

Data Management Support at Leiden University

Peter Verhaar | Leiden

7 september



Universiteit
Leiden

Bij ons leer je de wereld kennen

Leiden University : Data Management Policy

BEFORE

- ✓ Write a data management plan (DMP)

DURING

- ✓ Secure storage : integrity, availability, confidentiality

AFTER

- ✓ Data must remain available for **10 years**
- ✓ Data should be 'FAIR'
- ✓ Provide documentation, metadata, software for reuse
- ✓ Archive data in a certified archive

Obstacles for researchers

- ❑ Lack of knowledge and skills
- ❑ Multitude of requirements and policies
- ❑ Unclear career incentives

Obstacles for research support

- ❑ Differences between disciplines
- ❑ Dynamic area of expertise
- ❑ Combination of local, national and international services

University-wide Data Management programme

- Three projects:
 1. Formulation of data management policy
 2. Training and education
 3. Development of catalogue of data management services

- Programme led by department of Academic Affairs; Project team includes employees from Leiden University Libraries and ICT Shared Service Centre
- Sounding board with researchers from all faculties
- Faculties and research institutes have actively provided feedback on data management policy

Catalogue of data management services

- Description of ca. 60 characteristics of DM services
- Description model largely based on policies from funders and publishers, evaluation protocols and code of conduct
- Concepts and terminology borrowed from models: Research data Life Cycle (UKDA), Data Curation Continuum (ANDS), Data publication pyramid and Curation Life Cycle Model (DCC)
- At present, ca. 50 services have been described
- Descriptions are all in English

Leiden Research Data Information Sheets

Search this site...

Leiden Research Data Information Sheets About this website Help desk Manual

The services in the list below can be used before (B), during (D) and after (A) research projects. Their suitability during these three general phases is indicated using the symbols below.

✓ Meets all requirements ? Partly meets all requirements ✗ Does not meet all requirements ■ Not applicable

Local	National	International
B D A ✓ Bulkstorage	B D A 4TU.ResearchData	B D A B2DROP
? ✗ Dataopslag Cell Observatory	? ✗ BeeHub	? ✗ B2FIND
✓ ✗ Departments	✓ CLARIN INL Portal	? ✗ B2SAFE
✓ Template DMP Leiden	✓ DANS Dark Archive	? ✗ B2SHARE
✓ ? Virtual Research Environments	✓ De Digitale Koepel (Meertens Instituut)	? ✗ B2STAGE
✗ Workgroups	✓ Dutch Dataverse Network (DDN)	✓ ✗ DataVerse Network
	✓ EASY	? DataFirst
	✓ EDNA	✓ DCCD
	Essentials 4 Data Support	✗ DDMoRe - Drug Disease Model Resources
	✓ NWO datamanagementplan	✓ DMP Online
	✗ SURF Data Archive	? Dryad
	? ? SURFdrive	? ✗ Figshare
	✗ SURFResender	✓ ICPSR
	✓ Surveydata Nederland	✗ ✗ Infrared Space Observatory data archive
	✓ The Language Archive	MANTRA
		✗ MycoBank
		✓ NESSTAR
		Open Machinery Learning
		✗ OpenMIRL
		✓ ? SeaDataNet
		TalkBank
		World Data Center for Earth and Environmental Sciences

<https://vre.leidenuniv.nl/vre/lrd/>

Information sheets

Data about, amongst other things

- Organisation
- Legal aspects
- Storage
- Target groups (disciplines and faculties)
- Phase in the research process (before / during / after)
- Funding
- Depositor and user agreements
- Metadata schemas
- Supported formats
- Costs
- Preservation strategy

