

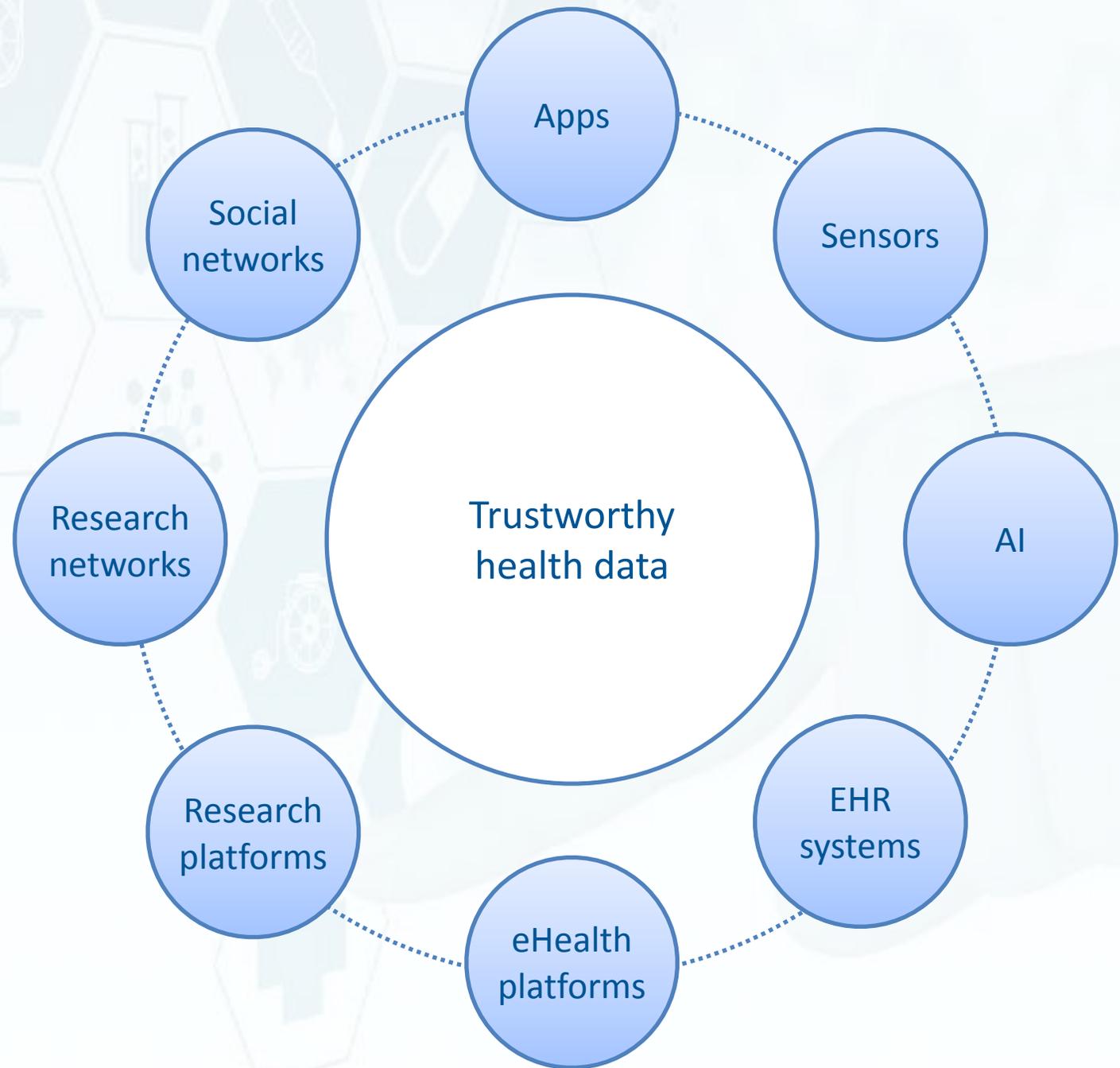
The European Institute For
Innovation Through Health Data

A RECIPE FOR TRUSTWORTHY DIGITAL HEALTH

Professor Dipak Kalra
President of i~HD

Trusted!
Quest for data-driven and fair health solutions
SITRA
3rd February 2021

The digital health toolbox



The spectrum of data use: from care to research

Individual level health data

EHR systems, apps, sensors, genomics,
Clinical Decision Support, AI

Used for:

