



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



# Data Science Competence Center (DSCC)

**A service centre for the entire federal administration**



**Prof. Dr. Bertrand Loison**

Vice-Director at the Federal Statistical Office

UN Global Working Group on Big Data for Official Statistics

New York and Bern (online), 15 October 2020



# Agenda

- 1. Historical context**
2. Federal Council mandate
3. Why a Data Science Competence Center at the FSO?
4. Data Science Competence Center (DSCC)
5. Next steps

# Timeline: Data Innovation at Federal Statistical Office (FSO)

## Phase 1:

Creation of an internal working group at the FSO dedicated to the **Big Data** theme

## Phase 2:

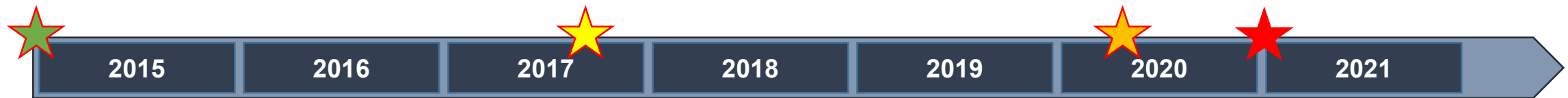
**Data Innovation Strategy v. 1.0**

Launch of **5 pilot projects**

Ad-hoc **Data Science lab**

## Phase 3:

Official mandate from the **Federal Council** to create a **Data Science Competence Center (DSCC)**



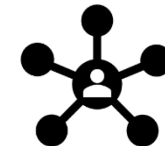
Phase 1:

**Exploration**



Phase 2:

**Concept and Proof of Concept**



Phase 3:

**Institutionalization**

# Agenda

1. Historical context
- 2. Federal Council mandate**
3. Why a Data Science Competence Center at the FSO?
4. Data Science Competence Center (DSCC)
5. Next steps



# Federal Council mandate

**FCD of 13.05.2020 [Item 1]:** According to the legislature plan 2019-2023, the FDHA (FSO) will create from 1 January 2021, a **centre of excellence for data science** that will **provide the federal administration with services in the field of data science and promote knowledge building in the federal administration and communication with the cantonal and communal authorities.**

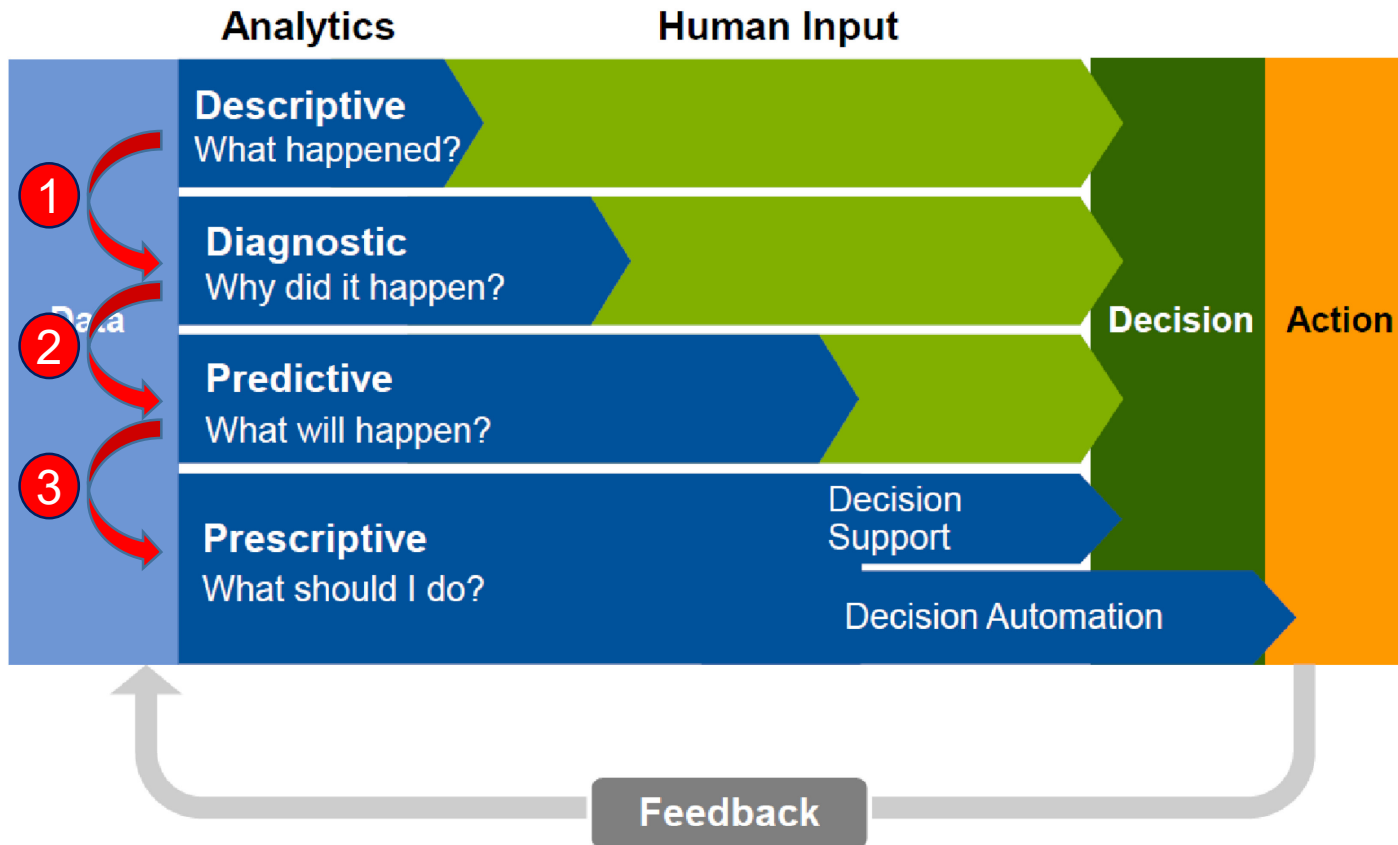
**FCD of 13.05.2020 [Item 2]:** In cooperation with the **other Departments and existing committees**, the FDHA (FSO) will prepare recommendations for the **governance** of the Data Science Competence Center (DSCC), which it will submit to the Conference of General Secretaries by the end of June 2021.

