The value of open data

– Economic impact assessment of open public sector information
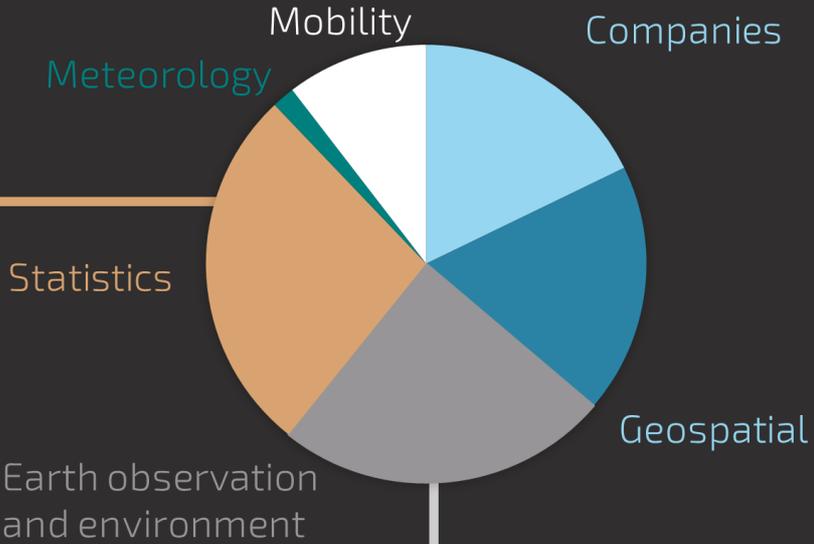
05.06.20
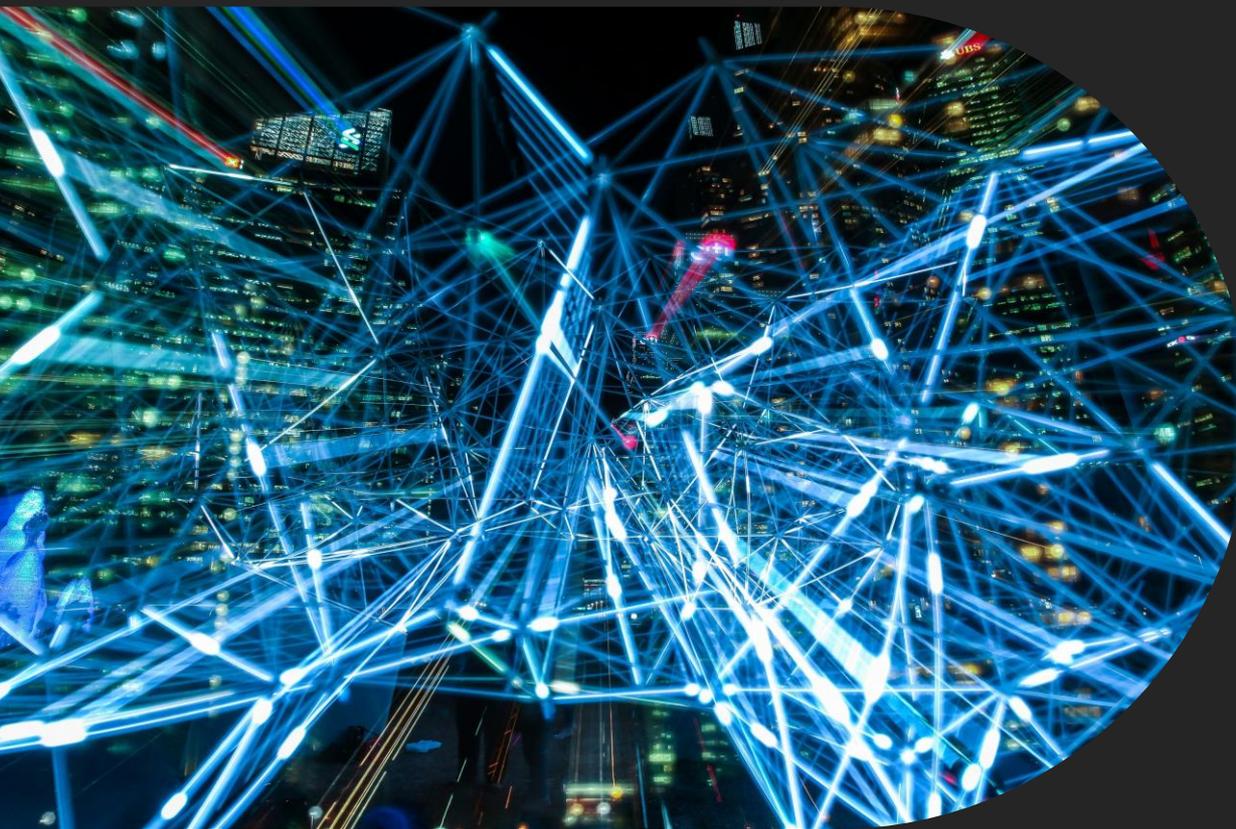
The value of open data

