The role of Persistent Identifiers
for interoperability in EOSC

Dr Maggie Hellström

ICOS Carbon Portal & Lund University

PT presidency e-IRG workshop May 25, 2021

Key aspects

- To support and enable research that is FAIR, identifiers must be globally unique, persistent and resolvable (PID-> GUPRID)
- Three important functions:
  - Unambiguous identity (tied to fixity)
  - Indirection to "landing page" and/or resource itself
  - Basic metadata at registry level ("kernel" profiles)

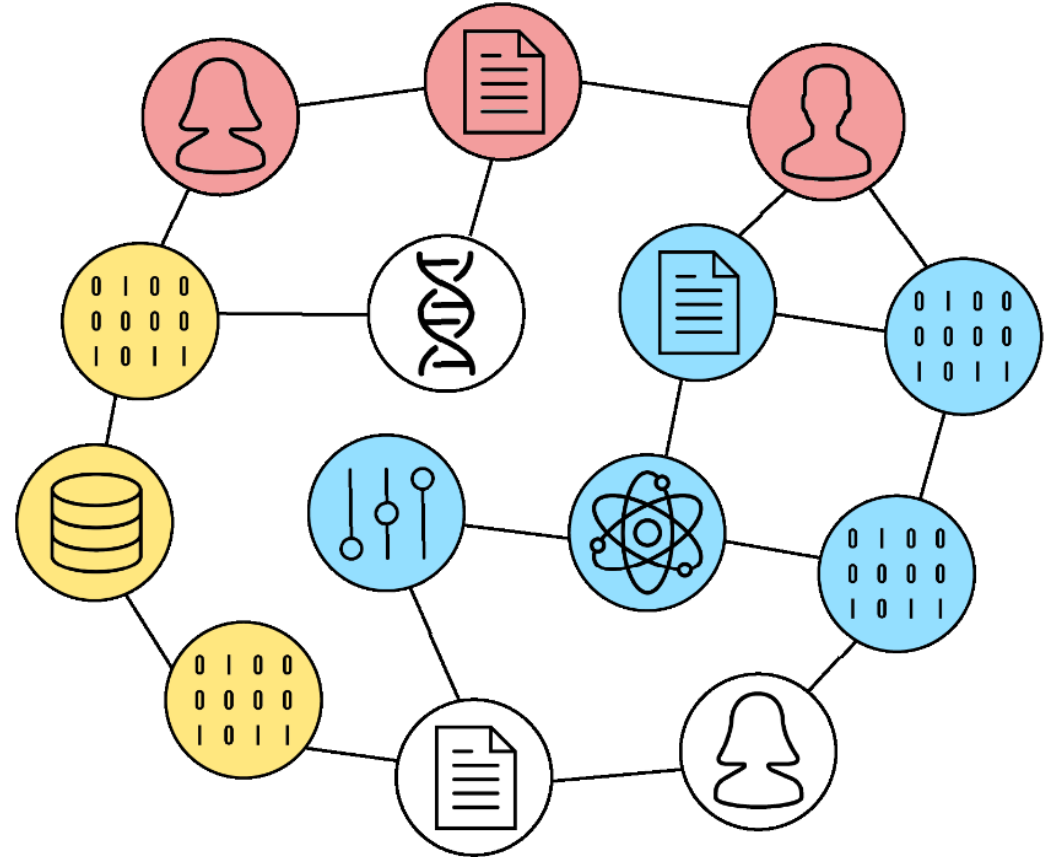# Resolution gives access to DOs & metadata
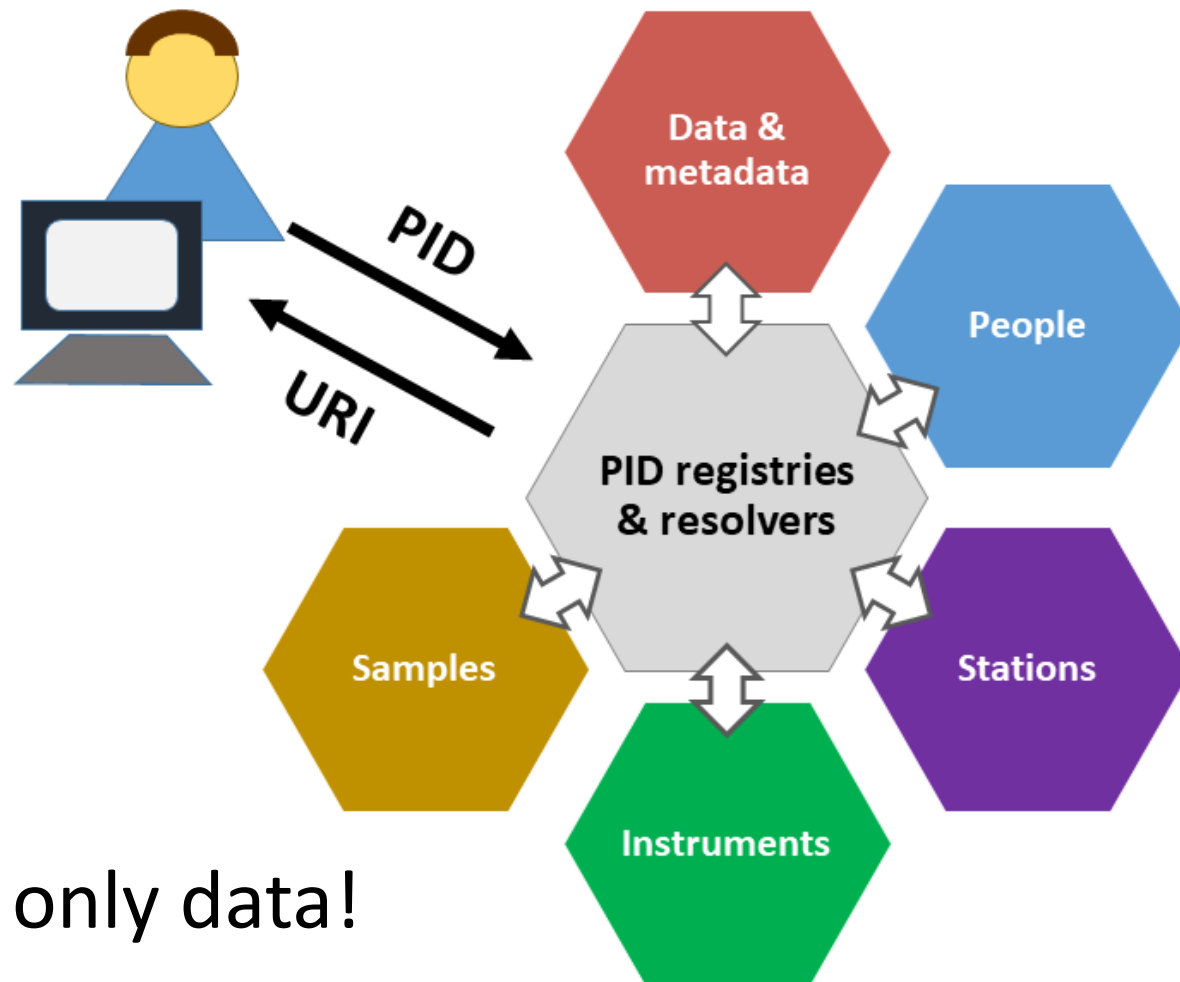
