

FAKIN

(Forschungsdatenmanagement an kleinen Instituten)



Gefördert durch:



Bundesministerium
für Bildung
und Forschung



No of doc: 1 Version: 1 Date: 8.5.2017	MANAGEMENT SYSTEM Statement	KOMPETENZZENTRUM WasserBerlin
--	--------------------------------	----------------------------------

Verpflichtungserklärung „Gute Wissenschaftliche Praxis“

Die Anwendung hoher Sorgfalt bei den im Kompetenzzentrum Wasser Berlin (KWB) angewandten wissenschaftlichen Methoden und bei der Dokumentation der wissenschaftlichen Daten ist Grundlage für das Vertrauen unserer Gesellschafter, Forschungspartner und der Öffentlichkeit in unsere Arbeit. Daher sind wissenschaftliche Redlichkeit und die Beachtung der Grundsätze guter wissenschaftlicher Praxis unverzichtbare Voraussetzungen für unser wissenschaftliches Arbeiten. In unserem seit 2010 jährlich nach DIN EN ISO 9001:2015 zertifizierten Qualitätsmanagementsystem werden diese Grundsätze berücksichtigt.

Das Kompetenzzentrum Wasser Berlin begrüßt daher die Empfehlungen der Kommission „Selbstkontrolle in der Wissenschaft“ durch die Mitgliedseinrichtungen der Deutschen Forschungsgemeinschaft (DFG) und setzt diese daher auch um.

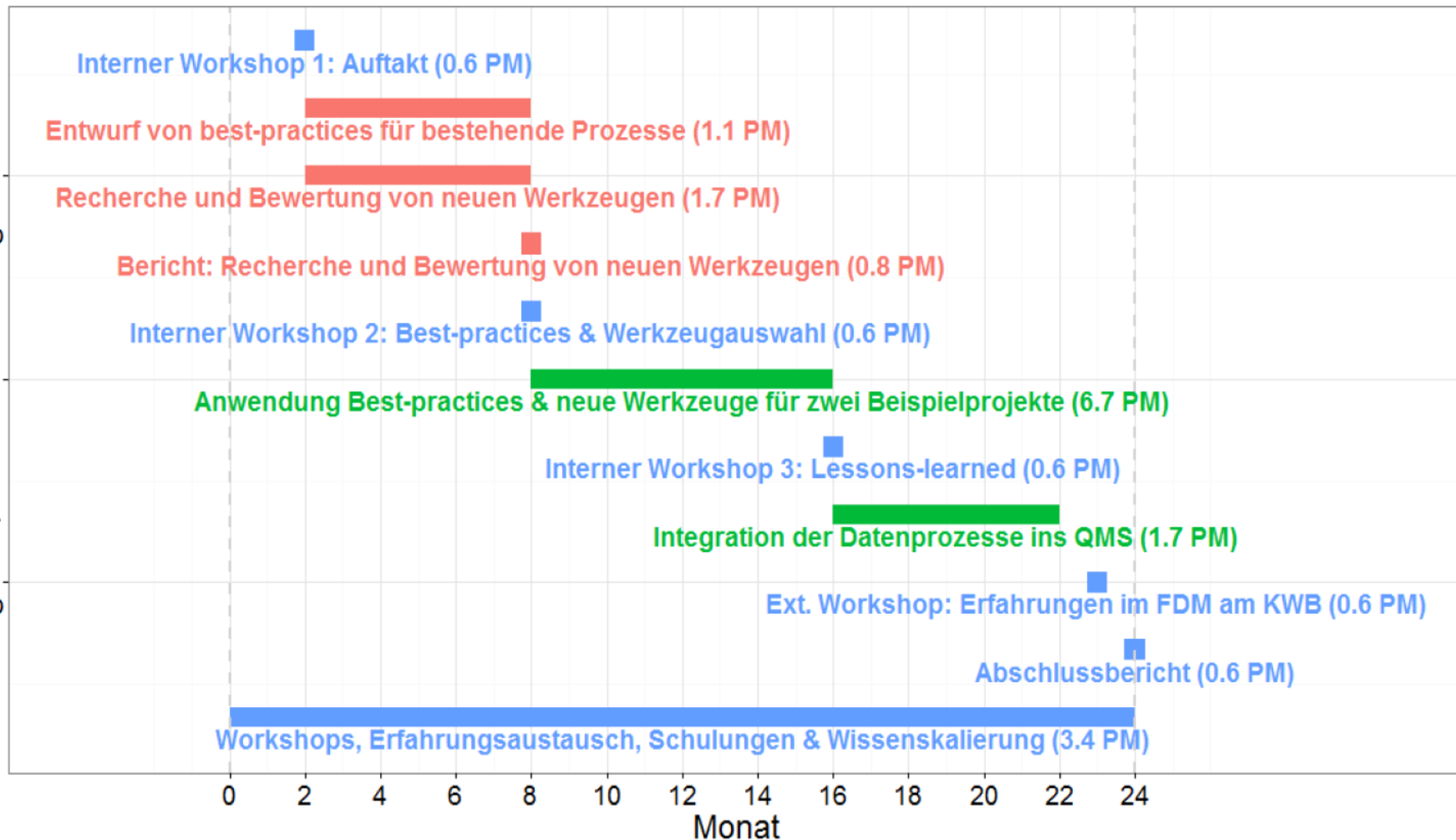
Allgemeine Prinzipien unserer wissenschaftlichen Arbeit

- Dokumentation aller Daten (inkl. Beschreibung der Datengewinnung z.B. Versuchsaufbauten, Verfahrensparameter) sowie der daraus gewonnenen (abgeleiteten) Resultate (z.B. durch numerische Auswertungen oder mathematische Modellierungen)
- Gewährleistung der Reproduzierbarkeit und Richtigkeit der Daten
- Aufbewahrung der für Veröffentlichungen relevanten Primärdaten auf Datenträgern in den Instituten für 10 Jahre, sofern dieses vertretbar und möglich ist, dazu gehören alle für die Reproduzierbarkeit der Ergebnisse notwendigen Daten
- Systematischer Skeptizismus an den Daten und Resultaten, unter anderem bei vorab angestrebten Ergebnissen

Wie gehen wir vor?