**FCD of 13.05.2020 [Item 3]:** Based on experiences made by the FSO and the other federal offices of the **first years of operation**, an assessment will be submitted to the Federal Council at the end of March 2024, that also details the **cost-benefit ratio**.

# Agenda

1. Historical context
2. Federal Council mandate
3. **Why a Data Science Competence Center at the FSO?**
4. Data Science Competence Center (DSCC)
5. Next steps

... because analytics (i.e. **learning from data**) is the FSO's core business



«16. **Big data breaks with the traditional method to search for causality.**

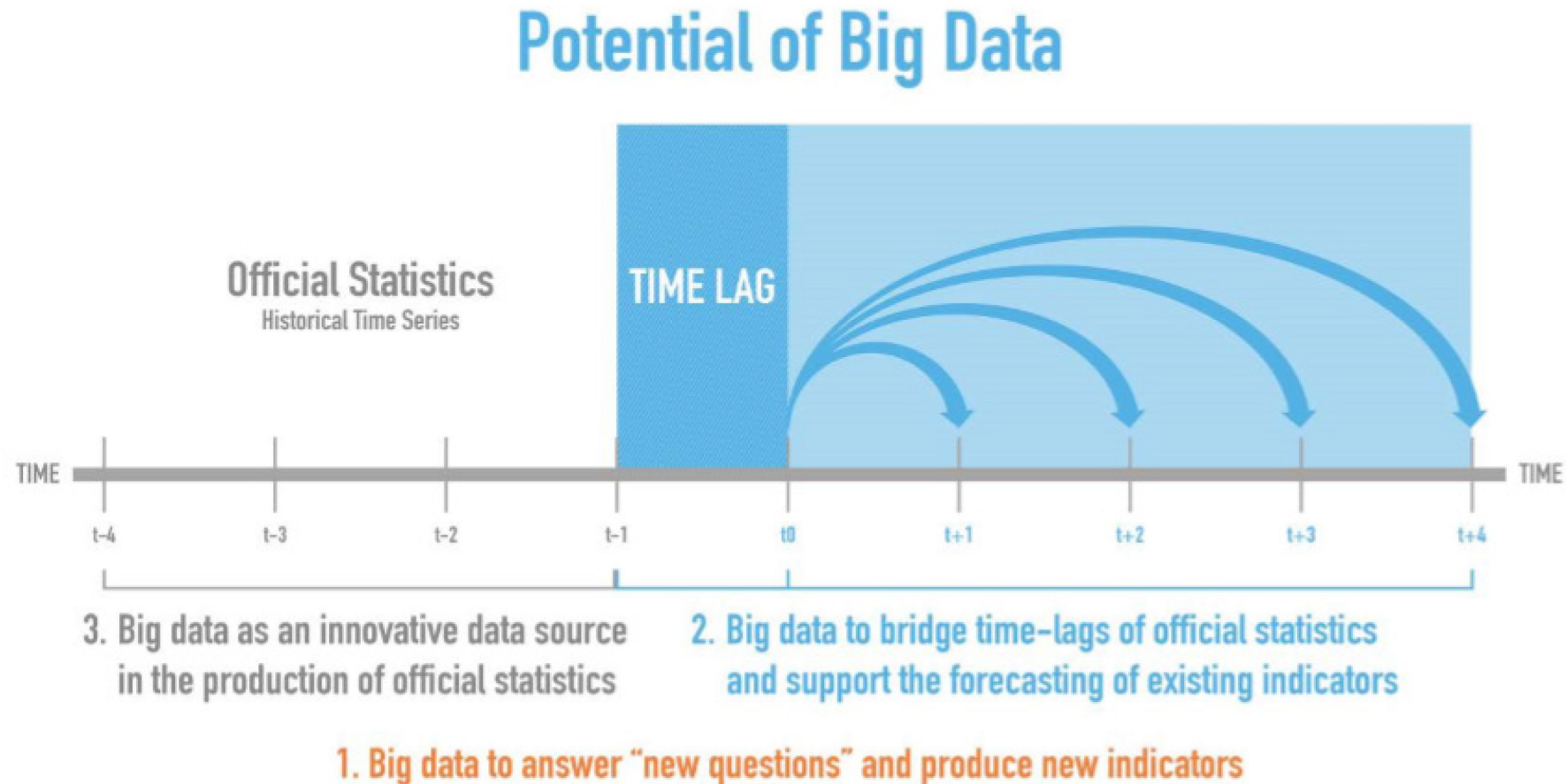
*Working with big data implies seeking patterns and correlations that may not tell us why something is happening, but rather alert us that it is happening [...].*

*In this vein, new indicators can be developed to obtain real-time correlations and to establish a more comprehensive early-warning system (Kitchin 2015) that can monitor the buildup of country specific as well as systemic risks in the real, external, fiscal, and financial sectors. »*

Source: Big Data: Potential, Challenges, and Statistical Implication, IMF, (2017). <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2017/09/13/Big-Data-Potential-Challenges-and-Statistical-Implications-45106>, pp. 11 - 12.

Source: João Tapadinhas, VP Business Analytics and Data Science, Gartner, June 2014 (goo.gl/YmjFPB)

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



Source: Big Data: Potential, Challenges, and Statistical Implication, IMF, (2017). <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2017/09/13/Big-Data-Potential-Challenges-and-Statistical-Implications-45106>, pp. 11.