– Economic impact assessment of open public sector information

Daniel Ekström

Jesper Johannesson

6 Categories

About 300 datasets

Mobility

Meteorology

Companies

Statistics

Geospatial

Earth observation and environment

The meta analysis show socio-economic benefits that amount to

11,1
billion SEK for four data categories

32% subject to charge

5,4 billion SEK

2,2

2,7

0,8

Geospatial data

Address data

Company information

Geological information
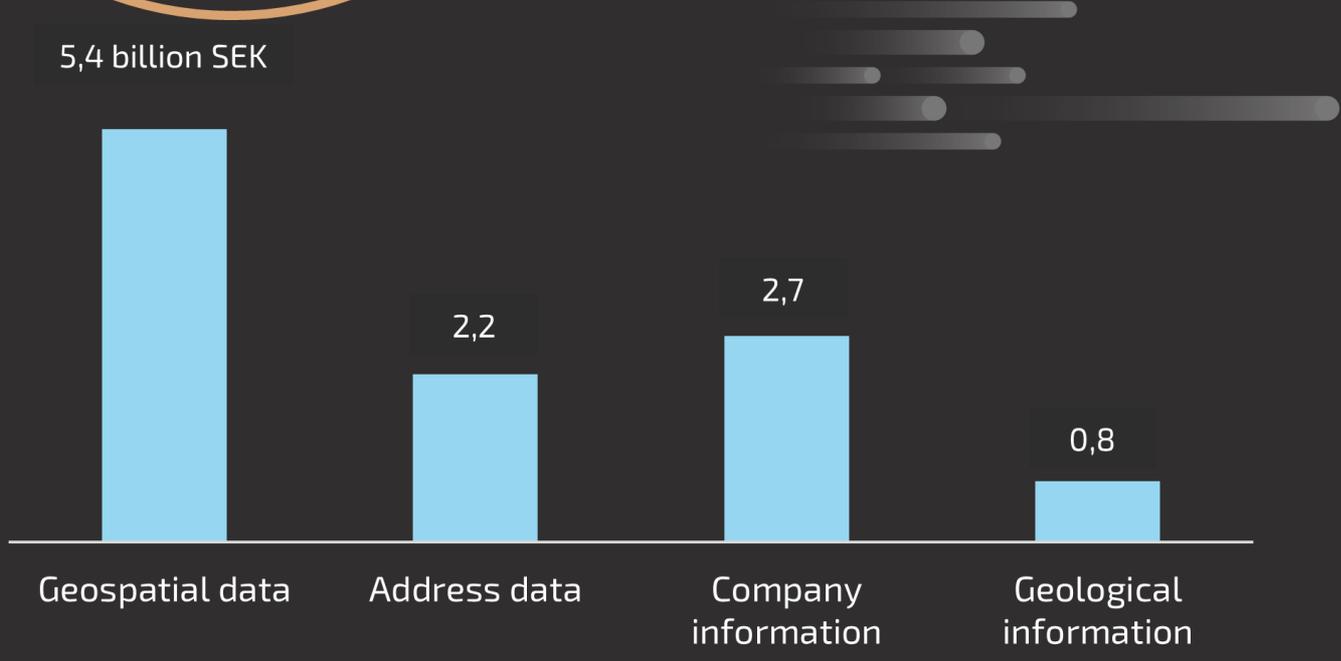
Agriculture and forestry

