Processes Behind the Research Data Management Life Cycle

Ville Tenhunen, James A.J. Wilson

EGI Conference, 2.11.2020

Agenda

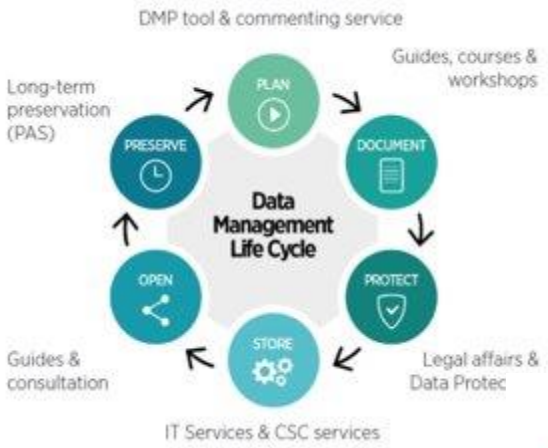
- Process mapping as a tool
- Processes behind research data management life cycle
- RDARI survey – insight to the real life
- Conclusions

EUNIS
EUROPEAN UNIVERSITY INFORMATION SYSTEMS

This presentation based on the EUNIS 2020 conference paper:

https://www.eunis.org/wp-content/uploads/2020/06/80-Tenhunen_Wilson_eunis2020_full_paper_final_2020-05-15.pdf

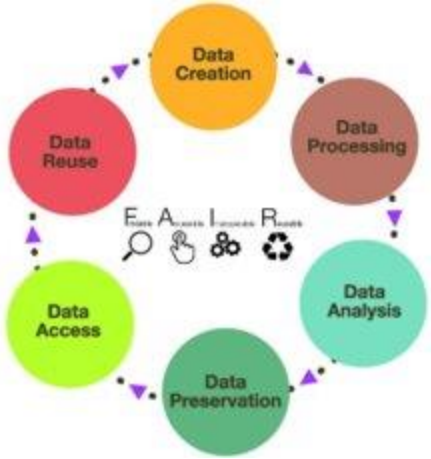
Introduction



1

3

5



2

4

Sources:
1) University of Helsinki; https://www.helsinki.fi/en/research/research-environment/research-data-management,
2) Peter Doorn, Eliane Fankhauser, Mustapha Mokrane; FAIR data in practice: From FAIRy tale to FAIR enough, Webinar, 11 December 2018
3) University of California, Santa Cruz; https://guides.library.ucsc.edu/datamanagement
4) University of Ottawa; https://biblio.uottawa.ca/en/services/faculty/research-data-management/what-research-data-management
5) Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC), vs.0.8 18 October 2020 (from the FOSTER project) https://www.eoscsecretariat.eu/sites/default/files/eosc-sria-v08.pdf

