

Wie etablieren wir FDM an unserem kleinen Wasserforschungsinstitut ?

(Michael Rustler, Projektleiter FAKIN)




Gefördert durch:



Bundesministerium
für Bildung
und Forschung

kleines Forschungsinstitut

Spezifika



keine IT-
Abteilung

Projekt-
bezogen

Einfache
Struktur

kleines Forschungsinstitut

Spezifika

keine IT-
Abteilung

Projekt-
bezogen

Einfache
Struktur

Folgerungen

Keine FDM-
Richtlinien

Schnitt-
mengen?

Informeller
Austausch

kleines Forschungsinstitut

Spezifika

keine IT-
Abteilung

Projekt-
bezogen

Einfache
Struktur

Folgerungen

Keine FDM-
Richtlinien

Schnitt-
mengen?

Informeller
Austausch

FAKIN-Projektziele

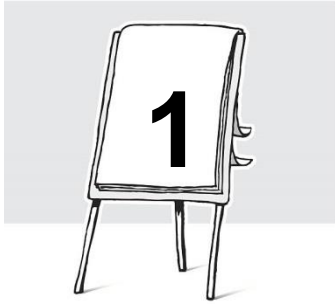
Best Practices
-> QMS

Tools

Unsere (Forschungs-) Datenreise


Auftakt

(Sept. 2017)



1. Finden und
Bekommen
2. Analysieren,
Visualisieren,
Kommunizieren

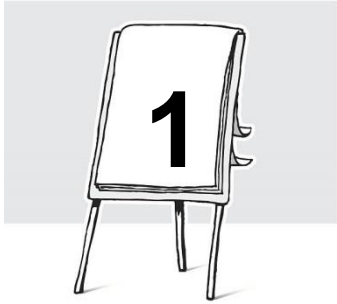
Tools:

- „Data Literacy Matrix“ der
Datenschule
<https://datenschule.de/workshops/>
- Online Umfrage  LimeSurvey

Unsere (Forschungs-) Datenreise

Auftakt

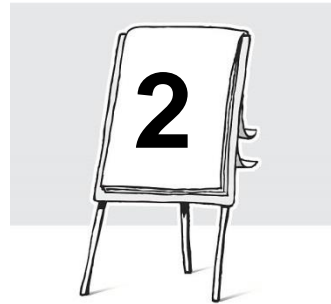
(Sept. 2017)



1. Finden und Bekommen
2. Analysieren, Visualisieren, Kommunizieren

Best-practices

(Jan. 2018)



- Rohdaten



- Ordnerstruktur



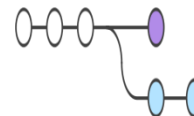
- Nomenklatur



- Metadaten



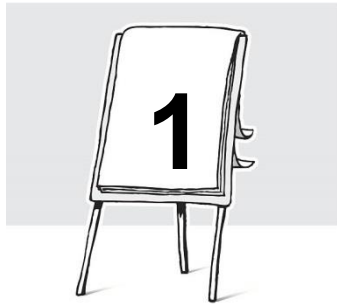
- Versionierung



Unsere (Forschungs-) Datenreise

Auftakt

(Sept. 2017)



1. Finden und Bekommen
2. Analysieren, Visualisieren, Kommunizieren

Best-practices

(Jan. 2018)



- Rohdaten

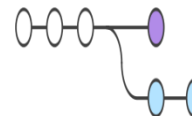
- Ordnerstruktur

- Nomenklatur

- Metadaten

- Versionierung

Testprojekte



Lessons-learned

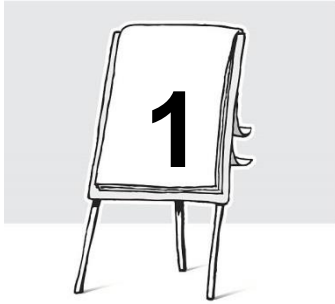
(März 2019)



