Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Bundesamt für Statistik BFS
Office fédéral de la statistique OFS
Ufficio federale di statistica UST
Uffizi federal da statistica UST

SDSC
EPFL ETH zürich

Data Science throughout the Swiss government

ROAD to BERN

Data Science Competence Center for the entire federal administration

Prof. Dr. Bertrand Loison

Vice-Director General at the Federal Statistical Office (FSO)

Head of Data Science Competence Center (DSCC)

United Nations WORLD DATA FORUM
BERN, SWITZERLAND
3–6 OCTOBER 2021

Webinar Data Science and Official Statistics

New-York (online), 30 June 2021

DSCC's Storyline

2015

Internal FSO working group on big data

2017

Data Innovation Strategy Pilot projects

2019

2021

Data Science Competence Center (DSCC)

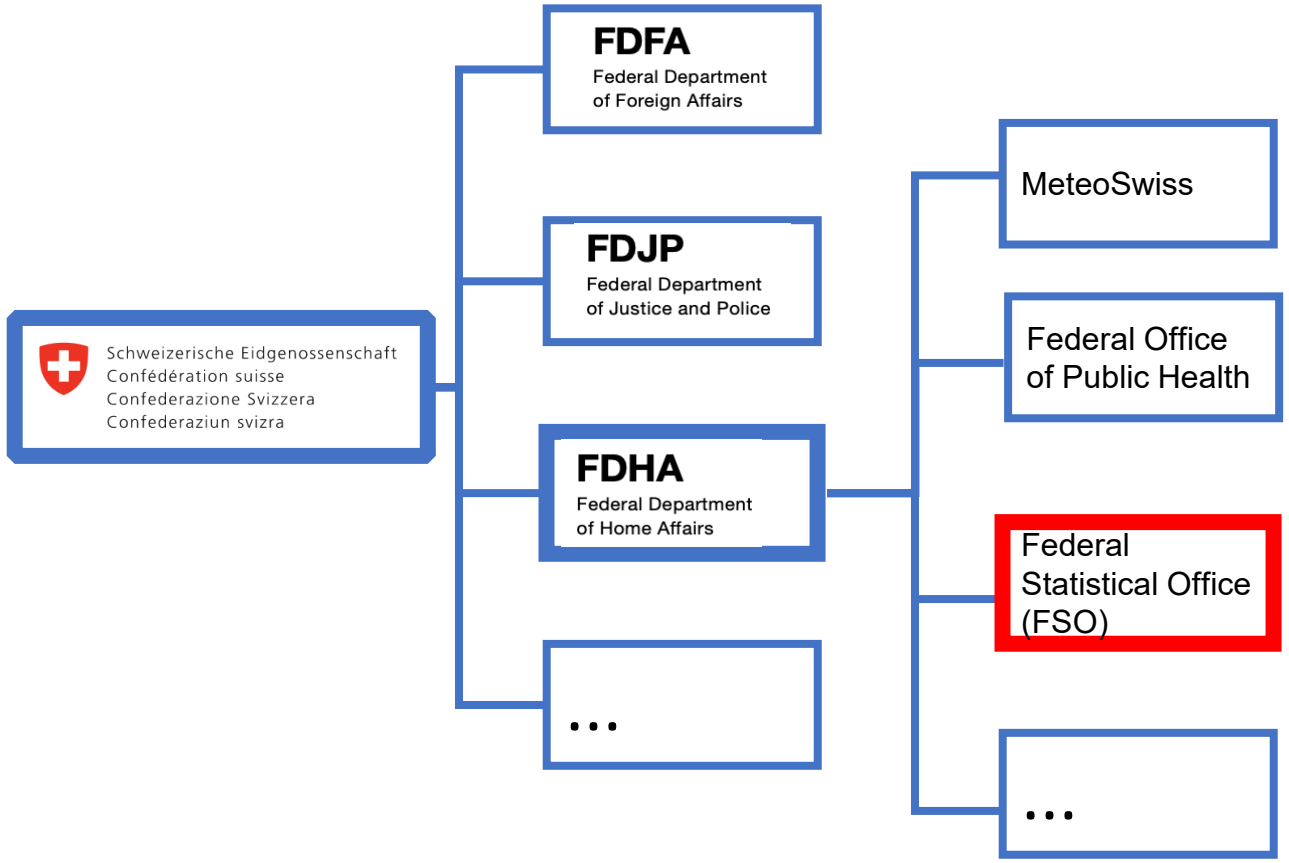
2024-...