They are useful, but…

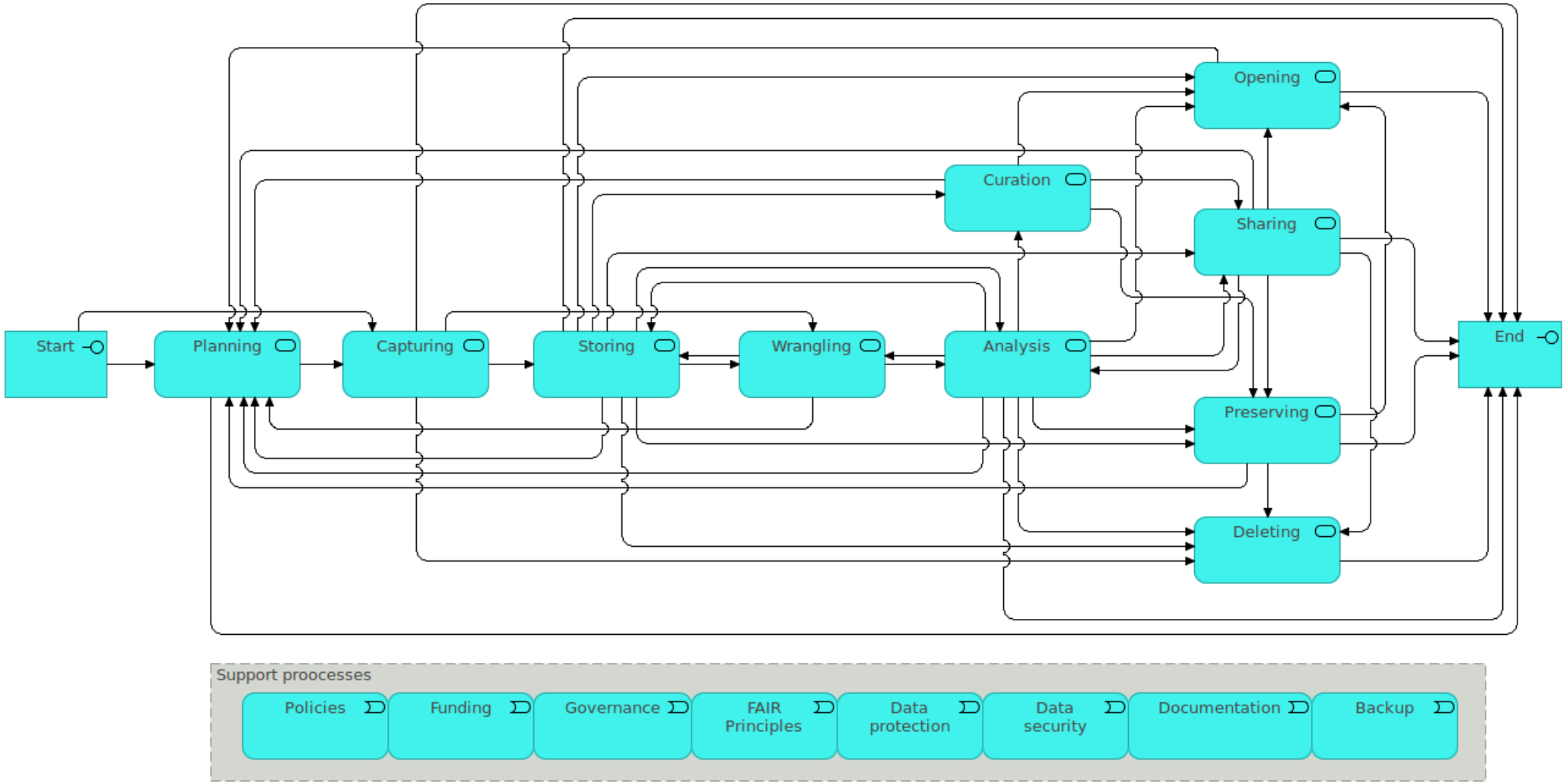
- Life cycle models give us one way to visualize the steps involved, but…
  - Cycles hide several data use cases of researchers
  - Real-world processes are not always – if ever – circles
  - Researchers work processes does not always include all life cycle phases
  - The research data does not always follow them to the next cycle

=> More informative and specific models are needed when develop real life services

