Meeting Notes: Data Fabrics - Digital Objects and Linked Data

Thursday, April 22 2021,
Breakout 8
This document: https://docs.google.com/document/d/11WNC3PqN5k1NLaHH3gY2cJn1uaNTHfGBT6qR0r3GPBg/edit

Meeting agenda:
(Slides will be available at https://www.rd-alliance.org/group/data-fabric-ig/wiki/monthly-meetings)

● 10 Minute Introduction
● 10 Minute Overview of DO Approach
● 15 Minute Discussion of DO Approach
● 10 Minute Overview of LDP Approach https://www.youtube.com/watch?v=BqJFPmEo9PA
● 15 Minute Discussion of LDP Approach
● 25 Minutes of moderated discussion about interoperation in an environment with both Approaches
● 5 Minutes wrap and action items

Attendance list:
Please sign in:

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainer Stotzka</td>
<td>Karlsruhe Institute of Technology, Germany</td>
<td><a href="mailto:rainer.stotzka@kit.edu">rainer.stotzka@kit.edu</a></td>
</tr>
<tr>
<td>Stian Soiland-Reyes</td>
<td>The University of Manchester; University of Amsterdam</td>
<td></td>
</tr>
<tr>
<td>Maggie Hellström</td>
<td>ICOS Carbon Portal and Lund University</td>
<td><a href="mailto:margareta.hellstrom@nateko.lu.se">margareta.hellstrom@nateko.lu.se</a></td>
</tr>
<tr>
<td>Larry Lannom</td>
<td>CNRI</td>
<td><a href="mailto:llannom@cnri.reston.va.us">llannom@cnri.reston.va.us</a></td>
</tr>
<tr>
<td>Ray Plante</td>
<td>NIST</td>
<td><a href="mailto:raymond.plante@nist.gov">raymond.plante@nist.gov</a></td>
</tr>
<tr>
<td>Hanna Koivula</td>
<td>CSC - IT Center for Science</td>
<td><a href="mailto:hanna.koivula@csc.fi">hanna.koivula@csc.fi</a></td>
</tr>
<tr>
<td>Zachary Trautt</td>
<td>NIST</td>
<td><a href="mailto:zachary.trautt@nist.gov">zachary.trautt@nist.gov</a></td>
</tr>
<tr>
<td>Name</td>
<td>Institution/Project</td>
<td>Email</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Anton Pirogov</td>
<td>Forschungszentrum Jülich, Germany</td>
<td><a href="mailto:a.pirogov@fz-juelich.de">a.pirogov@fz-juelich.de</a></td>
</tr>
<tr>
<td>Alex Hardisty</td>
<td>Cardiff University / DISSCo</td>
<td><a href="mailto:hardistyar@cardiff.ac.uk">hardistyar@cardiff.ac.uk</a></td>
</tr>
<tr>
<td>Robert Ulrich</td>
<td>re3data / Karlsruhe Institute of Technology, Germany</td>
<td><a href="mailto:robert.ulrich@kit.edu">robert.ulrich@kit.edu</a></td>
</tr>
<tr>
<td>Ulrich Schwardmann</td>
<td>GWDG Goettingen, ePIC</td>
<td><a href="mailto:uschwar1@gwdg.de">uschwar1@gwdg.de</a></td>
</tr>
<tr>
<td>Soizick Lesteven</td>
<td>CNRS/CDS (France)</td>
<td><a href="mailto:soizick.lesteven@astro.unistra.fr">soizick.lesteven@astro.unistra.fr</a></td>
</tr>
<tr>
<td>Tom Honeyman</td>
<td>Australian Research Data Commons</td>
<td><a href="mailto:tom.honeyman@ardc.edu.au">tom.honeyman@ardc.edu.au</a></td>
</tr>
<tr>
<td>Xiaolei Xia</td>
<td>Computer network information Center, CAS</td>
<td>xlxiao@cnic</td>
</tr>
<tr>
<td>Jonathan Clark</td>
<td>The DOI Foundation</td>
<td><a href="mailto:j.clark@doi.org">j.clark@doi.org</a></td>
</tr>
<tr>
<td>Ahmad Zainul Ihsan</td>
<td>Forschungszentrum Jülich, Germany</td>
<td><a href="mailto:a.ihsan@fz-juelich.de">a.ihsan@fz-juelich.de</a></td>
</tr>
<tr>
<td>Peter Pennefather</td>
<td>gDial Inc</td>
<td><a href="mailto:p.pennefather@gmail.com">p.pennefather@gmail.com</a></td>
