

# Can a Digital Object Architecture help ICOS streamline its data service provisioning?

Maggie Hellström

ICOS Carbon Portal  
& Lund University

GEDE Digital Objects workshop  
Brussels, September 26, 2018

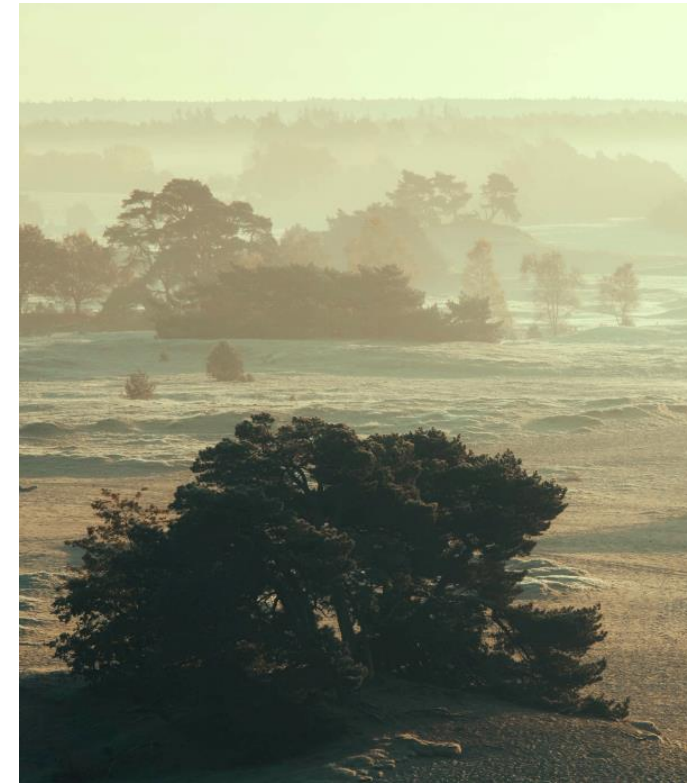


# ICOS in a nutshell

Integrated Carbon Observation System – ERIC since 2015

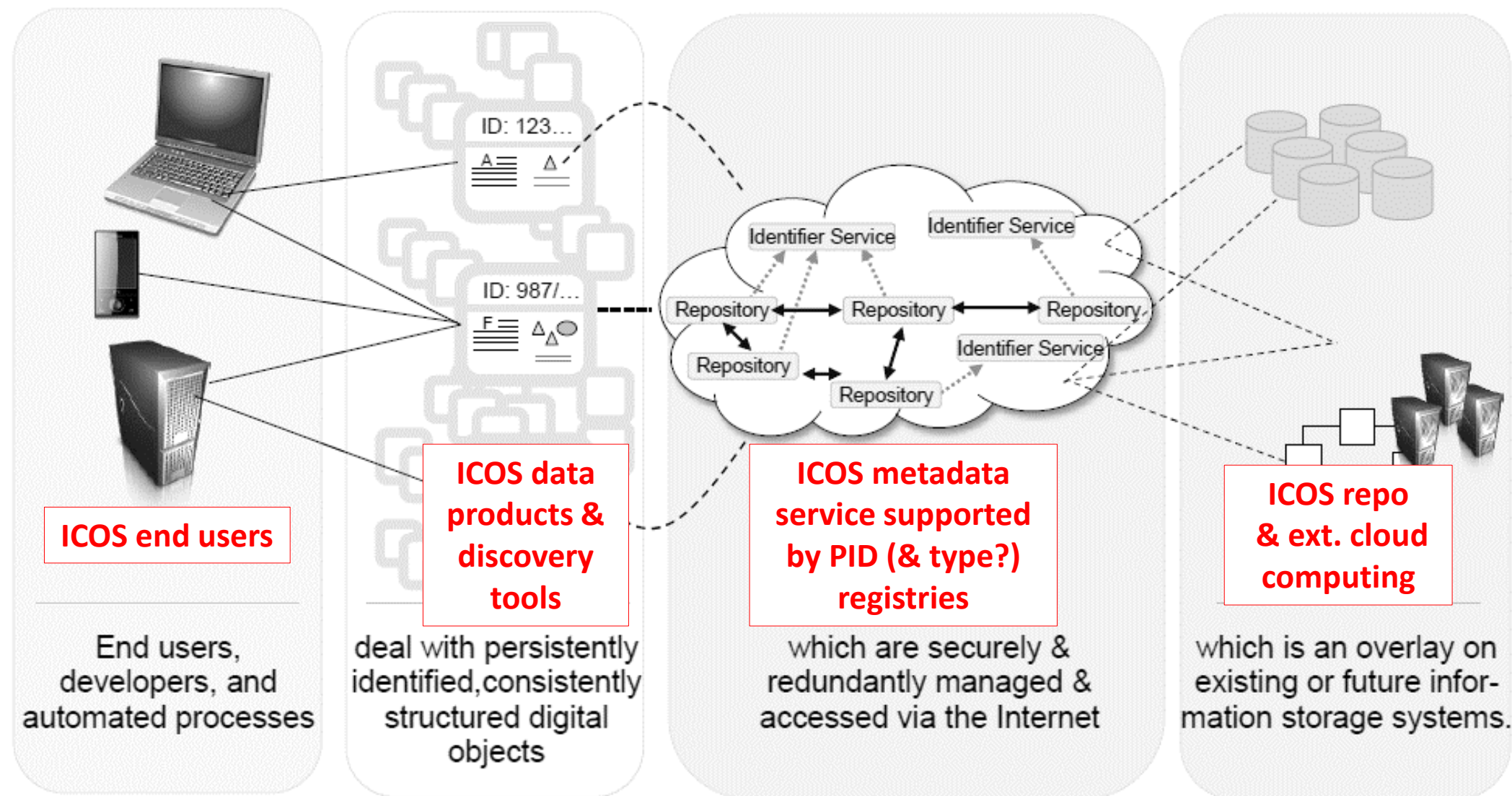
- Producing harmonised European-wide measurements on carbon cycle, on greenhouse gas emissions and on atmospheric concentrations of greenhouse gases
- Integrating atmosphere, ecosystem and ocean observations across the Europe and beyond
- Central Facilities for data processing, quality control, calibration, instrument development and training
- Reliable data curation and open data access through the ICOS Carbon Portal
- State-of-the-art infrastructure for the European research community, policy makers and public

