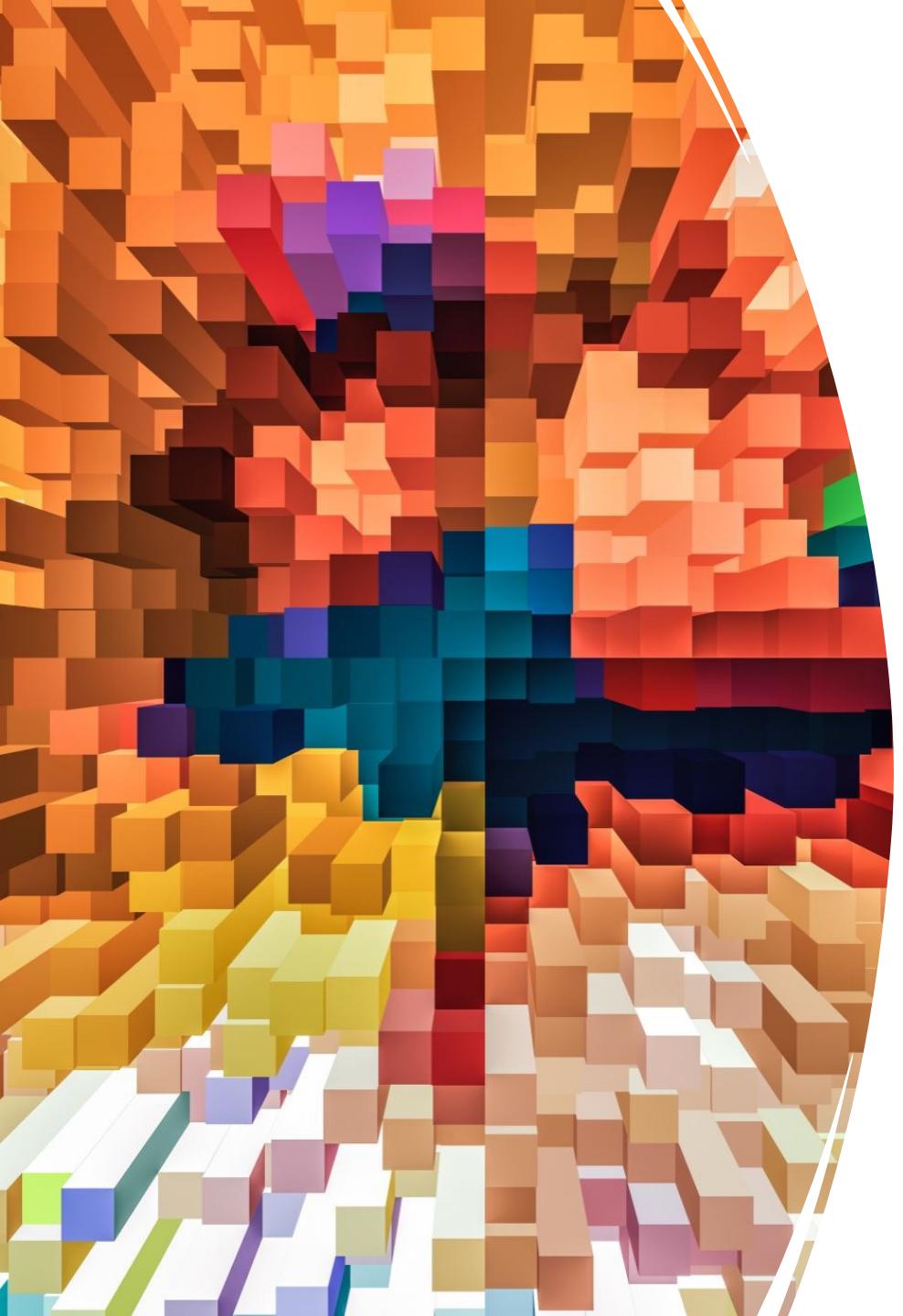




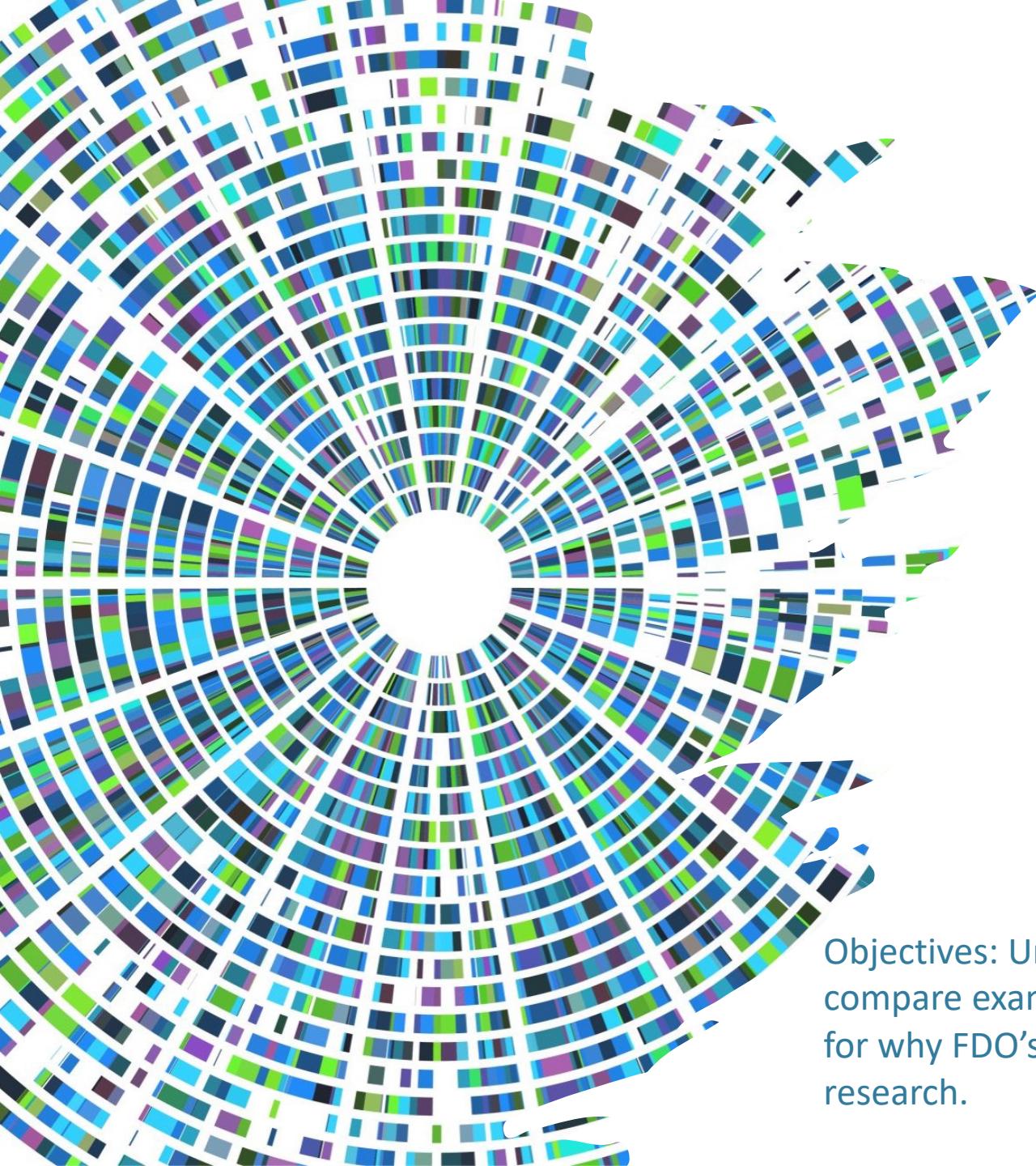
FDO - FAIR Digital Objects

Open Science Course
Friday, 17 November 2023
16:15 – 17:00



Let's start a brief discussion with neighbors:

What are some digital replacements for what were previously physical objects?



Today's FDO – FAIR Digital Objects session outline

- Brainstorming
- Historical context of **Digital Objects**
- What are **Digital Objects**?
- Making a **DO FAIR (FDO)** discussion
- An example of a hypothetical **(F)DO**
- Why **FDOs** matter to us
- Implementation of **FDOs**

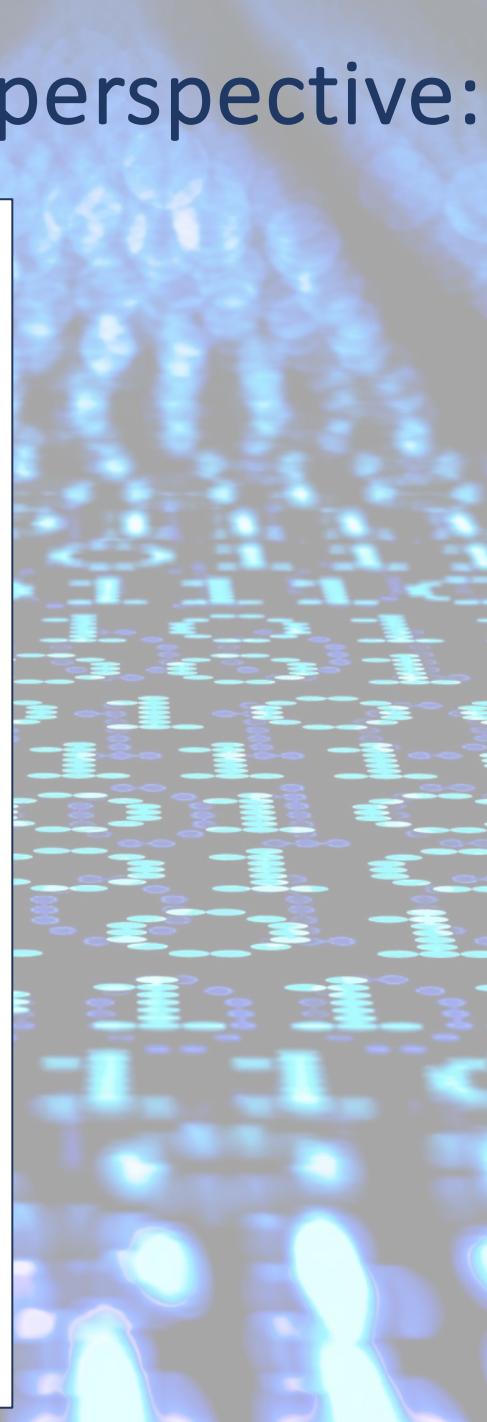
Objectives: Understand the general features of a DO & FDO, to be able to compare examples of each in the context of STEM research. Discuss reasons for why FDO's are being implemented, and how they will influence STEM research.

Gaining perspective: the historical context of Digital Objects

- Recollect computers started before the internet



- The world wide web posed new challenges for keeping order and communication across geopolitical boundaries
- Bob Kahn, one of the fathers of the internet, saw the need for an architectural framework (support) for the web
- Revolutionary for the "data landscape" of the web

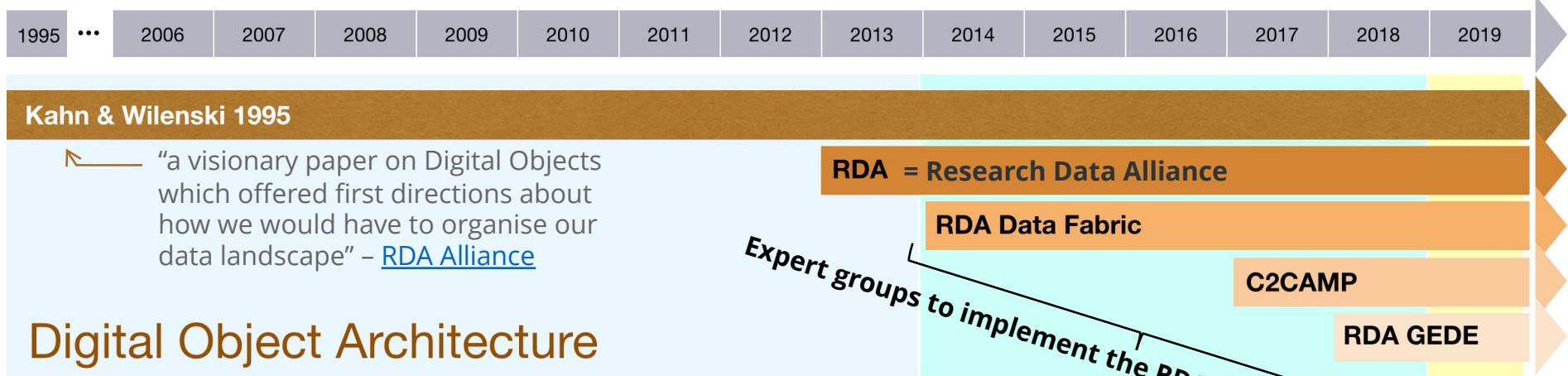


Gaining perspective: the historical context of Digital Objects

- Recollect computers started before the internet



- The world wide web posed new challenges for keeping order and communication across geopolitical boundaries
- Bob Kahn, one of the fathers of the internet, saw the need for an architectural framework (support) for the web
- Revolutionary for the "data landscape" of the web



Digital Object Architecture

→ basic principles of the internet

Timeline of developments in the convergence of various approaches to data infrastructures: conceptualization (blue background), design (green background), and implementation (yellow background)

Gaining perspective: the historical context of Digital Objects

- Recollect computers started before the internet



- The world wide web posed new challenges for keeping order and communication across geopolitical boundaries
- Bob Kahn, one of the fathers of the internet, saw the need for an architectural framework (support) for the web
- Revolutionary for the "data landscape" of the web



Kahn & Wilenski 1995

→ "a visionary paper on Digital Objects which offered first directions about how we would have to organise our data landscape" – [RDA Alliance](#)

Digital Object Architecture

→ basic principles of the internet

RDA = Research Data Alliance

RDA Data Fabric

C2CAMP

RDA GEDE

Expert groups to implement the RDA objectives / procedures

RDA - What can we learn from the Internet?

Home > Webinar > RDA - What can we learn from the Internet?

08 November 2016 | 4234 reads | Facebook | Twitter

Type:

Webinar, Digital Objects, PIDs, Digital Object Identifiers

Date/Time:

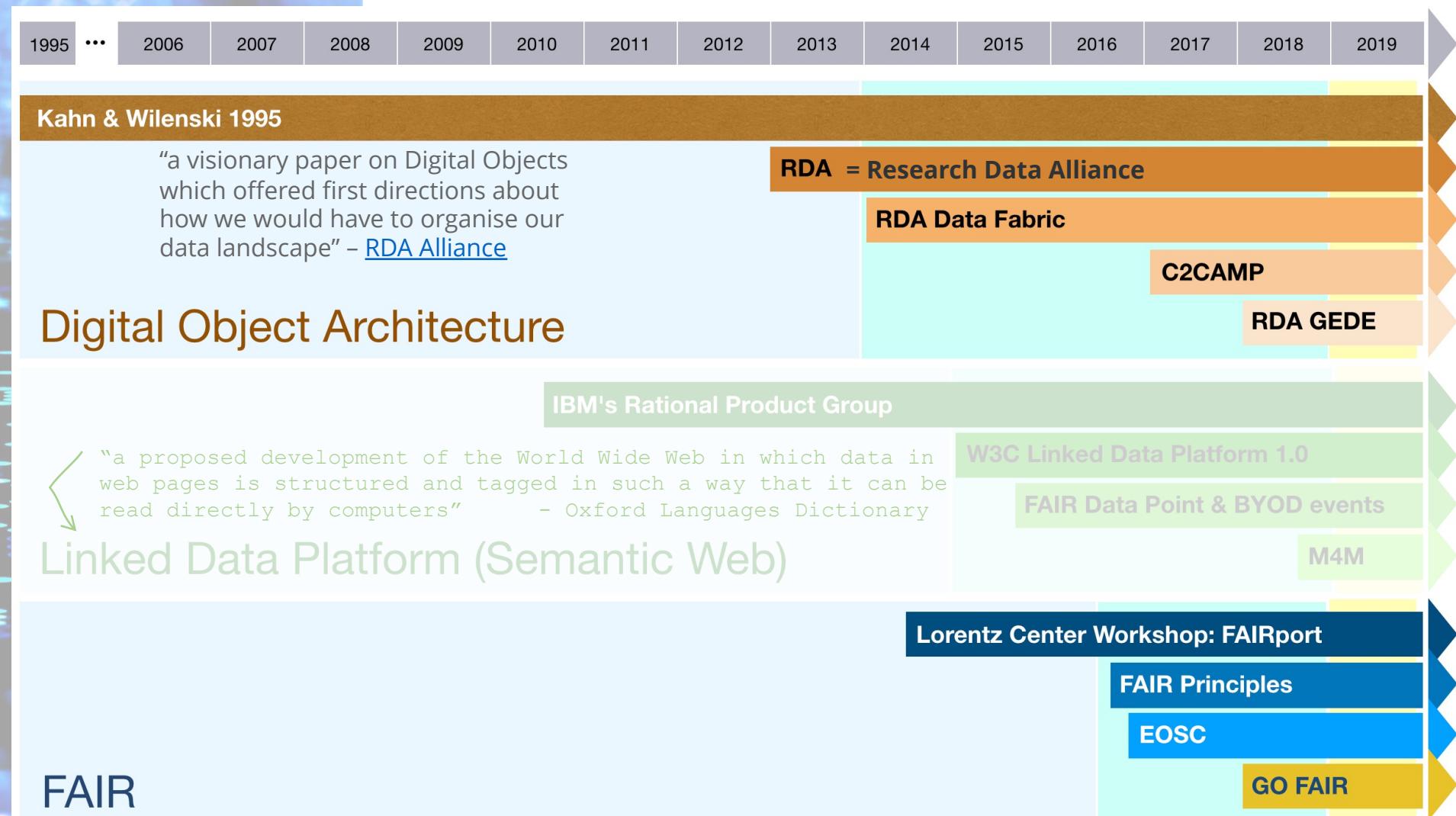
06 December 2016 - 16:00 CET, 15:00 UTC



(If interested, [a 2016 interview with Bob Kahn](#); ~1 hour 11 min)

