OASC MIMs

Helsinki Position Paper

From MyData Declaration to Proactive and Human-Centric Services

How does the City of Helsinki implement better services for citizens using MyData principles?

Mikko Rusama, Chief Digital Officer, City of Helsinki

Mika Huhtamäki., Deputy CEO, Vastuu Group Oy.

City of Helsinki is building up MyData operator capabilities? Why? What does it mean?

The City of Helsinki's digital transformation has an ambitious goal: to enable a transition from a reactive to a proactive city on citizens' terms. There are two major paradigm shifts dependent on each other; firstly transforming from a reactive to a proactive mode of operation and secondly giving people more power with their personal data. The goal is to provide better and easier-to-use services with better privacy. Proactive services cannot be implemented without implementing MyData operator capabilities.

EARLY DRAFT VERSION, WORK-IN-PROGESS,

Authors: mikko.rusama@hel.fi and mika.huhtamaki@vastuugroup.fi

Comments are welcome!

Towards sustainable and human-centric digital services

Data is a vital ingredient in most of our everyday services. Data can be used to drive the creation of faster, better and more personalised services in the public and private sectors. Rapid technology development has also given us a glimpse of what can be achieved if all the relevant data is offered to support better services, rather than storing it in isolated silos. In this brief introduction we would like to highlight some of our key views on this topic as

well as share some further insights into the reasoning underpinning these views.

Data regulation has not followed in lock-step with the incredible and ever-accelerating speed of digital development. This is causing some issues with offering better services to citizens. In some cases, these issues are valid and beneficial for adjusting the direction of public and private service development, but they can also cause major bottlenecks. One example of these issues is the gap in ways of using personal data between the digital and traditional paper-based worlds. Many of our rights to use and process personal data are accepted in “paper era processes”. However, the same rights may not always apply in digital contexts. This gap of accepted practicalities between the old and the new eras is something that has to be tackled and developed into common practice without limitations of the underlying new technology in contrast with the older paper-based implementations. Accepting digital as a norm is a key enabler for faster development of human-centric services in both private and public sectors. A person’s right to access and use their own data should not be dependent on the medium.

A second consideration in human-centric digital services derives from privacy and sustainability needs. Understanding these topics requires more than a basic understanding of the technical security and legal requirements. GDPR provides a good overarching framework and data security audits are important elements that should be implemented. However, we have to raise the bar higher to gain true sustainability beyond the GDPR regulations. Trust between all stakeholders is the main requirement that has to be taken into account in all areas. Trust is required between people, data registry holders, services, processes, public stakeholders as well as private sector players. This will require a range of activities in diverse areas of development, including business, legal and technology implementations in which it is necessary to define how trust is achieved. Earning and maintaining trust plays an essential part in the development and maintenance of human-centric services.

Here we aim to provide a wide yet concise overview of the current state, limitations and needs for digital service development, especially in smart cities from both the public and private sector points of view. By sharing our experiences in developing human-centric services we seek to encourage, provide inspiration and sow ideas for building a better digital future for all of us!

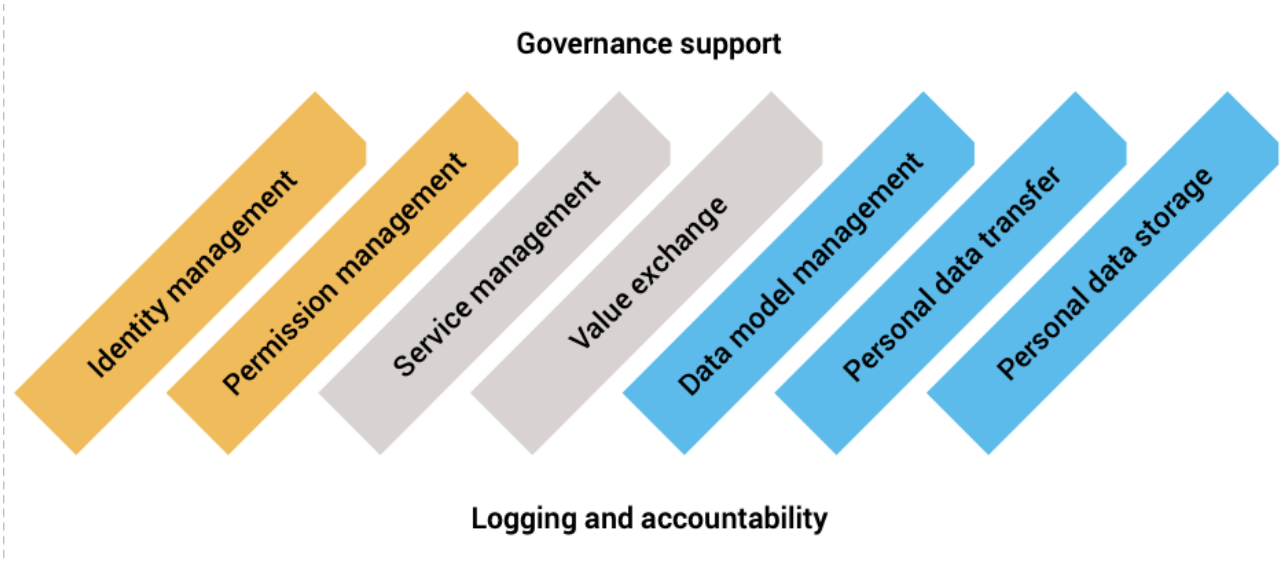
About MyData

MyData is both an alternative vision and guiding technical principles to empower individuals with their personal data.

The human-centric paradigm is aimed at a fair, sustainable, and prosperous digital society, where the sharing of personal data is based on trust as well as balanced and fair relationships between individuals and organisations.

The concept of MyData operators was introduced in the MyData white paper (Poikola, Kuikkaniemi and Honko, 2015) and the MyData declaration (MyData Global Network, 2017). The Understanding MyData Operator (2020) white paper further introduced the reference model of the MyData Operator.

The MyData operator reference model describes nine core functional elements of operators. Applying this model, Helsinki is focusing on the implementation of permission management, identity management and personal data transfer. Following text is a direct quote from MyData Global's "Understanding MyData Operator (2020)" document (page 16).



Identity management handles authentication and authorisation of individuals and organisations in different, linked identity domains and links identities to permis- sions.

Permission management enables people to manage and have an overview of data transactions and connections and to execute their legal rights. It includes maintain- ing records (notices, consents, permissions, mandates, legal bases, purposes, prefer- ences etc.) on data exchange.

Service management uses connection and relationship management tools to link operators, data sources, and data using services. Data can be available from different sources and can be used by multiple data using services.

Value exchange facilitates accounting and capturing value (monetary or other forms of credits or reputation) created in the exchange of data.