- Health status monitoring
- Continuity of care (including the patient and caregivers)
- Care pathway tracking, clinical workflow management
- Real-time feedback and guidance to patients and clinicians
- Personalised medicine
- Disease interception, prevention and wellness

Population level health data

EHR systems, regional & national
eHealth infrastructures

Reused for:

- Quality and safety, care pathway optimisation
- Medical device and algorithm refinement
- Pharmacovigilance
- Public health surveillance
- Public health strategy
- Health services and resource planning

Big health data

national & international research
infrastructures,
federated query platforms
+ cross-sectoral services

Reused for:

- Digital innovation: devices, sensors, apps
- AI development
- Personalised medicine and bio-marker research
- Diagnostics development
- Drug development
- Disease understanding and stratification

TOWARDS TRUSTWORTHY HEALTH DATA ECOSYSTEMS

How the reuse of data can create new services for the benefit of all

Hannu Hämäläinen
Senior Advisor
Sitra

Saara Malkamäki
Specialist
Sitra

Ilkka Räsänen
Senior Lead
Sitra

Jaana Sinipuro
Project Director
Sitra

Artur Olesch (editor)
Digital Health Journalist

Data can improve patient outcomes, foster research and accelerate the development of new health services only if it is shared between stakeholders and reused while respecting privacy, strengthening people's control over their data and guaranteeing transparency.

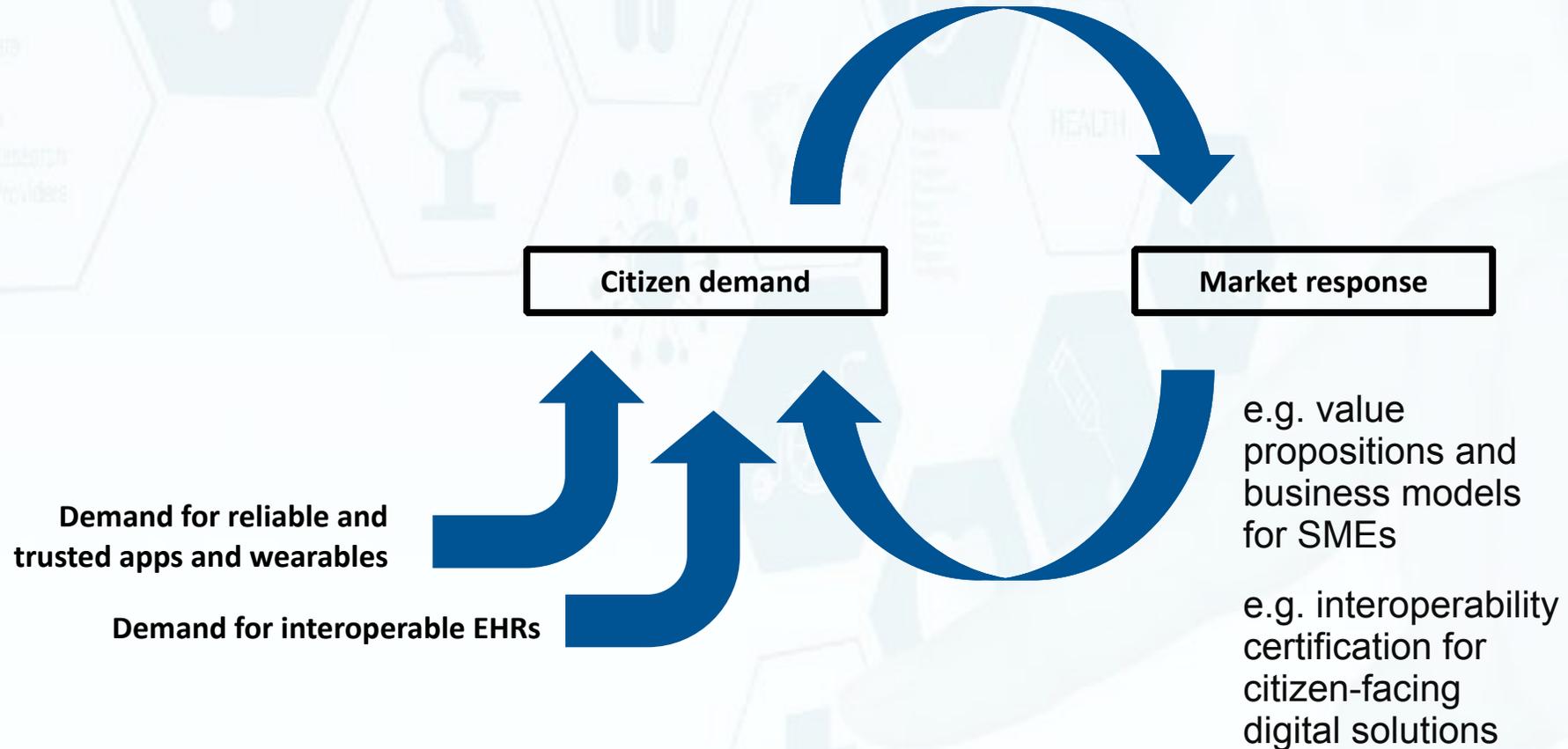
The working paper from the Finnish Innovation Fund Sitra provides guidelines for the future strategic use of sensitive data within European health systems. It addresses ethical, legal and self-regulatory frameworks for balanced data ecosystems, discusses the current challenges and establishes a set of recommendations for a new health and well-being data paradigm.

