

OPEN DATA IMPACT CASE STUDIES

DENMARK'S OPEN ADDRESS DATA SET

Consolidating and Freeing-Up Address Data



By Juliet McMurren, Stefaan Verhulst and Andrew Young¹

December 18, 2015



¹ Special thanks to Akash Kapur who provided crucial editorial support for this case study, and to the peer reviewers odimpact.org/about who provided input on a pre-published draft.

Summary: A broad variety of services depend on accurate, up-to-date address data, including emergency services; the police; postal and transport services; and GPS systems. In 2005, the Building and Dwelling Register of Denmark started to release its address data to the public free of charge. Prior to that date, each municipality charged a separate fee for access, rendering the data practically inaccessible. Other entities, such as the Cadastre, the Land Registry, the Central Business Register and the utilities sector, collected their own address information, leading to significant discrepancies between the content held across different databases. A follow-up study commissioned by the Danish government estimated the direct financial benefits alone for the period 2005-2009 at EUR 62 million, at a cost of only EUR 2 million.

Dimension of Impact: Improving Government – Improving Services

Key Takeaways:

- In highly open, transparent societies, the opening of government data can be successfully justified and driven by expected economic benefits from increased efficiency and innovation within government, without the need for political arguments for transparency.
- Relatedly, efforts to collect and open data can uncover major gaps and deficiencies in official government data sources, creating benefits to internal efficiency and recordkeeping, often with little to do with improving transparency or accountability.
- There is a need to establish robust and carefully planned governance models for data collection systems, and skills for data collection among those charged with collecting it, to ensure the resulting data set is clean, accurate, usable and widely used.

I. Context and Background

Denmark is a high-income¹ constitutional monarchy in northern Europe with a population of 5.6 million in 2015.² It is ranked 10th on the United Nations Human Development Index.³ Locally, the country is divided into five regions, which are further subdivided into 98 municipalities, reduced from 270 in 2007.

Denmark is considered, both nationally and internationally, among the most highly open and free countries in the world. The country was ranked second on the 2014 Global Open Data Index by the Open Knowledge Foundation,⁴ ninth on the 2015 Open Data Barometer⁵ and fourth on the World Justice Project's Open Government Index.⁶ The Danish media is considered free by Freedom House,⁷ and the country is ranked third in the World Press Freedom Index by Reporters Without Borders.⁸ Consequently, corruption is not widely considered a problem in Denmark.⁹

Denmark is a signatory to the Open Government Partnership. It began the process to join in September 2011, and is currently implementing its first Action Plan and developing its second, which has open data as one of its four themes. The Action Plan expresses the government's intention to develop an Open Data Innovation Strategy (ODIS), to draw attention to the potential of public data and make it accessible. ODIS will also operate the public data catalogue and facilitate communication between public sector and private reusers of data.¹⁰

Despite its wealth, Denmark has found itself under increasing pressure to find ways to control government spending. Danish citizens are among the most highly taxed in Europe,¹¹ but there is little public appetite for government program cuts.¹² Nonetheless, according to OECD data from 2011, the Danish government spent over a quarter of its GDP (26.1 percent) on social spending, compared to an average of 21.7 percent across OECD member countries.¹³ The OECD also found that 26.8 percent of the labor force was employed in the government sector (excluding public corporations) in 2012, and compensation for these employees consumed 19 percent of GDP, the highest percentage in the OECD.¹⁴ Faced with an aging population and fewer taxpayers, a real prospect of government spending reaching

1 <http://data.worldbank.org/country/denmark>

2 <http://denmark.dk/en/quick-facts/facts/>

3 http://hdr.undp.org/sites/all/themes/hdr_theme/country-notes/DNK.pdf

4 <http://index.okfn.org/place/denmark/>

5 <http://www.opendatabarometer.org/report/analysis/rankings.html>

6 <http://worldjusticeproject.org/open-government-index/open-government-around-world>

7 <https://freedomhouse.org/report/freedom-press/freedom-press-2015#.Vf29Nt9Viko>