Data model management is about managing the semantics (meaning) of data, including conversion from one data model to another.

Personal data transfer implements the interfaces (e.g. APIs) to enable data exchange between the ecosystem participants in a standardised and secure manner.

Personal data storage allows data to be integrated from multiple sources (including data created by a person) in personal data storage (PDS) under the individuals' control.

Governance support enables compliance with the underlying governance frameworks to establish trustworthy relationships between individuals and organisations.

Logging and accountability entails keeping track of all information exchanges taking place and creating transparency about who accessed what and when.

The City of Helsinki is now transforming the MyData declaration into a real-world implementation together with Vastuu Group and other Finnish Cities. The City of Helsinki together with cities of Espoo, Oulu, and Turku received 2,25 M€ of funding towards implementing MyData operator capabilities.

From a reactive to a proactive city on citizens' terms

The vision of Helsinki, Finland, is to be the world's most functional city that makes the best use of digitalisation. The City of Helsinki's digital transformation has an ambitious goal: to enable a transition from a reactive to a proactive city on citizens' terms. There are two major interrelated paradigm shifts that we believe in:

1. From a reactive to a proactive mode of operation

If the city has 1) an obligation to provide a statutory service, 2) a citizen is eligible for the service, and 3) the city has all the required personal data and information regarding the citizen's likely service need, the city should provide the service automatically or de minimis recommend the most suitable service.

Our hypothesis is that proactive and personalised services are not only improving life and wellbeing of citizens but also saving costs. Solving problems early on is better than waiting problems to escalate and getting more complex.

2. From outright exploitation to human-centric use of personal data

Our digital footprint - the information about you and me - has dramatically increased. A significant number of data points that would have been considered private 20 years ago are

no longer private. There are many examples on how major social media platforms have misused personal data for commercial purposes eroding trust of citizens.

As of today, citizens have little power or even understanding regarding how their personal data is used. Even if required by law, “the right to be forgotten” is difficult, if not impossible, to achieve on the Internet. Revoking access to data is impractical if consent has been confirmed through paper forms.

Cities are data-rich and, as public bodies, they should lead the way towards a more human-centric Internet where citizens have better control over their personal data. Citizens should be able to provide and revoke consent relating to processing their personal data under specified purposes.

In the first phase, the City should provide citizens information on all the consents that they have given in a digital format. In the second phase, a mechanism should be implemented that enables to give, deny or revoke consent, as explained in Figure 1.

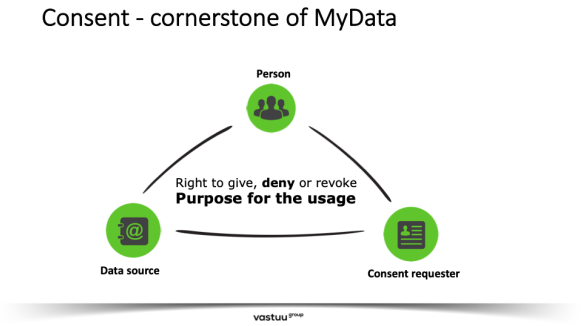


Figure 1. Towards human-centric data use

Implementing MyData operator capabilities is essential to achieving a more ethical and human-centric way of utilising data. Implementing MyData principles enables new proactive and personalised services as explained in the following

Preventative health care

The idea of preventative health care is not new. For example, in Finland, for many years women of 30-60 years of age are invited to take part in cervical cancer screening every five years. It has been estimated that screening avoids more than 250 cancer deaths in Finland each year. https://cancerregistry.fi/screening/cervical-cancer-screening/

The City of Helsinki's Social Services and Healthcare Division has developed a Health Benefit Analysis (HBA) tool in collaboration with The Finnish Medical Society Duodecim. This tool analyses a patient's health care data and applies a set of rules and 300 criteria in order to recommend the appropriate treatment.

All data is pseudonymised and derived from existing health records, giving medical professionals an overview of a patient's test results, previous diagnoses, medication history, and more, all without disclosing their identity. The data is analysed to highlight any "care gaps" that may exist where a patient is not receiving the treatment expected based on their health record. High-risk patients are prioritised based on need of intervention and invited to discuss their health issues with a doctor.
https://assets.bbhub.io/dotorg/sites/33/2020/10/BA_CityTools.pdf

Citizens should have the right to understand what social and health care information has been stored about him or her; and give consent to the City to proactively contact him or her if a health risk has been identified. As of today, patients can ban further analysis of proactive contacts after the initial contact by the health care professionals. Opt-out information is currently recorded to the old healthcare system's (Pegasos) risk registry.

Finnish health care data is also available in the national OmaKanta or MyKanta service in which citizens can browse their medical records and prescriptions and order repeat prescriptions in the online service. In OmaKanta, citizens can also

Table 1. Preventative health care

Use case description	City is analysing social and health care data and identifying care gaps and high-risk patients, and then proactively contacting these patients for treatment
Keychain (identifiers)	Social security number
What data is needed?	Health care data: age, gender, diagnosis, medication, laboratory results, information about the health care professionals
Purpose for the data use	Several use cases: • identification of the high blood pressure that may cause brain or heart attack; • Identification of the high blood sugar level indicating potential diabetes High-risk patients are prioritised based on need of intervention and invited to discuss their health issues with a doctor. According to GDPR, specific and legally valid reason is needed for data use
Permission (legal bases for data processing)	• Health care law • Consent for proactively contacting. NB: citizens should have right to deny (opt-out) analysis of their personal data aiming at analysing health risks and care gaps
Permission type	Permanent permission for proactive contacts (or one-time permission)
Data provider	City of Helsinki's Social and Health Care Division
Data consumer	City of Helsinki's Social and Health Care Division
Open issues	• Tietosuojavaltuutetun lausuntoa odotellaan huhti-toukokuussa • Pilottia laajennetaan • Integration to Apotti

Automatically checking the driver's licensing status

Helsinki City Construction Services, Stara, is providing services in the fields of construction, environmental management and logistics. In order to drive City of Helsinki governed vehicles - cars, heavy trucks or others- the employee is required to complete a training and possess an applicable driving licence and applicable permits.

Old way (As-is)

Currently, Stara's employees need to complete the mandatory training and physically visit their supervisor to show their driving licence. Supervisor validates the permits' validity in the face-to-face meeting.

New way (To Be):

Stara has developed a digital driving journal with an integrated booking calendar for each vehicle. Upon booking the vehicle, the validity of the driving licence and applicable permits is automatically verified directly from the registry holder.

