

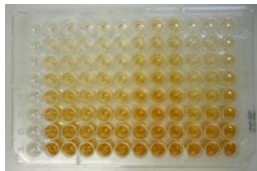


# Research data publication – why, how and where?

**Ulrike Stahl & Anja Hühnlein**  
**Information Centre and Library, Quedlinburg**

# Research data at the JKI

*“Research data are the result of a scientific working process or are produced throughout it. They encompass many different types and formats, for example measurements, audio data, texts, images, maps, software-code, simulations, surveys and samples .”*

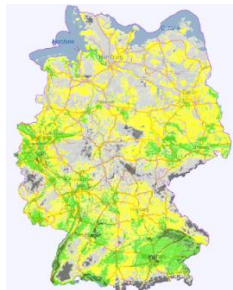
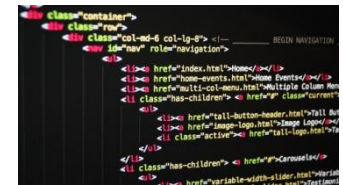


**photometric data**  
detection of molecules

**collections**  
herbaria  
sera  
preparations



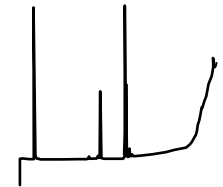
**software code**



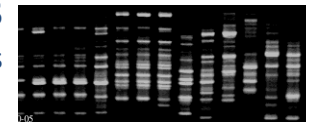
**geographic maps**  
distribution  
nutrition state  
growing area



**chromatographic data**  
analysis of chemical compounds



**PCR- products**  
gene analysis



**qualitative data**  
infestation  
sensory evaluation  
stress

**images**  
phenotyping



**quantitative data**  
soil data, weather data,  
machine data, numbers,  
volumes, lengths, areas,  
mass of plants and animals



# Why? – everyone requests it

funding agencies, scientific organizations, publishers and councils request open access to data

*„... supports[...] open access to research data from publicly funded research”*



*„authors are required to make materials, data, code, and associated protocols promptly available to readers [...]"*



*„... research data should be made publicly available as soon as possible”*



*„... research data should be published in suitable repositories and service centers”*

# Why? – the JKI supports it

Data Policy of the JKI was passed on 22nd of March 2017



*„The JKI supports and promotes open access to any research data that results of publicly funded research and recognizes any effort of the scientists to publish their data and make them citable.“*

[https://www.julius-kuehn.de/media/IB/PDF/IB-JKI-Leitlinien\\_Forschungsdaten.pdf](https://www.julius-kuehn.de/media/IB/PDF/IB-JKI-Leitlinien_Forschungsdaten.pdf)

# Why? – the real reason



Data Sharing and Management Snafu in 3 Short Acts

## [video of NYU Health Sciences Library](#)

Karen Hanson, Alisa Surkis and Karen Yacobucci (2012), Data Sharing and Management Snafu in 3 Short Acts, NYU Health Science Library

- research data becomes **easily accessible and re-usable for me and others**
  - no looking for, no guessing, no problems with reading and opening anymore
- research becomes **transparent** and will be **verifiable**
- research will be **more efficient** and **networked**
- **reputation** of the institutes and the scientists **increases**
- **integration** in national and international infrastructure becomes possible

# How? – publish data „FAIR“

*FAIR- the most accepted set of guiding principles to make data **F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable to date*

Wilkinson, M. D. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18, <https://www.force11.org/group/fairgroup/fairprinciples>

- F**indable: a **persistent identifier** (e.g. DOI) is assigned  
**metadata** (data about the data) are assigned  
**registered** in a searchable resource (e.g. DataCite, re3data.org)
- A**ccessible: **retrievable by their identifier**  
**metadata are accessible**, even when the data are no longer available
- I**nteroperable: a formal, accessible and broadly applicable language is used  
**links to other** resources (data, paper, projects) are given
- R**e-usable: a clear and accessible **data usage license** is released  
**provenance information** is associated  
domain-relevant community **standards** are

# Where? - the praxis to publish data

## Data paper

a peer-reviewed article describing a dataset but lacks any analysis or conclusions, data is deposited in a repository with thoroughly detailed rationale and collection methods

e.g. [ScientificData](#)

## Repository

an independent documentation of data that underlies a journal article and links to it or is unpublished to date, i.e. „stand alone“ publication

e.g. [Pangaea](#) or [Dryad](#)

## Supplement

the traditional journal article accompanied by underlying data in the supplements, data can be hosted by the journal directly or deposited externally and is linked then