8 <http://index.rsf.org/#/>

9 http://transparency.dk/?page_id=1258

10 The Danish Government, "Open Government National Action Plan 2013-2014." October 2013. http://www.opengovpartnership.org/sites/default/files/Denmark_Open%20Government%20Action%20Plan%202013-2014_ENG_1-sided_print.pdf, pp 14-15; "Open Data Innovation Strategy (ODIS)." Digitaliseringsstyrelsen. April 22, 2013. <http://www.digst.dk/ServiceMenu/English/Policy-and-Strategy/Open-Data-Innovation-Strategy-ODIS>

11 "Tax reforms in EU Member States 2012: tax policy challenges for economic growth and fiscal stability." Taxation Papers 34 and European Economy 6. http://ec.europa.eu/europe2020/pdf/themes/02_taxation.pdf

12 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

13 <http://www.oecd.org/social/expenditure.htm>

14 "Human Resources Management Country Profiles: Denmark." OECD. December 6, 2012. <http://www.oecd.org/gov/pem/OECD%20HRM%20Profile%20-%20Denmark.pdf>

unsustainable levels, and a population reluctant to embrace budget cuts, the Danish government has had no option but to look for ways to improve government efficiency.¹⁵

II. Project Description and Inception

Before 1996, address data was registered and collected individually by the 270 Danish municipalities. Although public data was available, organizations wanting to access the data had to make separate access and pricing agreements with each municipality, rendering the data practically inaccessible. The lack of an accessible, unified public data set resulted in the development of several private databases of varying quality.¹⁶ Furthermore, there was no consensus on the definition of an address, or who (if anyone) held the reference address file, with users such as Post Danmark and the emergency services each holding rival data sets.¹⁷

Appreciating that there would be considerable practical benefit to a single standardized address system, with each address datum attached to a unique geographic point, officials at the Danish National Survey and Cadastre (now the Danish Geodata Agency) began a pilot program coordinating this already-collected data into a single body. Morten Lind, business manager of the Danish Address Program at the Danish Geodata Agency, describes the origins of the program: “We started as a pilot project in 1996, where we said, there’s a potential in having good address data, and in taking addresses and giving them a geographic coordinate so that we know where each address is. If every government agency has its own separate address database, we are not sure that we can actually compare things. So we had this idea of making a common reference data set where we could say: These are the addresses that exist. [And] it was clear that the only party that could create this reference file of addresses was the local authorities.”¹⁸

In order to do this, the Danish National Survey and Cadastre decided to adapt the existing building and dwelling register, with minor adjustments, as the primary address register. The local authorities would collect the data, assign and register addresses and street names, and would have responsibility for and ownership of the data, while the Danish Survey and Cadastre would coordinate the project.¹⁹ At this point, the submission of data was voluntary, and municipalities could not be compelled to supply address data, but some financial support was offered to help municipalities with compliance.²⁰ Lind and a colleague from the Danish organization of local authorities also met with local authorities throughout the country to promote the project.²¹

15 Jetzek, Thorhildur. “Managing complexity across multiple dimensions of open data: the case of the Danish Basic Data Program.” *Government Information Quarterly*, November 24, 2015. <http://www.sciencedirect.com/science/article/pii/S0740624X15300186>

16 Lind, Morten. “The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge.” *Danish Enterprise and Construction Authority*. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

17 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

18 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

19 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

20 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

21 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

The potential interest and utility of the data set quickly became apparent to those working with it. A project manager at the Ministry of Housing, Urban and Rural Affairs commented: “It did not take many years after starting to work with the public data in around 1993 to realize that these data could be used for many interesting things. And we felt it was a shame ... that the data weren’t used more.”²²

By 2001, the data had been collected in a single register owned by the local authorities. Potential users such as the police, emergency services, public health agencies, public transportation companies and agencies like the Danish Ministry of the Environment and county administration were aware of and eager to use the data to provide their services armed with a more accurate understanding of the locations of homes, businesses and other entities, but, as Lind explained, the next hurdle was finding an agreement on how it could be accessed and used. “The police want[ed] to use it, the Ministry of the Environment want[ed] to use it, but ... we had almost 300 local authorities that could decide their own pricing models on a completely individual basis. We couldn’t foresee that an agreement could be reached [on] a price [for] address data. So we were in a deadlock. ... We had a lot of very good data but nobody used it.”²³ Some government agencies tried to access the data via individual agreements with each local authority: “a very cumbersome process even for a large organization.”²⁴