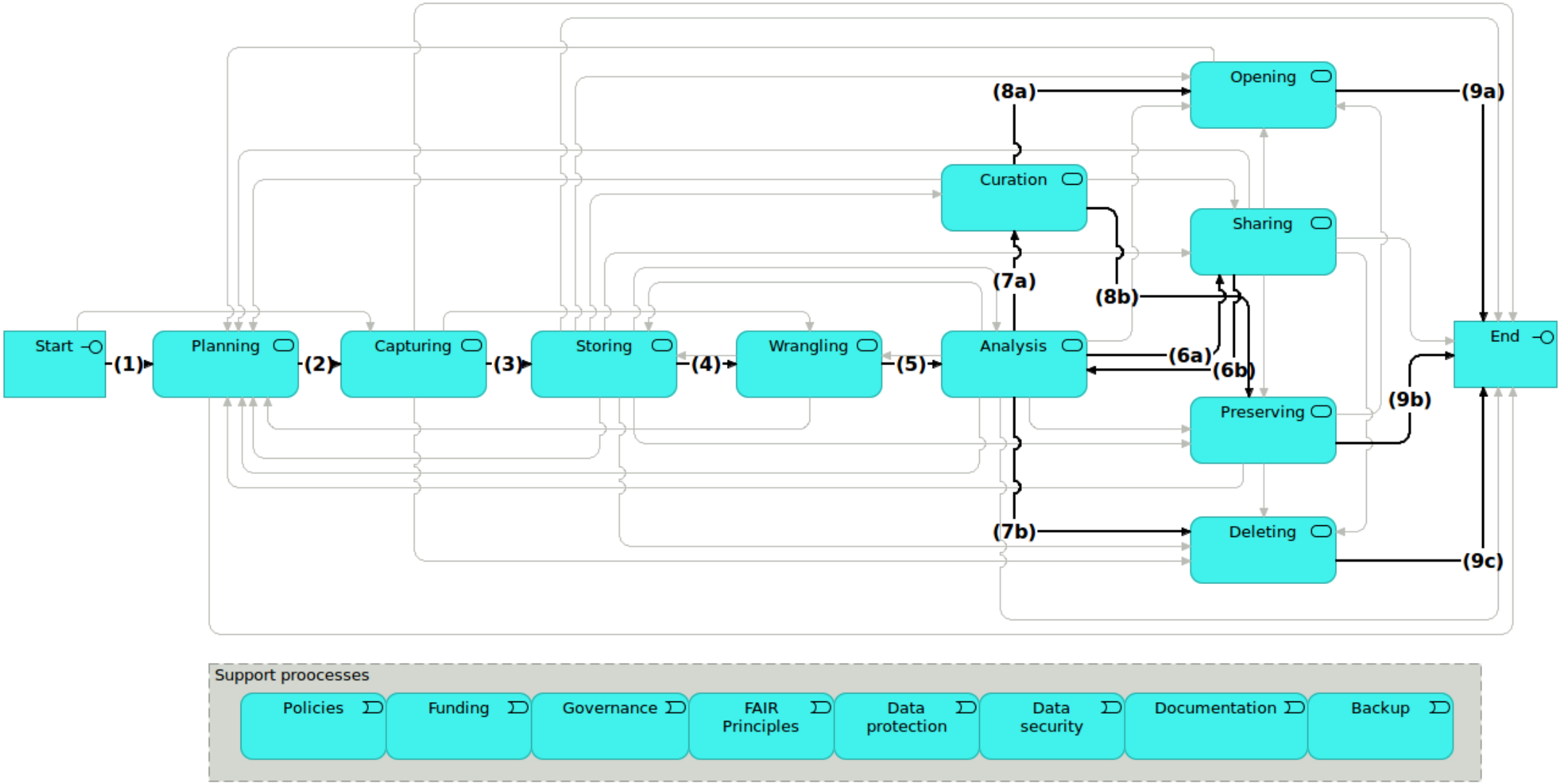
Process mapping

- The main purpose behind process mapping is to assist organizations in becoming more effective
  - In the context of the continuous improvement, organisations have to understand how each process relates to other processes and how those interactions impact quality management
  - A clear business process map allows look at whether or not improvements can be made to the current process
- Simplest process map includes processes and their relations
  - It is possible to evolve the map by adding more information like process owners, actors, process breakthrough times, delays, (lean) wastes etc.