Timeline of developments in the convergence of various approaches to data infrastructures: conceptualization (blue background), design (green background), and implementation (yellow background)

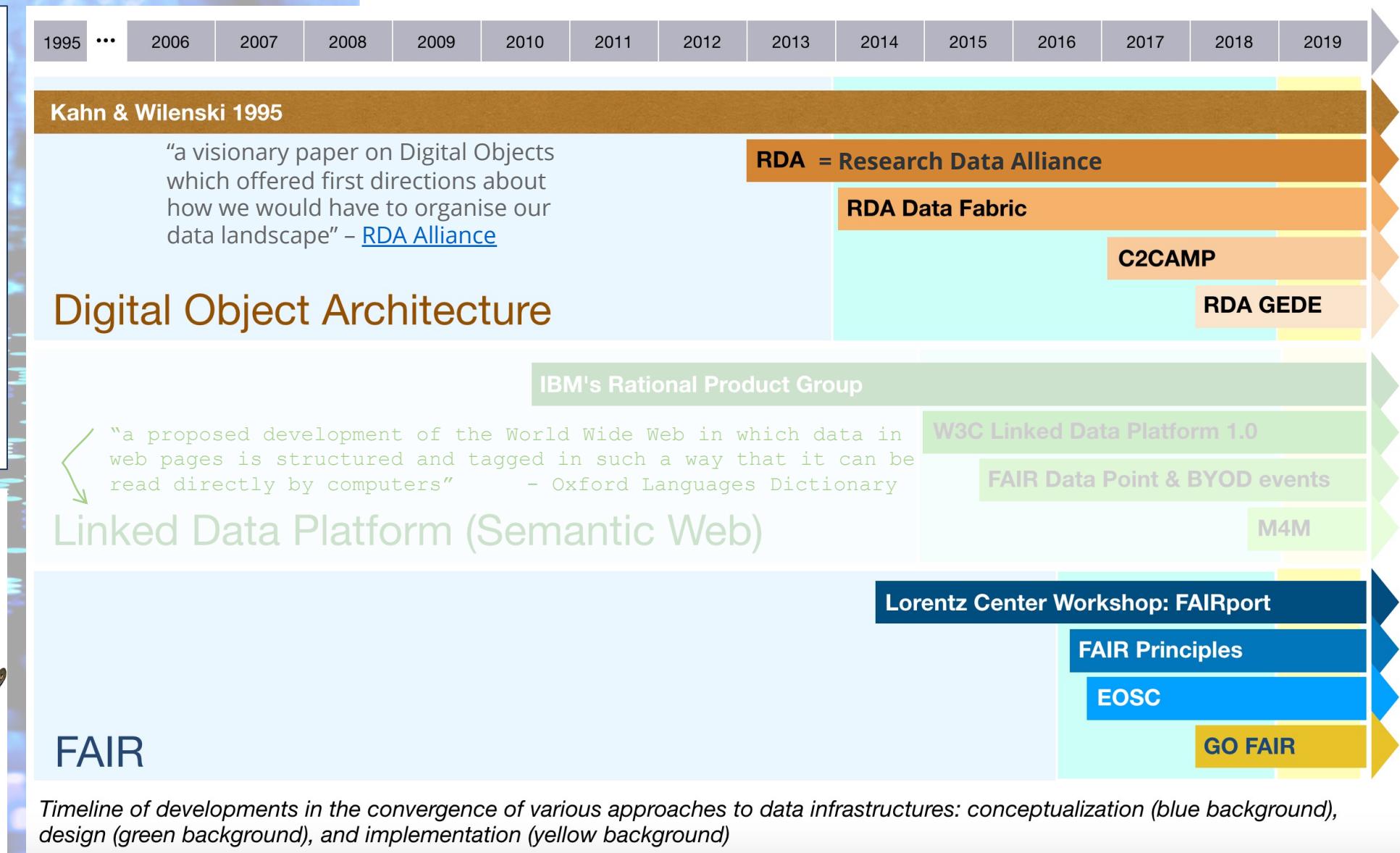
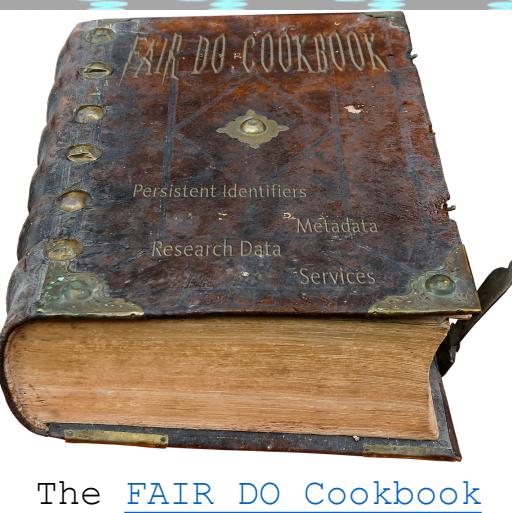
Gaining perspective: the historical context of Digital Objects



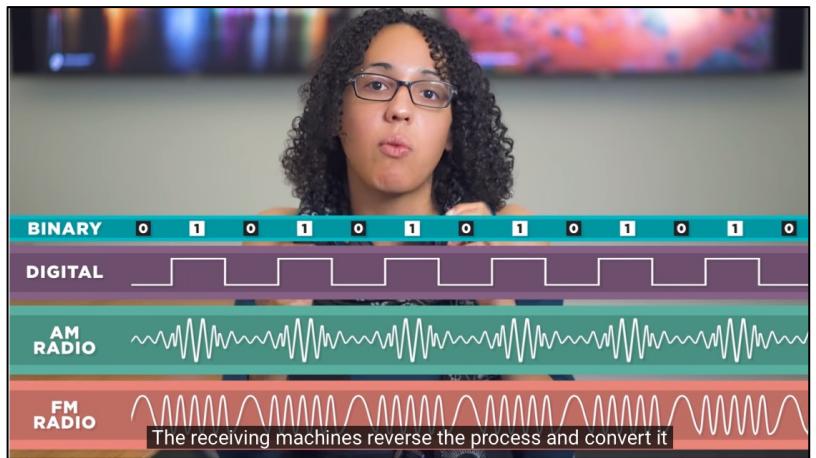
Timeline of developments in the convergence of various approaches to data infrastructures: conceptualization (blue background), design (green background), and implementation (yellow background)

Gaining perspective: the historical context of Digital Objects

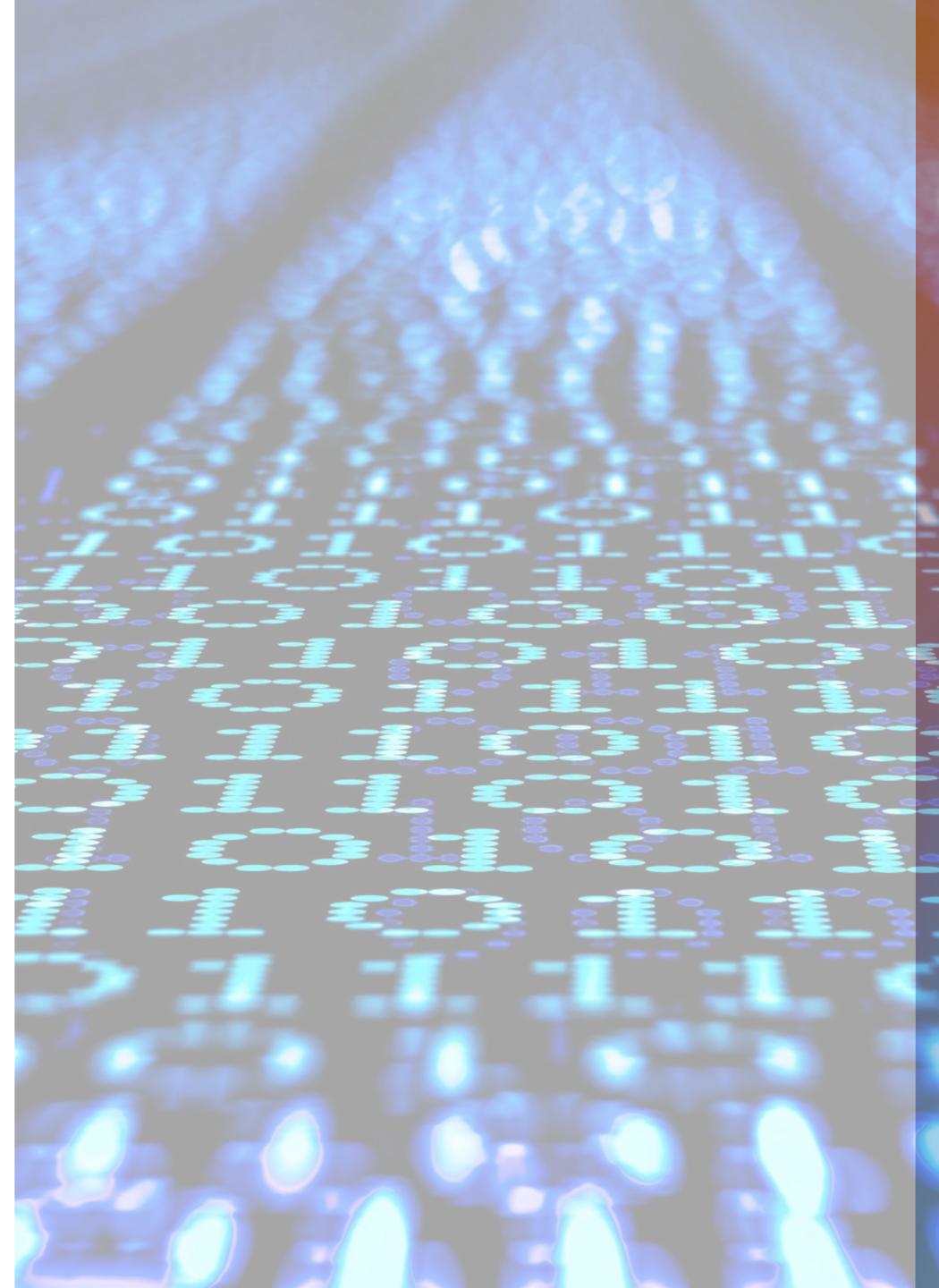
- More recently being incorporated with FAIR principles = FAIR Digital Objects (FDOS)
- To be FAIR, all Digital Objects (data, software, knowledge, etc.) need Persistent Identifiers (PIDs), e.g., Digital Object Identifiers ([DOIs](#))



What are Digital Objects?

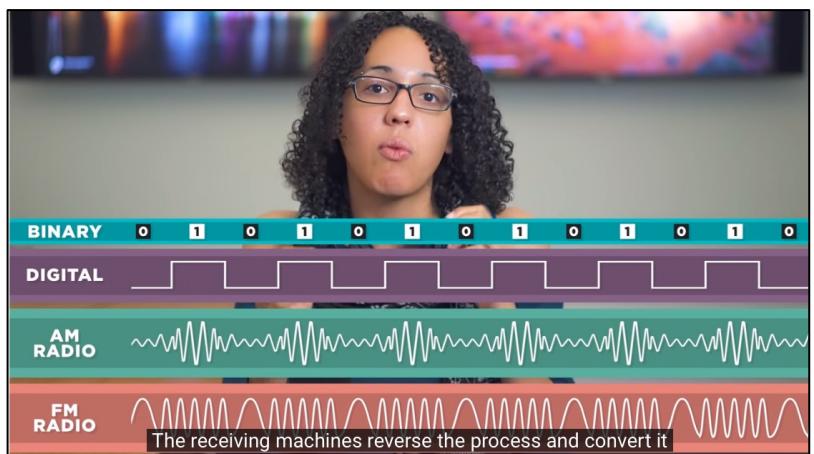


Bit sequences are what forms information transfer, computation & data storage

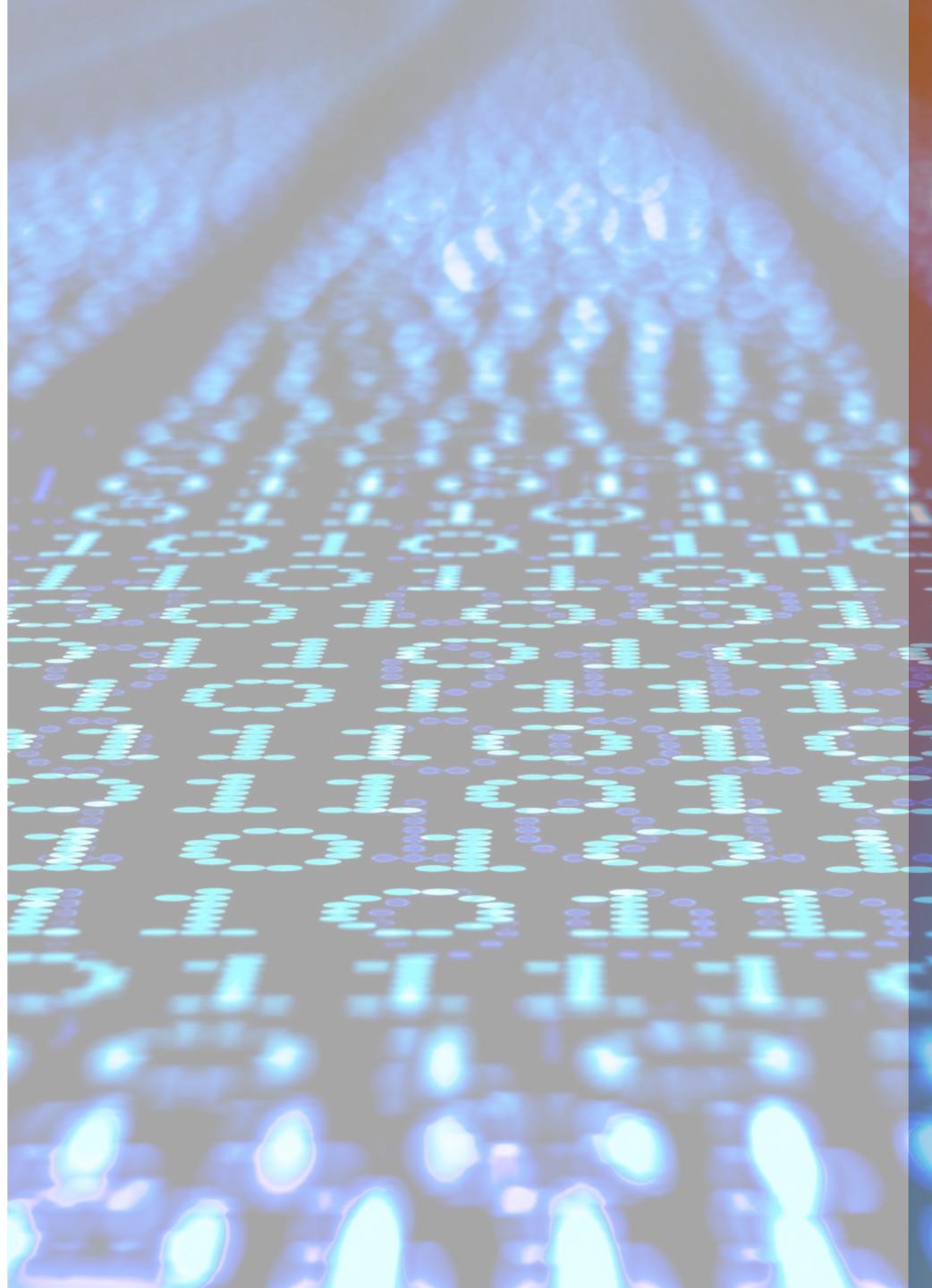


What are Digital Objects?

