established for the public sector, anyone may view geodata and download if data are available as open data.

§ 8.4.4.3 Usable

The usability of the researched datasets scored 2.9 overall. Our desk research resulted in significant differences between the usability of traditional geodata and non-geodata.

Reliability and clarity

In general, we found only limited metadata (if any) available and often only in Dutch. Most data suppliers provide additional documentation online but only in Dutch. Because of limited metadata, users find it difficult to check the quality of the data; however, this is alleviated to some extent by additional documentation. Apart from incomplete metadata, users perceived problems with no metadata updates, metadata not machine-readable or not describing the data content.

Manageable

For most geodata there are multiple web services and versions available (*e.g.* area selection, different formats), with most often, two or three options available. For some open data, there is a limit to the maximum number of objects that may be downloaded ('fair use policy'). The researched healthcare data were published as an 'as-is' data dump without options and only available as viewing service. Viewing services developed by local governments or NGOs lacked user-friendliness and speed as such services use an open source interface developed some time ago, whereas most users are more familiar with "Google-like" interfaces. Users perceive the level of detail not always to be sufficient or experience gaps in land-covering data.

Up-to-date

We found that for all researched datasets, a recent version was published, although the interviewed users indicated that often, the most recent version is not timely published. For some datasets, *e.g.* aerial photography, historical versions were also available. We could not find any online commitment to ensure the (long-term) availability of the researched datasets, although the interviewed users assume this is the case for key register data. For fee-based data, there is often a contractual clause pertaining to data availability.

Sustainable publication

Most of the researched open datasets were published in a structured format, although not always in an open format. Some of the open data, *e.g.* health data tables, are only published in PDF format. Fee-based data are often available in an open format as well as propriety format. Thus, the researched data scored either 0, 1 or 3 stars in the TBL model. A number of users indicated that open geo-formats (*e.g.* GML or XML)

were useful, whereas other users indicated that these formats were too complicated and preferred either a general open format (CSV) or a proprietary format (shape files). The lack of URIs and linked open data were perceived to be a missed chance. Figure 8.6 provides a summary of our findings.

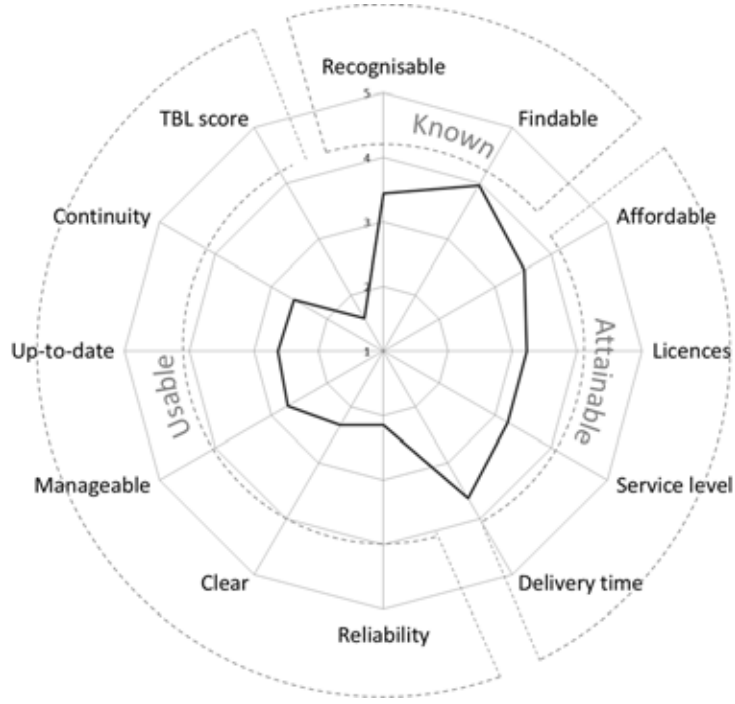


FIGURE 8.6 The Open Data State of the Netherlands in 2014 for the Top 20 Most Wanted Datasets

§ 8.5 Application of the framework to open data in the Netherlands: governance

In this section, we describe our findings of applying the governance part of our assessment framework. To assess the governance aspects of open data, we interviewed

seven OGP holders[66] using semi-closed questions related to governance and their experiences with user interaction. The interviewees were managers on operational level in charge of implementing open data policies. In addition, we asked the interviewed users what their experiences were related to communication with the government and their involvement in policy-making.

For the governance part of our assessment framework, we adapted the five elements of the maturity matrix for geographical infrastructures (*cf.* van Loenen, 2006) to determine the governance of open data provision (see Appendix B for the detailed indicators):

1. Vision: to provide a common goal, to avoid a fragmented approach and to stimulate cooperation between stakeholders.
2. Leadership and control: open data need to have a problem owner who will stimulate and coordinate open data activities. Awareness creation and capacity-building may lead to political support for open data, which is an important success factor (Craglia *et al.*, 2002), as is work floor support.
3. Communication channels: with whom, how and what is communicated. In the initial stages, this will be mostly internal communication and in later stages, also external communication.
4. Self-organising capacity: the way in which supply matches demand. In the initial stages, it will be mostly data providers requiring answers to specific questions and pro-actively promote open data. In later stages, matching supply and demand will increasingly become a part of the organisation's culture.
5. Sustainable financing: should extend beyond the initial stages and become embedded in the organisation's budget for data management and infrastructures.

§ 8.5.1 Vision

The general vision of the Dutch national government is formulated in the 'Vision Open Government and Action Plan' policy document, based on the OGP framework of "open, unless" for data that are already public. The Ministry of the Interior and Kingdom Relations (BZK) is responsible for the open data agenda. The Ministry of Infrastructure and the Environment (I&M), holder of many open datasets, has formulated a more extended open data policy for its agencies. Not only the most recent version must

66 These data holders represented the Ministry of the Interior and Kingdom Relations; Ministry of Education, Culture and Science, the Netherlands' Cadastre, Land Registry and Mapping Agency; the Province of South-Holland; the Municipality of Rotterdam; and the Water Information House, a portal for water information of the provinces, water councils and the Department of Public Works and Water Management.

be published but also a minimum of four previous versions (if applicable); once-off published datasets are to be maintained for at least five years; and a deadline is set for publishing all suitable data as open data.

We found that open data policies are firmly established within the government organisations and that there is broad political support. We also found that most government organisations follow the extended policy of the Ministry of I&M rather than the general open data policy, although each organisation had their own interpretation of "open, unless". There are variations in the decision on how to publish (pro-actively versus passively); what (all data versus only data not affecting the financial model); and which licence conditions (CC0 declaration, CC-BY licence or a non-standard "open" licence).

§ 8.5.2 Leadership

Open data are promoted for different reasons and consequently, there are vast variations in the perception of which organisation is actually providing leadership. The Ministry of BZK promotes open data from a transparency view, whereas the Ministry of I&M advocates open data to improve their data quality and more effective reuse between the agencies, and the Ministry of Economic Affairs (EZ) promotes open data to stimulate economic value-adding. Most users perceived the Ministry of I&M to provide leadership because of their extended open data policy, whereas most data holders considered their own organisation to be a leader in the open data field, or pointed to specific open data champions or open data activist. It was conspicuous that none of the other organisations viewed the Ministry of BZK as an open data leader although this Ministry is responsible for the open data agenda.

Although there is some cooperation on strategic level between government organisations with similar public tasks (*e.g.* between Provinces and Water District Boards for water management), there is almost no coordinated cooperation between ministries and municipalities.

§ 8.5.3 Self-organising capacity

To determine the self-organising capacity, we have assessed which strategies are employed to promote/stimulate open data and match supply and demand of open data. The interviewed data holders are all involved in open data stimulation

activities, such as conference presentations, organising workshops and hackathons, and offering innovation prizes. The Ministry of EZ organises so-called Open Data Relay events centred on specific themes, *e.g.* Energy or Agro-food, in cooperation with the private sector. The aim of an Open Data Relay event is to match specific questions to available data.

The employed strategies concentrate on matching open data supply to demand or to improve internal data sharing. None of the data providers mentioned the government as a launching customer – as suggested by a number of users – as a potential stimulation measure. However, this may in part be due to governments having to adhere to the legal framework for public procurement conform the EU Public Contracts directives[67], which data holders view to be complex and a barrier to outsourcing.

§ 8.5.4 Communication

Within the various government departments and agencies communication about open data takes place both on formal (ad hoc) and informal level during domain-specific meetings, via personal contacts, social media (Twitter, LinkedIn) and during open data events. Most communication concerns legal issues, best practices and exchange of experience and knowledge. A point of concern is that announcements of specific open data research commissioned by one government organisation and the ensuing results are not disseminated to other government organisations.

Communication between government and external users occurs both formally via user group meetings held on a regular basis, usually annually, with professional users of a specific dataset, and informally via personal contacts, social media, and meetings. On formal level, there is ad-hoc communication related to strategic level goals. On operational level, there is no formal communication with the exception of a few municipalities cooperating closely with the private sector. Most of the communication concerns data updates, open data best practices, and event announcements. Users indicated that they appreciated this form of communication.