# ... because official statisticians are dealing **at international level** with Big Data as well as Data Science for years

United Nations » Department of Economic and Social Affairs » Statistics Division

**BigData** UN Global Working Group

HOME TASK TEAMS MEETINGS PROJECTS



## UN Big Data For Official Statistics

Exploring the use of Big Data for official statistics to meet the expectation of society for enhanced products and improved and more efficient ways of working

**NEWS**

International Symposium on the use of Big Data for official statistics  
16 - 18 October 2019, Hangzhou, China // [EVENT PAGE](#)

The symposium on the use of Big Data for official statistics was opened by Mr. Xian Zude, the Deputy Commissioner of the National Bureau of Statistics (NBS) of China, and Mr. Stefan Schweinfest, the Director of the United Nations Statistics Division. The symposium was attended by about 25 international

**DOCUMENTS**

- UN GWG Terms of Reference
- Kigali Declaration
- Bogota Declaration

## Task Teams

1. [Access and Partnerships](#)
2. [Big Data and the Sustainable Development Goals](#)
3. [Mobile Phone Data](#)
4. [Satellite Imagery and Geo-Spatial Data](#)
5. [Scanner Data](#)
6. [Social Media Data](#)
7. [Training, Skills and Capacity-building](#)
8. [Committee on Global Platform for Data, Services and Applications](#)





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

## FSO's Experimental statistics

Experimental statistics are produced using new methods and/or new data sources and are therefore in line with the **FSO's data innovation strategy** and the Confederation's multi-annual programme for federal statistics. This site contains descriptions of the (pilot) projects currently being developed.

