SENSITIVE DATA ARCHITECTURE AND SERVICES

RDARI 25.10.2019, Espoo

Ville Tenhunen, @vtenhunen

AGENDA

A need

About terminology

Data protection and data security

Sensitive data could be FAIR

Architecture framework

Some possible solutions

A NEED

User need:

Researchers have a need to capture, store, share and distribute sensitive data

PERSONAL DATA

- "Personal data is any information that relates to an **identified or identifiable living individual**."
  - "Personal data that has been de-identified, encrypted or **pseudonymised** but can be used to re-identify a person remains personal data and falls within the scope of the GDPR."
  - "Personal data that has been rendered **anonymous** in such a way that the individual is not or no longer identifiable is no longer considered personal data. For data to be truly anonymised, the anonymisation must be irreversible."
  - "The GDPR protects personal data **regardless of the technology used for processing that data** – it's technology neutral and applies to both automated and manual processing, provided the data is organised in accordance with pre-defined criteria (for example alphabetical order). It also doesn't matter how the data is stored – in an IT system, through video surveillance, or on paper; in all cases, personal data is subject to the protection requirements set out in the GDPR."

https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

SENSITIVE PERSONAL DATA

"The following personal data is considered ‘sensitive’ and is subject to specific processing conditions:

- personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs;
- trade-union membership;
- genetic data, biometric data processed solely to identify a human being;
- health-related data;
- data concerning a person’s sex life or sexual orientation.

https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en

SENSITIVE DATA

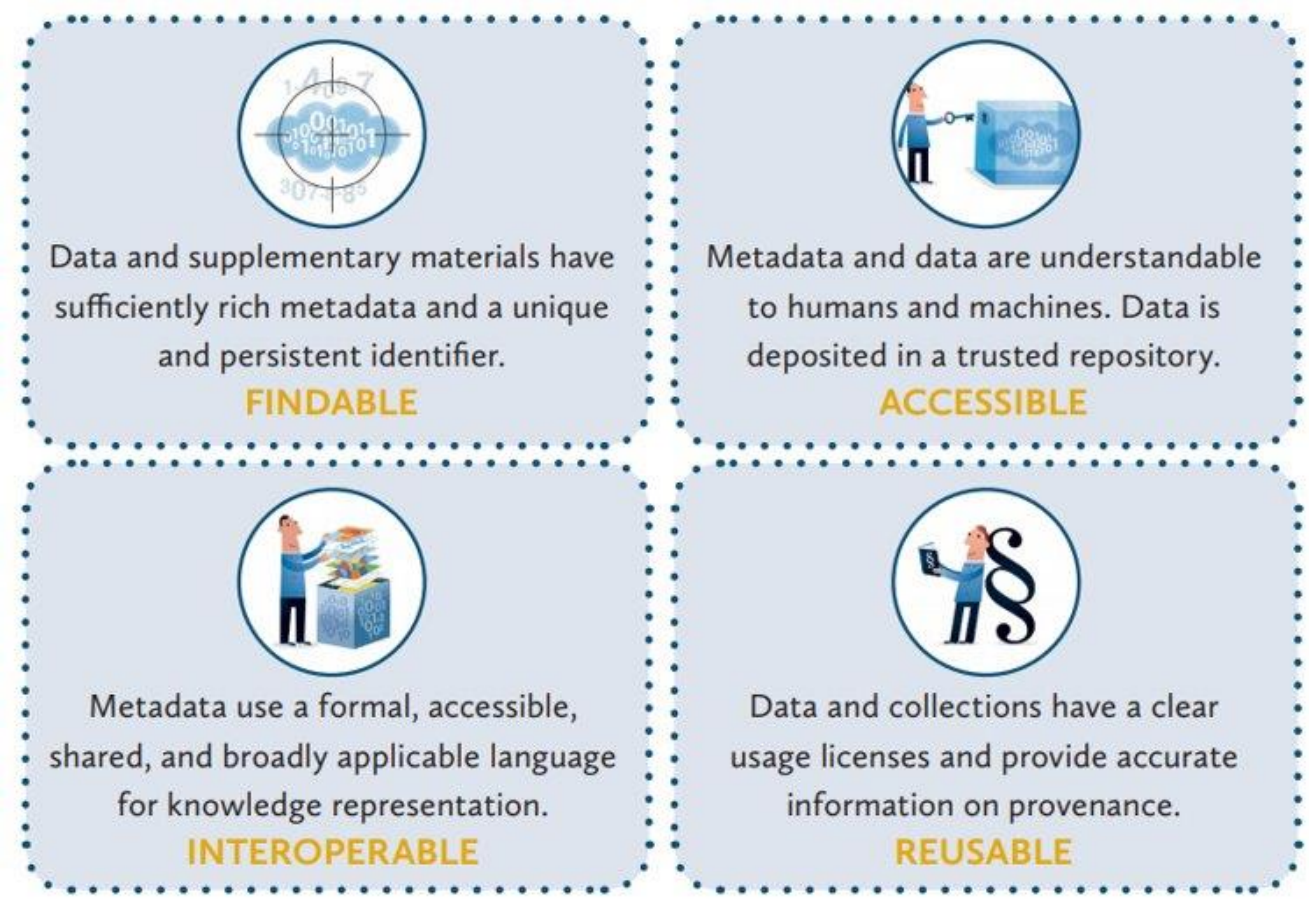
- Classified information by some law
  - For example:Habitats of endangered species
- Economical reasons
  - Patents
  - Innovations
- etc.