67 Directive 2014/23/EU of the European Parliament and of the Council of 26 February 2014 on the award of concession contracts, Directive 2014/24/EU of the European Parliament and of the Council of 26 February 2014 on public procurement and repealing Directive 2004/18/EC, and Directive 2014/25/EU of the European Parliament and of the Council of 26 February 2014 on procurement by entities operating in the water, energy, transport and postal services sectors and repealing Directive 2004/17/EC.

§ 8.5.5 Financing

Development and implementation of open data require on-going resources. Although government organisations all faced budget cuts across the board, all interviewees indicated that open data processes are financed as part of their primary processes. However, interviewees of self-funding organisations, having to generate revenue to cover part of their operating costs, expressed their concern about the long-term/ sustainable financing for key register datasets scheduled to be published as open data in the future instead of current fee-based data.

§ 8.5.6 User perspective of open data governance

Some users (mostly large companies) use open data to provide an added level of service to their current customers. Other users/developers are still struggling to develop a sustainable open data business model. These users, often start-ups and small companies, would prefer the government to act as a launching customer. Their message to the government is to stop organising hackathons, with data only being available during the hackathon, and to stop waiting for the "killer app" to be developed. Instead, the government should commission them to develop open data tools and applications required for a successful open data ecosystem. Users feel that they are better equipped to do so as they have closer ties to end-users and actually perceive the government's current initiatives to develop open data platforms and tools, etcetera to be unfair competition. Municipalities were perceived to be re-inventing the wheel related to open data platforms rather than reusing existing knowledge. Furthermore, if there was a platform/app store on data.overheid.nl showing products based on open data, users could see what had already been developed and, thus, save precious time.

Users perceived communication, both between government organisations and with the users, to be a key success factor for open data development. Although some professional users already participate in formal user group meetings for specific datasets, users indicated they would prefer to be included in more formalised and centralised communication with the government. This could be via a national open data user group, via an open data community, although social media are preferred for help on the fly.

Figure 8.7 provides a summary of the maturity level of open data governance in the Netherlands.

FIGURE 8.7 The maturity level of Open Data governance in the Netherlands in 2014

§ 8.6 Application of the framework to open data in The Netherlands: user characteristics

Part of our research was to explore the resources and characteristics users need to create and market a successful product based on open data. As this part of the research was qualitative, we have not developed quantitative indicators to assess the maturity of the user. From the interviews held with users, we found that most users finance their own activities, sometimes aided by subsidies. Most users indicate that open data apps do not generate revenue (yet) but may serve as calling cards for made-to-order applications. Large(r) companies often use open data to improve existing products and services.

Users also indicated that, depending on the type of product, having knowledge of geographical data formats and geocoding is a prerequisite, as are general ICT skills to process database extractions and transformations, and to develop programming code and scripts.

Most users indicated that the most important characteristics are the ability to think outside the box, be creative, and, above all, have perseverance. Part of that perseverance is the ability to accept that data are often imperfect and incomplete and, therefore, a user has to work with the data that are available.

§ 8.7 Conclusions and recommendations for further research

This article presented a holistic open data assessment framework addressing the quality of open data supply, open data governance, and the user perspective of the open data infrastructure. By adding the user's perspective to our framework, a holistic comprehensive approach to open data assessment is provided. Our holistic approach reuses elements of existing open data assessment frameworks, such as access through a portal, highlighted in the CapGemini framework, the openness aspect of a dataset of the OKI framework, and some of the parts of the maturity framework of ODI. We do this, however, from a user perspective. We found that in 2014 in the Netherlands, the supply side of open data scored, on a scale of 1 (low) – 5 (high), an average 3.41 and the governance of open data on average 2.71. These scores should be used as an indication to compare the maturity of the open data ecosystem over time and not as an absolute score.

In general, open data governance is well organised in some aspects but lagging in others. Although there is an open data vision on strategic level and the concept of open data no longer a point of discussion, there is no clear leadership outside each organisation. On an operational level, government organisations are struggling to apply the "open data, unless" policy to their specific data and would benefit if one organisation took control. This organisation should provide advice and hands-on tools to other organisations to make data suitable for open data and to coordinate consistency. Many organisations currently do not publish high-value data because they lack knowledge on how to adapt sensitive data suitable for open data publication. In addition, formal and structural communication (both intra-governmental and with users) should be established to match open data supply to demand as most of the current communication occurs on an ad-hoc and informal basis.

We applied the assessment model to the Dutch open data ecosystem to evaluate the state of the open data nation and to provide valuable information on (potential) bottlenecks. The model showed that "traditional" geodata scored significantly better than other government data. It maybe that the standardisation and implementation rules laid down by INSPIRE Directive may have been a catalyst for moving geodata to a higher maturity level (see Van Loenen and Gothe, 2014). The assessment model

provided policy makers with useful inputs for further development of the open data ecosystem and well-founded strategies, to ensure the full potential of open data will be reached. Since the publication of the State of the Nation in 2014, a number of the recommendations have already been implemented.

However, the assessment framework needs to be fine-tuned and made more user-friendly. The currently defined maturity stages need to be translated into concrete questions. Our results were based on researching a limited number of datasets and on a limited number of interviews. Therefore, the outcomes of the assessment may rely on some subjectivity. Although we considered the sample to be representative, the assessment model should be applied to assess more datasets. In addition, more users from a broader target group and more data providers, especially lower governments and NGOs, should be involved to validate the model. Once fine-tuned, organisations can use the model as a self-assessment tool to monitor the state of their open data ecosystem in cooperation with the actual users.

Assessing user needs in itself is complex and especially assessing user needs in open data since there is not one single user goal. As provided in the article, open data serves many masters and it is difficult if not impossible to model these masters in a single user need indicator. Therefore, we limited the indicators for communication as well as for usability to a generic, but still informative level.

The extent to which the Dutch case adheres to the ideal is informative for the specific data holder, but indeed questionable for the state of a country since we only obtained the data of nine interviewees. A survey approach may be needed to address this issue properly. We have added this issue in the recommendations

Although the holistic framework was only applied to the Netherland, its set-up is such that it can also be applied to other context and in other countries.

A further remark is that also the presented holistic assessment framework does not explain why open data cannot live up to its expectations. The addition of the user perspective including user characteristics is relevant for this assessment, but possibly also other aspects such as a critical mass of well-equipped users is equally of importance. Further research should look into this aspect of user (group) characteristics and its role in the performance of open data ecosystems.

Acknowledgements: The research this chapter is based on was conducted as part of the Dutch ICT Breakthrough Project Open Geodata, initiated and coordinated by the Ministry of Economic Affairs in cooperation with the Ministry of Infrastructure & the Environment, the Ministry of the Interior & Kingdom Relations, GeoBusiness Netherlands, and university researchers. We would like to thank members of the project team, and the people who agreed to be interviewed for their kind cooperation.

References

Attard, J., F. Orlandi, S. Scerri and S. Auer. 2015. A systematic review of open government data initiatives. Government Information Quarterly doi: http://dx.doi.org/10.1016/j.giq.2015.07.006.

Backx, M. 2003. Gebouwen redden levens. Toegankelijkheidseisen van gebouwgegevens in het kader van de openbare orde en veiligheid. [Buildings save lives. Accessibility requirements for buildings in the context of the public order and safety]. MSc. dissertation, Delft University of Technology.

Bemelmans, T.M.A. 1994. Bestuurlijke informatiesystemen en automatisering: Kluwer Bedrijfswetenschappen.

Brewer, B. 2007. Citizen or customer? Complaints handling in the public sector. International Review of Administrative Sciences 73 (no. 4):549-556. doi: 10.1177/0020852307083457.

Business Case Studies. 2016. Marketing theory. Creating strategies that meet customer needs (the marketing mix). Accessed June 17, 2016 http://businesscasestudies.co.uk/business-theory/marketing/creating-strategies-that-meet-customer-needs-the-marketing-mix.html#axzz4BqX2uhT2.

Capgemini Consulting. 2015. Open Data Maturity in Europe 2015. Insights into the European state of play. Accessed November 20, 2015 http://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n1_-_final.pdf.

Craglia, M., A. Annoni, R.S. Smith and P. Smits. 2002. Spatial Data Infrastructures: country reports. Accessed August 31, 2003 http://www.ec-gis.org/ginie/doc/SDI_final_en.pdf.

David-Barrett, E., P.M. Heywood and N. Theodorakis. 2015. TACOD Towards a European Strategy to reduce corruption by enhancing the use of open data. National research United Kingdom: RiSSC - Research Centre on Security and Crime - IT. Accessed January 28, 2016 http://www.tacod.eu/wordpress/wp-content/uploads/2015/04/National_Research_UK_def.pdf.