By publishing them we can involve users and partners at an early stage for both the development and consolidation of projects.

The aim of these statistical projects is to better meet users' needs in terms of efficiency, quality and speed. However, these statistics still have the potential to evolve, especially regarding their methodology, which is still being assessed. For this reason they are clearly marked as experimental and carry a logo that can easily be recognised.



## Published statistics



### Small area estimation (communes) of economic activity rate in the structural survey

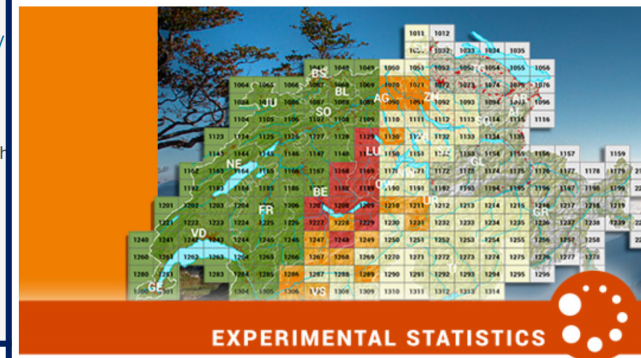
The structural population survey provides important information on the population, including information about work. The whole purpose of Small Area Estimation is to push the boundaries imposed by standard methods.

The study showed that it is possible to obtain reliable estimates for both annual economic activity rates for communes that had a sample of at least 100 people.

## Pilot projects within the data innovation strategy

On 21 November 2017, the FSO published its **data innovation strategy** [↗](#).

This document is the FSO's first response to the wider subject of digitalisation. More specifically, it focuses on the application of complementary analysis methods (e.g. predictive analysis using advanced statistical techniques, data science and machine learning) that enable the current production of official statistics to be increased or completed. Five pilot projects have been chosen to implement this strategy and are in progress. Each project is described below.



### Project "Area Statistics Deep Learning" (ADELE)

The FSO's land use statistics are an invaluable tool for long-term land observation. This project involves learning and mastering the use of artificial intelligence (AI) technologies to eventually automate (even partially) the visual interpretation of aerial images in order to detect and classify changes.



### Project "Automation of NOGA coding" (NOGAuto)

Automation of the coding of the economic activity of enterprises using Machine Learning methods applied to data already available within the FSO (data from surveys, descriptions in the commercial register, keywords, explanatory notes for classifications etc.) to support coding.

Source: Federal Statistical Office, (2019).  
<https://www.bfs.admin.ch/bfs/en/home/services/recherche/experimental-statistics.html>.

# ... because the FSO has to respect the **Code of Practice**

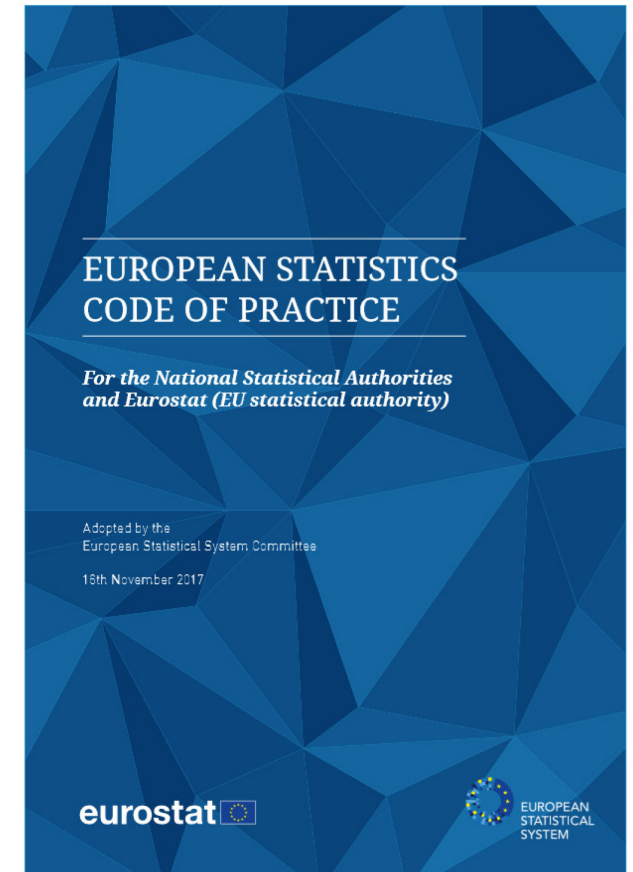
## European statistics: Code of Practice