The Sitra report highlights some important challenges

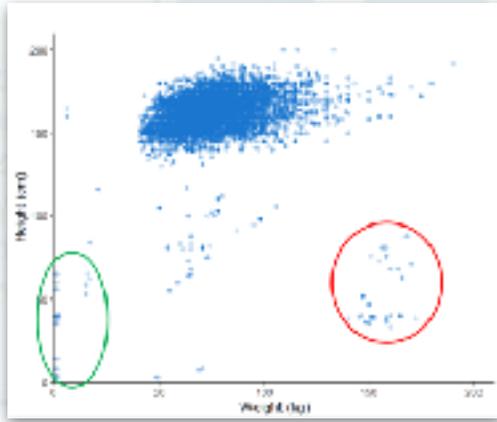
- Heterogeneous legislation
- Data incompatibility
- Delays in digitalisation
- Digital literacy
- Support for innovation
- Lack of trust and transparency
- Emerging threats
- Public health crises



Demand driven push for interoperability



Example data quality issues from hospitals



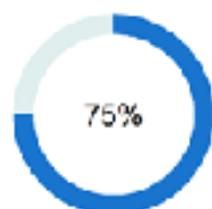
34% of weight errors led to medication-dosing errors
48% of these patients required additional monitoring, examination or treatment

Data Item	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7
Gender	100.00%	100.00%	100.00%	100.00%	100.00%	100%	100%
Case Status	99.87%	100.00%	60.00%	100.00%	100.00%	100.00%	100%
Date of Birth	100.00%	100.00%	99.00%	NA	100.00%	100%	100%
Admission date	100.00%	100.00%	100.00%	NA	100.00%	99.53%	58%
Diagnosis Text	50.46%	84.02%	100.00%	100.00%	98.05%	100.00%	14%
Diagnosis Code	50.46%	84.02%	100.00%	100.00%	98.05%	100.00%	14%
Discharge date	100.00%	100.00%	100.00%	NA	100.00%	100.00%	58%
Diagnosis Date	50.46%	84.02%	100.00%	100.00%	100.00%	NA	13%
Potassium in serum	52.38%	28.27%	100.00%	NA	100.00%	100%	49%
Sodium in Serum	52.38%	27.70%	100.00%	NA	100.00%	100%	49%
Platelets Blood	52.78%	33.14%	63.73%	NA	100.00%	100%	45%
SGPT (ALT) in serum	33.61%	22.29%	100.00%	NA	100.00%	100%	47%
Total Protein in serum	52.37%	14.96%	86.53%	NA	100.00%	100%	47%
Total Bilirubin in serum	33.03%	16.99%	100.00%	NA	100.00%	100%	47%

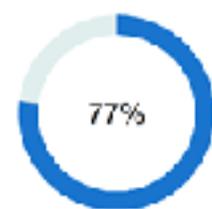
Data quality variation across example EU hospitals
- the data most needed for clinical research (partial table)

