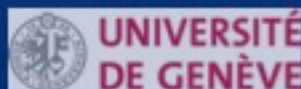


Données de la recherche: Enjeux et perspectives

20 June 2018 | HEG Genève

Presentation for "100 ans de la filière Information documentaire"

Pierre-Yves Burgi - Directeur SI adjoint



Research Data Management (RDM) is a “wicked problem” (1)

(McLeod and Childs, 2013)

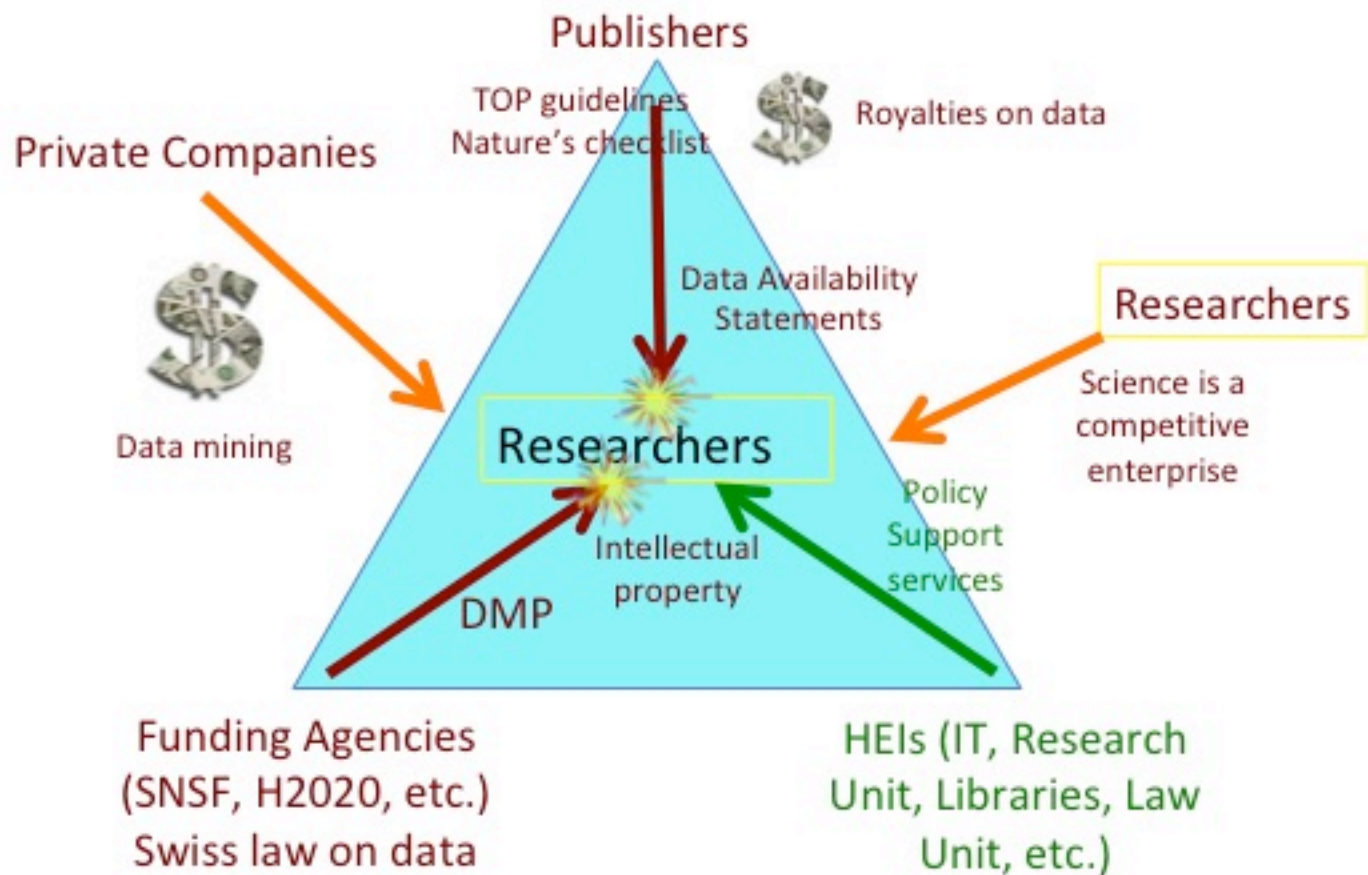
(1) ... evades easy answers and
may actually be insoluble ...