Unsere (Forschungs-) Datenreise

Auftakt

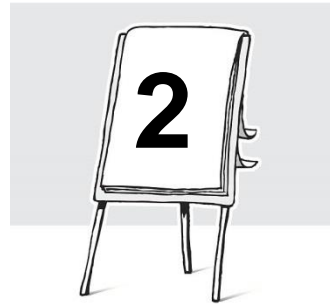
(Sept. 2017)



1. Finden und Bekommen
2. Analysieren, Visualisieren, Kommunizieren

Best-practices

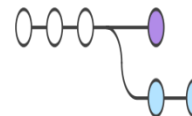
(Jan. 2018)



Testprojekte



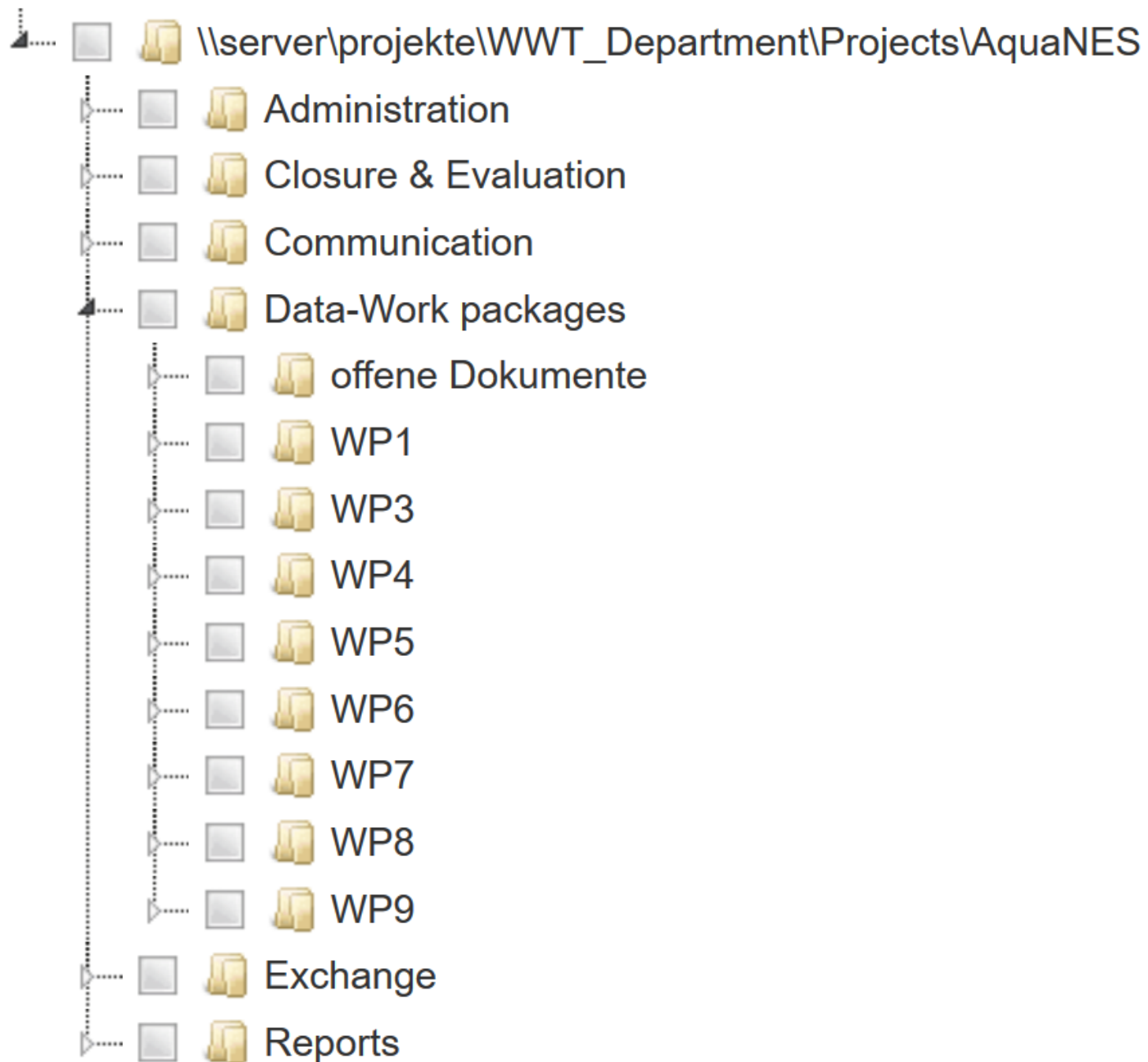
- Rohdaten
- Ordnerstruktur
- Nomenklatur
- Metadaten
- Versionierung