- '[An object composed of a set of bit sequences](#)' (CCSDS, 2012)
 - And the associated metadata
- A digital file can be considered a **digital object**
- A **DO** can (but need not) represent something physical (e.g., a digital object of a physical museum specimen)
- Two types:
 - *Simple digital objects* have limited elements: a text, pdf or image file
 - *Complex digital objects* contain multiple elements: like video (video track, audio track, container file & possibly others), websites, and digital books



Bit sequences are what forms information transfer, computation & data storage



There are many Digital Objects!

Digitized text

- Books, newspapers, journals ...



Digitized images

- Photos, drawings, posters ...
- Sheet music
- Maps



Sound and moving images

- Recordings, oral histories
- Film



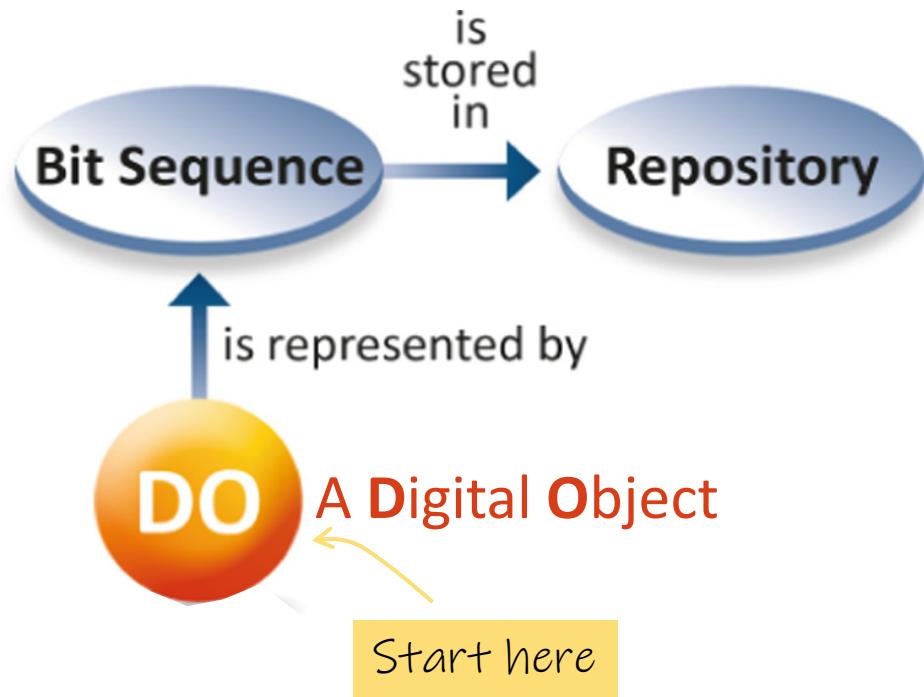
Born digital material

Scholarly preprints, data sets, dissertations
Archived Web sites

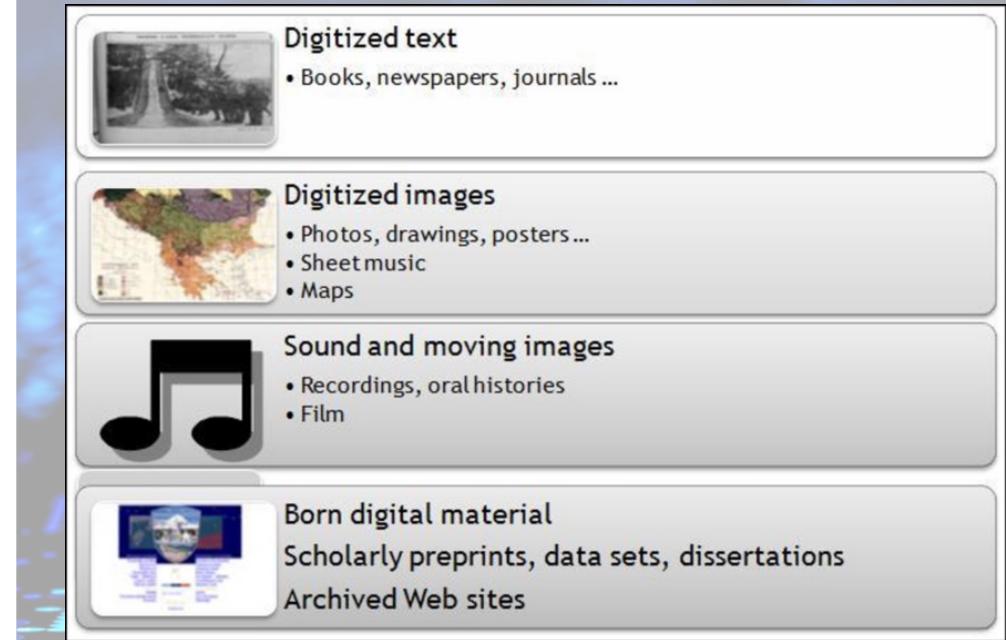


Types of Digital Objects. From: Calhoun, K., 2013. Supporting digital scholarship: Bibliographic control, library cooperatives and open access repositories. *Catalogue*, 2, pp.143-178.

There are many Digital Objects!

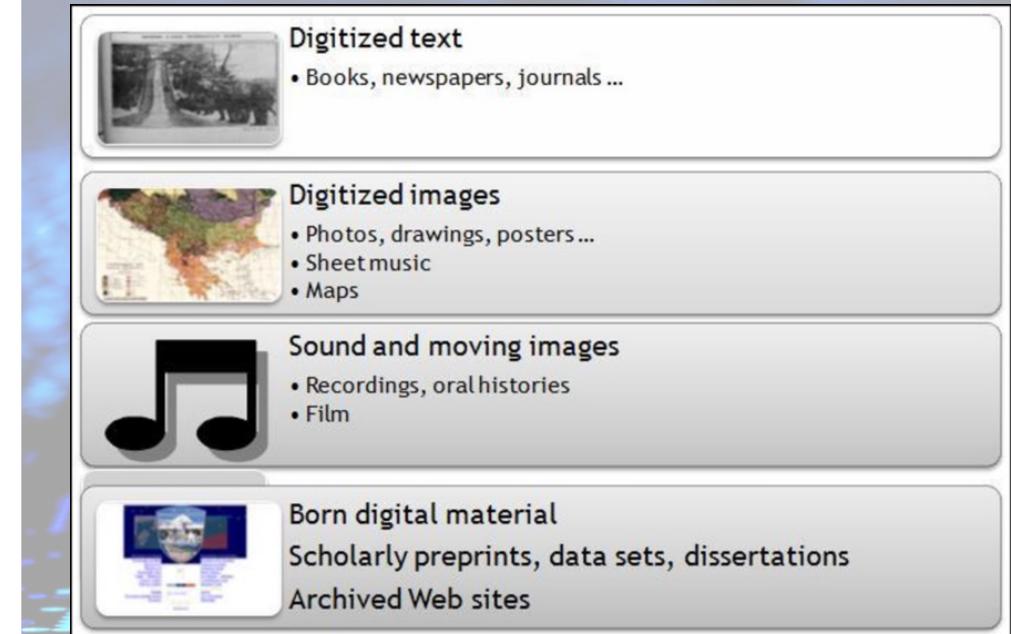
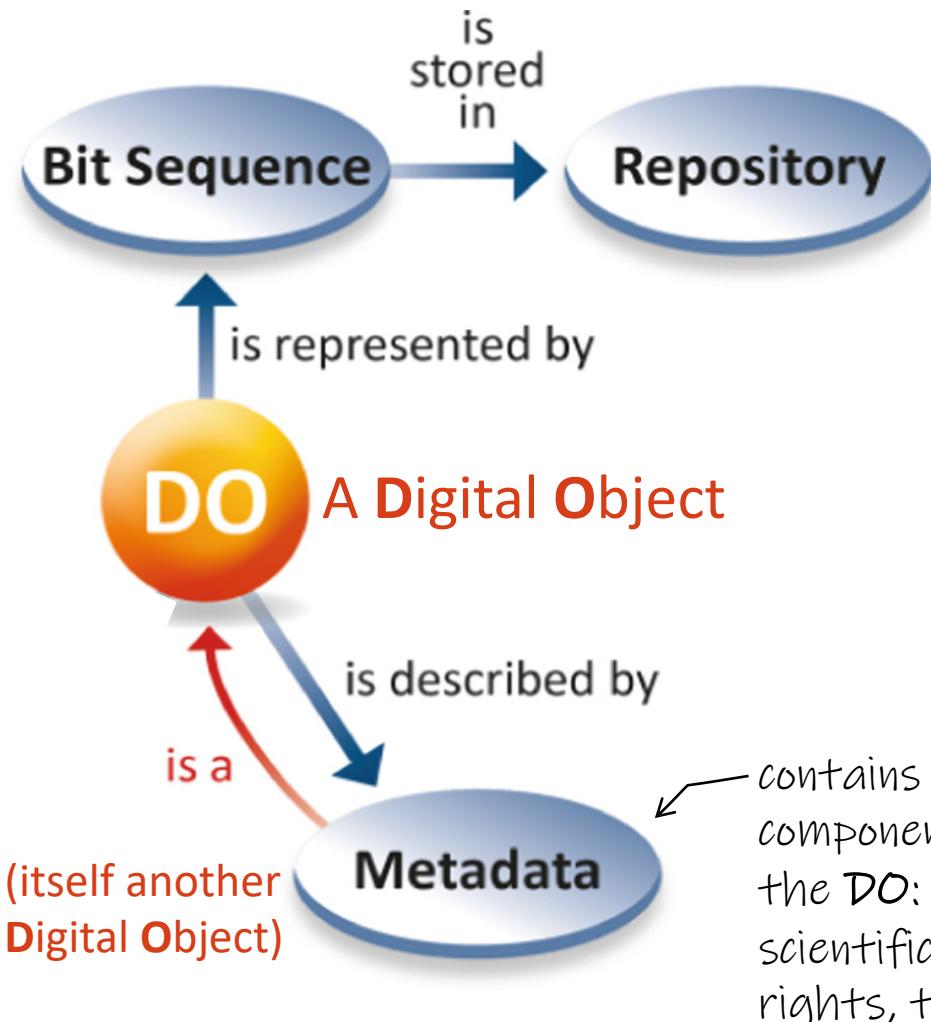


Start here



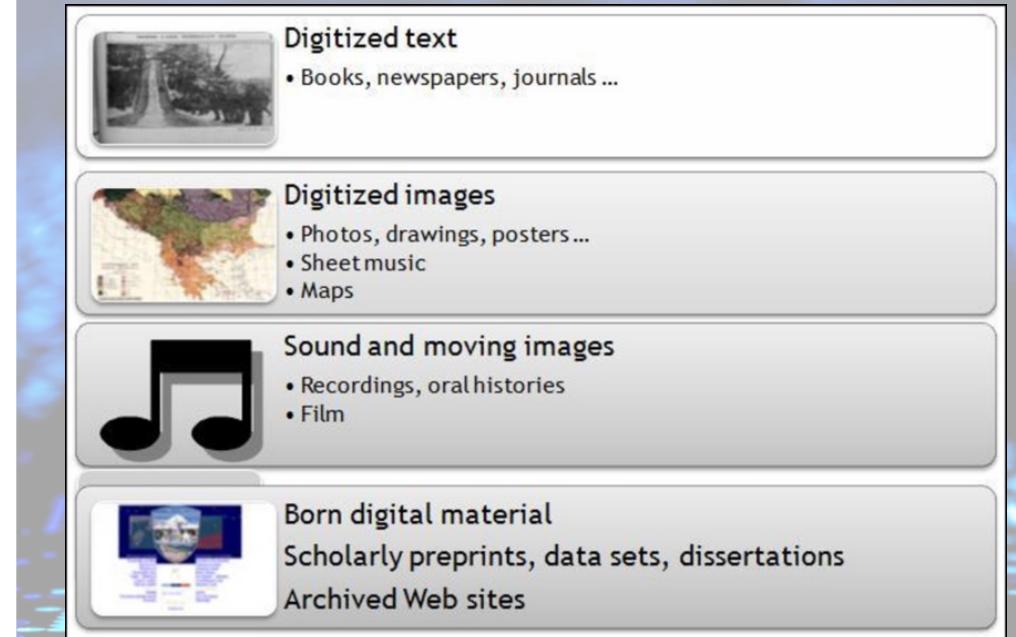
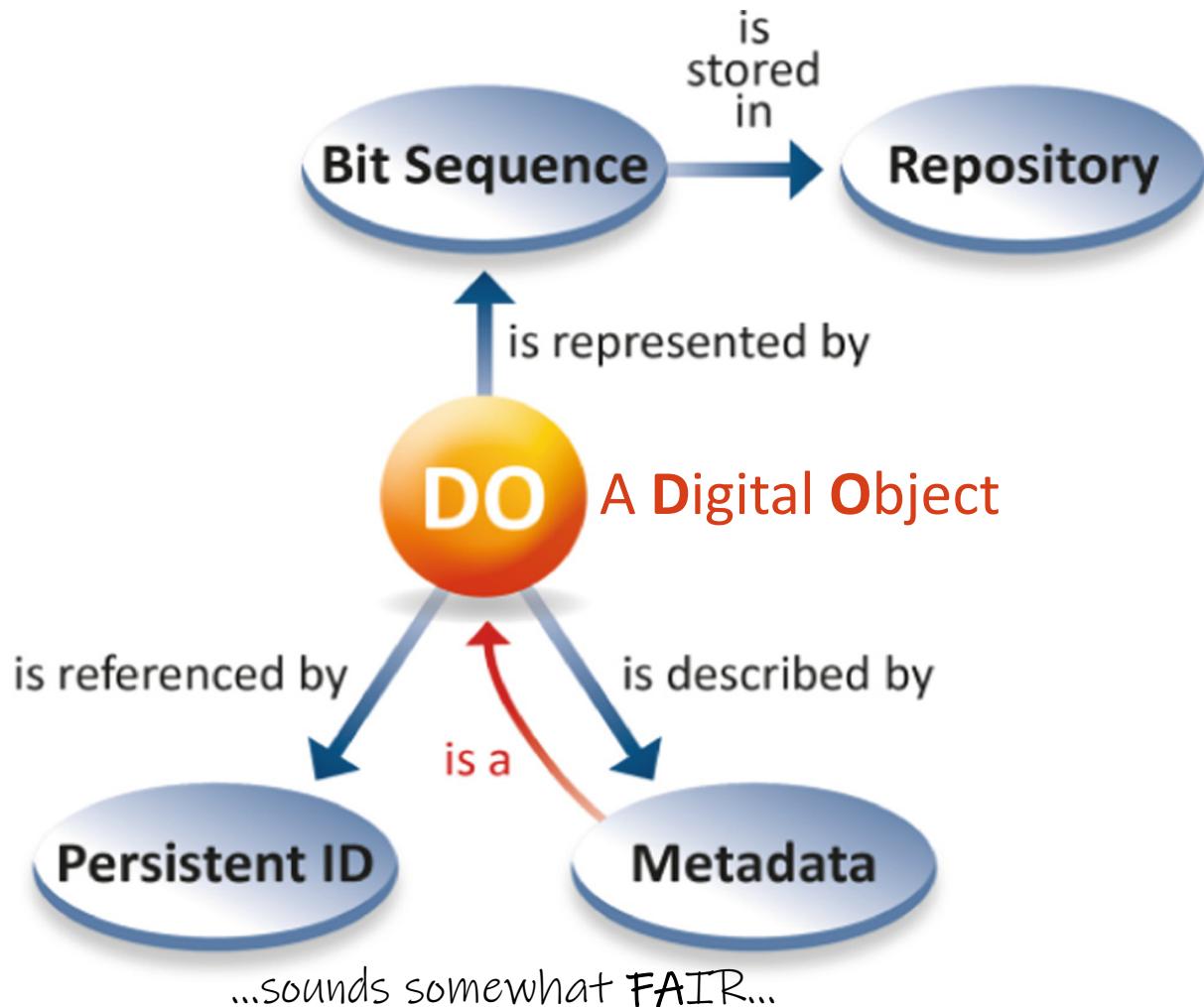
Types of Digital Objects. From: Calhoun, K., 2013. Supporting digital scholarship: Bibliographic control, library cooperatives and open access repositories. *Catalogue*, 2, pp.143-178.

There are many Digital Objects!

