GREAT

WORK PACKAGE 3 – TECHNICAL BLUEPRINT

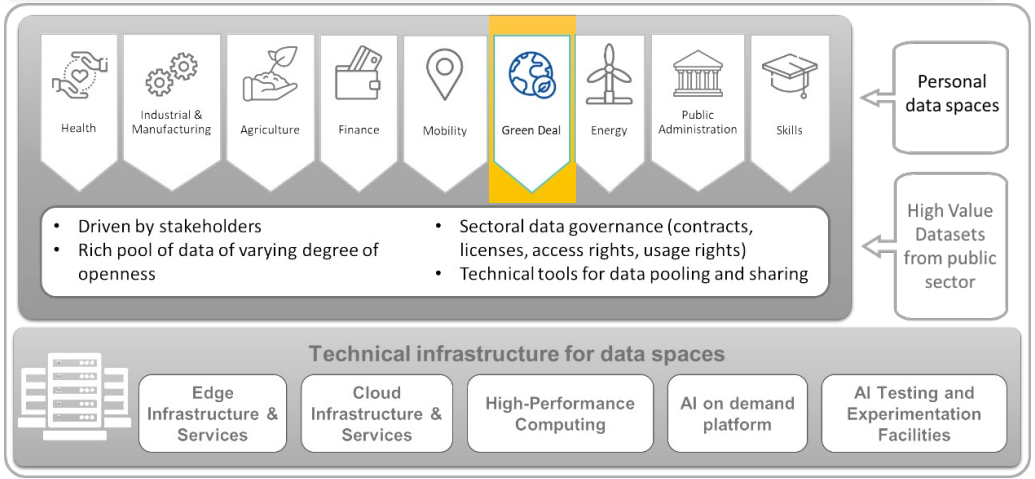
The Green Deal Data Space

GREAT

THE GREEN DEAL DATA SPACE

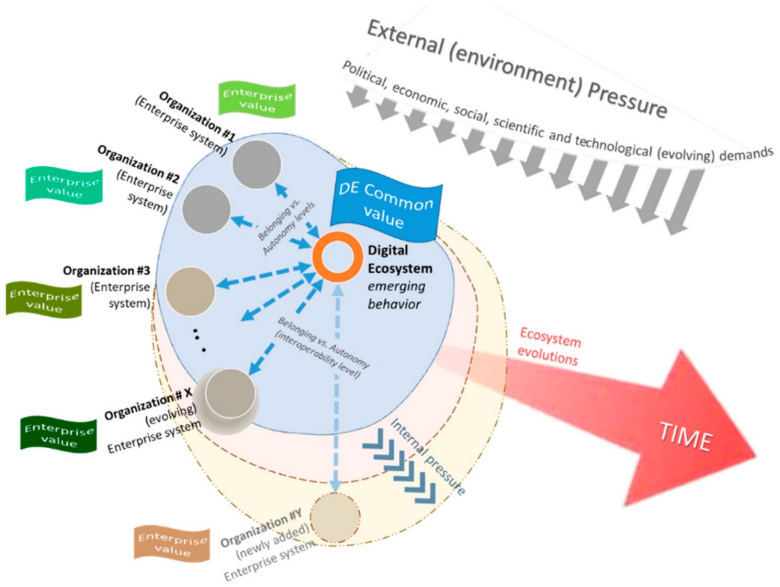
- will interconnect **currently fragmented and dispersed data from various ecosystems,** both for/from the private and public sectors.
- will offer an **interoperable, trusted IT environment,** for data processing
- will provide a **set of rules** of legislative, administrative and contractual nature that determine the rights of access to and processing of the data.

Define the **technical blueprint of the GDDS reference architecture** explaining how existing (and planned) data ecosystems (at national, regional, and local level) can be connected to provide an interoperable, secure data sharing environment which allows seamless discovery and use of available data
[GREAT Proposal]