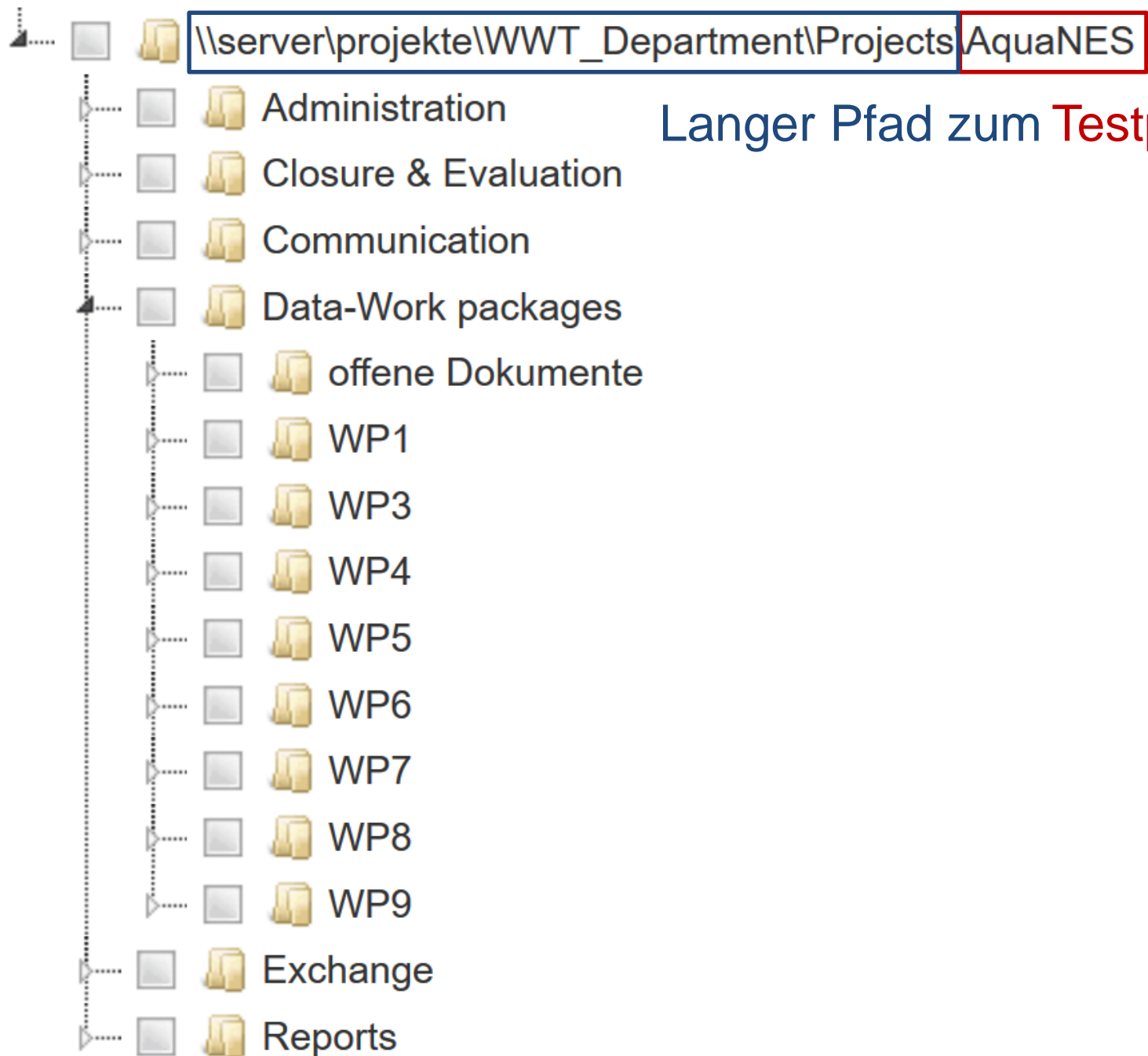


Lessons-learnt