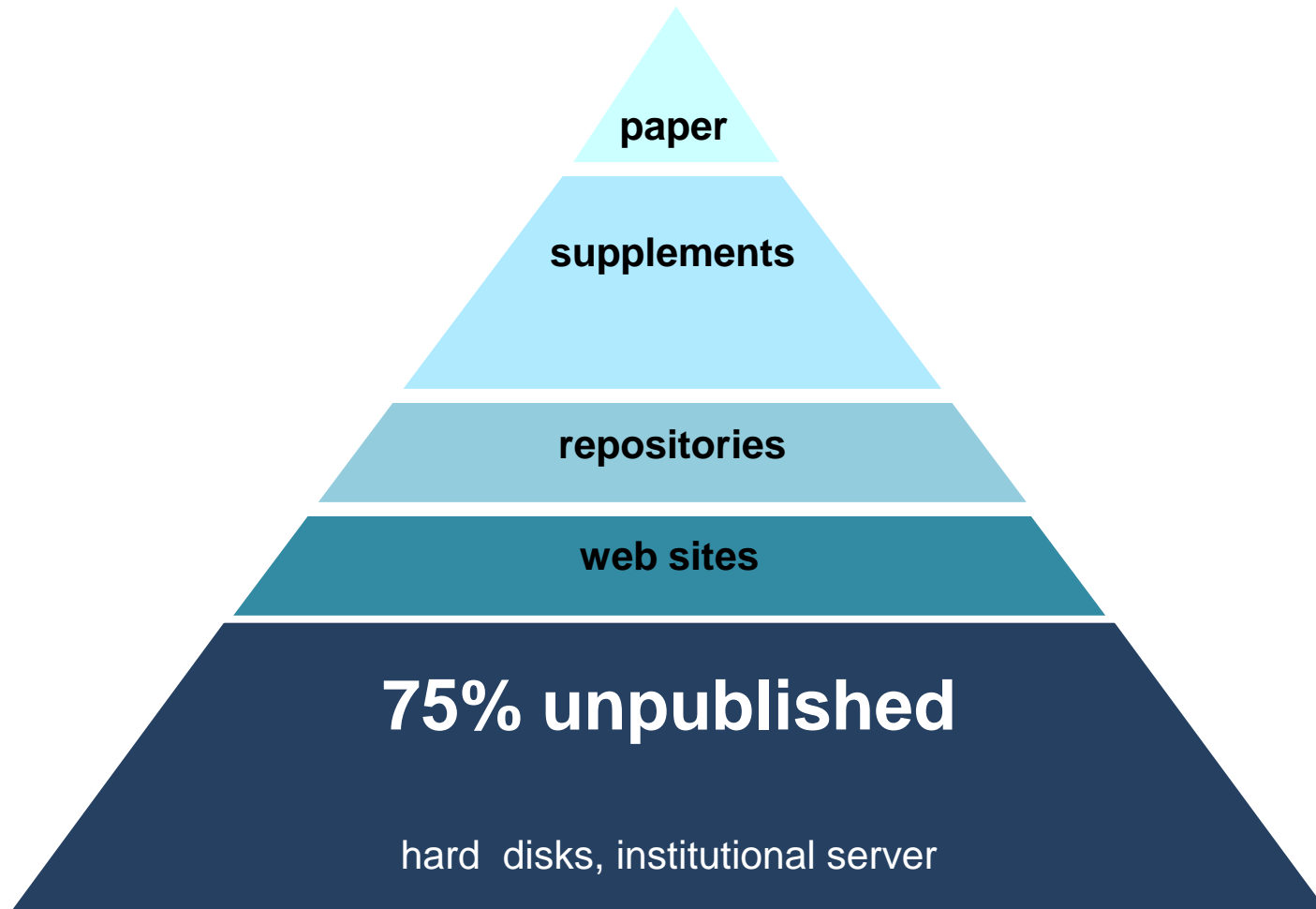
e.g. [Apidologie](#)

## Web site

any form of data release via the personal or institutional web site

















e.g. [spreadsheet](#)