Arbeitspaket ■ AP 1: Identifikation ■ AP 2: Implementierung ■ AP 3: Kommunikation

Aufgaben, Meilensteine & Ergebnisse



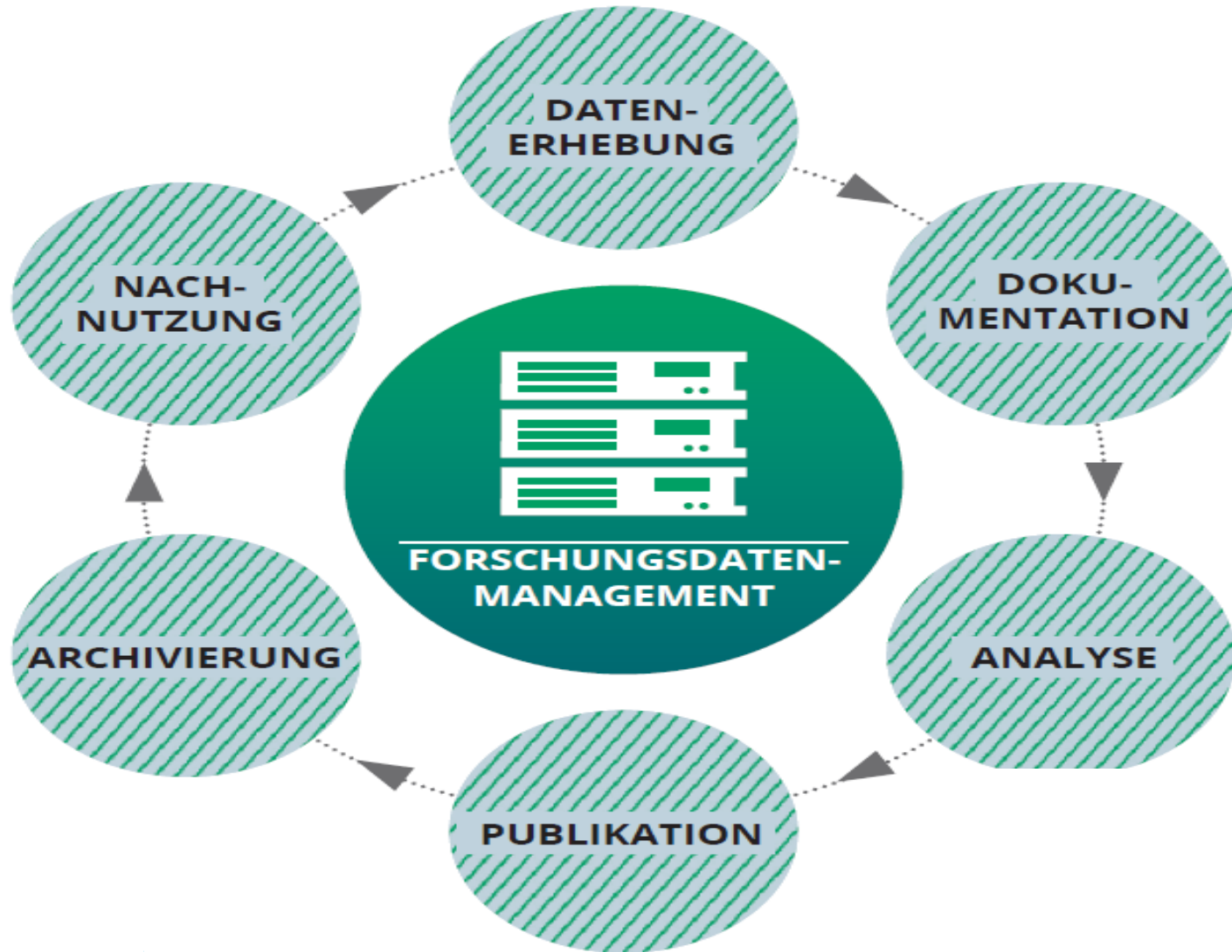
Was ist Forschungsdatenmanagement ?