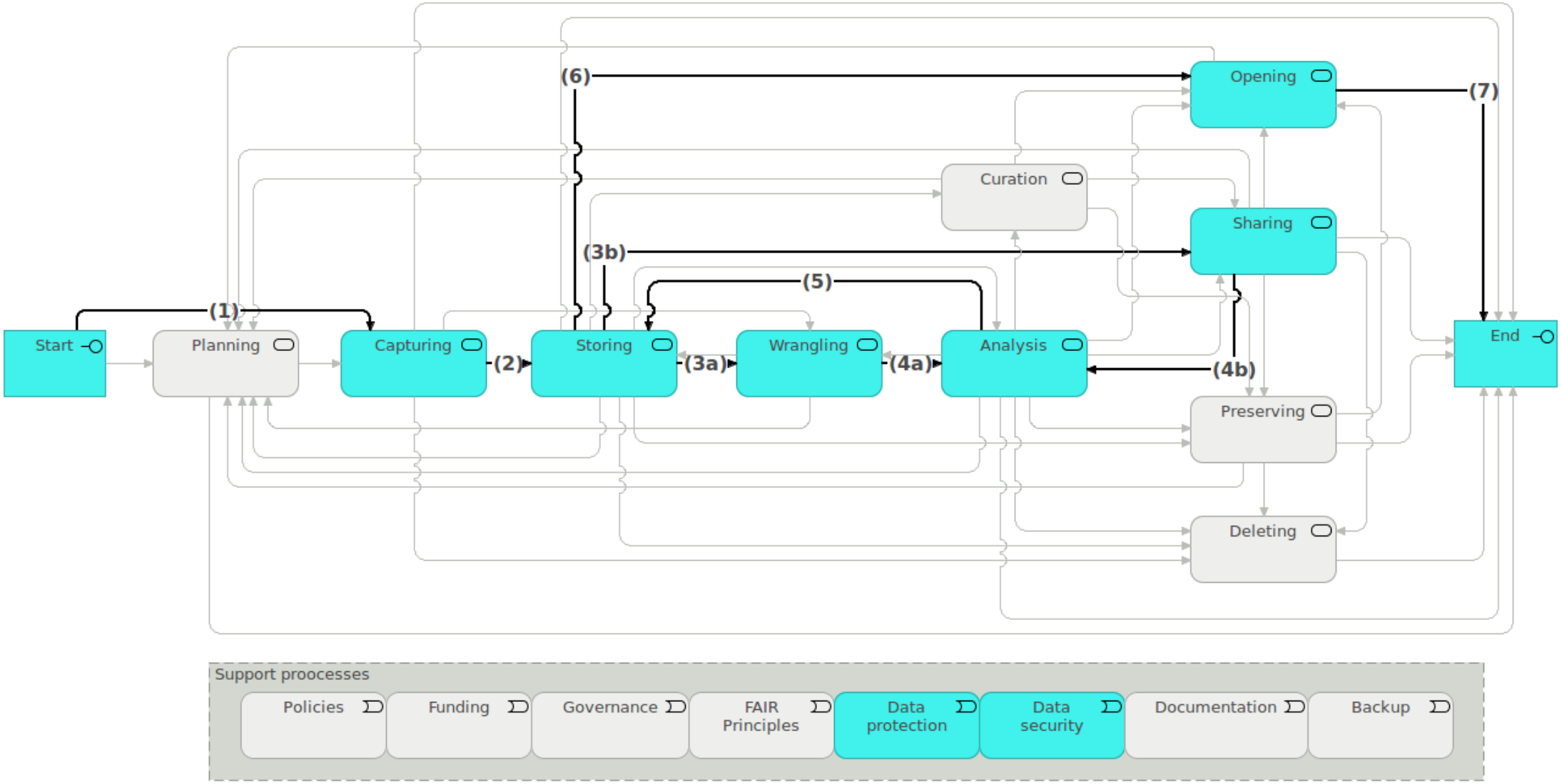
Big picture of the RDM processes



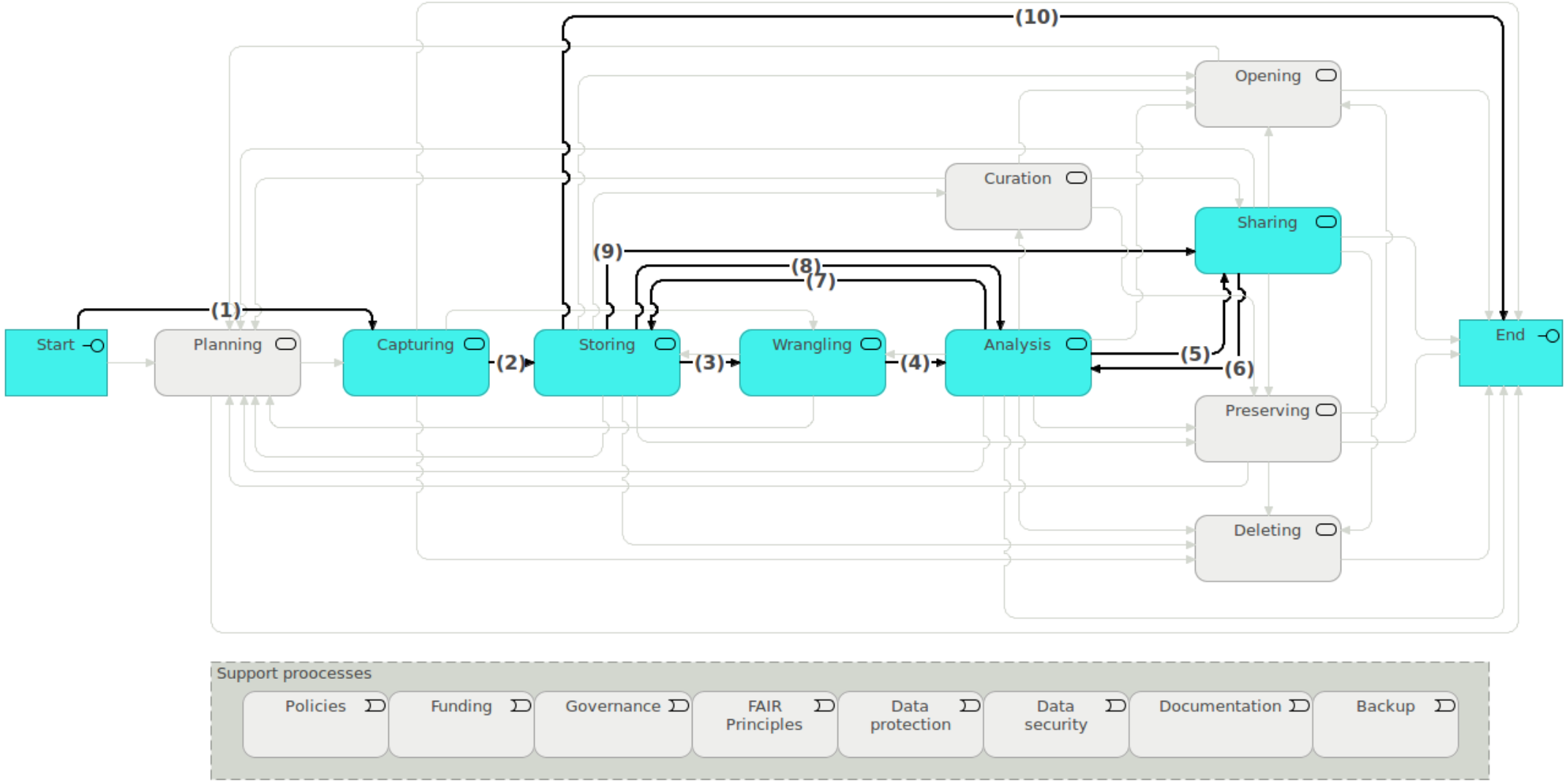
Ideal and FAIR data paths?