not only data!
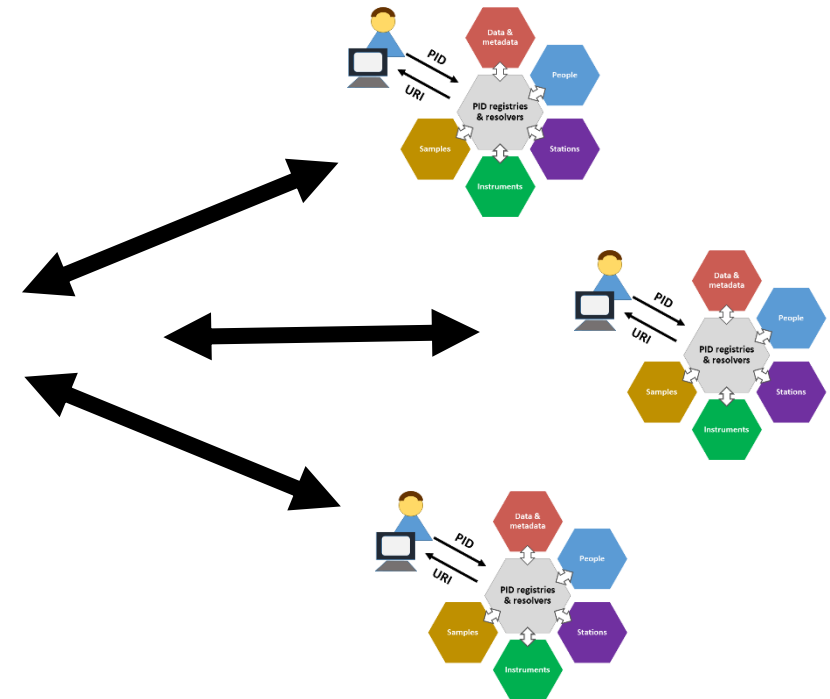
iterative lookups:
metadata, types, ...

PIDs as interconnectors

Research lifecycle

Web of objects

PID Policy for EOSC

- Principles, desired results and governance for a viable, trusted PID infrastructure suitable for the long-term sustainability of the EOSC

- Towards a future where PIDs can be used as the preferred method of referring to its assigned entity

- A functioning PID ecosystem requires interoperability both between PID Services and between PID Service Providers

- The Policy may be targeting service providers, but is relevant for all stakeholders!