Table 2. Use case description: checking the driving licence status

User story	As a vehicle driver my driving licence status and permissions are automatically checked with my own consent before every driving shift.
Keychain (identifiers)	Helsinki user ID mapped with the social security number (or Employee ID?)
What data is needed?	Driver's licence information

Purpose for the data use	Checking City vehicle users’ driving permission status
Permission (legal bases for data processing)	Consent or Contract MAY also be 'Legal obligation' in role of an employer. The jury is not out yet, according to Stara.
Permission type	Data flow - (vs. one time check/snapshot)
Data provider	Traficom - Finnish Transport and Communications Agency
Data consumer	Stara (City of Helsinki’s unit)

Personalised employment services

Unemployment has sharply risen due to Covid-19 pandemic. In the recent pilot, City has assumed the responsibility for the employment services from the state.

Current State (As-is):

A person visits an unemployment services’ career advisor who completes the situation assessment in a joint discussion. Career advisors help create the CV if that is not available and recommend suitable courses and job opportunities.

Future vision (to-be):

As an unemployed person, I can authorise the city to send me personalised recommendations on available courses as well as workplaces that help me take the next steps towards employment, reducing unemployment time.

With my permission, the City can use my data and AI to identify potential employment matches based on my CV data and help me to better describe my strengths to improve my

likelihood to succeed in the job application phase. A person consents their CV/resume information to be stored in the digital personal data storage. An algorithm analyses the content of the CV and suggests keywords to improve the matching probability.

A person consents their data to the city to recommend city's additional services (social, educational) for the benefit of the person.

User story	As an (unemployed) person I can give the city permission to analyze my work history, education, competences, certifications, CV contents and prior work seeking activities to recommend me suitable services, education and work opportunities to make finding new employment easier
Keychain (identifiers)	Social security number,
What data is needed?	CV containing: work history, education, competences, certifications category fit for social and educational services, data on prior service usage (eg. eligibility for helsinki-lisä)
Purpose for the data use	For unemployed people: increased likelihood of matching prospective employees and employment. Provide personalised recommendations on the city services based on the personal attributes. For employment services: prioritize customers to speed up process of taking up employment
Permission (legal bases for data processing)	Consent and agreements, law
Permission type	Periodical check
Data provider	TE-office

Data consumer	MATA, Elo (City of Helsinki's unit), Kasko, Sote

Subsidised daycare

In Finland, the daycare services are organised by cities, municipalities as well private operators. In 2019, over 26 000 children attended the City's day care services.

https://www.hel.fi/hel2/tietokeskus/julkaisut/pdf/20_12_31_tilastollinen_vuosikirja_2020.pdf

The family's income level affects the daycare fees as determined by the law.

Current state (As-is):

As of today, information of the family income has to be submitted by filling in the form and sending proof documents in pdf format either by mail or secure email as to the City of Helsinki's customer fee unit.

Future vision (To-Be):

As a citizen, I should be able to authorise the city to verify my annual income that may have a reducing impact on my child's daycare service fee. Automatically checking my income from the national income registry enables the city of Helsinki Customer fee unit employee to determine the correct applicable daycare fee with potential reductions for my child.

Table 3. Use case: subsidised daycare fee

User story	As a parent of a young child, I can get a recommendation on a subsidized childcare placement if my income is below the limit set by law. and I (together with my household adults) provide required justification documents required by the City's Unit in requested time limits.
Keychain (identifiers)	Social security number (parent and child)
What data is needed?	Income registry data, other income data
Purpose for the data use	Checking if the child is eligible for payment reduction.
Permission (legal bases for data processing)	Consent
Permission type	Periodical check (t (the eligibility is checked upon taking up daycare and every time family's financial situation or child's daycare needs situation changes.
Data provider	Incomes Register - Electronic Database for Incomes information by Finnish Tax Authority
Data consumer	Kasko (City of Helsinki's unit)

Cities are large-scale producers, providers and consumers of data

Cities of the 2020s are not just simple single operational units with administration, officials and statutory services, but a conglomerate of private and public companies, units, departments and subcontractors. The range of processed data is enormous from built environment, traffic, education, health to hobbies and personal preferences. This massive

amount of data is an increasingly valuable information source that can be used to improve our daily lives.

The City of Helsinki has approximately 500 statutory services, further to which it provides 200-300 voluntary tasks. City services are provided by four divisions: social services and health care, education, urban environment and culture and leisure division. An increasing proportion of city services are provided digitally, either wholly or in part.

Cities are massive data silos

The City of Helsinki has around 200 data registries, and approximately 450 technology systems containing personalised data. Thus, City controlled data is very siloed, sealed and isolated. The city itself has a challenge in even recognising all the available data silos.

Legacy systems and data interoperability

Before 2017, the City of Helsinki had 30+ independent agencies that for most part had a lot of independence and an IT unit of their own. City has an estimated 900 different technical systems, many of which are outdated legacy systems that do not provide API interfaces. Data is not easily moved from one system to another even if there is a valid permission to do so. Often data transfer is done via csv files that are moved from an ftp server.

Lack of digital and centralised permission management

As of today, citizens have little power or even understanding regarding how their personal data is used. In many cases, consent for data use has to be confirmed through paper forms. Even if the consent is given digitally, consent information is stored to different systems and there is no centralised place or service from which you can see all the different permissions given.

For example, health care data is available in the national OmaKanta or MyKanta service; as a citizen you can browse your own medical records and prescriptions and order repeat prescriptions in the online service. As a parent or guardian, you can also view the medical

records of a child under the age of 10.

In line with MyData principles, Citizens should be able to provide and revoke consent relating to their personal data that is needed for specified purposes. Maintaining this information in a digital format and providing a centralized service for viewing and changing the data permissions would be important.

Data exchange between cities and cities and public sector

To provide better services for the residents there is a need to interact between all the parties holding the person's data. Cities as a group also organize services for shared service domains where data sharing between the cities and the even at the state level is needed.

To enable needed data exchange it will be essential to agree common rules between the data sharing and using parties even when all the stakeholders represent the public sector.

To implement the human centric services the requirements for data exchange are identified from both city and state level. As an enabler city of Helsinki has recognized the need for a common rulebook for laying out the practises for the data sharing parties. Following examples reflects the data sharing needs in the Helsinki MyData Operator project.

Examples of data needs to exchange information with other cities:

- **Shared service domain organized as a joint venture or organization between cities**
- **Benefits governed by the city**
- **Educational and skills data governed by the city**

Examples of data exchange needa between cities and state:

- **Tax and income information from the Tax Authority is needed to get a discount from the daycare services**

- State level educational registries are needed to verify the student status
- State level information about the identity and identifiers
- Information about the trustee, mandates and caretakers

Providing transparent information

For human centric services transparency is a vital part for building trust between people and services. Providing transparency in an environment of different organizations is the vital point to earn trust from the citizens. Just the sheer number of different registries is overwhelming for the regular citizen. On top of that cities can have several processing criterias for the data, starting from consent and ending up to the public interest.

