Data Sharing Practice Examples

A Digest of Practice Examples published on the Support Centre for Data Sharing

Overview

To inspire the data sharing community, the Support Centre for Data Sharing (SCDS) selects and publishes practice examples of initiatives that successfully overcome the barriers to data sharing. The examples cover different sectors from Agriculture to Health and originate from across Europe and beyond. Despite their differences, there are significant parallels in how the organisations created solutions to share data or support data sharing. When retrospectively looking at the practice examples from the last 12 months, common challenges and approaches to address them become evident.

The leading theme of the examples is platforms as trusted third parties/ intermediaries that provide support in data sharing governance and legal frameworks to its members coming from selected interest groups. Predominant aims are to increase efficiency and facilitate research.

A main challenge in data sharing is trust, especially in the absence of an overarching legal framework. Therefore, trust in building solutions that emphasise data sovereignty are not unexpected. Trusted third parties are a successful approach. They can, however, increase the risk of limiting data liquidity and encouraging silos. The creation of a joint terminology and joint standards within the initiative that differ from other initiatives can become a silo, limiting interoperability between different stakeholders. A focus in the future is to link stakeholders to further facilitate sector-wide and general governance and data sharing standards, whilst still harnessing the advantages of situation-specific, decentralised data sharing and fragmentation.

In order to support the reader in finding quickly the Practice Examples at the crossings of their interests, the table below maps them against the sector they focus on, and the role the organisation that was interviewed is covering in that sector.

We have identified three different types of role:
- **Data space catalysts**: these are organisations aimed at making spaces for data sharing happen, by addressing a sector or context holistically, generally or targeting one specific territory (e.g. JoinData for agriculture in the Netherlands, or Maas Madrid for the Spanish city's mobility);

- **Policy / legal facilitators**: these are organisations whose focus is addressing the legal challenges of sharing (e.g. Technology Industries Finland), or businesses that, through their experience, found themselves offering insight in that space (e.g. IBM); and

- **Technology enablers**: mostly unrelated to one sector or another, these are organisations whose effort is spent towards building the technology or advising in relation to the technology that makes data sharing possible (e.g. Ctrl-Shift).

Sector	Data space catalysts	Policy / legal facilitators	Technology enablers
Agriculture	• JoinData (B2B)		• CDQ • Ctrl-Shift • DataVaccinator (B2B) • NIIS (G2G and G2C) • SODA
Government	• NIIS (G2G and G2C)		
Manufacturing		• Technology Industries Finland (B2B)	
Mobility	• MaaS Madrid (X2X)		
Research	• Social Science One (B2Academia)	• Ctrl-Shift	
Transport and Logistics	• iShare • iShare Foundation		
No sector focus	• BDVA	• IBM Watson	
Other	• Data Pitch		

In the table below, the Practice Examples are summarised in the order by which they were documented and are described holistically. Following that, each example is added in full, as published on eudatasharing.eu/data-sharing-practice-examples.

Name	Sector	Country	Data Sharing model	Impact
1. MAAS Madrid	Mobility	Spain	Private and public organisations in Data Sharing Partnership offering their services jointly to citizens.	MAAS Madrid enables more efficient urban transport, reducing pollution.
2. JoinData	Agriculture	Netherlands	Farmers, suppliers, service providers share data using a platform as mediator/TTP while farmers decide how and by whom their data can be used.	JoinData facilitates insight creation for R&D for more efficient and sustainable farming.
3. NIIS	Government	Finland and Estonia	Finland and Estonia share public data using a jointly founded company and software solution (X-road) as an intermediary to overcome interoperability and governance challenges.	NIIS aims to facilitate seamless digital public services for citizens.
4. Ctrl-Shift	No sector focus	No country focus	An intermediary to support compliant sharing of personal data to realise opportunities for innovation and growth.	Ctrl-Shift is supporting organisations in value creation through personal data sharing and citizens in creating fairer value exchange for their data.

5. Data Pitch	No sector focus	No country focus	Public and private organisations share data with SMEs and Start-ups to "practice" data sharing and open innovation as the new way of working.	At Data Pitch, the data users, SMEs and start-ups use the data made available to create solutions for predefines challenges.
6. Technology Industries Finland	Manufacturing (Electronics, electrotechnical, mechanical and metals industries)	Finland	Support for data sharing by developing model terms and conditions for data sharing, addressing the limitations of nondisclosure agreements and intellectual property rights.	Companies have a base or addendum for their data sharing contractual agreements that e.g. allows them to retain learnings originated from using the shared data.
7. SODA	No sector focus	No country focus	A software intermediary that enables privacy-preserving analytics of information from multiple data assets using multi-party computation techniques.	These techniques allow multiple parties with private inputs to jointly compute data without revealing the input to each other.
8. Social Science One	Research	No country focus	Social Science One acts as an independent intermediary between researchers and the data provider, forming a data sharing partnership.	Social Science One enables academics to gain access to and analyse information from the private industry in a manner that is responsible and socially beneficial.
9. BDVA	No sector focus	No country focus	Research association of members from academia, public bodies, industries and private citizens that raises awareness and promotes the value of data.	BDVA's focus lies on challenges and best practices of data sharing, focusing on governance and legal questions.
10. iShare	Transport and Logistics	Netherlands	Collaboration between the Ministry of Economic Affairs and the Ministry of Infrastructure and the Environment as a trusted decentralised intermediary.	iShare provides a trusted data sharing framework to increase efficiency in logistics reducing congestion and preserving data subject's sovereignty.
11. iShare Foundation	Transport and Logistics	Netherlands	Collaboration between the Ministry of Economic Affairs and the Ministry of Infrastructure and the Environment as a trusted decentralised network facilitating data sharing governance within the network.	iShare Foundation established joint standards for its members to facilitate sure and trusted data sharing. They provide identification, authentication and authorisation of its members.

12. IBM Watson	No sector focus	No country focus	IBM's Watson processes clients' data outside of the borders of the office using AI to extract insights.	IBM aims at finding legal solutions for the novel cchallenges around intellectual property rights to learnings from processing confidential or personal data in Watson.
13. CDQ	No sector focus	No country focus	A trusted third party that manages and provides crowd sourced, open and shared commercial data for its members.	CDQ minimises the efforts to maintain and update clean business partner data, especially for large international enterprises by joining forces and sharing the data with the community.
14. DataVac cinator	No sector focus	No country focus	DataVaccinator is an open source software for secure and safe data storage and personal data protection using pseudonymisation of personal data at the point of its generation.	DataVaccinator facilitates sharing of pseudonymised data that is stored separately from its identifiable attributes. It also protects data against abuse or errors that lead to data breaches violating GDPR.

1   Maas Madrid: Data sharing for smart mobility

Link: https://eudatasharing.eu/examples/data-sharing-smart-mobility

"We aim to improve access to all public transport and shared mobility offers holistically with the ultimate goal of convincing people that there is no need to use a private car"

Sergio Fernández Balaguer, Project Manager of EMT Madrid

Mobility in major cities
In every city across the world, there are numerous public and private mobility companies that specialise in different services or transport "modes" – i.e. subway, buses, trains, trams, bicycles and cars. This is especially true in metropolitan cities such as Madrid, Amsterdam, Prague or Helsinki.

Several studies have found that in most metropolitan cities many people are not aware of the different public transport offers in their city. For example, the Spanish association for public transport operators – [Asociación de transportes públicos urbanos y metropolitanos](#) (ATUC) – published a [report in 2017](#) that found that in cities with more than 1 million inhabitants there is a high percentage of people that are either unaware, or have limited knowledge, about the public transport services offered in their city. This lack of awareness or easy understanding of different public transport options in a city can act as a barrier for citizens to use it, thereby increasing private car usage and leading to road congestion and poor air quality.

Mobility in Madrid
In Madrid, there are about 70 mobility operators with services such as buses, trains, trams, cars, e-scooters, bicycles and motorcycles. Of these, 41 are public transport operators while the others are private shared-mobility operators such as Coup, Car2Go, Muving, Emov and VOI. Coordination is provided by the regional transport authority: the [Consorcio Regional de Transportes de Madrid](#) (CRTM), a public body responsible for planning the mobility infrastructure . Currently, to enable smoother co-ordination between the different mobility operators (public and private), CRTM is collaborating with the [municipality](#)'s public transport company [Empresa Municipal de Transportes de Madrid](#) (EMT Madrid) under the framework of [CIVITAS](#) – an EU H2020 project – to develop a multi-modal platform: [Mobility as a Service Madrid](#) (MaaS Madrid).

MaaS Madrid: coordinating mobility data between private and public organisations in Madrid
[MaaS Madrid](#) offers the official mobile app of EMT Madrid and aims to facilitate sustainable mobility in the city. It is a shared initiative between businesses and government bodies that allows users to access information in real-time from the aforementioned mobility operators on a single integrated platform.

We are currently at the end of MaaS Madrid's first phase. The goal of this first phase is to create a first version of their data aggregator that will help integrate all the mobility data from the different transport providers into the app, that is already available to use on iOS and Android devices.

A win-win business model
There is great potential in sharing mobility data as a business model. If the involved private and public organisations share their data, then the citizens and tourists using transport services can better discover and use them. New data-driven services can provide (potential) users with information on the length and duration of a trip, the cost of the different options

and how to combine them for the fastest or cheapest route. With this overview and information, citizens and tourists will thereby find less convenient to use a private car.