(billion SEK)

1,1 – 2,9

ICT

2,6 – 6,4

Bank & Finance

2,0 – 3,4

Planning and real estate

3,0 – 6,2

Public Sector

1,1 – 1,7

The case-based analysis shows socio-economic benefits of

10 – 21

billion SEK across five sectors

Summary

Introduction

DAMVAD Analytics was appointed by Lantmäteriet (The Swedish mapping, cadastral and land registration authority) to analyse the impact of providing open access to several suggested datasets. The purpose of this analysis is to describe how providing open access to the datasets results in socio-economic utility. By combining a meta-analysis based on experiences from comparable countries with an impact analysis based on case studies, we describe the value of open access to the proposed data as well as the impact on specific sectors.

The report is part of a government assignment assigned to Lantmäteriet in May 2019. The suggested datasets have been defined by Lantmäteriet, in cooperation with relevant government actors, as being of particularly high value. The government assignment is part of Sweden's preparations for the so-called PSI- directive (EU) 2019/1024 on open data and the re-use of public sector information. Through the new PSI Directive, countries will jointly define particularly valuable datasets which are to be published without fees and with minimal restrictions.

Through the government assignment, a proposal of which datasets to consider as being High-Value Datasets (HVDs) from a Swedish perspective has been prepared. In total, 298 datasets within six categories have been identified. To what degree the proposed datasets are currently available (March 2020) varies. Therefore, part of the study is concerned with an analysis of the proposed datasets. The impact analysis shows the benefit of providing open access to the datasets currently (March 2020) not free of charge.

This study is, to the best of the author's knowledge, the most extensive impact assessment of open data in Sweden. A large number of representatives of public organisations, businesses and interest groups, as well as open data experts, have been interviewed. In total, sixty interviews were conducted during the study. The base for the meta-analysis consists of a dozen impact assessments from neighbouring Nordic countries. At the same time, some 140 case studies were analysed within the scope of the case-based impact study. Further, a comprehensive material of internal official documents and other sources of information are included in the analysis.

This is a summary of the findings from the full report, available in Swedish via DAMVAD Analytics and Lantmäteriet.

The value of open data

The value of open data has been subject to intense debate over the past decade, and free access to public sector data has previously been associated with great value. The increased interest in re-using public sector information can be related to the emergence of data as an important and strategic resource in the digital economy. Despite a comprehensive international debate, only a few previous Swedish studies concerning the aggregated value of open data have been conducted.

Because of various definitions of open data and differences in evaluation methods used to quantify the benefits, the results from previous studies vary. In general, a distinction can be made between studies focusing on direct impact and studies analysing the dynamic effects of open data. Impact assessment concerning the dynamic effects generally identifies higher values since they also consider

indirect benefits in the form of new products and services, developing as a result of open data access. Moreover, there is a significant difference between studies with regard to how conservative, respectively optimistic, the definitions and estimates used to calculate benefits are.

The outline of the study

Within the scope of this report, a meta-analysis and a case-based impact analysis is conducted. The meta-analysis is based on realized values from open data in countries comparable to Sweden. The case-based impact analysis is based on conservative estimates of direct and indirect values following from providing open access to the proposed datasets. The following paragraphs briefly outline the set-up of the study.

The meta-analysis aims to present an estimation of the socio-economic value of four datasets – *geospatial data*, *address data*, *company information and geological information*. The study is mainly based on previous assessments, conducted after the data was made available in Sweden's neighbouring countries. The results from the studies are recalculated according to Swedish conditions.