- Most EHR data is captured by busy junior staff, using various EHR systems
- Staff have no access to training in data quality
- Patients also have no training! (but their data is becoming increasingly important)

Completeness



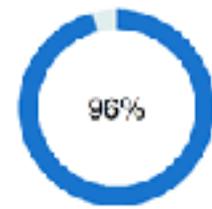
Consistency



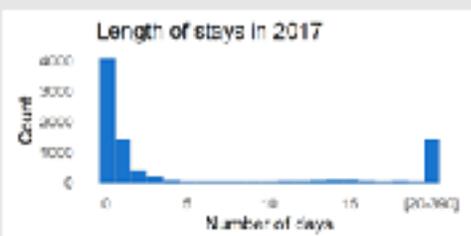
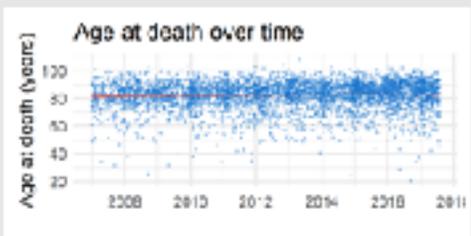
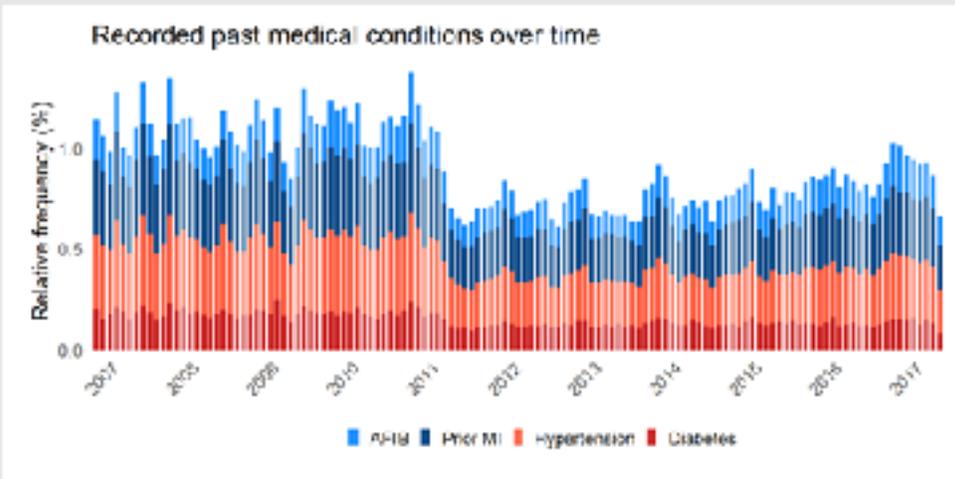
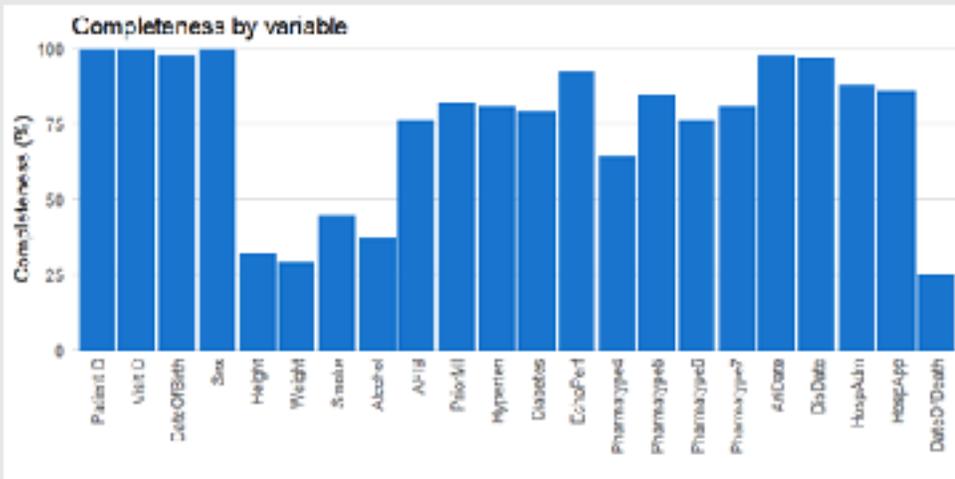
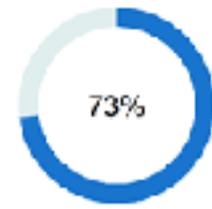
Correctness