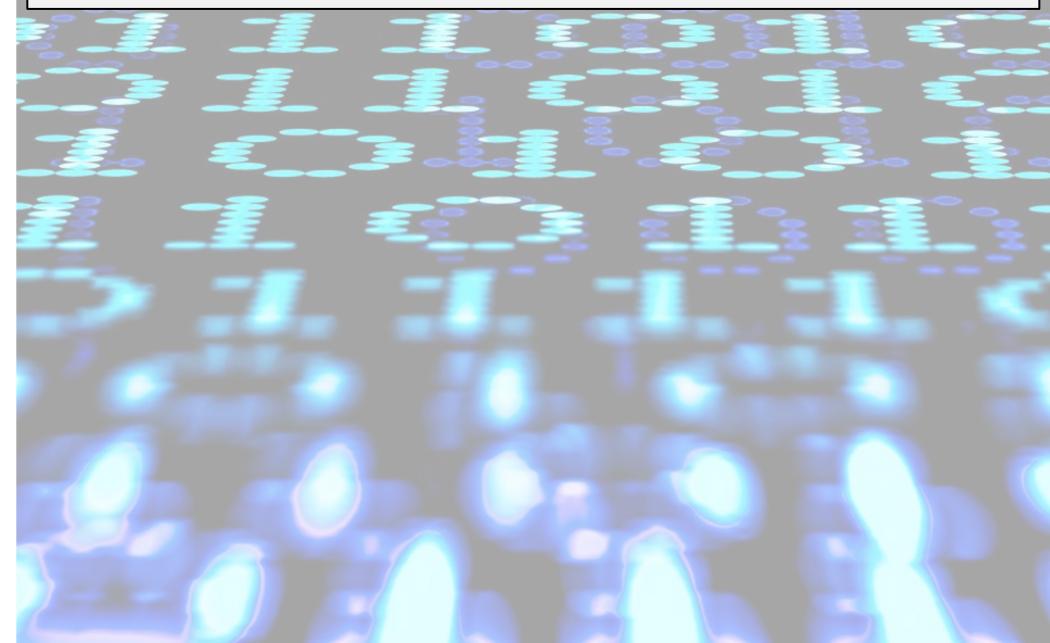


Types of Digital Objects. From: Calhoun, K., 2013. Supporting digital scholarship: Bibliographic control, library cooperatives and open access repositories. *Catalogue*, 2, pp.143-178.

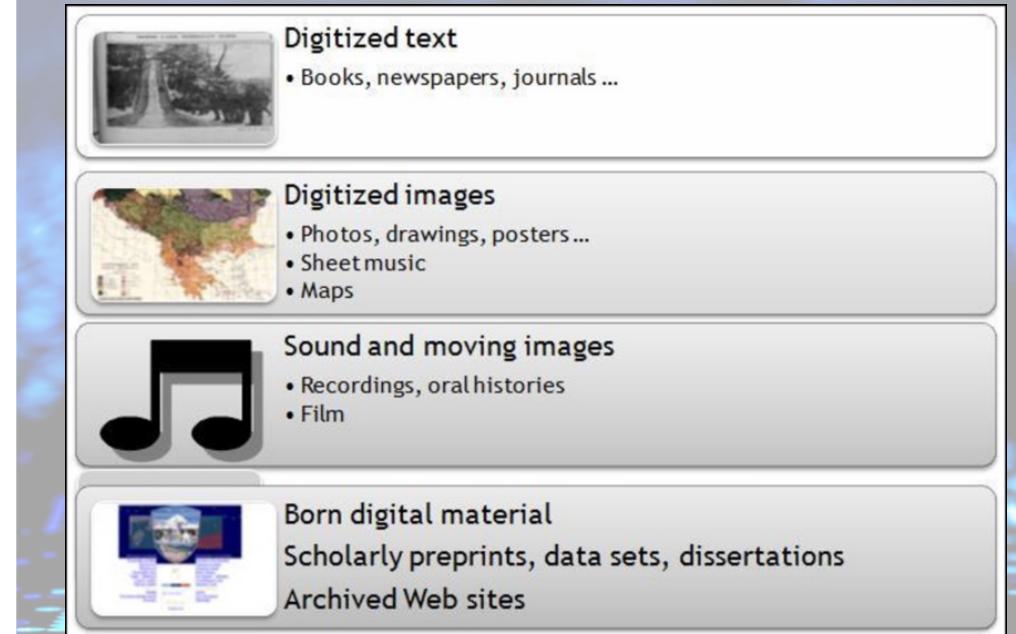
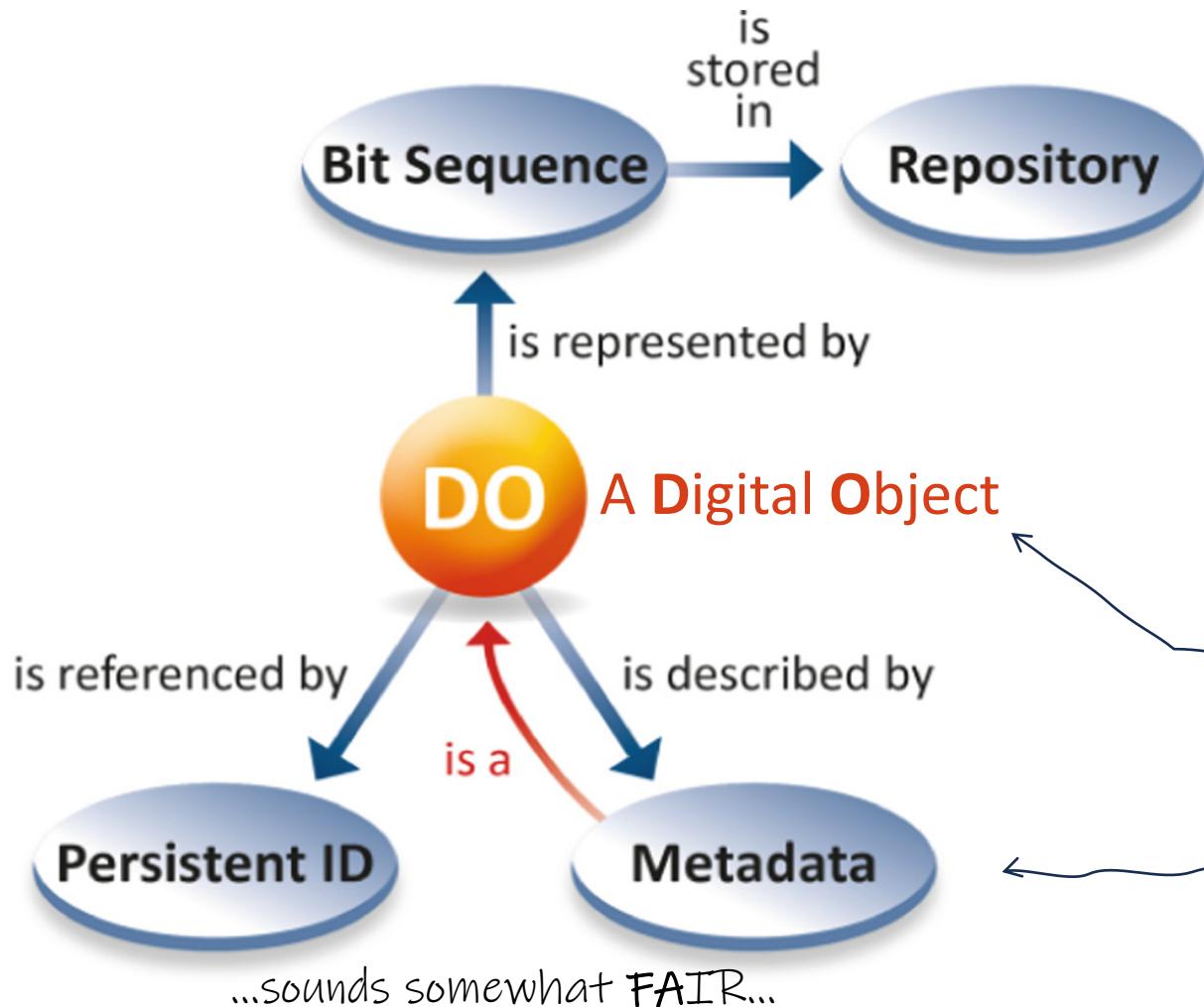
There are many Digital Objects!



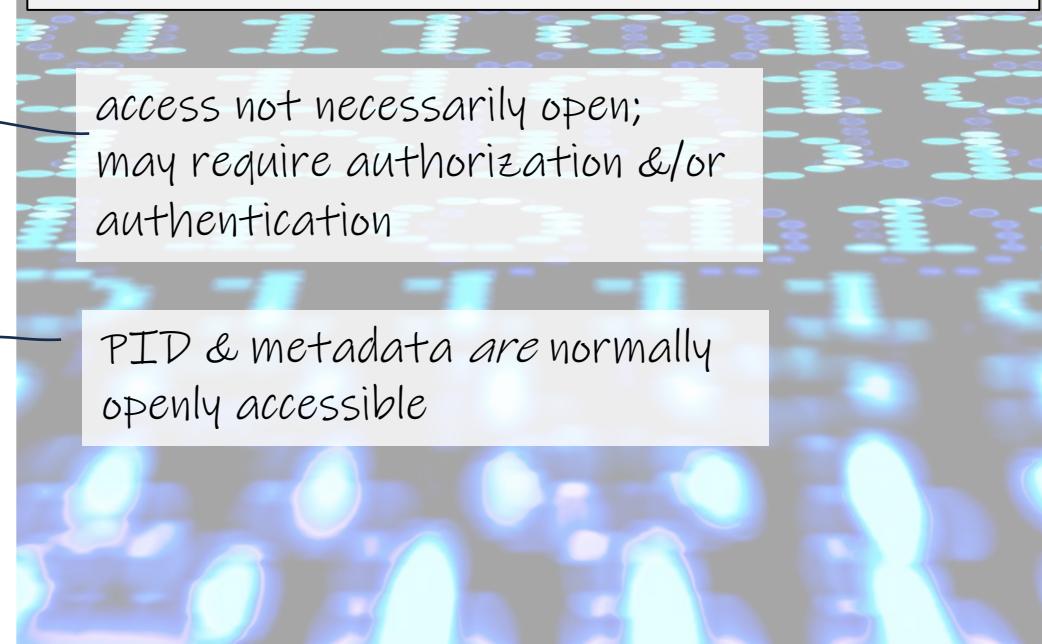
Types of Digital Objects. From: Calhoun, K., 2013. Supporting digital scholarship: Bibliographic control, library cooperatives and open access repositories. *Catalogue*, 2, pp.143-178.



There are many Digital Objects!

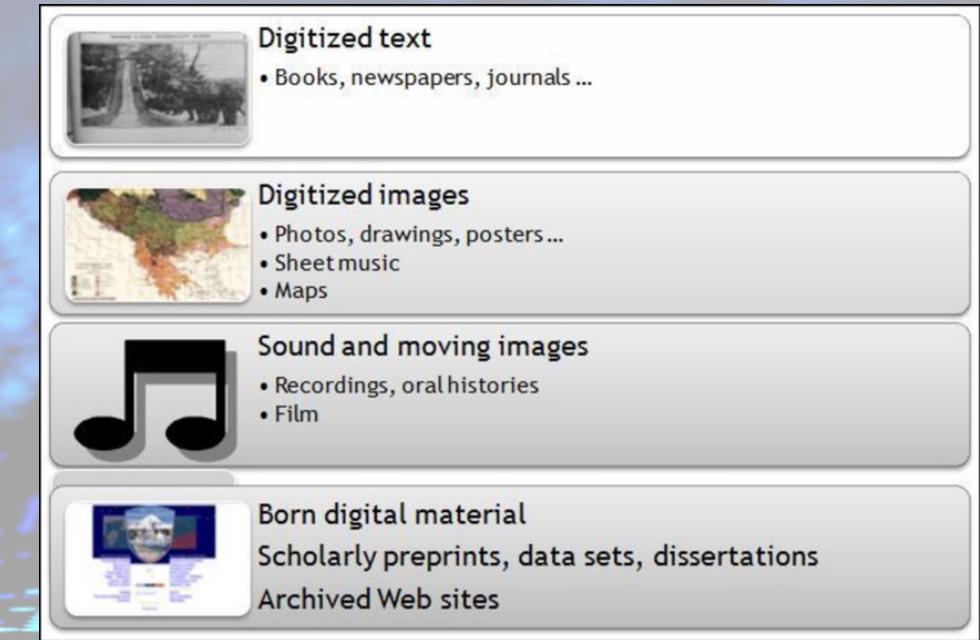
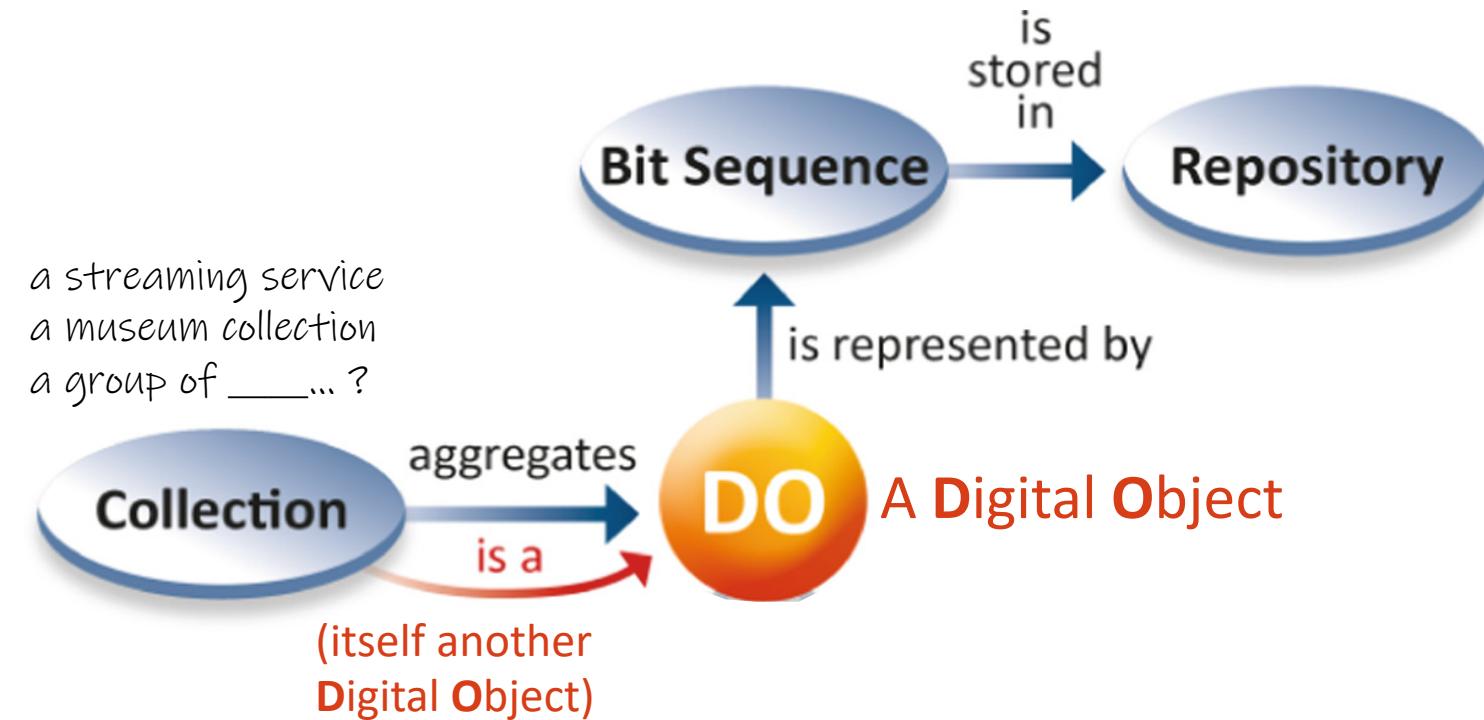


Types of Digital Objects. From: Calhoun, K., 2013. Supporting digital scholarship: Bibliographic control, library cooperatives and open access repositories. *Catalogue*, 2, pp.143-178.

