

Manifesto of the C2CAMP GO FAIR Implementation Network

C2CAMP is a group of international experts from relevant research institutions and infrastructure initiatives committed to build interoperable components that are based on specifications from RDA working groups and other comparable initiatives such as OAI, W3C, ISO, and IETF and to integrate them into flexible testbeds. As is the case with RDA, so is C2CAMP committed to the FAIR principles. The group will collaborate closely on what is called the Digital Object model which can be seen as one way to implement significant aspects of the FAIR framework¹. Scientific practices increasingly depend on high quality and interlinked digital information and respective online services. To enable data infrastructures to provide the expected level of service, it is imperative that we underpin their critical path of operations. A globally agreed-upon framework for structuring, curating and communicating Digital Objects is urgently needed.

The big challenge for data infrastructure builders is to increase the efficiency of working with data, which currently is not a given. Three reports from RDA EU, MIT, and CloudFlowers indicate that between 75 and 80% of scientists' time is wasted with so called data wrangling, include the steps needed before the actual analytics work can be done. Basically this is due to missing or bad specifications of the data organisation and a lack of data/metadata quality.

A Digital Object has content expressed in bit sequences stored in some trustworthy repositories, and is associated with a PID and different types of metadata such as location of the bit sequences, its checksum, the PIDs of its descriptive metadata and provenance information, its access permission record, its type to enable automatic processing, etc. A DO can include data, software, workflows, configurations and others entities. A number of RDA groups have been working on the DO concept, including Data Foundation & Terminology, PID Information Types, Data Type Registry, Kernel Information Types, Practical Policies, and Data Citation,² to just mention a few. Some software has already been developed and tested according to the RDA specifications and some communities follow already a DO centred approach in infrastructure building. It is now time to test it out in configurations bringing a number of the results together including components from other initiatives such as ResourceSynch from OAI.

The method of work will be as follows:

- the C2CAMP core partners will push projects which will embed DO-based components together with those from other focus areas
- the C2CAMP core partners will take care that these projects will fit with the C2CAMP guidelines extracted from the RDA and other specifications
- the C2CAMP core partners will ensure that there is a compliance with the FAIR principles where applicable
- the C2CAMP core partners will take care that new challenges for specification will be addressed in the realm of RDA, i.e. it will use the possibilities that this international, cross-disciplinary platform offers

The reason to participate in the GOFAIR initiative as an IN is clear. We want to have a close discussion link to ensure not only FAIR compliance but also to interact with other implementation groups to increase interoperability where possible.

Guiding Purpose: Increased Interoperability through Digital Objects

DOs as one way to implement FAIR partly will help to increase the efficiency of data science enormously.

Deliverables

C2CAMP will regularly develop documents that summarise and evaluate component specifications and their characteristics and interfaces, describe new components needed, identify gaps and weaknesses. We will share these documents within the GOFAIR community and raise discussions.

¹ This DO concept does not address semantic interoperability for example, although it might facilitate it.

² The information about these groups can be found on the RDA web-Site

We will also suggest starting a "components" subgroup within the GOFAIR implementation pillar to keep a close documentation of available and developing components and stimulate cross-fertilisation.

Overarching Principle of Operation

We commit to comply with the Rules of Engagement of GO FAIR Implementation Networks.

The primary objectives of C2CAMP thus include:

1. Developing components based on RDA specifications (and others) to make data intensive science more efficient and disseminate the knowledge about such components.
2. Push projects that implement such components in an integrated way to gather much more experience.
3. Aim at a global testbed project to test out the various components.
4. Indicate missing components and other specifications and look for ways to overcome the gaps by engaging RDA (and others).
5. Take care that the components are compliant with existing specifications and push projects to test compliance.
6. Take care that FAIR compliance is given.

The C2CAMP core network is thus a network of experts that share the same objectives. With respect to GOFAIR C2CAMP is currently represented by : Dimitris Koureas and Peter Wittenburg.

Founding Members are

- Larry Lannom CNRI (US)
- Peter Fox, Mark Parsons DCO, RPI (US)
- Robert Quick Indiana U (US)
- Eric Nienhouse NCAR (US)
- Jonathan Clark, Paul Jessop DOI (Int)
- Peter Wittenburg, Raphael Ritz MPCDF (DE)
- Dieter van Uytvanck, Koenraad de Smedt CLARIN ERIC (NL)
- Dimitris Koureas DiSSCo, Naturalis (NL)
- Tobias Weigel DKRZ, ENES (DE)
- Per Öster CSC (FI)
- Ari Asmi, Werner Kutsch ICOS, U Helsinki (FI)
- Colin Mc Murtrie, Thomas Schulthess CSCS (HS)
- Daniel Mallmann, Thomas Lippert FZJ (DE)
- Tian Ye CNIC (CN)
- Anwar Vahed DIRISA (SA)
- Rainer Stotzka KIT (DE)

Some additional members have already indicated commitment to participate, but these will be mentioned on the coming web-site.

The C2CAMP Founding Members agree with this document.

Signed by:

Dimitris Koureas

Peter Wittenburg