Figshare		Edit this information sheet All information sheets
General information		
URL	http://figshare.com/	
Description	Figshare is a repository where users can make all of their research outputs available in a citable, shareable and discoverable manner. http://figshare.com/ Figshare allows users to upload any file format to be made visualisable in the browser so that figures, datasets, media, papers, posters, presentations and filesets can be disseminated in a way that the current scholarly publishing model does not allow.	
Organisation	Figshare http://figshare.com/contact Digital Science Macmillan Glasshouse Building 2 Trematon Walk Wharfedale Road London NW1 9SR (t) +44 (0) 20 7418 5573 (f) +44 (0) 20 7014 4180	
Type of service	Storage facility	
Legislation	UK law	http://figshare.com/terms
Usage and appreciation		
Support organisation	e-mail: info@figshare.com twitter: @figshare facebook: facebook.com/figshare http://figshare.com/contact	
Context		
Stage in the research project	During, After	
Position within the research process	4. Preserving data, 5. Giving access to data, 6. Re-using data	
Domain	1. Private Research Domain, 3. Shared Research Domain, 5. Public Domain	
Type of data	1. Raw data and data sets, 2. Data collections and structured databases, 3. Processed data and data representations, 4. Publications with data	
Data curation	5. Store, 6. Access, use and re-use	
Data classification	Klasse 3 voor openbare informatie, Klasse 2 voor interne informatie [voor een beperkte groep]	
Administrative information		
Funding	Figshare LLP has received investment from Digital Science. figshare is not owned by Digital Science and operates as an independent company. Digital Science is part of the Holtzbrinck Publishing Group. https://figshare.zendesk.com/hc/en-us/articles/201954013-Who-we-are-and-how-we-are-funded Digital Science has also acted as an incubator for figshare since its inception, providing invaluable business and strategic advice, particularly with regards to sustainability. Our sustainability model involves providing services for publishers such as PLOS, Wiley and Nature, as well as institutions such as Stockholm University and the University of Melbourne.	

Leiden Research Data Information Sheets

Search this site...

Leiden Research Data Information Sheets About this website Help desk Manual

Lists
All information sheets
View by
Faculty
Phases in the research project
Research data Life Cycle (UKDA)
Data Curation Continuum
Data publication pyramid
Curation Life Cycle Model (DCC)
Policies
Funders
Standard Evaluation Protocol
Publishers
European Data Protection Directive
Leiden University
Add a new sheet
All Site Content

UKDA Research data Life Cycle

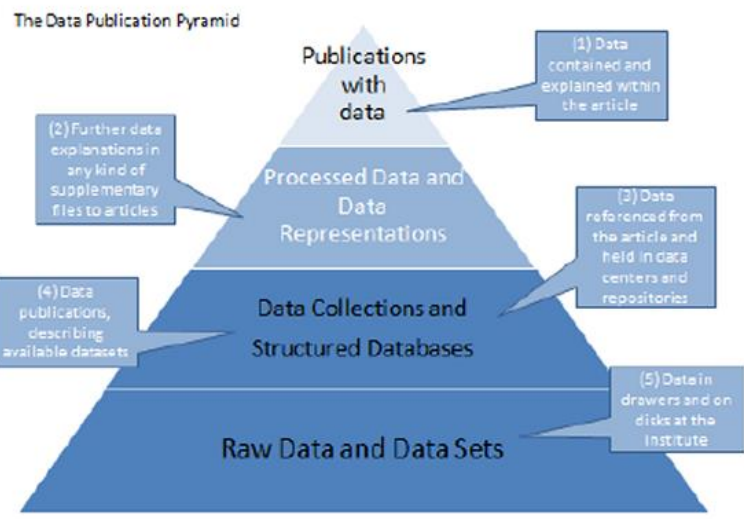
Het UKDA Life cycle model is geschikt om de datamanagement voorzieningen te positioneren ten aanzien van het onderzoeksproces: welke voorziening(en) staat/staan de onderzoeker in een bepaalde fase van het onderzoek ter beschikking. Het helpt bovendien om bij het inventariseren en beschrijven van de voorzieningen de activiteiten te benoemen waarvoor de voorziening kan worden gebruikt. Dit kunnen activiteiten uit meerdere fasen zijn.

De datacyclus is een aanvulling op de onderzoekscyclus. In de data levenscyclus worden onderzoeksprocessen en -activiteiten geïntegreerd met concepten als het cureren, preserveren, publiceren en delen van data. Het is gebaseerd op de gedachte dat de meeste data een langere levensloop hebben dan het project dat de data genereert: onderzoekers verken zelf verder met hun data in follow-up onderzoek of andere onderzoekers hergebruiken de data. Hiervoor is het van belang dat de data goed worden beheerd, gecensureerd en toegankelijk gemaakt. Een lijst van veel voorkomende activiteiten voor databeheer staat in bijlage 1. De data levenscyclus helpt, zoals Jake Carlson stelt, om al deze activiteiten op een visueel aantrekkelijke manier weer te geven.