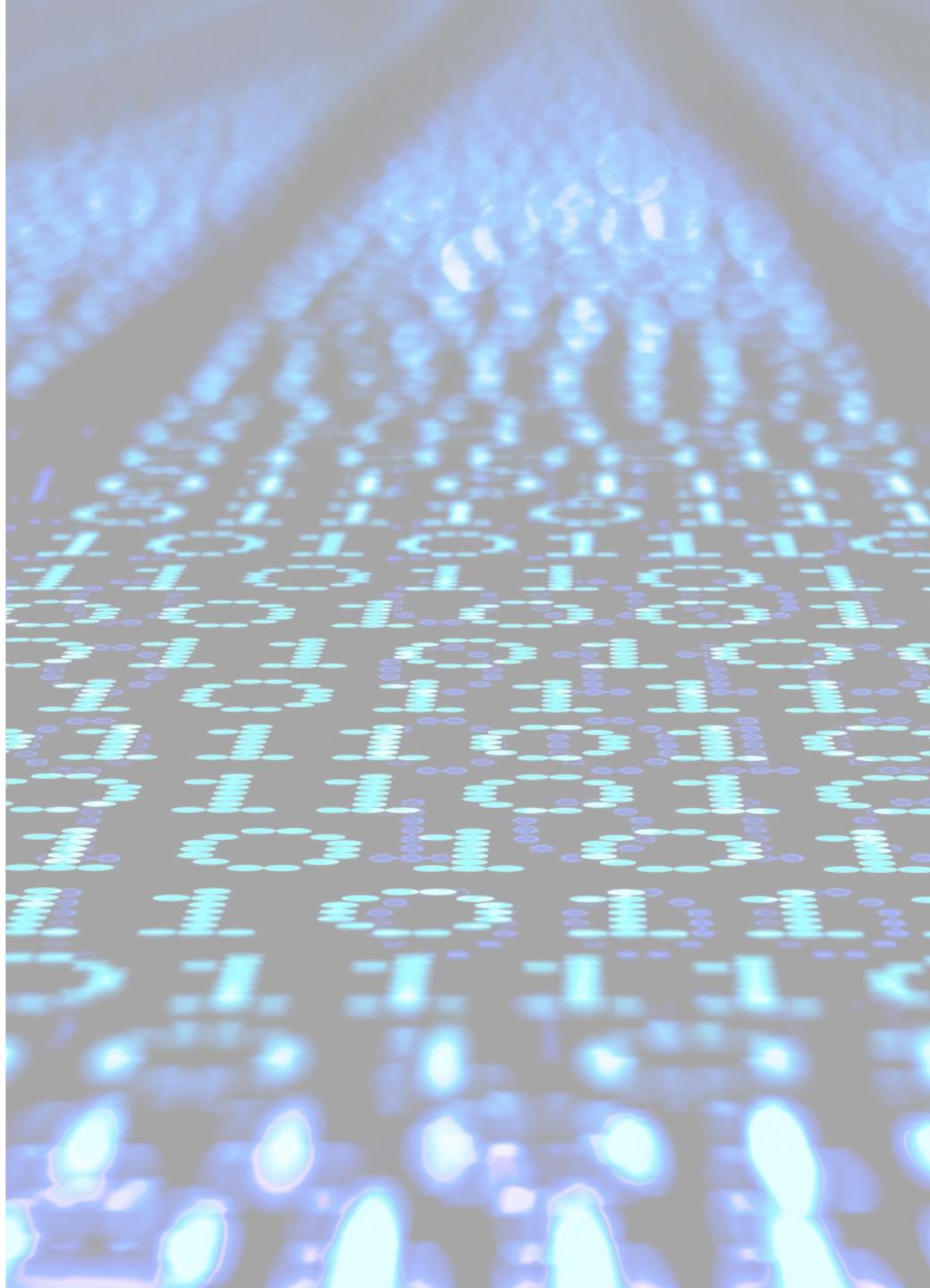
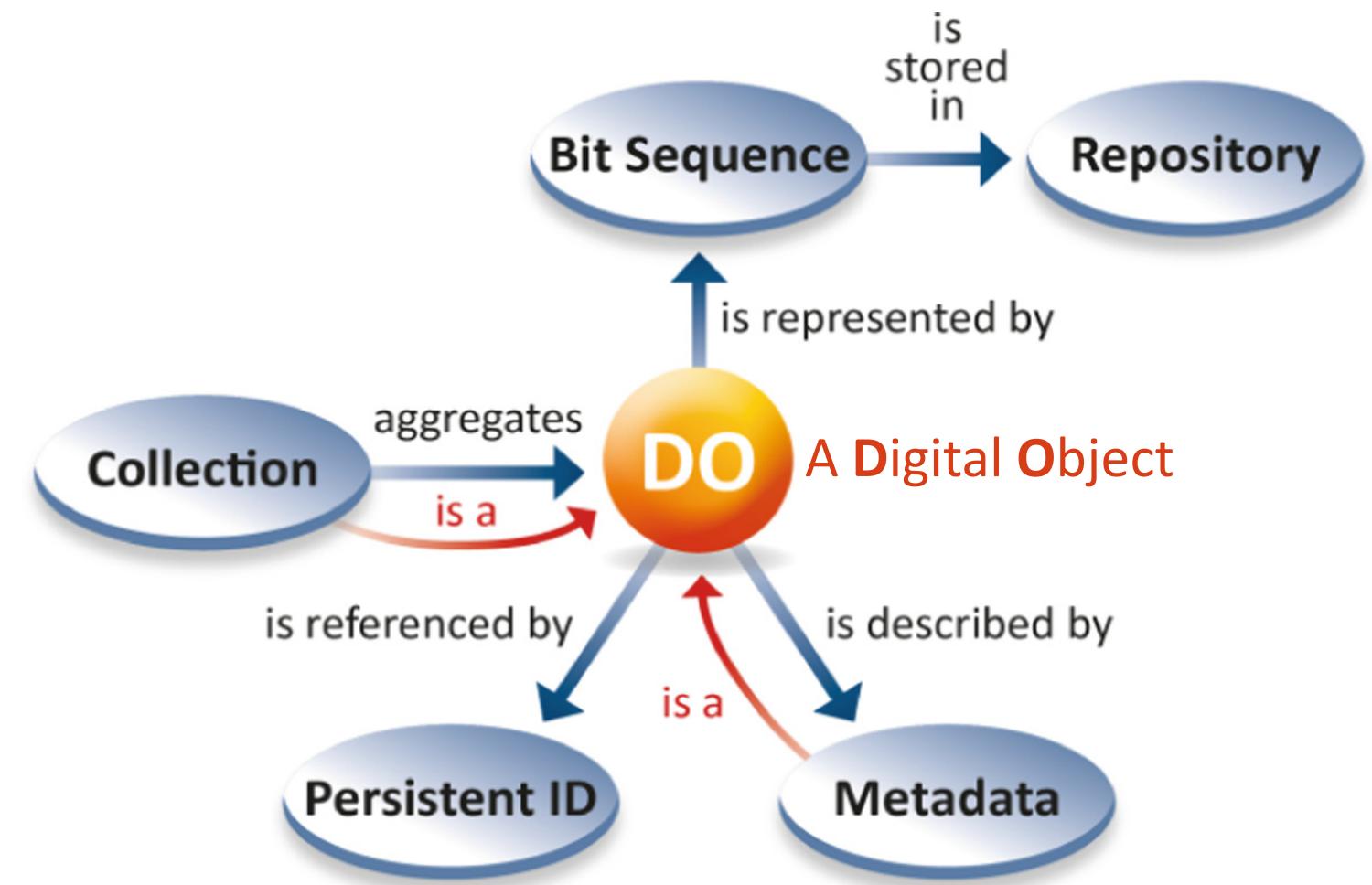


What are Digital Objects?

a streaming service
a museum collection
a group of _____ ?



Types of Digital Objects. From: Calhoun, K., 2013. Supporting digital scholarship: Bibliographic control, library cooperatives and open access repositories. *Catalogue*, 2, pp.143-178.



OK, but is this a FAIR Digital Object?

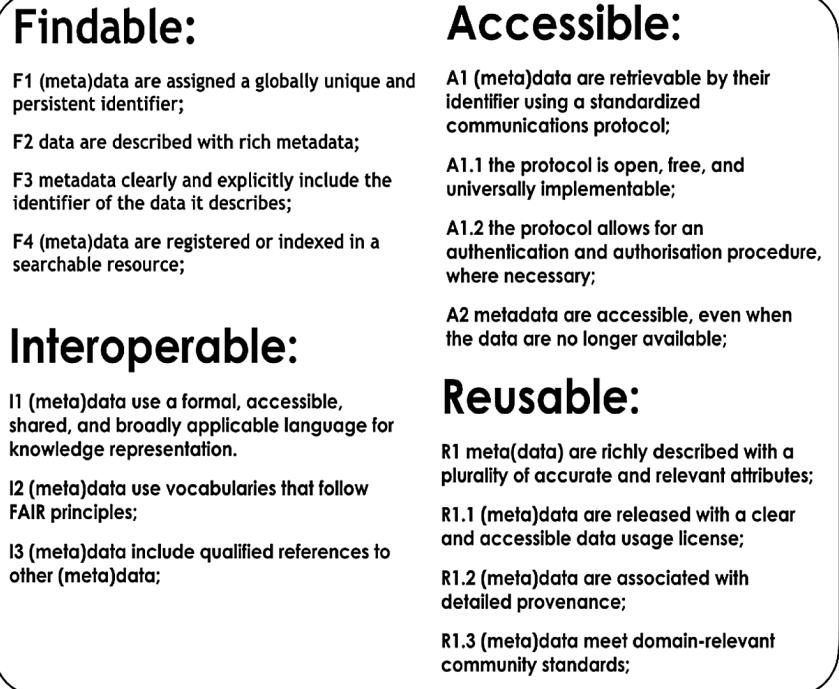
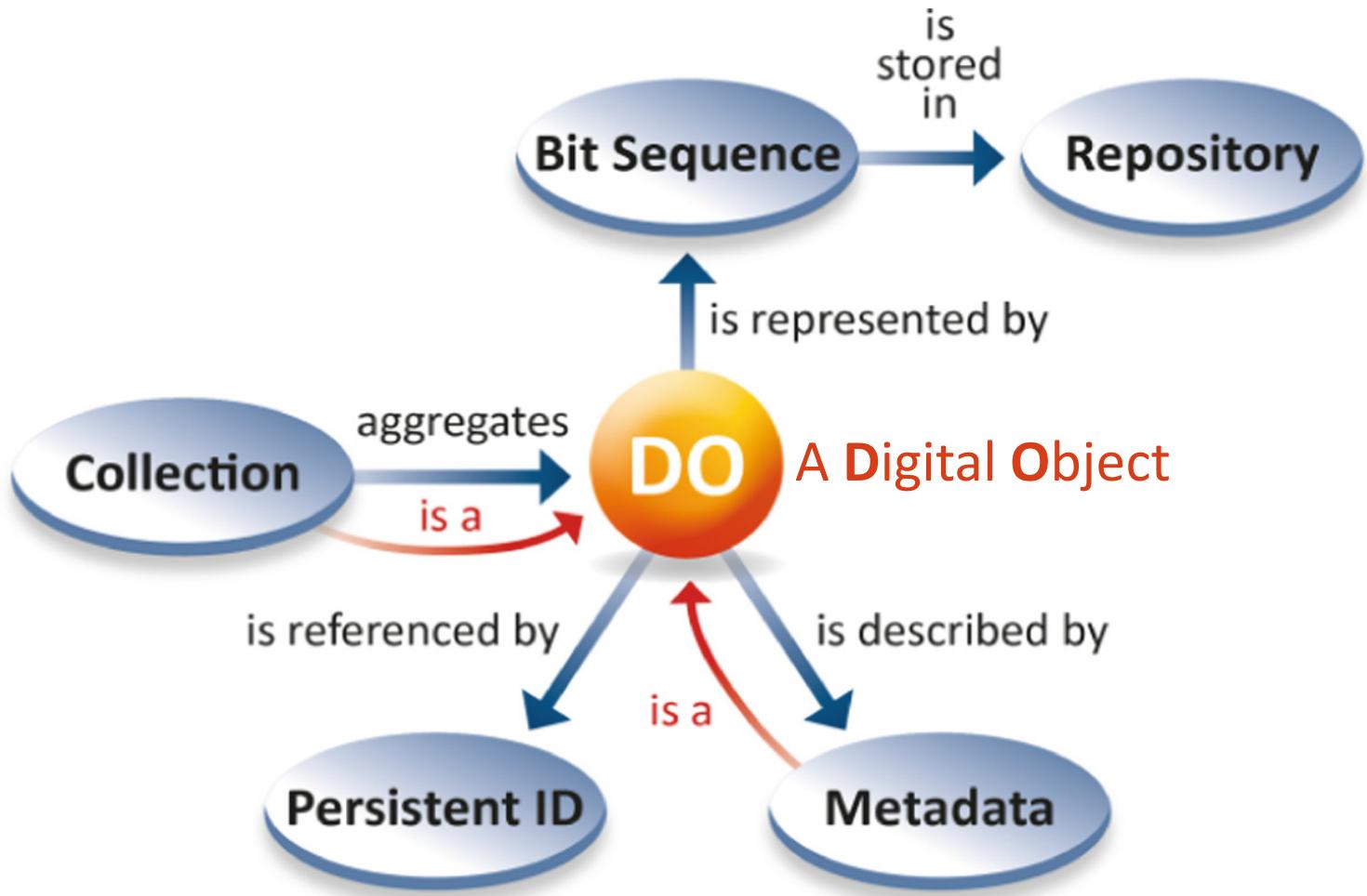
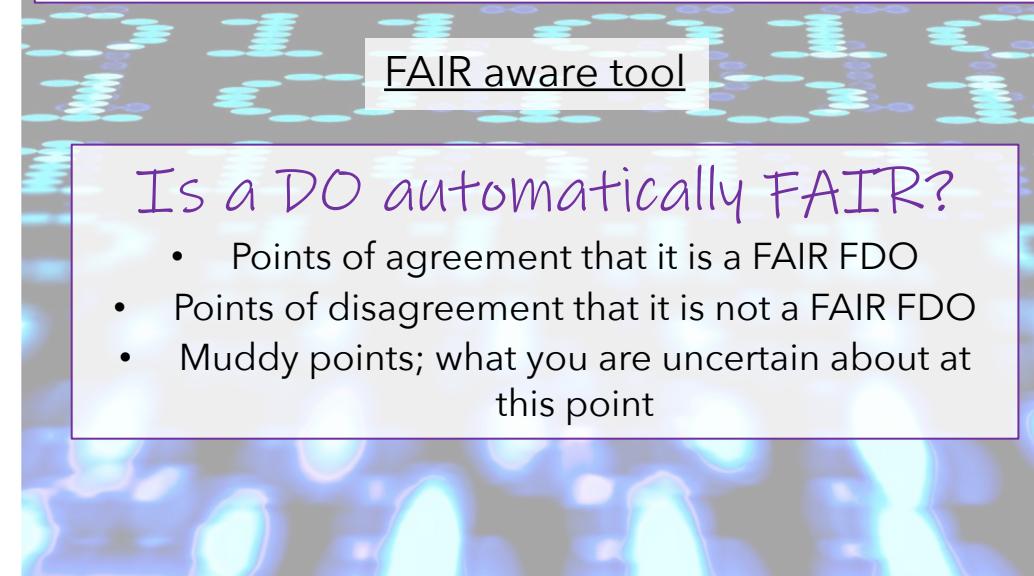


Fig. 1. The 15 FAIR principles ensuring machine findability, accessibility, interoperation and re-use of digital resources.

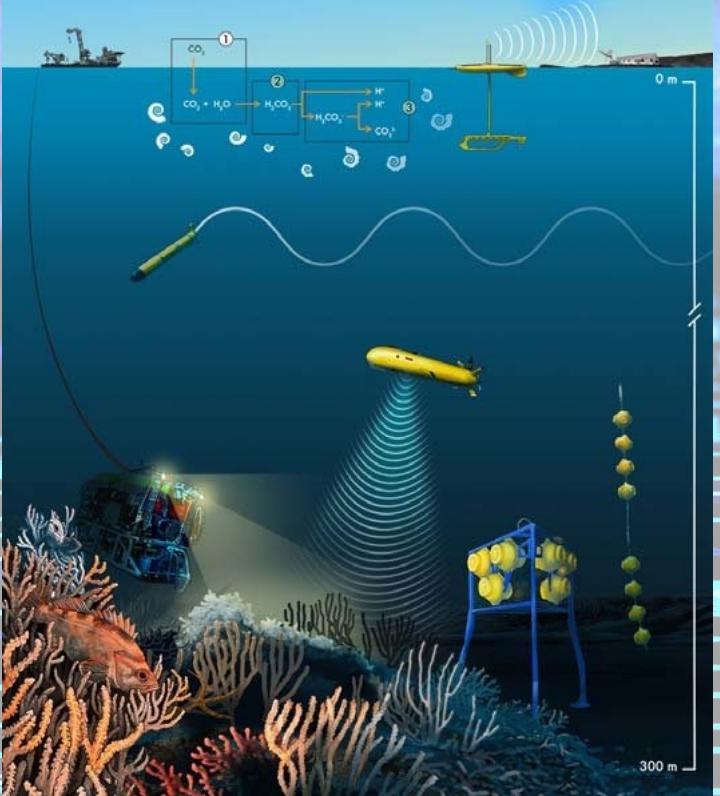


Is a DO automatically FAIR?

