FDOs for the Citation of Digital Artefacts

Peter Wittenburg

FAIR **DIGITAL OBJECTS** FORUM

Collection Reference in 2006/7 using DOBES material

Language Resources – since 2000 - https://dobes.mpi.nl/

- in 2000 the need for responsibility, accountability and persistence due to cultural heritage (endangered languages)
- in 2002 the well-defined IMDI schema
- in 2005 decision of MPG to take care of persistent infrastructure components
  - in 2005 a PID service was set up (became the ePIC system)
  - in 2005 the big MPG data centers got the task of long-term services
- in 2006 a collection builder ready
- **in 2007 a first dissertation with one reference to all data used**

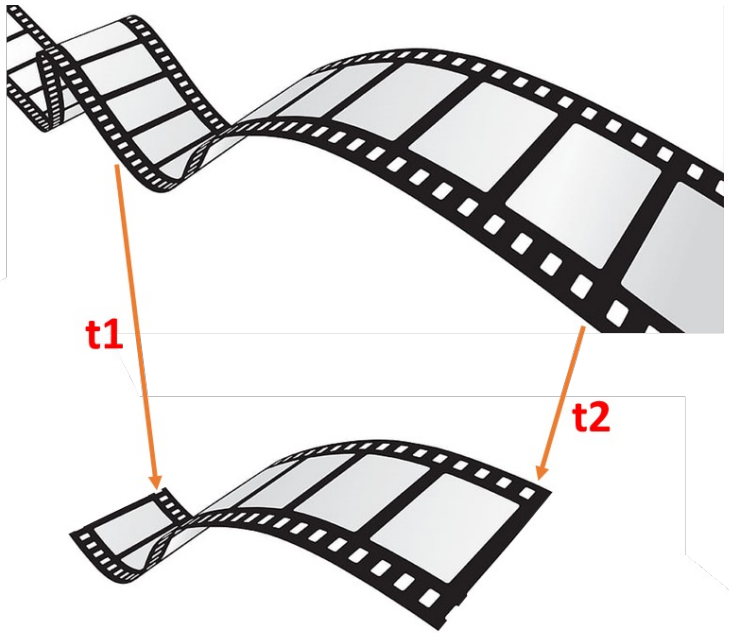
What was right / what was missing

- excellent that we had a PhD who
  - did not complain but just used the editor, created rich metadata and added a PID to all IMDI bundles (**almost FAIR in 2006/7**)
  - used the collection builder, put all digital artefacts (audio/video recordings, multiple types of annotations, word lists, etc.) into the collection and gave it a PID (Handle)

- missing at that moment
  - no agreements on how to build a collection (just a bunch of references or today better to use RO Crate or RDA recommendations)
  - no agreements how to define a PID record

Note on Granularity

Example: 2 hours recording of an event



other examples:

- slices in brainimages
- lexicon
- etc.

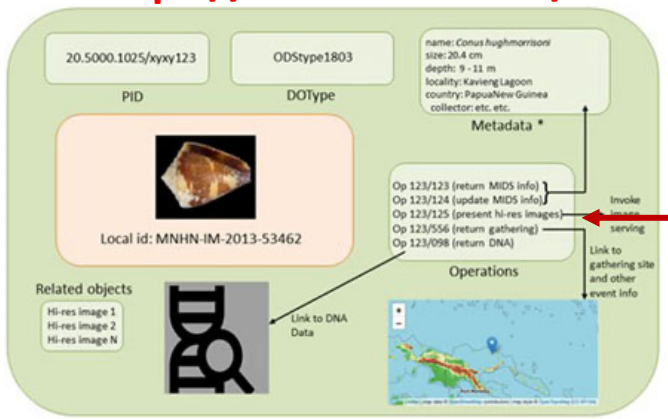
- for the creator this is the Digital Object, i.e. a PID and metadata is associated
- for some users only fragments are of interest
  - either you cut and create a new DO
  - or you use a fragment identifier **<myprefix>/<mysuffix>#<t1:t2>**
- you can combine several fragments -> collection

Whatever one does,

- **context and provenance need to be maintained**
- **policy statements need to explain practices**

