

February 3rd, 2022

DGFiP data strategy and how Data Management helps building it

Su YANG

DGFiP

1

Data governance, data management

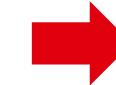
Difficulties and needs

Absence of governance framework



- Absence of **an agile and horizontal organization** allowing us to deal with urgent demands.
- Roles and responsibilities not defined and difficulty to find the right entry point for a specific data.
- Business units and IT department organised in silos.

Knowledge



- **Absence of a common language.**
- Difficulty to identify golden sources and the right business experts.

Quality and master data



- **Lack of visibility on the level of quality and certainty.**
- Remediation process not adequate and time-consuming.
- Applications have different level of support to shared master data..

Data opening



- **Increasing number of access demands to fiscal data**
- Data exchange are subject to different processes
- How to valorise Open Data

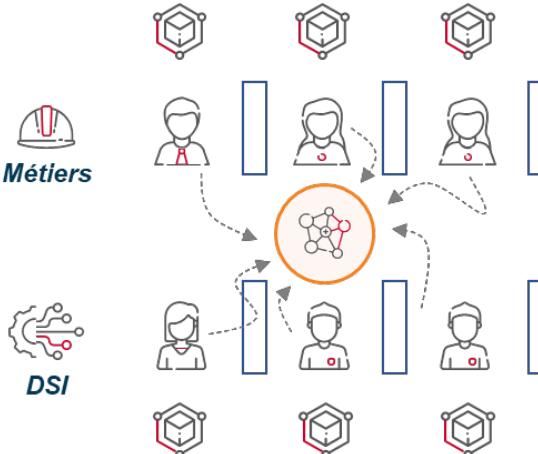
Tools and culture



- **Change management related to new tools: data catalogue and data visualization for instance**
- Business units are not autonomous in using those tools.

Data governance

Data Governance represents the modalities of decision making and responsibilities definition revolving around a governance framework, the whole intended to ensure the protection and fair usage of data such as in creation, management, control and data valorisation
(adapted from Gartner Glossary definition)



The objective of data governance is to « **break the wall** » between different business units and open up the ability to communicate altogether in order to create new usages of data and data sharing

To
govern/manage
data

Data science
Data vizualisation...



Difference between data governance and data management

Data management designates the practices necessary to build and to operate the framework for data collection, storage, usage, exploration and archiving. Data management adopts the policies and process defined within the governance framework thus allowing data valorisation.

What is the link with the governance of the Data Information System?

Data management therefore **represents** the operational side, **in particular related to information technology**, of data governance.

DATA GOVERNANCE VERSUS DATA MANAGEMENT

DATA GOVERNANCE

Decision making, monitoring and enforcement body that has authority over an organization's data management

Data governance is the strategy of data

Involves establishing processes, practices and theories

Focuses on the safety, reliability and trustworthiness of data

DATA MANAGEMENT

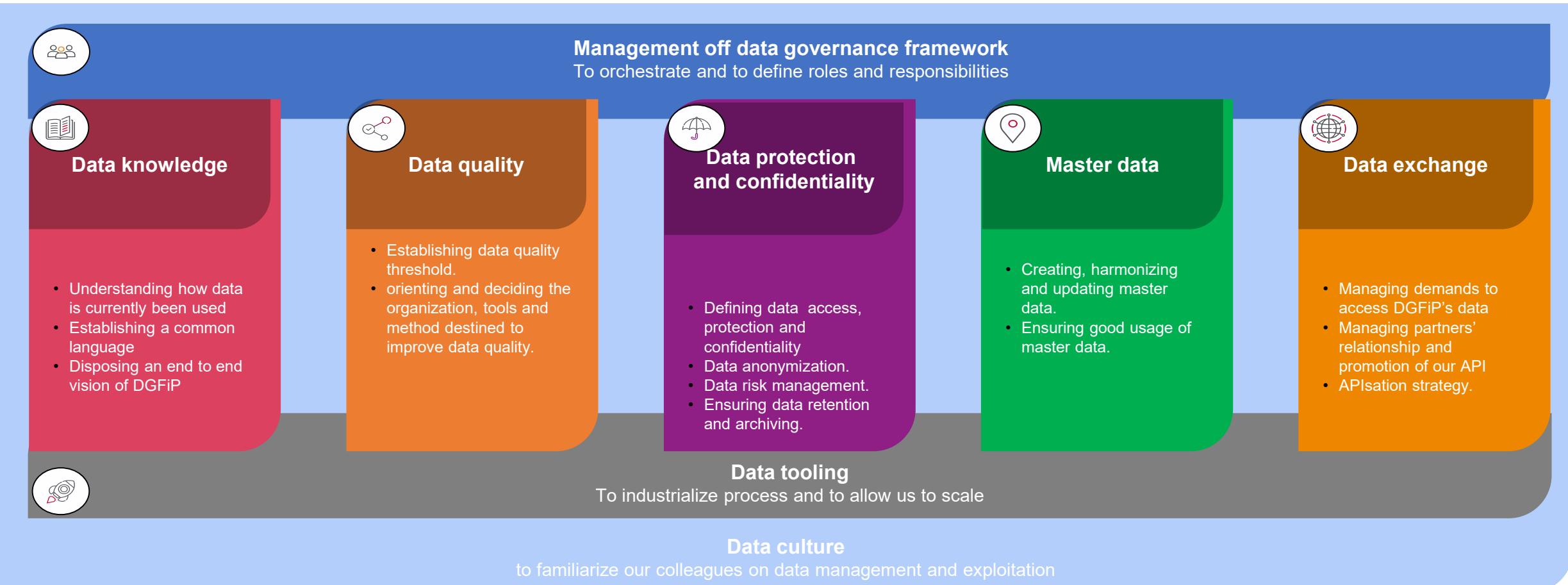
Process of controlling data architecture, quality, security, policy, practices and procedures