The deadlock was eventually broken through the intervention of the Minister of Finance. Analysis by the Ministry of Finance’s eGovernment Taskforce had identified address data as one of the data sets with the most potential use, leading the Minister of Finance to broker an agreement on access and use with municipalities. This agreement, officially called the Better Access to Public Data, but more widely known as the “free of charge agreement,” would make available data from the Cadastre and municipal property and dwelling registers, which comprised address data and their associated geographic coordinates, free of charge through a government portal, with those accessing it paying only the cost of distribution.²⁵ Although the Danish National Survey and Cadastre had advocated this solution, the outcome was, Lind concedes, surprising, since the government was at that stage committed, by policy and legislation, to selling its data.²⁶

The efforts and influence of Lind were crucial in advancing the cause of opening the data.²⁷ He had succeeded in convincing his superiors at the Danish National Survey and Cadastre of the necessity of opening the data without charge, since an agreement on pricing seemed elusive, and this became the agency’s position. He and his colleagues had involved potential users and stakeholders in pressing for the release of the data through a series of meetings and conferences at which the benefits and opportunities of the data set were explained. “It was very important for us to do that because we had to put ourselves on the agenda for the Minister of Finance, [who] has many important issues to solve, and this one was only a little one.”²⁸ Post Danmark was also, according to Lind, a major driver behind the

22 Jetzek, Thorhildur. “Managing complexity across multiple dimensions of open data: the case of the Danish Basic Data Program.” *Government Information Quarterly*, November 24, 2015. <http://www.sciencedirect.com/science/article/pii/S0740624X15300186>

23 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

24 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

25 Jetzek, Thorhildur. “Managing complexity across multiple dimensions of open data: the case of the Danish Basic Data Program.” *Government Information Quarterly*, November 24, 2015. <http://www.sciencedirect.com/science/article/pii/S0740624X15300186>

26 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

27 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015; and GovLab Interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21 2015.

28 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

decision to open, having declared early on that they wanted to rely on public data. Significant users such as Post Danmark and the emergency services have subsequently provided important feedback on how the data set could be improved.²⁹

The agreement came into effect on January 1, 2003, but legal issues delayed its full implementation until 2005, when an amendment to the law governing the Public Data Server removed the legal restrictions on the distribution of address data to third parties, as well as the fee for distribution.³⁰ The release was not conceived of as an open data project, according to Lind: “‘Open data’ ... was not a term we used then. We used the term ‘data free of charge’ ... [but] really it lives up to almost all of the definitions of open data today.”³¹ There would be no restrictions placed on the use or redistribution of the data beyond those required to satisfy the requirements of the law, particularly in regard to personal data or product marketing. Steps should continue to be taken, however, to ensure that such restricted, privacy-eroding activities are not being undertaken by internal or external users of the address data.

The municipalities were compensated EUR 1.3 million for loss of income from sales of data for the three years after the agreement was reached. At that point, the agreement offered them the opportunity to renegotiate for further compensation, but no further negotiations were sought by the municipalities.³² The compensation did not fully recompense the municipalities for the costs of data collection, but, as Lind notes, the data already had to be collected for their own use. “Our position was that the data was there, and of course we needed to have an assurance that the data will be updated, and of course there would be costs associated with this updating ... but this is something that the local authority has to do, agreement or not.”³³ The money offered was, instead, a compensation for no longer being able to sell the data. In addition, Danish Enterprise and Construction Authority (DECA), the agency responsible for road names and addresses in Denmark, calculated that the municipalities would realize savings from no longer having to negotiate data purchase agreements, deliver data or enforce licenses.³⁴

Beyond the local authorities themselves, which were the biggest users of the address data, the primary intended users of the data set were the public sector. Emergency services, public health disease prevention programs, national and county environmental agencies, public transport companies and Post Danmark were all expected to make use of the data set. Journey Planner, a then groundbreaking Web app allowing people to plan journeys using public transportation from one address to another, also relied on the address data set.³⁵ Input from intended users was sought through nationwide conferences and informal meetings between the Danish Survey and Cadastre and potential user groups such as the police.³⁶

29 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

30 Morten Lind “The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge.” Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

31 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

32 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

33 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

34 Morten Lind “The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge.” Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

