

# Research Data Management in the Cloud: Current Practices, Tools, and Challenges

Working on data stewardship? Meet your peers!

October 3<sup>rd</sup>, 2017

Break-out session

Rob Gommans

(Radboud University – Behavioural Science Institute)

Mariëtte van Selm

(University of Amsterdam & Amsterdam University of Applied Sciences)



Data storage & transport

Documentation & metadata

Controlled sharing

Data management planning

Data collection

Data carpentry & analysis

Data publication

Data archival

Medium term ( $\leq 10$  years)

Long term ( $> 10$  years)

Data registration

Data storage & transport

Documentation & metadata

Controlled sharing

Alone

With internal partners

With external partners

With students

Medium term ( $\leq 10$  years)

Long term ( $> 10$  years)

Data management planning

Data collection

Data carpentry & analysis

Data publication

Data archival

Data registration

Data storage & transport

Documentation & metadata

Controlled sharing

Data management planning

Data collection

- Alone
- With internal partners
- With external partners
- With students

Data carpentry & analysis

- Alone
- With internal partners
- With external partners
- With students

Data publication

Data archival

- Medium term ( $\leq 10$  years)
- Long term ( $> 10$  years)

Data registration

Data storage & transport

Documentation & metadata

Controlled sharing

Alone

With internal partners

With external partners

With students

Alone

With internal partners

With external partners

With students

Data collection

Data carpentry  
& analysis



## What is it about and what NOT

- **Yes**, collaboration and sharing with external researchers while conducting your research.
- **No**, registering and archiving your data after publication for re-use purposes.
- **No**, ready-made solutions for collaboration and sharing, no one size fits all.



## Challenges

- Project management
  - How to manage data properly?
  - How to keep things in sync?
  - How to deal with versioning and file change conflicts?
  - How to provide access?
- Data protection
  - How to prevent data loss?
  - How to prevent unauthorized access (data leakage)?
  - Backup / Data retention
  - Reliability / Trustworthiness cloud storage provider
- Privacy and confidentiality
  - Informed consent for collaborating and sharing
  - Data processing agreements
  - Privacy and data protection laws USA vs EU
  - Institute and national policies and regulations
- Technical
  - How to ensure sufficient storage space?
  - How to limit data transfers between locations?
  - Interoperability with other systems / tools



## Possible solutions

- Many cloud storage providers





## Possible solutions

- Some online project management tools



## In the ideal world...

- Each research project registered in project management system (PMS)
- Scalable server- or cloud-based storage for PMS
- Local (read-only) copies of data in PMS
- Automated synchronization between local and PMS
- Automated encryption
- Check-in / Check-out system
- Build-in versioning
- Access control per collection (subfolder) / data object
- Master User access
- Possibility to add (custom) meta-data and comments
- Possibility to branch/fork
- Possibility to deposit final version (after branch/fork)
- Connectivity with existing (archiving) systems
- Advanced reporting at the project and organizational level



## Discussion / Examples

- How do you or your colleagues collaborate with external partners?
- What is offered and allowed in your organization?
- What other challenges do we face?
- ...



## Thank you for your attention!

- Please contact us if you have any questions.
- Rob Gommans, [r.gommans@bsi.ru.nl](mailto:r.gommans@bsi.ru.nl)
- Mariëtte van Selm, [m.vanselm1@uva.nl](mailto:m.vanselm1@uva.nl)