1. Professional independence
2. Mandate to collect and access data
3. Appropriate funds and resources
4. Quality commitment
5. Confidentiality of statistics and data protection
6. Impartiality and objectivity
7. Robust methodology
8. Suitable statistical procedures
9. Reasonable burden on respondents
10. Cost efficiency
11. Relevance
12. Accuracy and reliability
13. Topicality and timeliness
14. Coherence and comparability
15. Accessibility and clarity

Environment

Process

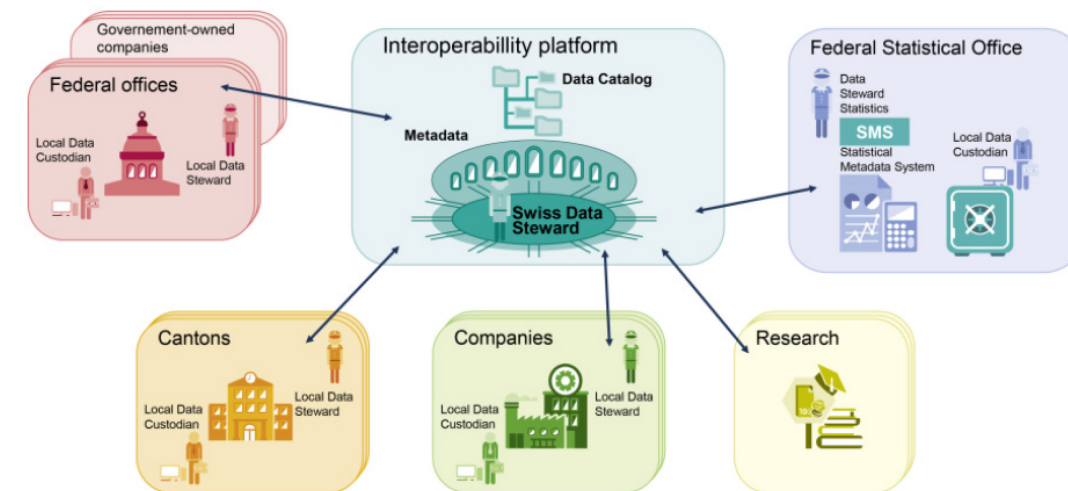
Output



# ... because the FSO is in charge to build and manage the **interoperability platform** and the **national data management program (NaDB)**

- The Federal Council expects to make data management in the public sector easier and more efficient by **reusing data**: Persons and businesses will only need to report certain information once.
- In order to promote the reuse of data in the long-term, a data catalogue is used in which the type of data and the quality of these data can be seen. An **interoperability platform** is developed as a system that is available to all participating offices.
- The coordination of the standardisation and harmonisation tasks falls within the role of the **Swiss Data Steward**. The joint storage of metadata with decentralised data storage (local data custodian) is considered a future-looking procedure to establish the reuse of the Administration's data.
- FSO, as **Swiss Data Steward**, is in charge to build and manage the interoperability platform.

Interoperability platform





# Agenda

1. Historical context
2. Federal Council mandate
3. Why a Data Science Competence Center at the FSO?
- 4. Data Science Competence Center (DSCC)**
5. Next steps