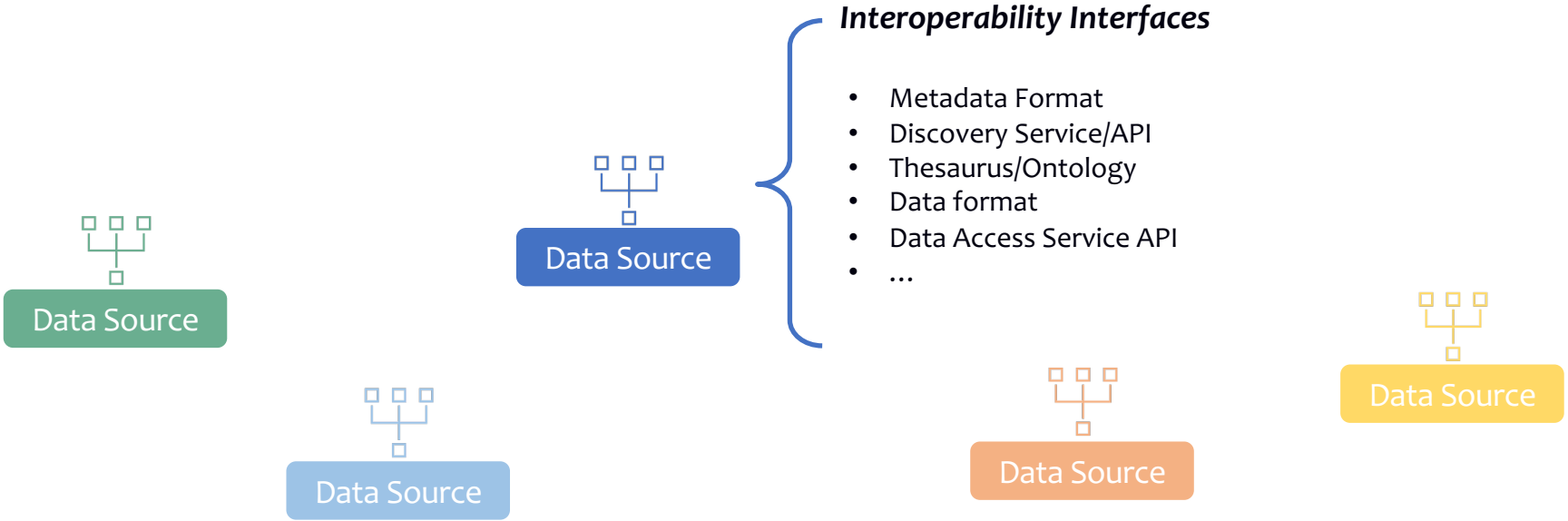
DIGITAL ECOSYSTEMS

- A Digital Ecosystem emulates Natural Ecosystems
  - Multiple ‘species’ (autonomous entities) collaborating and competing
  - In a (digital) ‘environment’
  - Carrying out different functions
  - Contributing to a ‘service’ for the human society
  - To be protected (governance)
- No fixed set of participants (‘species’)
- No fixed set of requirements, only one or more general ‘services’ (e.g., generating Earth Intelligence, secure sharing)
  - Ready to changes
- Participants can enrich the DE providing tools and services on top of the existing ones
  - Security and trust
  - Generation of knowledge for Earth Intelligence
  - …



Source: Nativi, S.; Mazzetti, P.; Craglia, M. Digital Ecosystems for Developing Digital Twins of the Earth: The Destination Earth Case. Remote Sens. 2021, 13, 2119. https://doi.org/10.3390/rs13112119

Building on existing (and Future) Capacities



Basic Principles

Inclusiveness	We can expect a high heterogeneity of data systems in terms of supported metadata content and formats, data encoding, coordinate reference systems, ontologies. At least part of this heterogeneity is justified by the specificity of the community that generates and uses those data. Since the driving benefit of a data space is to share *all* the valuable datasets, data systems cannot be excluded only due to their diversity.
Fairness	We can expect high heterogeneity also in terms of 'species' including big companies, SMEs, public administrations, research and academic organizations, intergovernmental institutions, citizens. A data space should be the common ground where collaboration and competition take place for the benefit of the 'species' but, overall, for the ecosystem to serve data for generating knowledge. Therefore, no privileged access should be granted to anyone at the risk of changing the fairness of the data space.
Autonomy	We expect that some data sources are already part of other SoS or ecosystems with their own mandate and governance – e.g., European Research Infrastructures, Copernicus Services, Space Agency ground segments, Public Administration systems including INSPIRE. It is necessary to respect such autonomy without imposing, de-iure or de-facto, the exclusive participation in the data space. This is strictly related to the autonomy vs. belonging conflict that will affect any data system. In a Common European Data Space, belonging should be encouraged through soft means mostly based on the overall value of the data space.