<https://www.icos-ri.eu>

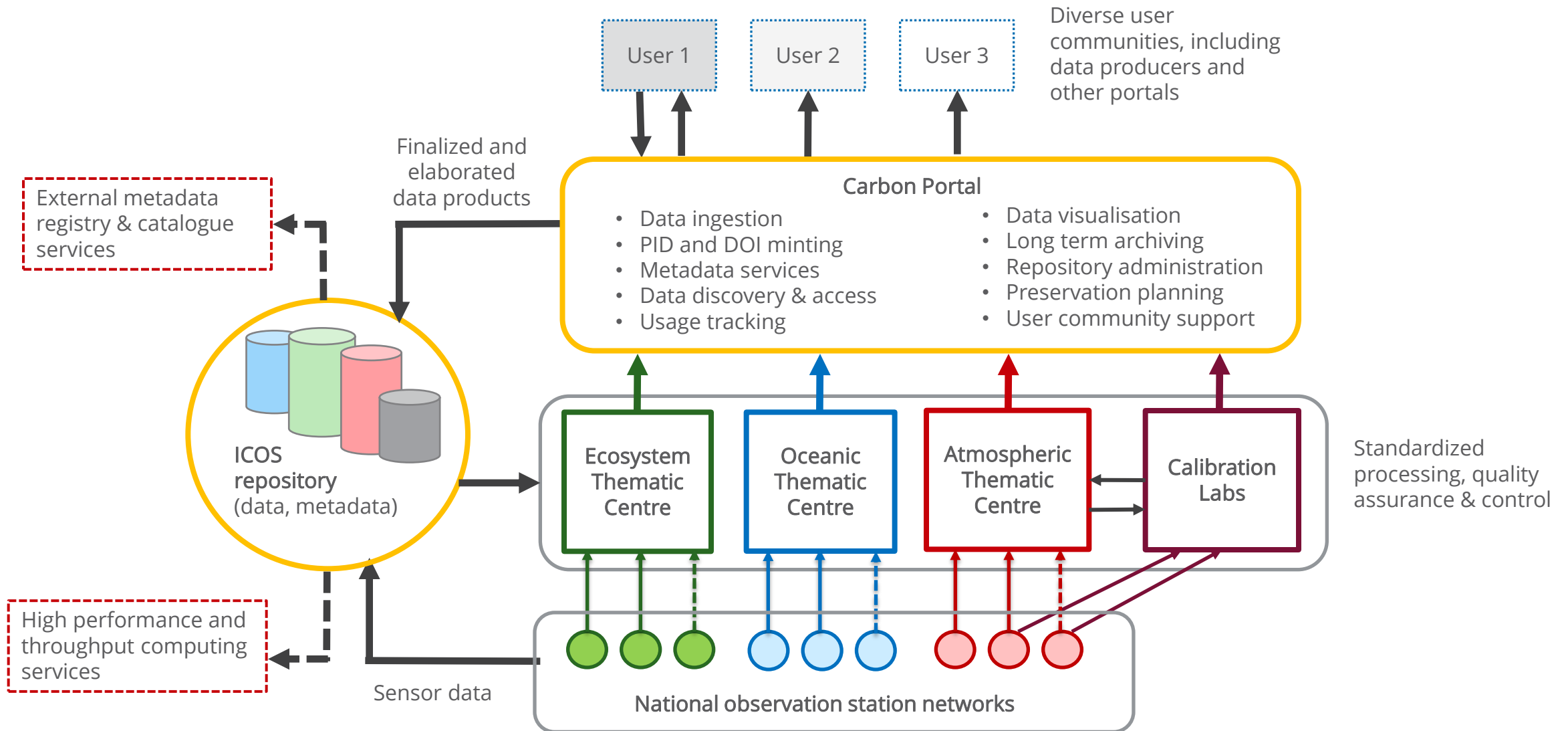


*Photographs © Konsta Punkka (2017-2018)*

# The GDO~~C~~N vision applied to ICOS...



# ICOS Data Flow





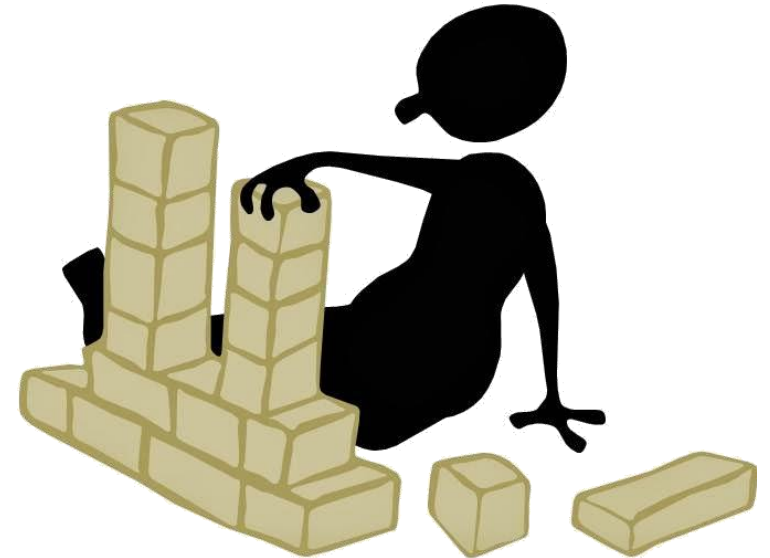
# What is needed?

To support its end users ICOS needs to provide

- FAIR data, metadata and services
- Persistent identifiers – wherever needed!
- Metadata store – fast & efficient, (mostly) open to “the world”
- Scalable services for data & metadata, to ensure functionality 24/7
- Documentation – targeting “designated” communities & others

The Carbon Portal is ICOS data center

<https://www.icos-cp.eu>



# Persistent identifiers

- ICOS uses Handle system-based PIDs
  - ePIC for raw & intermediate data, DataCite DOI for finalized data
- ICOS suffix is checksum of data (=built-in fingerprint)
- PIDs resolve to landing page URLs
- Landing pages support content negotiation, providing a means for machine-interpretability (if you know the code)
- Citation recommendations based around PIDs/DOIs
- Usage statistics can be inferred from bibliometrics