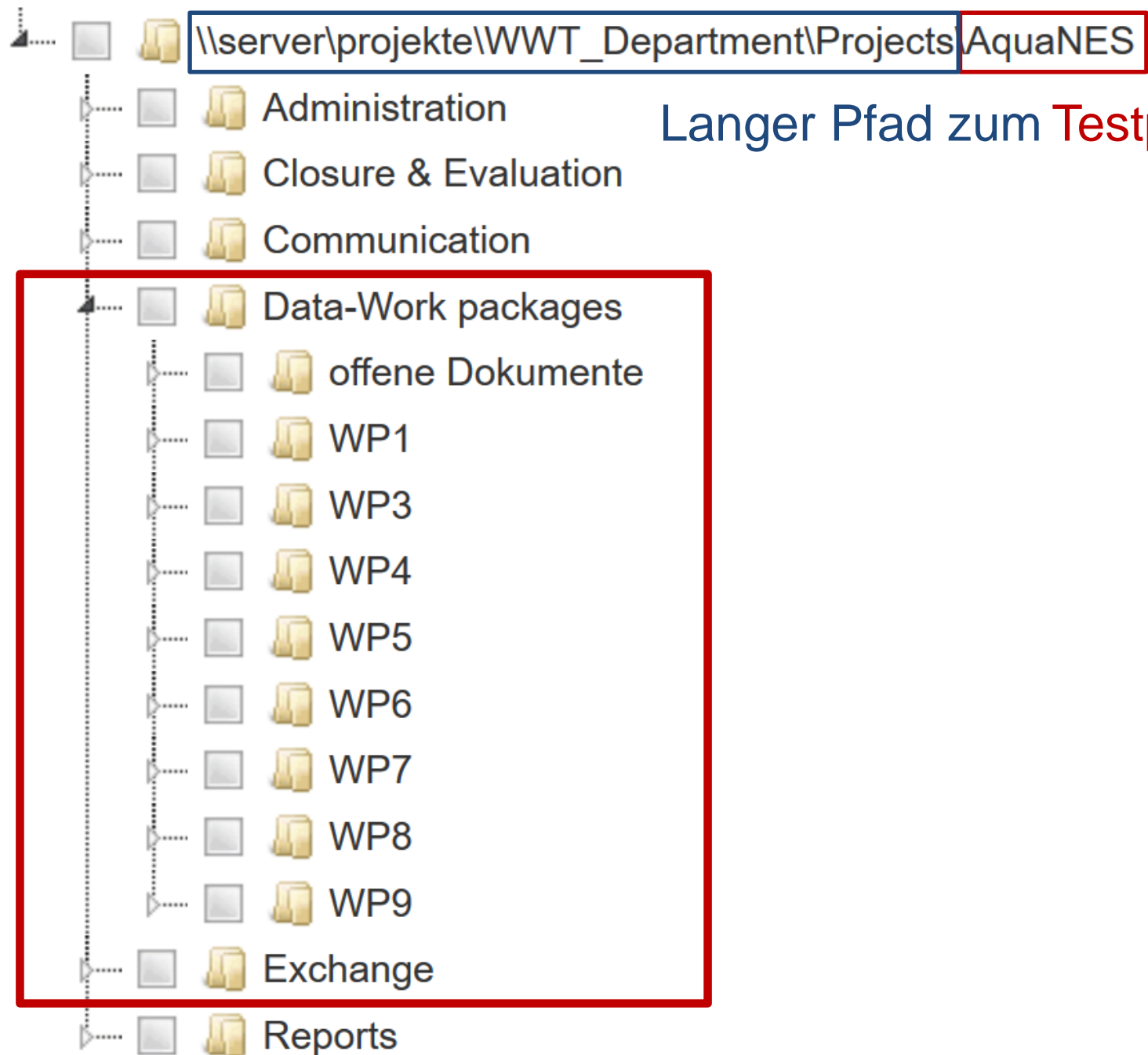
(März 2019)