GDDS DE Soft Infrastructure

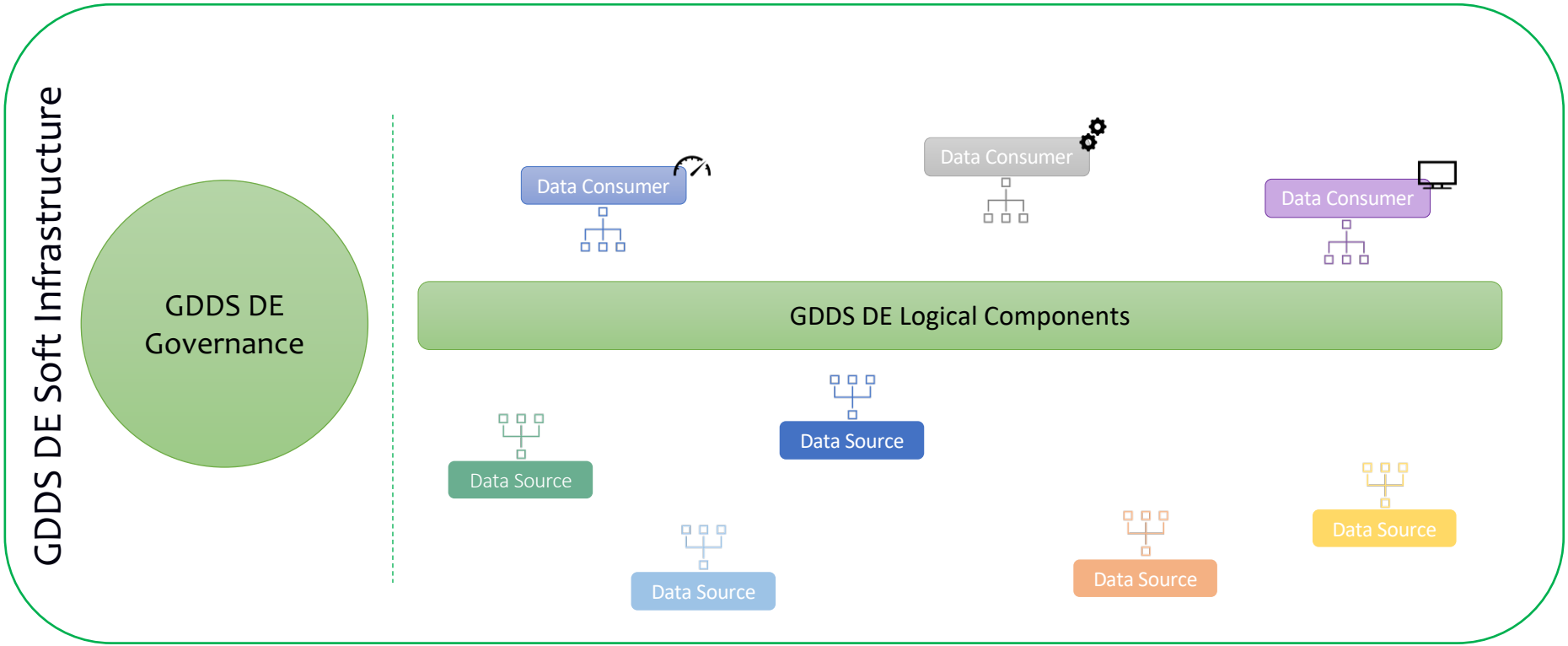
A soft infrastructure is invisible, made up of technology neutral agreements and standards, on how to participate in an ecosystem.

The GDDS is characterized by a high level of heterogeneity, with many already existing data sharing initiatives that offer their resources to diverse consumers, which mirrors the current state of (geospatial) data sharing globally.

Establishing a single "common format" is not possible in a multidisciplinary context like GDDS.

The challenge is how to transform a collection of disparate systems that use different technical standards into a digital ecosystem. This requires a minimal set of logical components that enable the ecosystem's digital environment.

