



# NINA Data Infrastructures Planning

Copernicus Brukerforum - Oslo – 4 Mai 2022

- Matteo De Stefano



# Contents

---

- DM = Data Management
- Intro about NINA DM
- A time perspective:
  - ▶ The past and present (in shift)
  - ▶ Planning the future DM
- Questions about possible processing solutions
- A desired goal
- Opportunity for cooperation

# NINA DM Considerations

---

- Managing data in a context like NINA is complex -> increasing!
- Many data types, formats, workflows, needs...
- DM historically based on project requirements
- DM practices evolve constantly
- Hot questions:
  - ▶ How to organize the information?
  - ▶ Where data is stored? How accessed?
  - ▶ How to process data in the most efficient way?

# NINA DM – a time perspective

---

- The (Past and) PRESENT – Varied solutions, mostly centralized, currently reorganizing
- The FUTURE - To more shared, cloud, interoperable and standardized solutions!

# PRESENT DM - storage

---

- A variety of storage solutions
  - ▶ Some centralized
    - Monolithic postgis DB
    - Sqlserver spatial DB
    - Geospatial data from different sources organized in the intranet file system following Inspire nomenclature
    - IPT for biodiversity data
    - GRASS data
  - ▶ Some based on specific project requirements

# PRESENT DM - processing

---

- A variety of processing solutions
  - ▶ GIS desktop and online tools
  - ▶ Google Earth Engine
  - ▶ Own scripts in various languages
  - ▶ SQL
- Heterogeneous data, which has to be downloaded and managed manually

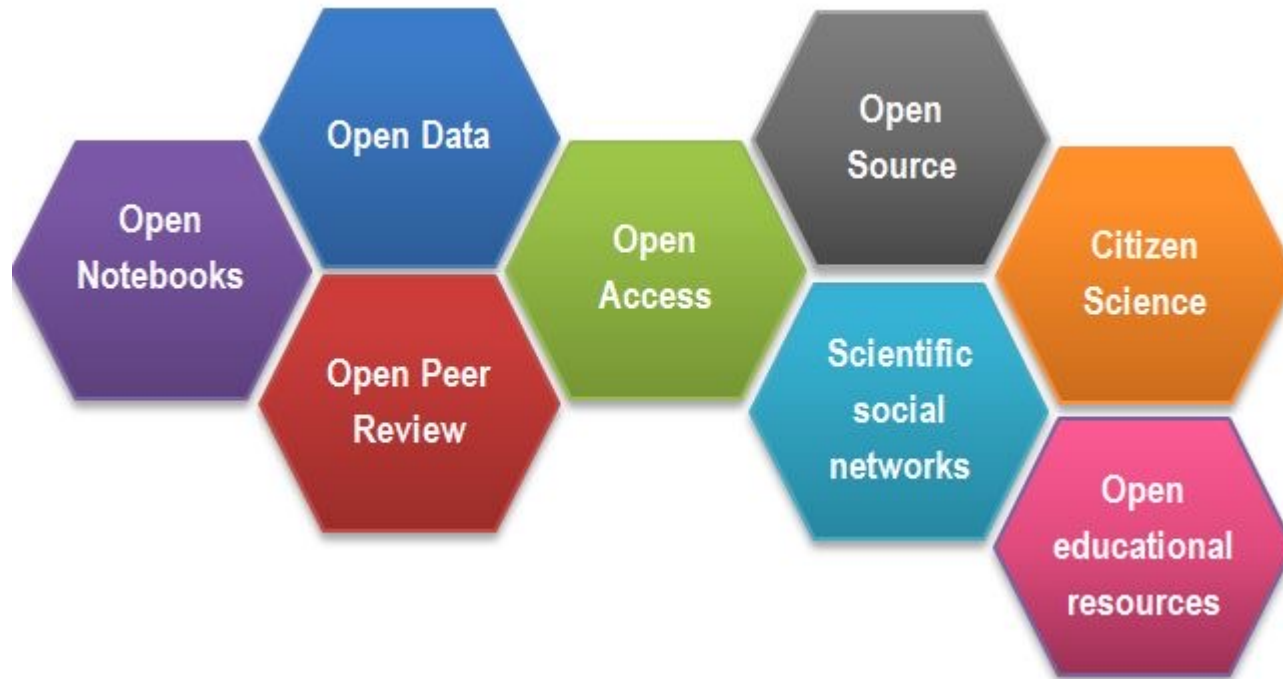
# FUTURE DM (?)

---

- Standardization!
- FAIR catalogues and portals
- Increase sharing of data and metadata on national facilities
- Finding new data processing infrastructures and workflows
- Improved cooperation between national stakeholders

# Cultural and technical shift

Going towards Open Science, FOSS, and FAIR data





# Standards

---

- Standardize internal workflows
  - ▶ New tools for data collection
  - ▶ Initiatives for data harmonization
- Standardize data formats and types
- Manage metadata – keep it close to data!
  - ▶ Choose the best metadata standard/s
- Use vocabularies (ontologies?)
- Design infrastructures with linked data as a final goal

# Building FAIR infratructures

---

Aiming at standard and FAIR Data Management practices in NINA, we are building:

Generic

Spatial

External



Internal



# New processing solutions?

---

- Adapt to different formats and workflows – Satellitdata and OpenDAP?
- New OGC API including web processing? Suits all requirements?
- Use DIAS? How to deal with national data widely used?
- Increase in house computational hardware?
- Keep going with current solutions? GEE and similar involve data transfer costs

# Possible improvements?

---

- Exploring massive data caching – avoid manual downloading data, use data provided by other institutions (ex: GeoNorge) and cache data on request locally, ready to be used.
- preferred:

• SIGMA 2!

# Why Sigma 2?

---

- It is a huge national e-infrastructure
- Huge storage and computational capacity
- Would be an ideal national infrastructure for RS and Copernicus storage and processing
- High flexibility (potential) in the processing tools to be used (kubernetes infrastructure, which allows the use of custom containers with any type of software)

# Sigma 2- challenges?

---

- It would require a huge coordination effort between institutions
  - ▶ It is worth, we believe!
- There could be some political resistance
  - ▶ Sometimes some competition between institutions with similar mandates
- Technical solutions, policies and business plans are not ready for the use we would need.
  - ▶ They could be changed provided a strong and clear enough request
  - ▶ Good catalogues are crucial...

# National cooperation

---

- Many ongoing initiatives
  - ▶ Forum for Geografisk Informasjon (FoU)
  - ▶ Copernicus Brukerforum
  - ▶ Living Norway Ecological Data Network (Biodiversity data)
  - ▶ S-Enda project (Dynamic geodata Interoperability)
- Opportunity for cooperation and lobbying
  - ▶ Unify efforts to request truly useful national infrastructures for data storage and processing
  - ▶ Win win perspective!

# Feedbacks?

---

- Mostly questions, perspectives and opportunities.

What do you think?



# Cooperation and expertise for a sustainable future



Photo: A. Staverløkk