In addition, once MaaS Madrid is fully functional after the launch of the next and second phase in the project, the City of Madrid will start collecting aggregated anonymous real-time data from the app to help in city planning. This aggregated data will enable the municipality to improve Madrid's infrastructure, reduce road congestion and update their public transport services to better suit the needs of their citizens and tourists.

The future of MaaS Madrid

As of September 2019, MaaS Madrid is still in its second phase and the new version of the app will be launched in the late autumn or early winter of 2019. Once this stage is complete, too, MaaS Madrid will be fully functional and the app include a travel planner for users to find the best travel options to get from point A to B in the city that is based on their preferences, purchase history and payment options.

Moreover, in the future the app also aims to be more interactive and provide incentives for users to use mobility services over private cars. These incentives are not yet finalised; however, they will aim to stimulate citizens to use public and shared mobility services. For example, if a user uses the combination of a bus and a shared bike, a reward scheme may give them points that they could use to get a free coffee at a coffee shop at the public transport station.

For further insights into MaaS Madrid listen to our interview with the Project Manager of EMT Madrid, Sergio Fernández Balaguer here.

Name	MaaS Madrid
Sector	Mobility
Region	Southern Europe
Country	Spain
Time	2018 - ongoing
URL	https://play.google.com/store/apps/details?id=com.emt.maas&hl=en
Business model	Business-to-Government
Partici-pants	The majority of mobility services in Madrid – including public transport and complementary services of shared mobility: BiciMAD, Car2Go, eCooltra, Emov, Ioscoot, Muving, Obike, OFO and Zity, radio taxi, etc.
Type of organisa-tions	Public and private organisations that specialise in mobility services, i.e. trams, busses, and shared car services.
Data shar-ing model(s)	Data partnership
Core im-pact	Increased efficiency of urban transport and improved seamless mobility experience for citizens.
Context	All mobility services in Madrid are available on a single tool connecting over 3,000 vehicles in operation across Madrid. Authorities in Madrid have encouraged car sharing through a series of initiatives and incentives.

2 JoinData: Data sharing in the agricultural sector

Link: https://eudatasharing.eu/examples/data-sharing-agricultural-sector

"Data is the key to sustainable innovation in the agricultural sector"
Sener Celik – Director of JoinData

Data in the agricultural sector
Modern farms create a huge amount of data. Machines are increasingly equipped with sensors that measure, for example, product quality, temperature, and water usage. Smart sensors on fields examine ground conditions and pest developments, and GPS-guided tractors produce data about the dimensions and location of the fields, the position of obstacles, and exact crop position. Data produced on farms can be used for a wide variety of purposes that ultimately benefit farmers in their decision-making processes and help them running businesses that are more efficient and effective. Field sensor data, for example, can be combined with weather data and applied to artificial intelligence to predict droughts or diseases that could damage harvests; data from milk machinery can be translated into insights that show farmers at one glance the health of their cows. Data collected through ear tags on animals can be used to estimate birth dates for pregnant sheep. Data offers a lot of opportunities for farm businesses and has the potential to enable sustainable innovation and growth.

A twofold problem with data sharing
A premise for sustainable innovation in the agricultural sector is that the data needs to be shared. Whereas data about the number of sick dairy cows on an individual farm might reveal nothing remarkable, when combined and integrated with data from thousands of other farms, patterns might be discovered, and steps can be taken accordingly. Data sharing in the sector is already taking place, according to Sener Celik – director of the data sharing platform JoinData. Parties such as application developers, hardware suppliers, agribusinesses, and banks are interested in farm data and take part in data sharing initiatives. However, data sharing in the Dutch agricultural sector is relatively fragmented: it mainly takes place in silos, e.g. between local groups and between parties that already have a tradition of sharing information with each other. Besides that, farmers have no or little control over who uses the data that is produced on their farms and for what purposes. Data from their milk machinery, for example, could be shared with the machinery's makers without the farmer knowing exactly what it is being used for or even without the farmers being aware of it. JoinData was founded against this background of data silos and a lack of transparency.

JoinData: distribution of agricultural data
JoinData is a non-profit organisation that was founded in 2017 by farmer cooperatives who aim to make data sharing in the agricultural sector more efficient and transparent. It provides a digital platform that facilitates secured data sharing along the value chain. On the platform, data providers can share data with parties who are interested in using this data. Laboratories, for example, can share their data with application developers, who then can combine it with other farm data to develop new software. But what does a data sharing platform mean for the farmer? JoinData enables farmers to be in charge of which data is shared with whom and for what purpose. When participating in JoinData, parties such as milk machinery makers that hold data about the health of dairy cows need permission from farmers before they can share this data with others. Farmers grant permissions by issuing explicit authorisations. Parties using the data are only allowed to use it according to terms and conditions that are agreed upon in the authorisation. By doing so, farmers retain control over the data that is produced on their farms.

Data distribution as a business model

JoinData does not own any intellectual property over the data that is shared on the platform. Neither do they use or modify it; they only distribute it from one party to another. Distribution is at the core of their business model: companies interested in using the data from the platform pay a fee to JoinData for their service. As a non-profit organisation, JoinData does not aim to make profit. Therefore, revenues are re-invested in the platform for further development. Currently, most revenues are invested in the technical aspects of the platform, such as the web-based user interface and security. Besides that, JoinData invests its time and effort in creating new partnerships, connecting all parties in the value chain and communicating the benefits of the service. An obvious question is of course: who benefits from JoinData? According to Sener Celik: the whole value chain. Farmers, for example, benefit because they have more insights and control over data. In addition, they benefit from the innovation that is developed thanks to the data, helping them run their business more efficiently and effectively. Application developers benefit from JoinData because they get access easily to thousands of different data sources, which enables them to make new data combinations. Data providers such as machinery makers no longer need to create their own data collection and processing platforms but can do the same through JoinData.

The future of JoinData
As JoinData is already expanding their activities in livestock, arable and horticulture, they soon want to cover the whole agricultural sector in the Netherlands. Their ultimate ambition is to facilitate secured data sharing on an international level. The biggest challenge is to connect all parties in the value chain. Some parties might be hesitant due to a wide variety of reasons, such as perceived missed revenues, or wanting to maintain exclusive access to specific data. Sener Celik, however, is optimistic about the future of data sharing. In addition, the Dutch government encourages data sharing between businesses and state in their Dutch vision on data sharing between businesses report that they prefer to encourage data sharing by means of facilitation but, if necessary, they will do so by regulation, too. This, for example, has already happened at the European Union level with the PSD2 regulation that requires banks to be able to share data with third parties for the development of new value added financial services and beyond. According to Sener Celik, initiatives as JoinData show that the agricultural sector is willing to innovate and unlock the potential of data to revolutionise the agricultural sector.

Name	JoinData
Sector	Agriculture
Region	Western-Europe
Countries	The Netherlands
Time	2018 - ongoing
URL	https://www.join-data.nl/?lang=en
Business model	Business-to-business
Participants	JoinData is a cooperative for the agriculture sector: companies, knowledge institutes, application developers, farmers, and agricultural entrepreneurs work together.

Type of organisations	Private organisations
Data sharing model(s)	Data sharing intermediary
Core impact	JoinData gives farmers greater control over their data since farmers can decide who gets access to the data that is produced on their farms. By enabling exchanging, combining and re-using data, Join-Data helps farmers and other parties in the agricultural sector to run their businesses more sustainable and efficiently.
Context	Parties in the agricultural sector can exchange data via the JoinData platform based on clear agreements about access to and use of the data. The data is being used by developers to create innovative applications that give farmers and other agricultural parties more insights in their businesses. JoinData aims to encourage innovations, which will eventually result in improved performance in terms of sustainability, profitability and welfare.

3   NIIS: The story of two countries sharing data

Link: https://eudatasharing.eu/examples/story-two-countries-sharing-data

"The Nordic Institute for Interoperability Solutions (NIIS) aims to develop and maintain standards and software that enable secure data transfer between various governmental and private data sources."

Petteri Kivimäki - CTO of the Nordic Institute for Interoperability Solutions (NIIS)

Data sharing between national governments: Estonia and Finland
The story of the Nordic Institute for Interoperability Solutions (NIIS) is one of two European countries that throughout history joint forces to collaborate and face challenges together. In 2013, the challenge to overcome was data sharing in and between national governments. Estonia and Finland decided to find mutually beneficial solutions together. The framework for the collaboration was set up in 2017 and called the Nordic Institute for Interoperability Solutions.

Coordinating technical interoperability
From another point of view, the story of NIIS is one about technical interoperability. Interoperability is "the ability of computer systems or software to exchange and make use of information".

European governments are getting more digital and citizen centric. To provide seamless electronic services, government departments and their information systems must exchange data. Therefore, the different IT systems have to be technically interoperable. The interoperability between systems can be facilitated by a standardised and secure exchange layer, for example X-Road.

Co-developing X-Road
X-Road is originally an Estonian open source software for standardised and secure data sharing which can be used and modified by anyone. X-road is the base for many digital services enabling and shaping Estonia, the national digital government initiative in line with the Once Only Principle. Beyond national interoperability, the X-Road open standard enables cross-border data sharing between governments.