Data Science & AI Competence Center (DSCC)

Agenda

1. **Organizational perspective**
2. **Analytic services perspective**
3. **Data perspective**

What does it mean to be a competence center ?



Structure of the Federal Administration

- **7 government departments**
- **Close to 90 federal agencies**
- **38,000 employees**

Topics covered by the Federal Administration

- **Public transport**
- **Energy**
- **Migration**
- **Finance**
- **IT**
- **Health**
- **Social security**
- **Justice**
- **Telecommunication**
- **Statistics**
- **…**

Coordination with regional administrations

- **26 cantons**
- **2200 communes**

Why Data Science and AI for Public Good at the FSO?

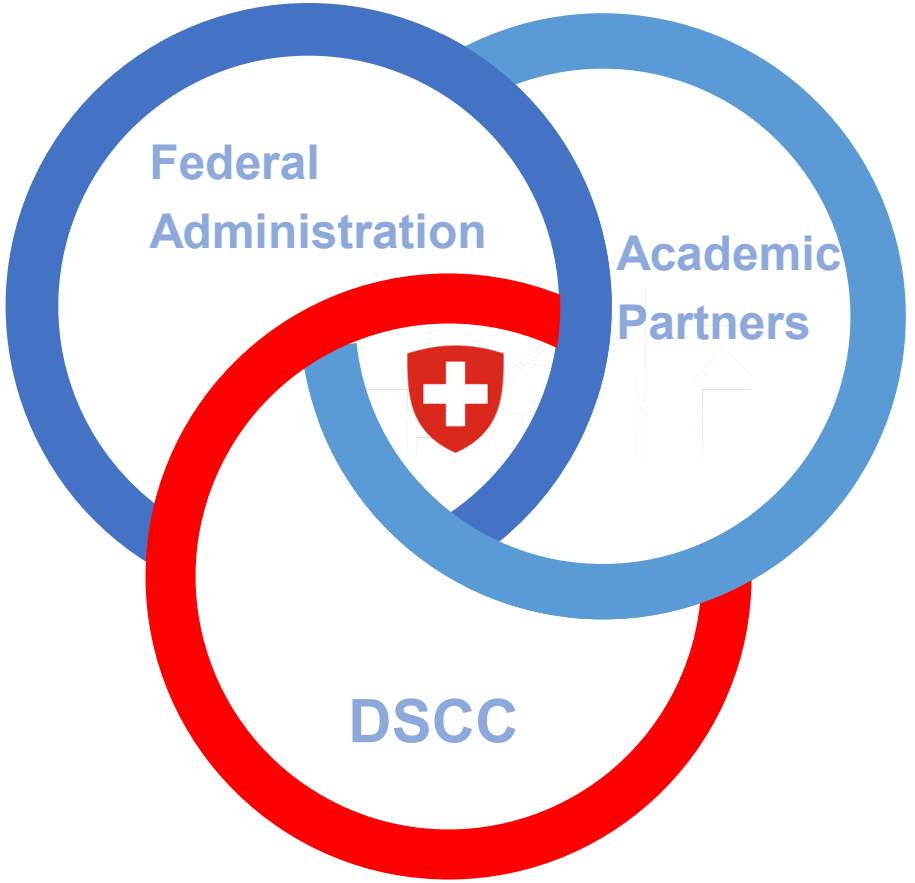
Because analytics (*learning from data*) is FSO's **core business.**

Because the FSO has **much to gain** itself from big data and data science.