DATA PROTECTION =! DATA SECURITY

- It is not possible to lead security demands from the data protection regulations and vice versa
- "Data protection is a fundamental right that safeguards the rights and freedoms of data subjects when personal data is processed."[1]
  - There is not any technological tools
- Data security based on the risk assesments and content of the data
  - Technology
  - Organisational security
  - Physical security
  - Governance
  - etc.

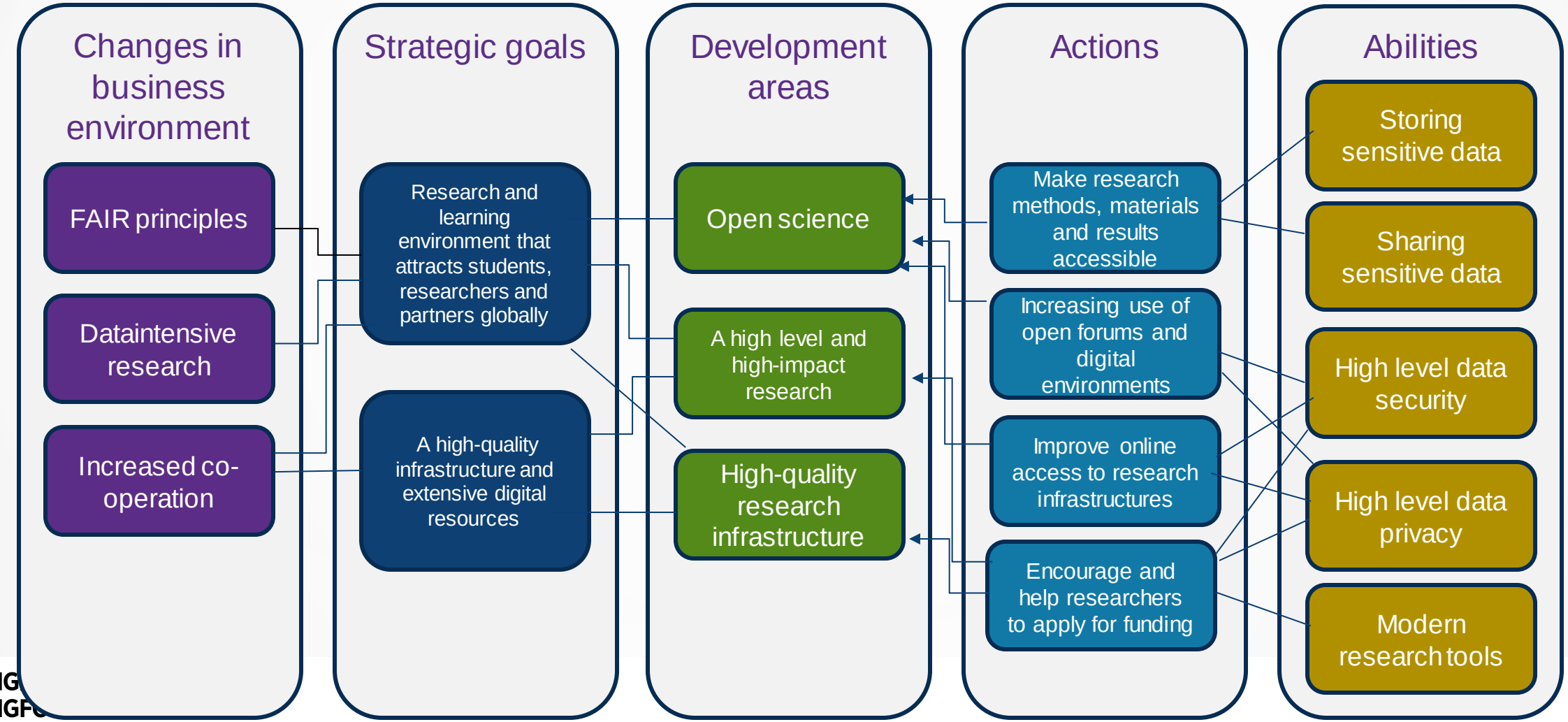
1) https://tietosuoja.fi/en/data-protection

SENSITIVE DATA COULD BE FAIR



Source: Implementing FAIR Data Principles: The Role of Libraries https://libereurope.eu/wp-content/uploads/2017/12/LIBER-FAIR-Data.pdf

ARCHITECTURE FRAMEWORK



SOME QUESTIONS

- AAI
  - Access and authorisation
  - Collaborators can't always visit in Helsinki for the registration and identification – how to arrange
- Logs and user tracking
  - Who can see and what
- Encryption
  - Server side and/or end to end; Where are keys?
  - How much CPU is enough?

SOME POSSIBLE SOLUTIONS

- Umpio; NAS solution
  - Encryption
  - No haring possibilities
- Ceph + Nextcloud
  - End to end encryption?
  - Server side encryption?


THANK YOU!