In all these categories cities can bring additional layers of trust through transparency of the data usage by using MyData Operator as a tool to provide transparency for all processing criterias even when the person cannot prevent the personal data usage.

Registries and processing principles of Helsinki

The City of Helsinki has about 200 data registries, file descriptions of which according to the EU General Data Protection Regulation are publicly available here.

Businesses and organisations that process personal data must provide individuals with information on the type of processing that is taking place and who is carrying it out. At a minimum, this information must clearly state:

1. **Who you (the organisation) are**
2. **Why you are processing the data**
3. **What legal basis you rely on to legitimise the processing**
4. **Whether or not the data will be transferred on to other organisations or individuals**
5. **How long the data will be stored**
6. **The existence of the individual's rights under data protection, including the rights to access, correction, erasure, restriction, objection and portability**

The City of Helsinki's data strategy sets a goal that "Data produced by Helsinki is the world's most usable and used city data by 2025". There are many legal and regulatory considerations in utilising data.

Legal and regulatory constraints on personal data use

The General Data Protection Regulation (GDPR) is the overarching regulatory framework that governs data protection and privacy in European Union (EU) countries. This regulation aims to provide individuals control over their personal data in a unified way across the EU.

What is personal data?

According to GDPR, 'personal data' means any information relating to an identified or identifiable natural person ('data subject'). In other words, "data that can be used to identify a person directly or indirectly, such as by combining an individual data item with some other piece of data that enables identification, are personal data. Persons can be identified by their name, personal identity code or some other specific factor."

Examples of personal data:

- **E-mail address, such as firstname.lastname@company.com**
- **Telephone number**
- **Identity card number**
- **Car registration number**
- **Positioning data (e.g. from a mobile phone)**
- **IP address**
- **Patient records**
- **A pet's veterinary records**
- **Data on the hereditary diseases of the person's great-great-grandparents**

In order to process personal data in Finland (and within the wider European Union) you must have a permission - a lawful basis - to do so.

According to Article 6 of the GDPR there are six available lawful bases for processing personal data:

1. **Consent - the consent of the individual that should always be freely given**

2. **Agreement-based - Performance of a contract;**
3. **Legal obligation - processing is necessary for compliance with a legal obligation to which the controller is subject**
4. **Public task - the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller**
5. **Vital interest - to protect the vital interests of the data subject (person) or of another natural person**
6. **Legitimate interest - the most flexible lawful basis for data processing, but not necessarily the most appropriate. Public authorities can only rely on legitimate interests if they are processing for a legitimate reason other than performing their tasks as a public authority.**

Data cannot be transferred from the data provider to the data consumer without a valid permission. A MyData operator explained in the following Chapters maintains information on the legal basis for the data processing.

Figure 2. Lawful bases for personal data processing

Legitimate Interest as a processing criteria

Law is one of the strongest forms of permission type for personal data processing.

For example, in Finland, women of 30-60 years of age are invited to take part in cervical cancer screening every five years. Some municipalities also invite women aged 25 and/or 65 for screening. It has been estimated that screening avoids more than 250 cancer deaths in Finland each year. https://cancerregistry.fi/screening/cervical-cancer-screening/

The processing of personal data is based on the following laws and regulations https://cancerregistry.fi/information/data-protection-notice/ :

- **Article 6(1)(c) of the EU General Data Protection Regulation (2016/679) - processing is necessary for compliance with a legal obligation to which the controller is subject**
- **the Act on the National Institute for Health and Welfare (668/2008) and the Government Decree on Screenings (339/2011).**
- **The processing of special categories of personal data referred to in Article 9(1) of the General Data Protection Regulation is based on Section 6(1)(2) of the Data Protection Act (1050/2018), the Act on the National Institute for Health and Welfare (668/2008) and the Government Decree on Screenings (339/2011).**

Agreement-based data processing

To obtain a library card, you need to make an agreement with the library. Personal data recorded in the register is processed for the purpose of organising library services including monitoring circulation, debt collection, statistics, communications, providing online and mobile services as well as self-service library and other library services and user identification.

https://www.helmet.fi/fi-FI/Info/Asiakkaana_kirjastossa/Rekisteriseloste(773)

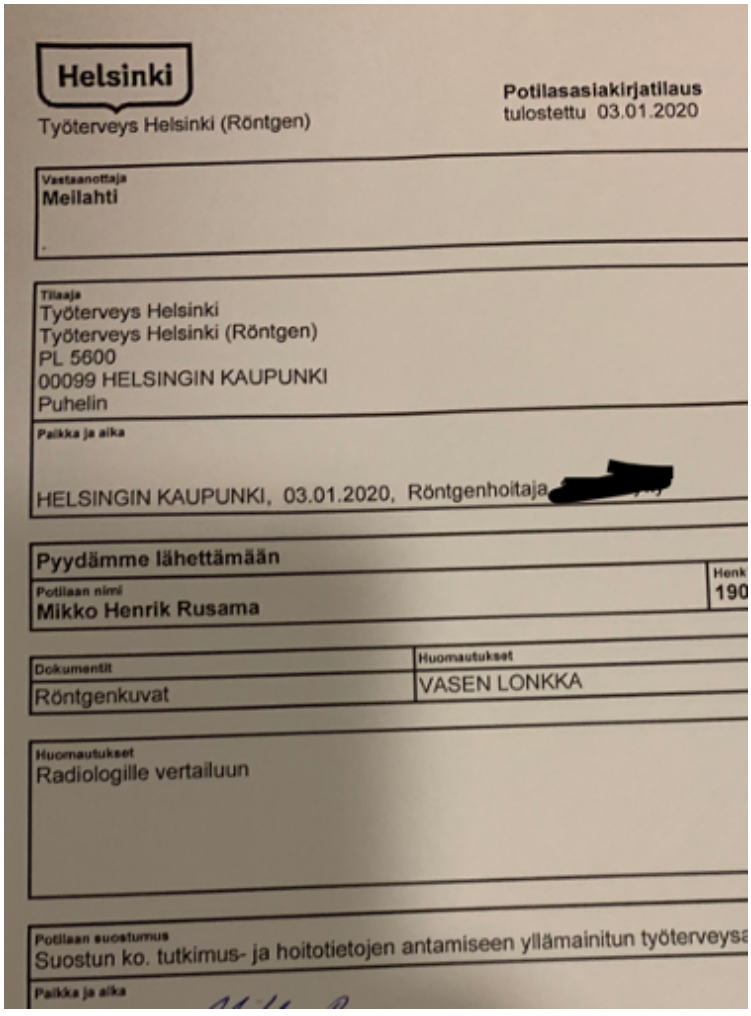
Data processing based on a freely given consent

Consent is defined in Article 4(11) of GDPR as:

“any freely given, specific, informed and unambiguous indication of the data subject’s wishes by which he or she, by a statement or by a clear affirmative action, signifies

agreement to the processing of personal data relating to him or her".