In 2013, Finland started to implement X-road, too. However, country-specific aspects and national legislation made the implementation of X-Road in the Finnish public sector challenging and called for further modification. The two countries decided to co-develop the standard and the software. To make the collaboration more effective and to share the costs of development, the governments joined their skills and resources and coordinated the software development in the non-profit organisation NIIS.

The Nordic Institute for Interoperability Solutions for an interoperable EU
NIIS is further developing X-Road to enable digital cooperation in and between national governments and organisations. NIIS is equally funded and governed by its members and contribute their resources and skills. Currently, Estonia and Finland are the only members. However, Iceland and the Faroe Islands are already part of the NIIS partner programme and in the process of becoming NIIS members.

To put the cross-border feature of X-Road to good use, NIIS is welcoming new members and partners from Europe, connecting more national environments and developing new interoperability solutions. The wider vision is an interoperable Europe, integrating existing interoperability initiatives by the European Union, like eDelivery, to be compatible with X-Road. Another goal for the near future is the possibility to share the insights of data analysis instead of the data itself using X-Road.

NIIS and their achievements in further developing X-Road are a good example of strategic enablement of data sharing, that, by being successful, is almost invisible to most of its beneficiaries and the wider public. "I do not think citizens know NIIS" says Petteri Kivimäki – CTO of NIIS. Therefore, we invite you to listen to the Support Centre for Data Sharing interview with Petteri Kivimäki who elaborates on the background and the future plans of NIIS and X-Road for a data sharing ecosystem in Europe. Information about the upcoming NIIS community meetings can be found in the [Support Centre for Data Sharing event list](#) or [directly at Niis.org](#).

Name	Nordic Institute for Interoperability Solutions
Sector	Government & Public Sector
Region	Nordics and Baltics
Countries	Members: Estonia, Finland. Partners: Iceland and Faroe Islands
Time	2013 - ongoing
URL	https://www.niis.org
Business model	Government-to-Government
Participants	NIIS was founded and is owned by the governments of Finland and Estonia. Iceland and the Faroe Islands joined recently, respectively in 2018 and 2019.
Type of organisations	Public organisations
Data sharing model(s)	Multi-party data sharing agreement
Core impact	By further developing and maintaining X-Road, NIIS enables digital cooperation in the Nordic and Baltic region and any other country that decides to adopt the standard.
Context	NIIS is a cooperation between Estonia and Finland that was recently joined by Iceland and the Faroe Islands. The cooperation aims to develop and maintain standards and software that enable secure connectivity, data search and data transfer between various governmental and private data sources. NIIS focuses on practical collaboration, sharing of experience, and promoting innovation.

4   Ctrl-Shift: Data Mobility Infrastructure Sandbox
Link: https://eudatasharing.eu/examples/data-mobility-infrastructure-sandbox

"Personal data is currently not being harnessed. If it is, we can create and deliver new services that can aid and deliver new value for citizens."
Liz Brandt – CEO and Co-Founder of Ctrl-Shift

Personal Data Mobility
Personal data - the data we as individuals generate every day - is becoming increasingly more detailed and valuable, making it one of the fastest expanding markets in the digital economy. Examples of personal data are people's financial profiles, their electronic health record (EHR), social media "likes" and posts, social activity in the physical world and fitness habits and performance as captured by activity trackers. By harnessing this information, governments and businesses have the potential to create new services for social impact or economic growth and shareholder value and can support citizens in recognising what is happening in their lives and managing them.

The market of personal data is still new and involves a different type of data sharing relationship between the stakeholders involved. Currently, businesses, government bodies and citizens are still trying hard to understand what the data sharing economy entails, let alone personal data sharing, and are more at risk of making mistakes, e.g. putting themselves and others in danger if they share personal data. This risk leads to the concerns on how to share data easily while safely and how to increase transparency and trust.

Personal Data Mobility Sandbox
Ctrl-Shift is a business innovation consultancy that specialises in the strategic value of trusted personal data. The consultancy helps organisations to realise the opportunity in personal data and navigate through pitfalls by creating strategic and sustainable solutions to deliver business growth and value in customers' lives. This is exemplified in the report that collects the results of their Ctrl-Shift Personal Data Mobility Sandbox project.

The Ctrl-Shift Personal Data Mobility Sandbox completed with the publishing of the report in June 2019. It aimed to explore how to enable organisations and citizens to access the economic and social benefits of what they called "Personal Data Mobility". The project was an experiment completed in partnership with a series of mostly UK based organisations: Barclays (a bank), the BBC (a media organisation), BT (a telecom company), Centrica (an energy company), Facebook and digi.me in the role of "Data Facilitator" over 5 months. In the experiment, the organisations focused on implementing one data sharing model – where data is shared from an "exporting organisations" (an incumbent), to a Data Facilitator and then out to an "importer".

There are several findings from this report. Personal Data Mobility:
- has the potential to unlock significant value for individuals through the enhancement of pre-existing services or the creation of valuable new ones;
- can create additional value for businesses in existing markets;
- acts as an accelerant for governments in the development of new growth engines for the digital economy and ensures that additional value is created and shared;
- can create opportunities for consumers and citizens to collaborate with businesses on how to use Personal Data Mobility to create fairer value exchange in delivering the benefits of data sharing and guard citizens wellbeing; and
- needs to be facilitated by someone on behalf of individuals. In the experiment, this individual was called the Data Facilitator. Currently, Data Facilitators have not yet found their role in the market but are recognised as a core part of the model to ena-ble individuals feel more comfortable and able to share data.

Personal Data Mobility Challenges
To enable and support personal data sharing, Ctrl-Shift identified five challenges in the report that need to be addressed. These challenges are:

1. Infrastructure,
2. Consumer know-how,
3. New services,
4. Business models and
5. Regulatory standards.

Of these challenges, addressing infrastructure and standards is essential because they ensure that the market is safe, easy to use and can generate value for everyone involved. By having solid infrastructure and standards in place, the market can move forward, and it will be clear who is accountable for the data being shared so everyone involved knows in what stage they are responsible and – eventually – liable.

Future research on personal data sharing
As mentioned, infrastructure is critical to personal data sharing and can become a massive market in its own right. Therefore, Ctrl-Shift will continue developing the Personal Data Mobility Infrastructure Sandbox work and publish a series of reports in the coming months.

These reports will focus specifically on the infrastructure needed for personal data sharing and will ensure that personal data can be shared safely between public and private organisations in a way that is GDPR compliant. Here, individuals will be able to get access to their own data if they want to and share it with other organisations in a safe manner where their privacy is ensured. Moreover, the reports will look at liability in data sharing models, privacy and communication standards.

For further insights into the key findings of the Personal Data Mobility Sandbox, the challenges to personal data sharing and Ctrl-Shift's series on Personal Data Mobility Infrastructure Sandbox, listen to our interview with the CEO and Co-Founder of Ctrl-Shift, Liz Brandt, in the video above.

Name	Ctrl-Shift
Sector	Economy & Finance
Region	Northern Europe
Countries	United Kingdom
Time	2008 - ongoing
URL	https://www.ctrl-shift.co.uk
Business model	Business-to-Business
Participants	The "Data Mobility Infrastructure Sandbox" is a project facilitated by Ctrl-Shift that includes partners Barclays, British Telecom (BT), Centrica plc, BBC and Facebook.
Type of organisations	Business Innovation Consultancy
Data sharing model(s)	Data sharing intermediary
Core impact	In June 2019 Ctrl-Shift completed the first phase of their "Data Mobility Infrastructure Sandbox" for partners Ctrl-Shift that includes partners Barclays, British Telecom (BT), Centrica plc, BBC and Facebook, which examined how to deliver safe data sharing,

	and investigated how new markets can be unlocked by making data sharing safe for individuals and organisations. The <u>report is available on Ctrl-Shift</u> and describes the current enabling capabilities for data sharing and defines the gaps that need to be filled to make it safe and valuable.
Context	Ctrl-Shift is a business innovation consultancy that specialises in the strategic value of trusted personal data. The consultancy helps organisations realise the unprecedented growth opportunity in personal data by creating strategic, sustainable and practical solutions that deliver new value in people's lives. Ctrl-Shift is powered by an extensive network of expert strategy consultants, business analysts, consultants and researchers and is headed up by the team at the consultancy.

5   Data Pitch: The Data Sharing Toolkit
Link: https://eudatasharing.eu/examples/data-sharing-toolkit

> *"[Data sharing is] one of the most exciting topics in data and AI in the last decade and probably in the next decade to come."*
> Elena Simperl – Principal Investigator of Data Pitch

Data Pitch
Data Pitch is an EU H2020-funded open innovation programme that brings together corporate and public-sector organisations holding valuable datasets with start-ups and SMEs working with data. It is centred around a competition with several challenges and a virtual accelerator programme of 6 months to help start-ups and SMEs develop solutions to meet those challenges.

At the 2019 European Big Value Data Forum, we interviewed Elena Simperl: professor of computer science at the University of Southampton working in user-centric data science and AI. She is a Turing Fellow and a Fellow of the British Computer Society (FBCS), as well as a director of the Southampton Data Science Academy of the Web Science Institute. She has contributed to more than 20 research projects, often as a principal investigator – as for Data Pitch - or as project lead.