Because, internationally, **official statisticians** have been dealing with big data and data science for years.

Because the FSO has **practical experience** in statistics and data science.

Because the FSO has to respect the **Code of Practice** and **UN's Fundamental Principles of Official Statistics**.



What is the DSSC's Unique Selling Proposition (USP)?

"Support for government use of AI correlates moderately with trust in government. Trust in institutions is essential if governments are to gain the support needed to roll out AI capabilities."

Source: Boston Consulting Group (BCG), The Citizen's Perspective on the Use of AI in Government

DSSC's Core values

In the implementation of all data science services, utmost importance is attached to the core values of information security, data protection, data security, data governance, non-discrimination, explainability, transparency, reproducibility, neutrality, objectivity, ethical handling of data and results, and public confidence.

Promoting community and knowledge building

Academic partners

Partnerships with the Swiss Data Science Center (EPFL/ETHZ) & Universities:

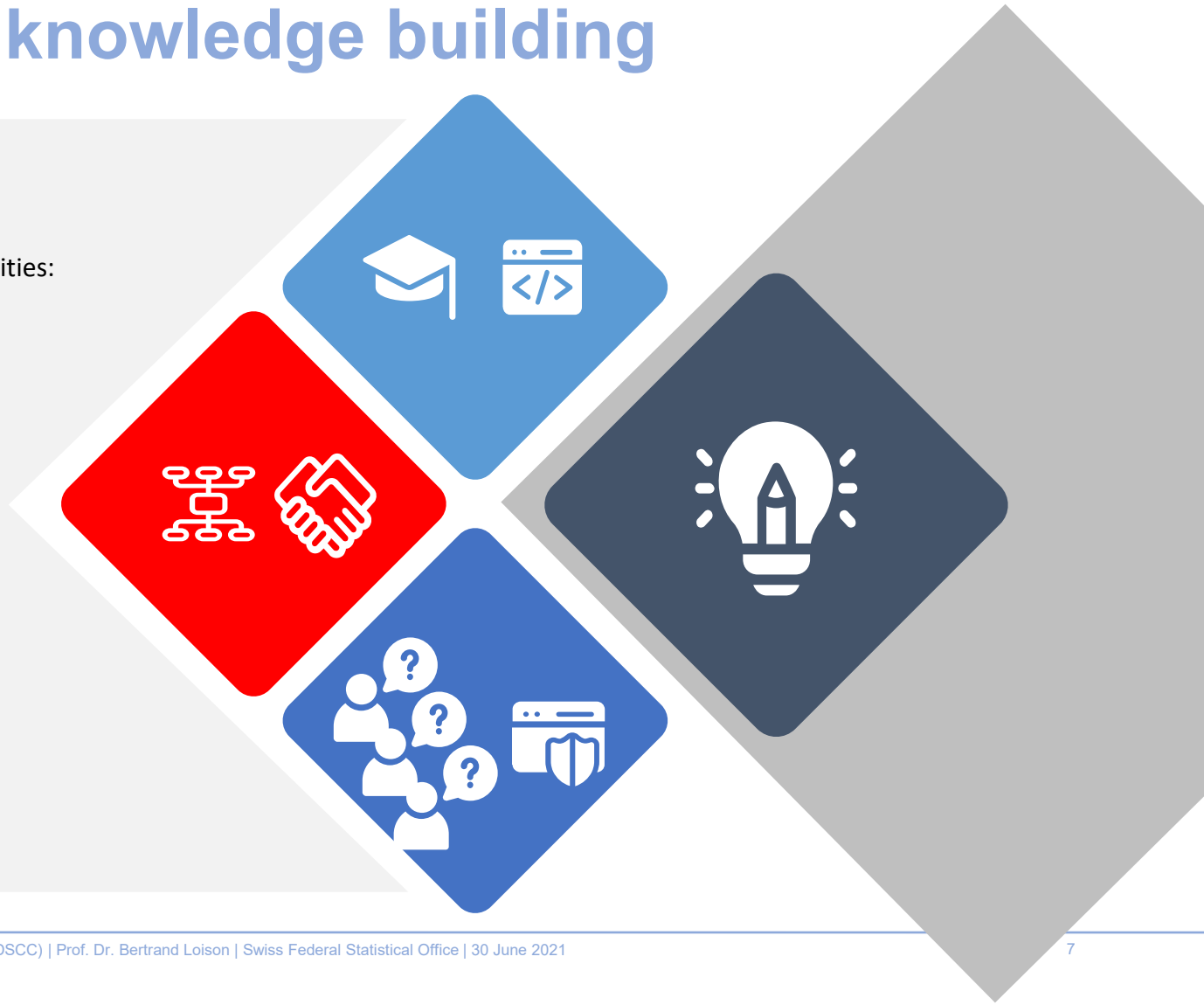
1. Data science expertise
2. Platforms as a Service for collaborative data science