- Points of agreement that it is a FAIR FDO
- Points of disagreement that it is not a FAIR FDO
- Muddy points; what you are uncertain about at this point

An example process of becoming an FDO

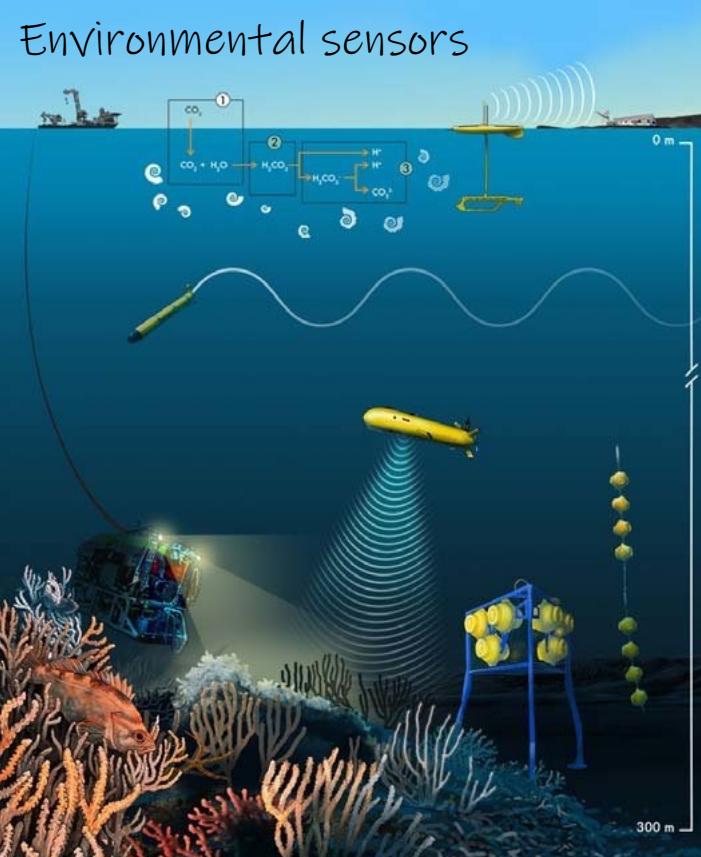
Environmental sensors



Data are produced as bit sequences
(along with associated metadata (time, location, sensor type,...))

Figure 9.1: Secretariat of the Convention on Biological Diversity (2014)

An example process of becoming an FDO

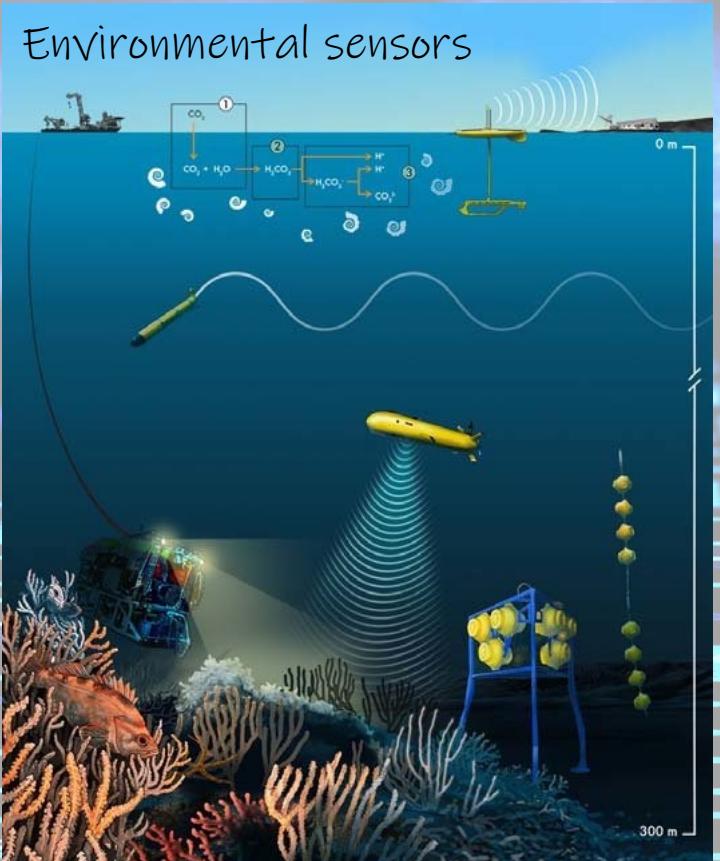


Data are produced as bit sequences
(along with associated metadata (time, location, sensor type,...))

sent to

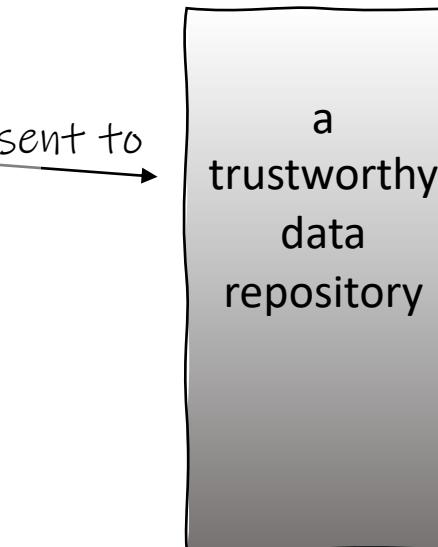


Environmental sensors



Data are produced as bit sequences
(along with associated metadata (time, location, sensor type,...))

An example process of becoming an FDO



Transparency

To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence

User Focus

To ensure that the data management norms and expectations of target user communities are met

Responsibility

To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service

Sustainability

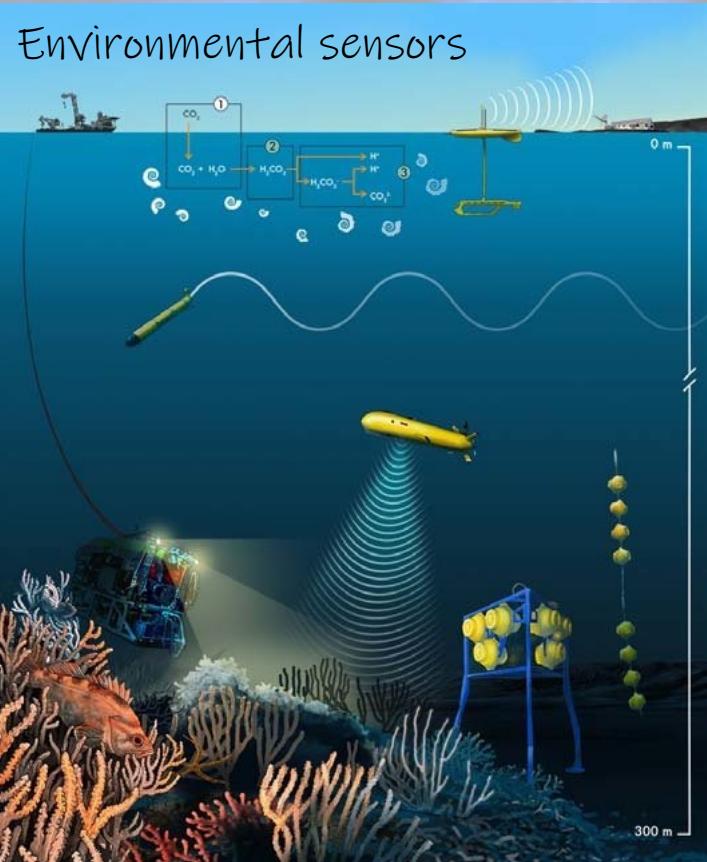
To sustain services and preserve data holdings for the long-term

Technology

To provide infrastructure and capabilities to support secure, persistent, and reliable services

- guiding principles for achieving data repository quality
- some repositories certify their trustworthiness – takes extra effort and is not currently required
- most popular certification: [CoreTrustSeal](#)

<https://doi.org/10.56367/OAG-039-10749>
<https://doi.org/10.1038/s41597-020-0486-7>



Environmental sensors

An example process of becoming an FDO

Data are produced as bit sequences
(along with associated metadata (time, location, sensor type,...))

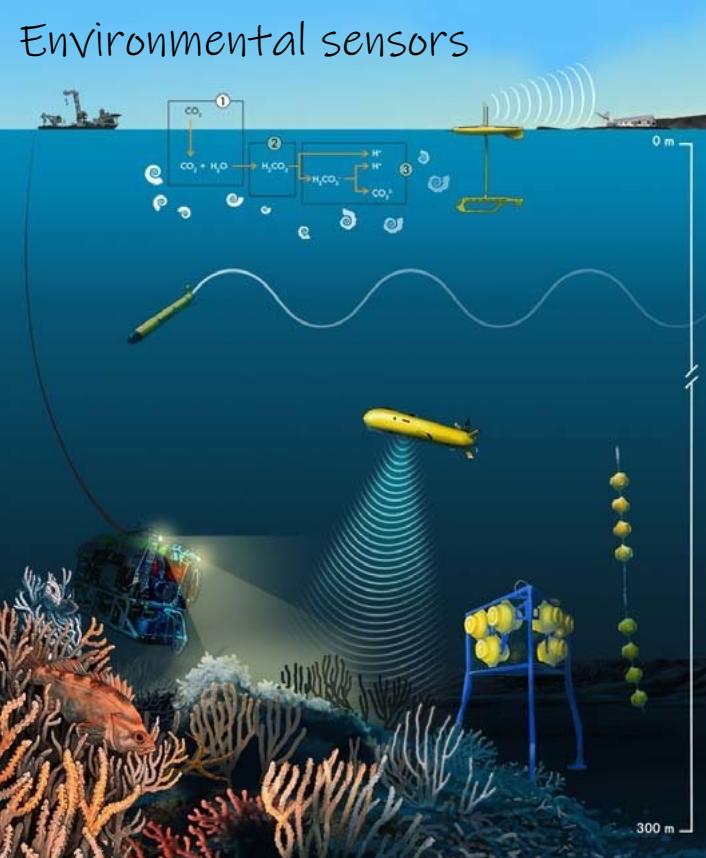
sent to



where to find data repositories?

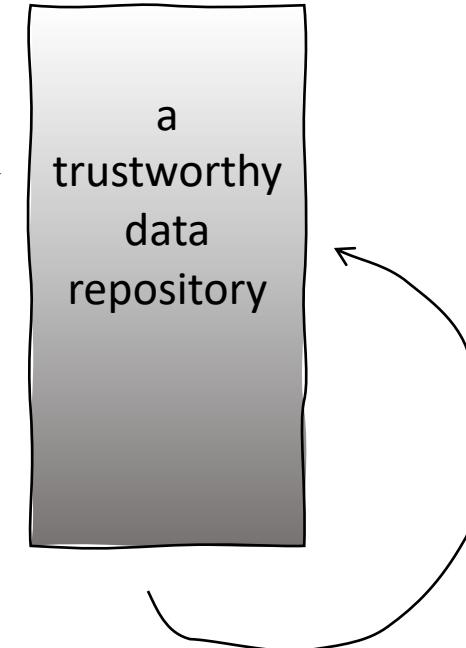
- re3data.org and FAIRsharing.org are global resources for finding data repositories (not all are certified)
- most journals seem to now recommend repositories, e.g.,
 - PLOS One
 - Springer Nature
 - Scientific Data
 - Taylor & Francis
- universities and research institutions are also often helping us find resources, e.g., from NTNU and UiO
- not necessarily FAIR

An example process of becoming an FDO



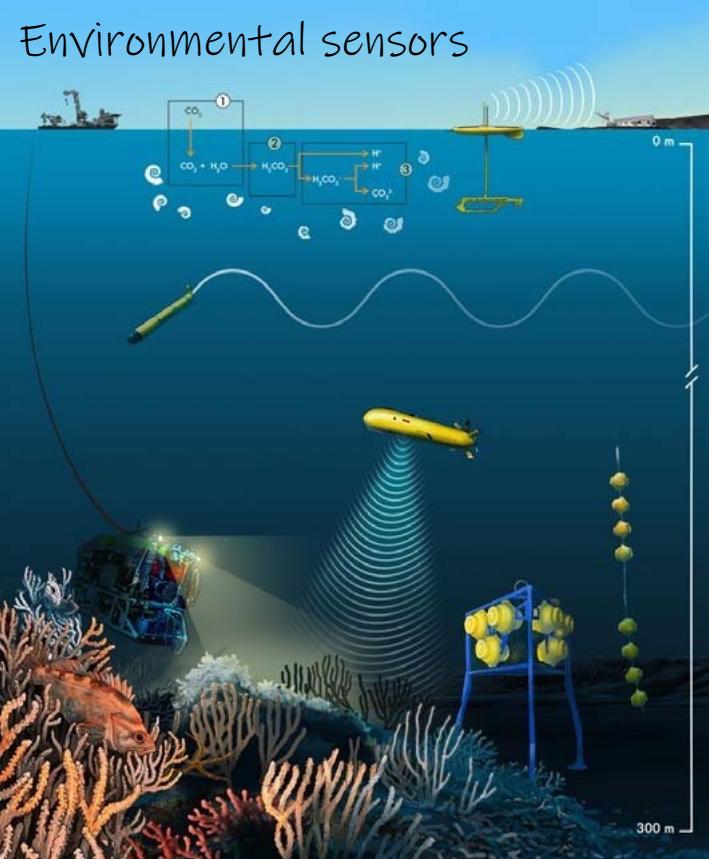
Data are produced as bit sequences
(along with associated metadata (time, location, sensor type,...))

sent to



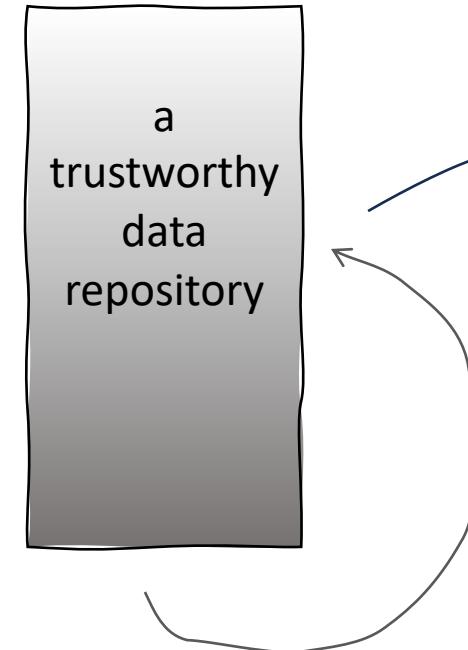
- data + metadata bundled into a Digital Object
- ideally a FAIR DO
 - assigned an FDO type with associated operations
 - has a Persistent ID
 - may extend metadata for permissions policies, licenses & access controls
 - adheres to a unifying protocol for interoperability

An example process of becoming an FDO



Data are produced as bit sequences
(along with associated metadata (time, location, sensor type,...))

sent to



- can be propagated online to further FDOs, with associated metadata to guarantee interoperability
- if legally allowed

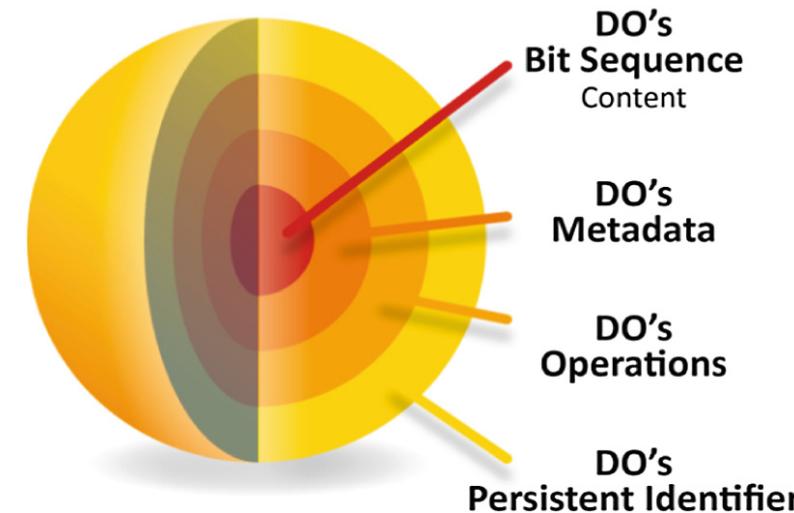
- data + metadata bundled into a Digital Object
- ideally a FAIR DO
 - assigned a FDO type with associated operations
 - has a Persistent ID
 - may extend metadata for permissions policies, licenses & access controls

"a stable [machine] actionable unit that bundles sufficient information to allow the reliable interpretation and processing of the data contained in it"

De Smedt et al. (2020) <http://dx.doi.org/10.3390/publications8020021>

Another way of seeing FAIR Digital Objects:

Data
Metadata
Software
Asssertions
Knowlets
etc

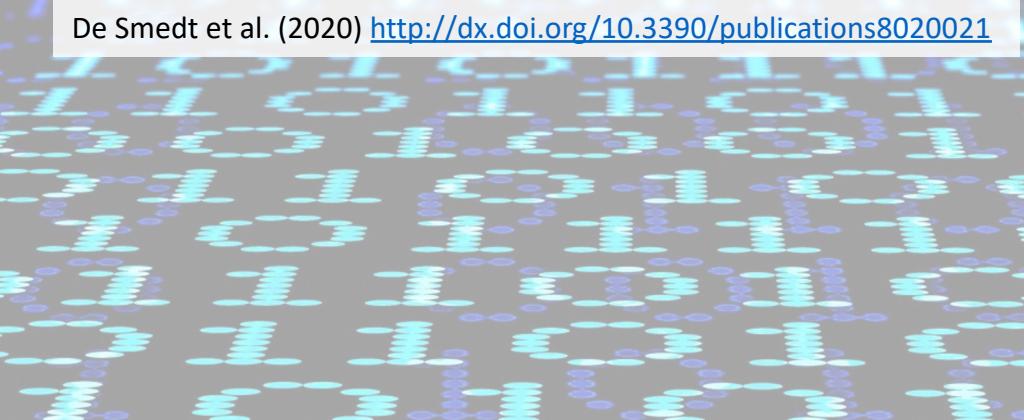


Schultes & Wittenburg (2019) https://doi.org/10.1007/978-3-030-23584-0_1

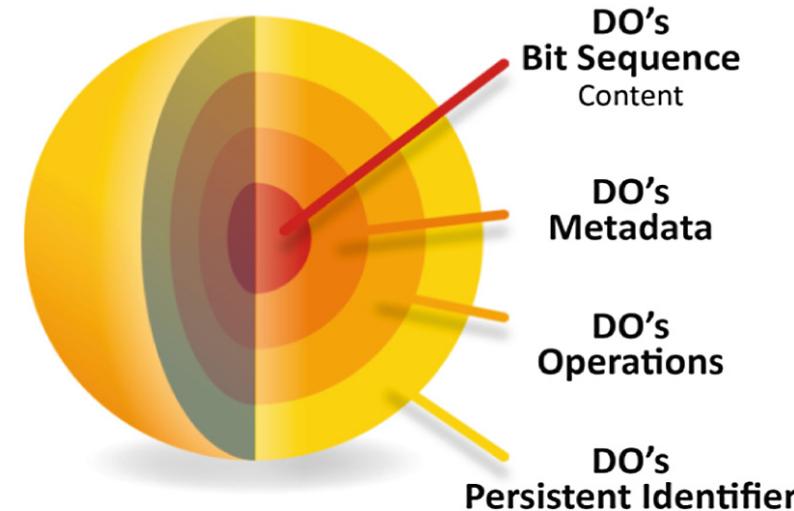
Another way of seeing FAIR Digital Objects:

"a stable [machine] actionable unit that bundles sufficient information to allow the reliable interpretation and processing of the data contained in it"

De Smedt et al. (2020) <http://dx.doi.org/10.3390/publications8020021>



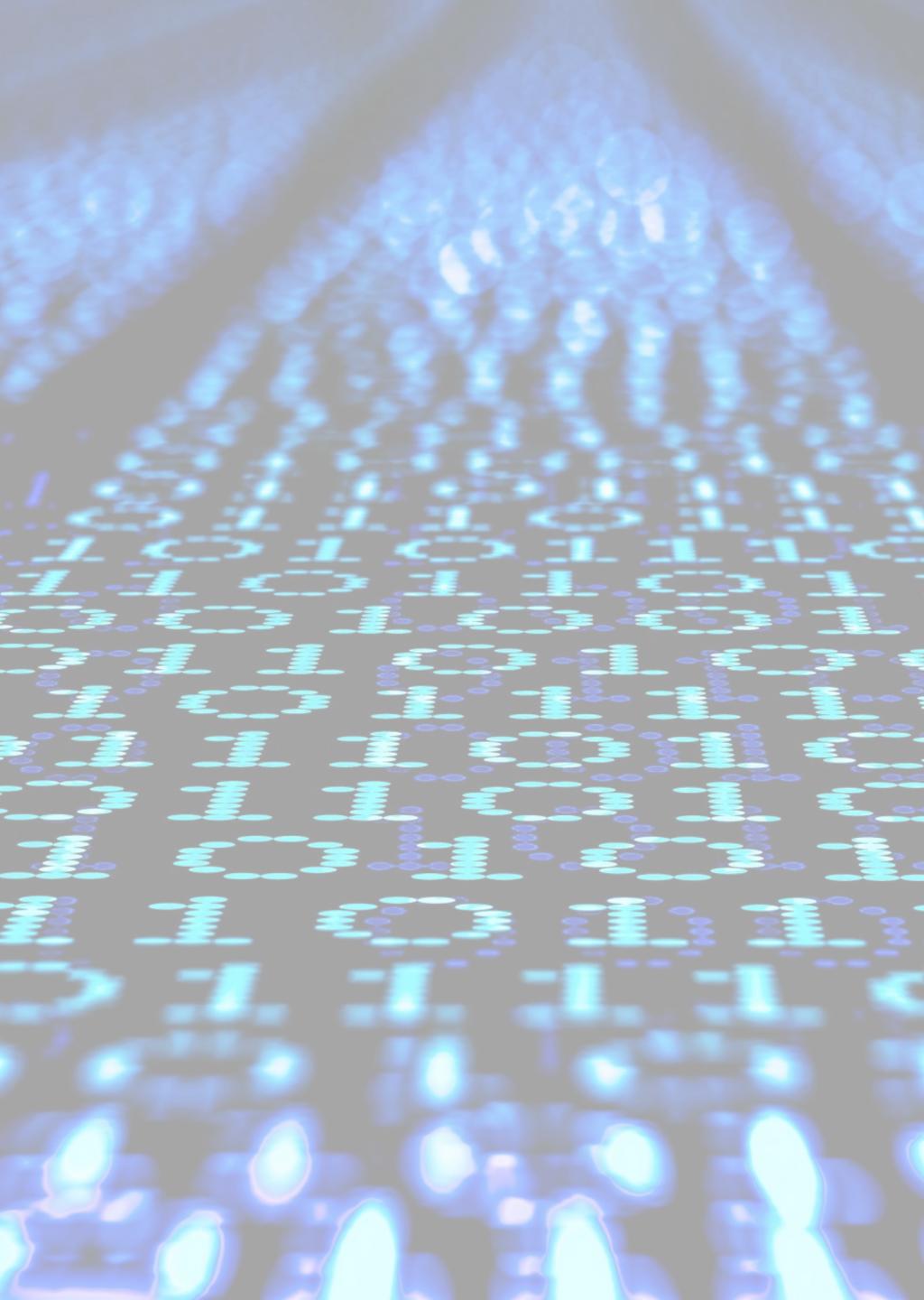
Data
Metadata
Software
Asssertions
Knowlets
etc



"Data Foundation and Terminology" 2014 Digital Objects summary:

- Core of data organization: binds entities necessary for stable & reusable data domain
- The bit sequence (content) can be stored in various repositories
- Referenced by unique and persistent identifier(s) (PID(s)) -- issued by a *trustworthy globally available resolution system*
- Described by various types of metadata (e.g., descriptive, system, access rights, license, contractual, transactional, etc.)
 - The metadata are also DOs
- Can combine to collections, which are also DOs
- Can include all kinds of digital information such as data, software, configurations, representations of persons, institutions, semantic concepts, etc.

Schultes & Wittenburg (2019) https://doi.org/10.1007/978-3-030-23584-0_1



Why does it matter to us if a Digital Object is FAIR?

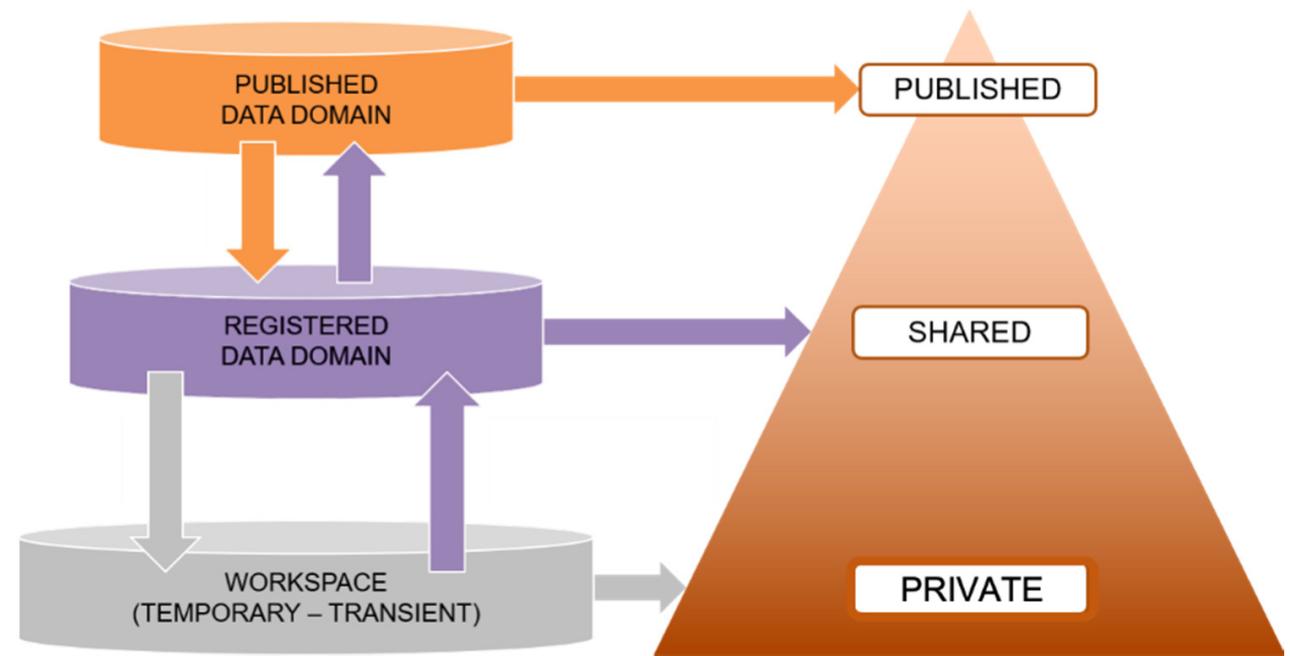


Figure 1. Layers of data with some data being published, more being shared for reuse in labs and collaborations, and a large amount residing in transient storage.¹

1. Cross-disciplinary data availability
 - Promote transparency
 - Increase trust in the data integrity and quality (through maintaining the context)



Why does it matter to us if a Digital Object is FAIR?

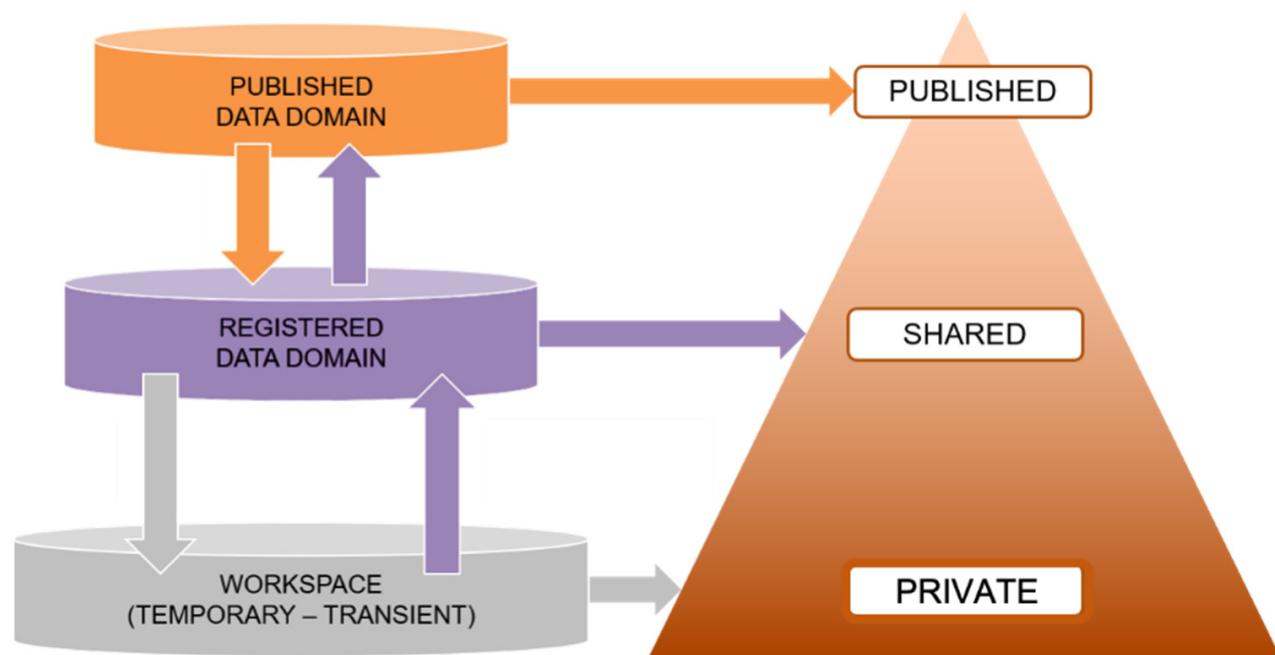
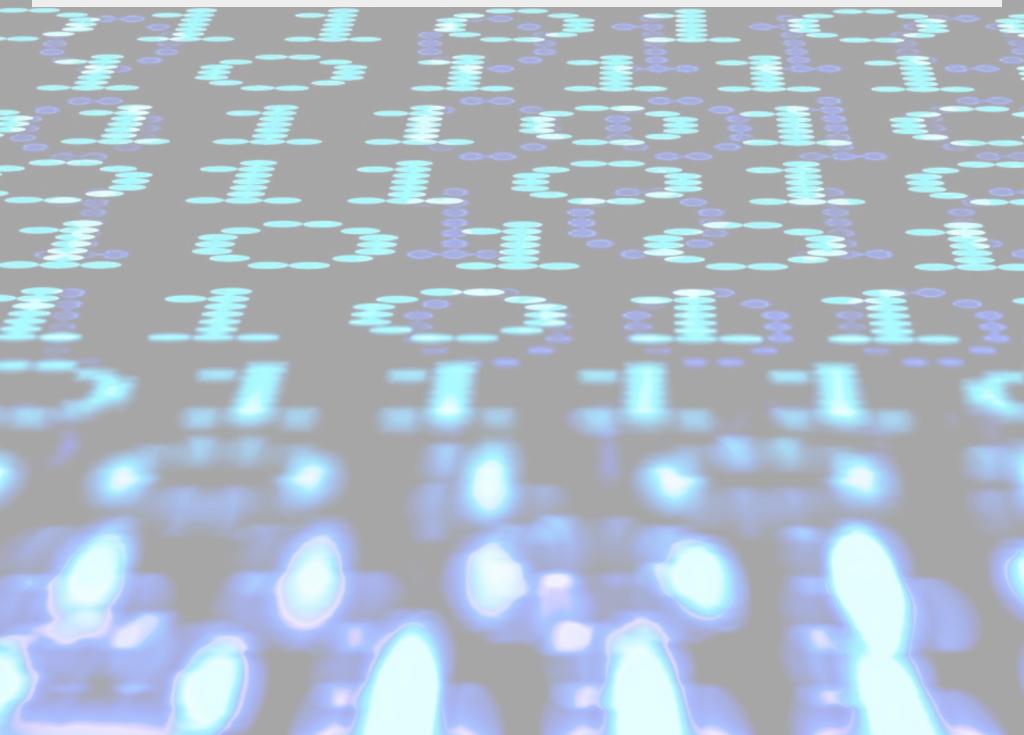


Figure 1. Layers of data with some data being published, more being shared for reuse in labs and collaborations, and a large amount residing in transient storage.¹

1. Cross-disciplinary data availability
 - Promote transparency
 - Increase trust in the data integrity and quality (through maintaining the context)
2. Stop the short “shelf-life” of data
 - Otherwise, knowledge becomes lost with time
 - We lose the context, origination & other details - that could instead become part of the metadata



Why does it matter to us if a Digital Object is FAIR?

