

GenAI for Research Support: a case to DMPs pre-filling

Davide Nardi, Data Steward at TU/e, d.nardi@tue.nl

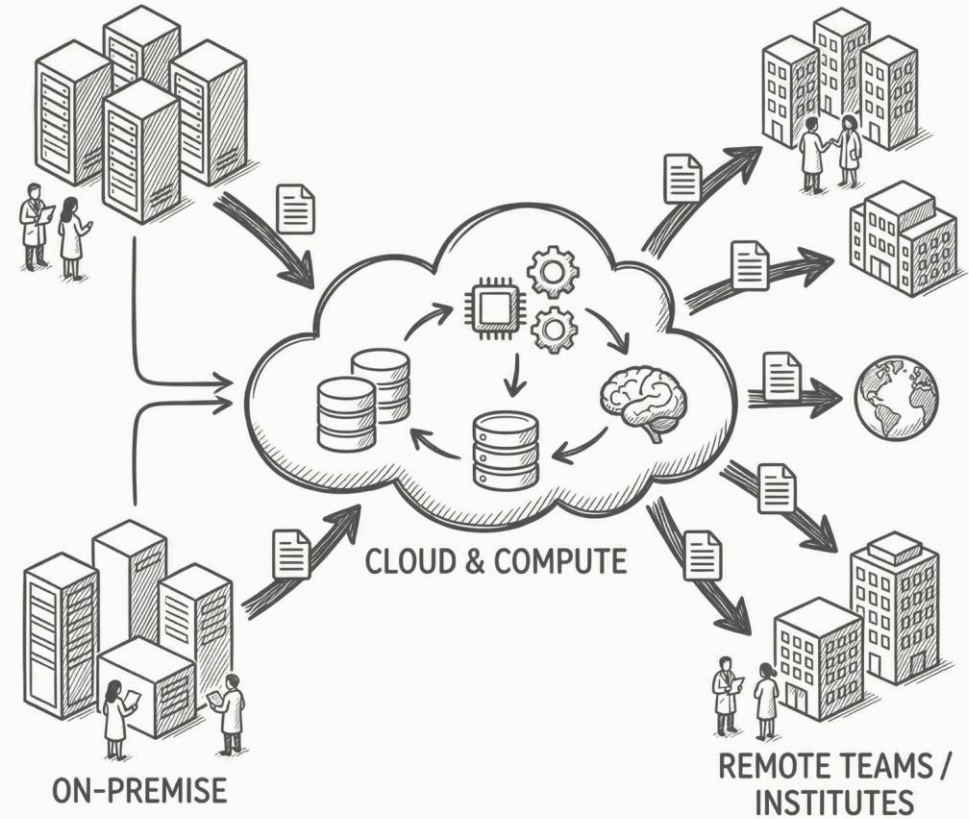
Liz Guzman Ramirez, Data Stewards Coordinator at TU/e, l.guzman.ramirez@tue.nl

Erwin Hoogerwoord, Product Owner Research Workflows team at TU/e, e.hoogerwoord@tue.nl



Context:

Pictures created with AI (Nano Banana Pro)
Catch the mistake...



**Research Data
Management**

RDM



Research Cockpit



RDI

**Research Data
Infrastructure**



Traditional DMP preparation workflow at TU/e

Researchers compile DMP via the Research Cockpit

- Data steward provides feedback
- Researchers review the DMP
- Data steward approves the DMP

- Researcher submit DMP to the funder
- Research updates the DMP during and at the end of the research