</tr>
<tr>
<td>Anders Conrad</td>
<td>Danish e-Infrastructure Cooperation (DeIC)</td>
<td><a href="mailto:anders.conrad@deic.dk">anders.conrad@deic.dk</a></td>
</tr>
<tr>
<td>Peter Mutschke</td>
<td>GESIS, Germany</td>
<td><a href="mailto:peter.mutschke@gesis.org">peter.mutschke@gesis.org</a></td>
</tr>
<tr>
<td>Christian Pagé</td>
<td>CERFACS, France</td>
<td><a href="mailto:christian.page@cerfacs.fr">christian.page@cerfacs.fr</a></td>
</tr>
<tr>
<td>Constanze Curdt</td>
<td>GEOMAR Helmholtz Centre for Ocean Research, Germany</td>
<td><a href="mailto:ccurdt@geomar.de">ccurdt@geomar.de</a></td>
</tr>
<tr>
<td>Andreas Pfeil</td>
<td>Karlsruhe Institute of Technology, Germany</td>
<td><a href="mailto:andreas.pfeil@kit.edu">andreas.pfeil@kit.edu</a></td>
</tr>
<tr>
<td>Binh T. Nguyen</td>
<td>re3data / Karlsruhe Institute of Technology, Germany</td>
<td><a href="mailto:binh.nguyen@kit.edu">binh.nguyen@kit.edu</a></td>
</tr>
<tr>
<td>Gail Hodge</td>
<td>IIA</td>
<td><a href="mailto:ghodge@iiaweb.com">ghodge@iiaweb.com</a></td>
</tr>
<tr>
<td>Name</td>
<td>Organization</td>
<td>Email</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Doug Fils</td>
<td>Ocean Leadership</td>
<td><a href="mailto:dfils@oceanleadership.org">dfils@oceanleadership.org</a></td>
</tr>
<tr>
<td>Rob Quick</td>
<td>Indiana University</td>
<td><a href="mailto:rquick@iu.edu">rquick@iu.edu</a></td>
</tr>
<tr>
<td>Yann Le Franc</td>
<td>e-Science Data Factory</td>
<td><a href="mailto:ylefranc@esciencefactory.com">ylefranc@esciencefactory.com</a></td>
</tr>
<tr>
<td>Barbara Magagna</td>
<td>Environment Agency Austria</td>
<td><a href="mailto:barbara.magagna@umweltbundesamt.at">barbara.magagna@umweltbundesamt.at</a></td>
</tr>
<tr>
<td>Joan Damerow</td>
<td>Lawrence Berkeley National Lab</td>
<td><a href="mailto:JoanDamerow@lbl.gov">JoanDamerow@lbl.gov</a></td>
</tr>
<tr>
<td>Marina Sánchez</td>
<td>UC3M</td>
<td><a href="mailto:marinasa@pa.uc3m.es">marinasa@pa.uc3m.es</a></td>
</tr>
<tr>
<td>Fajar J. Ekaputra</td>
<td>TU Wien</td>
<td><a href="mailto:fajar.ekaputra@tuwien.ac.at">fajar.ekaputra@tuwien.ac.at</a></td>
</tr>
<tr>
<td>Jose Benito Gonzalez</td>
<td>CERN</td>
<td><a href="mailto:jose.benito.gonzalez@cern.ch">jose.benito.gonzalez@cern.ch</a></td>
</tr>
<tr>
<td>Ralph Müller-Pfefferkorn</td>
<td>TU Dresden / Germany</td>
<td><a href="mailto:ralph.mueller-pfefferkorn@tu-dresden.de">ralph.mueller-pfefferkorn@tu-dresden.de</a></td>
</tr>
<tr>
<td>David Elbert</td>
<td>Johns Hopkins University</td>
<td><a href="mailto:elbert@jhu.edu">elbert@jhu.edu</a></td>
</tr>
<tr>
<td>Wouter Addink</td>
<td>Naturalis/ DiSSCo</td>
<td><a href="mailto:wouter.addink@naturalis.nl">wouter.addink@naturalis.nl</a></td>
</tr>
<tr>
<td>Mariya Dimitrova</td>
<td>Pensoft/BAS</td>
<td><a href="mailto:m.dimitrova@pensoft.net">m.dimitrova@pensoft.net</a></td>
</tr>
<tr>
<td>Emanuel Soeding</td>
<td>GEOMAR / HMC</td>
<td><a href="mailto:esoeding@geomar.de">esoeding@geomar.de</a></td>
</tr>
<tr>
<td>Limor Peer</td>
<td>Yale University</td>
<td></td>
</tr>
<tr>
<td>Alex Henderson</td>
<td>University of Manchester</td>
<td><a href="mailto:alex.henderson@manchester.ac.uk">alex.henderson@manchester.ac.uk</a></td>
</tr>
</tbody>
</table>