35 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

36 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

Although the developers were aware of the potential benefits of the address data set for the private sector, Lind maintains that this benefit was always secondary to the desire to improve public sector efficiency and services. “We were of course aware that [this data] could have large benefits for the private sector in geomarketing, logistics, transport planning, route planning, and analysis of where to locate a new store, and so forth ... but that was not our selling argument. Our argument was the public sector. We have data that enables the public sector to be more efficient and to provide better services for citizens and for business, and it’s stupid that we don’t use that data.”³⁷

As far as its developers were aware, the Danish address data set was the first of its kind to be compiled and opened free of charge. Lind reports that they were influenced by a similar data set, Address-Point, compiled in the United Kingdom by the Ordnance Survey. Address-Point took the Royal Mail’s Postcode Address File (PAF), and attached geographic coordinates, creating a unique geolocated identifier for each address. Address-Point was, however, a commercial data set owned by the Ordnance Survey, and was not open.³⁸ Moreover, the PAF which enabled the development of Address-Point was controversially sold at the time of the privatization of the Royal Mail. In March 2014, the chairman of Parliament’s Public Administration Committee argued that, “The sale of the PAF with the Royal Mail was a mistake. Public access to public sector data must never be sold or given away again.”³⁹ Lind recalls, “I think we were totally on our own in deciding that the only possible model could be opening the data free of charge. I’m proud of that situation.”⁴⁰

37 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

38 <https://www.ordnancesurvey.co.uk/business-and-government/products/address-point.html>

39 Hogge, Becky. “Open Data: Six Stories About Impact in the UK.” Omidyar Network. November 2015. https://www.omidyar.com/sites/default/files/file_archive/insights/Open%20Data_Six%20Stories%20About%20Impact%20in%20the%20UK/OpenData_CaseStudies_Report_complete_DIGITAL_102715.pdf

40 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

III. Impact

The opening of the address data set was, almost universally, enthusiastically received, and the data set widely used. According to DECA, in 2009 address data was delivered to a total of 1,236 public and private parties, including 286 full, nationwide address sets. This response was based on feedback from only 14 of the 22 data distributors, and therefore underrepresents the total number of deliveries. Of these deliveries, about 70 percent were to the private sector, 20 percent to central and regional government, and 10 percent to the municipalities. Of the 286 nationwide address sets delivered, 12 were for IT products with more than 1 million end users.⁴¹

We have data that enables the public sector to be more efficient and to provide better services for citizens and for business, and it's stupid that we don't use that data.
– Morten Lind, Danish Geodata Agency

The exceptions to this warm response were a small number of private vendors with agreements to supply competing address data to municipalities, who were unhappy with the decision to open the public data, since it cut off a source of revenue. The majority, according to Lind, embraced the opportunity to get out of data collection. “The more future-minded of these companies said, we are happy not to use our efforts in collecting these data. We want to use our efforts in making applications [and] value adding onto the data, ... and making something more intelligent.”⁴² GPS and navigation companies were likewise happy to abandon data collection and adopt the open address data set, according to Lind.⁴³

DIRECT FINANCIAL BENEFITS

According to a 2010 study commissioned by the DECA, the direct financial benefits to society of the open address data for the period 2005-2009 totaled EUR 62 million through especially improved government back-end capabilities and more efficient service delivery. The total cost for the program over the same period was EUR 2 million, comprised of EUR 1.3 million in compensation to municipalities and the balance in costs to distribute the data through the Public Data Server (PDS).⁴⁴ The benefits of the agreement in 2010 were expected to be EUR 14 million, while total costs would be EUR 0.2 million for distribution of the data through the PDS. Approximately 30 percent of the benefit was to the public sector, with the remainder to the private sector – including, notably the nongovernment Post Danmark, which saw major efficiency increases based on access to the data.⁴⁵ The study only measured direct

41 Lind, Morten. “The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge.” Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

42 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

43 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

44 Lind, Morten. “The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge.” Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

45 Lind, Morten. “Addresses and Address Data – Experiences from Denmark.” August 11, 2014. <http://www.slideshare.net/Mortlin/addresses-and-address-data-experiences-from-denmark-37872471>

financial benefit to those receiving address data via one of the PDS's 22 distributors, and did not attempt to place a value on the supplementary benefits derived further down the distribution chain through secondary and subsequent reuse, except to suggest that they are "very considerable."⁴⁶