Davies, T. 2014. Open data in developing countries - Emerging insights from Phase I. Berlin: The World Wide Web Foundation, Accessed July 15, 2014 http://www.opendataresearch.org/sites/default/files/publications/Phase%201%20-%20Synthesis%20-%20Full%20Report-print.pdf.

de Vries, M., L. Kapff, M. Negreiro Achiaga, P. Wauters, D. Osimo, P. Foley, K. Szkuta, J. O'Connor and D. Whitehouse. 2011. Pricing of Public Sector Information Study (POPSIS) - Summary Report. Brussels: Deloitte Consulting CVBA. Accessed January 3, 2012 https://ec.europa.eu/digital-single-market/en/news/pricing-public-sector-information-study-popsis-summary-report.

Dodds, L. and A. Newman. 2015. A guide to the Open Data Maturity Model. Assessing your open data publishing and use: Open Data Institute. Accessed March 31, 2015 http://theodi.org/guides/maturity-model.

dos Santos Brito, K., M.A. da Silva Costa, S.C. Garcia and S. Romero de Lemos Meira. 2015. Assessing the Benefits of Open Government Data: The Case of Meu Congresso Nacional in Brazilian Elections 2014. 16th Annual International Digital Government Research Conference, Phoenix, Arizona, May 27-30, 2015, doi: 10.1145/2757401.2757422.

European Commission. 2014. Report on high-value datasets from EU institutions. Brussels. Accessed September 12, 2014, http://ec.europa.eu/isa/documents/publications/report-on-high-value-datasets-from-eu-institutions_en.pdf.

European Commission. 2011. Open Data: An engine for innovation, growth and transparent governance. Brussels. Accessed June 10, 2013, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF.

Fornefeld, M., G. Boele-Keimer, S. Recher and M. Fanning. 2008. Assessment of the re-use of Public Sector Information (PSI) in the Geographic Information, Meteorological Information and Legal Information sectors - Final Report. Study commissioned by EC in 2007. Dusseldorf: MICUS. Accessed January 28, 2009 http://ec.europa.eu/newsroom/document.cfm?doc_id=1258.

Granickas, K. 2014. Open Data as a tool to fight corruption: European Public Sector Information Platform Topic Report no. 2014/04. Accessed January 29, 2016 https://ofti.org/wp-content/uploads/2014/05/221171136-Open-Data-as-a-Tool-to-Fight-Corruption.pdf.

Grimmelikhuijsen, S.G. 2012. Transparency and trust. An experimental study of online disclosure and trust in government. Ph.D. Dissertation, Utrecht University. ISBN 978-90-8570-971-2.

Groot, N., A. Bregt, J. Crompvoets and Y. Pluijmers. 2007. Toegankelijkheid publieke geodata vooral gehinderd door juridische barrières [Accessibility public geodata mostly hampered by legal barriers]. Vi Matrix, Dec. 2007-Jan. 2008, 6-8.

Gurstein, M. 2011. Open Data: Empowering the empowered or effective data use for everyone? First Monday vol. 16 (no. 2). doi: 10.5210/fm.v16i2.3316.

Helbig, N., A.M. Cresswell, G.B. Burke and L. Luna-Reyes. 2012. The dynamics of opening government data. A White Paper. Albany: Center for Technology in Government, University of Albany, Accessed July 1, 2014 http://www.ctg.albany.edu/publications/reports/opendata/opendata.pdf.

Huijboom, N. and T. van den Broek. 2011. Open Data: an international comparison of strategies. European Journal of ePractice no. 12 (March/April 2011):4-16. Accessed June 30, 2012 http://share-psi.eu/papers/TNO-epractice.pdf.

IRM [Independent Reporting Mechanism]. 2015. Aligning supply and demand for better governance. Washington: Independent Reporting Mechanism (IRM). Accessed October 19, 2015 http://www.opengovpartnership.org/sites/default/files/IRMReport-OpenData.pdf.

Janssen, K., J. Crompvoets and J. Dumortier. 2011. When is providing Spatial Information a public task? A search for criteria. Spatial Data Infrastructures in Context: North and South, edited by Zorica Nedović-Budić, Joep Crompvoets and Yola Georgiadou, 3-20. Baton Rouge, FL: CRC Press. Taylor & Francis Group.

Janssen, M., Y. Charalabidis and A. Zuiderwijk. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management 29 (4):258-268. doi: 10.1080/10580530.2012.716740.

Jetzek, T. 2013. The value of Open Government Data. Geoforum Perspectiv 12 (23 (2013)):47-56. doi: http://journals.aau.dk/index.php/gfp/article/view/489/400.

Jetzek, T., M. Avital and N. Bjørn-Andersen. 2014. Generating Sustainable Value from Open Data in a Sharing Society. IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2014, Aalborg, March 2014, doi:http://dx.doi.org/10.1007/978-3-662-43459-8_5.

Kok, B. and B. van Loenen. 2005. How to assess the success of National Spatial Data Infrastructures? Computers, Environment and Urban Systems 29 (6):699-717. doi: 10.1016/j.compenvurbsys.2004.02.001.

Koski, H. 2015. The Impact of open data – a preliminary study. Helsinki: ETLA, Accessed January 28, 2016. https://www.w3.org/2013/share-psi/wiki/images/6/67/Impact_of_Open_Data_in_the_Public_Sector_Koski_2015.pdf.

McClean, T. 2011. Not with a Bang but a Whimper: The Politics of Accountability and Open Data in the UK. APSA 2011 Annual Meeting. Accessed January 28, 2016 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1899790.

McKinsey Global Institute. 2013. Open Data: Unlocking innovation and performance with liquid information: McKinsey Global Institute, Accessed October 31, 2013 http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information.

McLaughlin, J. and S. Nichols. 1994. Developing a National Spatial Data Infrastructure. Journal of Surveying Engineering 120 (2):62-76. doi: http://dx.doi.org/10.1061/(ASCE)0733-9453(1994)120:2(62).

OECD [Organisation for Economic Co-Operation and Development]. 2011. Skills for innovation and research, doi:10.1787/9789264097490-en.

Omidyar Network. 2014. Open for Business: How open data can help achieve the G20 growth target. A Lateral Economics report commissioned by Omidyar Network. Accessed September 15, 2015 https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf.

Open Knowledge International. 2014. Global Open Data Index. Accessed December 20, 2015 http://index.okfn.org.

Open State Foundation. 2015. No substantial growth of Dutch government open data portal, Accessed April 15, 2015 http://openstate.eu/2015/04/2204/.

Osterwalder, A. and Y. Pigneur. 2010. Business Model Generation. Hoboken, NJ: John Wiley & Sons.

Peled, A. 2011. When transparency and collaboration collide: The USA Open Data program. Journal of the American Society for Information Science and Technology 62 (11):2085-2094. doi: 10.1002/asi.21622.

Rajabifard, A., M.-E.F. Feeney and I.P. Williamson. 2002. Directions for the future of SDI development. International Journal of Applied Earth Observation and Geoinformation 4 (1):p.11-22. doi: 10.1016/S0303-2434(02)00002-8.

Rhind, D. ed. 2014. What is the value of open data? What is the Value of Open Data? APPSI Seminar on 28 January 2014, London, April 2014. Accessed December 4, 2015 https://www.nationalarchives.gov.uk/documents/meetings/20140128-appsi-what-is-the-value-of-open-data.pdf.

Sunlight Foundation. 2010. Ten Principles for Opening Up Government Information. Sunlight Foundation, Last Modified August 11, 2010. Accessed August 31, 2015 https://sunlightfoundation.com/policy/documents/ten-open-data-principles/.

Termeer, C., A. Dewulf, H. van Rijswick, A. van Buuren, D. Huitema, S. Meijerink, T. Rayner and M. Wiering. 2011. The regional governance of climate adaptation: A framework for developing legitimate, effective, and resilient governance arrangements. Climate Law 2:159-179. doi: 10.3233/CL-2011-032.

The Cabinet Office. 2013. The Government Response to Shakespeare Review of Public Sector Information. Accessed December 5, 2013 http://data.gov.uk/sites/default/files/library/130529%20Government%20Response%20to%20Shakespeare%20Review%20of%20Public%20Sector%20Information%20v6%201.pdf.

Tinholt, D. 2013. The Open Data Economy. Unlocking economic value by opening government and public data: Capgemini Consulting, Accessed February 27, 2013 http://www.capgemini-consulting.com/sites/default/files/resource/pdf/opendata_pov_6feb.pdf.

Ubaldi, B. 2013. Open Government Data: OECD Publishing, doi:http://dx.doi.org/10.1787/5k46bj4f03s7-en.

Uhlir, P. F., ed. 2009. The socioeconomic effects of public sector information on digital networks. Towards a better understanding of different access and reuse policies - Workshop Summary. http://www.nap.edu/openbook.php?record_id=12687. Washington: National Academies Press.

van Loenen, B. 2006. Developing geographic information infrastructures: The role of information policies Ph.D. Dissertation, OTB Research Institute for the Built Environment, Delft University of Technology.

van Loenen, B., J. Crompvoets and A. Poplin. 2010. Assessing geoportals from a user perspective. 2nd workshop on Value of Geoinformation, 30 September - 2 October 2010, Hamburg, 2010. Accessed April 11, 2012 http://www.spatialist.be/download/pub/IO0906039_CrompvoetsVanLoenen_2010_Geovalue_Hamburg.pdf.

van Loenen, B. and M. Grothe. 2014. INSPIRE Empowers Re-Use of Public Sector Information. International Journal of Spatial Data Infrastructures Research 9 (2014):86-106. doi: 10.2902/1725-0463.2014.09.art4.

van Loenen, B., K. Janssen and F. Welle Donker. 2012. Quest for a global standard for geo-data licences. Refereed book chapter GSDI 13 World Conference Global Geospatial Conference 2012: Spatially enabling government, industry and citizens, Québec City, 14-17 May 2012. doi: http://repository.tudelft.nl/view/ir/uuid%3Aca7682be-9bca-47d9-8d29-0db047051cb5/.