**Discussions:**
- **Talk 0:** Here is the slide that was missing from the introduction talk with Guiding Documents and links:
  - Provides a high-level conceptual framework to support Digital Object management and service development
  - [http://dx.doi.org/10.15497/RDA00026](http://dx.doi.org/10.15497/RDA00026)
  - Consolidated assertions on Persistent Identifiers (2017/2018)
  - Provides a cross-disciplinary collection of statements on PIDs, reflecting commonalities across disciplinary perceptions and use cases
- http://dx.doi.org/10.15497/RDA00027
- FAIR Digital Object Forum
- o https://fairdo.org/
- o Spring 2022 FAIR DO Conference

- Talk 1: What is Digital Object Architecture trying to solve?
  - Overwhelming data amounts
  - APIs, multiple storage, security issues
  - Making data available - but Fragmentation challenges
  - FAIR Guidelines attempt to address - mostly oriented for machine-readability
    - Easy to implement for publications
    - Lab work is much more challenging to FAIRify
  - Add abstraction level for persistently identified digital objects
  - File system abstraction is familiar. For Digital Objects:
    - Globally unique/actionable identifiers
    - Every object is typed
    - Every object has associated metadata
    - Every object has queryable set of operations
    - Every object addressed via Digital Object Access Protocol
  - What are minimal DO requirements?
    - Identifier + Resolution
    - Types, type registry
    - Digital Object repositories (incl. Metadata, repositories)
    - Mapping/brokering systems for existing systems to map to DO
      - Don’t assume existing stuff will go away!
    - Protocol for repositories to implement
  - Need to talk common protocol
  - (Showing interactions of DOIP operations)
  - Digital Specimen Object example
    - Digital twin of physical sample
    - Identifier for the digital object different from identifier of physical specimen
  - FAIR Digital Object Forum
    - https://fairdo.org/
    - register at https://fairdo.org/register/
    - Spring 2022: FDO conference in Europe
  - Combine:
    - FAIR
    - FDO
    - LDP (Linked Data Platform)

- Discussion:
  - URLs from chat:
    - https://www.dona.net/
    - https://www.dona.net/specsandsoftware#block-digitalobjectinterfaceprotocol
    - https://www.w3.org/TR/ldp-primer/
○ Stian: Need for a minimal baseline protocol and type operations, e.g. CRUD on HTTP - as needed by LDP. Need for negotiation for support of additional, optional DOIP implementations.
○ Laurence: Digital Objects all the way down, types, even methods are identified Digital Objects.
  ■ (Mint a DOI for https://tools.ietf.org/html/rfc7231#section-4.3.1 ? :)
○
○ One type are variables: for this we would have a proposal: The I-ADOPT framework see I-ADOPT Framework 1.0.0 RDA 17 and the related presentation
○
  ● Luiz Bonino: FAIR Digital Object Framework (FDOF)
  ○ Luiz video: https://www.youtube.com/watch?v=BqJFPmEo9PA
  ○ Requirements
    ■ Identifiers, separation of metadata/objects, relations, machine-actionable metadata
  ○ Informational unit - sequence of bits
  ○ Current draft: https://fairdigitalobjectframework.org/
  ○ Discussion: Peter points to the FAIR Digital Object Framework definitions & other related materials at https://github.com/GEDE-RDA-Europe/GEDE/tree/master/FAIR%20Digital%20Objects
○
  ●

Chat Dump

Alex Henderson and 50 others joined the room - 2:01 pm
Stefanie
Welcome! Please add your name to the collab notes document at https://docs.google.com/document/d/11WNC3PqN5k1NLaHH3qY2cJn1uaNTHfGBT6qR0r3GPBg/edit
Stephanie Holmgren joined the room - 2:01 pm
Brian Westra joined the room - 2:03 pm
Robert Szczodruch joined the room - 2:04 pm
Daniel Mohr joined the room - 2:04 pm
Clement Jonquet joined the room - 2:05 pm
Stian
Is HTTP such a common protocol?
David Elbert joined the room - 2:06 pm
Megan Force joined the room - 2:06 pm
David Elbert joined the room - 2:08 pm
Christian Pagé joined the room - 2:08 pm
Stefanie
Welcome! Please add your name to the collab notes document: https://docs.google.com/document/d/11WNC3PqN5k1NLaHH3qY2cJn1uaNTHfGBT6qR0r3GPRg/edit

