Active DMP Interest Group at RDA plenary 9 in Barcelona

The notes below record discussion and decisions from the Active DMP IG held on Thursday 6th April at 9:30-11am. The session description and agenda are available on the RDA website.

Original googledoc with attribution on comments.

Meeting agenda

1. Summary of conclusions from CERN, IDCC and other relevant workshops

Presentation slides

Sarah Jones gave an overview, reporting on the CERN workshop in June 2016, an IDCC workshop in February 2017, and a white paper published in RIO Journal with proposed next steps. The IDCC workshop and white paper focus on machine-actionable DMP use cases.

2. Presentation of potential work items, outline specification and stakeholders

Presentation slides

Based on the overview and ideas presented in the white paper, Kevin Ashley proposed three potential areas of work (as noted in the slides) and invited suggestions from others. Two further options were proposed by Marta Teperek and Neil Chue Hong during the session.

- **DMP common standards**: define a standard for expression of machine readable and actionable DMPs
- **Exposing DMPs**: develop use cases, workflows and guidelines to support the publication of DMPs via journals, repositories or other routes to making them open
- **Domain/infrastructure specialisation**: group to explore disciplinary tailoring and the collection of specific information needed to support service requests and use of domain infrastructure
- **Funder liaison**: group to engage with funders, support DMP review ideas and develop specific use cases for their context
- **Software management plans**: group to explore the remit of DMPs and inclusion of different output types e.g. software and workflows too

3. Group work to refine this and define scope of new working group(s)

There was sufficient interest in all 5 topics so the remaining time focused on discussing the potential remit and outputs of new working groups.

DMP common standards*:*
Requirements:
- Extensible
- Permanent identifiers for both questions and answers (FAIR DMPs)
- Ingest data and export data
- Versioning of the plans with a track change
- Judgement / balance between known use cases and future requirements
- Protocols used for the actual exchange - use well supported e.g. http, json, xml
- Also semantic interoperability and expression

Interested parties: Paul Walk, Tomasz Miksa, Raphael Ritz,(Max Planck Germany), Rob Hooft, M-Christine Jacquemot (Inist-CNRS, France), Jens Ludwig (State Library Berlin, Germany), Cristina Ribeiro (INESC TEC, Porto Portugal), Daniel Mietchen, Antonio Sánchez-Padial(INIA), Mark Leggott (RDC)

Exposing DMPs (less defined than 'publishing')
- Link to draft case statement - comments, contributions very welcome
- Machine readable
- Human readable
- Inter-disciplinary readability? (transdisciplinary semantic) /ratio effort / efficiency of maDMP? (suggestion from CNRS France - Romain DAVID)
- Identifiers available (who is registering PIDs for these?)
- Output - opt in and opt out with decision based on funders etc. when in flow, reason, who gets to make that decision, etc
- Potential to include reviewing DMP in this too? Suggestion from Elsevier. Also EC's public review of plans - https://ec.europa.eu/futurium/en/search/site/%2522data%2520management%2520plan%2522
- Desire to publish DMP alongside data so people can understand where data has come from (Keith Russell comment on GoToMeeting) but must take in account timing; dmp is way before data collection, analysis and deposited. But a link back and forth - it would be nice to have demo as early in the flow as possible
- What do you share and when? May only want to share certain aspects of DMP due to sensitivities, or share with certain groups e.g. your institution or publicly (feature in DMPTool)

Interested parties: Angus Whyte, John Chodacki, Elena Zudilova-Seinstra (Elsevier), Romain David (CNRS, France), Cristina Ribeiro (INESC TEC, Porto Portugal), Daniel Mietchen, Mike Brown (CEH), David Carr, Robert Kiley, Aki MacFarlane (Wellcome Trust)

Domain and infrastructure
- Large domain/topics with framework/template that are restrictive but also more palpable to researcher
- Open ended process for those long tail/ less categorized topics/fields
- Lessons learned from other systems
- *Would software as a domain go here?*
- *Technical outputs* term for humanities
- {like some questions being automatically filled for projects that are fulfilling certain criteria}
- Question from Jamie Shiers about how well this scales
- Idea to have a minimal DMP for research council requirements and then extend for domain info

Interested parties: Peter Doorn, Susanna Sansone, Mercè Crosas, M-Christine Jacquemot, Daniel Mietchen, Antonio Sánchez-Padial (INIA), Birger Jerlehag, Ilze Lace

Definitely willing to follow: Rob Hooft, Romain David

Funder liaison
- Specific group to work with funders
- Negotiate and educate funders on actionable dmps
- Build on existing work in US & UK to develop evaluation rubrics
- Understand what the funders want/need
  - comment from a funder perspective: need to consider general/minimum requirements for a DMP vs. requirements for a given data type or discipline
- Could link with discussion today - next breakout (not tomorrow!)IG Data policy standardisation and implementation: Standardisation of policies for publishing research data

Interested parties: Marta Teperek, Sarah Jones, Stephanie Simms, Lisa Zilinski, Kylie Emery, Daniel Mietchen

Software management plans / output / process plans
- Q: Extending machine actionable DMPs to also cover software management plans, e.g. as per https://ssi-dev.epcc.ed.ac.uk/smp-service ? A: Yes, there should be close integration of data and software management

- Develop 'Output Management Plans' covering all outputs. Perhaps output is the wrong word though as it focuses on what is published and shared. Need to ensure emphasis is placed on data management in general

Interested Parties: Neil Chue Hong, Andi Rauber, Tomek Miksa, Patrick Aerts, Mike Brown

Preservation of plans? Storage or repository or at least recommendation?

The discussions remind me of discussions that i have heard around Laboratory Information Management Systems (LIMs) and electronic lab notebooks. These systems can be very useful and active, but the more functionality added, the less transferable to other systems/labs/situations they are (clearly a broad statement, but these are problems that I've seen). Need to ensure that broad usability is not compromised.

4. Summary, actions, next steps

The Active DMP IG will circulate these notes as a record of the meeting and to inform the community of ongoing work. Those who have noted an interest in each of the proposed WGs will be contacted to begin to develop case statements. An update on progress for each group and whether their ideas were pursued further will be presented at plenary 10 in Montreal.

Administrative updates related to the interest group were also done at the close of the meeting. To expand the international coverage of group chairs, Stephanie Simms from UC3 was voted onto the group. We are seeking one further co-chair since Helen Glaves wishes to demit her role following her election to TAB. We have one nomination already; any further nominations should be made to Kevin Ashley at kevin.ashley@ed.ac.uk by April 20th.

Comment [11]: It's not just outputs - inputs and the process need to be managed as well, which is why some have proposed the term "Project Management Plan" (see slide 10 in https://www.coar-repositories.org/files/7_DMP_Vienna.pdf)

Comment [12]: slide 10 is Process Management Plans :)

Comment [13]: More acronyms PMPs...

Comment [14]: Instead of or in addition to "management", some people also use "stewardship", to express a more long-term perspective (starting before data come in, preserving way beyond duration of project)

Comment [15]: linked to semantic evolutions and readability