“...umfasst alle Methoden und Verfahren, die zur Sicherung der langfristigen Nutzbarkeit von Forschungsdaten angewendet werden.“

Quelle: Bertelmann et al. 2014

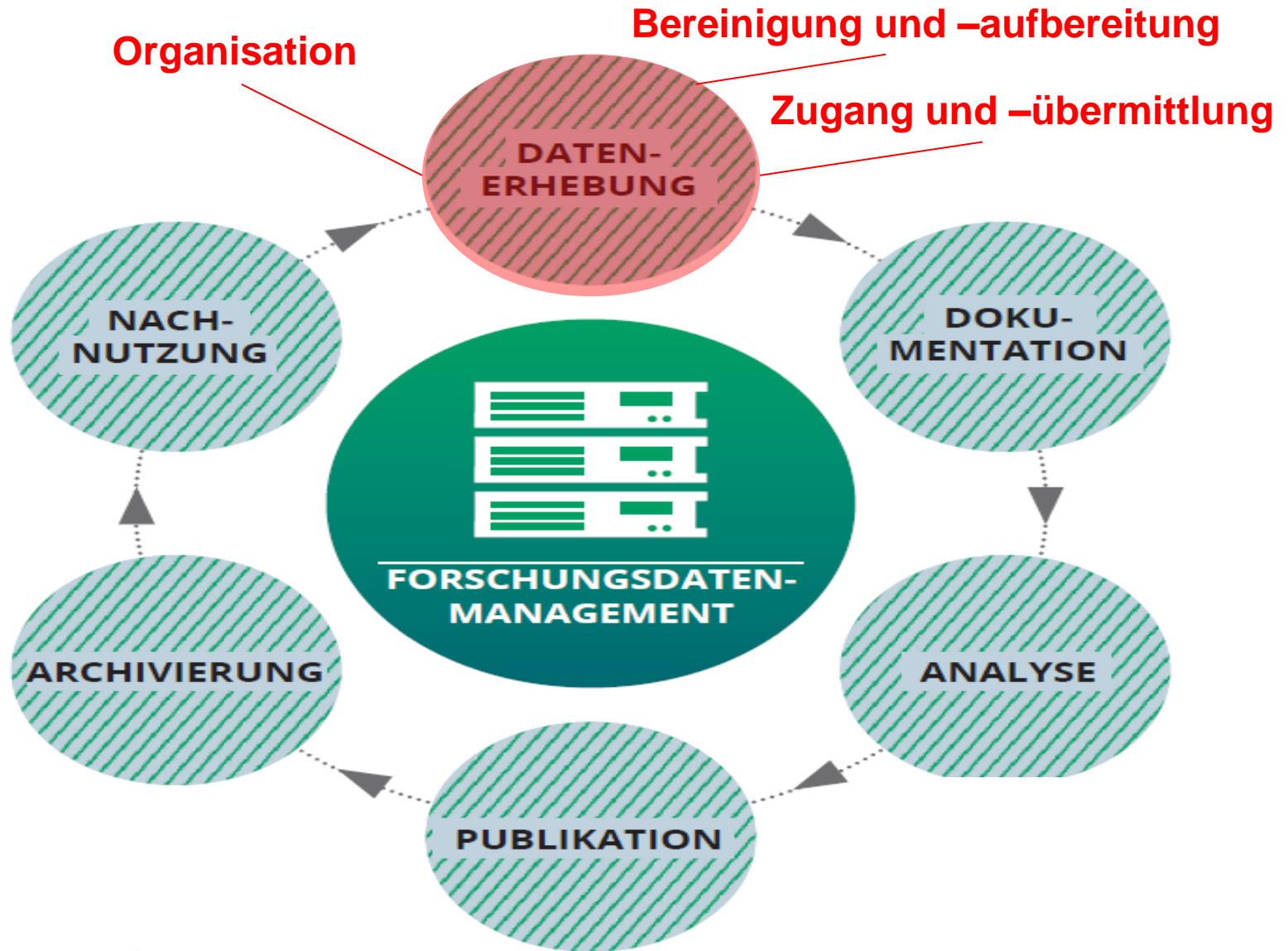


Der Lebenszyklus von Forschungsdaten



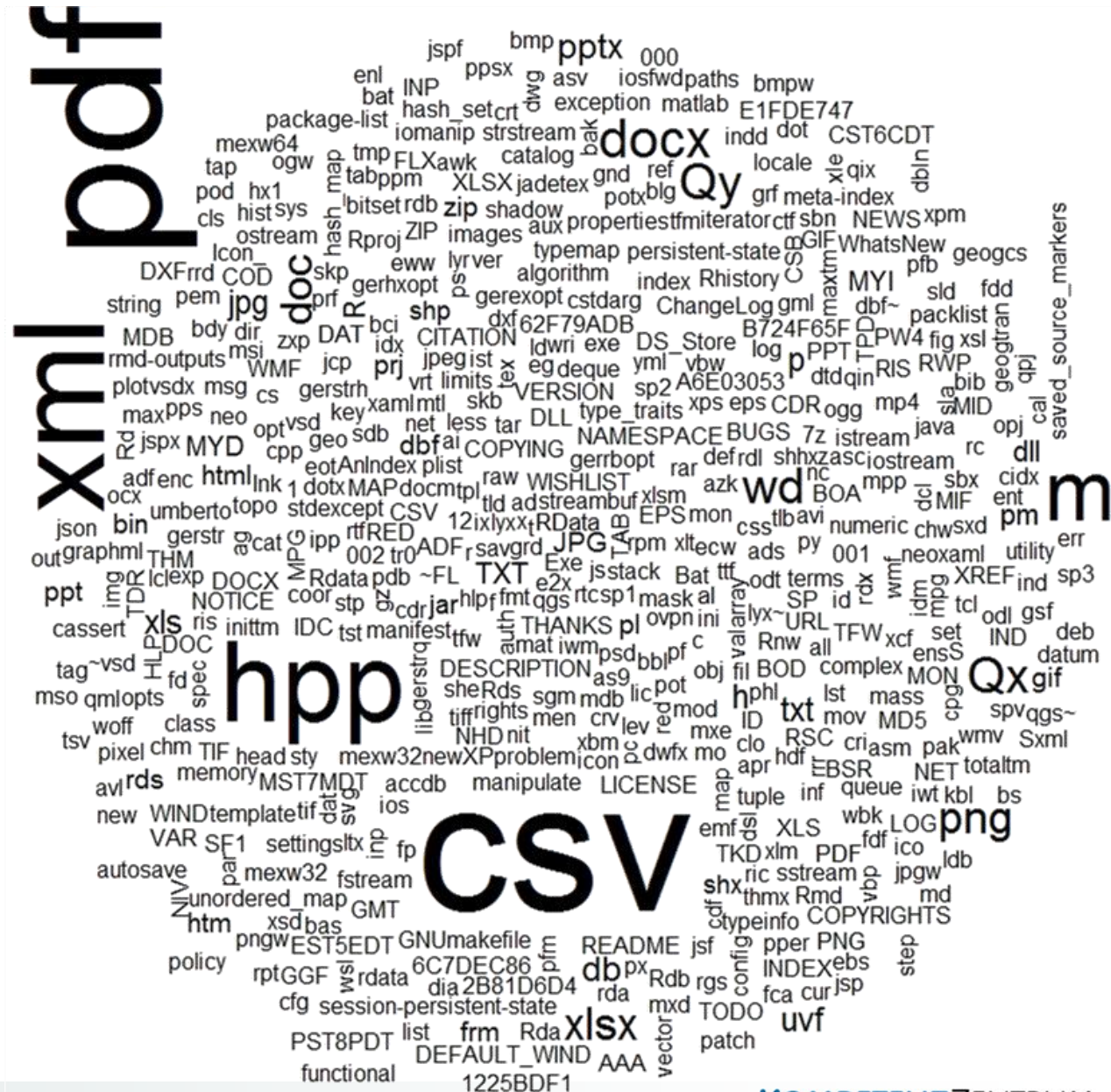
Quelle: Bertelmann et al. 2014

Der Lebenszyklus von Forschungsdaten



Quelle: Bertelmann et al. 2014

**Über 1324
verschiedene
Dateiendungen**
(die mindestens
2x vorkommen)



Datenorganisation: Wie gut sind wir schon ?

- **Administrativ:** erste Erfahrungen mit FAKIN

Datenorganisation: Wie gut sind wir schon ?

- **Administrativ:** erste Erfahrungen mit FAKIN

Email von Bodo vom 30.06:

“.... Bitte unbedingt auch für Fakin die für alle Projekte festgeschriebene Folderstruktur verwenden!!!!”

[\\poseidon\projekte\\$\R&D Templates\Templates_document-current\FolderStructure_template\Project_folder_template](#)

Datenorganisation: Wie gut sind wir schon ?

Beispiel: FAKIN

**Bodo hatte
recht! Nur 25%
umgesetzt! ☹️**

1	Administration
2	Administration/Budget
3	Administration/Contracts
4	Administration/Meetings-Minutes
5	Administration/Meetings-Minutes/Team
6	Administration/Memos