Currently, a consent is often provided through a paper-based form which data subjects can forget about. A further limitation is that there is no opportunity to revoke consent through the current consent process. For example, a consent is needed from the patient visiting the City's occupational health care unit to fetch X-ray images from Helsinki University Hospital (HUS) as depicted in Figure x. Another consent is needed, if HUS wants to use new x-rays taken by the Occupational Health Care unit (Työterveys Helsinki)



Helsinki

Työterveys Helsinki (Röntgen)

Potilasasiakirjatilaus
tulostettu 03.01.2020

Vastaanottaja
Meilahti

Tilaaja
Työterveys Helsinki
Työterveys Helsinki (Röntgen)
PL 5600
00099 HELSINGIN KAUPUNKI
Puhelin

Paikka ja aika
HELSINGIN KAUPUNKI, 03.01.2020, Röntgenhoitaja [redacted]

Pyydämme lähettämään

Potilaan nimi
Mikko Henrik Rusama

Henk
190

Dokumentit	Huomautukset
Röntgenkuvat	VASEN LONKKA

Huomautukset
Radiologille vertailuun

Potilaan suostumus
Suostun ko. tutkimus- ja hoitotietojen antamiseen yllämainitun työterveysa

Paikka ja aika

Figure. Consent form to fetch x-ray images for HUS

As another example, Bolt Works is a company offering gig-based job opportunities for drivers. If a digital consent is given for checking the driving licence information, job

opportunities can be provided faster. If a digital consent cannot be given, it is necessary to visit the company's offices to provide a copy of driver licencing documentation.

https://s3.eu-central-1.amazonaws.com/q.bolt.works/files/Bolt-Rekisteriseloste-Vuokrahenkilostorekisteri.pdf

Consent means giving people a genuine choice and real control over how you use their data. If the individual has no real choice, consent is not freely given. In this situation, the consent can be considered to be invalid.

Consent (of citizens or data subjects) cannot be regarded as freely given if there is a clear imbalance between the parties. Consent should not be bundled up as a condition of service unless it is necessary for that service:

If a public authority like the City is the controller of the data and asking for a consent, special focus should be paid on the question whether consent is "freely given". By default, there is always an imbalance of power if giving consent is the only way of getting the required service. Thus, if citizens are not willing to give consent for an automated service provisioning, there should always be an alternative way of getting the service - even if more tedious and manual.

https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/consent/what-is-valid-consent/

Ethical considerations

Ethical considerations in the city context has top cover all the relevant stakeholders and their view and expectations towards fair and sustainable data sharing. Obvious stakeholder in the middle is the person (data subject) but other parties have to be taken into account as well. Data source (data controller)

For example, there are both legal and ethical issues in the preventative health care

1. **What is the legal basis for data processing; and**
2. **Does the city have the right to proactively contact patients?**
3. **If the city has information about a potential health risk and does not act upon the data, is the city guilty of negligence?**
4. **Are citizens treated equally in the physical and digital worlds? If a person loses consciousness in real life, a bystander can call an ambulance without having explicit permission to do so. In the digital world, privacy concerns may prevent us from contacting citizens.**

Provided that a valid legal basis for preventative health data analysis can be found, citizens should have the right to understand what social and health care information has been stored about him or her; and give consent to the City to proactively contact him or her if a health risk has been identified.

Overview of the data and identity ecosystem

Currently, personal data management involves several roles defined in GDPR including data controller and data processor. In this document these roles are referred to as data sources that are responsible for collecting and managing the data.

However, GDPR does not specify identity providers (IDP) that provide either strong or weak identity services for 3rd parties. Weak authentication refers to the level of assurance of the identity and does not refer to the data security or privacy. Security and privacy may be at the same level in both weak and strong IDPs. The data using service that uses the personal has to identify the person, with the help of IDPs.

Digital Trust Infrastructure is a new shared capability for data verifications between data issuer and data verifier. This tech-legal infrastructure is developing rapidly and it will be an essential element to gain distributed trust in the future. Trust Over IP is one example of a fully functional trust framework.

MyData operator is positioned in between identity providers, data sources, data using services and data subjects. MyData operator's core task is to manage permissions but it can include features from IDPs or Personal Data Storages.

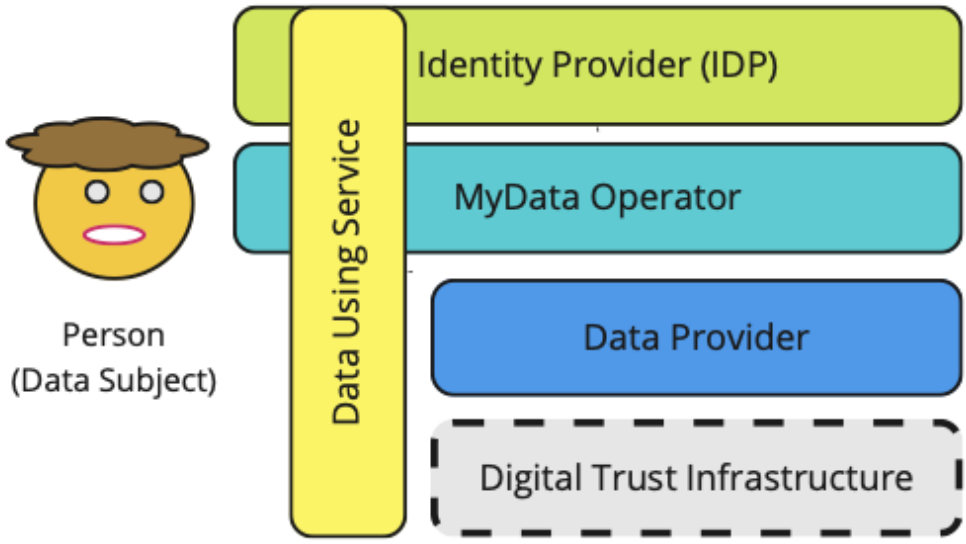


Figure: Overview of the current and new parties and their positioning

Roles and stakeholders

Mydata has grown from the need to get a better control and understanding of usage of personal data in digital services. Primary stakeholders of personal data are easy to identify but it is much harder to get a grasp on how data is used or distributed and how it is managed. This problem applies in both, public and private sectors.

Person (data subject)

The core actor in the human-centric approach is a person that is called a data subject in GDPR. Person manages the use of his/her personal data, for their own purposes, and maintains relationships with other individuals, services or organisations. Person is the single point of integration that enables or denies the data access.