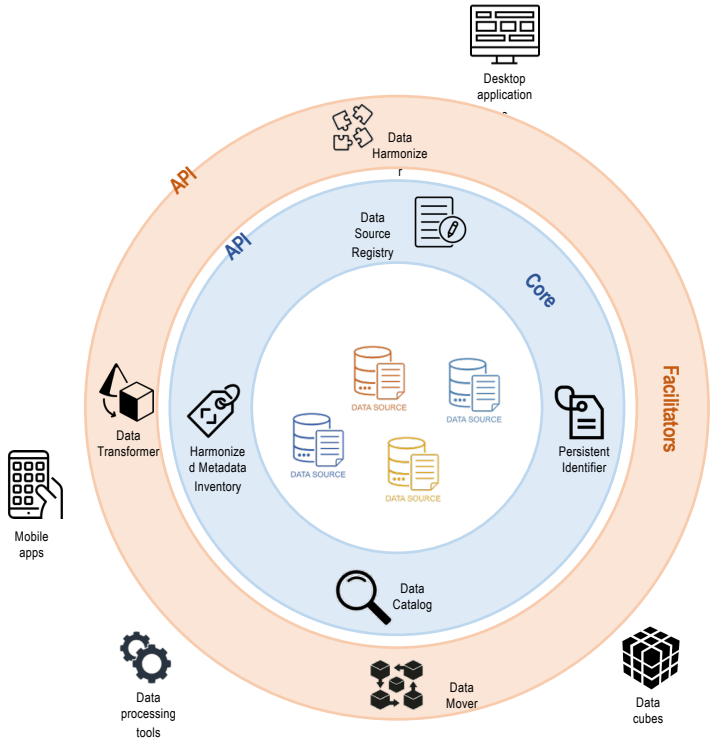
GDDS DE Soft Infrastructure



Logical Components

The **Core** components are expected not to evolve at a rapid pace, they constitute the foundation of the GDDS DE and are expected to be relatively stable in terms of basic functionalities.

Facilitators are designed to enable an as seamless as possible use of the GDDS content. These components are expected to evolve (both in number and in functionalities) more rapidly in response to both users' needs and the emergence of new technologies. In fact, the GDDS DE technical blueprint must be able to cope with a rapidly changing technological environment where we expect the emergence of new technologies, enabling now unpredictable scenarios.

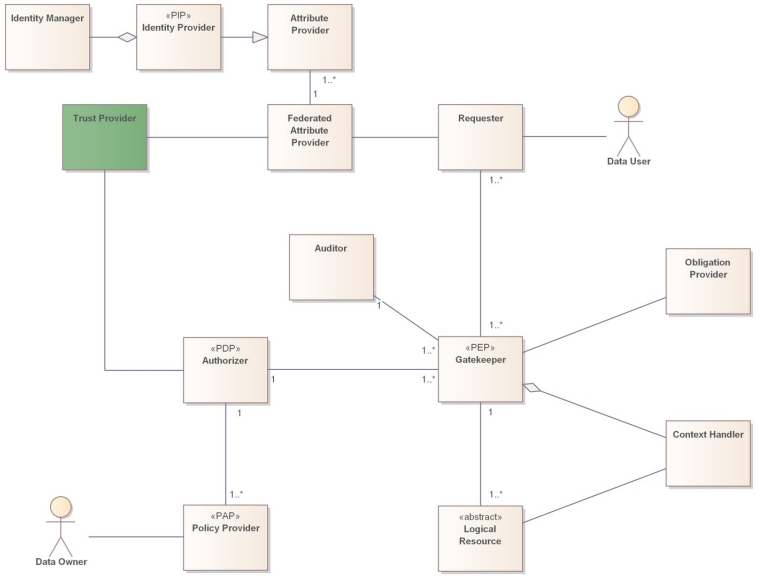


Security

Each provider of the GDDS must be able to define its own data policies and these must be supported at the GDDS level. This requires supporting a highly heterogenous set of data policies, resulting in a very complex and difficult to maintain access control framework.

Decoupling of Authentication and Authorization: the business logics for authentication and for authorization are separated. This is a good practice in general, but even more in a distributed system like GDDS.

Authorization Framework: the authorization (i.e., access control) framework is based on the Remote Access Control approach and compliant with XACML framework. This choice is driven by the recognition that such an approach has a minor impact on GDDS DE participants



Main Benefits

GREAT

Final Users

Mediation/Harmonization functionalities to:

- Facilitate access to data from multiple communities
- Enable data exploitation in new scenarios

Intermediate Users

Flexibility to support changes in technological context

Data Providers

Low entry barrier for existing Data Providers

Support Data Owners' data policies

Thank You