Vickery, G. 2011. Review of recent studies on PSI re-use and related market developments. Paris: Information Economics, Accessed November 1, 2011 http://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments.

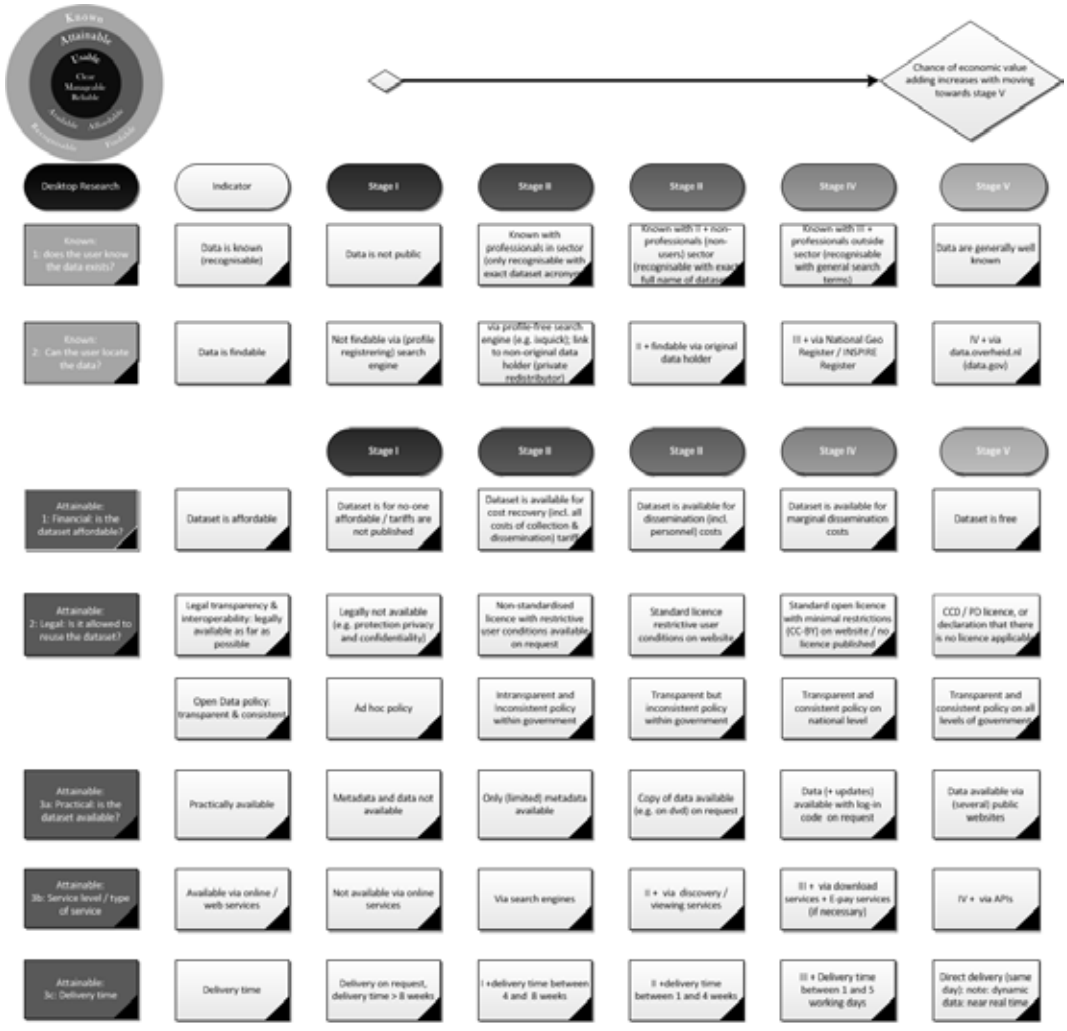
Welle Donker, F., B. van Loenen and A. K. Bregt. 2016. Open Data and Beyond. ISPRS International Journal of Geo-Information 5 (4). doi: 10.3390/ijgi5040048.

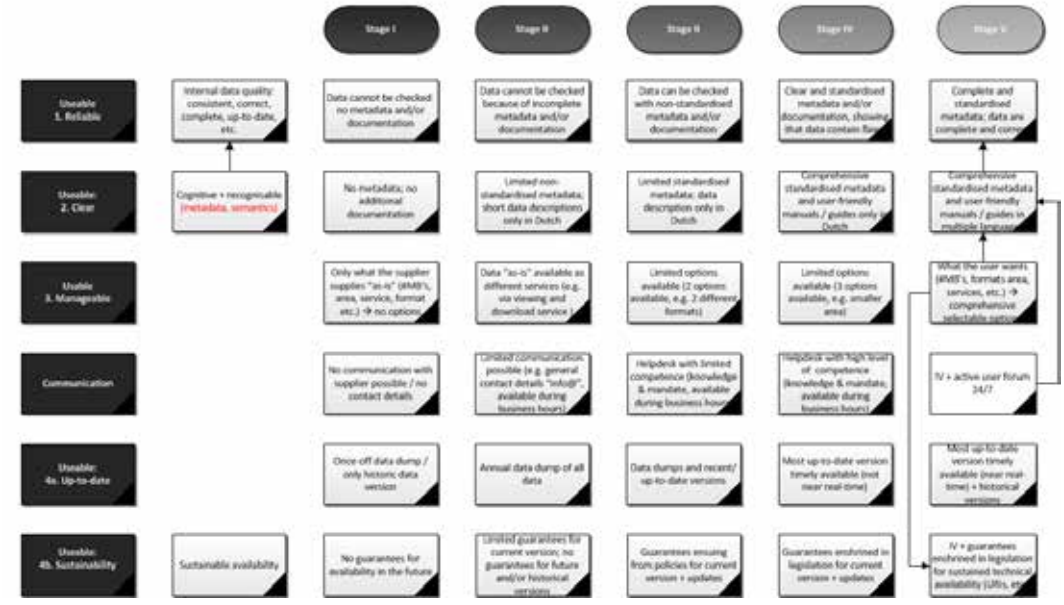
WISE Institute. 2014. Big and open data in Europe. A growth engine or a missed opportunity? Warsaw: WISE Institute [Warsaw Institute for Economic Studies] and demosEuropa. Accessed November 29, 2014 http://www.bigopendata.eu/.

World Wide Web Foundation. 2015. Open Data Barometer Global Report. 2nd Edition, Accessed January 31, 2015 http://barometer.opendataresearch.org/assets/downloads/Open%20Data%20Barometer%20-%20Global%20Report%20-%202nd%20Edition%20-%20PRINT.pdf.

Zuiderwijk, A.M.G. 2015. Open data infrastructures: The design of an infrastructure to enhance the coordination of open data use Ph.D. Dissertation, Delft University of Technology.

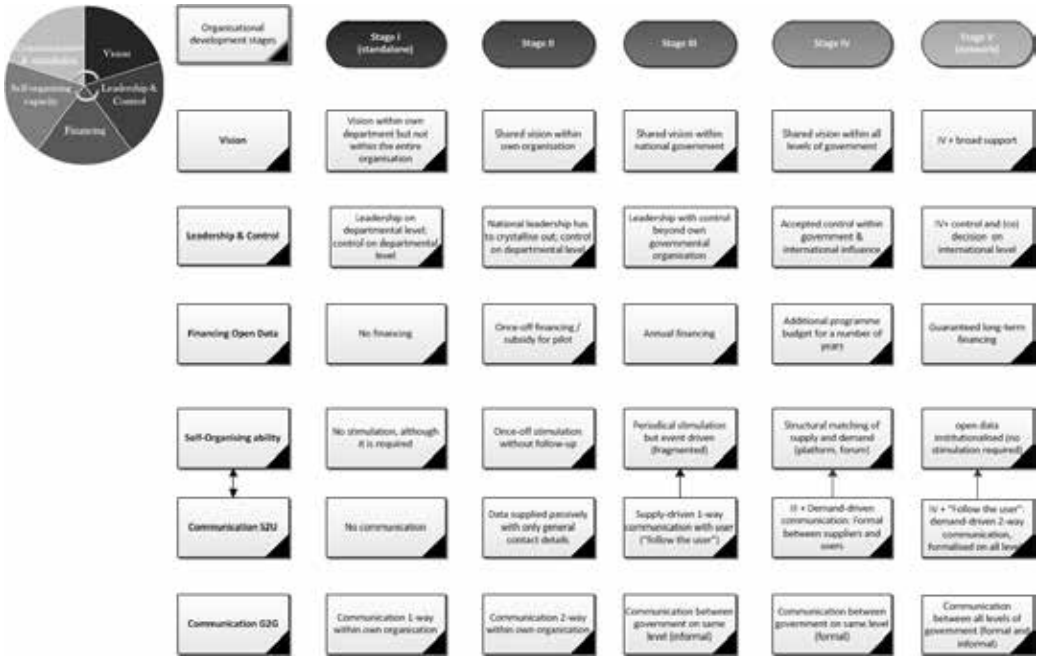
Appendix A Indicators used for "Top 20 Most Wanted Datasets" desktop research




