Openness and environmental data sharing: a JRC perspective
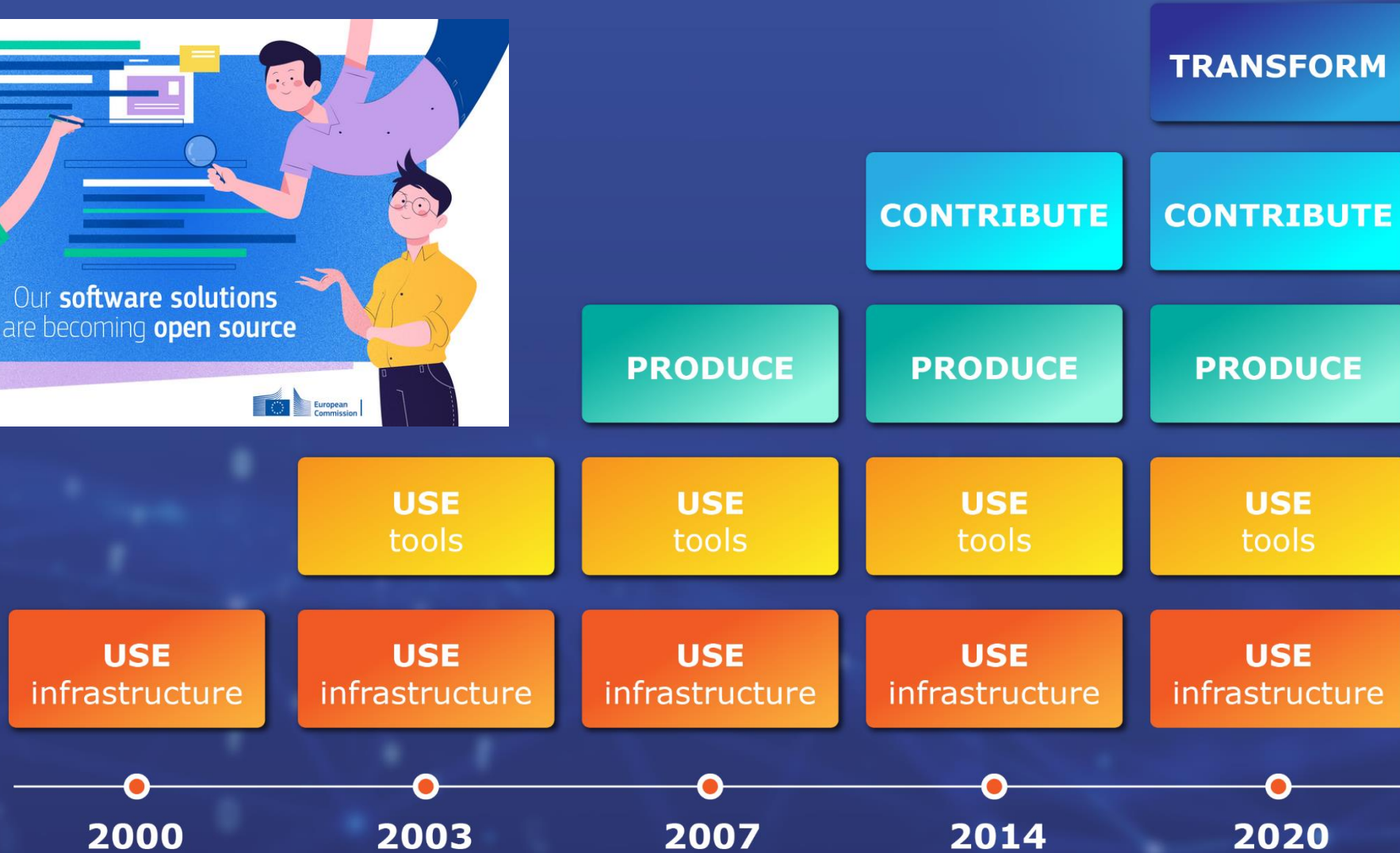
Marco Minghini

2022 Open Innovation for EO Programmes Workshop – November 2-4, 2022

Outline

- Open source at the European Commission
  - strategy & policy
  - tools & actions
- Experiences at JRC – Openness in data sharing
  - INSPIRE
  - JRC contribution to GEO and EuroGEO
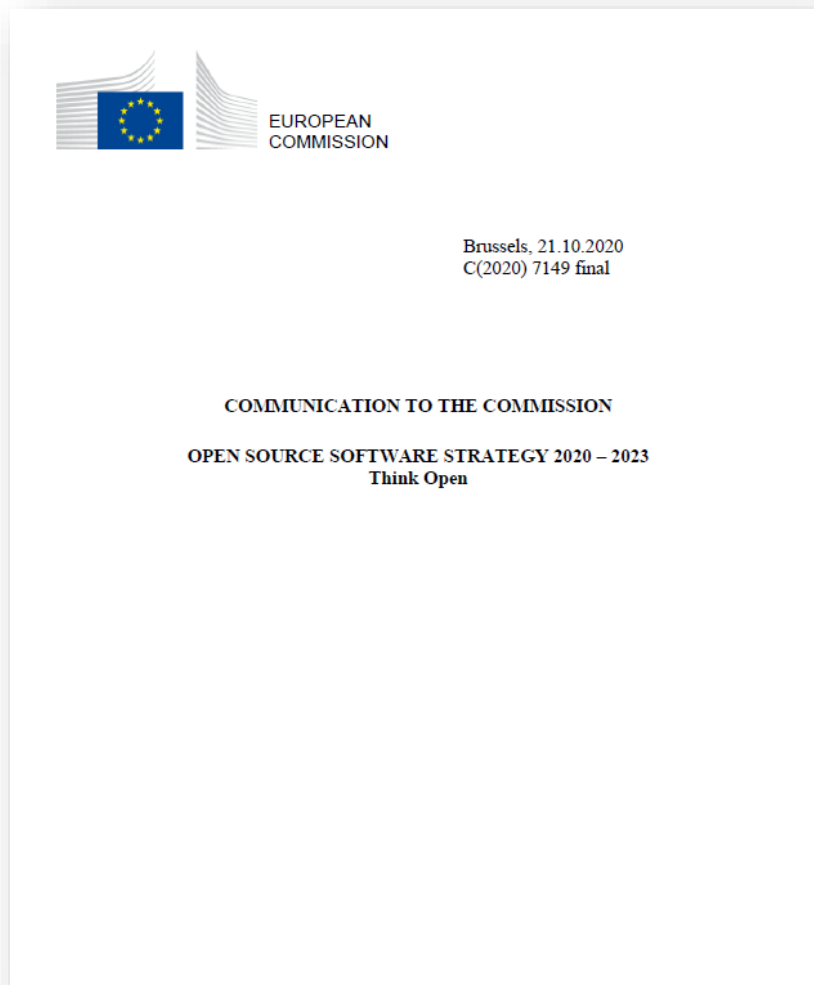- Final recommendations

European Commission

Open source at the European Commission

Open source at the European Commission

Our **software solutions** are becoming **open source**

European Commission

TRANSFORM

CONTRIBUTE   CONTRIBUTE

Drupal™

PRODUCE   PRODUCE   PRODUCE

code.europa.eu

USE tools   USE tools   USE tools   USE tools

USE infrastructure   USE infrastructure   USE infrastructure   USE infrastructure   USE infrastructure