The Data Sharing Toolkit
Over the last three years, and in the context of [open innovation](#) that characterised the programme participants' collaboration, Data Pitch has collected a significant wealth of experience, that is now collected in the "Data Sharing Toolkit", available from the programme's website at https://datapitch.eu/datasharingtoolkit/. The short report collects lessons learned, resources, and recommendations that span from the access, exploitation, and evaluation of data on one side, to the creation of sustainable businesses on the other.

By dealing with multiple projects and challenges, different models by which data sharing was performed were documented. It is important to study each of the start-ups and their projects to learn about their choices. All start-ups investigated in the study are documented online at https://datapitch.eu/data-pitch-startups/.

By being "in flight" during the introduction of GDPR in Europe, the programme also had to learn and adapt to the new regulation. The Data Sharing Toolkit includes a "Legal and privacy toolkit" that collects the programme's learnings specific to these aspects and the impact on the data sharing models they experimented with.

In the interview, Prof. Simperl highlighted how the start-ups demonstrated themselves to be quite advanced technically, but experienced difficulty in recruiting talent. During the programme, they showed how access to subject matter expertise and skills – even just good practices in documenting data – can be an element of more significant friction than the technology *per se*.

For data holders, friction was often found in changing their way of working. Unless they had previous experience of open innovation – whether through hackathons or participating to previous challenges etc. – a significant part of the 6 to 12-month duration of a Data Pitch challenge was spent just changing the culture and finding the "champions". These are the right people in the organisation that could feel comfortable working against the new models and promote them to the others. According to Prof. Simperl, if an organisation was already

effective at sharing data internally, that was a positive sign of the same organisation being able to share data with others.

The interview closed by discussing the differences between the knowledge developed around the practices for open data and the ones for data sharing, and the – perhaps surprising – characteristics of the latter that make it *simpler* than publishing in the open.

Name	Data Pitch
Sector	Miscellaneous, including automotive, energy, health and wellness, manufacturing, retail, tourism, transport but also cross-sector.
Region	European Union
Countries	Denmark, France, Germany, Greece, Ireland, Italy, Latvia, the Netherlands, Portugal, Romania, Serbia, Spain, United Kingdom
Time	2016 - ongoing
URL	https://datapitch.eu/
Business model	Business-to-Business
Participants	Beta-i, DAWEX, the Open Data Institute, the University of Southampton and the >40 companies and organisations across the EU that participated to the challenges.
Type of organisations	Data-centred technology start-ups, data holder companies of any kind
Data sharing model(s)	Open Innovation, Challenge
Core impact	Through the release of the "Data Sharing Toolkit", the programme shares with the public the experience developed over three years of experimentation with start-ups and data holders. The toolkit is made of case studies, checklists and other tools that will support future data sharing initiatives.
Context	Data Pitch is an EU H2020-funded open innovation programme that brought together corporate and public-sector organisations that hold valuable datasets with start-ups and SMEs that work with data. It is centred around a competition with several challenges, and a virtual accelerator programme of 6 months to help start-ups and SMEs develop solutions to meet those challenges.

6   Technology Industries Finland

Link: https://eudatasharing.eu/examples/technology-industries-finland

> *"Data is [already] being shared, there are technical solutions (…) and there's business interest."*
> *Jussi Mäkinen – Head of digital regulation, Technology Industries Finland*

Technology Industries Finland

The combination of electronics, electrotechnical, mechanical and metals industries, engineering consulting and information technology is the most important export industry in Finland, with operations constituting over 50% of all exports.

Technology Industries of Finland[1] (TIF) is the advocacy organisation for the companies in these industries, with over 1,600 members. They promote competitiveness and the operational preconditions for the industry. Among their activities, they safeguard the industry' interests in national and EU-level decision-making and negotiate collective agreements on behalf of their members.

Over the last year TIF has dedicated substantial effort into observing the use of data in the industry. Data is today instrumental to its effectiveness and the way Finnish companies use its data is changing[2]. It is TIF's analysis that companies store large amounts of various data, but so far there have been no established practices for agreeing on the rights to use data, particularly when sharing it across organisations. To address this need, TIF has developed model terms and conditions for data sharing, that is now being promoted in the country and beyond.[3]

In October 2019 the Support Centre for Data Sharing had the opportunity to interview in Helsinki TIF's Antti Poikola, co-lead of TIF's AI Accelerator, and Jussi Mäkinen, head of digital regulation.

Addressing the limitations of nondisclosure agreements and intellectual property rights models

TIF's project aimed to create model terms and conditions and was initiated in early 2019 in response to the realisation that data sharing was already happening in the technology industries. More than half of Finnish companies stated in a survey that they share data already in one form or another. Both the business interest and the technological means to deliver data sharing are already there. However, the contractual terms under which data sharing was taking place were dominated by either non-disclosure agreements (NDA's) or intellectual property rights (IPR) assurance models. Those are not ideal because they are, respectively, too restrictive or based on concepts of property and exclusivity. Research indicated that no other more suitable practices were common.

At the same time, the legal teams in organisations cannot be expected to write original contracts, as it is a complex activity to be performed by teams that are usually challenged by capacity and workload.

---

[1] https://teknologiateollisuus.fi/
[2] https://teknologiateollisuus.fi/en/ajankohtaista/press-release/how-we-use-data-evolving
[3] https://teknologiateollisuus.fi/en/ajankohtaista/news/model-terms-and-conditions-data-sharing-solution-socially-economically-and

TIF's model terms and conditions for data sharing
TIF realised that writing even just industry-agnostic standard model terms that could cater for most legal elements and apply to most situations would have been significant progress already. The terms are also intended to be compatible and live along competition and personal data protection legislation in the European Union such as GDPR. In addition to that, they decided that the most suitable framing for those terms would have been to be an addendum to the (most likely) pre-existing agreement between the parties involved in data sharing. The terms are also suitable to extend an agreement to additional third parties, e.g. an AI analytics company that joins the collaboration later, as the need for their contribution is identified.

Finally, learning from the consulting industry, a key characteristic of the model terms is the right for participating parties to retain their learning of professional skills and experience that originated from using the shared data.

What happens next
The model terms are available commercially from Teknova's webstore[4] and are now in a roll-out stage. Their use is being monitored in trials and learnings collected by TIF to feed into potential future iterations of the work.

Name	Technology Industries Finland's "model terms and conditions for data sharing"
Sector	Technology industries such as electronics, electrotechnical, mechanical and metals industries, engineering consulting and information technology.
Region	European Union
Countries	Finland
Time	2019, trials ongoing
URL	https://teknologiateollisuus.fi/en/ajankohtaista/news/model-terms-and-conditions-data-sharing-solution-socially-economically-and
Business model	Business-to-Business
Participants	n/a
Type of organisations	Any
Data sharing model(s)	Any
Core impact	Industry-agnostic standard model terms cater for most legal elements for most situations and most companies as they enter data sharing agreements.
Context	Technology Industries of Finland is the advocacy organisation for the companies in these industries. They promote competitiveness and the operational preconditions for this.

---

[4] https://teknologiainfo.net/en/node/22768

7   SODA: A secure way to share data

Link: https://eudatasharing.eu/examples/secure-way-share-data

"We always have to remain critical and careful that we keep protecting our privacy"
Niek Bouman - post-doctoral researcher 'Secure Multiparty Computation'

A secure way to share data

At the European Big Data Value Forum 2019 in Helsinki, we interviewed Niek Bouman. He is a post-doctoral researcher 'Secure Multiparty Computation' at the Eindhoven University of Technology and involved in the Scalable Oblivious Data Analytics (SODA) project. This is an EU H2020-funded project that aims to enable privacy-preserving analytics of information from multiple data assets using multi-party computation techniques. These techniques allow multiple parties with private inputs to jointly compute data without revealing the input to each other.

Privacy concerns

Organisations generate a huge amount of data during their daily operations. Combining and analysing data from different sources is becoming increasingly important and offers great opportunities for value creation. Despite the promises of data sharing, many organisations are reluctant to share their data with others. One of the concerns around data sharing is the protection of personal data. In health care, for example, there is a huge potential in sharing the data that healthcare entities gather about their patients. At the same time, personal health information needs to be protected. Therefore, a technology is needed that allows for data sharing without compromising on privacy.

Secure multi party computation

A solution for a secure way of sharing data without endangering privacy, might be multi party computation (MPC). MPC is a toolbox of cryptographic techniques that enable different parties to carry out joint computations on their input while no party can see the other parties' data. This means that data from multiple providers can be analysed, without requiring them to share their data with anybody else. Computations are distributed across different data bases, which means that each server performs computations on its small part of the data, without disclosing the data. The distribution of data eliminates the need for a trusted third party to analyse the data. The participating parties determine who is allowed to view the outcome of the computations.

Applying MPC in practice

Going back to the healthcare example. Analysing data that different healthcare entities gather about their patients can be of great value for medical research and can reveal patterns that cannot be found within one single healthcare entity. With MPC technologies, multiple parties can provide data, such as data about peoples' age, lifestyles, health, diagnoses, and treatments. The data is split into separate pieces, encrypted, and computations are made on multiple servers. Participating healthcare entities only receive the output of computations, such as the average age of people diagnosed with a specific disease, while they cannot see each other's data. Each party can contribute personal data to the analysis while being technically guaranteed that the data cannot be de-anonymised, decrypted, or used for any other purpose than the intended one.