The case-based impact analysis describes how providing access to the proposed datasets may affect different sectors. The analysis accounts for direct as well as indirect values and aims to describe the complex dynamics through which open data create socio-economic value. The study mainly consists of a number of interviews and literature studies and is conducted in three steps:

- **Identification** – 140 different applications of open data are identified and evaluated.
- **Evaluation** – a number of case studies are carried out to identify how the application of the proposed datasets can contribute to efficiency and increased productivity.
- **Aggregation** – calculating the potential of the identified applications on a national level and compared to the results from other studies.

Identify > Evaluate > Aggregate

Analysis of the proposed datasets

In total, 298 datasets have been defined as particularly valuable by the joint work led by Lantmäteriet. The proposed datasets are sorted into six different categories: *companies, geospatial data, earth observations and environment, statistics, meteorology and mobility*. As a first step, an analysis of these datasets was conducted to understand how they may be impacted by open access.

To what extent the proposed datasets are currently available varies and is visualised in the figure below. The analysis shows that usage of approximately a third of the proposed datasets is restricted by fees. There are however significant differences between the categories; where statistics, mobility and metrological data, to a higher degree, are free of charge. Fees are most common within the

categories; *companies*, *geospatial data*, and *earth observations and environment*. The impact analysis is mainly focused on the added value from removing fees and providing open access to these datasets.

The proposed datasets and their current accessibility

Available free of charge   ■ Not available free of charge
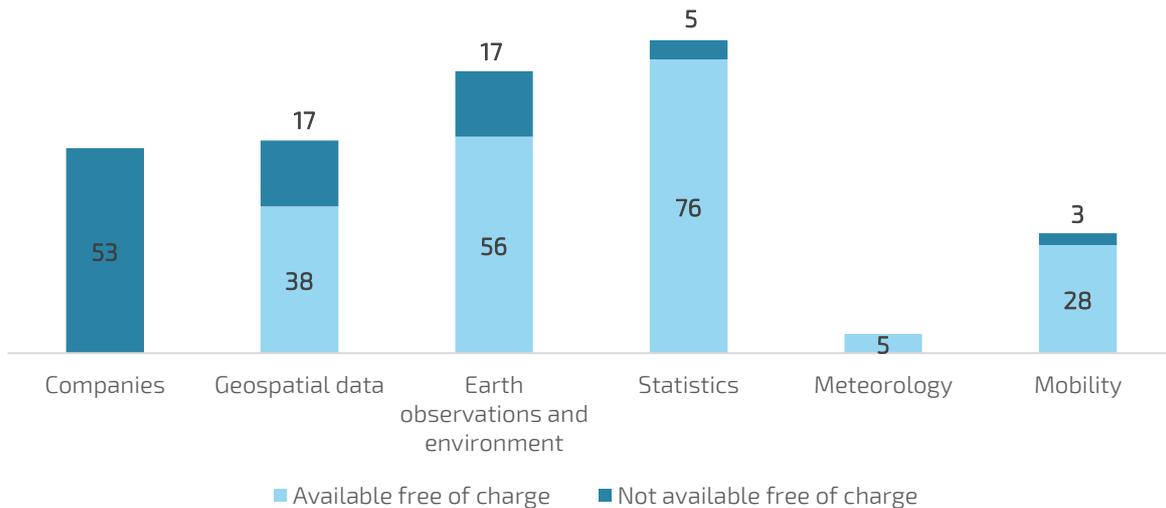
Meta-analysis

Several neighbouring countries to Sweden have already provided open access to datasets similar to those identified as HVDs by Swedish authorities. In many of these countries, socio-economic studies have been conducted both before and after providing open access to the data. Through a thorough examination of the results from these studies, lessons regarding the benefits of providing open access can be learnt.

Based on studies from comparable countries, a meta-analysis is conducted focused on the value of four specific datasets*: geospatial data*, *address data*, *company data* and *geological data*. These sets of data are currently, to a high degree, restricted by fees in Sweden and providing them as open data is expected to result in significant socio-economic impact. By calculating the realised effects from providing open access to these datasets, the estimated value adds up to 11,1 billion Swedish krona (SEK).