7	Administration/Project development-Proposal
8	Closure & Evaluation
9	Communication
10	Communication/Docs
11	Communication/Logos
12	Communication/Pictures
13	Communication/Presentations
14	Data-Work packages
15	Exchange
16	Reports

Datenorganisation: Wie gut sind wir schon ?

	dep_proj_name	qms_dirs_total	qms_dirs_percent
1	GRW_FAKIN	16	100.00%
2	SUW_DSMT	16	100.00%
3	SUW_FLUSSHYGIENE	16	100.00%
4	SUW_netWORKS4	16	100.00%
5	SUW_NITROLIMIT	16	100.00%
6	WWT_AquaNES	16	100.00%
7	WWT_CWPharma_Seed	16	100.00%
8	WWT_OEMP	16	100.00%
9	WWT_POWERSTEP	16	100.00%
10	WWT_TestTools	16	100.00%
11	GRW_SAPH-PANI	15	93.75%
12	WWT_ASKURIS	15	93.75%
13	WWT_IST4R_LCA	15	93.75%
14	WWT_nurec4org	15	93.75%
15	WWT_SMART-Plant	15	93.75%
16	SUW_KURAS	14	87.50%
17	SUW_SEMA-Berlin	14	87.50%
18	GRW_DEMEAU	13	81.25%
19	SUW_NITROLIMIT 2	13	81.25%
20	SUW_OGRE	13	81.25%
21	WWT_NewFert	13	81.25%
22	SUW_SEMA	12	75.00%
23	WWT_E-VENT	12	75.00%
24	SUW_PROCOL_Berlin-Bogota	11	68.75%
25	SUW_PREPARED	7	43.75%
26	WWT_DEMOWARE	7	43.75%

Datenorganisation: Wie gut sind wir schon ?

- **Administrativ:** erste Erfahrungen mit FAKIN
- **„Data-Work packages“:** Erfahrungen aus dem Surface Water Department



And the **WINNER** is.....