DSCC at the FSO

1. Data Science as a Service for the entire Federal Administration
2. Building connections with similar international initiatives

Federal Offices

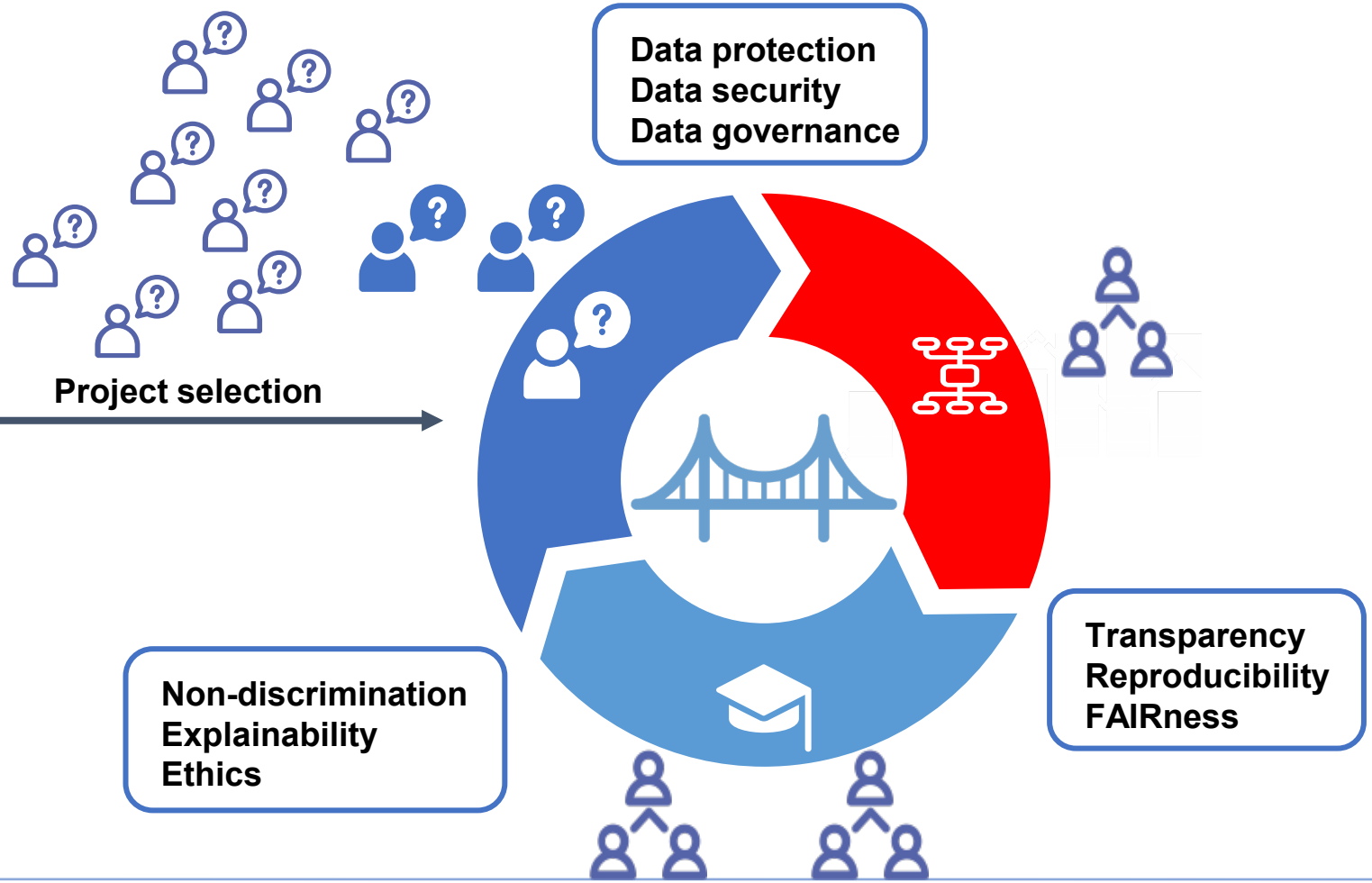
1. Data science questions and requests for collaborations from Federal Offices, Cantons, Communes, ...
2. Compliant IT platforms to deliver Data Science as a Service