RDM's Definition

“Not exactly a delimited concept, RDM is an umbrella term for activities related to the creation, organization, structuring and naming of data; to their backup, storage, conservation and sharing, and to all actions that guarantee data security.”

(Schöpfel et al. 2018)

RDM's Actors



RDM's Stakes: (1) Data Sharing

Benefits

- Brings new questions → advances research and innovation
- Improves data integrity
- Enhances transparency and reproducibility
- Recognition from peers
→ Theory of Reasoned Action (motivations)

Issues

- Publish first is the sharing condition
- Trust in others' data → "confidence meter"
- Difficulty to interpret shared data without considerable documentation (metadata, context, software code, etc.)
- Adequate citation → standard not yet well practiced by researchers
- Loss of control over data
- Lack of knowledge on curating and archiving datasets prevents from sharing
- Making data available ≠ making data usable

RDM's Stakes: (2) Data Re-use

Benefits

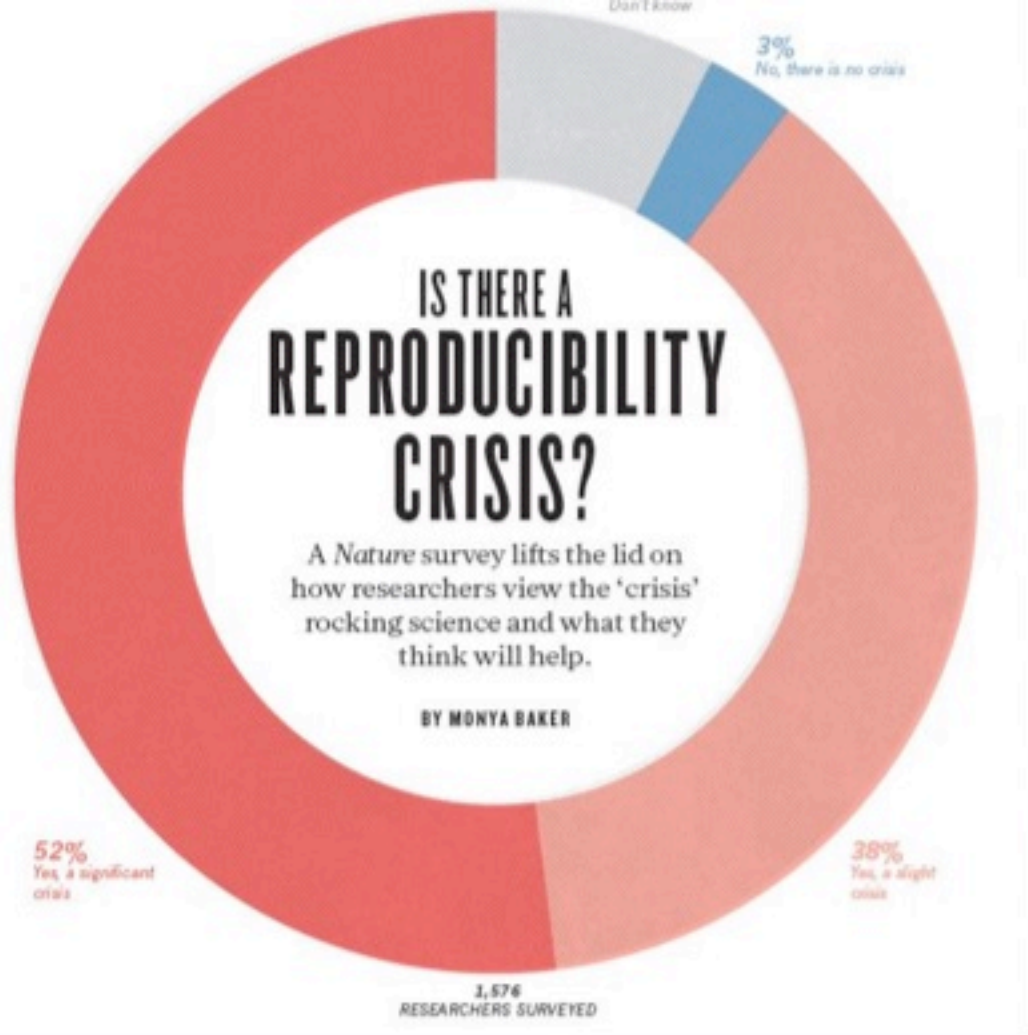
- Costs: eg. Remotely sensed data via automated equipment → easy to reuse (little embedded IP)
- New profession: data scientists → provides science as a service (SClaaS)
- Citizen science