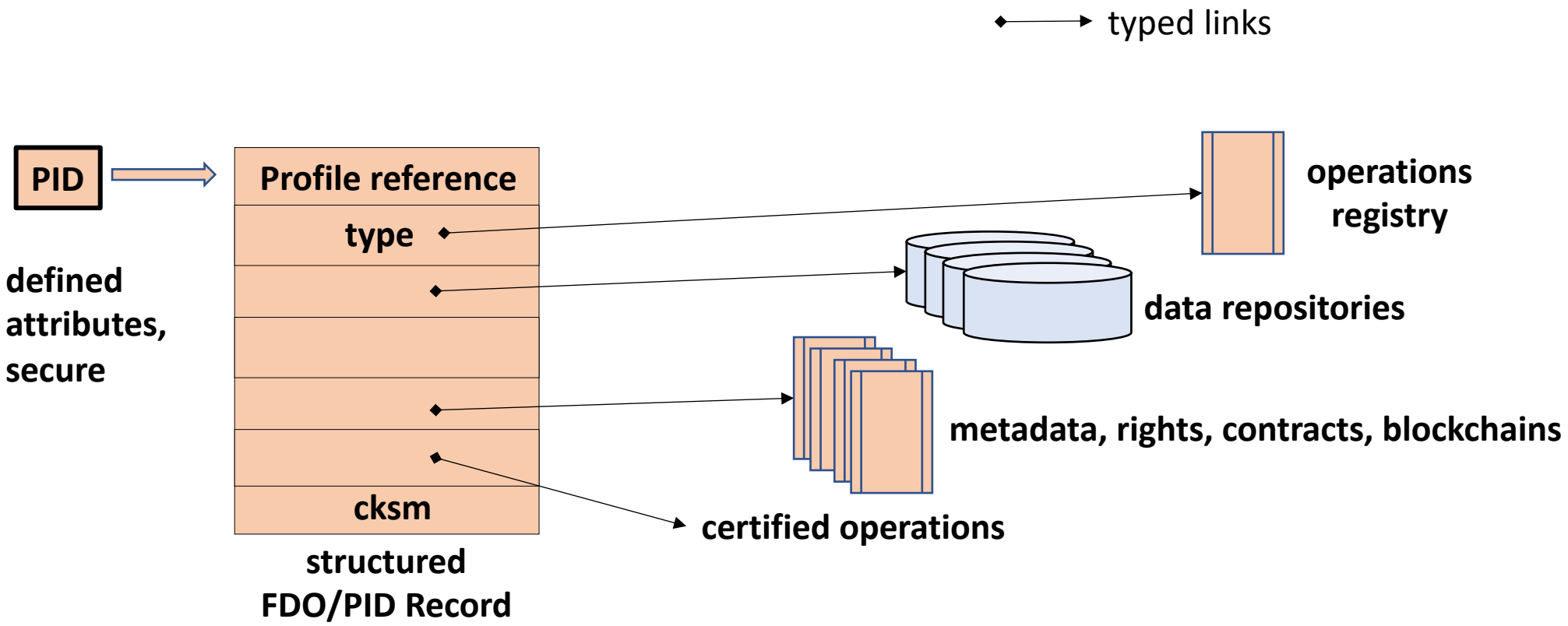
DiSSCO ESFRI – Referencing a Digital Specimen FDO

https://www.dissco.eu/



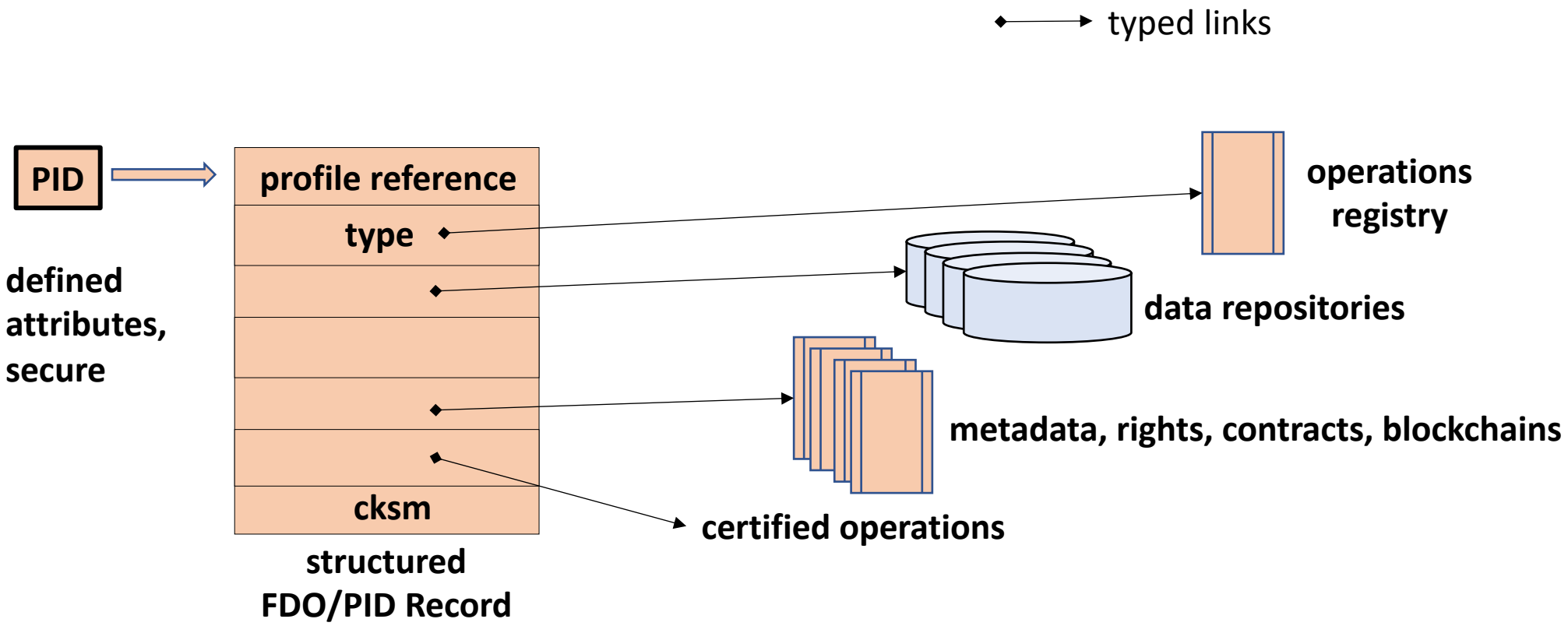
- global biodiversity-infrastructure
  - 120 natural science institutes in EU
  - 1000 collections, 3 billion objects, 2 million standards
  - **trillions of relations**
- Task: build an integrated dataspace to bundle all information to one virtual object
- **FDOs as a means to structure this space and preserve relations!**
  - **for each physical object a digital twin**
  - **all entities are FDOs – self-contained virtual entities**
  - **PID attributes contain references to tested operations**