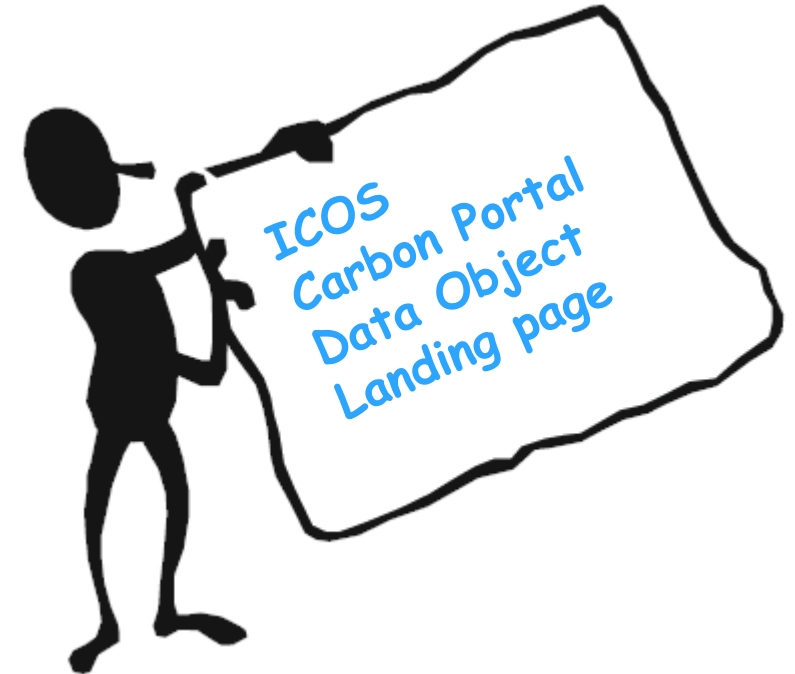
# Open (versioned) linked data store

- Semantic web (WEB 3.0), open linked data, the web is the database, everything is a URL
- Ontology adapted for ICO needs
- Stored in a non-SQL, RDF database
- Versioned meta data store, allowing roll-back & time dependent queries
- All data and metadata accessible through standard http(s), no drivers required (via SPARQL queries)
- Any portal or portal of portals can link to ICOS (meta)data and vice versa!



# Meta service

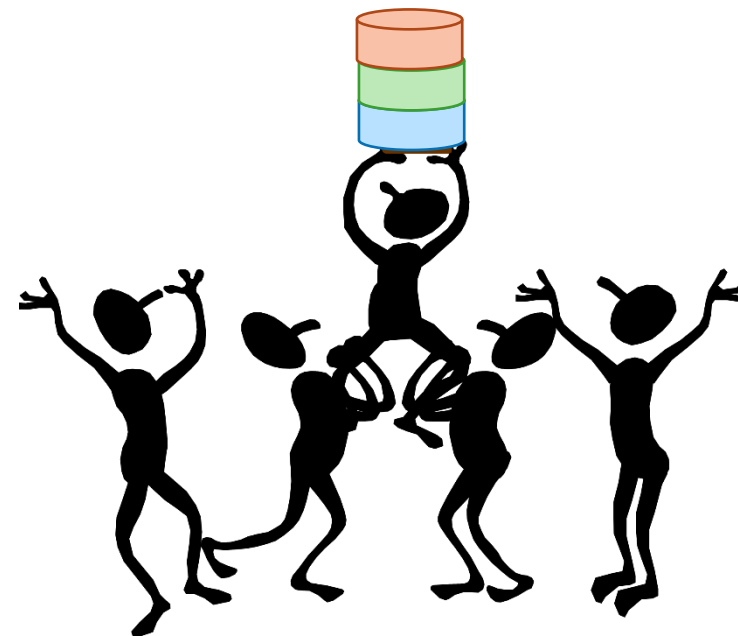
- Generates dynamic landing pages for all defined concepts, showing linkages
- Dynamic landing pages for all PIDs and DOIs, supporting content negotiation
- Ontology informs further on
  - Data Level, data format, license, DOI (if minted), ...
- User interfaces for editing of the ontology or metadata
- SPARQL endpoint for both humans and machines





# Data services

- Discovery of ICOS data products, supporting FAIR
- Access of data object link triggers licence check, usage count, https download
- Data links can be harvested and linked transparently into other portals
- Fully interactive search frontend (REST)
- Data cart (in user profile)
- Preview interactive charts/maps (REST)
- Supports versions & collections (subsetting planned)
- <https://data.icos-cp.eu>



# Some conclusions

- All data objects get Handle-based PIDs
- Metadata store is semantic web-based
- Data & metadata services are RESTful
- ICOS is based around “data as DO”
- Positive to the DO network approach
- Willing to participate as use case
  - HTC/Cloud computing aspects ??
  - LOD vs DTR ??