Stage I
Stage II
Stage III
Stage IV
Stage V
Usable
1. Reliable
Internal data quality: consistent, correct, complete, up-to-date, etc.
Data cannot be checked no metadata and/or documentation
Data cannot be checked because of incomplete metadata and/or documentation
Data can be checked with non-standardised metadata and/or documentation
Clear and standardised metadata and/or documentation, showing that data contain fit
Complete and standardised metadata; data are complete and corre
Usable:
2. Clear
Cognitive + recognisable (metadata, semantics)
No metadata; no additional documentation
Limited non-standardised metadata; short data descriptions only in Dutch
Limited standardised metadata; data description only in Dutch
Comprehensive standardised metadata and user-friendly manuals / guides only in Dutch
Comprehensive standardised metadata and user-friendly manuals / guides in multiple languages
Usable:
3. Manageable
Only what the supplier supplies "as-is" (#MB's, area, service, format etc.) → no options
Data "as-is" available as different services (e.g. via viewing and download service)
Limited options available (2 options available, e.g. 2 different formats)
Limited options available (3 options available, e.g. smaller area)
What the user wants (#MB's, formats area, services, etc.) → comprehensive selectable option
Communication
No communication with supplier possible / no contact details
Limited communication possible (e.g. general contact details "info@"; available during business hours)
Helpdesk with limited competence (knowledge & mandate, available during business hours)
Helpdesk with high level of competence (knowledge & mandate, available during business hours)
IV + active user forum 24/7
Usable:
4a. Up-to-date
Once-off data dump / only historic data version
Annual data dump of all data
Data dumps and recent/ up-to-date versions
Most up-to-date version timely available (not near real-time)
Most up-to-date version timely available (near real-time) + historical versions
Usable:
4b. Sustainability
Sustainable availability
No guarantees for availability in the future
Limited guarantees for current version; no guarantees for future and/or historical versions
Guarantees ensuing from policies for current version + updates
Guarantees enshrined in legislation for current version + updates
IV + guarantees enshrined in legislation for sustained technical availability (URIs, etc

TOC

Appendix B Indicators used for assessment of Governance



TOC

9 Analysis and conclusions

This dissertation has provided an overview of the open data developments of the past decade. The dissertation started with an example of the problems a re-user may run into when wanting to re-use public sector data. The example was based on real-life experiences by people that were interviewed as part of the research carried out for the Space for Geo-Information (RGI) projects between 2006 and 2008. The main barriers to re-use of public sector information were identified as a lack of legal, financial, organisational and technical interoperability. When this research started, the debate focussed on access policies for public sector information and how to create a level playing field for re-users of public sector information. In spite of the so-called EU PSI Directive[68], which envisaged better accessibility of public sector information, access policies did not change significantly in the last decade in most European nations. The main research question of this dissertation was:

How can the accessibility of public sector information for re-users be improved?

In this chapter, the societal relevance and the scientific relevance will be provided. In addition, further research topics are recommended.

§ 9.1 Moving goalposts during this research: from restricted data to open data

Since this research started in 2006, a lot has happened. Open data happened. This research was unique in that it does not often happen that a number of variables are changed during the course of a qualitative research. However, after the Obama memorandum on Transparency and Open Government[69] in 2009 and the Digital Agenda for Europe[70] in 2010, accessibility of public sector information has gained momentum. More and more public sector information becomes available as open data. Thus, of the four barriers to re-use of public sector information data cited in

68 Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information.

69 https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment

70 http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Asi0016

Chapter 1 (legal, financial, technical and organisational barriers), two barriers appear to have been lifted due to open data. The shift from a cost recovery regime to an open data regime provided an excellent opportunity to test if the main barriers for re-users of public sector information are indeed restrictive licences and high fees as suggested by earlier research.

§ 9.2 Reflection on the theoretical framework

Chapter 2 showed that by 2008, although many Member States transposed the 2003 PSI Directive into a national framework with considerable delay, the effects of the PSI Directive were slowly starting to emerge. A number of Member States reviewed their access policies and some Member States made more public sector information available for no more than dissemination costs or reduced their fees significantly. Where fees were reduced, the number of regular re-users increased significantly and total revenue even increased in spite of lower fees.

Although the INSPIRE framework Directive paved the way for technical interoperability by providing guidelines for web services and catalogues, neither the INSPIRE Directive nor the 2003 PSI Directive the EU directives had tackled the issue of legal interoperability. Chapter 2 demonstrated that a major barrier to the envisioned level playing field for the private sector was the result of a number of public sector bodies acting as value added resellers (VARs). These public sector bodies had developed products and services based on their own data in direct competition with the private sector. Using their economy of scales advantages and employing restrictive licence conditions for the private sector re-using the same data, the public sector bodies appeared to be acting as monopolists.

Chapter 3 researched the aspect of legal interoperability. After an exploration of the legal barriers faced by re-users of public sector information, one of the most prominent barriers was caused by complex, intransparent and inconsistent licence agreements, especially when re-users wanted to combine data from different sources. The conclusion of this chapter was that the introduction of a Creative Commons inspired concept, such as the Geo Shared concept, would be a step towards legal interoperability for public sector information. Both Creative Commons and Geo Shared licence suites enable harmonisation of licence conditions and thus, promote transparency and consistency as re-users can see in one glance which restrictions are applicable . Geo Shared licences were considered as a serious option to be included to the INSPIRE framework and were implemented in the Dutch National GeoRegister.

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

In Chapter 4, the interoperability of open data licences, as recommended in the 2013 Amended Re-use Directive was analysed. The chapter concludes that even though using open licences for public sector information should have addressed legal barriers to re-use, in practice, different types of open licences currently in use may not be so interoperable after all. Effectively, only a public domain declaration, such as a CC-0 declaration is suitable for open data re-users requiring with cross-border data sets.

Chapter 1 and 2 indicated that employing a cost recovery regime for dissemination of public sector information might impede re-use by the private sector. However, in 2008, there were still many advocates for maintaining a cost recovery regime, especially for self-funding agencies to ensure sufficient finances to maintain a sufficient level of data quality. In Chapter 5, business models for INSPIRE web services were explored. Although, depending on the type of web service, and type of re-user, there might be an argument for employing subscriptions as a pricing mechanism, the conclusion was that business models based on generating revenue from data licences were not viable in the long run for public sector information bodies and were not in the spirit of the INSPIRE Directive. This research concluded that public sector information web services with different pricing regimes were counterproductive to achieving financial interoperability.

Business model for open data providers were revisited in Chapter 6. Self-funding government agencies are under increasing pressure to implement open data. Business models of self-funding agencies that supplied open data and of one agency that is under pressure to supply (some) open data in the near future were analysed to see which adaptions may be necessary to ensure the long-term availability of high quality open data and the long-term financial sustainability of self-funding agencies. The case studies confirmed that providing (raw) open data does not necessarily lead to losses in revenue in the long term as long as the organisation has enough flexibility to adapt its role in the information value chain and where revenue from data provision only represents a relative small part of the total revenue. The case studies indicated that open data has led to internal efficiency gains. In practice, it is difficult to quantify internal efficiency gains solely due to open data in isolation as the researched organisations continuously implement measures to increase efficiency. However, the reported decreases in internal and external transaction costs due to open data are in line with the case study carried out in Chapter 7.

During the course of this research, the introduction of open data provided an excellent opportunity to assess the effects of open data ex ante as baseline measurements could be carried out. To develop both quantitative and qualitative indicators to assess the success of a policy change is a challenge for open data initiatives. In Chapter 7,

Liander's intention to provide some of their data as open data offered an opportunity to apply the developed assessment model, and provide an insight into internal, external, and relational effects on Liander. A benchmark was carried out prior to release of open data and a follow-up measurement one year later. The follow-up monitor indicated that Liander open data are used by a wide range of users and have had a positive effect on energy consumption visualisation applications although it remains a challenge to quantify such societal effects. The follow-up monitor also indicated that Liander open data were used to improve existing applications and work processes rather than to create new products. The case study demonstrated that private energy companies can successfully release small-scale energy usage information open data. The case study also showed that Liander served as a best-practice case and had an open data flywheel effect on companies within the same sector. The monitoring model developed in this project was assessed to be suitable to monitor the open data effects.