Issues

- Fitting already-collected data to new research questions remains difficult
- Communication problems among heterogeneous communities
- Needs DMP and Domain Data Protocols of quality
- Needs FAIR compliant datasets → the "I" is complex
- Plagiarism (especially for raw data)

Data re-use and data sharing are not linked:
“Some researchers may be data-sharers but not re-users, or the other way around”

(Curty et al. 2017)



RDM's Stakes: (3) Reproducibility

More than 70% of researchers have tried and failed to reproduce another scientist's experiments, and more than half have failed to reproduce their own experiments

(Baker and Penny 2016)



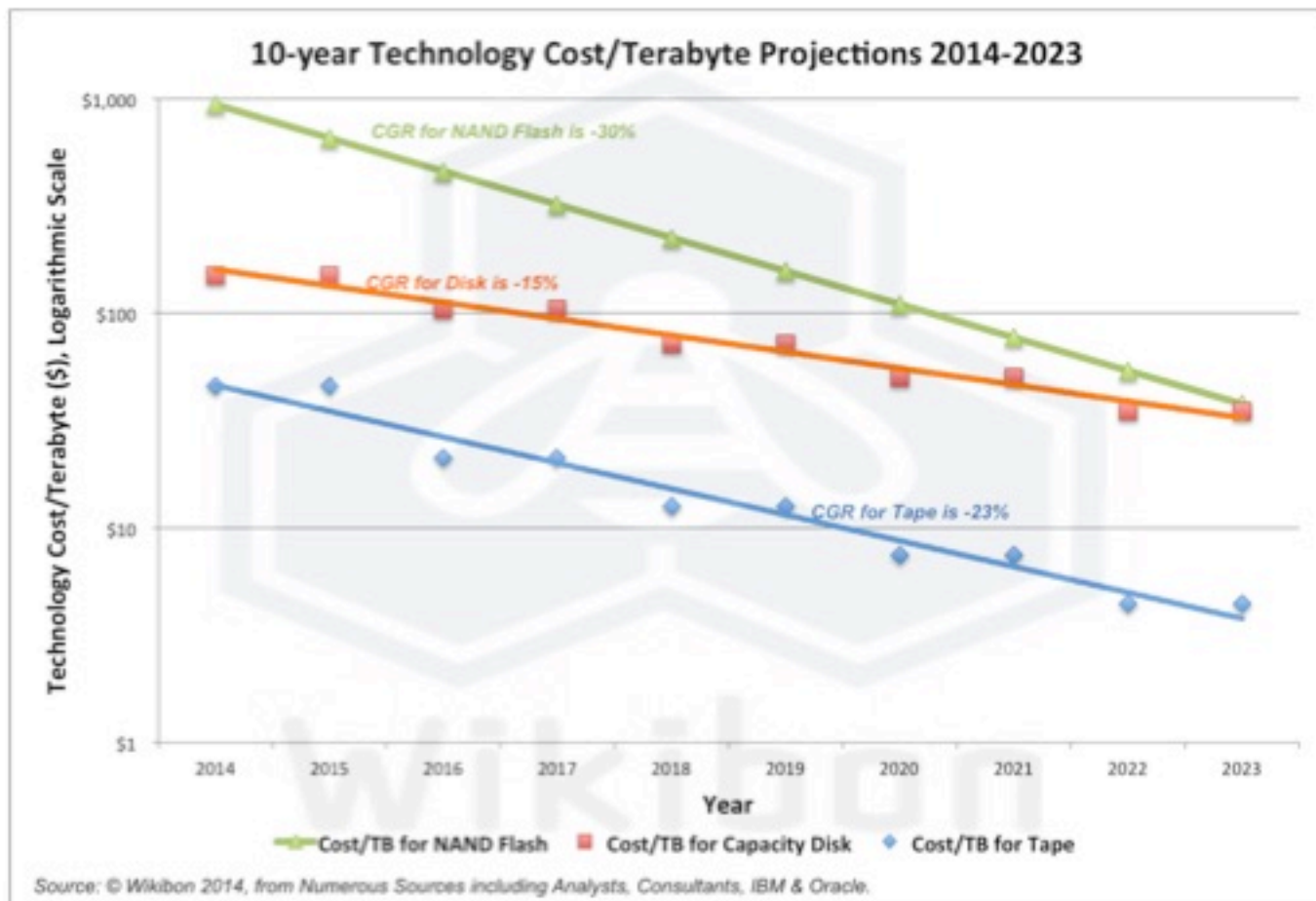
Reproducibility failures are essential to scientific inquiry