# Vision and mission of the Data Science Competence Center (DSCC)

## Vision

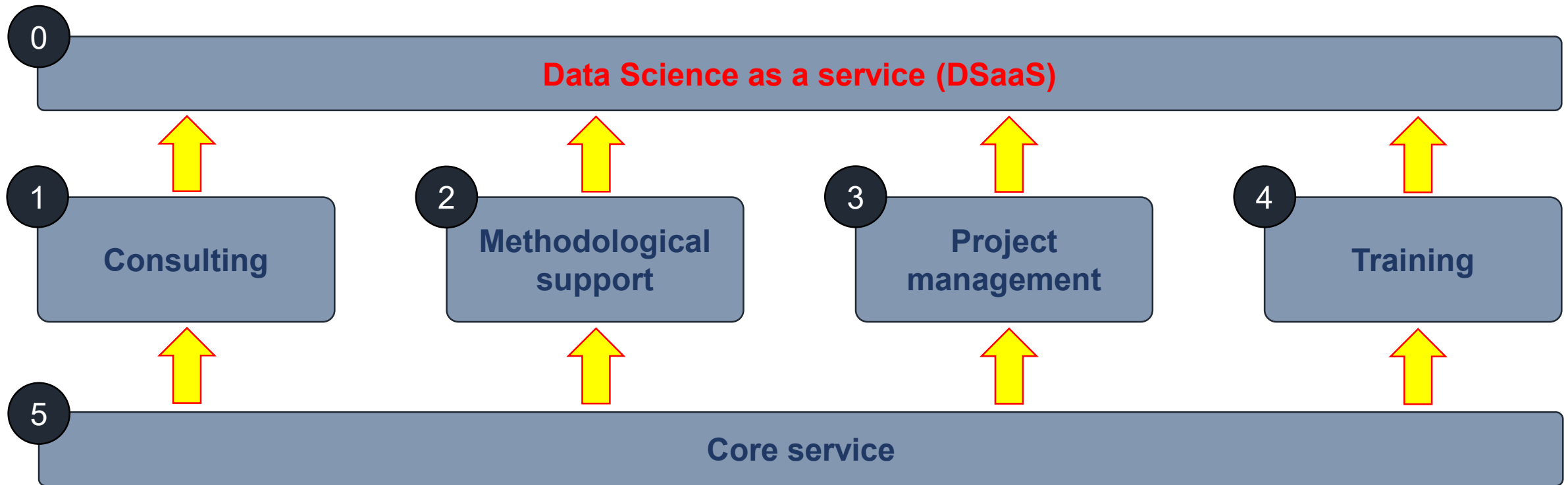
“We use data science and develop skills for the **public good** in Switzerland.”

## Mission

“We work **at the crossroads between data science and artificial intelligence** – we build skills and apply the appropriate methods, techniques and practices to obtain a new understanding and to improve decision-making for the public good.”

# Service portfolio of the Data Science Competence Center (DSCC) - #1

As a service provider, the centre of excellence will provide services within the administration in the field of data science - “Data science as a service (DSaaS)”.



# Service portfolio of the Data Science Competence Center (DSCC) - #2

0

## Data Science as a service

As a service provider, the centre of excellence will provide services within the administration in the field of data science - “Data science as a service (DSaaS)”. There are five possible services in the field of data science.

1

## Consulting

Consulting related to strategical, tactical and operational application of (for the federal administration) innovative data science methods and procedures (e.g. potential analysis of methods from extended statistics, machine learning and the field of artificial intelligence).

2

## Methodological support

Methodological support (coaching - ‘on the job training’) in executing relevant application-oriented, externally procured or internal projects by Federal Offices and/or transferring concrete results and findings from these projects into existing administrative processes (if required).

# Service portfolio of the Data Science Competence Center (DSCC) - #3

3

## Project management

Complete execution of relevant data science requests.

4

## Training

Application-oriented training ('off the job training') on data science methods, techniques and practices and the necessary (IT) technologies and tools.

5

## Basic service

The centre of excellence is also tasked with developing quality standards, guidelines on respecting data protection and basic infrastructures ('sandboxes') for data science applications.

# What are the core values of the Data Science Competence Center (DSCC)?

## Core values

Great importance is attached to the core values of information security, data protection, data security, data governance, non-discrimination, explainability, transparency, reproducibility and public confidence in the implementation of all data science services.

For example, the results of the individual projects (as far as permitted by data protection laws and other provisions) are documented in a transparent manner and made available.