Challenges of MPC

A challenge on the road to large scale application of MPC, is the current performance of the technology. Each computation needs to communicate with each participant in the network, which effects the speed of the computation. Applications build on MPC could therefore potentially have a lower speed than applications build on the cloud. A solution is to do pre-computations locally, e.g. run specific analyses on datasets on a single server, before running

the multiple party computation. Compressing the data locally before processing the data jointly will increase the speed. It is also expected that with better networks and computers, the performance of MPC will automatically increase.

The future of MPC
MPC is not new. It has evolved from a theoretical curiosity in the 1980 to a technique to build privacy-preserving applications today. According to Niek Bouman, it is now time to use the technology to transform how organisations handle sensitive data. The technology is ripe enough for commercialisation and can be used as a solution to real business problems. The SODA project uses the healthcare as a first use case and aims to apply MPC to protect people's privacy while at the same time harvest the benefits of healthcare data. As a co-founder of Roseman Labs, Niek Bouman is also looking for business cases that can benefit from MPC to unlock the value of data.

Name	SODA project- Scalable Oblivious Data Analytics
Sector	Any
Region	European Union
Countries	Any
Time	2017 - ongoing
URL	https://www.soda-project.eu/
Business model	Any
Participants	The SODA project is run by a consortium of: Philips Electronics Netherlands B.V., Eindhoven University of Technology, Aarhus University, The Alexandra Institute A/S, and Georg-August Universitaet Goettingen
Type of organisations	Any
Data sharing model(s)	Any
Core impact	Building Big Data and Artificial Intelligence applications based on Secure Multiparty Computation (MPC)
Context	SODA is an EU H2020-funded project that aims to tackle data protection and anonymisation issues. The project enables privacy-preserving analytics on Big Data with MPC.

8   Social Science One and the Facebook URLs dataset
Link: https://eudatasharing.eu/examples/social-science-one-and-facebook-urls-dataset

"The ivory tower (of governments and academia as custodian of data) has to be broken down. We actually have to integrate more into the rest of society, understand what it is they're doing, why they're doing it, showing them that they can make a big difference in the world."
Prof. Gary King - Social Science One and Harvard University

Soon after Social Science One announced on February 13 2020 the readiness of their "Facebook URLs dataset" inaugural project, the SCDS team had the privilege to interview prof. Gary King, one of its founders.

A new form of partnership to share data
As described on the organisation's website, Social Science One implements a new type of data sharing partnership between academic researchers and private industry to advance the goals of social science in understanding and solving society's challenges. The partnership enables academics to gain access to and analyse information from the private industry in a manner that is responsible and socially beneficial. In addition, it ensures that the privacy of the people described in the data is protected while gaining societal value from academic research. Finally, it enables the companies willing to offer their data to support research and produce social good, without compromising their competitive positions.

The first project
Social Science One's inaugural project in partnership with Facebook can be used to explain its model. Following the 2016 presidential election in the United States, the effects of social media - and of Facebook in particular - on democracy and elections have become a hot topic. The platform wanted to collaborate with researchers to limit speculation and fully understand the phenomenon. However, at the same time, there was no obvious way for the company to share the so called "URLs dataset".

The URLs dataset is the collection of about 38 million URLs shared by Facebook users worldwide that triggered at least 100 interactions (viewed, shared, liked, reacted to, shared without viewing...) between 1 January 2017 and 31 July 2019. It also includes detail such as in which country they were shared and whether they were fact-checked or flagged by users as hate speech, and the aggregated data concerning the types of people who interacted with the URLs.

How does it work?
Social Science One was created to act as an independent intermediary between the researchers and the data provider. The provider - e.g. Facebook - entrusts Social Science One to evaluate the academics' research proposals according to their academic merit and to release authorisations accordingly. The board reviewing the applications is a group of academics, who are committed not to research the topics for the duration of their participation to the board, as a guarantee of independence. In prof. King's words, they "take one for the team".

The challenges of the data and the model
It took 20 months from the original idea to the moment Social Science One was ready to start releasing authorisation to researchers. There were many challenges, including legal, technical and related to how to communicate the project.

From a legal perspective, among the challenges were creating first the best form of organisation for Social Science One and supporting Facebook's legal department to get to a point where they could feel comfortable to share their data.

From a technical perspective, the main challenge - besides the complication and cost of maintaining a dataset this size - was to protect the privacy of Facebook users. Conventional anonymisation was unfortunately not an option, as user behaviour on a social platform offers itself to re-identification more easily than other kinds of data. Strong aggregation was also not suitable, as it would have compromised the detail of the original information. Social Science One opted for "differential privacy": a technique that manipulates a dataset in order to preserve its statistical characteristics while ensuring the anonymity of the Facebook users.

Finally, even just communicating the project is a challenge, and not a secondary one. During our interview, prof. King highlighted how it is a key element of success as much as the other elements that make data sharing possible. Differential privacy, for example, is a relatively new technique most people do not know about. Presenting Social Science One's proposition every time needs to be accompanied by a long explanation of how it was made possible: legally, technically, and in terms of respecting people's privacy.

Is this the future of social science research?
Governments and universities used to be where the best methods and the most valuable data for social science research were - highlights prof. King - but this is no longer true. Successful digital platforms today have a better visibility of social behaviour than anyone else. Academia needs to leave its ivory tower to work with them, showing them that they can make a big difference in the world.

Name	Social Science One
Sector	Academic research - Social sciences
Region	USA / World
Countries	Any
Time	2019 - ongoing
URL	https://socialscience.one/
Business model	Non-profit
Participants	Social Science One is a non-profit organisation being incubated at Harvard's Institute for Quantitative Social Science. Founders were prof. Gary King and Stanford University's Nathaniel Persily.
Type of organisation	Non-profit. More detail on the organisational structure is explained in the paper "A New Model for Industry-Academic Partnerships" by Gary King and Nathaniel Persily.
Data sharing model(s)	Academic-private partnership for the independent administration of access to confidential data for research purposes.
Core impact	The project enabled the release to social science research of what is likely the largest ever dataset relevant to studying user behaviour on social networks, the

	distribution on news on such platforms, disinformation etc.
Context	Social Science One implements a new type of partnership between academic researchers and private industry to advance the goals of social science in understanding and solving society's challenges. The partnership enables academics to analyse the information available to private industry in responsible and socially beneficial ways. It ensures that the public maintains privacy while gaining societal value from scholarly research. And it enables firms to enlist the scientific community to improve their business and produce social good, while protecting their competitive positions.

9   BDVA: Towards a Pan-European Data Sharing Space
Link: https://eudatasharing.eu/examples/towards-pan-european-data-sharing-space-0

> "There is an understanding that data sharing has to happen, and it is already happening. The more data there is available to use and experiment with, the better we can develop AI for everyone."
> Simon Scerri of Fraunhofer IAIS and Irene de López de Vallejo of BluSpecs from the European Big Data Value Association

Big Data Value Association
The [Big Data Value Association](#) (BDVA) is an international non-profit organisation that consists of members from academia, public bodies, industries and private citizens. The association aims to raise awareness and promote the value of data across Europe and to enable data and AI-driven digital transformation in Europe to:
1. Deliver maximum economic and societal benefit; and
2. Achieve and sustain Europe's leadership on big data value creation and artificial intelligence (AI).

Recently, the BDVA has shifted its efforts towards AI and in making more data available to support it – including using open data and shared data that cannot be published as open. In the beginning, the BDVA prioritised publishing and supporting the promotion of open data to accomplish this shift. Now, they are broadening their scope as they realised that what has worked for open data can also be adapted for non-open data (data sharing). These datasets, both open and shared, can contribute to develop AI for everyone.

Within the BDVA's network, there are organisations - including small-, medium- and large-sized industries – that are already incorporating data sharing into the heart of their activities. These entities started pursuing data sharing because they saw the potential and advantages it could bring them, such as opportunities to create more solutions and products and the related revenue streams. The BDVA recognised this potential and decided to take a mediating role to help converge its network's effort on data sharing. One of the first outputs of this process was to produce a position paper highlighting this potential.

Towards a Pan-European Data Sharing Space
The BDVA's position paper titled "[Towards a European Data Sharing Space: Enabling data exchange and unlocking AI potential](#)" was published in April 2019. It is a first step to converge data sharing practices of businesses and institutions from different sectors to inspire others and demonstrate what is possible. As stated, the paper aims to inspire the European community - such as researchers, industry representatives, public bodies and private individuals - to share data with one another by showcasing examples and opportunities and highlighting Europe's strengths.