Christian Pagé joined the room - 2:11 pm

Ulrich
hand

Alex

Alex
SDK: https://www.dona.net/specsandsoftware#block-digitalobjectinterfaceprotocol

Stian
Combined with LDP would presumably be a HTTP-based implementation

Alex
DONA Foundation: https://www.dona.net/

Alex
FAIR Digital forum: https://fairdo.org

Stian
Linked DAta Platform https://www.w3.org/TR/ldp-primer/

Peter
time Stian - then it would be a duty I guess.

Peter
LDP is relying on the Web protocol stack

Ian and 1 other joined the room - 2:14 pm

Riina Luoma joined the room - 2:15 pm

Stian
with optional other DOIP access. Perhaps there could be a bridge as well. It sounds like some negotiation is needed..

Krishna Patel joined the room - 2:16 pm

Ulrich
but an operation is a special object wrt to just bitstreams, similar to functions wrt parameters

Barbara
May I say something about this variability?

Peter
sounds interesting Barbara

Peter
please send the link

Barbara
https://docs.google.com/document/d/1NGC5BQ9tKsGWnh_IFtMLBWu-BIT8vpulAanF6zYEr2I/edit and the related presentation

Barbara
The I-ADOPT framework

Barbara
https://padlet.com/barbaramagagna/sogprqszse1bgd24

Maria Stoica joined the room - 2:24 pm
sorry here the presentation
https://docs.google.com/presentation/d/16IETFQkeUttEvfENTVHDDN2jdwMb1mLzILmL-04/edit#slide=id.gd0d27f88ae_0_12
Ulrich
what the I-ADOPT group did, really looks promising.
Alex Henderson
+1 for I-ADOPT
Nicola Knight joined the room - 2:27 pm
Zachary
can you post the link to youtube?
Christian Pagé joined the room - 2:27 pm
Rainer
Video: https://www.youtube.com/watch?v=BqJFPmEo9PA
Limor Peer joined the room - 2:28 pm
Günther Gerrit joined the room - 2:29 pm
Stefanie
collab notes document for those who just joined:
https://docs.google.com/document/d/11WNC3PqN5k1NLaHH3gY2cJn1uaNTHfGBT6qR0r3GPB/edit
Siddeswara Guru joined the room - 2:29 pm
David Elbert joined the room - 2:30 pm
Barbara Magagna joined the room - 2:32 pm
Maria Luiza Campos joined the room - 2:32 pm
David Elbert joined the room - 2:33 pm
Peter Pennefather joined the room - 2:35 pm
Christophe Parisse joined the room - 2:36 pm
David Elbert joined the room - 2:36 pm
Dieter Van Uytvanck joined the room - 2:37 pm
Christian Pagé joined the room - 2:38 pm
Christian Pagé joined the room - 2:38 pm
Christian Pagé joined the room - 2:38 pm
Christian Pagé joined the room - 2:38 pm
Tom Honeyman joined the room - 2:38 pm
Stian
I'll have to go, but have to say how I like the FDOF approach as exactly the kind of "base types",
and "base methods"
Stian
and base protocol is then HTTP
Maria Luiza Campos joined the room - 2:39 pm
Krishna Patel joined the room - 2:39 pm
Peter
FDOF basics defined here:
https://github.com/GEDE-RDA-Europe/GEDE/tree/master/FAIR%20Digital%20Objects
Christian Pagé joined the room - 2:42 pm
Christian Pagé joined the room - 2:43 pm
Christian Pagé joined the room - 2:43 pm
Alex
\hand
Peter hand
Bryan Mc Alister joined the room - 2:46 pm
Christian Pagé joined the room - 2:47 pm
Charles Schmitt joined the room - 2:47 pm
Christian and 1 other joined the room - 2:48 pm
Peter
what about DO contents based identifier. The original DO will be owned by someone who could self-curate the persitence
Zachary
RO-Crate was speaker hosted in a monthly meeting:
Christian Pagé joined the room - 2:48 pm
Wouter
https://w3id.org/ro/crate
Ulrich
at ePIC we have a special kind of short live Handles that have prefixes like 21.Txxxx with T for test or temporari ly
Christian Pagé joined the room - 2:49 pm
Christian Pagé joined the room - 2:50 pm
Christian Pagé joined the room - 2:50 pm
Maggie
Of course we still need to support the equivalent of "tombstones", where the DO may be unavaiolable, but its related metadata live on
Christian Pagé joined the room - 2:51 pm
Peter Wittenburg joined the room - 2:51 pm
Christian Pagé joined the room - 2:51 pm
Christian Pagé joined the room - 2:52 pm
Christian Pagé joined the room - 2:52 pm
Christian and 1 other joined the room - 2:52 pm
Peter
Maggie is right: at that moment when people start referencing to an FDO you need to maintain the PID
Christian Pagé joined the room - 2:52 pm
Raymond
Objects that are created as a result of an operation will often have a "short-lived" identifier
Christian Pagé joined the room - 2:53 pm
Peter
it depends on the reuse of such an object
Christian Pagé joined the room - 2:54 pm
Christian Pagé joined the room - 2:54 pm
Christian Pagé joined the room - 2:56 pm
Peter
finally a PID does not cost much - just some bytes, so add a tombstone note
Anders
I think Larry put it right: take everything away to begin with, that doesn't break the architecture
Thomas Jejkal joined the room - 2:56 pm
Maggie
PIDs in themselves are cheap - but maintenance (and associated quality control and quality assurance) of their kernel metadata record may be quite costly!