Data path: Fast track, not so FAIR



Data path: Loops without publishing

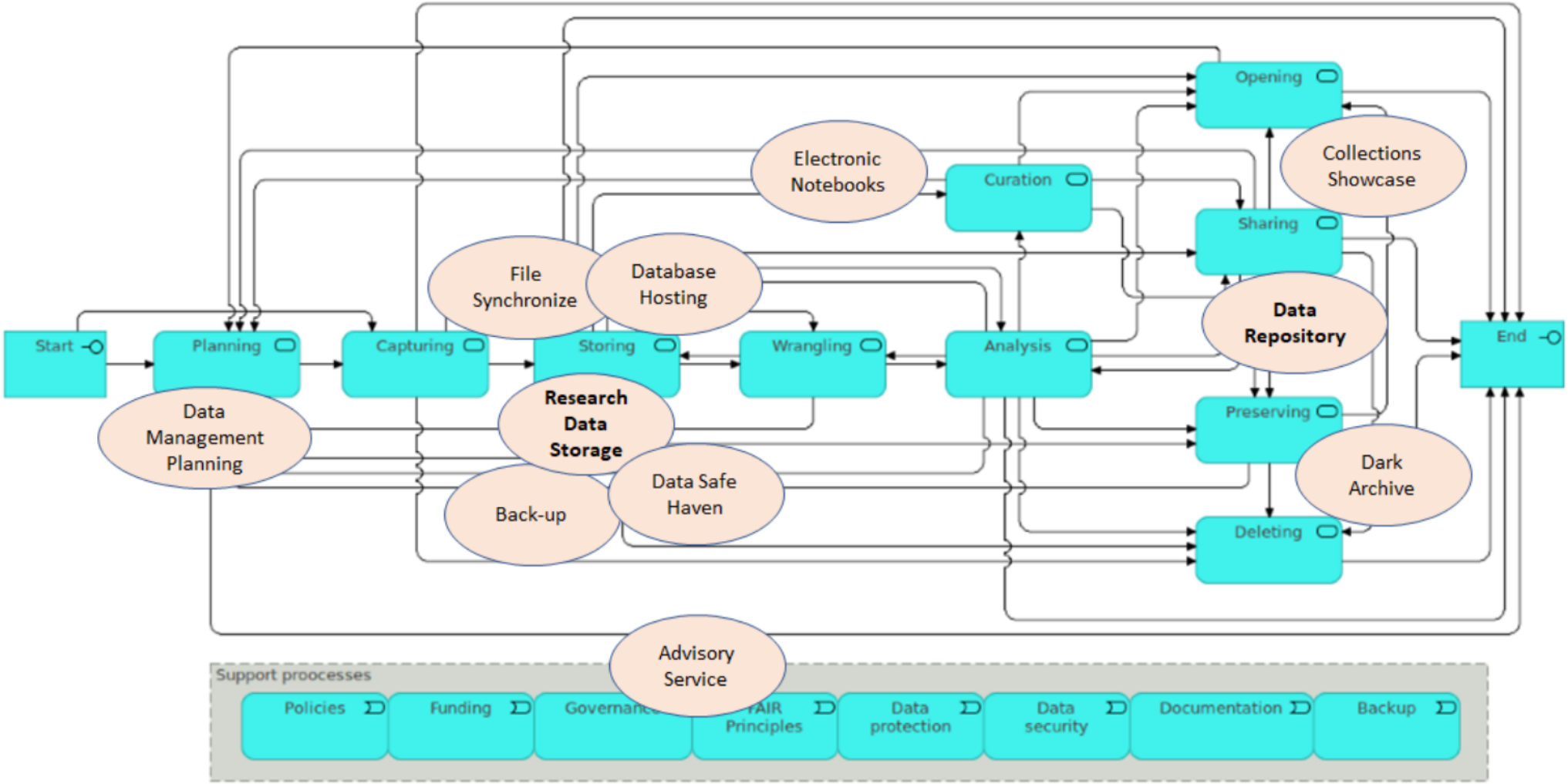


Real life insight - RDARI Survey

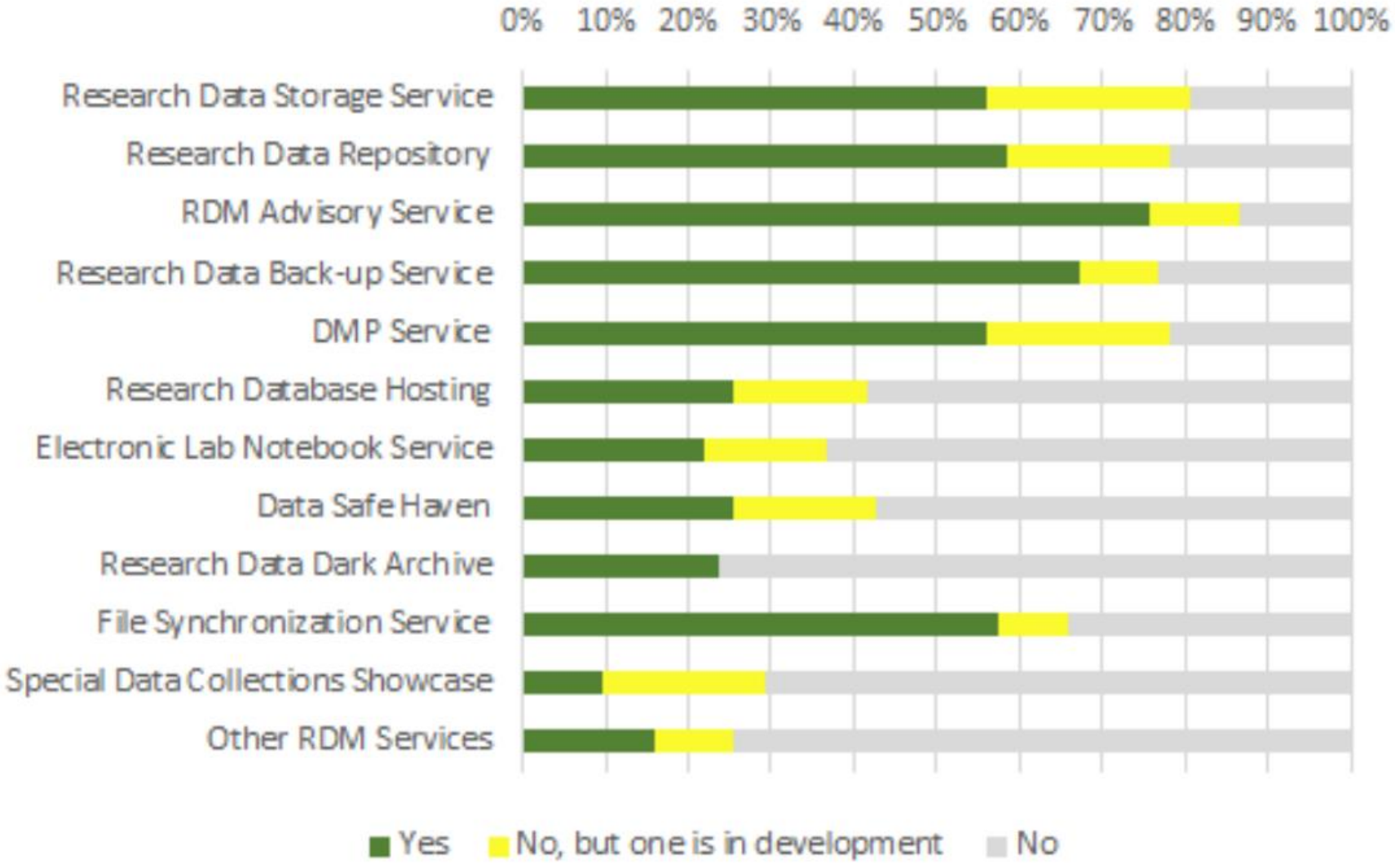
- The Research Data Architectures in Research Institutions (RDARI) Interest Group of the Research Data Alliance (RDA) Survey of Institutional Research Data Services was conducted between July and November 2019
- The aim of the survey was to build a picture of the contemporary state of research data management service at universities and other research-conducting institutions across the world, both to assist with benchmarking and also to help put people in touch with each other so they could exchange notes about their approaches (assuming they were happy to share

More information:
- Research Data Architectures in Research Institutions IG: https://www.rd-alliance.org/groups/research-data-architectures-research-institutions-ig
- Wilson, James A J; Tenhunen, Ville; Russell, Keith (2020): RDARI International Survey of Institutional Research Data Services 2019. Figshare. Dataset. https://doi.org/10.5522/04/10283540.v1