Data management is the logistics of data

Involves collecting, validating, storing, organizing, protecting, processing and maintaining data

Focuses on making data valuable

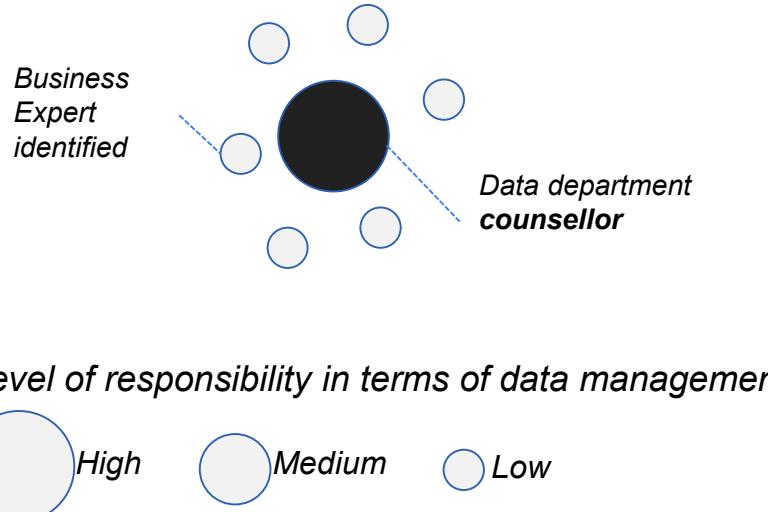
Missions within our data governance framework