Governance and trusted insights: Data Science Bridge

Data sources

- Text
- Images
- IoT, sensors
- Health
- Social media
- ...



Agenda

1. **Organizational perspective**
2. **Analytic services perspective**
3. **Data perspective**

What are the main motivations for using data science & AI?

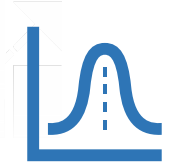
Generation of new insights

- Fraud detection
- Development of new scenarios
- Pattern recognition
- Contextualisation and quality assurance of decision bases
- Etc.

Generation of productivity gains

- Supervision
- Development of new data sources (data matching)
- Data preparation and data plausibility checks
- Support of application-oriented research projects
- Chatbot
- Etc.

Data science and AI are based on statistical models.



Code of practice of official statisticians enables e.g. explainable data science & AI.

A center of excellence in DSaaS for the Federal Administration

Data Science as a Service (DSaaS)

1 **Consulting**
Innovative methods in data science and AI

2 **Methodological support**
On the job training (coaching)

3 **Project execution**
Complete project realisation

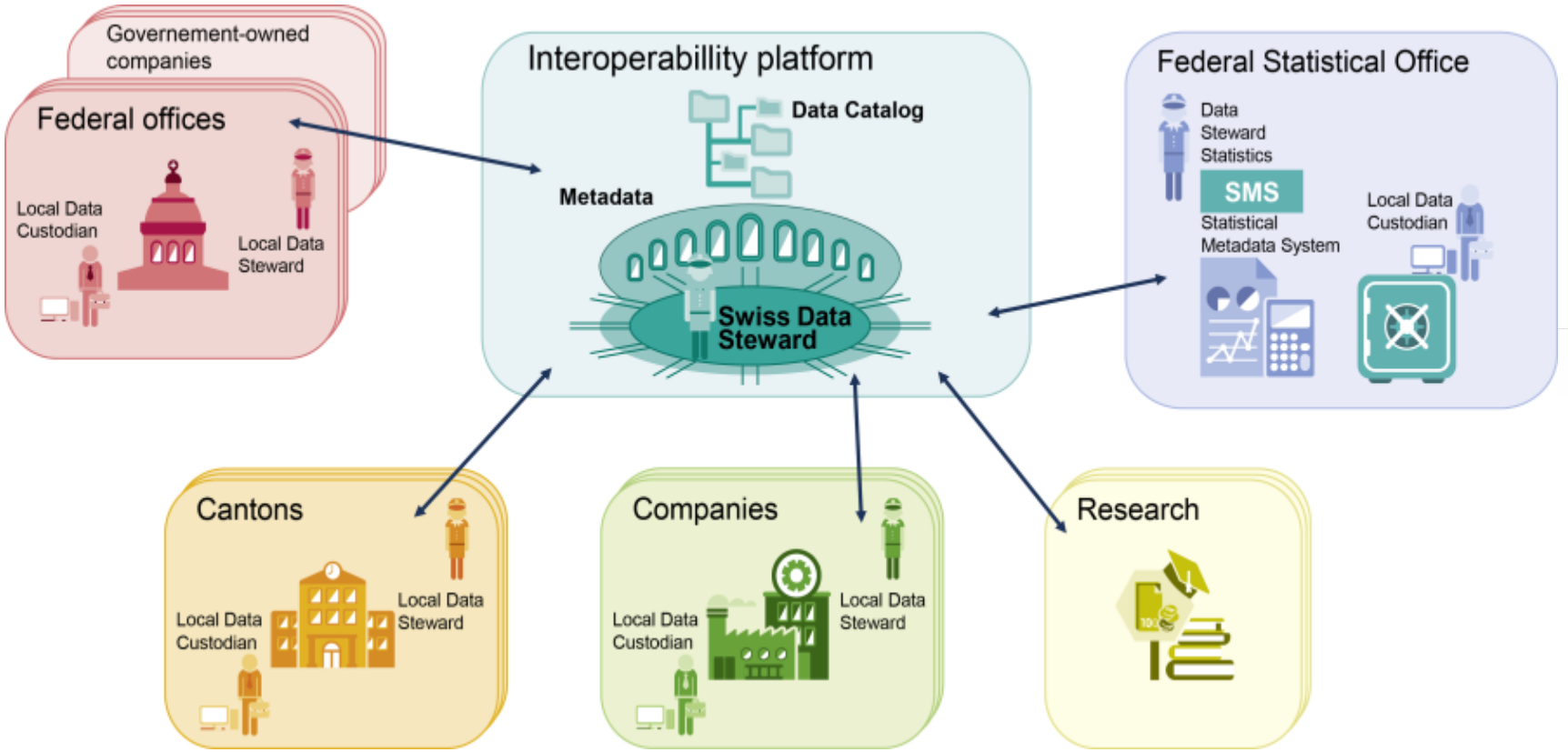
4 **Training**
Off the job training

0 **Basic services**
Quality standards, guidelines on data protection and basic infrastructures (“sandboxes”)

Agenda

1. **Organizational perspective**
2. **Analytic services perspective**
3. **Data perspective**

National interoperability



Solved problems

- To promote the reuse of data in the long-term, a **data catalogue** is used.
- An **interoperability platform** provides a system that is available to all participants.