# Where? - the praxis to date

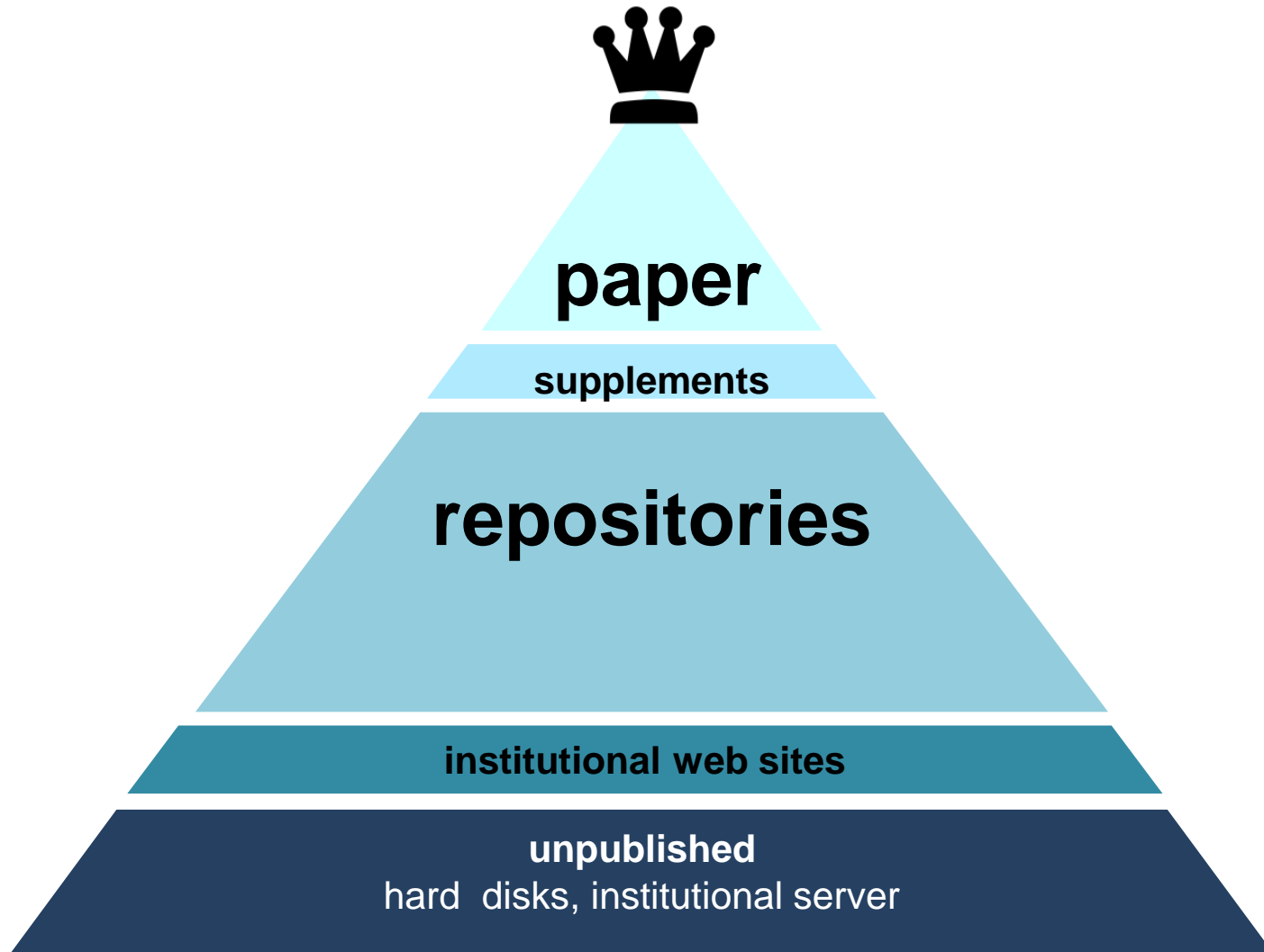




# Is the praxis „FAIR“?

	Data Paper	Repository	Supplement	Web site	
F	indable			 <ul style="list-style-type: none"> <li>• no metadata</li> <li>• no identifier</li> <li>• not registered</li> </ul>	 <ul style="list-style-type: none"> <li>• no metadata</li> <li>• no identifier</li> <li>• seldom registered</li> </ul>
A	ccessible			 <ul style="list-style-type: none"> <li>• as restricted as the article</li> </ul>	 <ul style="list-style-type: none"> <li>• broken links if moved</li> </ul>
I	nteroperabel			 <ul style="list-style-type: none"> <li>• no standards</li> <li>• often size limited</li> </ul>	 <ul style="list-style-type: none"> <li>• no standards</li> <li>• no detailed description</li> </ul>
R	eusable			 <ul style="list-style-type: none"> <li>• only reusable if OA, or explicitly stated</li> </ul>	 <ul style="list-style-type: none"> <li>• seldom free visible license</li> </ul>