- Längster Pfad
- Tiefste Verschachtelung



Längster Pfad mit 321 Zeichen:

```
01 //server/  
02  projekte$/  
03    SUW_Department/  
04      Projects/  
05        NITROLIMIT/  
06          Data-Work packages/  
07            Massnahmenkatalog/  
08              Literatur/  
09                Endnote/  
10                  NITROLIMIT.Data/  
11                    PDF/  
12 + 13                      (*)
```

(*) Kopie von Reinigungsleistung_von_zentralen_Anlagen_zur_Behandlung_von_Abluessen_Kasting_2002-3006138113/
Reinigungsleistung_von_zentralen_Anlagen_zur_Behandlung_von_Abluessen_Kasting_2002.pdf



Tiefste Verschachtelung mit 23 Ebenen:

```
01 //server/  
02   projekte$/  
03     SUW_Department/  
04       Projects/  
05         FLUSSHYGIENE/  
06           Data-Work packages/  
07             AP2 Monitoring/  
08               Daten/  
09                 RData/  
10                   R/  
11                     win-library/  
12                       3.1/  
13                         file19d478bc2c82/  
14                           BH/  
15                             include/  
16                               boost/  
17                                 spirit/  
18                                   home/  
19                                     support/  
20                                       detail/  
21                                         lexer/  
22                                           parser/  
23                                             tokeniser  
23                                               tree
```



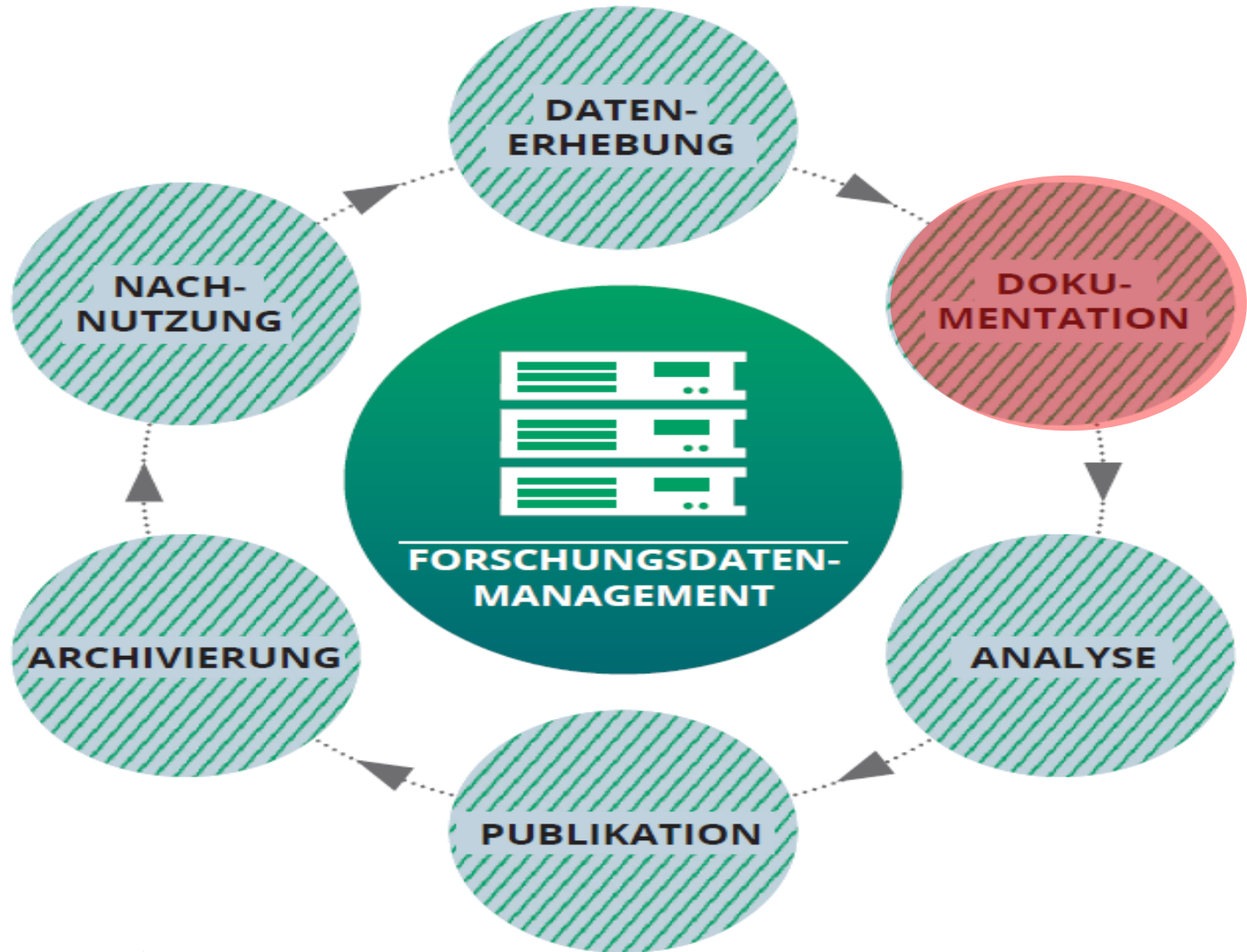
And the WINNER is.....

- Längster Pfad: 321 Zeichen
- Tiefste Verschachtelung: 23 Ebenen

Wie können wir das verbessern?