PID Architecture for EOSC

- Components, Actors and Activities involved in PID "ecosystem"
- Performance and Scalability
- Security, Reliability and Resilience
- Needs for certification
- Example implementations
- Gaps in the Existing PID Landscape
- Challenges for Investigation and Development

European Commission (2020): PID Architecture for the European Open Science Cloud.
https://doi.org/10.2777/525581

Actors in the PID ecosystem

e-Services

Producers

Users

Stewards
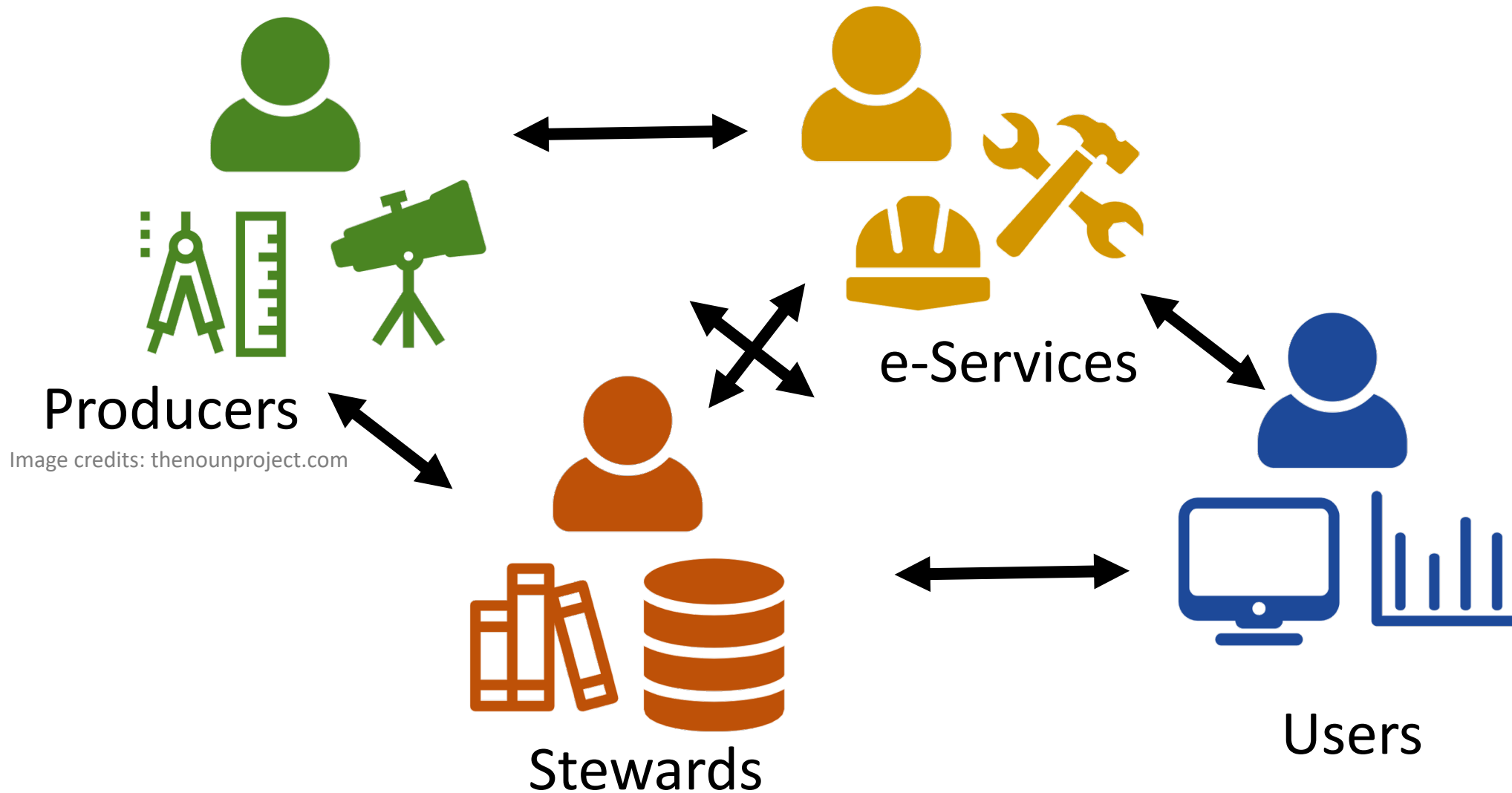
Challenges ahead

- Existing PID service landscape is rich, but not fully developed

- Need for sustainable services, also for non-data resources, including basic system support PIDs (types, metadata models, ...)

- A number of challenges, technical and metadata-related
  - Generic PID minting (free at point of use)
  - Common mechanisms for (complex) querying across PID systems
  - Comprehensive application of PID kernel metadata profiles and types
  - Performance & scalability of PID registries & resolvers
  - Security, reliability & resilience