2000   2003   2007   2014   2020

European Commission's open source policy

COMMUNICATION TO THE COMMISSION

OPEN SOURCE SOFTWARE STRATEGY 2020 – 2023
Think Open

- **Open Source Software Strategy 2020-2023 – October 2020**
  - impacts the entire organisation
  - links EC policy goals to open source
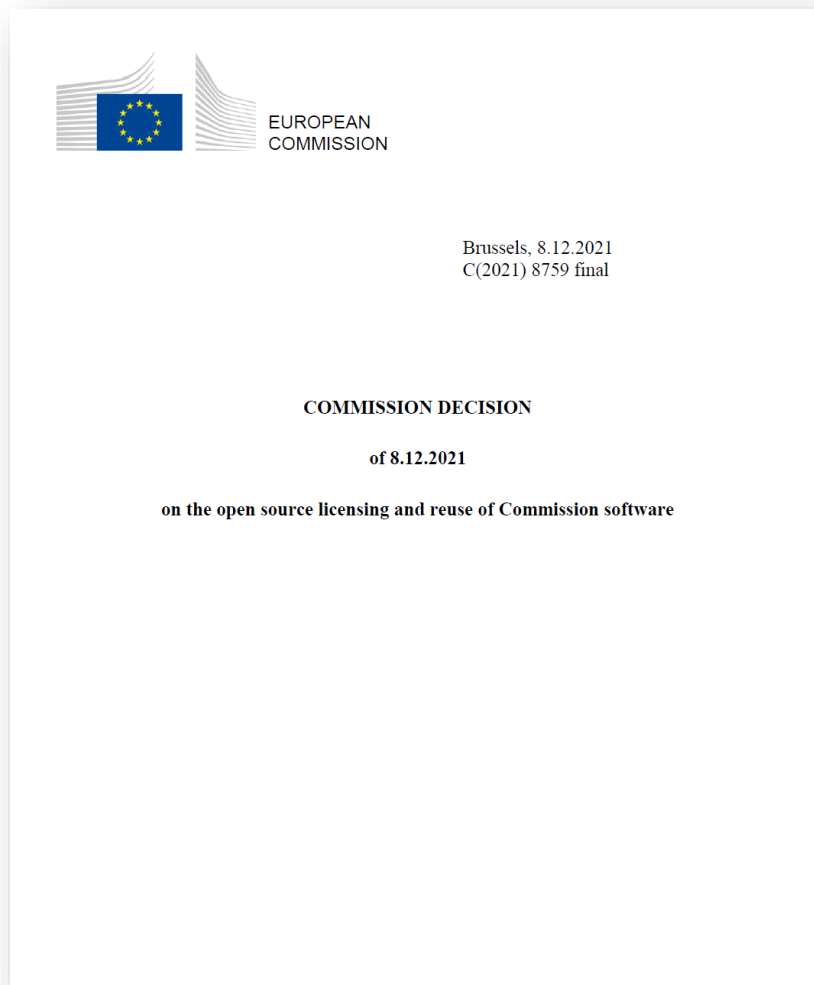 - 6 governing principles

    think open   contribute
    transform   secure
    share   stay in control

European Commission's open source policy

EUROPEAN
COMMISSION

Brussels, 8.12.2021
C(2021) 8759 final

COMMISSION DECISION

of 8.12.2021

on the open source licensing and reuse of Commission software

- Commission Decision on the open source licensing and reuse of Commission software – December 2021
  - simplifies publication of software as open source
    - default license: EUPL
    - alternative open source licenses can be used
  - requires a repository as a single point of access to Commission software
  - allows Commission services to contribute to open source projects
    - including transfer of the ownership of IP rights on the contributed code

European
Commission

European Commission's open source policy
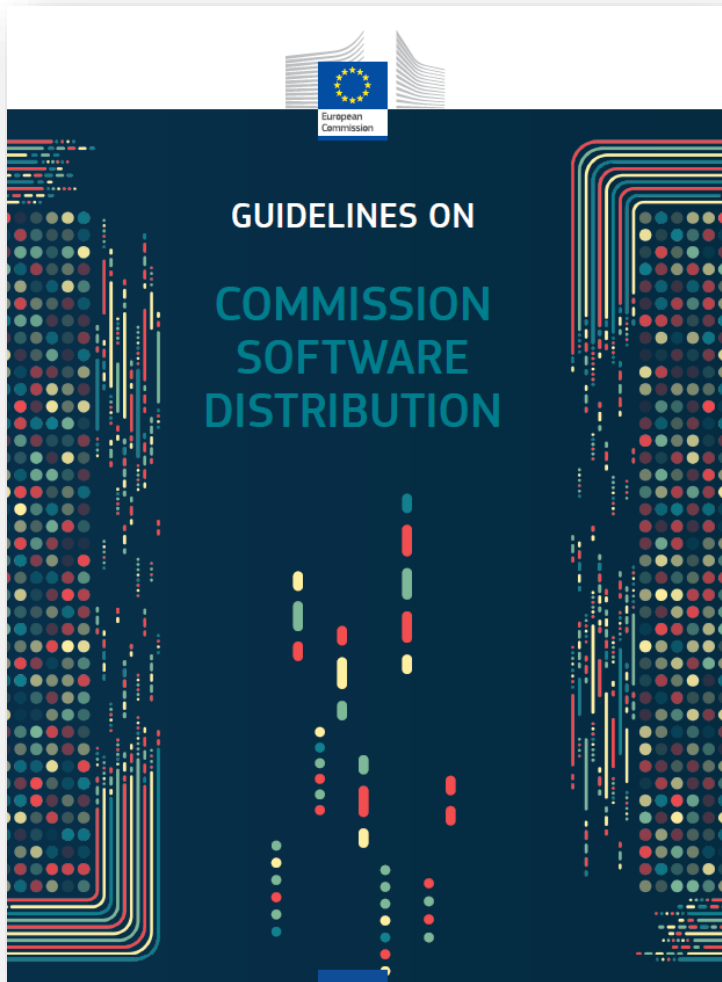
- **Guidelines on Commission software distribution – October 2022**
  - internal guidelines for open source distribution
    - software identification
    - IP clearance
    - vulnerability assessment

https://code.europa.eu/info/about/-/blob/master/guidelines/code-europa-eu-guidelines.md

European Union Public License

- Designed by the European Commission to be fully compliant with EU law
  - open source, approved by the OSI, copyleft
  - latest version v.1.2 (2017) available in 23 EU languages
- Highly compatible license
  - GPL v.2 & v.3, LGPL v.2.1 & v.3, AGPL v.3, CeCILL v.2.0, v.2.1, OSL v.2.1 & v.3.0, EPL v.1.0, MPL v.2

EUPL-1.2   European Union Public Licence, Version 1.2 or later (EUPL)

Can	Use/reproduce, Distribute, Modify/merge, Sublicense, Commercial use, Use patents, Place warranty
Must	Incl. Copyright, Royalty free, State changes, Disclose source, Copyleft/Share a., SaaS/network, Include licence
Cannot	Hold liable, Use trademark
Compatible	GPL, Other copyleft, Linking freedom, Multilingual, For data, For software
Law	EU/MS law, Licensor's law, Venue fixed
Support	Strong Community, Governments/EU, OSI approved, FSF Free/Libre

Licence comment:

Official Licence of the European Union (EC Decision, part of European law). The licence is **interoperable** (no restrictions on linking in order to facilitate the integration of multiple components), **reciprocal** (third parties distributing improvements or derivatives must publish and provide back the modified source code) and **compatible**: no global relicensing permitted, but the source code could be reused in other projects under GPL/AGPL, EPL, LGPL, MPL, OSL, CeCILL, LiLiQ. EUPL covers SaaS / network distribution. EUPL covers "the Work" (software and ancillary data). Original in 23 EU languages. Replaces EUPL-1.1 for works "Licensed under the EUPL" without specifying licence version, or adding "or later". Applicable law and court: licensor seat in EU (or specific additional agreement), otherwise Belgium. Support from the Joinup.eu community. Free legal support provided.

https://joinup.ec.europa.eu/collection/eupl/eupl-text-eupl-12
https://joinup.ec.europa.eu/sites/default/files/custom-page/attachment/eupl_v1.2_en.pdf

Open source repository

https://code.europa.eu

Outreach to communities

Bug Bounties: Collabora Online, Dynare, jupyter, XWIKI, odoo, CryptPad, LibreOffice, moodle, element, zimbra

Hackathons: Nextcloud, jitsi, X-ROAD, moodle

European Commission

Outreach to communities

Bug Bounties

Collabora Online
Dynare
jupyter
XWIKI
odoo
CryptPad
LibreOffice
moodle
element
zimbra

Hackathons

Nextcloud
jitsi
X-ROAD
moodle

OSS4SDG
Open Source Software for Sustainable Development Goals

Sustainable Smart Cities

UNITED NATIONS OFFICE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY

European Commission

OpenStreetMap

https://ideas.unite.un.org/sdg11/Page/Overview

European Commission

Experiences at the JRC – Openness in data sharing

INSPIRE

DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 14 March 2007

establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)

- Directive from 2007 establishing a European Spatial Data Infrastructure for environment policies
  - 7k+ data providers from Member States public sector
  - currently about 90k datasets shared
- JRC is the Technical Coordinator
  - operation, maintenance & evolution of the infrastructure
  - technological and organisational focus

European Commission

JRC SCIENCE FOR POLICY REPORT

INSPIRE • A Public Sector Contribution to the European Green Deal Data Space
A vision for the technological evolution of Europe's Spatial Data Infrastructures for 2030

Alexander Kotsev, Marco Minghini, Vlado Cetl,
Friso Penninga, Joeri Robbrecht, Michael Lutz

Joint Research Centre

EUR 30832 EN

https://inspire.ec.europa.eu
https://publications.jrc.ec.europa.eu/repository/handle/JRC126319

INSPIRE open principles – Software

- INSPIRE (legal framework) as a catalyst for technological innovation:
  - central components based on reusable, open source software solutions

INSPIRE Geoportal, based on GeoNetwork opensource

INSPIRE Reference Validator, based on the ETF testing framework

INSPIRE Registry, based on the Re3gistry

INSPIRE open principles – Software

- INSPIRE (legal framework) as a catalyst for technological innovation:
  - central components based on reusable, open source software solutions
  - decentralised governance with multiple (open source) actors OSGeo
    - ETF and Re3gistry included in the OSGeoLive, ETF proposed as OSGeo Community Project

https://live.osgeo.org/en/overview/ETF_overview.html
https://live.osgeo.org/en/overview/re3gistry_overview.html

https://www.osgeo.org/projects/etf

European Commission

INSPIRE open principles – Software

- INSPIRE-related software for data search, provision & consumption
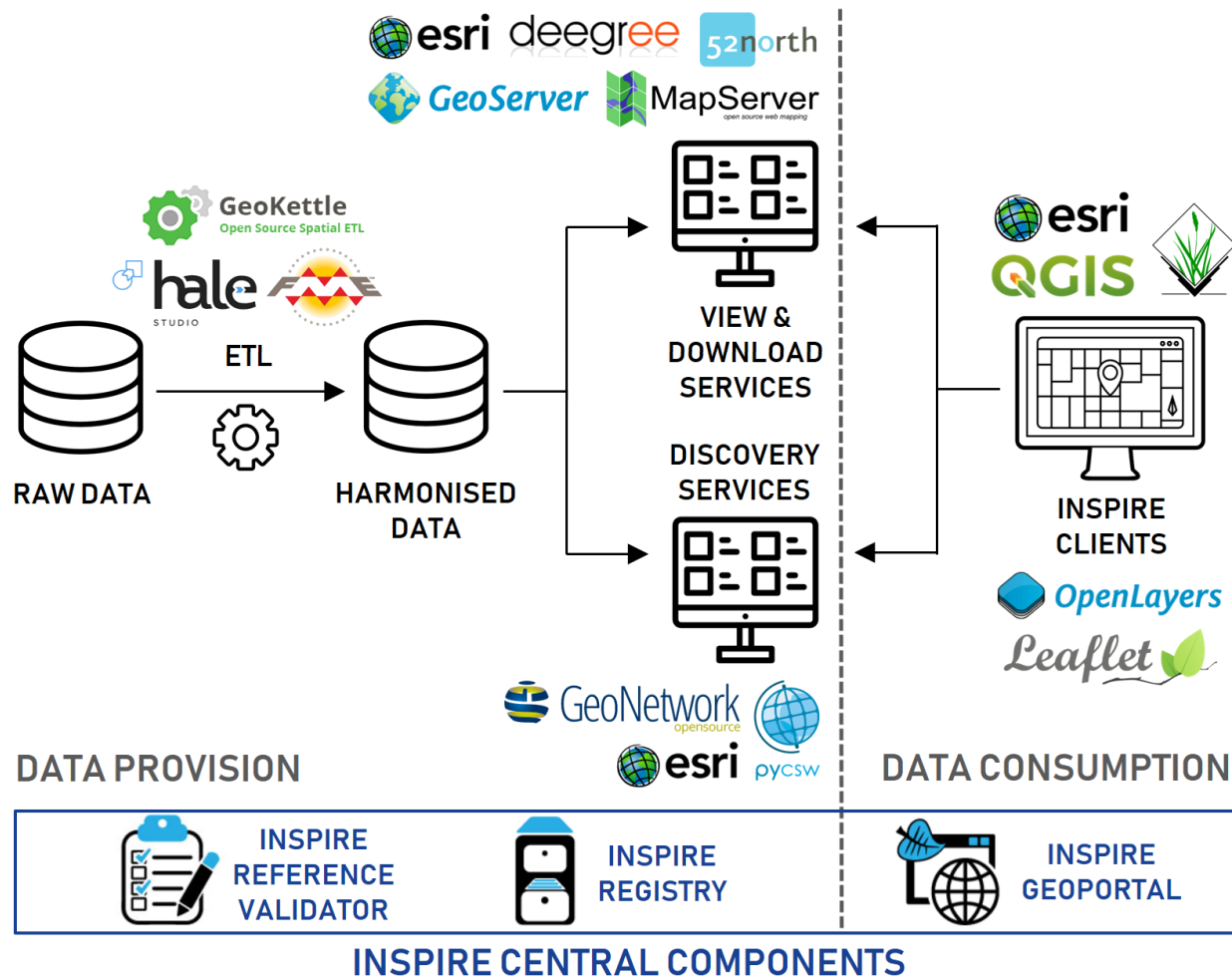
INSPIRE open principles – Standards