Person's main rights in the mydata:

- **Give the right to access personal data**
- **Revokes rights to access the personal data**
- **Denies the right to access the personal data**

Examples of a person's roles in the city context:

- **It can be a resident**
- **A tourist or visitor in the city**
- **A foreign student**

Typically, consent is regarded as the primary processing criteria in mydata but the permission can also be based on the agreement (legitimate interest) or for the performance of tasks carried out in the public interest, see Chapter X.

Identity Management (Identity provider, IDP)

Identity Management handles authentication and authorisation of individuals and organisations. A human-centric approach starts from the identification of the person.

Identity management is handled by an identity provider (IDP) which creates, maintains, and manages identity information and provides authentication services both internally and to 3rd parties.

Identity providers may also aggregate both strong and soft identity providers into a single service. In Finland, national Suomi.fi aggregates strong IDPs and provides aggregated authentication services for the public sector.

Data source (data controller, data processor)

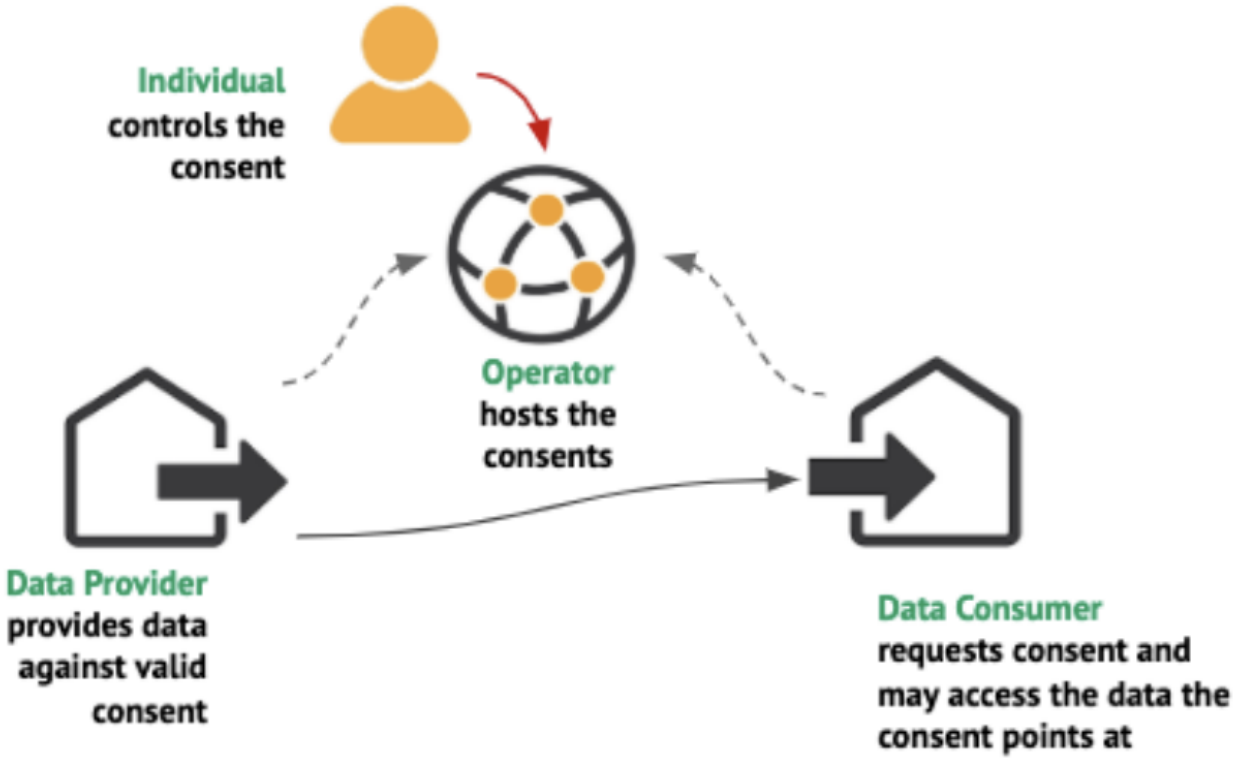
Data source collects, processes and distributes personal data. Typically, data source is an organization that needs the data to provide a service to its users. It can be a public organisation that stores personal data to offer a public service. In the public sector some organisations act solely as a controller of personal data without any services targeted for the citizens. Data controllers can act as a data source by granting access for 3rd party re-use with person's(data subject) consent.

In GDPR, the data processor processes personal data only on behalf of the controller. The data processor is usually a third party external to the company.

MyData operator (data intermediary)

MyData operator enables individuals to authenticate, securely access, manage and use their personal data, as well as to control the flow of personal data with, and between, data providers and data consumers.

The operator puts the person on the driver's seat doing the decisions on what the data should be used for. An operator enables entities to register as data providers and data consumers and provide this data connectivity as an ecosystem's service.



Categories of personal data

Personal data can be divided to three different categories:

- **Identifiers - typically fixed and always unique attributes that separate the data subject from others. Identifiers can be but are not limited to social security number, phone number, email, passport number or other similar personal attributes. Always verifiable.**
- **Identity data attributes - not unique but mostly fixed or almost persistent attributes such as date of birth, nationality or gender. Always verifiable.**
- **Profile data attributes - contextual data. This category covers all the other attributes of a data subject that are usually collected or generated for specific purpose or service needs. Profile attributes can be for example credit card usage information, heart rate, location, skills and qualification information. Please note that profile data attributes should not duplicate service related data, e.g. library related information**

like book loans should be maintained in the library service (service profile of the library)

A minimum set of core data is maintained by MyData operator

- **Identifiers are always unique. Differentiates data subjects from each others**
- **identity data attributes - fixed but not unique. Nationality, DOB**

Services are maintaining profile data attributes. However, service profiles should not store any any unique public identifiers but a generated contextual identifier key that is created by the MyData operator

Data clubs - ecosystem of data

MyData is human centric and allows people to control their personal data but it does not automatically mean that anyone can ask or receive consents to access all possible personal data. Initially it will be more probable that participants with similar regulation and goals form closed groups for sharing data. Data clubs is one way to describe closed data sharing ecosystems of data sources, data using services and operators.

Organizing rulebook controlled clubs allows better control over the usage and adds an additional layer of trust among the data sources and data using services. Data club is closely related to the interoperability and rulebook for managing data access and

Interoperability and Rulebook

Levels of interoperability at early stages in the smart city context are focusing on shared data sources and sharing the access to data sources between several operators. This will be governed within the Helsinki Trust Network Rulebook. Initial level of interoperability is defined through data source level connectivity which enables creation of a functional, controlled and agile data operator environment for future improvements. All connectivity related software and related agreement structures will be open sourced during the Helsinki MyData Operator project and more information can be obtained from the separate connectivity documentation.