- Vorschlag einer Nomenklatur und einfacher Regeln

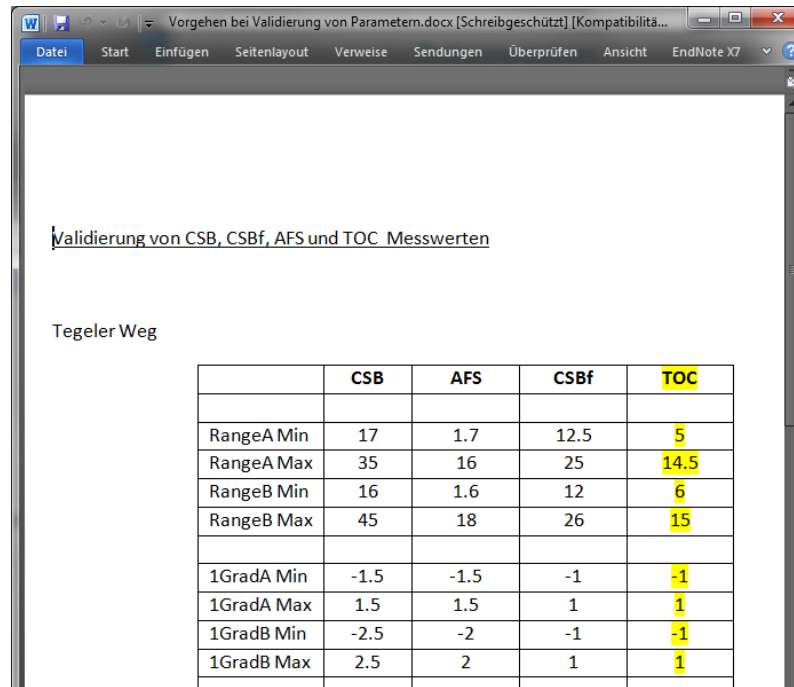
Der Lebenszyklus von Forschungsdaten



Quelle: Bertelmann et al. 2014

Dokumentation: Fundstücke aus MIA-CSO (.doc)

\\server\Archiv\$\Closed_Projects\MIA-CSO\MIA-CSO WP 2 Monitoring\
Data + Studies\Data management\Protocols\Vorgehen bei Validierung von Parametern.docx



The screenshot shows a Microsoft Word document window. The title bar reads 'Vorgehen bei Validierung von Parametern.docx [Schreibgeschützt] [Kompatibilität...]' and the ribbon includes 'Datei', 'Start', 'Einfügen', 'Seitenlayout', 'Verweise', 'Sendungen', 'Überprüfen', 'Ansicht', and 'EndNote X7'. The document content includes the heading 'Validierung von CSB, CSBf, AFS und TOC Messwerten' and the sub-heading 'Tegeler Weg'. Below these is a table with 5 columns: an empty header column, 'CSB', 'AFS', 'CSBf', and 'TOC'. The 'TOC' column contains yellow-highlighted values. The table rows are: RangeA Min, RangeA Max, RangeB Min, RangeB Max, 1GradA Min, 1GradA Max, 1GradB Min, and 1GradB Max.

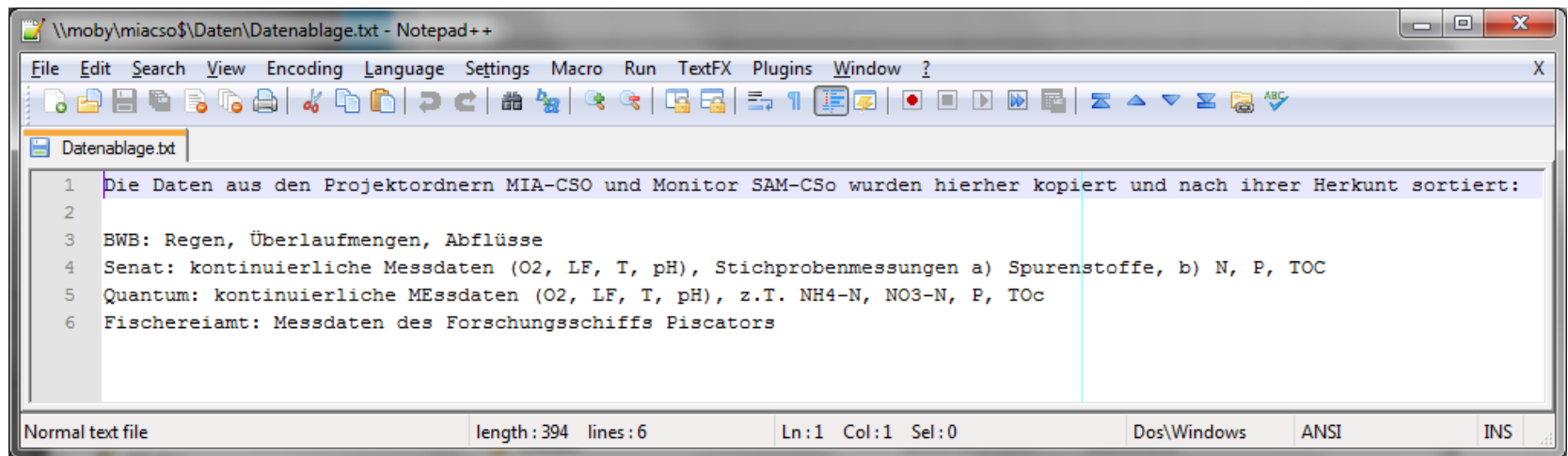
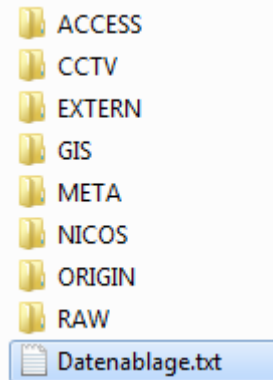
	CSB	AFS	CSBf	TOC
RangeA Min	17	1.7	12.5	5
RangeA Max	35	16	25	14.5
RangeB Min	16	1.6	12	6
RangeB Max	45	18	26	15
1GradA Min	-1.5	-1.5	-1	-1
1GradA Max	1.5	1.5	1	1
1GradB Min	-2.5	-2	-1	-1
1GradB Max	2.5	2	1	1

Fragen:

- Die Datei heißt „Vorgehen bei Validierung...“. Wo wird das Vorgehen beschrieben?
- Was bedeuten die Begriffe in der ersten Spalte?
- Was bedeutet die gelbe Markierung? Ist das noch nicht fertig oder kann das weg?