- By nature based on open standards
  - probably the largest uptake of OGC standards worldwide
  - benefits for all: data providers, users & standardisation bodies

INSPIRE open principles – Governance

- Active community
  - conferences, discussion forums, helpdesks
- Community-driven processes
  - inclusive approach since the beginning
  - INSPIRE Good Practices to introduce new approaches (standards, technologies, etc.) in INSPIRE
  - governance process to manage changes/updates to INSPIRE artefacts (Technical Guidelines, schemas and UML models)

https://inspire.ec.europa.eu/portfolio/good-practice-library
https://github.com/INSPIRE-MIF

Good Practice Library

Good Practice documents

Candidate	Endorsed
GeoPackage encoding of INSPIRE datasets	GeoDCAT-AP
Data-Service Linking Simplification	SDMX for Human Health and Population Distribution
	OGC API – Features as an INSPIRE download service
	OGC SensorThings API as an INSPIRE download service
	Building one access point to dispersed data sources
	Making spatial data downloadable via WMS services
	OGC compliant INSPIRE Coverage data and service implementation

JRC contribution to GEO & GEOSS

- Long-term contribution to GEO
  - defined in the Horizon Europe WP
  - EAG, Programme Board, WGs
- in close collaboration with other EC services (RTD, DEFIS, CNECT, ENV, etc.) and the GEO community
- multiple research outputs with a science for policy emphasis
  - datasets
  - services
  - analyses tailored to the needs of the GEO community
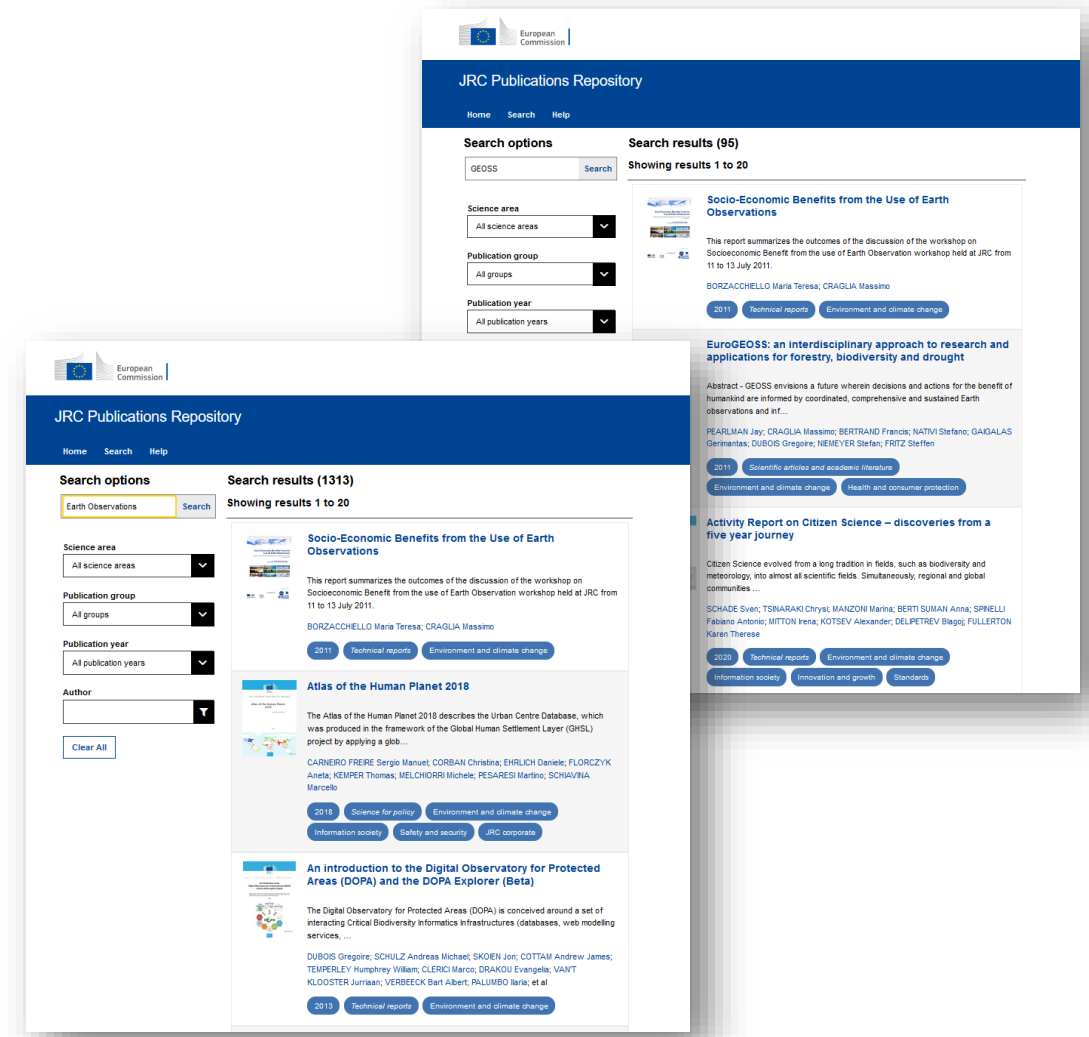
JRC contribution to EuroGEO

- European regional GEO initiative

  - umbrella framework to showcase and scale-up European user-driven products/services: Copernicus DIAS, INSPIRE, ESA TEPs, in-situ data sources, citizen observatories, etc.

  - EuroGEOSS virtual digital infrastructure as European contribution to GEOSS to address environmental use cases