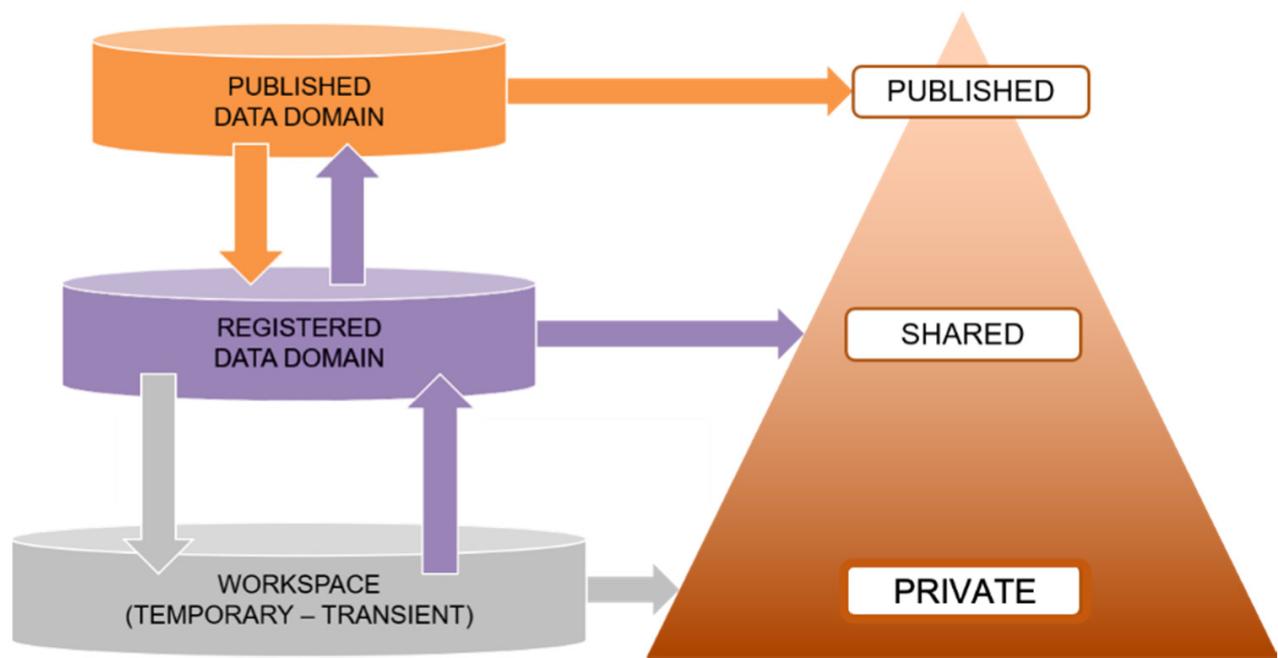


Figure 1. Layers of data with some data being published, more being shared for reuse in labs and collaborations, and a large amount residing in transient storage.¹

1. Cross-disciplinary data availability
 - Promote transparency
 - Increase trust in the data integrity and quality (through maintaining the context)
2. Stop the short “shelf-life” of data
 - Otherwise, knowledge becomes lost with time
 - We lose the context, origination & other details - that could instead become part of the metadata
3. Reusability, not repeatability
 - Reduce “wheel reinventions” when compiling, annotating, integrating, wrangling data
 - Often conducted by ±lower-position academic posts (e.g., graduate student, postdoc, researcher) – not all have this source of help
 - Time and money consuming
 - FAIR practices
 - Enable sharing and reuse of data-in-context of its origination / history

Why does it matter to us if a Digital Object is FAIR?

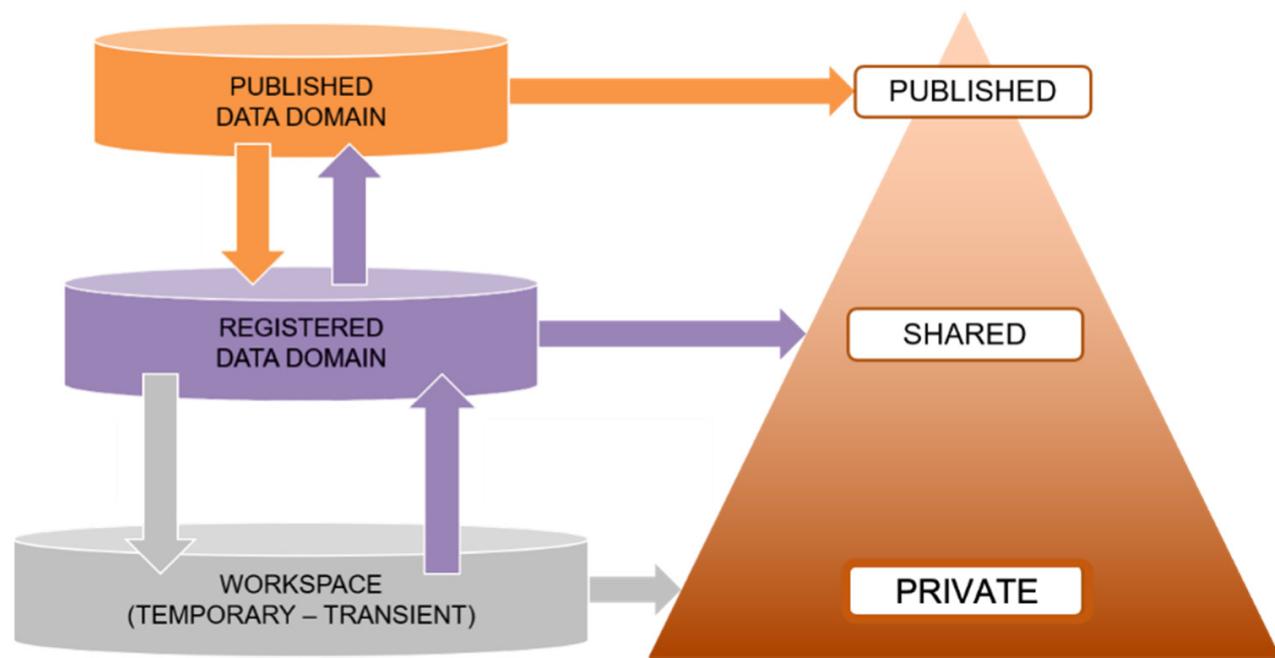


Figure 1. Layers of data with some data being published, more being shared for reuse in labs and collaborations, and a large amount residing in transient storage.¹

4. Machine actionability: AI

- Standardize data objects by using FAIR standards
- Automate the finding, extracting, evaluating and processing of digital data
- “Allowing automatic systems to make sense of the vast number of individual pieces of knowledge”

Why does it matter to us if a Digital Object is FAIR?

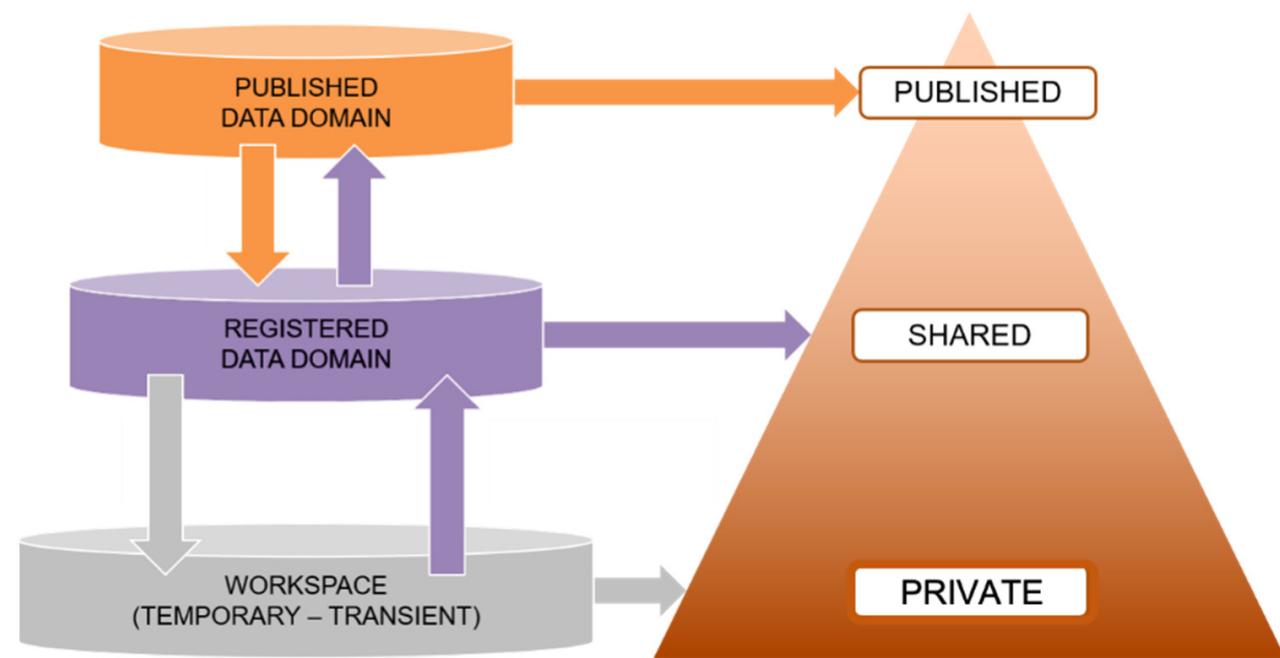


Figure 1. Layers of data with some data being published, more being shared for reuse in labs and collaborations, and a large amount residing in transient storage.¹

Why does it matter to us if a Digital Object is FAIR?

4. Machine actionability: AI

- Standardize data objects by using FAIR standards
- Automate the finding, extracting, evaluating and processing of digital data
- "Allowing automatic systems to make sense of the vast number of individual pieces of knowledge"

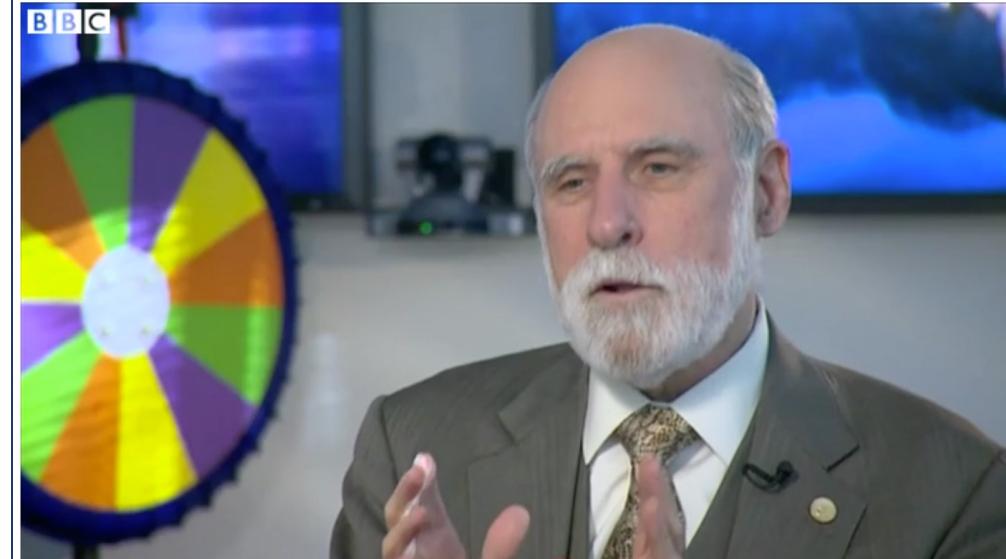
5. Visionary

- The web is currently ephemeral, e.g., with a lot of "link rot"; we risk sinking into a "dark digital age" without FDOs
- Prepare for the future while preserving the present (& past) data
- Transfer scientific knowledge to digital knowledge, instead of previously traditional forms (publications on library shelves)



Google's Vint Cerf warns of 'digital Dark Age'

© 13 February 2015



| Watch Pallab Ghosh's full interview with Vint Cerf, internet pioneer, on his "digital Dark Age" warning

By Pallab Ghosh

Science correspondent, BBC News, San Jose

Video: <https://www.bbc.com/news/science-environment-31450389> (3 min 12 sec)

Implementing FAIRness in Digital Objects

- Can be completed as an addition to existing internet protocols
- There is an [FDO Open Declaration](#) of support for a [FAIR Digital Object Framework](#)
- Annual summits bring interested participants together for discussions
- The 1st International Conference on FAIR Digital Objects (FDO2022) concluded by 130 people signing the [Leiden Declaration on FAIR Digital Objects](#). It appears that you could sign it, too 😊
- No one way: Multiple technical implementations promises interoperability

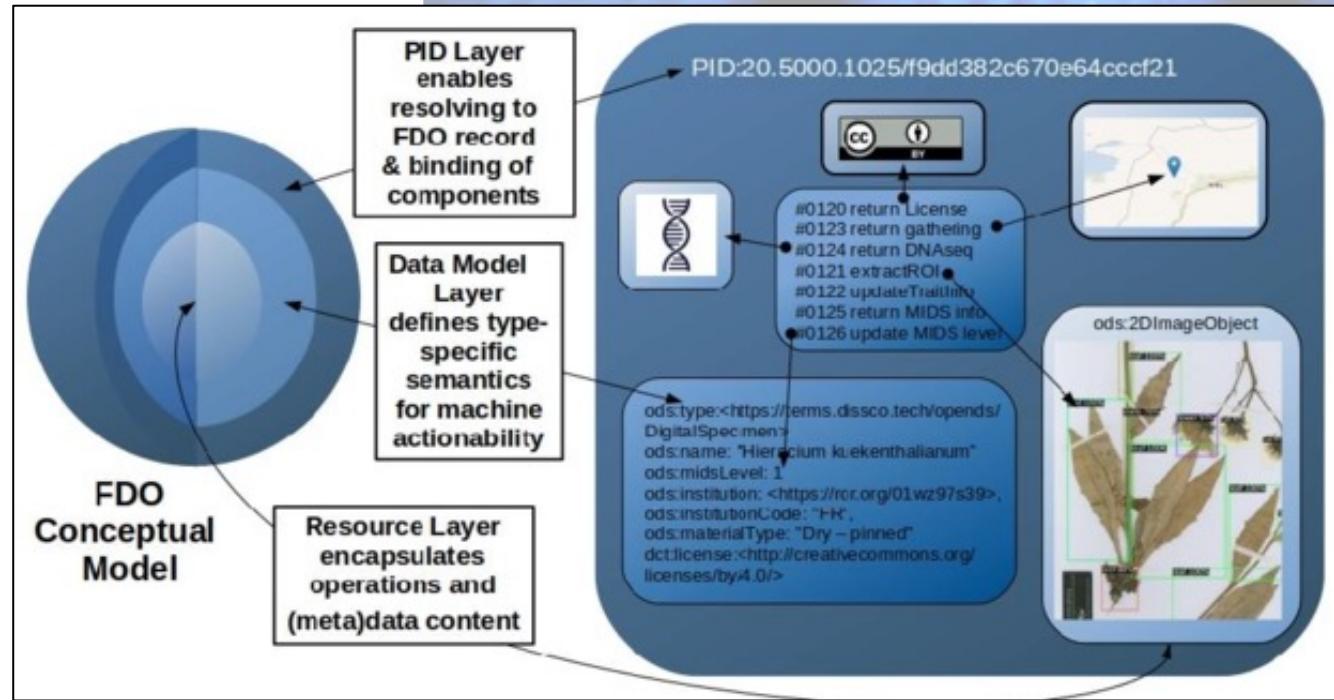


A screenshot of the FAIR DIGITAL OBJECTS FORUM website. The top navigation bar includes links for Home, Working Groups (highlighted in orange), News & Events, FAIR DIGITAL OBJECTS FORUM, Library, Get Involved, and About. A search icon is also present. The main content area is titled "Our working groups" and features a section for "Active groups". Three boxes are displayed, each representing a working group: "FDO TSIG II" (purple gear icon, FAIR Digital Object Technical Specification & Implementation Group (FDO-TSIG)), "FDO FIPP" (orange bar chart icon, FAIR Digital Object FAIR Implementation Profiles and Practice (FDO-FIPP)), and "FDO SEM" (green document icon, FAIR Digital Object Semantics Group (FDO-SIG)). Each box has a "MORE" link at the bottom. The background of the page features a blue and white abstract design.

Conclusion: we will continue to work more and more with FDOs

Organisations creating FDOs for specific tasks include:

- [DiSSCo](#) (biodiversity)
- [ELIXIR](#) (biomedical research)
- [E-rihs](#) (cultural heritage)
- [EISCAT](#) (atmospheric research)
- [ECRIN](#) (health initiatives)



"Exemplars of the corresponding digital objects are hosted in many institutions and labs worldwide. They are annotated, based on multiple information sources, taxonomies, and ontologies, and... the resulting digital objects will be part of workflows to generate derived data. Specifically designed collections serve as a basis for new theories. **Layers of digitally represented knowledge are thus created on top of the bottom layer of digital objects and form the incrementally growing scientific knowledge space, in a way similar to the knowledge network created by scientific papers with their references to other papers until now.** The inherent capabilities of abstraction, binding and encapsulation of FDOs based on stable identifiers will establish the **trust of researchers** to invest their time in **developing and maintaining these knowledge spaces** over the next decades. There is no doubt that researchers globally and in many disciplines are waiting on signals of convergence on FDOs, since this would mean that their investments will not be lost."

<https://doi.org/10.52825/cordi.v1i.263>