

Integrating maDMP with the data management process

Adil Hasan (Uninett Sigma2), Anne Fouilloux (UiO), Christine Jacquemot (DMP OPIDoR / Inist-CNRS France)

The data management process for a research project covers the collection, storage, transformation, analysis and archiving or deletion of the data. Each of these stages can involve some complex and detailed steps. Projects in the same discipline may have very similar, or even the same, series of steps in a given stage. For example, more than one project may make use of the same multi-purpose detector. For each project the transformation from the raw data into more physical quantities (such as pulse signals from a detector into three-dimensional coordinates) could be common. If one project were able to define such a workflow and include it as part of their DMP, other projects would be able to reuse it. Capturing these workflows in the DMP would more accurately reflect the data management process and provide some value to follow-up, or related projects that have similar workflows and which would be able to reuse these workflows.

What was the task/Issue defined for the hackathon?

The task we set ourselves for the hackathon was to take some specific workflows from a portion of the data management process for the Climate community in the EOSC-Nordic project and see how we could capture the workflow within the maDMP schema.

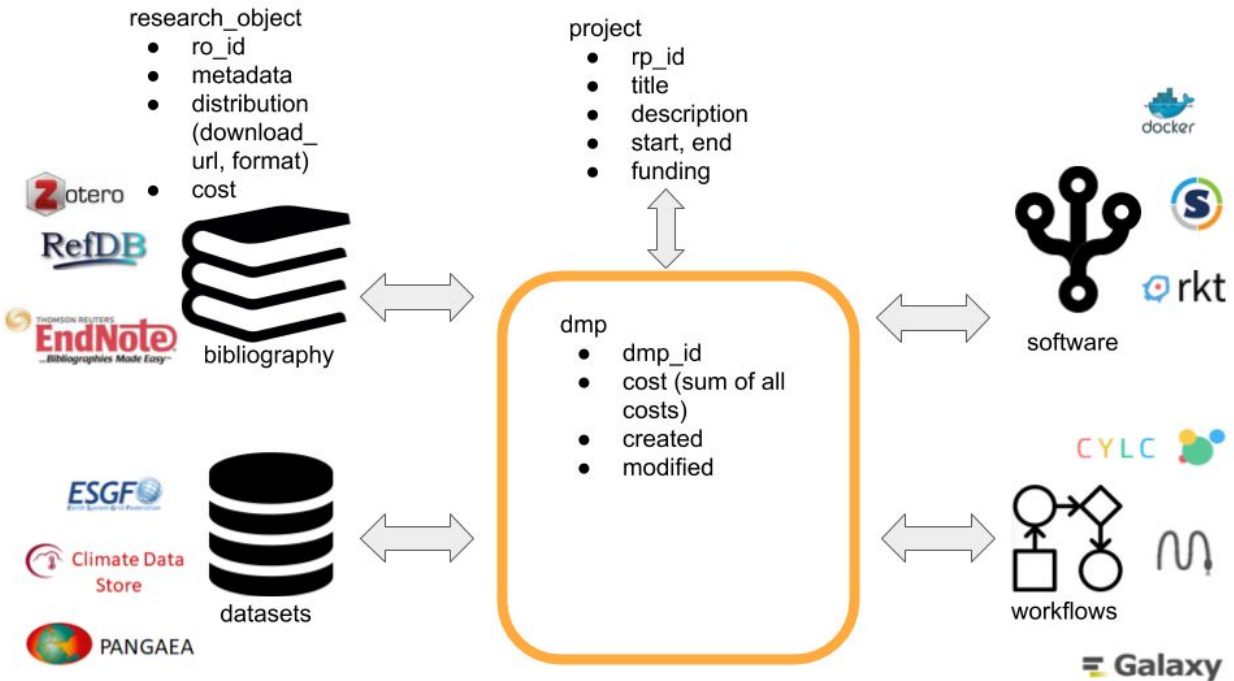
Why is this task important?

Being able to include sufficient detail of the data management workflow in the maDMP would: (a) provide an accurate description of the data management process in the DMP, (b) make it easier for other projects to adopt this workflow and create a similar one. We can imagine the DMP as the central hub containing all the information (or links to it) relevant to managing the data for a project, so similar projects would be able to reuse some or all of the information in their own processes. This would help domains to define best-practice for data management in their domain.

What did you achieve (and how) during the hackathon?

The hackathon provided an excellent opportunity to exchange ideas with, and learn from other people working on maDMP adoption. We set ourselves the task of mapping our workflows to the maDMP. We benefit from discussions as it helped us to view the maDMP as an exchange format, so we should not expect to be able to completely map to the schema. We could view our

schema as a superset (or extension) of the maDMP. We managed to identify the mapping shown in the figure below.



How will your work continue after the hackathon?

Our goal still remains to be able to capture and drive part of the data management process from the DMP. We have managed to identify some of the mapping and will now look at expressing some example workflows in the maDMP schema. We plan to then take the DMP and use it to create and drive a data management workflow demonstrating the ability to integrate a machine-actionable DMP with a part of the data management workflow. We noticed that there is, to some degree, an overlap with the team working on Issue #12 and we will look at their work to see if there is some information that we can adopt from them.