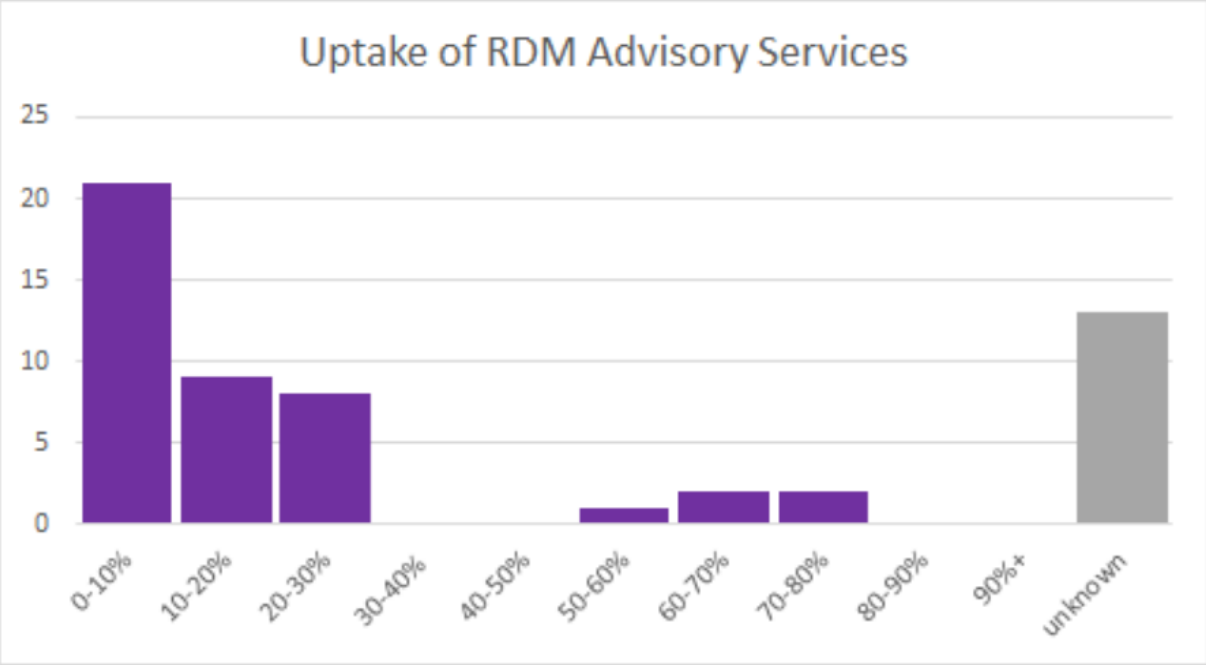
RDM Services loosely coupled to process map