Social and Behavioral Sciences

✓ Meets all requirements ? Partly meets all requirements ✗ Does not meet all requirements □ Not applicable **Recommended by this faculty**

Before	During	After
✓ DMP Online (International)	? B2DROP (International)	✗ B2FIND (International)
□ Essentials 4 Data Support (National)	? B2SAFE (International)	✗ B2SAFE (International)
□ MANTRA (International)	? B2SHARE (International)	✗ B2SHARE (International)
✓ NWO datamanagementplan (National)	? B2SHARE (International)	✓ DANS Dark Archive (National)
✓ Template DMP Leiden (Local)	? B2STAGE (International)	? DataFirst (International)
	? BeeHub (National)	✓ EASY (National)
	✓ Bulkstorage (Local)	✗ Figshare (International)
	✓ Data Verse Network (International)	✓ ICPSR (International)
	✓ Departments (Local)	✓ NESSTAR (International)
	✓ Dutch Dataverse Network (DDN) (National)	✗ OpenfMRI (International)
		✗ SURF Data Archive (National)
		✓ TalkBank (International)

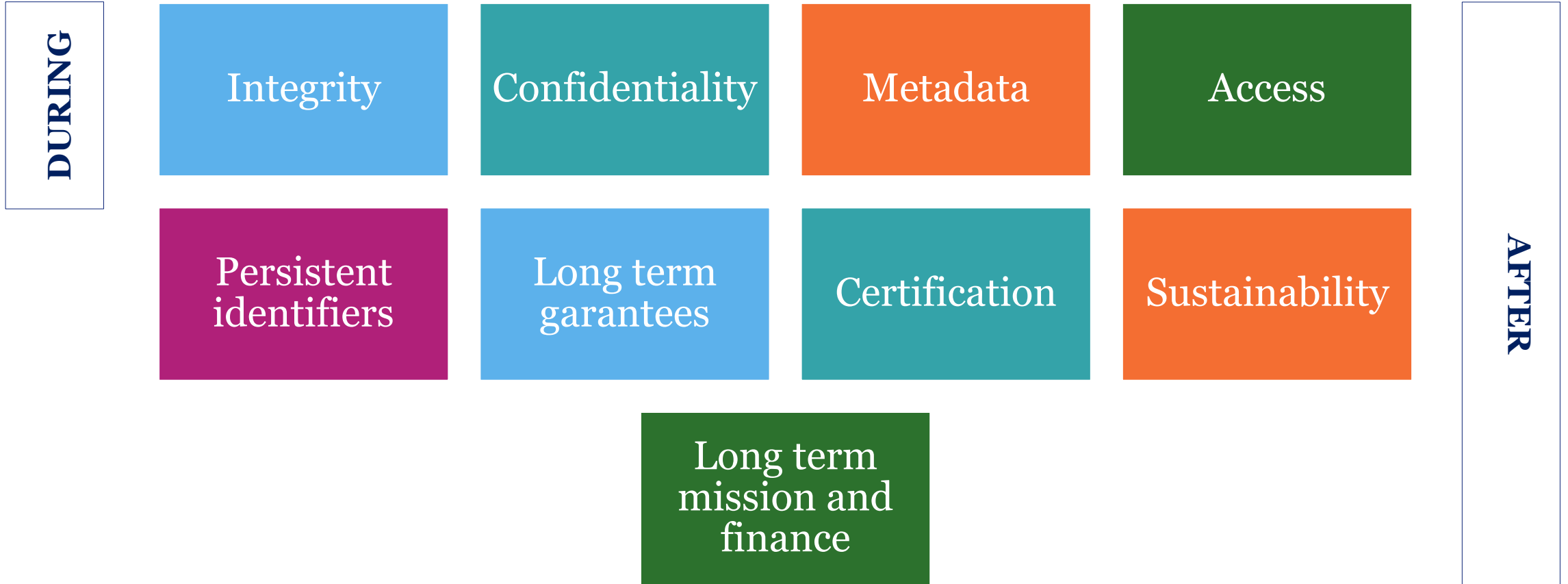


Research data Life Cycle: 3. Analysing data

✓ Meets all requirements ? Partly meets all requirements ✗ Does not meet all requirements □ Not applicable