To illustrate these subsequent benefits, Lind points to the ubiquity of GPS systems, used by nearly half of all Danish households, and reliant on the open address data set.⁴⁷ According to Lind, GPS companies interviewed by the Danish Geodata Agency report that they consider address data in Denmark the best in Europe. "We were able to improve the precision of car navigation systems in Denmark, and this has second-level impact on competitiveness and not using too much petrol."⁴⁸

NONFINANCIAL BENEFITS

The Danish emergency services, who were among the early users of the open address data set, introduced a system built on standard car GPS navigation systems into 1,200 emergency vehicles, in preference to more expensive mobile devices for alarm reception and tracking, with resulting significant reductions in cost, response times and fleet efficiency.⁴⁹ The opening of the address data also improved response accuracy for the emergency services within Denmark. Lind notes that the increased precision of the geocoded address set improved the capacity of emergency services to find the right address. "The public could be more confident that 1-1-2 [emergency] services would be able to find them, and ... we had a really great impact on that."⁵⁰

Not all anticipated benefits of the free-of-charge agreement were realized, however, although often these reflect a failure to embrace the data, rather than deficiencies within it. For example, the DECA analysis of the benefits of the address data set reported that adoption of the official address data set by the Central Business Register had been slow, and that in 2010 companies employing approximately 15 percent of the Danish workforce were registered at an address that was not in the official data set. As a result, DECA reports users found that it could take months or years for a new or changed commercial address to appear in their GPS. Given that emergency services as well as clients and suppliers rely on GPS to find commercial locations, the implications of such delays are potentially very costly.⁵¹

46 Lind, Morten. "The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge." Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

47 Lind, Morten. "The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge." Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

48 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

49 "112 in Denmark – a short overview of the Danish model." 112 Roundtable in Bulgaria, March 6, 2009. http://www.eena.org/ressource/static/files/2009_03_10_112_in_Denmark.pdf, p 9.

50 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

51 Lind, Morten. "The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge." Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

BENEFITS TO DEVELOPER COMMUNITY

The number of applications using national address data went from two or three in 2005 to over 500 by 2009, according to Lind.

“The multiplication of use was astonishing. There has been a lot of development ... in geomarketing and digitization, and each development shows new use cases for the address data. I think that’s the purpose of open data. Originally you collect it with a purpose – hopefully! – but when you put it out into the open, you will experience new usefulness of the data that you haven’t foreseen. Perhaps you will find that there are purposes that you haven’t even thought of, that have an even better business case [with] more benefit than you have estimated from your original [case]. By going into the open data agenda, you have to say to yourself: I don’t have the faintest idea what this can be used for! It’s not up to me to decide ... or to control.”

THE DANISH BASIC DATA PROGRAM

One of the most significant impacts of the release of the Danish address data was the impetus it gave to open other public data sets. Given the documented benefits of opening Danish address data, the Danish eGovernment strategy from 2011-2015 expanded the program with Good Basic Data for All,⁵² commonly known as the Danish Basic Data Program.

Basic data is essential information used repeatedly by the public sector in case processing, such as personal and company registration numbers, addresses, cadastral data and physical and political mapping data.⁵³ The Basic Data Program consists of a set of seven subagreements between multiple ministries and agencies, involving a total of nine registers, including business registers, the Cadastre, the Building and Dwelling Register, administrative and geographical boundaries, map and elevation data, place name and information register, and the register of property owners. A final subagreement covered the implementation of a common platform to disseminate the data from the other subprograms.⁵⁴

While personal data is protected by Denmark’s Act on Processing of Personal Data (2000), all other data within these registers was opened for access and reuse by individuals, the public sector and private companies as of January 1, 2013. Between 2013 and 2015, the agencies involved improved and expanded the nine official registers supplying basic data, with the aim of developing a “high quality common administrative foundation ... efficiently updated at (a) single site.” Implementing the Basic Data required the harmonization and conversion to a common format of all existing registers that would supply data; the addition of extra information to the basic data registers to better support public administration,

52 “Fact sheet – basic data.” Denmark Ministry of Finance. October 8, 2014. http://uk.fm.dk/~media/files/nyheder/press-releases/2012/10/basic-data/fact_sheet_basicdata.ashx?la=da