- Identification of approaches for modernising data sharing in EuroGEO

  - alignment with the European policy context around data spaces

  - tackling fragmentation of open EO infrastructures

  - streamlining end-to-end process from raw data to insights for decision makers

  - user-driven and demand-driven

  - changing context – hyperscalers

European Commission

Mapping the existing landscape

- Categories of services offering by existing European EO digital platforms
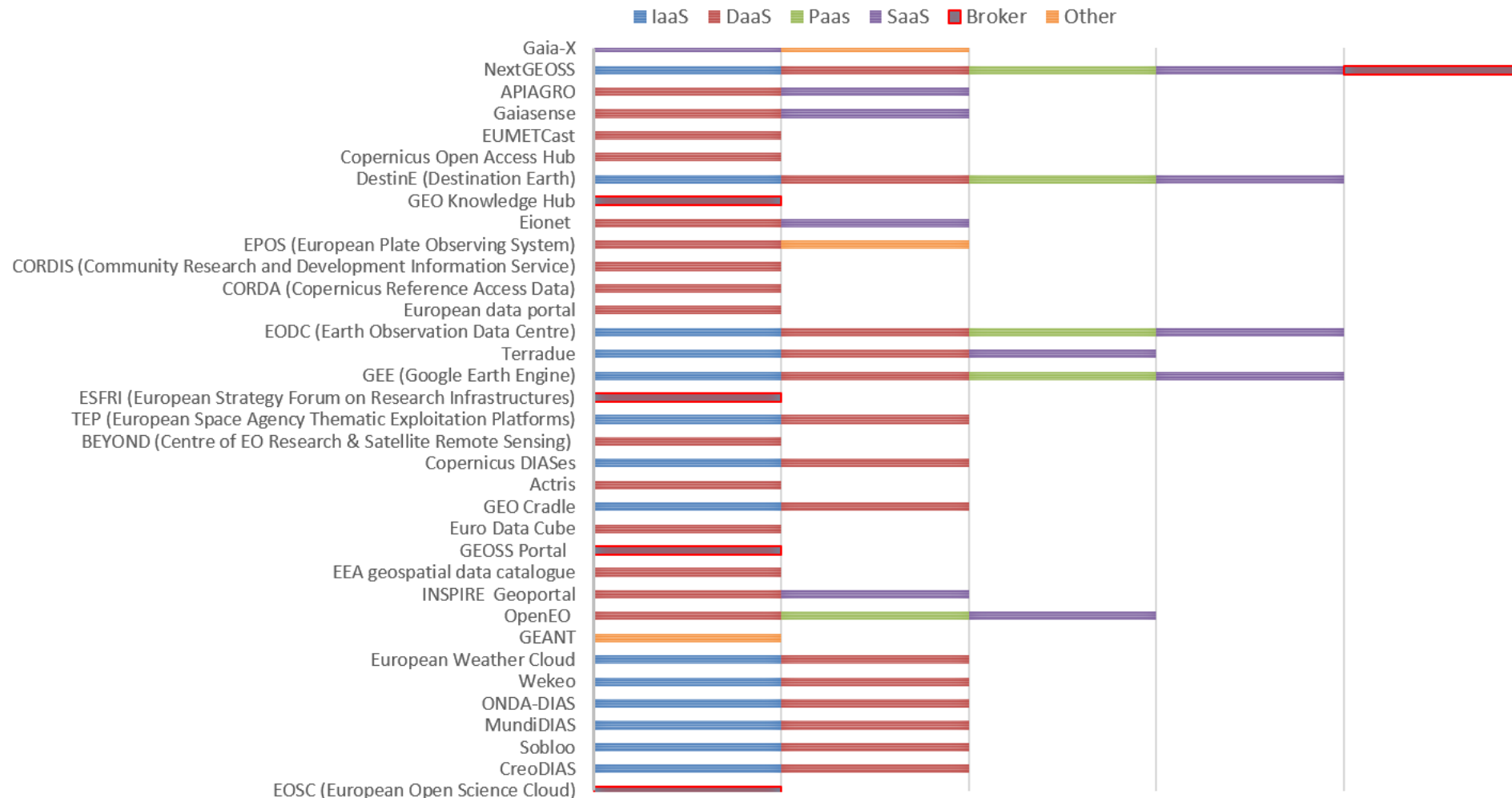
Mapping the existing landscape

- Platform usage by e-shape pilots, classified by showcase
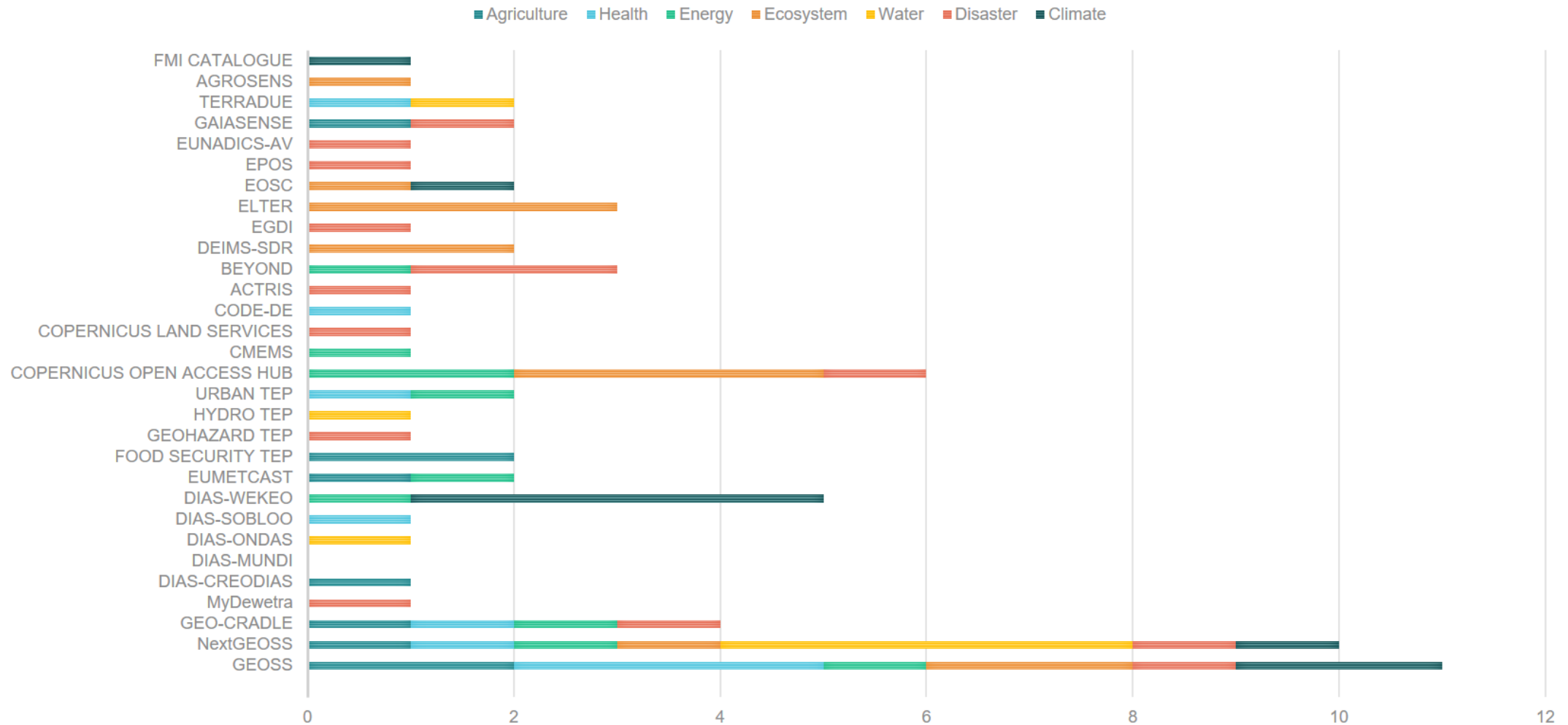
Mapping the existing landscape