Local	National	International
✗ Bulkstorage	✓ ? Dutch Dataverse Network (DDN)	□ ? B2DROP
? ✗ Dataopslag Cell Observatory	□ Essentials 4 Data Support	? ✗ B2SAFE
✓ ✗ Departments	✓ ✗ SURF Data Archive	? ✗ B2STAGE
✓ ? Virtual Research Environments	? ? SURFdrive	□ ? Dryad
✓ ✗ Workgroups	✓ ✓ The Language Archive	✗ ✗ Infrared Space Observatory data archive
		✓ ? SeaDataNet
		✓ ? World Data Centre for Soils (WDC-Soils)
		✓ ? Zenodo

Evaluation of services



National

B	D	A	
■	■	✓	4TU.ResearchData
■	?	✗	BeeHub
■	✓	✓	CLARIN INL Portal
■	■	✓	DANS Dark Archive
■	✓	✓	De Digitale Koepel (Meertens Instituut)
■	✓	?	Dutch Dataverse Network (DDN)
■	■	✓	EASY
■	■	✓	EDNA
■	■	■	Essentials 4 Data Support
✓	■	■	NWO datamanagementplan
■	✓	✗	SURF Data Archive
■	?	?	SURFdrive
■	✓	✗	SURFfilesender
■	✓	✓	Surveydata Nederland
■	✓	✓	The Language Archive

International

B	D	A	
■	?	?	B2DROP
■	■	✗	B2FIND
■	?	✗	B2SAFE
■	?	✗	B2SHARE
■	?	✗	B2STAGE
■	✓	✗	Data Verse Network
■	■	?	DataFirst
■	✓	✗	DCCD
■	✗	✗	DDMoRe - Drug Disease Model Resources
✓	■	■	DMP Online
■	■	?	Dryad
■	?	✗	Figshare
■	■	✓	ICPSR
■	✗	✗	Infrared Space Observatory data archive
■	■	■	MANTRA
■	■	✗	Mycobank
■	✓	✓	NESSTAR
■	■	■	Open Machinery Learning
■	■	✗	OpenfMRI
■	✓	?	SeaDataNet
■	■	✓	TalkBank
■	✓	?	World Data Centre for Soils (WDC-Soils)
■	✓	?	Zenodo

✓ Meets all requirements ? Partly meets all requirements

✗ Does not meet all requirements ■ Not applicable

Useful information for researchers:

- Which services can I use during the various stages of my research project?
- Which services are suitable for my discipline?
- Which services adhere my institutions data management policy?