Misunderstanding about the role of reproducibility to scientific inquiry:

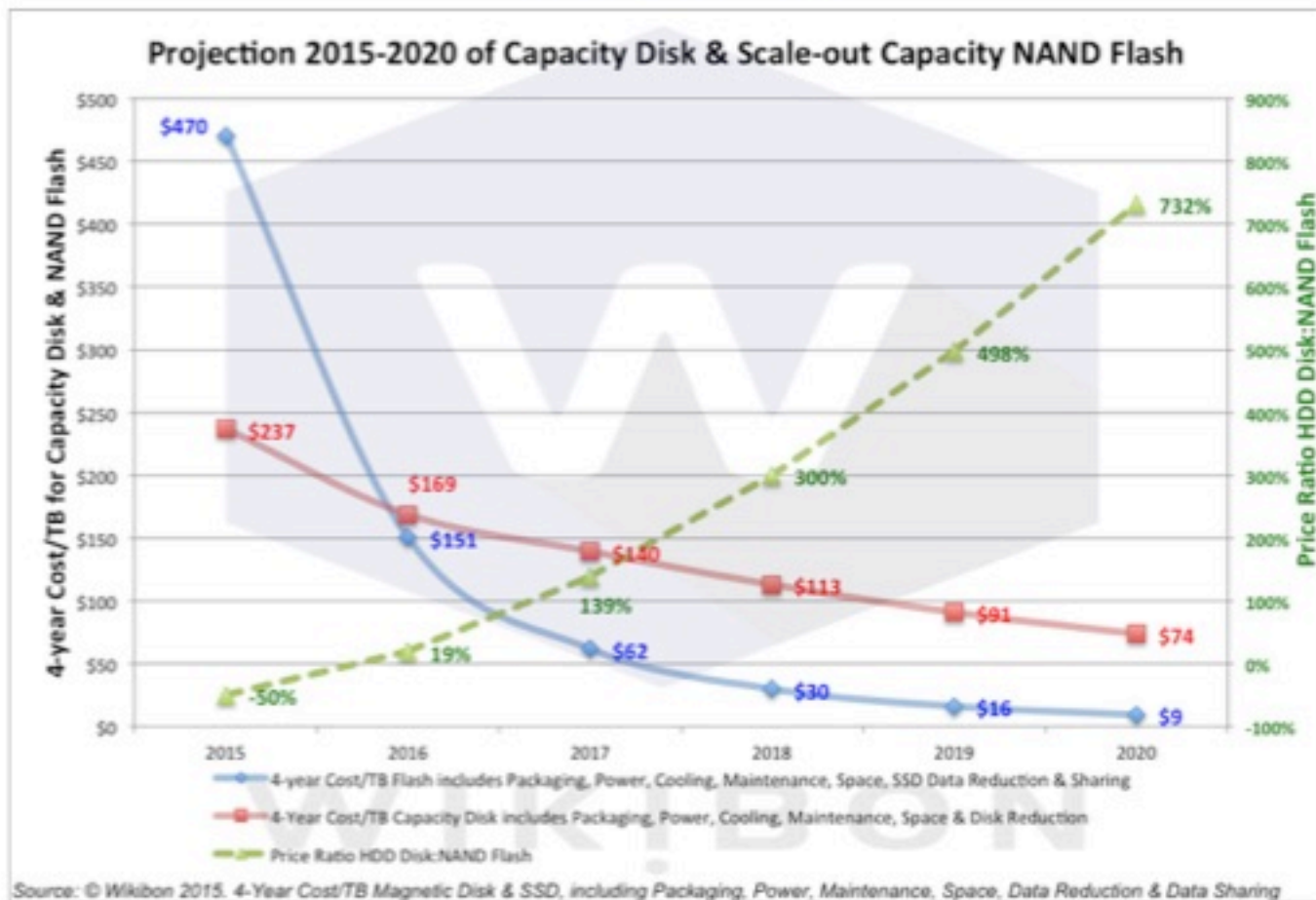
“Failures generates more reliable and enduring scientific knowledge by isolating and characterizing the crucial factors that underlie phenomena.”

(Redish et al. 2018)

Technology – storage costs (Tapes vs. HDD vs. SSD)



Technology (cont.)