# Cooperation as a cornerstone of the Data Science Competence Center (DSCC)

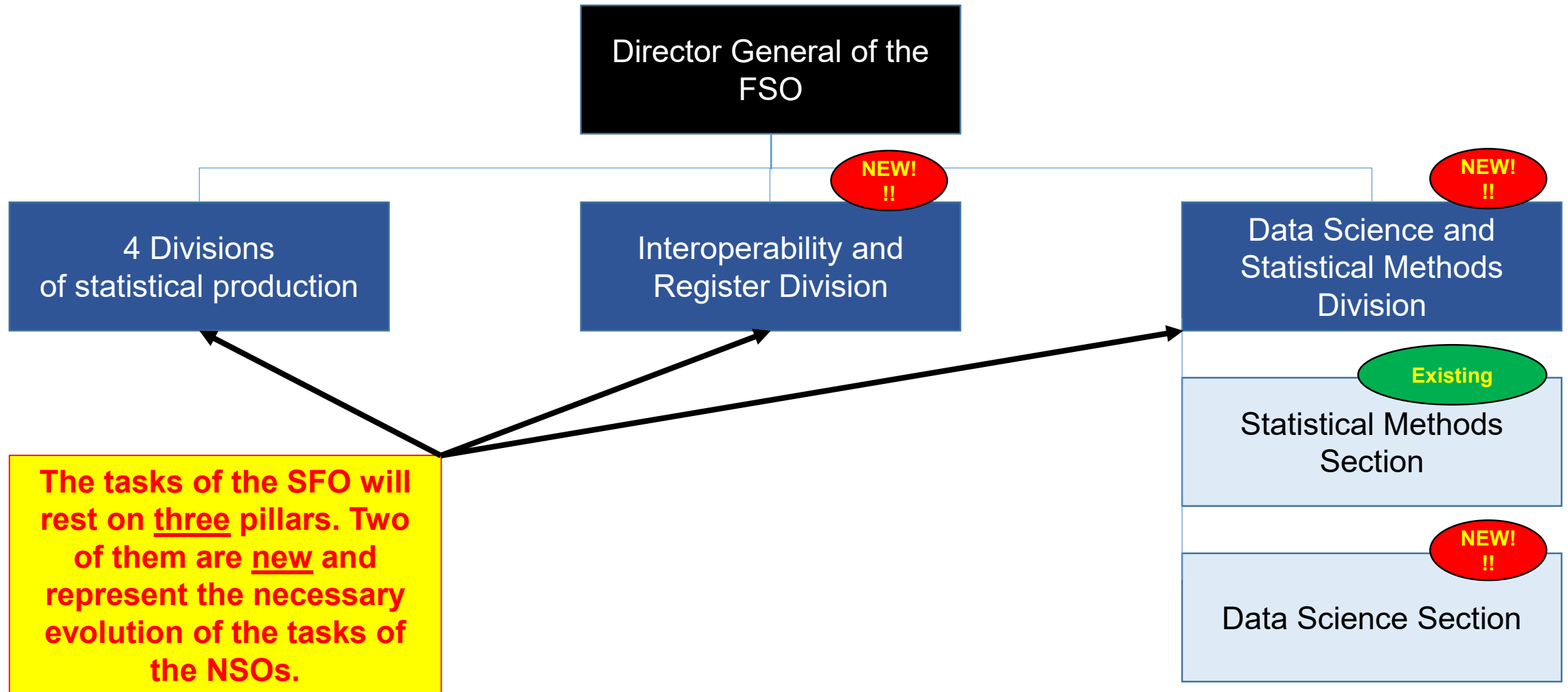
- The centre of excellence for data science **operates across disciplines** and encourages the exchange of experiences from previous projects (knowledge transfer).
- The centre of excellence **ensures the productive and sustainable connection between science** (e.g. Swiss Data Science Center, ETHZ, EPLC, universities, etc.), **research** (e.g. application-oriented research by higher education institutions) **and the practical application in the federal administration** (in particular Offices with complementary objectives, e.g. Swisstopo and MeteoSwiss).
- In this way, the need for new capabilities and skills such as in relation to data science methods and techniques, technologies and tools **can be assessed**, further developed where necessary and the findings **made available to the federal administration**. Key here will be **the exchange of knowledge between units** (that may not yet exist) in the Offices, that wish to adopt on a long-term basis a data science approach within their sector.
- Regarding the **further development of IT technologies**, close coordination with the **FOITT and other IT service providers** (ISCeco, ISC FDJP, AFC SO DDPS, FDFA IT, National Cyber Security Centre NCSC) is planned, as services in the field of data science are linked to ICT services.