- Source datasets used in e-shape pilots, classified by showcase
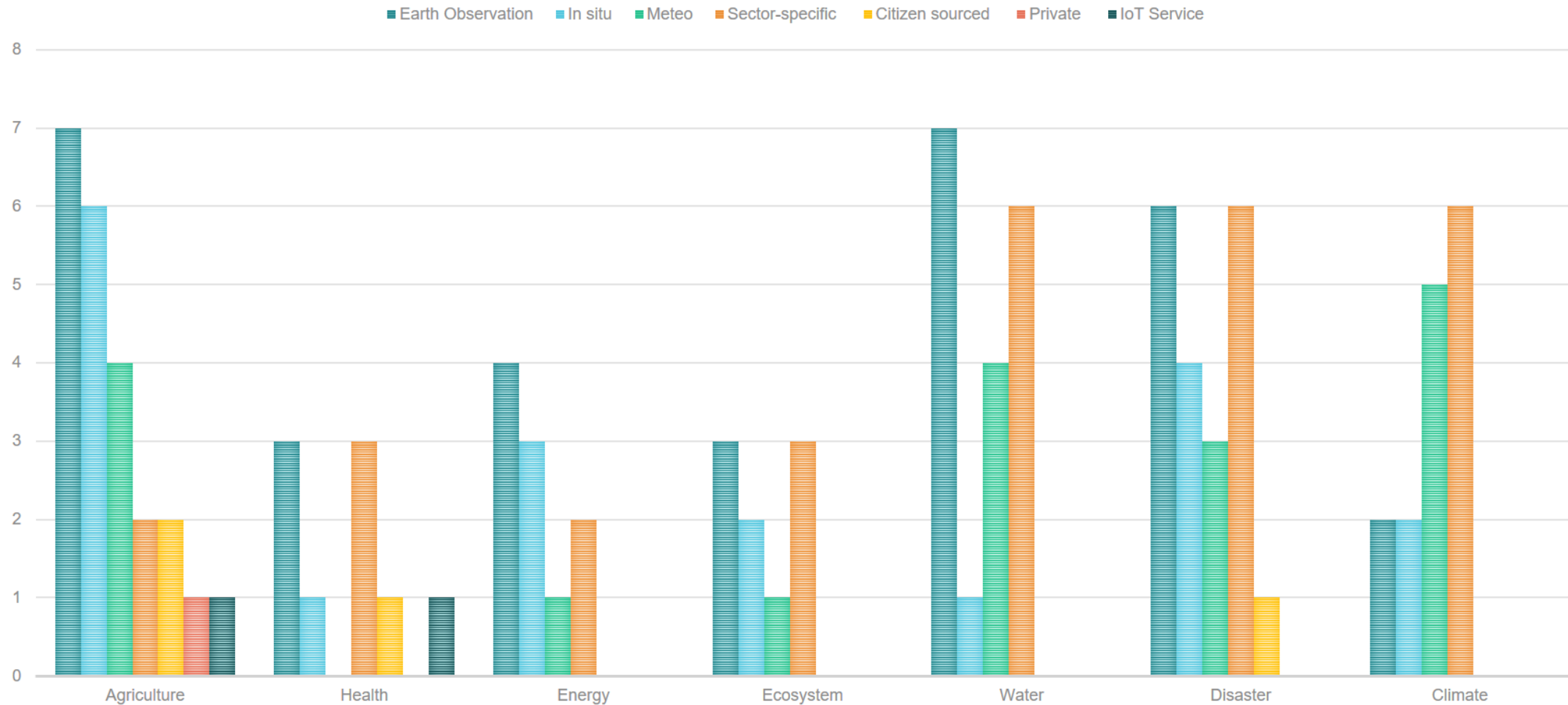
Mapping the existing landscape

- intended user categories for e-shape pilots, classified by showcase
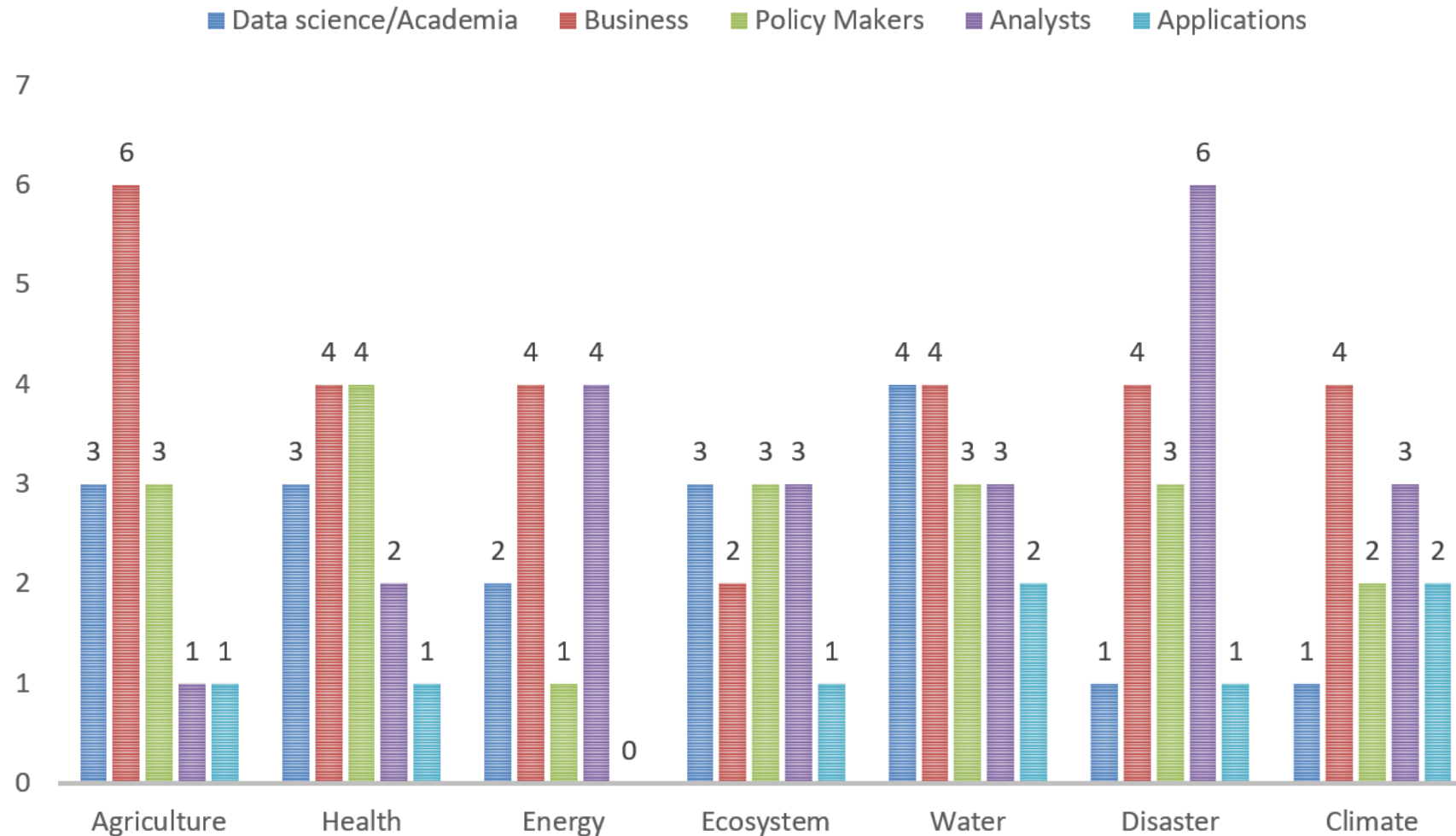
Development life cycle & available options