Training and education

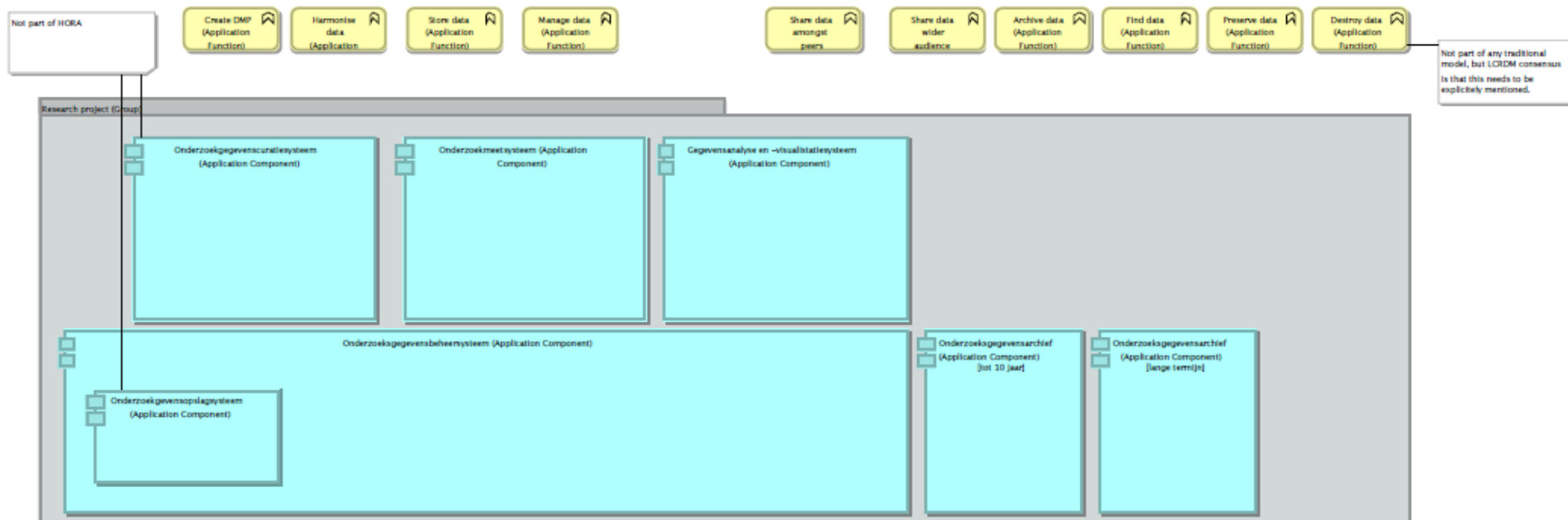
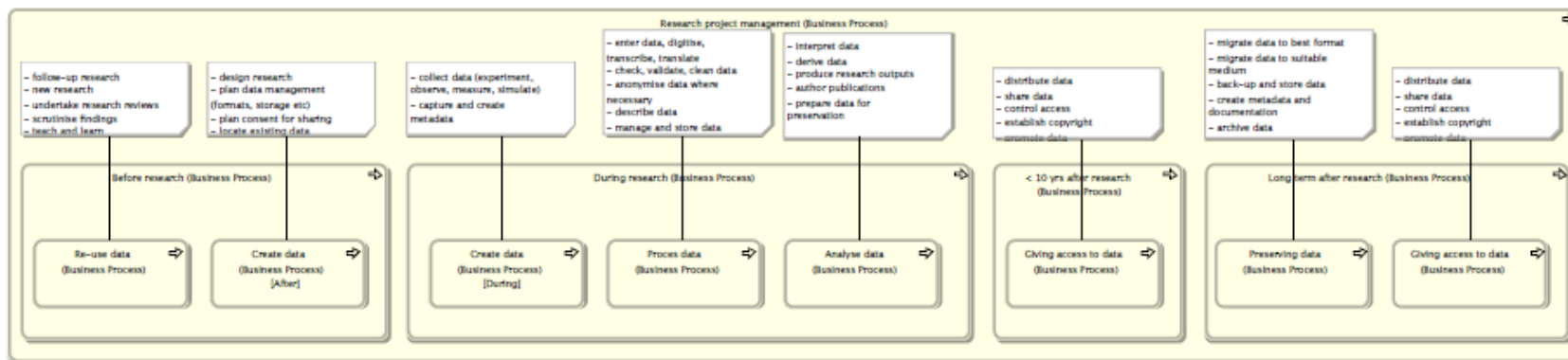
- Tailor-made interactive data management courses, for groups of PhD students, postdocs and other researchers
 - Sticks and carrots of data management planning, the principles of good data management and best practices in disciplines
 - Practical experiences with writing DMP
 - First courses taught to researchers at LU Centre for Linguistics and Leiden Academic Centre for Drug Research
- General introductory course “Data Management Policies and Planning”, open to all employees
- Many presentations and workshops on demand
- Educational materials



Next steps

- Faculties develop separate implementations of the central DM policy
- The implementation phase has a duration of 3 years (2016-2019)
- Catalogue of DM services will be developed further, in collaboration with national and international parties
- Development of local services
 - Collection and storage of data management plans
 - Encourage registration of data sets in CRIS
 - Promote the use of ORCID author identifiers

Architecture



Important principles

- Pillars of data management implementation: policy, infrastructure, education
- Combination of top-down and bottom-up approach
- Stakeholder involvement
- Emphasis not exclusively on requirements but also on benefits
 - Reporting becomes easier
 - Increase impact
 - Less duplication of efforts
 - Reuse of data

Questions?

<http://blogs.library.leiden.edu/researchdata/>

<https://vre.leidenuniv.nl/vre/lrd>

p.a.f.verhaar@library.leidenuniv.nl

datamanagement@library.leidenuniv.nl

Project team:

Laurents Sesink (Project Manager; Leiden University Libraries)

Michelle van den Berk (Leiden University Libraries)

Johan DeTollenaere (Science Faculty)

Floor Frederiks (Department of Academic Affairs)

Fieke Schoots (Leiden University Libraries)

Hans van Eck (ICT Shared Service Centre)

Peter Verhaar (Leiden University Libraries)