53 “Fact sheet – basic data.” Denmark Ministry of Finance. October 8, 2014. http://uk.fm.dk/~media/files/nyheder/press-releases/2012/10/basic-data/fact_sheet_basicdata.ashx?la=da

54 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

and the development of a common infrastructure to distribute the data, the Data Distributor.⁵⁵

The Basic Data Program did not come about as a result of public pressure, demand from potential users or a desire for transparency. A project manager at the Danish Agency for Digitization interviewed by Thorhildur Jetzek in 2012 noted: “We believed in [transparency] but it was not going to open any doors in Denmark. Because in Denmark we didn’t really have the feeling that our government wasn’t transparent enough. We never really seemed to regard that as a serious issue for Denmark.”⁵⁶

Instead, the Basic Data Program was developed as a result of a business decision to improve public sector efficiency and data quality, influenced by New Public Management reforms over the last two decades in which improving efficiency became the focus of efforts to control spending without increasing taxation.⁵⁷ Jetzek reports that initially there was resistance in some quarters to the idea of fully opening the data, rather than simply finding ways to improve the efficiency of its collection, management and use, while continuing with the prevailing model of offering data for sale, as mandated in Danish law. The decision-makers involved were ultimately persuaded to open the data by the experience of open data projects in the Netherlands, and by the work of Dutch analyst Marc de Vries, who argued that the income from data sales was typically a fraction of the actual costs of collection,⁵⁸ while the economic benefits of opening the data could be up to 17 times that cost.⁵⁹ The Pricing of Public Sector Information Study carried out by de Vries and others in 2011 for the European Commission concluded that only reducing charges to zero or cost-recovery increased data reuse by 1,000 to 10,000 percent, attracted new users such as small and medium enterprises, and increased economic activity, market dynamism, innovation, employment and efficiency, with little impact on cost.⁶⁰

The advent of Denmark’s Basic Data Program also coincided with a new government in 2012 that responded to the worldwide economic crisis with a focus on job creation and innovation. Although opening the data was a means to improve efficiency, rather than an end in itself, there was no conflict between efficiency, innovation, and open data ideals. Lind notes, “We could align the purpose of the Basic Data Program to make government more efficient with open data ideas [of] making the basic data open for the private sector and citizens to make smart things.”⁶¹

As intended, the Basic Data Program has provided opportunities for the private sector. Septima, a geodata startup, was founded around the same time as the Basic Data Program in 2013. It offers a range of products and services using the data, including data enhancement, such as calculating the distance between schools and houses; Web-based map components and search tools for public data; consulting to organizations like the Geodata Agency; and subcontracting to the large IT company involved in developing the distribution platform for the Basic Data Program.⁶²

55 “Fact sheet – basic data.” Denmark Ministry of Finance. October 8, 2014. http://uk.fm.dk/~media/files/nyheder/press-releases/2012/10/basic-data/fact_sheet_basicdata.ashx?la=da

56 Jetzek, Thorhildur (2015) Managing complexity across multiple dimensions of open data: the case of the Danish Basic Data Program. In press, pp 19-20.

57 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

58 See, for example, Sawyer, Geoff and Marc de Vries. “About GMEs and Data: Geese and Golden Eggs.” European Association of Remote Sensing Companies. December 2012. http://earsc.org/file_download/134/Open+Data+study+Final+report.pdf

59 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

60 de Vries, Mark, et al. “Pricing of Public Sector Information Study (POPSIS) Summary Report.” European Commission Information Society and Media Directorate-General. October 2011. <http://www.epsplatform.eu/sites/default/files/summary.pdf>

61 GovLab interview with Morten Lind, Danish Geodata Agency, September 15.

62 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

Bo Overgaard and Christian Fischer of Septima note that some of the company's work would have been possible without the Basic Data Program, but that the Program opened up possibilities for products and services that had not previously existed.