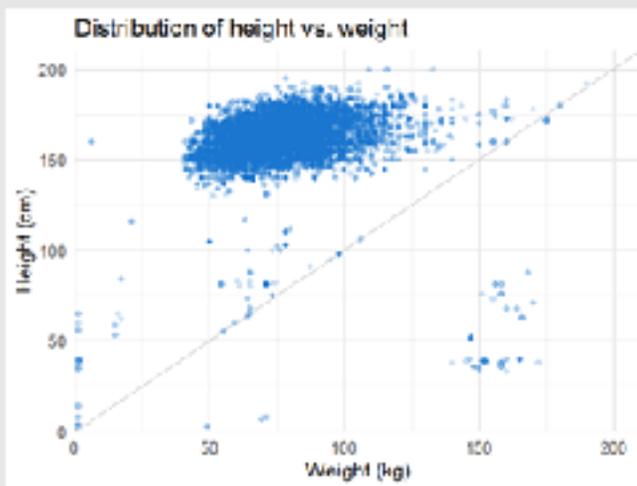
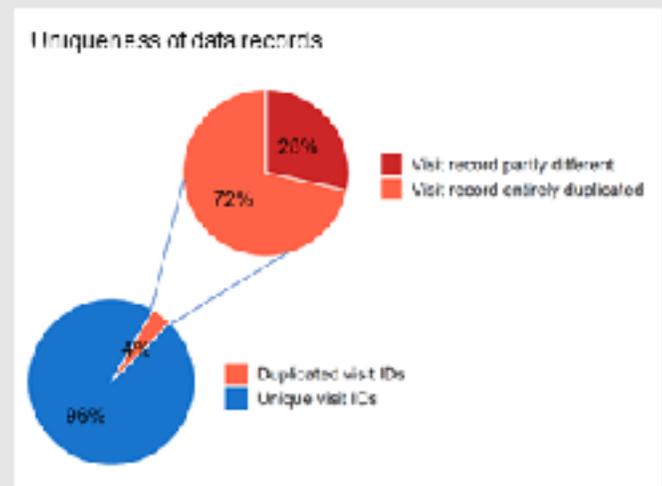
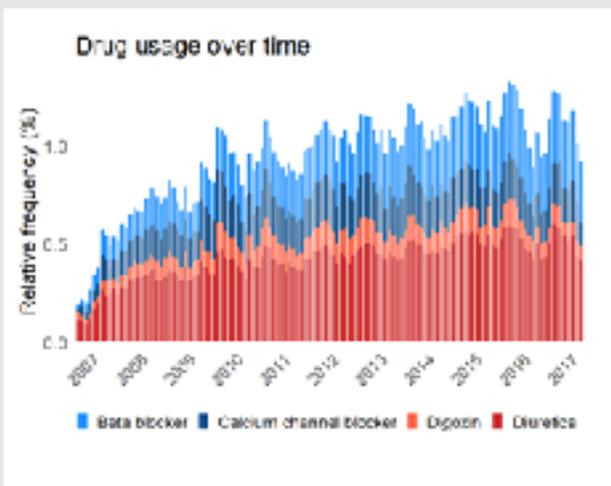
Uniqueness