For more information on the position paper, the Support Centre for Data Sharing (SCDS) spoke with two members of the BDVA and editors of the paper: Simon Scerri of Fraunhofer IAIS and board member of the BDVA, and Irene de López de Vallego of BlueSpecs. They stated that at the inception of the research for the report, several challenges were identified as the most common data sharing practices were documented. This included technical, business and organisational, legal compliance, and national and regional challenges. The two key challenges that were identified and discussed with Simon and Irene were data governance and regulations (absence or limited presence thereof). Technology was not identified as a challenge because many technical solutions are already mature and available. Instead, legal challenges – especially in the domain of policy and governance – were elaborated on. Several recommendations are presented in the report, too. Below are two

examples of recommendations discussed in the interview to alleviate some of the challenges:

- **Addressing data "ownership"**. Currently, the set of legal rights associated to data – commonly referred to as the "ownership" of data - is legally unclear. For example, there is no clarity or consensus when discussing how to transfer rights or how they can be constrained in time. To address this, data sovereignty needs to be guaranteed. Organisations that participate in data sharing feel more comfortable when they know who is responsible for what. This can be achieved by taking a more pragmatic approach to data sharing rights and management, and by incorporating technical and legal aspects, including governance – i.e. how governance and technology can support trust.
- **Incorporate in the data lifecycle**. Most data producers do not consider data sharing as an opportunity in any part of their data lifecycle. As data sharing becomes more prominent, entities including business, public bodies and individuals should become aware of how the data they are producing can be relevant and bring value to other organisations.

For more recommendations, explore the [BDVA's position paper](#).

The future of data sharing in Europe
The BDVA is at work in a second edition of the position paper. In the new edition, they will expand their scope to focus on additional aspects of data sharing across Europe. One of these will be data sharing models. The association will continue to have a business-centric focus and will research on different data sharing model with businesses, including with other businesses (B2B), with government bodies (B2G), with consumers (B2C), and with scientists and researchers (B2S).

The second edition is expected to be published in Spring 2020.

Name	Big Data Value Association
Sector	Academic research
Region	Europe / World
Countries	European countries
Time	2018 - ongoing
URL	http://www.bdva.eu/
Business model	Non-profit
Participants	The Big Data Value Association is an international non-profit organisation that consists of members from academia, public bodies, industries and private citizens. The association aims to enable data and AI-drive digital transformation across Europe. Two representatives from this association and editors of the position paper discussed were interviewed: Simon Scerri and Irene López de Vallego.
Type of organisation	Non-profit
Data sharing model(s)	Pan-European data sharing space
Core impact	The position paper sought to converge data sharing practices from different entities across Europe from multiple sectors to

	inspire other stakeholders to share data and create value and impact. In addition, the paper sought to showcase data sharing opportunities for different stakeholders and provide recommendations to overcome different challenges.
Context	A socio-technical shift is spreading across Europe. To match this shift, the BDVA is now focusing on making more data available to support AI by researching and showcasing how entities across Europe are using data – open and shared – to develop AI for everyone.

10 iShare: Sharing Dutch transport and logistics data
Link: <u>https://eudatasharing.eu/examples/ishare-sharing-dutch-transport-and-logistics-data</u>

> "*Organisations are starting to realise that in order to stay competitive and to increase their efficiency, they need to share data.*"

> Eefje van der Harst – Project Manager, INNOPAY

Data in the transport and logistics sector
The transport and logistics sector is currently booming in the Netherlands. It is one of the key sectors driving the Dutch economy, and the Dutch government is keen to sustain this growth and maintain its leading position in the global market. To achieve this, they have committed to continue improving and innovating the sector through supporting efficiency gains by making operations in the supply chain smoother. One way the government has done this is by creating and supporting the data sharing initiative <u>iSHARE</u> – a collaboration between the <u>Ministry of Economic Affairs</u> and the <u>Ministry of Infrastructure and the Environment</u> that supports the Dutch transport and logistics sector by improving its efficiency and reducing costs and carbon dioxide (CO_2) emissions through connecting different stakeholders with a trust framework.

iSHARE – data sharing in the Dutch transport and logistics sector
iSHARE went live in 2018. Before its launch, the Dutch logistics sector was inefficient with, for example, high levels of road congestions. A reason for this high congestion was that trucks were waiting in the Dutch harbour because it was unclear if the ship they needed was already in the harbour or where the specific location of their container was. This uncertainty was because the data about the ships and containers were inaccessible, fragmented, and not shared between partners. This meant that planning could not be optimised and that the supply chain remained inefficient. After iSHARE was created, conditions for data sharing were improved which led to a higher efficiency gain where congestion declined (and by extension also reducing CO_2 emissions) and businesses based in the Netherlands could save money, further boosting the Dutch economy.

During the initiative's inception, it was decided that instead of creating a single platform, iSHARE would be a trust framework. This means that iSHARE does not have a central infrastructure. Through this trust framework, companies in the transport and logistics sector can keep data as close to the source as possible. Moreover, this data can remain under the control of entitled parties, where they can decide what data is shared with whom, under what conditions and for what purpose. At most, iSHARE will have a list of participants – people and organisations that have entered the framework and have agreed and proven that they conform to the operational, technical and legal specification – in a database. This concept and the idea of "embracing fragmentation" resonated within the sector because it addressed trust issues between companies and government bodies who were cautious about how the data they share would be distributed and used. Thus, in 2017 INNOPAY received the opportunity to guide the sector and to co-create the iSHARE trust framework together with about 20 organisations from the transport and logistics sector. Examples of these organisations are <u>Dutch Customs</u>, <u>Portbase</u>, <u>Secure Logistics</u>, and <u>Hutchinson Ports ECT Rotterdam</u> (ECT). Moreover, it was agreed that the initiative's scope would be limited and focus only on:

1. **Identification** – determining who the individual is;

2. **Authentication** – proving an individual's identity; and

3. **Authorisation** – establishing how systems and organisations can identify and authenticate that an individual is authorised to access or publish certain data.

Data "ownership"

iSHARE participants can decide which other parties they will share their data with and under what conditions. Following the initiative's creation, a generic set of requirements for organisations in the sector was created. This set of requirements can be used by all organisations, irrespective of their size or modality. In addition, it was decided that iSHARE's framework would be limited with no discussion on data standards because it was domain specific. Instead, a generic infrastructure that is feasible for all organisations was prioritised. By using generic standards that create and support interoperability, organisations can more easily and efficiently create new connections and share data safely with multiple parties. Currently, we are in a world where organisations have unique and specialised connections with others, creating a whirlpool of different connections that is difficult to holistically maintain and secure. To navigate through this whirlpool, it is important to work with (international) standards so that systems can be integrated with other parties with certainty that data is being shared securely. Furthermore, usage of these standards ensures that only people who are authorised to access the data can view and use it.

Securing data sharing

As organisations discuss data sharing, they will look for ways to do so in a secure manner. This means that technical and legal standards are crucial. When joining iSHARE, all participants are required to sign an accession agreement, where they agree to adhere to the terms of use that are in accordance with Dutch contract law. This means that all participants are held to the same agreement and no separate bi-lateral agreements need to be made with other organisations. Once the agreement is signed with the scheme owner (now iSHARE Foundation), the participants are bound to the same contract with the other organisations.

In terms of licensing, iSHARE has standardised a few licenses that organisations can use during a transaction to provide instructions on how a service can be consumed or under which conditions data can be exchanged. Here, organisations can state what data can and cannot be shared. For example, organisations can choose to:

- Have no limitations on data sharing;

- Place a condition stating that this data can only be shared and used for internal purposes; or

- State that the data can be used for non-commercial purposes under the condition that it is not used to generate revenues.

The standards and licenses in iSHARE were leveraged from existing frameworks. One example is the Trust Services and Electronic identification (eIDAS) – an EU regulation on electronic identification and trust services for electronic transactions across Europe. Furthermore, in addition to the mandatory standards and licenses in the initiative, participants also need to adhere to generic law in Europe, including GDP and competition law.

Future of data sharing in the logistics sector and iSHARE

As stated, organisations across the world are realising that in order to stay competitive and increase their efficiency, they need to share data. Moreover, they need to be flexible and adapt to changing circumstances in the market, which can be achieved through collaboration and sharing information. Across Europe, stakeholders – including government bodies and businesses – are critical on the dominance of platforms and in centralising data. Thus, the

iSHARE trust framework perfectly fits in this growing momentum in Europe about the importance of data sovereignty and in keeping data de-centralised, meaning that data is only shared under the control of entitles parties.

As of 1 October, iSHARE is formally owned by the iSHARE foundation. Following the interview with Eefje van der Harst (the previous iSHARE Project Manager from INNOPAY), the SCDS team also interviewed the current Executive Director of the iSHARE Foundation: Gerard van der Hoeven.

Name	iSHARE
Sector	Transport and logistics sector
Region	European Union
Countries	The Netherlands
Time	2017 - ongoing
URL	https://www.ishareworks.org/
Business model	Government-to-Business
Participants	More than 20 logistics organisations, including Dutch Customs, Portbase, Secure Logistics, and Hutchinson Ports ECT Rotterdam (ECT).
Type of organisation	Any
Data sharing model(s)	Any
Core impact	Establishing a trust framework for organisations within the Dutch transport and logistics sector that enables participants to determine who their data is shared with, under what conditions and for what purpose.
Context	iSHARE is a trust framework for stakeholders in the Dutch transport and logistics sector. The initiative promotes efficiency gains along the supply chain in the sector.

11 iShare: Data sharing in the Dutch transport and logistics sector

Link: https://eudatasharing.eu/examples/ishare-data-sharing-dutch-transport-and-logistics-sector

This article is a follow-up on the practice example on iSHARE: Sharing Dutch transport and logistics data.