Geodata – including for instance; maps, aerial and orthophotos, positioning data and property registers – is estimated to contribute with 5 4000 million SEK to the Swedish economy when data is provided free of charge. Geospatial data is of particular value for the public sector, estimated to 3 400 million SEK. The data is also valuable for trade and industry, estimated to 2 000 million SEK.

Address data – including postcodes, are currently not open in Sweden. Providing open access to this register alone has the potential to generate a value of about 2 200 million SEK, where efficiency improvements within the public sector make up 1 500 million of the estimated value.

Company information – is a data category which is valuable to both the public and private sector. Providing open access to this information is estimated to result in a total value of 2 700 million SEK.

Geological information – primarily includes rock, soil and groundwater data from the Geological Survey of Sweden. These datasets are mainly used within the *mineral sector, groundwater-, environmental- and climate issue*s as well as research – and are estimated to a combined value of 800 million Swedish crowns.

The economic value of the four specified datasets, million SEK

	Public Sector	Business	Total
Geospatial data	3 400	2 000	5 400
Address data	1 500	700	2 200
Company information	-	-	2 700
Geological information	200	600	800
Total	-	-	**11 100**

Overall, the meta-study points to significant values in both the private- and public sector. The studies from Sweden's neighbouring countries indicate that providing open access to the data leads to an increase in the re-usage of data. In the private sector, providing open access to data leads to efficiency improvements and development of new business areas, resulting in faster growth for companies utilising open data compared to companies that do not. In the public sector, the usage of the data also increases through open access. The increased utilisation of data does not only occur as a result of removing fees but also because of the data being more accessible. In the public sector, values are foremost a result of labour savings.

Case-based impact analysis

As a complement to the estimated values from the meta-analysis, a case-based impact analysis is conducted. This analysis is focused on how the datasets currently restricted by fees may impact five different sectors. In contrast to the meta-analysis focusing on the realised values from our neighbouring countries, the case-based impact analysis estimates future values that providing open access data can contribute to. Open data is often a catalyst for innovation, which is why it is interesting to consider the indirect dynamic values which can arise when new business opportunities emerge.

*Dynamic effect*s entail situations when open data enables a transition of organisational and business processes. Open access public information data are often a prerequisite for digital services and advanced analyses. Many industries are facing digital transformations where open access to public data can make way for new innovative solutions, including AI applications.

The study points to 20 applicable areas within five different sectors where the added value of providing open access to the proposed data amounts to 10-21 billion SEK. The proposed datasets are

estimated to have the most significant potential within the ICT-sector and the real estate sector, where providing open access to the proposed data, for instance, enables more efficient IT-systems and planning processes. Overall, the analysis shows that significant values arise from increased use and combining of datasets, leading to both efficiency improvements and more informed decision-making.