Stability



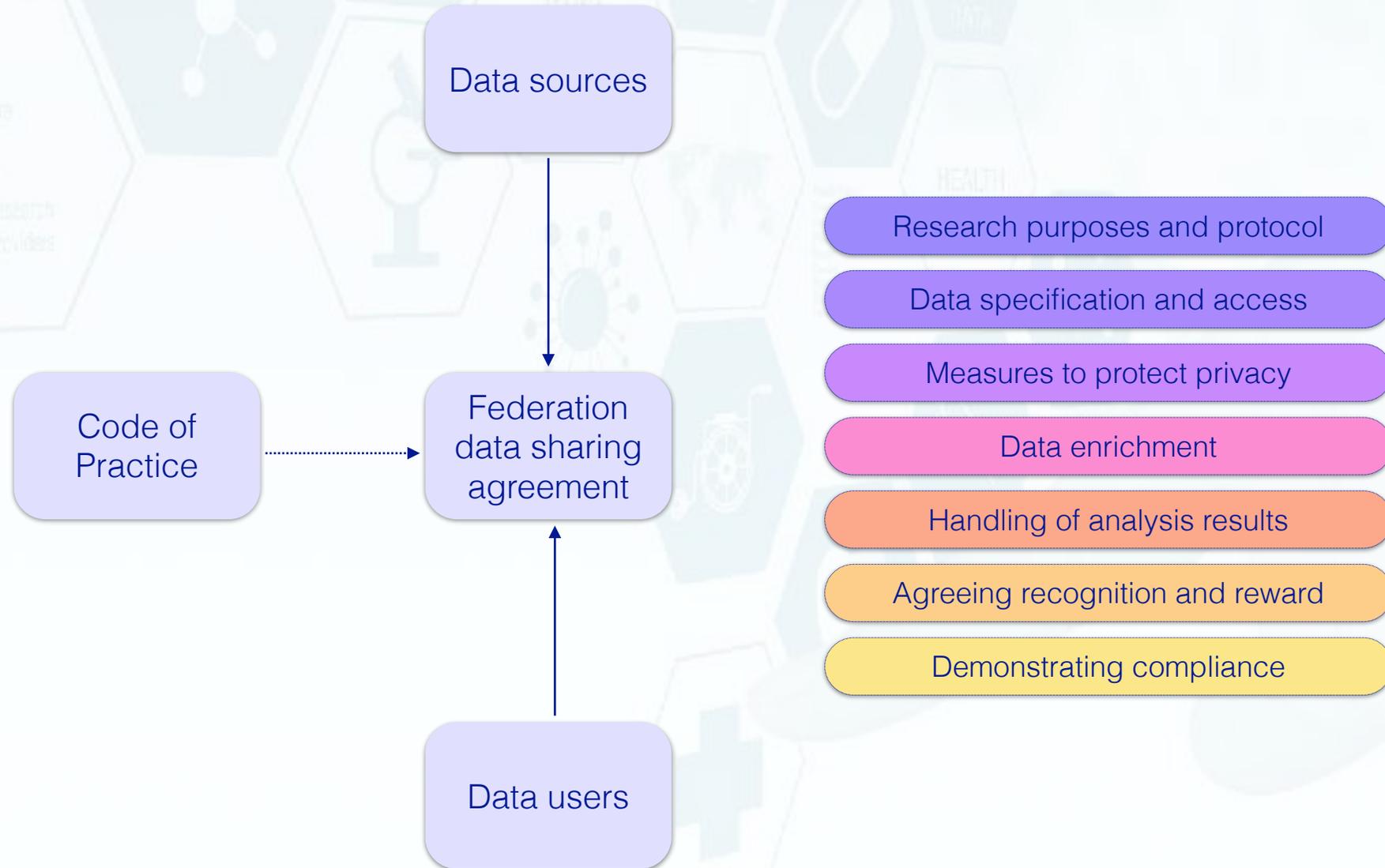
The European Institute for Innovation through Health Data



The FAIR principles: a commitment by data sources

- Data must be **Findable**
 - a searchable method to discover resources, with standardised metadata and a repository identifier
- Data must be **Accessible**
 - retrievable metadata, and potentially retrievable data via appropriate protocols and controls
- Data must be **Interoperable**
 - metadata is standardised, data conforms to relevant published standards
- Data must be **Reusable**
 - there is transparency about the terms under which the data may be reused

Data sharing agreements protect all parties



Some “grey” GDPR compliance areas researchers (and DPOs) struggle with

- Is new consent really needed when a large data set is reused? (every time?)
- How precisely should purpose be explained, when reuse is intended?
- How maximal can a data set for research reuse be?
- Can anonymised data retain one way linkage, for updates?
- What is acceptable anonymisation for rare disease groups, genomics, images...?
- What is required for a data set to be open?
- Can we have standards for what is “adequate” or “proportionate”?
- What data protection assurances can realistically be required, assessed, enforced?
- Can we create a level European playing field for those who wish to reuse data?