Rulebook and Connectivity component binds the data sharing group together:

1. **Connectivity to different data sources**
    - **Commonly agreed access infrastructure to the data sources**
    - **Shared technica data source access between the trusted operators**
    - **Enables easy onboarding of existing data sources without major integration projects**
2. **Helsinki Trust Network Rulebook between the data sources and data using services**
    - **Common rules between the parties of the data sharing network**
    - **Rules binds the data providers, operators and data using services into a logical trust group**
    - **Principles and criteria on how data providers and consumers can join the MyData network**
    - **Rulebook defines what condition data can be used.**

Architecture overview

Several key capabilities are needed to implement a human-centric approach to personal data. This chapter briefly explains the key capabilities and design principles made in Helsinki.

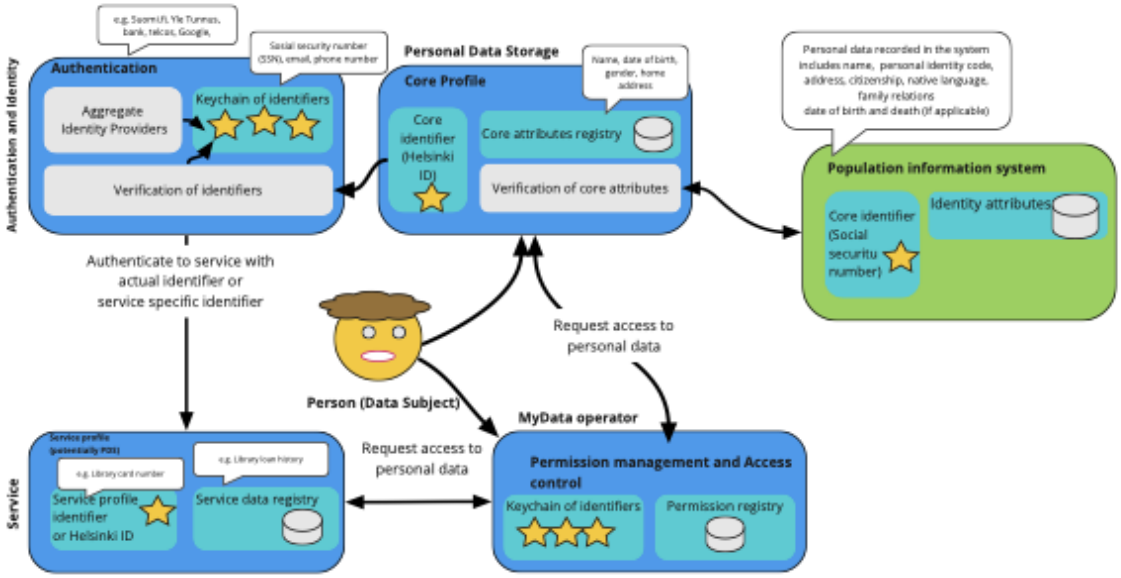


Figure: Architecture overview of the key capabilities

In the following the key capabilities are described in MyData terminology and in the parenthesis Helsinki name is used.

Authentication (Helsinki Tunnistus)

The centralized authentication service aggregates different identity providers. Aggregated identity providers includes but is not limited to:

- **Suomi.fi (strong national authentication services)**
- **Yle Tunnus (national soft authentication provided by the national broadcasting company**
- **Facebook connect (soft)**
- **Google login (soft)**

Authentication service verifies the identifier, generates identifiers if needed and manages and maintains the keychains. Different identifiers are collected to manage different identifiers in different services.

Personal Data Storage (Helsinki Profile)

Personal Data Storage (PDS) manages identifiers e.g. social security number, email and phone number and core attributes of the person (DOB), see "Chapter: Categories of Personal data.

Helsinki Profile is the PDS for every resident but can also be given to non-residents (like tourists). Helsinki Profile manages

- **Identifiers**
- **Core Attributes**
- **Preferences of the user**
- **Transaction information and audit trail from the services use (asiointitiedot)**

- **Embeds MyData operator capabilities and provides a holistic view on the permission information. Helsinki Profile integrates with the MyData Operator capabilities**

Helsinki profile is also a branded service and user interface to the information and functionalities as described earlier (compare with account.google.com)

MyData operator (scope in Helsinki)

In Helsinki, MyData operator provides the following services:

- **Identity management: aggregation of authentication methods and managing the keychain of verifiable identifiers such as Helsinki ID, social security number, email and phone number.**
- **Permission management: legal basis of the data usage (processing criteria) in the EU is based on:**
  - **Consent**
  - **Performance of a Contract**
  - **Legitimate Interest**
  - **Vital Interest**
  - **Legal Requirement**
  - **Public Interest**
- **Personal data transfer: MyData operator links the data sources and the data using services and ensures that the data using service has the permission to use the data, see legal basis for data processing.**
- **Service management (MIM4 connectivity) - enables the connection from the data source (data controller) to multiple data using services.**

  Service management will be described and specified in the MIM4 connectivity specification under the Open Agile Smart Cities association. The initial implementation called an Access Gateway will be published as an open source component. Access Gateway manages the data access audit trail from data source point of view. Includes some logging and accountability functions.
- **Service management (MIM4 connectivity) - enables the connection from the data**
- **Governance support**

Wallet is a user interface for centralised permission management that can also be embedded to services, e.g. in library service, users should only see relevant library-related consents.

Service Profile

City of Helsinki has 200+ registries in several service domains including Social Services and Health Care, Education, Culture and Leisure and Urban Environment. Personal data from those registries is used in more than 450 different service systems. Each service domain has specific requirements for storing personal data. Helsinki's enterprise architecture must support different domain specific requirements and data storages (registries) and promote mydata principles for sharing domain specific data.

Every service should maintain a service profile that contains:

- **Service Profile ID that is the primary key around which the data is modelled. By default, Helsinki ID should be used or a unique non-public identifier**
- **Service data registry - all the data the service is using or generating, e.g. book loans in the library service**

Service data should not be modelled around the public identifier like social security number, please see Design Principle X.

Service Profile can fetch additional personal data from the Helsinki Profile using the MyData service.

Design Principles

Principle 1. Centralized authentication service aggregates identity providers and creates a keychain of identifiers

A human-centric approach starts from the identification of the person. City of Helsinki's centralised authentication service "Helsinki Tunnistamo" aggregates both strong and soft Identity Providers into a single service for all service domains.

"Helsinki Tunnistamo" provides one single point of integration to all services of the city and enables the combination of different identity providers when needed. For example, a citizen can login to the service with her favorite social media credentials and once the strong identification is required she can update the session with the nationally supported strong identity providers.