# Where? - the praxis in the future



# Where? – suitable repositories and data journals



## ➤ ZB MED list (>230 checked repositories from life science)

<https://www.publisso.de/open-access-publizieren/forschungsdaten/forschungsdatenrepositorien/>

Last updated: 08/07/2017

	Select category ▾	Select category ▾	Select category ▾	Select category ▾	Select category ▾
Name	Subject area focus in the life sciences	Further subject area	Regionally limited content	Operator's location	Data publication only in connection with an article?
<a href="#">↗ 1000 Functional Connec-</a>	Neurosciences	Medicine	not limited	USA	no

## ➤ registry for repositories - re3data.org (>1900 repositories)

<https://www.re3data.org/>

The image shows the search interface of re3data.org. It features the 're3data.org' logo at the top left. Below it is a search bar with the placeholder text 'Search...'. To the left of the search bar is a 'Filter' section with two expandable categories: 'Subjects' and 'Content Types', each with a small square icon to its right. Below the search bar, there is a horizontal progress indicator with several segments, the second of which is highlighted in blue.

## ➤ list of data journals (> 20 journals)

[http://www.forschungsdaten.org/index.php/Data\\_Journals](http://www.forschungsdaten.org/index.php/Data_Journals)

# Where? – in OpenAgrar, a FAIR repository



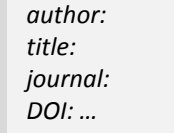
<https://www.openagrar.de/content/index.xml>



text, audio, poster, slides



research data



metadata of publications

**94.059** entries in total    **6775** from JKI    **3** research data

# Where? – in OpenAgrar, a FAIR repository

## view, metadata and DOI

Forschungsdaten Do, 27. Okt. 2016 CC BY-SA 4.0

### Riesling grapevine canopy images for 3D reconstruction


Kicherer, Anna <sup>GND</sup>; Rose, J. Christian

Image based 3D reconstruction of whole grapevine rows for automatic field phenotyping of grape bunches and berries. Images were acquired in VSP and SMPH trained Riesling rows (at JKI Geilweilerhof Siebeldingen, Germany) using the PHENObot (Kicherer et al.2015).

Wird referenziert bei

- Towards automated large-scale 3D phenotyping of vineyards under field conditions

Dateien

	Kicherer.7z	★	27.10.2016	1.25 GB
---	-------------	---	------------	---------

Einordnung

Datum der Veröffentlichung: 27.10.2016

DOI: [10.5073/jki-data.2016.2](https://doi.org/10.5073/jki-data.2016.2)

Sprache: Englisch

Ressourcentyp: Bild (Standbild)





Schlagwörter: Multi-View-Stereo; 3D point cloud; RGB images; viticulture; field phenotyping

DDC-Sachgruppe der DNB: 630 Landwirtschaft, Veterinärmedizin

Einrichtung: Julius Kühn-Institut, Institut für Rebenzüchtung

↑ auf die Merkliste    ⚙ Aktionen

#### Zitieren

 tweet   
  teilen   
  +1   
  i

Zitierform:


deutsche-sprache

Kicherer, A., & Rose, J. C. (2016). Riesling grapevine canopy images for 3D reconstruction (Version 1) [Data set]. Julius Kühn-Institut, Bundesforschungsinstitut für Kulturpflanzen. <https://doi.org/10.5073/jki-data.2016.2>

Zitier-Link kopieren

#### Rechte

Nutzung und Vervielfältigung:



#### Export

BibTeX, MODS, RIS, ISI, PICA, CSV, DC

meta data, data files  
 How to cite  
 license /rights  
 social media  
 export formats

author  
 title  
 abstract  
 resource type  
 publisher  
 date of publishing  
 file format, number, size  
 keywords  
 classification of German National Library  
 related resources

DOI registration

# Checklist to publish and cite data

- When to publish?**
- prepare data publication together with the scientific paper so that you can cite your data in your paper the first time
- Where to publish?**
- first look for a FAIR subject specific repository suitable for your data
  - second look for a FAIR general repository
  - contact IB for publishing in OpenAgrar ([ulrike.stahl@julius-kuehn.de](mailto:ulrike.stahl@julius-kuehn.de))
- What to consider for publishing?**
- prepare your data files (check abbreviations, column names, units, format)
  - prepare metadata according to the requirements of the repository (author, title, contributors, abstract, keywords, resource type, date created, geo location, linked documents, funding agency/project id ...)
  - think about a license (CCBY, CC0)
  - contact IB for the metadata necessary for OpenAgrar
- How and when cite data?**
- whenever a claim relies upon data, it should be cited
  - always cite it at the finest-grained level and cite the right version
  - cite it in text and put it in the reference list, only then it will be indexed