**We should strive for an EU aligned code of practice
The Data Governance Act may help**

The challenge with gaining public acceptance of health data reuse

Individual level health data

EHR systems, apps, sensors, genomics, Clinical Decision Support, AI guidance

Used for:

- Health status monitoring
- Continuity of care (including the patient and caregivers)
- Care pathway tracking, clinical workflow management
- Real-time feedback and guidance to patients and clinicians
- Personalised medicine
- Disease interception, prevention and wellness
- Healthcare provider reimbursement

Population level health data

EHR systems, regional & national eHealth infrastructures

Reused for:

- Healthcare provider performance and planning
- Quality and safety, care pathway optimisation
- Medical device and algorithm refinement
- Pharmacovigilance
- Public health surveillance
- Public health strategy
- Health services and resource planning

Big health data

national & international research infrastructures, federated query research platforms + cross-sectoral infrastructures & services

Reused for:

- Disease understanding and stratification
- Digital innovation: devices, sensors, apps
- AI development
- Personalised medicine and bio-marker research
- Diagnostics development
- Drug development

Decreasing public understanding of why and how data are used

Increasingly unfamiliar data users

Increasing distance of data results from the patient

Increasing time from data use to demonstrated value

Perceived lessening choice and greater cybersecurity risk = harder to trust

PUBLIC OPINION IN THE EU

I would like to know as much as possible about my DNA

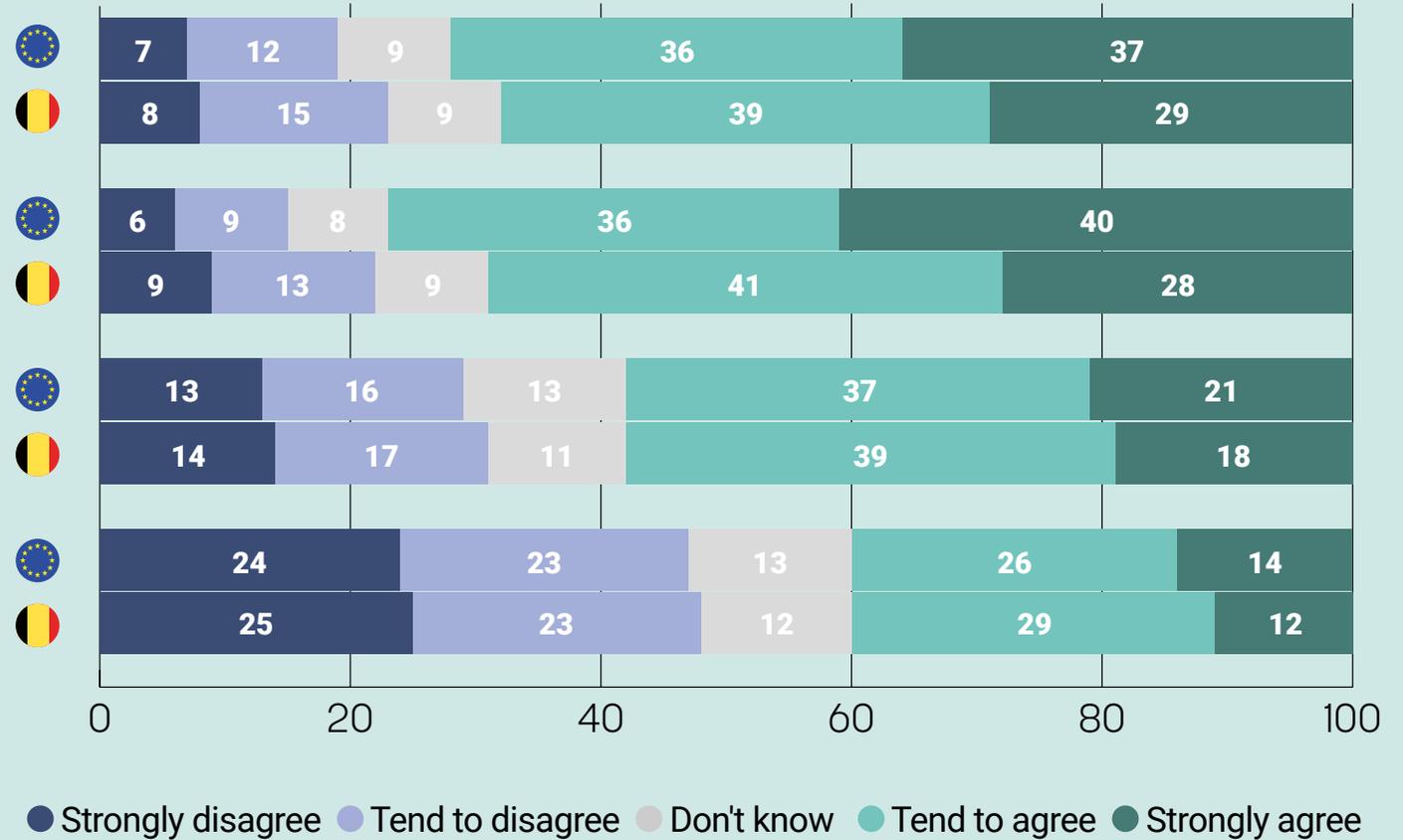
I'm willing to share my genetic information with my family