Some organization scenarios

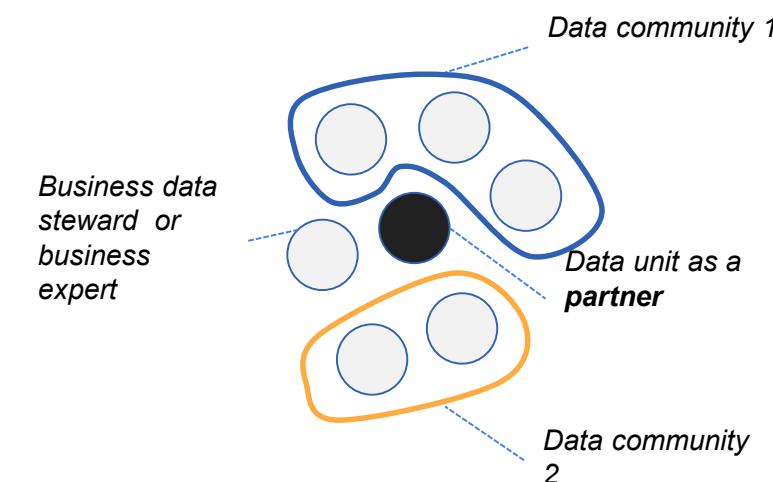
Scenario 1 – Council organization

A central unit that supports all the business units



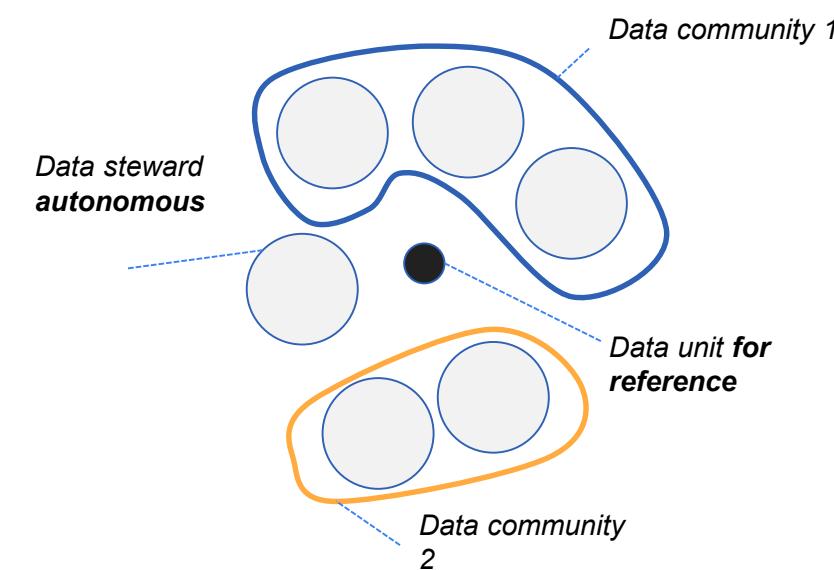
Scenario 2 – Partnership organization

A central unit serving as partner for data projects of business units



Scenario 3 – Reference organization

A central unit serving as a simple point of reference



Central unit for data governance

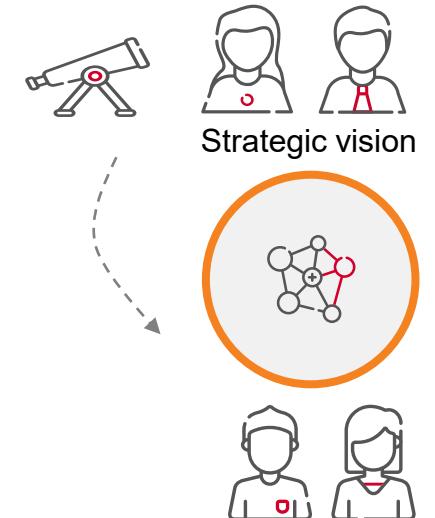
Business units

2

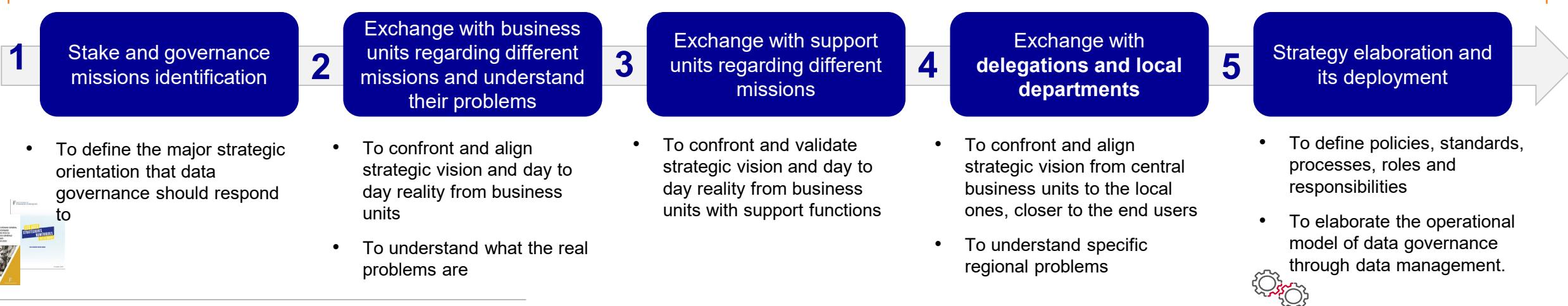
Our modus operandi

A 5 step process

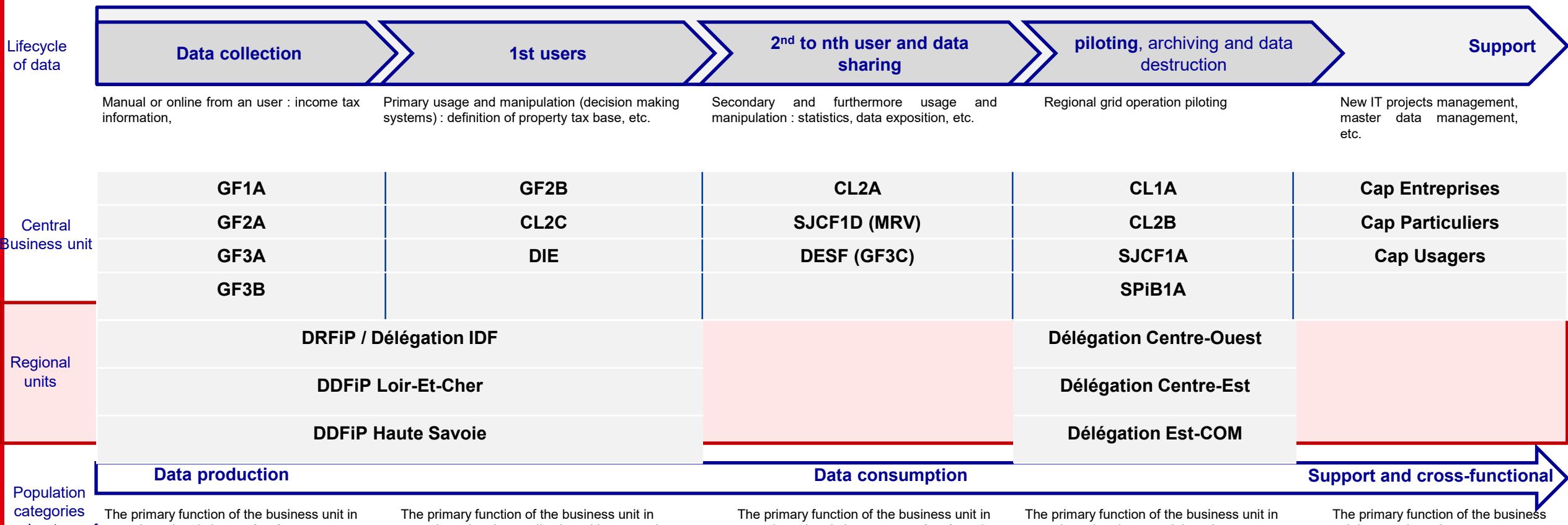
« A data governance model co-constructed with our different business units. »