The assessment model developed in Chapter 7 demonstrated the effects of open data on the organisation of an open data provider. The model is sufficient for a data provider to monitor the effects of open data on organisational level. However, to provide a broader picture of the state of open data and to assess if there are other barriers for re-users, a holistic approach is required. Therefore, an open data assessment framework, which assesses open data supply, open data governance and open data user characteristics holistically was developed and applied to the Dutch open data infrastructure in Chapter 8. This holistic open data framework assessed the maturity of the open data ecosystem in the Netherlands and proved to be a useful tool to indicate which aspects of the open data ecosystem are successful and which aspects require attention. The Holistic Open Data Maturity Assessment Framework showed that geographic data score significantly better than other government data. The standardisation and implementation rules laid down by INSPIRE Directive appears to have been a catalyst for moving geographic data to a higher maturity level. The assessment framework provided Dutch policy makers with useful inputs for further development of the open data ecosystem and development of well-founded strategies that will ensure the full potential of open data will be reached. Since the publication of the State of the Open Data Nation in 2014, a number of the recommendations have already been implemented.

§ 9.3 Practical results and research impact

From the start, this research has had a hands-on approach, due to the fact that the original research started as a one-year project, which was later extended. The first result of this research was the development of the Geo Shared licence suite based on the Creative Commons approach of harmonised and transparent licence conditions. The conceptual

model was enthusiastically received both nationally and internationally at various conferences, including the 12th EC-GI&GIS Workshop held in Innsbruck, Austria in 2006. The conceptual framework became a serious option for inclusion to the INSPIRE Directive as an annex. Unfortunately, the concept of one licence suite for the entire European Union came too early in 2006. The conceptual framework was further developed and adopted as the Geo Shared licence suite for the Dutch National GeoRegister. The Geo Shared licence suite was extremely useful as an interim licence suite, which created awareness of the necessity of harmonised and transparent licence conditions. The Geo Shared concept allowed self-funding public sector bodies to harmonise their licence conditions without sacrificing their cost recovery regime. As there was general support for the concept of harmonised licence conditions, it was a relatively small step to switch to Creative Commons licences after the adoption of the Digital Agenda in 2011. By 2016, many public sector bodies in the Netherlands use a Creative Commons licence requiring only source attribution (CC-BY) or a public domain declaration (CC0). In April 2016, there were only 4,847 datasets in the National GeoRegister with a Geo Shared licence, 4,836 of which published by Rijkswaterstaat, and the other datasets were mostly cultural heritage datasets.71 The research on business models for self-funding public sector bodies indicates that open data will not by definition endanger the long-term availability of high quality data, provided the organisation can adapt its role in the information chain and has a guaranteed income source, such as budged financing or revenue from legal instruments. The research formed an input into in a re-evaluation of the position of the Dutch Chamber of Commerce as a data provider of the Key Trade Register (van Loenen et al., 2016b). In July 2016, The Minister of Economic Affairs announced that an adopted version of the Trade Register will become available as open data within 12 months.72 Another result of this research is the development of the Holistic Open Data Maturity Assessment Framework. The State of the Open Data Nation report was well received in 2014 and served as an input for the Dutch government to fine-tune its open data policy. The Dutch government also carries out another recommendation: to establish an Open Data User Group to improve the governance of open data. A follow-up of the State of Open Data Nation was carried out in May and June 2016. The follow-up demonstrated that, overall, the supply side of open (geo)data had improved since 2014 and that steps were made to formally involve users in the governance of the open data ecosystem. However, to transform from a supply-driven open data ecosystem to a demand-driven open data ecosystem, more steps are needed, especially for non-geodata and sensor data (van Loenen, Welle Donker and Braggaar, 2016).

71 http://nationaalgeoregister.nl/geonetwork/srv/dut/search#fast=index&from=1&to=50&license=Geo Gedeeld licentie&relation=within (accessed 25 April 2016).

72 https://www.rijksoverheid.nl/documenten/kamerstukken/2016/07/04/kamerbrief-over-ontsluiting-handelsregister-als-open-data. In the open data version the register data will be anonymised so that individual businesses will not be identifiable by register number, name, address and postal code. The downgraded dataset will have a weekly update frequency.

§ 9.4 Main conclusions

This research has demonstrated that to achieve legal interoperability for cross-border re-use of (open) data, it is vital that only one licence suite is used in a transnational region. The Creative Commons licence suite is the best candidate as the Creative Commons licence suite are known worldwide and have been transposed into national legislation in many countries. This is in line with the recommendations of the European Commission for licencing public sector data (European Commission, 2014). For open data, only a public domain declaration, such as the CC0 declaration, should be used. The research carried out in Chapter 4 demonstrated that datasets cannot be combined, when organisations publish open data with their own version of a "CC-like" licence. A CC-BY licence, which required only source attribution, may seem a viable alternative for organisations that want to monitor re-use of data. However, it is essential that the version of CC-BY licences be updated when a new version is published, as demonstrated in Chapter 4. The other licences in the Creative Commons suite, *e.g.* a CC Share Alike licence or a CC Non-Commercial licence, are not suitable as these licences limit the creation of value-added derivative product and services. Therefore, for open data a CC0 declaration is the preferred, if not the only, option. Research in Chapter 8 also indicated that it is vital for re-users that the public domain declaration or open data licence is published in a prominent place to remove uncertainty of which licence conditions or declaration is applicable. If there is no licence or declaration published, re-use of open data is still impeded as re-users may not invest time in developing a value-added product or service when it is uncertain if the product/service can be marketed.

This research has also demonstrated that for open data to be successful, it is essential that self-funded public sector bodies be compensated for providing open data. This compensation may be budget financing from the central government to offset initial revenue losses due to open data. The self-funding public sector body may be provided with legal instruments to generate revenue. In the long-term, internal efficiency gains will offset the revenue losses due to open data and other sources of revenue will occur when a public sector organisation moves in the information value chain from data provider/aggregator to enabler.

This research presented a holistic open data assessment framework addressing the quality of open data supply, open data governance, and the user perspective of the open data infrastructure. By adding the user's perspective to the framework, a holistic comprehensive approach to open data assessment was provided. In general, open data governance in the Netherlands is well organised in some aspects but lags in others. Although there is an open data vision on strategic level and the concept of open data no longer a point of discussion, in 2014, there was no perception of clear leadership outside each organisation. On an operational level, government organisations were

struggling to apply the "open data, unless" policy to their specific data and would have benefitted if one organisation took control, especially for non-geodata. In addition, formal and structural communication (both intra-governmental and with users) should be established to match open data supply to demand as most of the current communication occurred on an ad-hoc and informal basis. It should be noted that the holistic open data assessment framework relied on interviews with nine users who were known in advance. To validate the framework, more users should be involved. However, as this research demonstrated, "the user" comes in many sizes and shapes and it remains a challenge for data suppliers to identify "the user" in an open data ecosystem.

§ 9.5 Further research

Open data is still in its infancy and the actual benefits of open data still have to materialise. However, measuring the effects and benefits of open data is no mean feat. Chapters 7 and 8 indicate that such benefits may initially be found in internal efficiencies for organisations. However, more case studies are needed to quantify the (longer-term) effects on data providers and on society. Such case studies should be carried out on an array of organisations, from self-funding agencies and decentralised governments to non-governmental organisations with a mandated public task; from commercial re-users to not-for-profit organisations and – if possible – citizens.
To ensure the long-term availability of high quality up-to-date open data, a governance framework should be developed which addresses not only aspects, such as long-term finances and commitments to manage open data, but also technical aspects, such as data interoperability, to ensure efficient re-use. More research is needed to formulate good governance of open data especially with current developments of linked open (meta)data, which may be beneficial to sustainable data availability in the long-term.

The Holistic Open Data Maturity Assessment Framework introduced in Chapter 8, needs to be fine-tuned and made more user-friendly. The currently defined maturity stages need to be translated into concrete questions. The results were based on researching a limited number of datasets and on a limited number of interviews. Although the sample was considered to be representative, the Holistic Open Data Maturity Assessment Framework should be applied to more datasets and to a broader group of data providers, especially lower governments and non-governmental organisations, to validate the model. Once fine-tuned, organisations can use the model as a self-assessment tool to monitor the state of their open data ecosystem in cooperation with the actual users.

In addition to more research on the effects of open data on data providers, more research is needed to assess the actual economic benefits of open data. Earlier research indicated that open data would lead to more creation of value-added products and services, which, in turn, would lead to additional revenue raised by governments, such as corporate taxes and value added taxes (cf. Weiss and Pluijmers, 2002; van Loenen, 2006; Pettifer, 2009; Ubaldi, 2013). This dissertation showed that the predicted free flow of value-added products and services based on public sector data still lag expectations. For instance, a UK study estimated that Ordnance Survey open data would generate a net growth in Gross Domestic Product of between £13.0 million and £28.5 million per annum by 2016 (Carpenter and Watts, 2013). However, at the same time there may also be a fear that 'open data could be immediately "swallowed up" by big global companies and not benefit the national economy' (Michael Fallon, UK Minister for Business and Enterprise, cited by PASC, 2014). If the costs of open data outrun the benefits, there is a fear that open data policies may stall or even be reversed if the benefits are not visible (*e.g.* World Wide Web Foundation, 2015). To assess the socio-economic effects of open data, a framework with economic indicators should be developed. It appears that the national economy does not benefit from open data re-used by global corporations, as corporate taxes may not be payable in the country where the data were re-used. It would be interesting to research what the economic effects are of taxes that are payable by global corporations, *e.g.* value added taxes on advertising, or what the effects would be of alternative forms of corporate taxation, such as unitary taxation (see *e.g.* Faccio, 2016).