Dokumentation: Fundstücke aus MIA-CSO (.txt)

\\server\miacso\$\Daten



Fragen:

- Wo finde ich denn die Daten von BWB, Senat, Fischereiamt?
- Was bedeuten die ganzen Ordner?

Dokumentation: Metadaten in MIA-CSO

Geräte und Aktionen

Messstelle: **Kanal Stallstr. (Sta)**

Gerät: Sta: Betriebsmotor

Filter Aktionstypen: **zugeordnete Aktionen**

Betrieb
Reinigung
Sonstige
Wartung

Allgemeiner Kommentar
Ölstandsprüfung
Ölwechsel
Schmierung

Markierung aufheben

Gerät-Info

Allgemein... Gerät-Aktionen...

Aktions-Info

durchzuführen alle: 6 Monat(e) Wartungsplan...

zuletzt durchgeführt: 06.07.2017 13:45

wieder durchführen: 06.01.2018 Historie...

Hinweise:
Die rot markierten Teile sind zu schmieren.

Aktion wird durchgeführt

Sta: Betriebsmotor: Schmierung

Wann?

☐ Jetzt Zeige "Jetzt" Jetzt

☒ Am/um: 06.07.2017 13:45 bis:

Von wem?

HS ...

☒ Warnung bei Wiederholung

Bemerkung

ganz schön schmierige Angelegenheit

Vorhandene Bemerkungen:

durchzuführen am: 06.07.2017

PMP	Schlauchersatz
AMO	Membranersatz

heute durchgeführt: 06.07.2017

06.07.2017 13:45	MOT	Schmierung

Aktion eintragen Aktionskalender

Fragen:

- Wo liegt das Programm, wie „installiere“ ich es?
- Was mache ich, wenn es einen Fehler gibt, wer wartet das Ding?
- Und natürlich der Klassiker: Warum muss das alles so kompliziert sein?

Das Orakel

www.onlineorakel.com

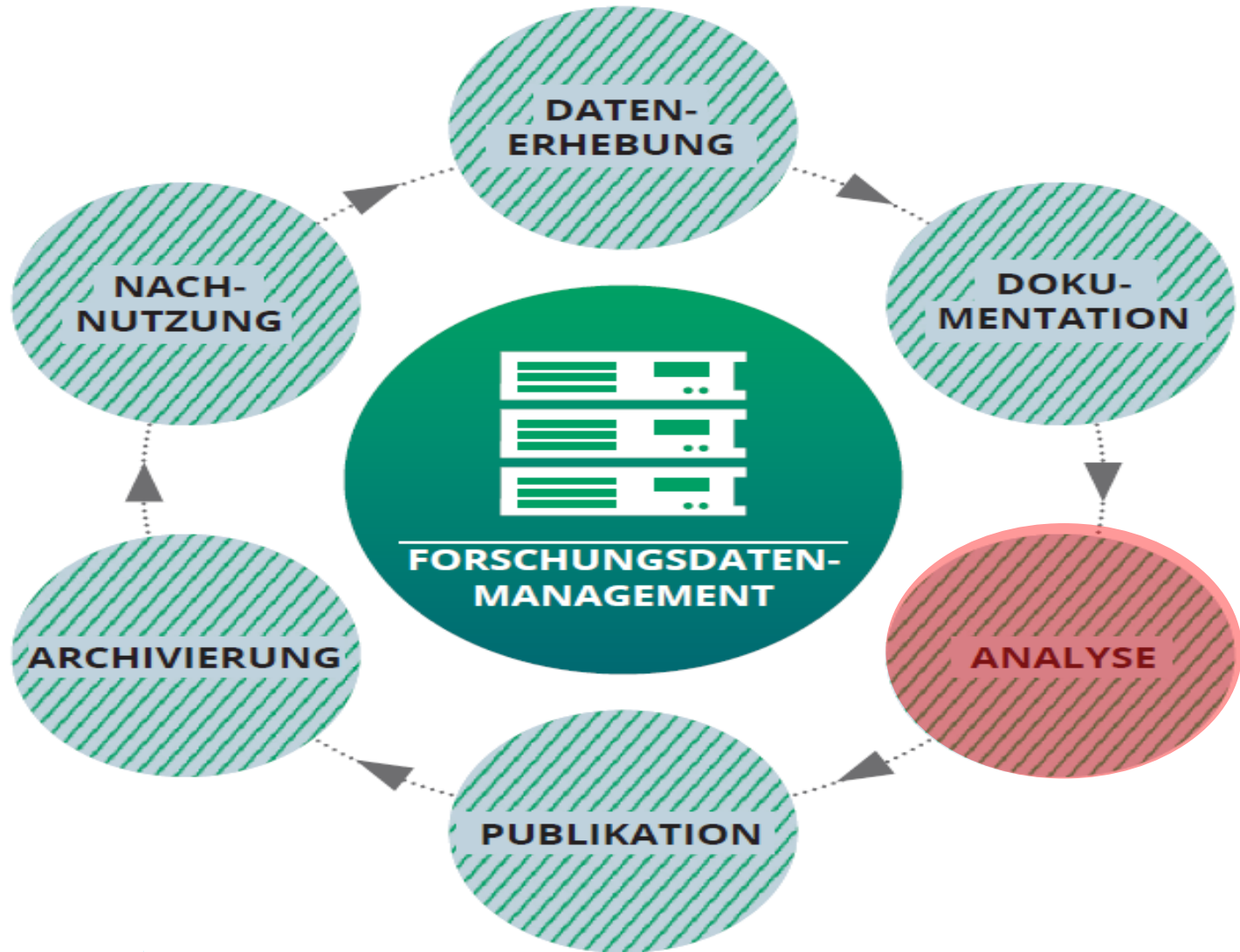
Gib eine Frage ein, die sich mit JA oder NEIN beantworten lässt.