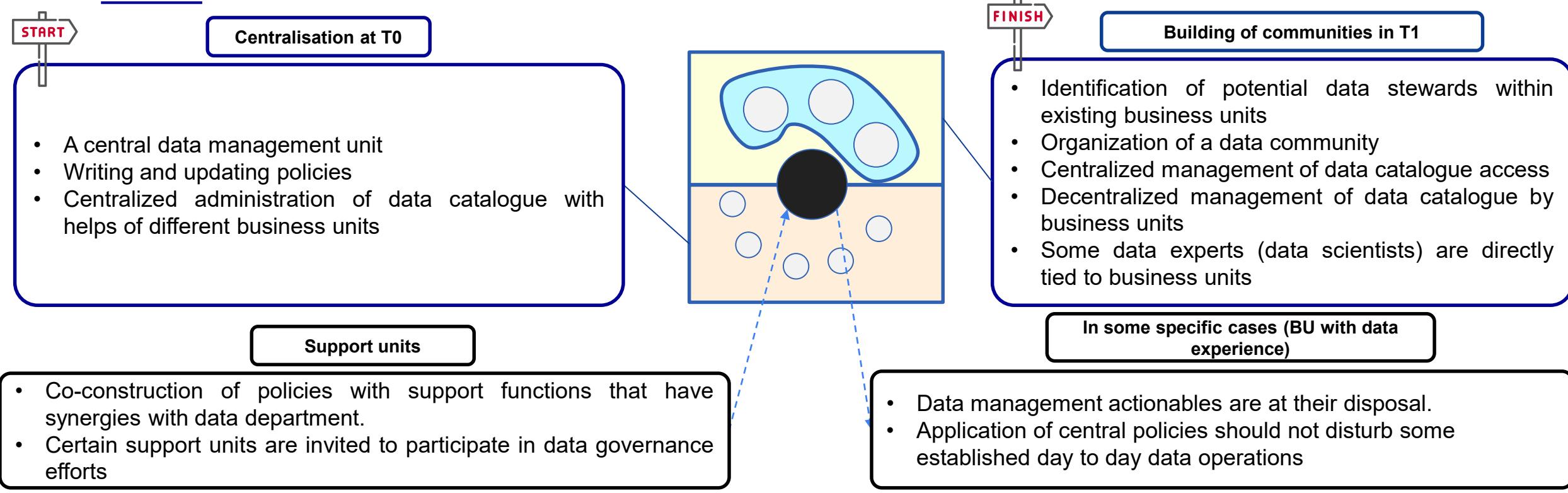
Business units and IT
(day to day reality: pain points, needs, etc.)



An approach to engage all actors



The chosen scenario



Points of attention in data governance deployment

**Data governance
should be everyone's
matter**

Most of us produce or consume data, it's in everyday operations where master data is used, that the meaning of data is important and that quality of data can be ensured and can impact operations. Better defining roles and responsibilities and equipping everyone with the right tools is very important.

Change management is absolutely crucial and is a major part of the process.

**It is especially crucial
if one wants to scale
its data operations**

An estimate of 1 months to 3 months is observed with our 19 data visualisation and 5 data science projects of 2021 because of data experts' difficulties to find out what each variables means.
Roughly the same data is used in average in 3 projects and some master data in more than half of it.
Their transportation, prepping and pre-analysis are all matters of months that could have been saved.
In some cases, half of the data has to be rejected during joining effort because of the difficulties to matching data.

Data management allows us to reduce frictions and rationalize our data resources.