- The **joint storage of meta-data with decentralised data storage** is considered a procedure to establish the reuse of the data.
- Access to the data depends of the every single law. Data privacy techniques should be used.

International interoperability: the civil servant perspective

Legal interoperability

- Civil servants and not only official statisticians at a country level need to have «access» to data of public intests (energy, tourism, etc.) **stored outside of their law jurisdiction**.
- Data protection is crucial because the purpose for having access to these data is not the same for every Federal agency. It depends of every single law's purposes (statistical vs administrative purposes).

Organisational interoperability

- Civil servants of all federal agencies need to agree on a set of needed variables (master data set) to negociate/cooperate efficiently with the companies - «Once only» principle!

Semantic interoperability

- Non identifiable secondary data sources are often not harmonised. The need for harmonised metadata and classifications of master data set is increasing (differential privacy needs it).

Technical interoperability

- Civil servants need to embrace new technical possibilities for working with data without having to «visualise» them (differential privacy techniques, …)

Unsolved problems

- Data stored on **platforms** (AirBnb, Uber, etc.) outside of the Swiss law jurisdiction.
- Ongoing needs to access to this platforms (sharing economy).
- Fully automated data pipelines as final goal (RAP).
- Every single federal agency hat its **own** data needs.
- Stronger coordination at national and international levels is needed.