Befrag das Orakel

Wurde in MIA-CSO CSB gemessen?

Der Lebenszyklus von Forschungsdaten



Quelle: Bertelmann et al. 2014

Der Lebenszyklus von Forschungsdaten



Quelle: Bertelmann et al. 2014

Beispiel: R Pakete



KWB-R

**Collection of KWB
R packages**

Statistics

12 public repos
2 members

kontakt@kompetenz-wasser.de

Recently updated [View All on GitHub](#)

[kwb.utils](#) Jul 4, 2017 · 0 watchers · 0 forks

[aquanes.report](#) Jun 28, 2017 · 0 watchers · 3 forks

[fakin.blog](#) Jun 20, 2017 · 0 watchers · 0 forks

[kwb.utils](#)

some basic functions used by
other kwb packages

[aquanes.report](#)

Collects, aggregates and
visualises operational analytical
data from water suppliers
(including a standardised
reporting document)

[fakin.blog](#)

Blog on FAKIN project status

[kwb-r.github.io](#)

Landing page of KWB-R
community

[kwb.default](#)

Get and Set Function Argument
Defaults

[kwb.wtaq](#)

Interface to WTAQ Drawdown
Model (<http://water.usgs.gov/ogw/wtaq/>)

[kwb.test](#)

Test whether Different Functions
Return the Same

[kwb.qmra](#)

QMRA (quantitative microbial
risk assessment) for water
supply systems

[kwb.hantush](#)

Calculation of groundwater
mounding beneath an infiltration
basin

[HydroServerLite](#)

[dc_kwb.hantush](#)

GitHub (aber mehr für Programmierer)

KWB-R/kwb.hantush: Calcula x +

GitHub, Inc. (US) | <https://github.com/KWB-R/kwb.hantush> 130% Suchen

README.md

kwb.hantush

An R package for calculating groundwater mounding beneath an infiltration

build passing

CRAN 0.2.1 downloads 4838 downloads 132/month downloads 19/week

Impact factor: Depsy 38th percentile

Cite as: DOI 10.5281/zenodo.61613

Publikation, Archivierung & Nachnutzung

Zenodo (Langzeitarchivierung in der Schweiz und DOI)

September 6, 2016

kwb.hantush (v.0.2.1)

Rustler, Michael

An R package for calculating groundwater mounding beneath an infiltration basin based on the Hantush equation

Tutorial: <http://kwb-r.github.io/kwb.hantush/>

Software Open Access

Available in

GitHub

Publication date:
September 6, 2016

DOI:
[DOI 10.5281/zenodo.61613](https://doi.org/10.5281/zenodo.61613)

Related identifiers:
Supplement to:
<https://github.com/KWB-R/kwb.hantush/tree/v.0.2.1>

License (for files):
[GNU General Public License 2.0](#)

Share

Cite as
Rustler, M. (2016, September 6). kwb.hantush (v.0.2.1). Zenodo. <http://doi.org/10.5281/zenodo.61613>

Start typing a citation style...

Export
[BibTeX](#) [CSL](#) [DataCite](#) [Dublin Core](#) [JSON](#)

Preview

kwb.hantush-v.0.2.1.zip

- KWB-R-kwb.hantush-bbfea52**
 - .Rbuildignore 134 Bytes
 - .travis.yml 27 Bytes
 - DESCRIPTION 698 Bytes
 - LICENSE 18.0 kB
 - NAMESPACE 67 Bytes
 - NEWS 320 Bytes
 - R
 - kwb.hantush.R 11.2 kB
 - verification.R 7.2 kB
 - README.md 2.2 kB
 - cran-comments.md 476 Bytes
 - inst
 - extdata
 - modelComparison.csv 710 Bytes
 - kwb.hantush.Rproj 355 Bytes
 - kwb_hantush.png 4.5 kB
 - man

Files (30.2 kB)

Name	Size	Preview	Download
kwb.hantush-v.0.2.1.zip	30.2 kB		
md5:43e800f2100a382d261ea8df8d5de828			

Ziele: abteilungsübergreifend beantworten:

Wo stehen wir?

- Wie wird in den verschiedenen Projekten mit den Forschungsdaten gearbeitet?
- Was läuft gut / schlecht?

Wo wollen wir hin?

- Anforderungen an das Datenmanagement (z.B. einfache Auffindbarkeit von Daten) definieren
- Verbesserungsbedarfe identifizieren & priorisieren

Ergebnis:

Hoch priorisierte Themen wollen wir in FAKIN angehen!

Bertelmann, Roland, Petra Gebauer, Tim Hasler, Ingo Kirchner, Wolfgang Peters-Kottig, Matthias Razum, Astrid Recker, D. Ulbricht, and Stephan Van Gasselt. 2014. "Einstieg Ins Forschungsdatenmanagement in Den Geowissenschaften." GeoForschungsZentrum Potsdam. doi:[10.2312/lis.14.01](https://doi.org/10.2312/lis.14.01).