# Agenda

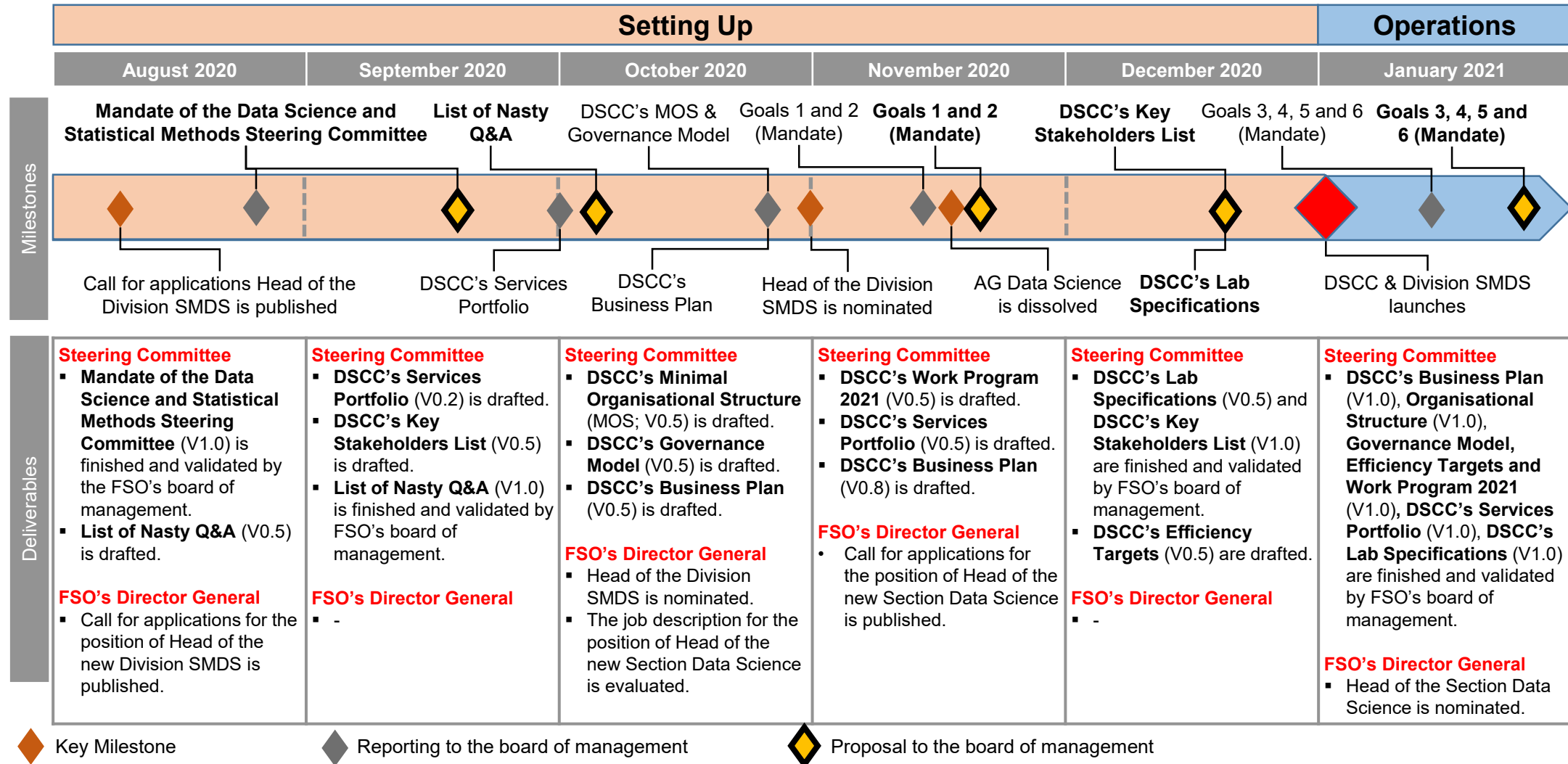
1. Historical context
2. Federal Council mandate
3. Why a Data Science Competence Center at the FSO?
4. Data Science Competence Center (DSCC)
- 5. Next steps**



# Organisation of the FSO from 01.01.2021



# Roadmap defined by the FSO for setting up the DSCC



# Roadmap defined by the Federal Council

- 01.01.2021:** Launch of the centre of excellence for data science.
- from 01.01.2021:** Need-based and gradual set-up:
- Phase I:** for the FSO,
  - Phase II:** for the federal organisational units that are members of Fedestat (producers of official statistics),
  - Phase III:** for the entire federal administration (for administrative and not statistical purposes).
- To ensure the exchange of experiences, as of Phase II, discussions will be held at cantonal and communal level, primarily with Regiostat members.
- 30.06.2021:** Recommendations for the management of the centre of excellence are approved by the Conference of General Secretary.
- 31.03.2024:** Assessment report (including cost-benefit ratio) from the first years of operation is approved by the Federal Council.

# Questions & Answers