Economic value of the specified datasets in five sectors, million SEK
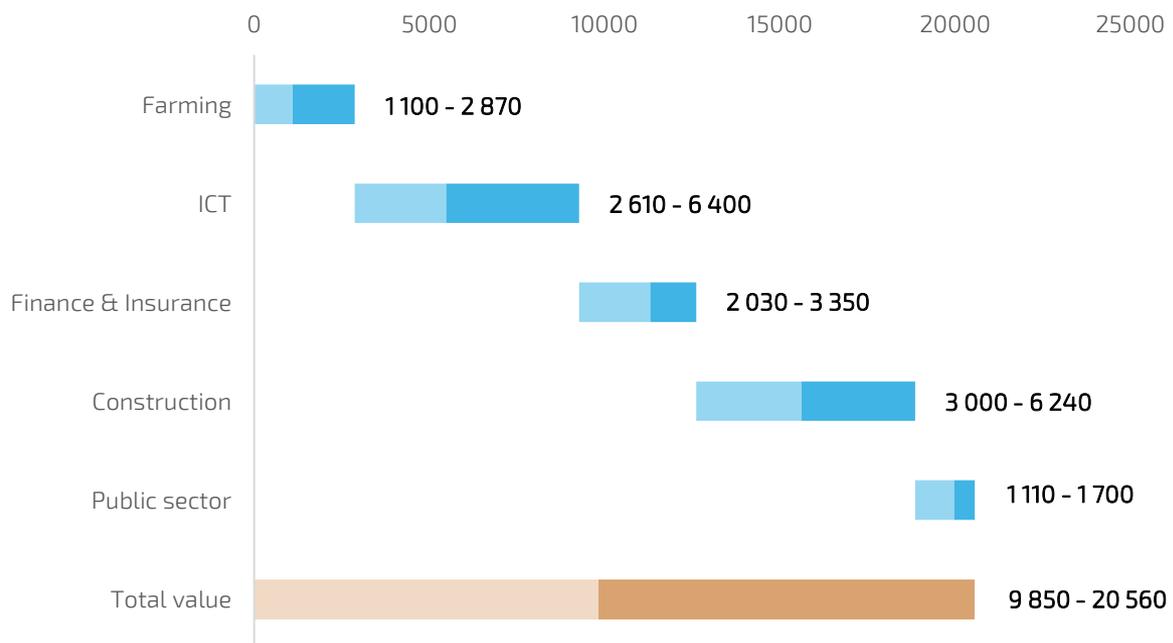
Farming – Within forestry, the proposed data will allow for improvements in planning, inventory and logging through labour savings and increased productivity. In agriculture, geospatial data and positioning data enable efficiency gains. The most significant societal benefit is achieved through more efficient usage of agricultural and forestry land, where the data increases the opportunities for joint cultivation, land exchange and real-estate transactions.

ICT – In the analysis of the ICT sector, significant value is identified from improved opportunities for analysis and more efficient operations of IT-systems. As a result of fees and technical obstacles, companies are currently forced to spend significant resources on compiling data. Increased access can also support start-ups through simplifying the process of taking an idea to *proof of concept* and further to investments. The study also suggests growth and increased plurality among business information companies resulting from providing open access to the proposed data.

The finance and insurance industry – The analysis points to four particular areas of use adding value within the finance and insurance industry. For instance, values arise through efficiency gains in the prevention of insurance frauds, improved risk analyses and efficiency gains in IT-system operations.

Construction – The proposed data can contribute to virtually all parts of the construction sector, from planning to the construction and maintenance of buildings. This is, for instance, achieved through

more efficient methods of analysis, significantly improving various planning processes. Data can also lead to better knowledge about potential clients, which enables more efficient usage of the existing property portfolios as well as demand-driven planning.

Public sector – Gains arise as a result of diminished costs of administration and IT-systems through improved coordination of IT-systems. If the proposed data is made available without charge, data purchase from other authorities will naturally decline. Improving the accessibility to data could also contribute to improved public services, including more efficient crisis management.

Environmental effects

Out of the proposed data, property data, geodetic data, earth data, species sightings, endangered animals and plants, satellite images as well as image and height data, could, in particular, contribute to the work towards achieving environmental and climate goals. Providing access to the proposed datasets contribute to the national and international efforts towards these goals. Domestically, the effects can be seen, for instance, in the form of preserving biological diversity, increased efficiency gains regarding labour used in areas such as forest and land as well as improved opportunities for individuals to access the nature and its resources. Internationally, the effects can be seen in the form of simplified collaborations and increased access to information since data from different countries can be combined to solve the global environmental and climate challenges. Therefore, the proposed datasets make up an essential piece of the puzzle in solving global environmental and climate issues.

AI

The proposed datasets have potential usage within artificial intelligence with several characteristics making them particularly valuable – for instance, the fact that they describe valuable resources such as companies, estates and vehicles.

Several factors make data particularly useful in AI-algorithms. Through the PSI Directive, data is to be provided with minimal restrictions and in a machine-readable format - a prerequisite for many different AI applications. The proposed datasets contain detailed and updated information which can also be connected to other datasets. Overall, the proposed datasets are suitable for many different AI applications entailing good prospects of providing access to the data contributing to the Swedish development of AI.