- Definition of a developer journey
  - considering emerging technologies & trends
  - based on some principles:

```
┌──────────────────┐
│ Definition of idea│
└──────────────────┘
         │
         ▼
┌──────────────────┐
│   Data search    │
└──────────────────┘
         │
         ▼
┌──────────────────┐
│  Data ingestion  │
└──────────────────┘

┌──────────────────┐
│ Model development│
└──────────────────┘

┌──────────────────┐
│       Test       │
└──────────────────┘

┌──────────────────┐
│   Deployment     │
└──────────────────┘
         │
         ▼
┌──────────────────┐
│      Scale       │
└──────────────────┘
```
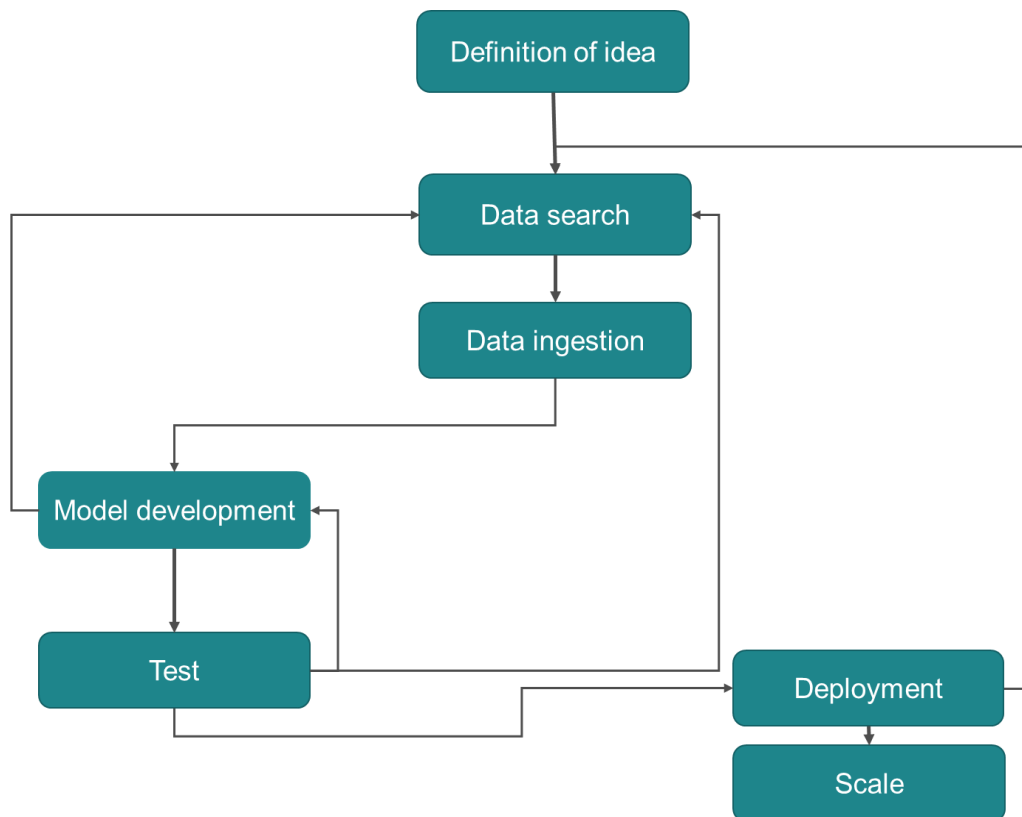
- prefer open source software
- avoid vendor/technology lock-in
- avoid approval processes for user projects as much as possible
- limit data movement as much as possible
- reuse existing tools/modules
- combine data from different infrastructures/services
- adhere to FAIR principles

European Commission

Development life cycle & available options

- Definition of a developer journey

- High-level infrastructure landscape/analysis

  - GEO ecosystem vs. DIASes vs. GEE
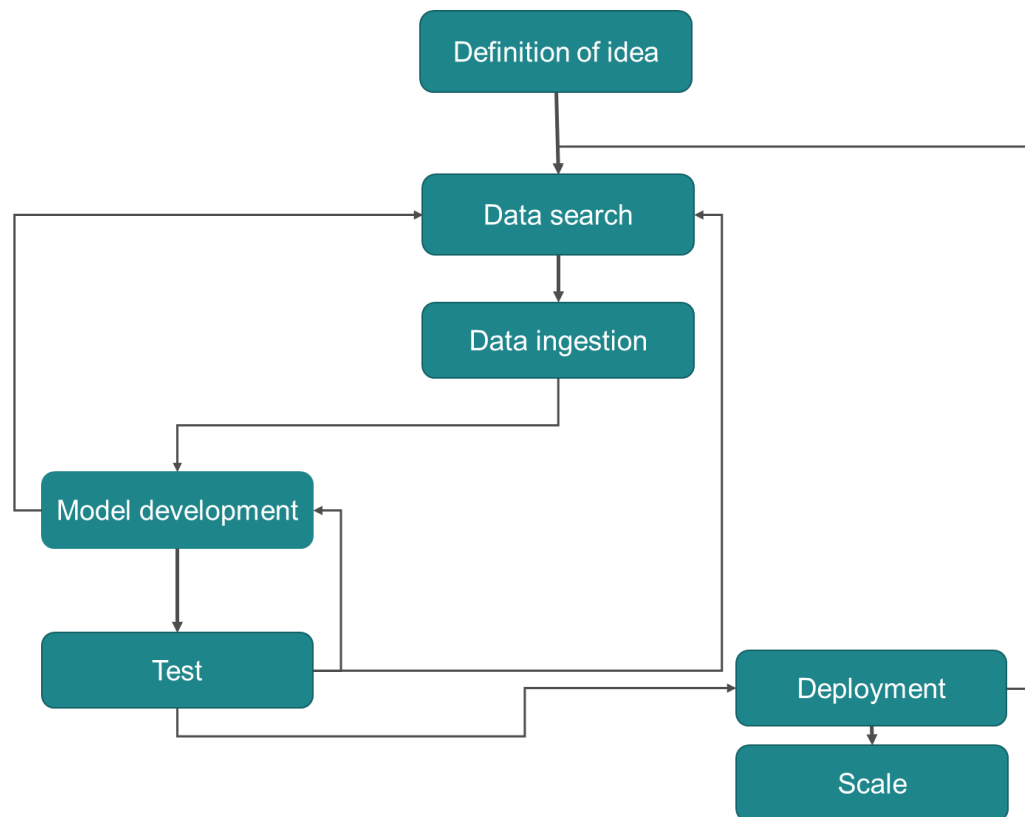
	Google Earth Engine	GEOSS ecosystem	DIASes
Data search	Google Datasets	GEODAB	EO browser (Copernicus data only)
Data ingestion	Earth Engine APIs	GEODAB	S3, NFS, APIs, OGC
Model development	Integrated development env or Google colab	ESA TEPs, Online Jupyter Hubs/Labs	Jupyter Hubs/Labs
Test			
Deployment	GCP (VM or serveless), Google AI Platform	ESA TEPs, VLAB, other thematic platforms. DIASes	Manual deploy
Scale	Autoscaling can be managed through GCP	In some environment should be possible to handle it programmatically	Not all DIASes support automatic scaling.