# Helpful links and further read

## ➤ The FAIR principles

Wilkinson, M. D. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 <https://www.nature.com/articles/sdata201618> <https://www.force11.org/group/fairgroup/fairprinciples>

## ➤ The Joint Declaration of Data Citation Principles

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 [<https://www.force11.org/group/joint-declaration-data-citation-principles-final>]

## ➤ How to cite datasets

Ball, A., & Duke, M. (2015). 'How to Cite Datasets and Link to Publications'. *DCC How-to Guides*. Edinburgh: Digital Curation Centre. Available online: <http://www.dcc.ac.uk/resources/how-guides/cite-datasets>

## ➤ Overview about data publication initiatives underway and the current conversation

Kratz J and Strasser C. Data publication consensus and controversies [version 3; referees: 3 approved]. *F1000Research* 2014, 3:94 (doi: [10.12688/f1000research.3979.3](https://doi.org/10.12688/f1000research.3979.3))

## ➤ Study about datasharing practice based on a survey of more than 1500 mostly german scientists

Fecher, B., Friesike, S., Hebing, M., & Linek, S. (2017). A reputation economy: how individual reward considerations trump systemic arguments for open access to data. *3*, 17051. doi:10.1057/palcomms.2017.51 <https://www.nature.com/articles/palcomms201751#supplementary-information>

## ➤ Websites – comprehensive information to research data management

<http://www.forschungsdaten.org/index.php/Hauptseite>  
<https://www.forschungsdaten.info/en/>

## ➤ Organization - Research Data Alliance (RDA)

a community-driven organization launched in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation with the goal of building the social and technical infrastructure to enable open sharing of data, now with over 6,000 members from 130 countries <https://www.rd-alliance.org/about-rda>



**Thank you for your attention!**



# References of pictures

- ELISA test (license: [Public Domain](#), <https://commons.wikimedia.org/wiki/File:ELISA.jpg>), slide 1
- map (author: JKI, Deutscher Sojaförderring, license: all rights reserved, [http://geoportal.julius-kuehn.de/map?app=soja\\_neu](http://geoportal.julius-kuehn.de/map?app=soja_neu)), slide 1
- grapes (license: [CC0](#), <https://pixabay.com/de/trauben-detail-beeren-henkel-1612607/>), slide 1
- carrots (license: [CC0](#), <https://pixabay.com/de/m%C3%B6hren-bund-m%C3%B6hren-lebensmittel-1160683/>), slide 1
- insects (license: [CC0](#), [https://cdn.pixabay.com/photo/2016/06/03/05/54/aphids-1432792\\_960\\_720.jpg](https://cdn.pixabay.com/photo/2016/06/03/05/54/aphids-1432792_960_720.jpg)), slide 1
- PCR (author: Olaf Hering, license: [CC0](#)), slide 1
- chromatogram (author: Olaf Hering, license: [CC0](#)), slide 1
- sample (author: Olaf Hering, license: [CCBY](#)), slide 1
- source code (license: [CC0](#), <https://pixabay.com/de/code-html-digital-codierung-web-1076533/>), slide 1
- database icon (license: [CC0](#), <https://openclipart.org/detail/183938/icon-data-base>), slide 1
- USB-drive icon (license: [CC0](#), <https://openclipart.org/detail/104053/usb-flash-drive>), slide 1
- PC screen icon (license: [CC0](#), <https://pixabay.com/de/laptop-computer-computer-pc-1155173/>), slide 1
- container icon (license: [Public Domain](#), [https://commons.wikimedia.org/wiki/File:Replacement\\_filing\\_cabinet.svg](https://commons.wikimedia.org/wiki/File:Replacement_filing_cabinet.svg)), slide 1
- crown icon (license: [CC0](#), <https://pixabay.com/de/krone-icon-adel-silhouette-k%C3%B6nig-1332785/>), slide 10
- smiley icon, smiling (license: [CC0](#), <https://pixabay.com/de/smiley-gelb-cartoon-1156497/>), slide 9
- smiley icon, sad (license: [CC0](#), <https://pixabay.com/de/traurig-icon-smiley-emotion-chat-2192410/>), slide 9
- smiley icon, neutral (license: [CC0](#), <https://openclipart.org/detail/30229/tango-face-plain>), slide 9
- csv icon (license: [CC0](#), <https://openclipart.org/detail/169752/fileiconcsv>), slide 12
- pdf icon (license: [CC0](#), <https://openclipart.org/detail/169750/fileiconpdf>), slide 12