Factors contributing to the proposed datasets being suitable for AI applications

Linkable	Detailed	Up-to-date	Machine-readable

1   References

Literature

ACIL Tasman, *The value of spatial information*, 2011

ACIL Tasman, *The value of spatial information*, 2018

ASEDIE, *2019 Infomediary Sector*, 2019

Bhagwat, S.B., Ipe, V. C., *Economic benefits of detailed geologic mapping of Kentucky*, 2000

Capgemini Consulting, *Creating value through open data*, 2015

Capgemini Consulting, *The economic impact of open data – Opportunities for value creation in Europe*, 2020

ConsultingWhere & ACIL Tasman, *Assessing the value of OS OpenData to the Economy of Great Britain*, 2013

Craglia, M., Campagna, M., *Advanced regional spatial data infrastructures in Europe*, 2009

Damvad Analytics, *Erhvervslivets brug af Kortforsyningen*, 2019

Damvad Analytics, *Måling af brugen og værdiskabelsen af Danmarks Adressers Web API (DAWA)*, 2019

Damvad Analytics, Værdiberegning af DMI's data – Effekt på samfundet og den grønne omstilling, 2019

Data Foundation, Standard Business Reporting: Open Data to Cut Compliance Costs, 2017

Deloitte, *Market assessment of public sector information*, 2013

Deloitte, *Effekten af de frie geografiske grunddata*, 2014

DIGG, Främja den offentliga förvaltningens förmåga att använda AI, 2020

The European Commission, European innovation scoreboard, 2019

FOI - Totalförsvarets forskningsinstitut, *Klimatförändringar, migration och konflikter: samband och förutsägelser*, 2011

Gartner, *En rapport för Regeringskansliet – Öppna data och datadriven innovation*, 2018

Häggquist, E., *The economic value of use of geological information*, 2017

Internationella Röda Kors- och Röda Halvmånefederationen, *The cost of doing nothing*, 2019

Koski, H., *The impact of open data – a preliminary study, Ministry of Finance (Finland)*, 2015

Koski, H., *Does marginal cost pricing of public sector information spur firm growth*?, 2011

Lantmäteriet, *Delredovisning av uppdraget "Effekter och konsekvenser av öppna data"*, 2017

Lantmäteriet, *Ekonomisk nytta av ett samlat nationellt tillgängliggörande av geodata i samhällsbyggnadsprocessen*, 2019

Lateral Economics, *The economic value of data assets under alternative policy regime*s, 2016

Lakomaa, E., *Samhällsekonomisk effekt av öppna geodata*, 2016

Lakomaa, E. & Kallberg, J., *Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneur*s, 2013

McKinsey & Company, *An AI Nation? Harnessing the opportunity of artificial intelligence in Denmark*, 2019

McKinsey & Company, *Open data: Unlocking innovation and performance with liquid information*, 2013

Myndigheten för digital förvaltning (DIGG), *Främja den offentliga förvaltningens förmåga att använda AI*, 2019

OpenELS, *The Socio Economic Impact of Open ELS*, 2018

PwC Danmark, *Effekten af de frie geodata – Eftermåling*, 2017

PwC Sverige, *Öppna data i Sverige*, 2014

Shakespeare, S., *An independent review of public sector information*, 2013

Skogsstyrelsen, *Nytta av öppna data*, 2016

SOU, *Till sista utposten – En översyn av postlagstiftningen i ett digitaliserat samhälle 2016:54*, 2016

SOU, *Framtidens miljö: Allas vårt ansvar SOU 2000:52*, 2000

Spatineo, *The economic value of spatially enabled services in Finland -including the impact of the Geospatial Platform*, 2018

Statskontoret, *Hinder för att använda myndigheternas öppna data*, 2018

Svensk Försäkring, *Försäkringar i Sverige 2019*, 2019

Svensk Försäkring, *Försäkringsmarknaden – Kvartalsvis redovisning av statistik, Kvartal 4 2019*, 2020