The iSHARE Foundation is an independent, transparent and non-profit body that is focused on enforcing agreements, managing the processes that are associated with joining the network and facilitating changes in the scheme. The iSHARE scheme includes identification, authentication and authorisation for the transport and logistics sector in the Netherlands. This set of agreements enables stakeholders to share logistics data with other parties, including with new and previously unknown partners. It is simple and controlled, because the owner of the data always retains the final decision about who has access to which data and how. The number of iSHARE participants in the logistics sector has been growing steadily ever since the launch in 2018.

The SCDS team interviewed Gerard van der Hoeven, the executive director of iSHARE Foundation, for more information about his strategy for the future of iSHARE.

Use case example: Hutchison Ports ECT Rotterdam
One of the new companies joining the growing network of iSHARE is Hutchison Ports ECT Rotterdam (ECT). ECT joins the iSHARE network after a successful pilot which launched in 2019. During the pilot, ECT shared data with the third parties, the (Dutch) fruit and vegetable supply companies Nature's Pride and Total Produce. ECT decided to join the iSHARE network after the pilot was a success in April 2020.

ECT and its customers can securely share container data in logistics chains under ECT's control with the iSHARE data sharing scheme as of May 2020. The iSHARE data sharing scheme is used for data sharing between parties that are not directly related to each other. ECT shares information about container activities at its ECT Delta and ECT Euromax terminals with numerous third parties every day, often at the request of a customer.

Shippers benefit from the accurate and timely information on the status of their container(s) during the unloading process in the terminal(s). The container's actual unloading time is particularly important, as it enables shippers to further optimise their own planning and processes. In the case of fresh food shipments from companies such as Nature's Pride and Total Produce, the container's status even determines the timing of the goods' release into the market.

Currently, shippers or forwarders enter various data, including a PIN code, in order to obtain information from a terminal on a specific container. This information is not automatically available to them or the other parties involved, because the information sharing rights between the various parties are not known at the outset. Not only does this process generate manual work, but there are often delays in issuing PIN codes to the various parties.

Through the iSHARE data sharing scheme, organisations can share data with parties such as Nature's Pride and Total Produce at an earlier stage using data from Portbase's Port Community System. With the owner's consent, Portbase shares information on the container (i.e. location and time of arrival) with the terminal. That information is then automatically

made available to authorised users through ECT's new "MyTerminal portal". ECT and Portbase have aligned their messaging with iSHARE's agreements system for this purpose.

Before providing controlled access to the requested data, ECT wants to be sure that the information will be received by the right person within the right company for the right purpose. By joining the iSHARE scheme, ECT can use the identities from iSHARE Identity Provider Secure Logistics to gain the necessary technical and legal certainty for itself and its customers. One of the applications of iSHARE is a single sign-on on the ECT app.

Companies in the logistics and transport sector, such as ECT can join iSHARE in the Netherlands. The executive director of the iSHARE Foundation plans to expand to more countries in Europe as well because the trading routes between European countries, especially in the Schengen area countries, are interwoven.

Practical guide to start with data sharing

The use case of ECT shows that data sharing can help several companies in the logistics and transport sector to work more efficiently. However, data sharing might not be beneficial for all companies and different rules and regulations apply for different types of data. Therefore, before starting to share data with third parties, iSHARE recommends following a five-step plan:

1. Businesses should consider making the supply chain more efficient in terms of costs and time. Also, security and safety of the data needs to be taken into consideration when drafting a data sharing strategy.
2. The next step is to identify the role of the organisation in data sharing, different types of data sharing and partners. For example, data can be shared from machine-to-machine or from human-to-machine.
3. Business professionals need to think about compliance with the rules and regulations of data sharing and include their IT director and legal advisors in this process. For example, in order to comply with the iSHARE scheme, professionals need to adhere to the technical, functional, operational and legal agreements. Compliance with technical agreements possibly includes (new) digital certificates. As for the operational agreements, technical tests of the systems and updates are likely to be necessary. The legal advisors can check for compliance with data sharing legal agreements.
4. Start the implementation of data sharing through a scheme, such as iSHARE, by registering your company and contact details.
5. Start with data sharing once the scheme has approved your registration.

It is advised that everyone who starts data sharing in Europe should draft a plan and take into consideration the rules and regulations for data sharing of the European Union, such as the General Data Protection Regulation (GDPR) and EU competition law. In this way, stakeholders such as business officials and government bodies in the logistics and transport sector, for example, can get the most out of sharing data.

Overall, the benefit of data sharing in the transport and logistics sector is greater insight into the supply chain, enabling you to work faster, more efficiently and more cost-effectively.

Name	iSHARE
Sector	Transport and logistics sector
Region	European Union
Countries	The Netherlands
Time	2017 - ongoing
URL	https://www.ishareworks.org/
Business model	Government-to-Business
Participants	More than 20 logistics organisations, including Dutch Customs, Portbase, Secure Logistics, and Hutchinson Ports ECT Rotterdam (ECT).
Type of organisation	Any
Data sharing model(s)	Any
Core impact	Establishing a trust framework for organisations within the Dutch transport and logistics sector that enables participants to determine who their data is shared with, under what conditions and for what purpose.
Context	iSHARE is a trust framework for stakeholders in the Dutch transport and logistics sector. The initiative promotes efficiency gains along the supply chain in the sector.

12 IBM and Data Sharing
Link: https://eudatasharing.eu/examples/ibm-and-data-sharing

> *"After [the AI] does an engagement, (…) just like the people (…) going to a customer and consulting with them, you walk out of the engagement and you're a different person, you've learnt something from your experience in solving that customer's problem (…). The machine, basically, is doing something similar, or at least analogous."*
> Chris O'Neill, Associate General Counsel for IP at IBM

IBM is not new to the open technologies, from quantum and blockchain to containers, AI, and operating systems, data and software. The company is actively involved in many mainstream projects[5], and, more recently, its interest extended to the space of open data and data sharing. SCDS's Hans Graux and Gianfranco Cecconi have had the opportunity to interview Chris O'Neil: IBM associate general counsel, intellectual property law, specializing in data-related matters and law.

Specifying the value of data is a challenge. The experience being developed around the topic is limited, and often bound to old-fashioned information technology models. Before the Cloud, companies, used to operate resources they owned physically: the servers, the data stored on those servers, the software processing it, the networks. The "borders" surrounding all these assets were clear.

Today those borders are often immaterial. Servers, data and software are used every day that could be anywhere in the world. Their physical location lost great part of their meaning, if not in court for the applicability of a legislative framework or another. The word "ownership" itself can be misleading with data and talking about intellectual property rights and licences is much more meaningful than about "data ownership". The problem becomes even more articulated when focusing on the value of the *insight* generated from data, or the *learning* gathered by that insight. What rights does the owner of an artificial intelligence (AI) has, for example, to exploit the learning from data held by a client?

Who owns the insight? Who owns the learning?
IBM faced this riddle when they introduced their Watson family of products. With Watson, the company offers a service that – in short – processes the clients' data outside of those safe, known borders of the office. The legal aspects related to the data being processed that way, and to the data being *produced* as the outcome of the process – the insight, the learning – were novel.

To make the picture more complicated, with AI there is no net separation between the software and the data being used to train the AI. The Watson software, owned and operated as a service by IBM, becomes smarter thanks to the data held by the clients, data that is most times proprietary and confidential. How could this conundrum be solved, and in a way that would stand the challenge of law?

Artificial intelligence like human consultants
One of IBM's intuitions in dealing with the issue was to consider AI's as similar to human consultants, learning as they work for their clients, and leaving the projects as "different persons", more skilled and capable.

It is commonly accepted how this category of professionals can capitalise on its learning, build on it, and re-use it as they work for the next client. They don't take with them the source's commercial secrets, but rather develop language, skills, and experience. What

---

[5] https://www.ibm.com/opensource/

happens, though, when that next client is a competitor of the first? For human professionals, it is not an issue: work ethics will prevent the consultant from sharing confidential information from the source. However, they will be free to re-use the acquired skills.

Indirectly, the consultancy industry educated the business to a culture of sharing. We all benefit from the accumulated experience of those professionals, as they perform their work across many different industries and organisations.

... and the need for "data spaces"
Healthcare and pharmaceutics were among the first industries IBM decided to develop Watson in. Chris O'Neil observed how, in most industries, the value associated to data is to be found in the *insights* that the data can produce, but healthcare was different. IBM realised that, in the sector, *the data itself* has commercial value. For example, the data created as a new chemical component is developed, the history of its evolution, the clinical trials and the medical records of the patients involved are valuable *per se*, as it captures, as capital, the large research investment effort.

This capital can be exploited beyond the company that holds the data. Sharing is necessary to exploit its potential, to cross-over the data with other data resources held by others. At the same time, sharing the data openly, for anybody to re-use, is hardly commercially acceptable, if not simply an impossibility because of the need to protect the patients' privacy. One option is to anonymise or aggregate data before sharing, but that can hinder substantially its potential to support further research. It becomes inevitable then to implement models where access is restricted – e.g. to known researchers only - and the rights on that data – e.g. limiting the research to topics the patients provided consent to. These protected arrangement in Europe are often called "data spaces", e.g. in the European Commission's latest strategy for data[6].

The role for licences
All of the aforementioned legal elements of data sharing, from model to restrictions to rights to consent, are captured in licences and contracts. Through Chris O'Neil's work IBM has already contributed to the specifications of licences such as the Linux Foundation's "Community Data License Agreement"[7] (CDLA). The licence specification captures and acts upon many of the learnings from IBM's experience with Watson.