Perspectives

Team's Vision of the DLCM P5 (National) Project:

We support researchers in managing their data from research proposal to publication and data archiving



Partners (phase 1)



ETH zürich

Hes·so

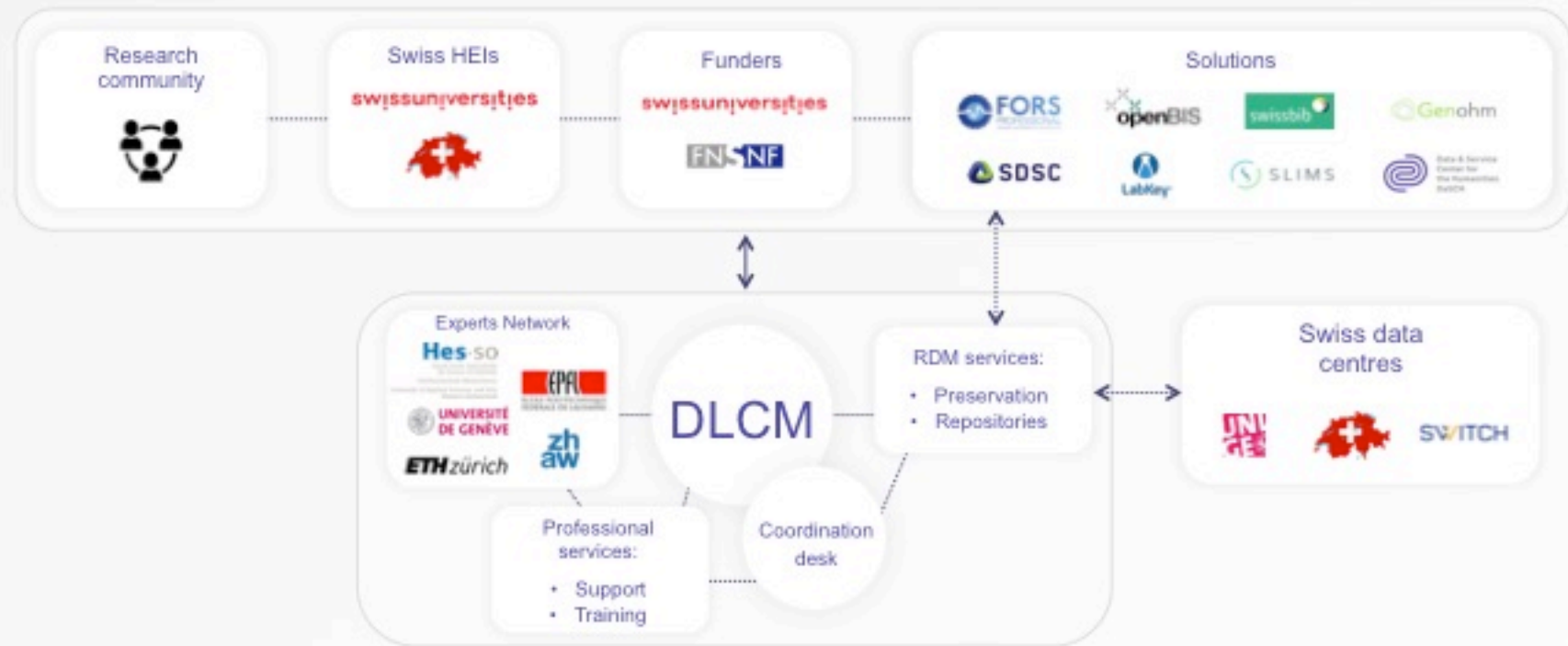
SWITCH






Swiss RDM

... creating an integrated Swiss ecosystem





DLCM Services



Coordination desk

Guaranteed answers within 2 working days in 3 different languages



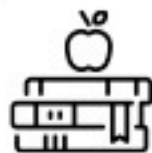
Training program

Available on site and at a distance



Pool of experts

Access to a pool of experts through our coordination desk



Knowledge base

FAQ online and interactive maps



DLCM Products

Summer 2018

2019



Preservation system

- On-premise
- Cloud
- Back-ends



Integration

- Solution providers
- Data centers
- Repositories



Data publication

- Unique Identifier (DOI)
- Referencing
- OAI-PMH



Web portal

- Submission
- Access
- Search
- Dashboard

Thank you

Contact us pierre-yves.burgi@unige.ch

Learn more on dlcm.ch