Sveriges Kommuner och Regioner, *Vidareutnyttjande av offentlig information*, 2012

SVT, *Sveriges osäkra Kurvor*, 2019

Sweco, *Modell för effektuppföljning av SGU:s geologiska information*, 2011

Söderström et al, *Avstånd mellan körspår – en jämförelse mellan traditionell spårmarkör och autostyrning med GPS*, 2008

Tillväxtanalys, *Företagens digitala mognad*, 2019

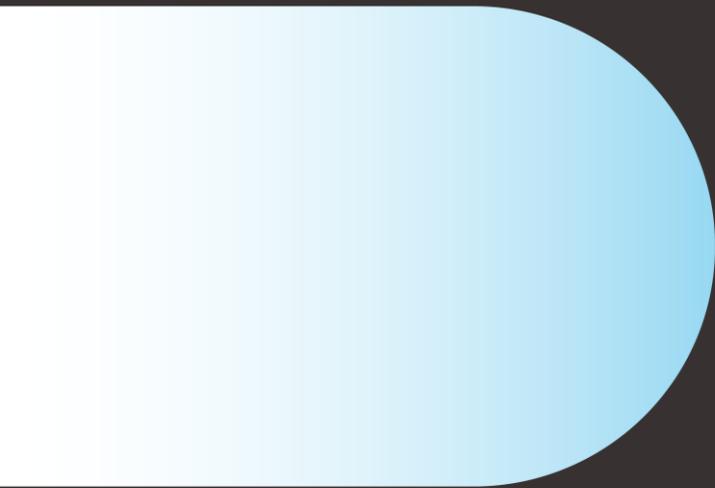Transportstyrelsen, *Användarundersökning uttagswebanvändare 2017*, 2017

Vista Analyse, *Verdien av gratis kart- og eiendomsdata*, 2014

WIPO, The Global Innovation Index 2019, 2019

WSP, *Samhällsekonomisk analys av geologisk information*, 2019

Interviewees (organisation)

Björn Hagström (konsult), Peter Mankenskiöld (TietoEVRY), Angela Yong (Tillväxtverket), Björn Lovén (Rymdstyrelsen), Lars Kristian Stölen (SGU), Annika Kindeberg (Sjöfartsverket), Johan Winell (Sjöfartsverket), Erik Lakomaa (Handelshögskolan), Elisabeth Häggquist (PTS), Anna Hermansson (Dataväxt AB), Lawrence Kay (Open data institute), Svante Eriksson (Governo), Christian Nordenskjöld (LRF), Jonas Svensson (Södra), Amir Mirbashi (SMHI), Marcus Flarup (SMHI), Beatrice Sablone (Storsthlm), Pierre Mesure (Civic Tech Sweden), Kristine Ulander (DIGG), Joakim Nyström (Bolagsverket), Marie Haldorson (SCB), Anders Frick (journalist), Claes Radojewski (Mobilityxlab), David Almstrom (Voi), Samir Sandberg (PwC), Ann Mohlin (Hack for Sweden), Marianne Leckström (SKR), Lars Olov Hjärp (Trafikverket), Christoffer Börjesson (Fastighetsägarna), Kristoffer Skjutare (CLOSER), Torsten Linders (Ocean Data Factory), Erik Borälv (Vinnova), Jonas Engström (RISE), Erik Willén (Skogforsk), Thomas Olsson (RISE), Martin Solberger (Svensk Försäkring), Emil Hagman (Skanska), Erica Olivius (Bisnode), Tanja Keisu (Biometria), Jenny Greberg (Svemin), Pär Nygårds (IT- och Telekomföretagen), Henrik Hermansson (Tillväxtanalys), Magnus Krantz (Bolagsverket), Ryo Yamazaki (Transportstyrelsen), Kerstin Konitzer (SGI), Simon Andersson (Spacemaker).

DAMVAD • ANALYTICS

DAMVAD • ANALYTICS