Rob(Q) agreed Maggie
Raymond @Maggie +1
Peter
An FDO should also point to infos saying how the 15 FAIR principles are addressed (eg reference to the metadata scheme, the usage licence etc)

Maggie
Thus, "sustainability" is a key thing - and we'll need to lobby stakeholders & policy makers to get enough funding!

Barbara
Peter: this again links than to FIPs...which are implementations for each of the FAIR principles for this resource

Peter @Barbara: Indeed. As I already said at the FAIR Convergence Symposium, I see nice relationships between the two approaches

Peter
Ideally, a FDO should contain a reference to a FIP, to make it more machine-actionable

Barbara
I agree
Alex
I agree we won't resolve today. Because we must think about what it is that conforms to any FDO specification. It's the system.

Peter
Folks - I do not know how FIPs relate in this context. FIPs somehow state the profiles of some provider which is another level - isn't it?

Maggie
So I think we'll need to collect a lot more use cases and examples on how all of this can be concretized

Alex
Rainer/Rob: Can you keep the chat for the notes please?

Stefanie
@Alex we can capture the chat and send it to the chairs

Rob(Q) yes, stefanie please send it to us and we'll publish it to the group

Peter @Peter: yes, but you can always break it down to the DO level

Rob(Q) @Rainer, maybe we can schedule a full presentation from Barbra in an upcoming Monthly meeting?

Peter
ok will have to look again, but we cannot leave the definition of what an FDO is to the individual providers. We need a "plug-in" structure with formal validators that check compliance

Rainer
@Rob: yes. Having Barbara presenting I-ADOPT in our monthly project shares would be great.

Tibor Kálmán joined the room - 3:12 pm
Maggie
Thanks all for an extremely interesting session, with a great dynamic exchange of views and ideas!
Barbara
thank you
Stefanie
It won't cut off :-)
Stefanie
As long as at least one speaker stays on, it won't close
Paul Stokes joined the room - 3:13 pm
Peter
indeed thanks to Rob and rainer
Alex
https://fairdo.org
Maggie
Be FAIR, or be square!
Alex
Thx Rob and Rainer.
Anders
Thanks, all!
Barbara
thanks to all
Ulrich
thanks
Jonathan
thanks - good session
Peter
Thanks all!
Yann
thanks for this great session
Christian
Thanks for this great session!!
Maggie and 1 other joined the room - 3:14 pm
Peter
Great session!
Emanuel
thank you!!
Alex and 1 other joined the room - 3:14 pm
Xiaolei
thanks!
Robert Ulrich joined the room - 3:14 pm
Alex Hardisty joined the room - 3:15 pm
Barbara
Thank you happy to join one of your next sessions then
Barbara Magagna joined the room - 3:15 pm
Emanuel Soeding joined the room - 3:15 pm
Q&A Dump (well the end of it anyway, not all questions were on screen)

Laurence
0 votes

Peter
Those associated with creation of DO have a stake their persistence. If the DO also has links to those who have authorship rights. They can self-curate DOs with content based identifiers that will be uniquely identifiable because they were created at particular time and place
0 votes

Peter
A FDO is by definition created for re-use. But the author of the FDO may want to choose when it is published to be examined a re-used by others. So there will be a time before which the same object becomes FAIR
0 votes