Data infrastructures are developed to publish and share open data in a variety of domains. However, many of these data infrastructures suffer from poor data searching and finding functionality, data analysis functionality, data visualisation, lack of interaction, and variations in data quality (Zuiderwijk, 2015). Zuiderwijk (2015) stresses the importance of resource and contextual metadata to enable the findability of data. This research has shown re-users of public sector data often perceive metadata to be of poor quality, incomplete, outdated, and often non-machine processable. This research also showed that not all re-users actually use the metadata files provided by the data holders. Instead they prefer metatags and/or additional documentation. In addition, users often bypass the data portals developed by governments and instead, use other strategies to locate the required data, which was confirmed in a follow-up of the State of the Open Nation in the Netherlands in 2016 (van Loenen, Welle Donker and Braggaar, 2016a). The Holistic Open Data Maturity Assessment Framework already identified that some user characteristics, *e.g.* lack of resources or IT skills, form a barrier to successfully market value added product and services. However, the assessment framework does not explain why open data cannot live up to its expectations. The addition of the user perspective including user characteristics is relevant for this assessment. However, it may be possible that other aspects, such as a critical mass of well-equipped users, are equally of importance. Further research should look into this aspect of user (group) characteristics and its role in the

performance of open data ecosystems. The framework should be applied to a broader and more heterogeneous group of re-users to assess the actual user needs and barriers to re-use. The results of on-going monitoring will become the basis of making data infrastructures more user-friendly and match open data demand to open data supply. It will be a challenge to identify re-users of open data when according to the open data principles, re-users should not have to register prior to using the data. The system developed by Labots (2016) which can be used to identify user groups through IP addresses, may be a starting point.

One of the main challenges of data infrastructures is to institutionalise good governance practices. To guarantee long-term financial and organisational commitment to ensure that high-quality data are updated regularly and remain available in a sustainable manner is quite a challenge. The European Location Framework (ELF) project provides a good case study to research the governance of a transnational data infrastructure beyond project funding.[73]

Legal interoperability for data relates to a compatible legal environment of laws, policies and agreements needed to allow seamless exchange, combination and re-use of data between different organisations and countries (*cf.* van Loenen, Janssen and Welle Donker, 2012). Legal interoperability is more than just interoperable licences; it also requires agreement on interoperable legal systems and agreement on the interpretation of legal access frameworks and aspects, such as data protection legislation. To achieve true legal interoperability, more research is needed.

The 2013 PSI Amended Re-use Directive encourages the implementation of open data policies for public sector data. This research demonstrated that for geographical data many implementation aspects, such as standardised metadata, have already been addressed. This research also showed that for non-geographical data, there is still a long way to go. There is a growing demand for non-geographical data, such as financial information, healthcare information and research data (see *e.g.* Ubaldi, 2013 and Omidyar Network, 2014). However, these types of information often contain microdata, *i.e.* data on the characteristics of units of a population, such as individuals, households, or establishments, collected by a census, survey, or experiment.[74] Although microdata need not necessarily contain personal data, there are other issues that need to be addressed. Microdata may contain data that could be classified as confidential information. In addition to a potential breach of confidentiality, publishing microdata may pose other threats, such as revealing poor data quality

73 See *e.g.* http://inspire.ec.europa.eu/events/conferences/inspire_2016/pdfs/2016_workshops/29%20THURSDAY_WORKSHOPS_A_9.00-10.30_____ELF-INSPIRE-%20love%20story-v1.ptt.pdf

74 https://stats.oecd.org/glossary/detail.asp?ID=1656

or misinterpretation of the data due to a lack of knowledge (Jackson 2012). Thus, microdata probably need to be processed in some way before they are suitable to be published as open data. In addition, there is a fear that although (micro)data may appear to be sufficiently anonymised or aggregated, the data may become personal data by combining it with other publicly available data or when it is de-anonymised (Kulk and van Loenen, 2012). As the 95/46/EC Data Protection Directive[75] is interpreted in different ways in Europe, protection of personal data may impede the implementation of open data policies, especially with increasing technological advances. The General Data Protection Regulation (GDPR)[76] will replace the Data Protection Directive in 2018. Although the new Regulation introduces new concepts such as the right to be forgotten, data portability, personal data breach notification, profiling, and easier access to their own data to strengthen citizens' fundamental rights (see, *e.g.* de Hert and Papakonstantinoua, 2012; Mantelero, 2013), the new Regulation does not fundamentally change the concept of personal data (van Loenen *et al.*, 2016b). The tension between open data and protection of personal data requires more research. This research was recently started with the PhD project 'Safeguarding Data Protection in an Open data World (SPOW)' at TU Delft and Tilburg University. However, the technical aspects of re-identifying anonymised or aggregated (micro)data need to be researched.

To assist public sector bodies to implement open data policies, a decision tree was developed. The decision tree can be used to determine if a dataset is suitable to be published as open data. The decision tree was originally developed for Rijkswaterstaat and proved to be a helpful tool for successfully implementing open data. The decision tree is available for re-use by other organisations via the Dutch open data portal data.overheid.nl.[77] The decision tree was fine-tuned and extended, and will be used by the Dutch National Institute of Public Health and the Environment (RIVM) as part of their data management framework. However, RIVM data are complex: the datasets are of a heterogeneous nature and collected in various domains, such as public health, infectious diseases and environmental (sensor) data. Not only are there challenges related to data formats and data size but there are legal challenges as well. For instance, data related to infectious diseases are subject to a complex and – at times – apparently contradictory international legal framework. Many of the datasets contain microdata and there are uncertainties related to the level of aggregation required to

75 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

76 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

77 https://data.overheid.nl/sites/default/files/Beslisboom%20open%20data%20v2016.pdf

comply with personal data protection regulations or to safeguard the concerns related to confidentiality. In addition, there are datasets that were co-created with other public sector organisations and/or private companies, or the data are collected by citizens on a voluntary basis, *e.g.* bird counts. The issues concerning microdata as open data were already discussed in the previous paragraph. The issue of data ownership of co-created data needs to be further researched as third party rights may be a barrier to publishing open data (van Loenen, Welle Donker and Ploeger, 2016c). Monitoring the implementation of open data for an institute such as RIVM will provide a good opportunity to fine-tune the decision tree based on user experiences.

§ 9.6 Final reflection

In the last decade the accessibility of public sector information has vastly improved. The intrepid student described in Section 1.2, would not face as many barriers faced in 2006. To develop the intended app, the student needed a national topographic map, municipal and provincial real-time public transport information, road maintenance data and crime statistics related to bicycle theft hotspots. The national topographic map has been available as open data since January 2012, as are the road maintenance data. Public transport data are published as open data, although this research demonstrated that it is not easy to locate municipal public transport data. Only the large municipalities publish public transport as open data via a central portal and not all data are published as real-time data. To locate and re-use municipal public transport data, a person needs domain knowledge and perseverance. Fortunately, there are community-driven platforms, such as OpenOV, that provide tools and documentation to aid re-users of public transport data. Crime statistics related to bicycle theft hotspots also require some domain knowledge. The Dutch police publish (near real-time) crime statistics available as a web mapping service[78], however, the available aggregated data relate to burglaries and attempted burglaries. The Police also publish crime statistics per category (including bicycle theft) as open data aggregated on municipal level.[79] Some municipalities publish interactive maps showing bicycle theft statistics based on open data.[80] However, these maps are based on older data

78 https://www.politie.nl/themas/misdaad-in-kaart.html

79 https://data.overheid.nl/data/dataset/criminaliteitsfrequenties-2015-per-categorie-en-gemeente

80 See *e.g.* https://projecten.versbeton.nl/fietsdiefstallen/ showing theft statistics in Rotterdam, which can be filtered on time of day and on the brand of bicycle or moped up to 2013, and and http://www.dordrechtopenda-ta.nl/fietsen/ showing theft statistics, which can be filtered on time of day and on bicycle or moped up to 2014.

and do not appear to be updated at all. Our intrepid student may be successful when he/she attends a hackathon where the required data may be available, at least for a limited time. Although hackathons have less value for professional re-users, hackathons are a means to establishing a network, which will be needed at a later stage to obtain required data.

Although many of the barriers identified in Chapter 1 have been alleviated to some degree with the emergence of open data, there are still barriers to re-use. Even though much public sector information becomes available as open data, the findability and the timeliness of the data are wanting. On national level, the governance of open data has improved, but decentralised governments appear to be struggling to implement open data policies. With the emergence of open data, we are slowly moving from access to re-use However, there is still a long way to go before we have moved from a supply-driven to a demand-driven data infrastructure.