*"[The Basic Data release] was part of the basis of us establishing our company, that now we [could] do some interesting things with all the new open data sets." These new opportunities included the possibility of speculative development, without the need for a customer. "[The Basic Data Program] has been a great help to us. Before, our clients had to have the rights [to the data]. Now we can develop our products without having a customer. We can get the data, [make] some nice products and show them to potential customers. It also [enables us] to do different business models. Before, the license was connected to a user, and if you wanted to make a mobile application, you had to buy a site license, which could be very expensive."*⁶³

Since the establishment of the Basic Data Program, the private sector market for this data has grown. More companies like Septima have been established, and existing companies that had not previously worked with geodata have moved into the sector. According to Overgaard and Fischer, this development has spawned a lot of interesting solutions. It has become standard for insurance companies, for example, to use elevation data to assess risks of flooding where they would previously have had to rely on aggregated risk by region. Municipalities are also using this data to model and plan for flooding risk as sea levels rise.⁶⁴

Jetzek observes that the impact on the private sector is harder to quantify. "There are a few companies in Denmark that you can say were founded on the basis of the Basic Data Program, but I think most of the use is, and will continue to be, companies that are using the data and a thousand other things in their daily processes. For most companies, it's hard to say exactly what effect it would have if the data were not available, or if [they] were more expensive."⁶⁵ She reports that the Danish Geodata has been attempting to track the impact of its data releases over time, with a baseline survey conducted in the year the Basic Data Program began, and follow-up surveys planned for future years.⁶⁶

The frequency with which data sets are updated varies according to the register. Some elevation data relies on aerial surveys that might be a decade old. Data in the address data set, by contrast, is continuously updated as new addresses are added by local authorities. It takes seconds after addresses are locally created for the information to become available as part of the national data set. "Fifteen seconds after an address is recorded [by] a local authority, it is available in our service, and that's very important if you are making e-government for GPS. A delay of one to two weeks would be a problem. For e-government, if you go to one part of government and make an address it should be available instantly for the next purpose."⁶⁷

63 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

64 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

65 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

66 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

67 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

IV. Challenges

DATA QUALITY

Once the decision to open data for the Danish Data Program was made, the challenges became those of data quality and governance across the range of agencies and registers. Jetzek notes that the nature of the data collection has created significant variability in quality.

“In Denmark, data collection is very distributed, and in most cases the data is just a part of what they do: a necessary part, but not the main part. Making some standards or blueprints for the organizations to follow when they’re going to model or open the data is very important. Some agencies like the Geodata Agency are very professional data custodians: They’ve done it for a long time, and there are standards across Europe, and there’s a lot of knowledge and professionalism. ... Whereas others ... don’t even know what the data model is. The data is just something that someone is keying in somewhere, and then they have this register. It’s very important to have some common understanding and some common skills about how ... we handle all this data ... [but] this skill [isn’t] common in the public sector. We haven’t really focused on data as a key resource.”⁶⁸

Overgaard and Fischer observe that accuracy appears to be higher in areas where data collection is directly connected with government revenue by comparison with areas in which data is an incidental by-product of another activity. “If money is directly involved, the quality tends to be good. The building and dwelling register, for example, is involved in taxation, and it’s pretty precise.”⁶⁹ The business registers, by contrast, which are not directly involved in revenue collection, are much more error-prone. Both Jetzek and Overgaard and Fischer cite weaknesses in the register of companies where the nature of the business is described: The company itself is listed, but the nature of its commercial activity is wrong. “In the business register, the businesses are supposed to say what they do, to classify the business, but you can look at the register and see that a pizza shop is classified as a shoemaker because the house used to belong to a shoemaker, and then it changed ... and nobody [corrected] the data.”⁷⁰ Some uses that Septima had hoped to make of the data from the business register, for example, turned out not to be possible because of deficiencies in the quality of the data.⁷¹

Having been open longer, the address data set has had longer to address these quality concerns. The aim of the address data subprogram was “to further improve the quality and accessibility of address data, and enrich ... the infrastructure behind the distribution of address data.”⁷² The Geodata Agency

68 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

69 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

70 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

71 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

72 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

has also collaborated with OpenStreetMap to identify and correct address errors in the data set. One of the greatest boons to accuracy, however, is simply opening the data and having it used, which provides opportunities for scrutiny and correction – what de Vries calls a partial outsourcing of quality control⁷³ – and raises the quality of the data set.⁷⁴ “Address data is very good in Denmark, and that’s because so many are using them. If no one uses the data, then no one finds the errors,” note Overgaard and Fischer.⁷⁵ Centralization of registration also allows for easier and faster correction, since users only have to report errors and omissions once,⁷⁶ through a portal that connects users with the responsible municipality.⁷⁷ With the Basic Data Program, the address data set also “took an extra step toward the open,” according to Lind, by removing the earlier distribution costs and making the data easier and cheaper to access and use. “Every citizen can now download a data set of every address free of charge, or access our services for a Web application.”⁷⁸