Use-case driven requirement analysis

- Setup of prototypical data-driven EO applications

    - based on the knowledge from existing European pilots & the development life cycle

    - using European infrastructure (GEO portal and DIAS platforms)

    - 2 use cases from different domains, addressed in an independent way

- Target

    - identify inefficiencies and bottlenecks

    - identify potential areas of improvement

    - document the utility, maturity and reusability of the technical stack

    - distill recommendations for GEO and EuroGEO
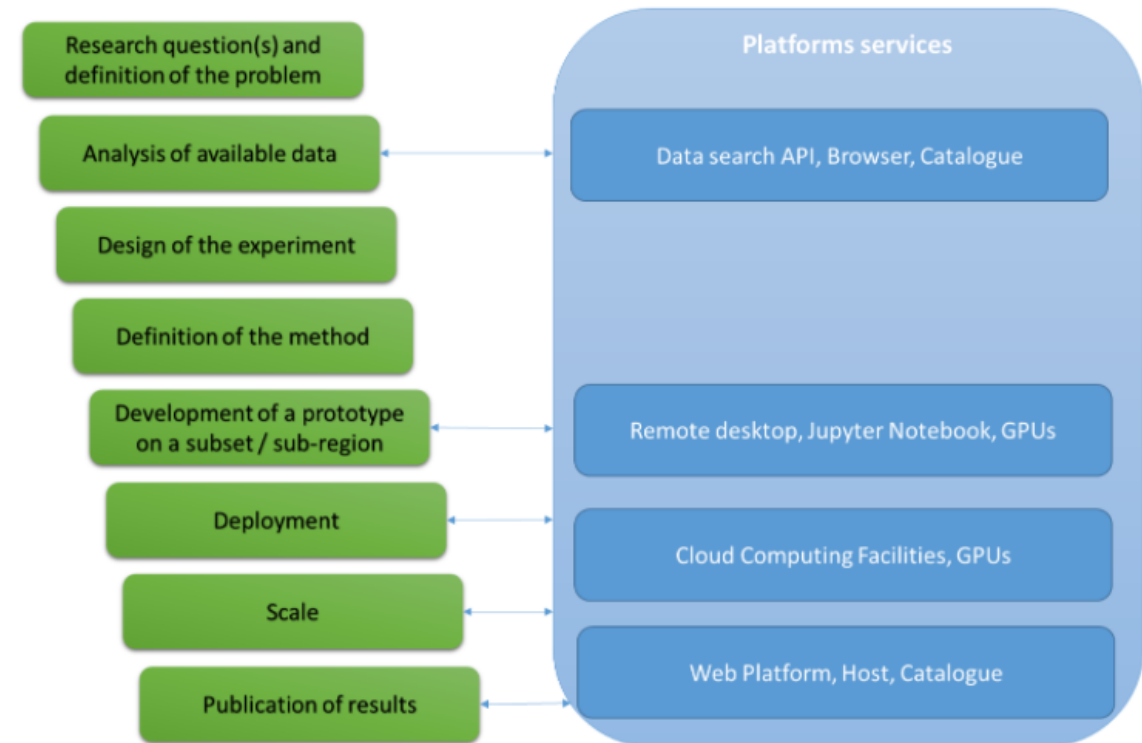
Final recommendations

Final recommendations

- When starting/developing an open source project
  - when possible, reuse the existing rather than developing new tools
  - choose an established license maximising compatibility and reuse (no custom licenses)
  - design with a user-driven & demand-driven approach
  - validate that user needs are satisfied
  - establish a governance and a community around the project to ensure sustainability

European Commission

Final recommendations

- When starting/developing an open source project
  - when possible, reuse the existing rather than developing new tools
  - choose an established license maximising compatibility and reuse (no custom licenses)
  - design with a user-driven & demand-driven approach
  - validate that user needs are satisfied
  - establish a governance and a community around the project to ensure sustainability

- Openness as a working culture
  - not only open source software, but open data, open standards, FAIR principles, etc.
  - partner with existing projects & communities (outreach, sponsor, etc.)
  - consider taking a role in the governance of software projects, open source communities, standardisation bodies, EO initiatives, etc.

European Commission

Thank you!

and thanks to: Alexander Kotsev, Josep Soler Garrido, Jordi Escriu, Margherita Di Leo, Nicholas Spadaro, Gijs Hillenius

✉ marco.minghini@ec.europa.eu

European Commission

Keep in touch

EU Science Hub: ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub – Joint Research Centre

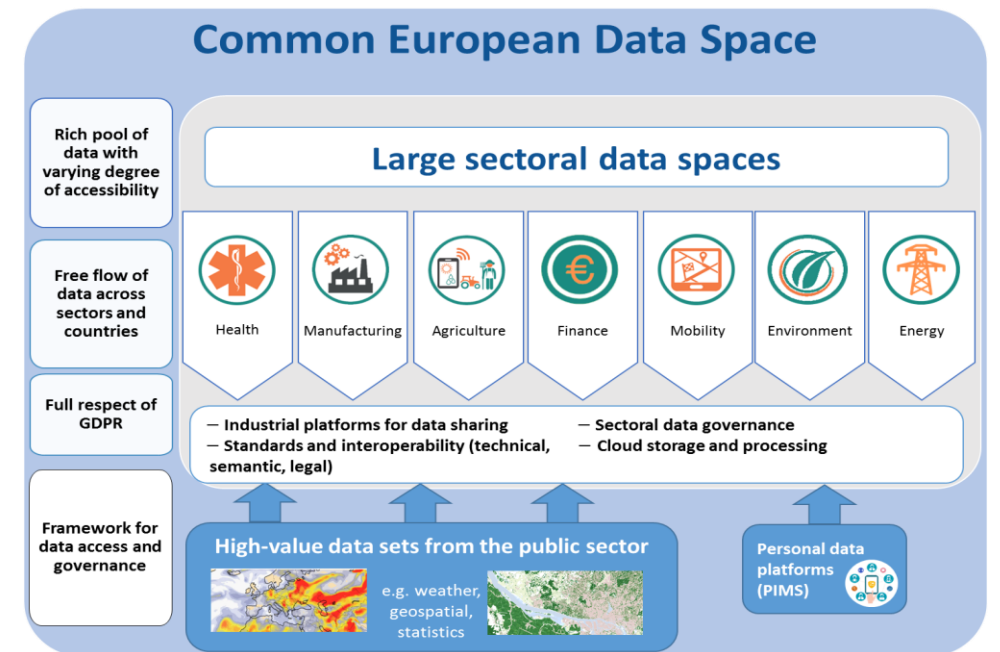EU Science, Research and Innovation

Eu Science Hub

European Commission

Common European data spaces

Brussels, 19.2.2020
COM(2020) 66 final

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN
PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL
COMMITTEE AND THE COMMITTEE OF THE REGIONS

A European strategy for data

https://digital-strategy.ec.europa.eu/en/policies/strategy-data

- European strategy for data – February 2020
  - vision: create a European single market for data
  - problems: data availability, licensing, governance, quality, interoperability, infrastructure, technology, skills, data literacy, cybersecurity

- Creation of common European data spaces
  - no formal definition, only characteristics
    - goal is to create value to make better decisions
    - data from all actors, personal & non personal
    - legal, organisational, semantical & technical focus
    - based on European values and legal framework

Common European Data Space

Large sectoral data spaces

Rich pool of data with varying degree of accessibility

Free flow of data across sectors and countries

Health | Manufacturing | Agriculture | Finance | Mobility | Environment | Energy

Full respect of GDPR

– Industrial platforms for data sharing — Sectoral data governance
– Standards and interoperability (technical, semantic, legal) — Cloud storage and processing

Framework for data access and governance

High-value data sets from the public sector
e.g. weather, geospatial, statistics

Personal data platforms (PIMS)

European Data Spaces Cookbook (coming soon)

- A practical guide to setting up, operating and maintaining European data spaces
  - target audience: stakeholders of data spaces
    - data providers, data users, data intermediaries, open source communities, standardisation bodies, etc.
  - Content:
    - technical requirements of data spaces
    - data sharing how-to's
    - JRC knowledge base