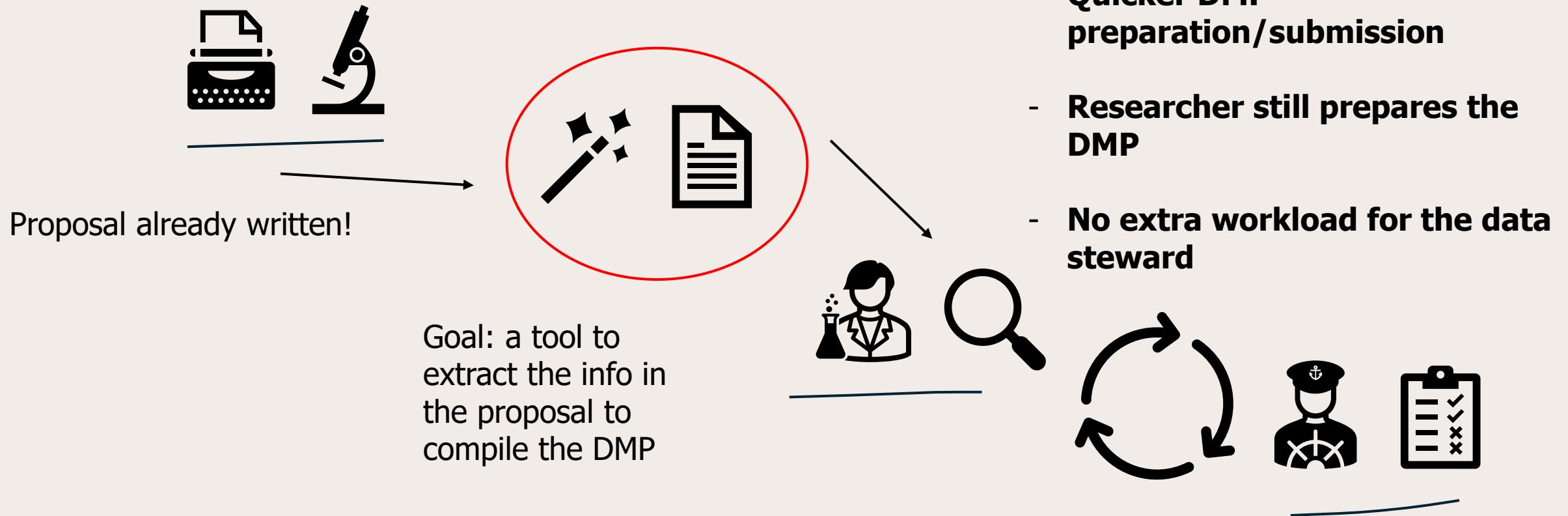


Traditional DMP preparation workflow at TU/e

Challenges

- Researchers: *Time consuming*
- Data steward: *No time to prepare DMPs for all the researchers*
- Support can come (most of the time) after the DMP is submitted

What can be done in advance?



LLM tool

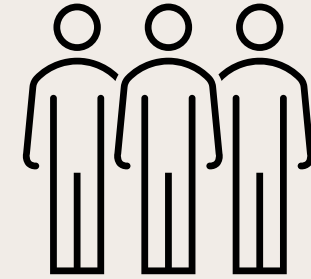


- Extract structured metadata from source data repositories
- User-defined schema specifying the desired metadata attributes (including a description per metadata attribute)
- The interface allows for quick configuration and testing of metadata attributes via a form-builder interface
- The system offers multiple extraction methods: Full mode, Smart modes, Agentic mode



User testing session

**3 participants
(only)**



Before the session

Do you think that compiling a DMP is a regular duty of your research activities?

Do you find writing a DMP a workload that you would like to avoid?

After the session

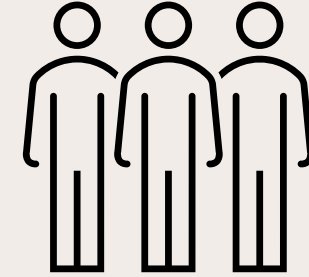
Very roughly: how many fields the LLM has correctly recognized? Provide a %

Has the review of the LLM compiled DMP eventually reduced your workload related to the standard DMP preparation?



User testing session

3 participants
(only)



Before the session

Do you think that compiling a DMP is a regular duty of your research activities?

- *"It comes and goes with assigned new projects. If a DMP has to be compiled it takes up a lot of time"*
- *"Yes. Compiling a DMP is easy. The difficult task is to implement it"*
- *"Only when the grant organization is asking it"*

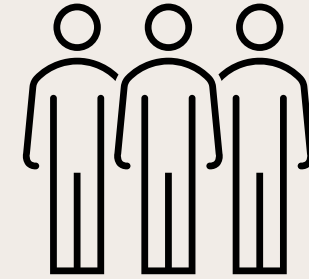
Do you find writing a DMP a workload that you would like to avoid?

- *"Yes, it is not immediately contributing to the application/project proposal"*



User testing session

3 participants
(only)



After the session

How many fields the LLM has correctly recognized? Provide a rough %

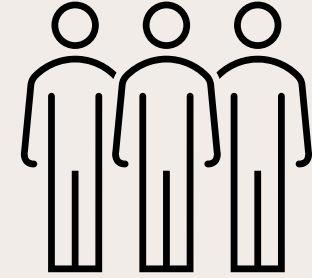
- "80%"
- *"50% correctly recognize; 5% incorrect; the other 45% from fields that were not compiled because those information were missing in the proposal"*
- "80/90%"

Has your review of the LLM compiled DMP eventually reduced your workload related to the standard DMP preparation?

- *"Yes but very limited. The questions were not difficult, and the information was already in the proposal"*
- *"Yes, as a lot of work goes into copy-paste work"*
- *"Yes, definitely"*



Future work



Extra feedback and future work

- More complex DMPs where more partners are involved -> more documents as input (not only the proposal) for an improved field recognition;
- To provide to the LLM a description of the research so that the LLM can suggest, for instance, the different alternatives for data storage, associated costs, formats, etc. -> ***Suggestion-based tool and not necessarily compiling-based.***
- LLM can also be useful to other groups (e.g., Data Steward, Ethical Board) to evaluate the applications they receive.



Thank you!