References

Carpenter, J. and P. Watts (2013). Assessing the value of OS Open Data to the economy of Great Britain - Synopsis. A synopsis of the Full Study Interim Report prepared for Department of Business, Innovation and Skill by ConsultingWhere Limited and ACIL Tasman. https://www.gov.uk/government/publications/ordnance-survey-open-data-economic-value-study: 32, accessed June 2013.

de Hert, P. and V. Papakonstantinoua (2012). The proposed data protection Regulation replacing Directive 95/46/EC: A sound system for the protection of individuals. Computer Law & Security Reviews 28(2): 130-142 DOI: 10.1016/j.clsr.2012.01.011.

European Commission (2014). Guidelines on recommended standard licences, datasets and charging for the reuse of documents, Official Journal of the European Union. C240/01: 10, https://ec.europa.eu/digital-agenda/en/news/commission-notice-guidelines-recommended-standard-licences-data-sets-and-charging-re-use

Faccio, T. (2016). Why Google (and other multinationals) are still not paying their fair share of corporation tax, in: Tax Justice Network, http://www.taxjustice.net/2016/02/16/why-google-and-other-multinationals-are-still-not-paying-their-fair-share-of-corporation-tax/, accessed: October 28, 2016.

Kulk, S. and B. van Loenen (2012). Brave New Open Data World? International Journal of Spatial Data Infrastructures Research 7: pp. 196-206. DOI: 10.2902/1725-0463.2012.07.art10, access date:

Labots, W. (2016). Gebruiker in kaart. Analyse van loggegevens van de PDOK-webservices om groepen gebruikers van open data te identificeren. Delft, Delft University of Technology: 33, Accessed February 2016.

Mantelero, A. (2013). The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten', Computer Law & Security Reviews 29(3), pp.229-235, doi: 10.1016/j.clsr.2013.03.010.

Omidyar Network (2014). Open for Business: How open data can help achieve the G20 growth target. A Lateral Economics report commissioned by Omidyar Network, https://www.omidyar.com/sites/default/files/file_archive/insights/ON%20Report_061114_FNL.pdf, accessed September 15, 2015.

PASC [Public Administration Select Committee] (2014). Public Administration Committee - Tenth Report. Statistics and Open Data: Harvesting unused knowledge, empowering citizens and improving public services. London, http://www.publications.parliament.uk/pa/cm201314/cmselect/cmpubadm/564/56402.htm, Accessed 17 March 2014.

Pettifer, R.E.W. (2009). PSI in European Meteorology - an unfulfilled potential, in 7th Eastern European eGov Days: eGovernment & eBusiness Ecosystem & eJustice Conference (European Projects and Management (EPMA) (CZ), Prague), http://www.primet.org/ckfinder/userfiles/files/EU%20PSI%20Working%20Groups/PSI%20in%20European%20Meteorology%20-%20an%20unfulfilled%20potential%20distribution%20copy.pdf, accessed: May 6, 2009.

Ubaldi, B. (2013). Open Government Data, OECD Publishing: 61, doi: http://dx.doi.org/10.1787/5k46bj4f03s7-en.
van Loenen, B. (2006). Developing geographic information infrastructures: The role of information policies Ph.D. Dissertation, OTB Research Institute, Delft University of Technology, Delft. doi: uuid:6781e9dd-1468-4b31-bf88-ee1fbfffe4a1.
van Loenen, B., K. Janssen and F. Welle Donker (2012). Quest for a global standard for geo-data licences, in GSDI 13 World Conference Global Geospatial Conference 2012: Spatially enabling government, industry and citizens (Québec City), doi: uuid:ca7682be-9bca-47d9-8d29-0db047051cb5.
van Loenen, B., F. Welle Donker and R. Braggaar (2016a). De stand in opendataland 2016, [The State of the Open Data Nation 2016]. Delft: Kenniscentrum Open Data, doi: uuid:b76cb0d8-e9b5-4125-8d7c-b452de1da0d0.
van Loenen, B., W. Korthals Altes, D. Groetelaers and F. Welle Donker (2016b). Ontsluiten handelsregister met open data nader belicht [Unlocking Trade Register as open data further discussed], Delft: Kenniscentrum Open Data, doi: uuid:2c053185-fb19-49cb-a6db-55efdd567647.
van Loenen, B., F. Welle Donker and H. Ploeger (2016c). RIVM open data Delft, Kenniscentrum Open Data: 71, 30 juni 2016. doi:uuid:c4df5c20-d2d8-4045-8bed-833163125f98.
Weiss, P.N. and Y. Pluijmers (2004). Borders in Cyberspace: Conflicting Public Sector Information Policies and their Economic Impact, in Public Sector Information in the Digital Age. Between markets, public management and citizens' rights. G. Aichholzer and H. Burkert (Eds), (Cheltenham, UK; Northampton, MA, US: Edward Elgar Publishing Ltd), pp. 137-159.
World Wide Web Foundation (2015). Open Data Barometer Global Report. 2nd Edition, http://barometer.opendataresearch.org/assets/downloads/Open%20Data%20Barometer%20-%20Global%20Report%20-%202nd%20Edition%20-%20PRINT.pdf, accessed: January 21, 2015.
Zuiderwijk, A.M.G. (2015). Open data infrastructures: The design of an infrastructure to enhance the coordination of open data use. Faculty of Technology, Policy and Management. Delft, Delft University of Technology. Ph.D. dissertation: 376. doi:http://dx.doi.org/10.4233/uuid:9b9e60bc-1edd-449a-84c6-7485d9bde012.

Curriculum vitae

Frederika Welle Donker was born and raised in Rotterdam in the Netherlands. After finishing high school, she migrated to Sydney, Australia. She worked for four years as a production manager in a factory producing electrical switchgear. During that time, she enrolled in a part-time Electronics Engineering Certificate course. To be able to graduate, it was necessary to do a traineeship. Therefore, Frederika moved to the National Building Technology Centre (NBTC), a research institute in North Ryde, to work as a Trainee Technical Officer. In the three years she worked there, Frederika had her first taste of carrying out research. This research into the flammability and combustibility of building and construction materials involved setting fire to curtains and floor coverings in the name of science. Although she would have liked to continue her career at the NBTC, unfortunately, after the NBTC was incorporated into the Commonwealth Science & Industrial Research Organisation (CSIRO), this was not an option anymore. Instead, Frederika moved to Cumberland College of Health Science at Lidcombe, a college of Sydney University. As a Technical Officer, she was instrumental in setting up the practical facilities for a new School of Medical Radiation Technology. During that period, Frederika also enrolled in an Associate Diploma of Electrical Engineering course. Upon graduation, she was promoted to Senior Technical Officer. As Laboratory Manager, she often assisted with practical classes for the undergraduate students. She also assisted M.Sc. and Ph.D. students who required guidance with radiation monitoring equipment, a second taste of research.

After 10 years at Sydney University, Frederika decided to move back to the Netherlands for personal reasons. As there was a backlog with working permit applications at that time, she was not allowed to work for about three months. To keep occupied, Frederika decided to enrol in University. A friend had just enrolled in a part-time degree at the Faculty of Technology, Policy and Management of Delft University of Technology. As the course consisted of one-year modules, the plan was to attend one year, obtain the certificate and move on. However, the course proved to be too interesting to leave after one year. Frederika finished the four-year degree course in three years. During that time, she also worked at the Faculty of TPM in the Secretariat where she was, inter alia, responsible for maintaining Blackboard.

After graduation, Frederika worked for three years as a research assistant at the Erasmus Medical Centre in Rotterdam. When the research project ended, Frederika applied for a position as a Ph.D. student at the OTB Research Institute for the Built Environment, Delft University of Technology. Although another applicant was considered to be more suitable, she was offered a one-year contract to carry out a research into the legal and financial aspects of geographical information accessibility.

This research led to the development of a licence suite for geographical information, the GeoShared licence suite. The one-year contract was extended and developed into a Ph.D. research into the barriers of re-use of public sector information. In 2012, Frederika joined the Geo-Information Governance Knowledge Centre, and the focus of her research switched to open data.

In the last eleven years, Frederika carried out many contract research projects in the Netherlands and participated in the European Location Framework (ELF) Programme, and has (co-)written many reports. A number of the research projects have resulted in publications in peer-reviewed journals and book chapters, which became the basis of this dissertation. In addition to her work as a researcher, Frederika participates in the Geo-Information Management Applications (GIMA) M.Sc. Programme, a cooperation between four universities in the Netherlands. She is a lecturer and an assistant coordinator for the GIMA Internships. After the defence of her dissertation, Frederika will continue as a researcher and lecturer at the Geo-Information Governance Knowledge Centre.

This dissertation has researched the barriers faced by re-users of government data in the past, present and future, and provides some tools to tackle these barriers. With the emergence of open data, a number of barriers have been lowered or lifted altogether. However, there is still a long way to go before the track is cleared.

Delft University of Technology,
Faculty of Architecture and the Built Environment,
OTB – Research for the Built Environment