The chosen approach also promotes international scalability by enabling a centralized support for the global identity provider services, such as Google, Apple, Alibaba or Facebook.

Principle 2. Helsinki core profile ("Helsinki profile") is the core personal data storage of every resident and non-resident.

Cities are controllers of a huge number of different data sources. Normally, most data sources share the person's core information. Consequently, the same personal data is asked for several times and maintained in multiple copies. This leads to unnecessary fragmentation and variations of the same core data. Without the centralised core profile, wrong or outdated personal data is being used in various services causing harm to citizens.

Centralized Personal Data Storage (PDS) for citizen's core data improves privacy, saves time and money for all stakeholders. Core data maintained in the PDS can be shared with the service domains. Maintenance and verification of the core data should be managed in a one centrally managed location where people can see, maintain and share their core data.

Helsinki Profile is a centralized data storage of personal core data for residents and non-residents from tourists to e-residents. Helsinki wants to avoid the misuse of public

identifiers and by following the city's data strategy, to ask for the personal data only once.

Helsinki Profile manages:

- **Identifiers are mostly persistent and always unique attributes that separate citizens (the data subjects) from others. Identifiers are always verifiable:**
  - **Helsinki ID (Helsinki Tunnus)**
  - **Social security number**
  - **Phone number**
  - **Email**
  - **Passport number**
- **Identity attributes are not unique but mostly persistent attributes that are always verifiable, e.g.**
  - **Name**
  - **Nationality**
  - **Gender**
  - **Address**
  - **Citizenship**
  - **Native language**
  - **Date of birth and death (if applicable)**
- **Additional identity attributes are usually attributes that can be shared with different service domains but are not specific to any particular service. e,g. location information, diet, hobbies and interests. Verification of these attributes is not required and can be difficult.**
- **Service related transaction and messaging - Helsinki Profile maintains a transaction log on all the services usage and acts as an "email box" for the service related messaging. Helsinki profile integrates with the national suomi.fi messaging service**

Principle 3. Service data is primarily modelled around the Helsinki ID

Typically, city services are mapped around the public national identifier (e.g. social security number) that is often the primary key. This approach has led to several privacy, maintenance and service accessibility related problems.

When social security number (SSN) is used as a primary key in services, it creates a potential privacy and security vulnerability. Showing or asking for a social security number is a privacy threat as it is widely used in all kinds of services both in public and private sectors. Identity thefts and misuse of the commonly used social security number is a relative common problem. People without the national SSN cannot use services in which SSN is used as a mandatory login key or core attribute.

Instead of using SSN as an identifier, non-public identifiers can save on maintenance. The planned changes to the SSN in Finland may require updating over 200 different registries just in Helsinki.

In the future service data will be primarily modelled around the Helsinki ID instead of the national social security number (SSN) or any other publicly used identifier. Even if the service needs the use of the public identifier, e.g. Health care services typically SSN, data should be modelled around the Helsinki ID that should be also the primary key. In other words, using and storing SSN to a service domain should be an exception; if absolutely necessary it should not be the primary key but an attribute. In services where privacy is vital, the Helsinki Tunnistamo can provide a pairwise ID that is unique to the specific service, and cannot be mapped with any other identifier between other service domains.

Principle 4. Internal and external parties have to agree on the rulebook of the “Helsinki Trust Network” in order to access data

Helsinki Trust Network is the network of trusted parties that comply with the rules as defined in the rulebook. Rulebook defines the rules for the data use in-line with the GDPR and MyData principles. Helsinki Trust Network and its rulebook ensures a transparent, ethical and human-centric way of using the data. All members of the “Helsinki Trust Network” have to comply with the rules of the rulebook. MyData operator enforces the rules as defined in the rulebook. If rules are violated, sanctions can be imposed and offenders removed from the network.

As an example, Helsinki Rulebook rules can dictate

- **All data sources and data using services have to comply with the GDPR, MyData principles and ethical guidelines**
- **All data using services need to have a legitimate reason for data processing as defined in GDPR that is checked and validated by the Mydata operator.**
- **Person should always be informed when his/her information has been accessed with a legitimate reason**
- **MyData operator must prevent access to data if there is no legitimate reason for data processing. If requested, a person should be informed when there has been an attempt to use his/her information without a legitimate reason.**
- **Only members of the Helsinki Trust Network can request consent from the person; giving consent for a non-trusted 3rd party is not allowed**
- **All legal terms must be human-readable and understandable explaining what does given consent mean personally**
- **If rules are violated, sanctions can be imposed and offenders removed from the network.**

Principle 5. Helsinki profile data can be shared with the trusted parties of the “Helsinki Trust Network”

With the person’s consent, data of the Helsinki Profile can be shared with the trusted parties of the Helsinki Trust Network. In addition to a freely given consent, data processing can be based on other provisions of the GDPR.

Principle 6. Service data can be shared with the trusted parties of the “Helsinki Trust Network”

With the person’s consent service data can be shared with the trusted parties of the Helsinki Trust Network. In addition to a freely given consent, data processing can be based on other provisions of the GDPR.

Principle 7. Agreements can be made between “Helsinki Trust Network” and other trust networks

Helsinki Trust Network can make agreements with other trust networks if both networks agree on the shared rulebook for technical connectivity and interoperability standards.

Helsinki Trust Network has a dedicated MyData Operator that, by default, all members should use. To avoid fragmentation and to favor interoperable and harmonised solutions, cities in data sharing should have one common national MyData operator network providing both standardized connectivity and permission management services. However, cities should be free to choose MyData operator of their own.

In addition, in a multi-operator environment technical connectivity and data harmonisation is needed. OASC's Minimum Interoperability Mechanism for Personal Data (MIM4) specification defines minimum requirements for all operators.

Principle 8. International data roaming is possible when the national trust networks agree on the common rulebook and connectivity standards

Roaming from one MyData operator network, e.g. Helsinki Trust Network is possible to another network if operator networks agree on the common rules, technical connectivity and interoperability standards. (comply with the rulebook).

In addition, in a multi-operator environment technical connectivity and data harmonisation is needed. OASC's Minimum Interoperability Mechanism for Personal Data (MIM4) specification defines minimum requirements for all operators.

References

MyDataShare White Paper:

https://kampanja.vastuugroup.fi/hubfs/MyData/MyDataShare-Whitepaper-v1.0.1.pdf

Kantara consent specification and adaptation in Finland:
https://tietomallit.suomi.fi/model/consent

Pan-Canadian Trust Network

https://www.trulioo.com/blog/pan-canadian-trust-framework

MyData Operators white paper: Understanding MyData Operator (2020)