Looking at the future, one of the challenges to effective data sharing is the potential fragmentation and extreme specialisation of licensing terms. Fragmentation creates friction to sharing and the exploitation of the opportunities arising from data. Initiatives such as the CDLA and the work of others in the licensing community – including the Support Centre for Data Sharing itself for the EU - will be instrumental to converge effort and maximise potential.

We recommend watching the full video interview to get Chris' insight on additional aspects related to data sharing, from the difference between a licence and a contract, to the intellectual property of *collections* of public domain data, or the "hub and spoke model" for liability and more.

---

[6] See the "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of Regions" of 19 February 2020, available online at https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf .
[7] https://cdla.io/

Name	IBM
Sector	Software and related advisory services
Region	USA / World
Countries	Any
Time	Ongoing
URL	https://www.ibm.com/
Business model	Commercial with open elements
Participants	IBM is a multinational corporation active in cloud, information technology, software and AI technologies.
Type of organisation	Commercial.
Data sharing model(s)	Miscellaneous, from open to shared.
Core impact	IBM has contributed to the development of licence agreements dedicated to data sharing, such as the "Community Data License Agreement" (CDLA) by the Linux Foundation.
Context	In the context of developing its Watson line of services, IBM needed to explore the legal aspects around the intellectual property rights of its clients and its own, when applying artificial intelligence technology to third parties' data. This brought to IBM's investment and contribution to the wider licensing community working on standardised terms for data sharing.

13 Corporate Data Quality: Sharing the efforts of quality business partner data

Link: https://eudatasharing.eu/examples/corporate-data-quality-cdq

> *"When companies accept that other companies have good data management as well and understand the value of joining forces, we can share the efforts of cleaning and updating client and vendor data, each and everyone keeping data sovereignty."*
>
> Kai Hüener, CTO at CDQ

International business partner data

International companies can have thousands of clients and vendors all over the world. Maintaining this data is crucial to minimise risks and administrative efforts of fixing, for example, wrong invoicing or shipping.

In principle, this data is publicly available and accessible. However, the complexity of business partner data comes from the global scope with ~200 countries, each with several national identifiers, address formats, legal forms, languages, standards, etc., and with the data being subject to change. In many countries, especially in Europe, national company registers are available as open data. These sources help to find and verify business partner data like addresses from clients, vendors or suppliers. However, keeping this information manually up to date is an immense effort and, to make it more complicated, many countries do not have open and well-maintained business registers.

The Competence Center Corporate Data Quality, a research initiative formed by academic institutions – among them the [Faculty of Business and Economics](#) (HEC – University of Lausanne), the [Institute of Information Management](#) (IWI – University of St. Gallen), and the [Institute of Accounting, Control and Auditing](#) (ACA – University of St. Gallen) – has further investigated this situation.   To answer the difficulty of keeping clean and updated, compliant, and harmonised business data, CDQ ([Corporate Data Quality](#)) was founded.

The CDQ community jointly tackles the challenge

The idea behind the initiative is to share the data management efforts of quality business partner data between CDQ community members, thus protecting data sovereignty. Kai Hüener, an evangelist for data sharing and CTO of CDQ, explains the idea in his interview with the Support Centre for Data Sharing.

CDQ uses external sources like open business registers, commercial sources and crowd sourced input from the members of the CDQ community to provide a pool of business partner data. Member companies can make use of this pool by integrating the [CDQ API](#) in their respective system, e.g. [SAP](#), and receive updates for their existing business partner data - before it causes difficulties. In return, the companies can decide to share their own updates with the CDQ community. CDQ validates the update and shares it with the community anonymously. Only public information is shared, no notes or comments, for example. If companies decide that they do not want to share their own business partner data, they can opt out and will not receive updates from the community. These are only available to the companies that agree to share their updates in return. It is however possible to share data but excluding strategic key vendors or suppliers.

In addition, member companies do not need to change their standard format of business data, e.g. which fields a dataset contains, how it is called or their order. CDQ works with different standards, forms and metadata vocabularies. In that way and for a service fee, CDQ enables the business data sharing community to share the efforts of clean, updated business partner data. This is particularly attractive for large international companies like Bayer or

Novartis, current members of the CDQ community. Each new member that joins the community improves the pool and the opportunities for AI to automate part of the process.

AI for automated validation of updates
Algorithms can already clean and structure addresses to make them comparable. They can also classify records after feedback loops from the community and find anomalies or patterns in updates that were rejected. In the future, algorithms can take over the role of validating and accepting updates. Using external sources, an updated VAT number from a CDQ community member, for example, can be validated in an automated way based on the information in the external source. The process of accepting the validated update can then also be automated.

Kai highlights that the basis for realising these new features is trust and the need for a change of mindset that will probably take another 2-5 years. Awareness for coping with data management jointly is increasing and more organisations see the advantage of crowd sourced business data management with a third trusted part for validation and organisation.

Name	CDQ
Sector	Software and related advisory services
Region	Switzerland / World
Countries	Any
Time	Ongoing
URL	https://www.cc-cdq.ch/
Business model	Commercial
Participants	The Competence Center Corporate Data Quality is a research initiative formed by academic institutions, including the Faculty of Business and Economics (HEC – University of Lausanne), the Institute of Information Management (IWI – University of St. Gallen), and the Institute of Accounting, Control and Auditing (ACA – University of St. Gallen). This research is what founded CDQ (Corporate Data Quality).
Type of organisation	Commercial
Data sharing model(s)	Trusted third party that manages and provides crowd sourced, open and shared commercial data for its members.
Core impact	The efforts to maintain and update clean business partner data are substantially minimised, specially for large international enterprises by joining forces and sharing the data with the community.
Context	International companies can have thousands of clients and vendors all over the world, each with several national identifiers, address formats, legal forms, languages, standards, etc., and with the data being subject to change. Keeping this information manually up to date is an immense effort. CDQ addresses this challenge.

14 DataVaccinator: Personal data protection by design via "industrialised pseudonymisation"

Link: https://eudatasharing.eu/examples/datavaccinator-personal-data-protection-design-industrialised-pseudonymisation

"Violation of personal data protection is not always with evil intent but often rooted in human errors. Storing data pseudonymised and separate from identifiable attributes protects from any kind of personal data violations."

Kurt Kammerer, Co-founder of DataVaccinator

Why

Protecting personal data has been a focus topic for several years. Sharing personal data can be as advantageous as it is controversial and complex. Medical records, for example, are sensitive personal data. At the same time, scientific research greatly benefits from analysing health records to better understand diagnostics and treatment. Identifiable attributes like the name of the patient are not important for research. Therefore, anonymisation and pseudonymisation are effective tools that enable sharing of personal data without the identifiable attributes. However, those tools can bare the risk of re-identification and leave the original pool of data that was used to disidentify the data, post-generation, at risk to leakages, breaches, and theft.

A pseudonymisation of data at the point of its generation mitigates this risk. If a person's medical record is stored separately from the identifiable attributes when it is generated, the medical record is no longer directly identifiable. Only with the knowledge about the XX can identity and content be combined, for example when going to the doctor. This makes storing and sharing medical records much safer.

How

DataVaccinator splits data into separate databases upon generation and is joined real-time in the local applications. Using encryption, DataVaccinator ensures that only eligible parties can access the sensitive joined data. The solution also lowers the costs of data pseudonymisation and management.

The software is available as opens source software. Developers can integrate ready-to-use components with open licence, state-of-the-art encryption and hashing algorithms. The source code can get viewed and downloaded on GitHub[8] (AGPL licenced) same as the JavaScipt Client implementation[9] (MIT licenced). The software can enable new applications but also re-engineer existing ones.

Organisations such as hospitals, universities, banks, insurers, retailers or manufacturers can also operate their own private or public instances of the DataVaccinator service.

Prospects

In the future, data protection will not be a current focus topic anymore but an established mode of data management. Generosity towards misuse, unawareness or mistakes will further decrease. At the same time the rick of **not** sharing depersonalized data will become clearer and more compliant and feasible solutions will be on demand to support a competitive Europe. Key sector for personal data protection solutions like DataVaccinator are Financial Services, E-Media, Smart home, E-Commerce, Industry 4.0, E-mobility, and of

---

[8] https://github.com/Kukulkano/dataVaccinato
[9] https://github.com/Kukulkano/vaccinatorJSClient

course E-health. To learn more about DataVaccinator please watch our interview with its co-founder Kurt Kammerer.

Name	DataVaccinator
Sector	IT and software services
Region	Luxembourg / World
Countries	Any
Time	Ongoing
URL	https://www.datavaccinator.com/
Business model	Open source, non-commercial
Participants	Spin off from Regify GmbH
Type of organisation	Non-commercial
Data sharing model(s)	Sharing of pseudonymised data that is stored separately from its identifiable attributes.
Core impact	Protection of data against abuse or errors that lead to data breaches violating GDPR. Separating the personal attributes from the data describing related content, like diagnostics is stored separately enabling the sharing of this information without sharing personal data. This is highly valuable for example for research in health and diagnostics.
Context	Open source software for secure and safe data storage and personal data protection using pseudonymisation of personal data at the point of its generation.