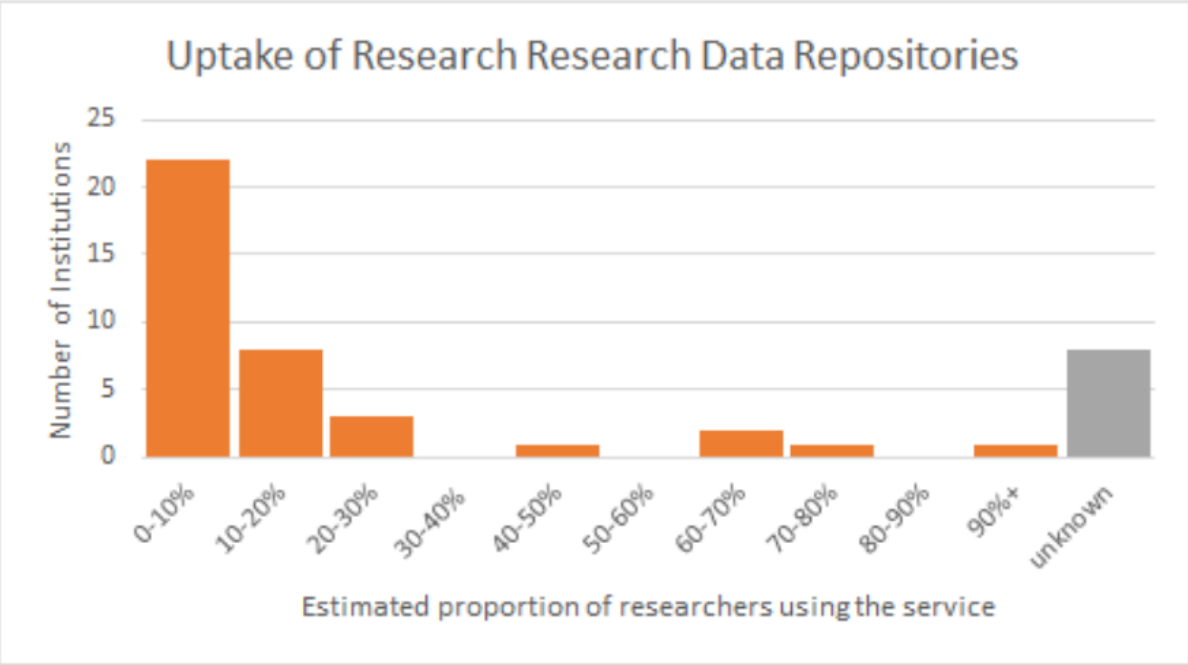


RDM Services Provided by Institutions

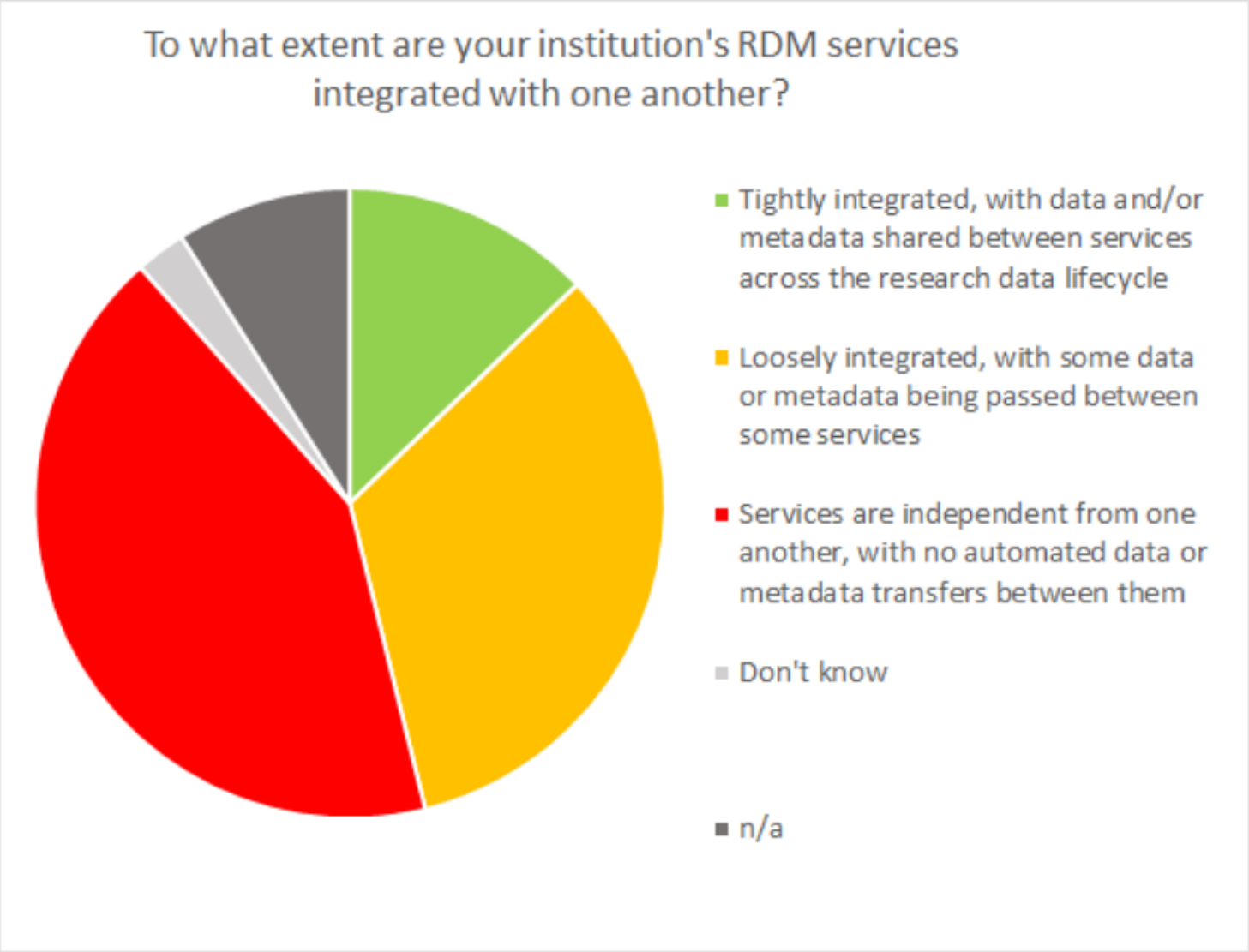


Uptake of RDM Advisory Services and Research Data Repositories





Integration of RDM services



Conclusions

- Usual research data management models give a general picture of the landscape
  - For the service development more specific tools for designing user driven data services are needed
  - Process mapping is potential option
  - Method might help to solve challenges when all research data processes are not suited to the ideal model but they might be useful and valuable for research and science
- RDARI survey gives real world data to describe services and their integrations in the research institutions
  - It also indicates there are gaps in services around the capture, wrangling, and analysis of data
  - It is understandable but noteworthy because these are important phases for the scientific results

Writers

- M. Sc. Ville Tenhunen (ORCID: 0000-0003-0217-0831) working in the EGI Foundation as Data Solutions Architect. Formerly he has worked as a team leader and project manager in the University of Helsinki more than 12 years. Last major projects has dealt with research data and its storages. Tenhunen has been also co-chair of the Research Data Architectures in Research Institutions IG of the Research Data Alliance (RDA). He is also member of the Architecture Working Group of the EOSC.

- Dr. James A.J. Wilson (ORCID: 0000-0002-8546-1142) has over ten years' experience in research data management and is currently Head of Research Data Services at UCL. He is the service manager of UCL's Research Data Storage Service and the Institutional Research Data Repository. He has formerly developed a database hosting facility and worked with a number of research teams on their research data workflows and documentation. His academic background is in the humanities. He is a co-chair of the Research Data Architectures in Research Institutions IG of the RDA.

Thank you!