I'm willing to share my genetic information with academic researchers

I'm willing to share my genetic information with pharmaceutical companies

n = 28,782

Weighted to be representative of the adult EU population



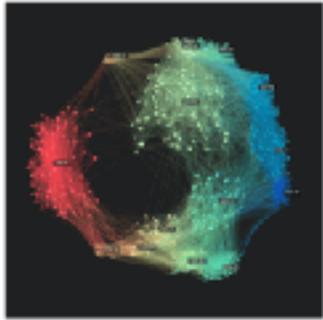
Citizens' Juries have highlighted...

Patients
Health Care
Hospital
Physicians
Clinical Research
Service Providers



...that the public are concerned about the use of health data even if their identify is not exposed

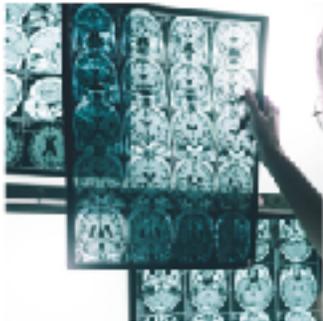
Transparency about the uses of health data



Big Data' used for the early identification of other diseases associated with cancer

A novel computer programme to help doctors and scientists to better understand which other diseases are likely to occur in patients with cancer.

[Learn more](#)



Helping the Fight Against Alzheimer's

Discovering a test to detect those at risk of developing Alzheimer's Disease

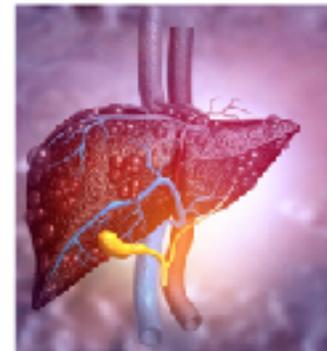
[Learn more](#)



Early Detection of a Cancer Killer

Helping identification of people at high risk of colorectal cancer from an existing blood test

[Learn more](#)



Using health data to explore a link between Body Mass Index and non-alcoholic fatty liver disease

Analysis of over 2 million anonymous patient records to establish how increasing body weight approaching obesity increases the risk of developing liver disease

[Learn more](#)

<https://datasaveslives.eu>

How do we reach societal acceptability?

- Data protection regulations prioritise the **rights of the individual** to privacy
- Clinical research can bring important benefits **to society**
- Many surveys indicate **patients are in favour** of their data being re-used for research
 - But the GDPR hype sometimes breeds fear (public, DPOs, CEOs...)
- The public need **greater transparency** about why and how health data are used, safeguarded, and the benefits of that use

We need to find the right balance between the rights of the individual and the benefits for society

A holistic approach is needed

Patients
Health Care
Hospital
Physicians
Clinical Research
Service Providers

Showcasing the benefits of research using health data

Public perceptions, preferences, priorities

Defining the elements of trust

Principles, codes of conduct, security controls

Consensus on
- kinds of health data
- purposes of use
- types of user

Capturing the outcomes from data use

Transparency in the uses of health data

Evidence and accountability to the public and professionals

Assuring the trustworthiness of data custodians and data users