FDO Data Model



- an FDO Profile defines the set of attributes used by a FDO/PID creator
- all (kernel) attributes in the FDO/PID record need to be defined and registered
- the first attribute needs to include the reference to the FDO Profile (also an FDO)
- all FDOs have a type according to a type registry entry

FDO Data Model



- an FDO Profile defines the set of attributes
- all (kerne
- the first a
- all FDOs have a type according to a type registry entry

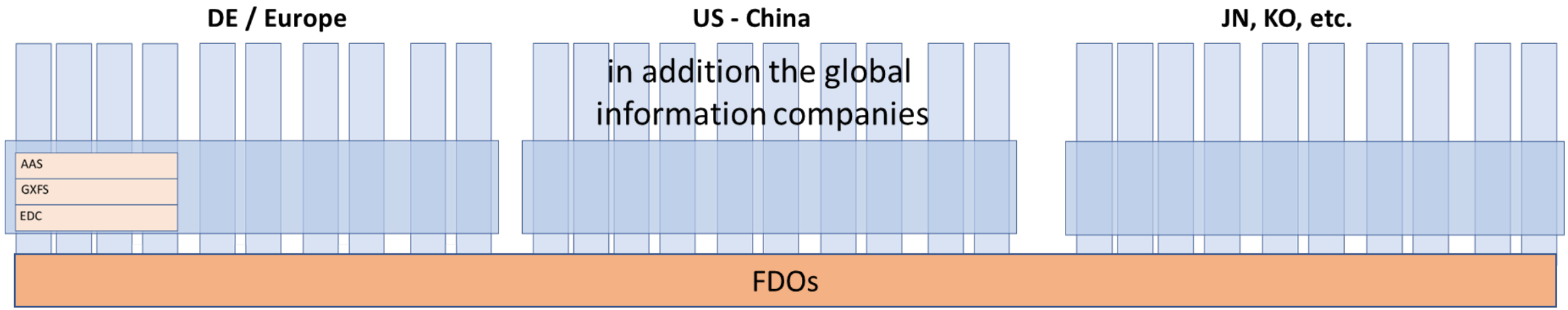
eator
nd registered
FDO Profile (also an FDO)

- to achieve predictability and machine actionability
- to enable relations to operations

What do FDOs help?

- at data management level we just have bit-sequences of a certain type
  (collections, recordings, measurements, semantic assertions, knowlets, configurations, software, etc.)
  - simplifying and systematising DM enormously (procedures a la CoreTrustSeal)
  - an implementation to maintain FAIRness

- **as autonomous, self-contained units of information they can travel through all the many emerging dataspaces independent of the technologies and rules applied**



Thanks for your attention.

FAIR **DIGITAL OBJECTS** FORUM

FAIR Digital Objects after FDO Conference

2nd Int. Conference on
FAIR Digital Objects
March 2024

- Basic -Specifications are done!
- different approaches are possible
  (DO-based, Signposting, etc.)
- starting a second phase
  - open workshops to tackle open issues
  - focus is on integrated Reference Implementation
  - **turn specs into an international Standard (DIN, NIST)**
- FDO Forum
  - to be a Legal Entity until March 2024
  - simple governance
    4 co-chairs, Exec Com. (9), Steering Committee (30), AGs

DONA Foundation - Geneva

Foundation Geneva

2014 DONA Foundation.
Guiding the work on DO Architecture
ITU Umbrella

- operation of a redundant, distributed, secure and persistent Handle/DOI root network
- Switzerland is accepted as a neutral place
- issuing prefixes for trustworthy service providers (publishers, film industry, ePIC, national centres, global initiatives, etc.)

- generic PID Interface Request Protocol – DO-IRP
- generic DO Interface Protocol – DO-IP