GOVERNANCE OF THE DATA PROJECT

Project management issues also arose over the governance of the program, which was initially similarly distributed and collaborative. The number of entities involved in the governance of the project created major complexities and management challenges. Delays occurred when the development and implementation of the Data Distributor was much more complicated than originally expected, overran its timeline and began to impede progress elsewhere. “This kind of project needs to be handled like infrastructure,” says Jetzek. “If you want to have an initiative that covers a lot of different data sets, it [is] really important to think about how you govern this. It’s hard to do it top-down; it’s also almost impossible to do it bottom-up. There has to be a middle way.”⁷⁹ Jetzek notes that, in response to the difficulties encountered, the governance structure of the Basic Data Program was altered to make it “less democratic.”⁸⁰

73 de Vries, Mark, et al. “Pricing of Public Sector Information Study (POPSIS) Summary Report.” European Commission Information Society and Media Directorate-General. October 2011. <http://www.epsiplatform.eu/sites/default/files/summary.pdf>

74 de Vries, Mark, et al. “Pricing of Public Sector Information Study (POPSIS) Summary Report.” European Commission Information Society and Media Directorate-General. October 2011. <http://www.epsiplatform.eu/sites/default/files/summary.pdf>

75 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

76 Lind, Morten. “The value of Danish address data: social benefits from the 2002 agreement on procuring address data etc. free of charge.” Danish Enterprise and Construction Authority. July 7, 2010. http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf

77 GovLab interview with Bo Overgaard, Director, Septima, and Christian Fischer, Senior Advisor, Septima, September 22, 2015.

78 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

79 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

80 GovLab interview with Thorhildur Jetzek, Department of IT Management, Copenhagen Business School, September 21, 2015.

V. Looking Forward

EXPECTED BENEFITS OF THE BASIC DATA PROGRAM

At launch, the expected benefits of the Basic Data Program were efficiencies to government of about EUR 34.8 million from 2020 through decreased administrative costs, of which two-thirds of the benefit (EUR 23.5 million) would be to municipalities. The release of the data was also expected to generate returns in the private sector of up to EUR 67 million, through the encouragement of new digital services and products, and the removal of bureaucratic barriers and data costs. Finally, individuals would benefit from increasingly improved and integrated public services.⁸¹

ADDRESSING CAPACITY ISSUES

Lind identifies secure funding as the main future challenge of the Basic Data Program and projects like it. “One of the issues with open data is that you have to have a stable, sustainable source of funding ... to be able to update, and to ensure the long-term quality of the open data. We [were] lucky that the Danish Basic Data Program came at a time when we needed funding of this kind. If we could have done anything differently or better, we should have focused on getting better [long-term] funding. This reorganization of government process we have just been through shows how easily things can change, and suddenly you can be without any source of funding.”⁸²

For Jetzek, the major future challenge is developing skills in data collection among public servants. “Getting the data to be really clear and easy to use, so that it’s used widely across the public and private sectors and in business-critical processes, is a huge accomplishment on its own. If you want the data to be really used across the public and private sectors for more than just creating an interesting visualization or one-off something, I think you really need to focus on [data collection].”⁸³

The Danish effort to consolidate and share address data has already had major impacts on the country, particularly in inspiring a much larger effort to open and consolidate all types of public data. To enable ongoing impacts, though, work must be done to engender readiness – in the form of adequate funding and employee training – within the public service, a common theme across global efforts to create impacts through the opening of data.

81 “Fact sheet – basic data.” Denmark Ministry of Finance. October 8, 2014. http://uk.fm.dk/~media/files/nyheder/press-releases/2012/10/basic-data/fact_sheet_basicdata.ashx?la=da

82 GovLab interview with Morten Lind, Senior Advisor, Danish Geodata Agency, September 15.

83 GovLab interview with Thorhildur Jetzek, September 21, 2015.