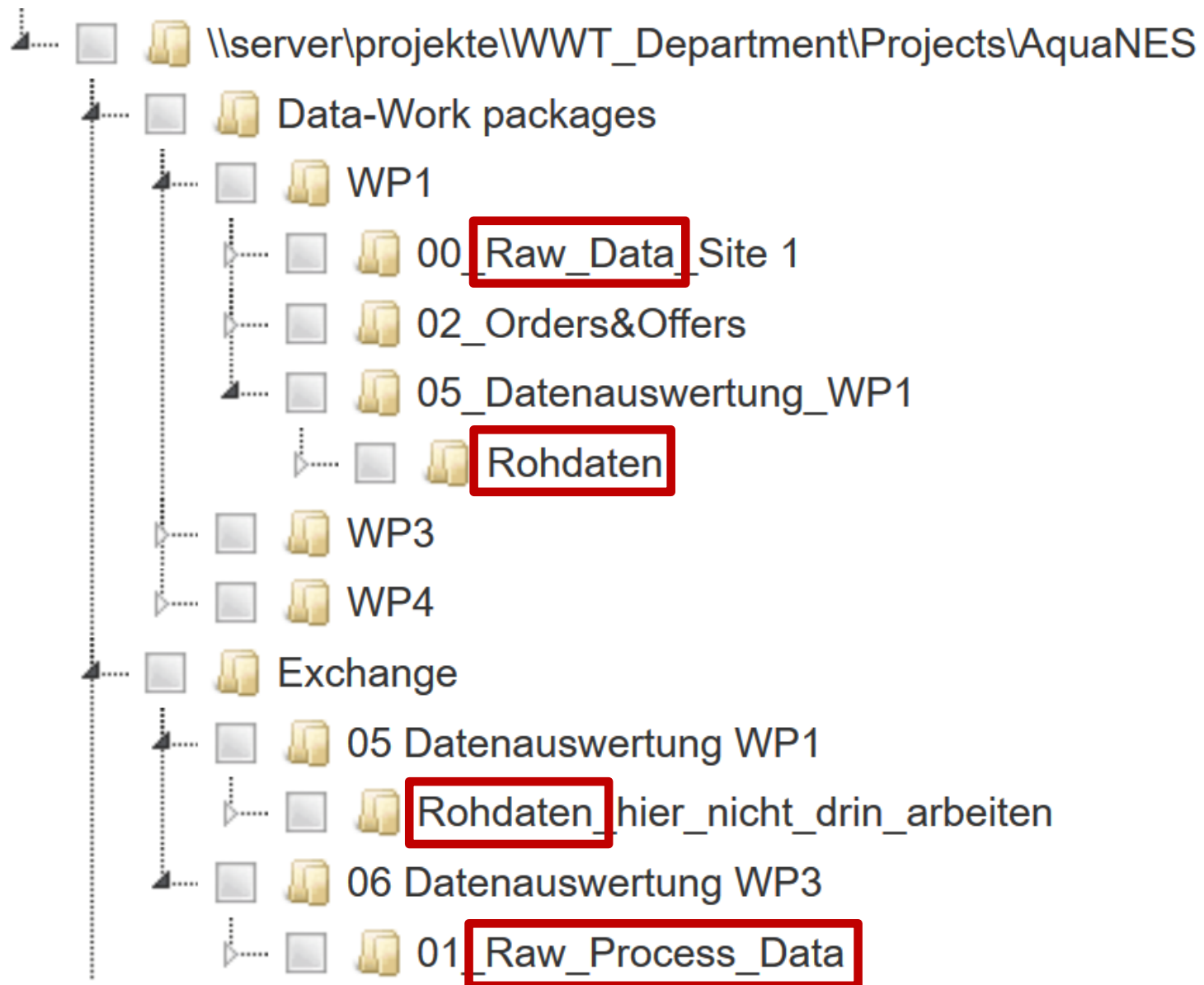


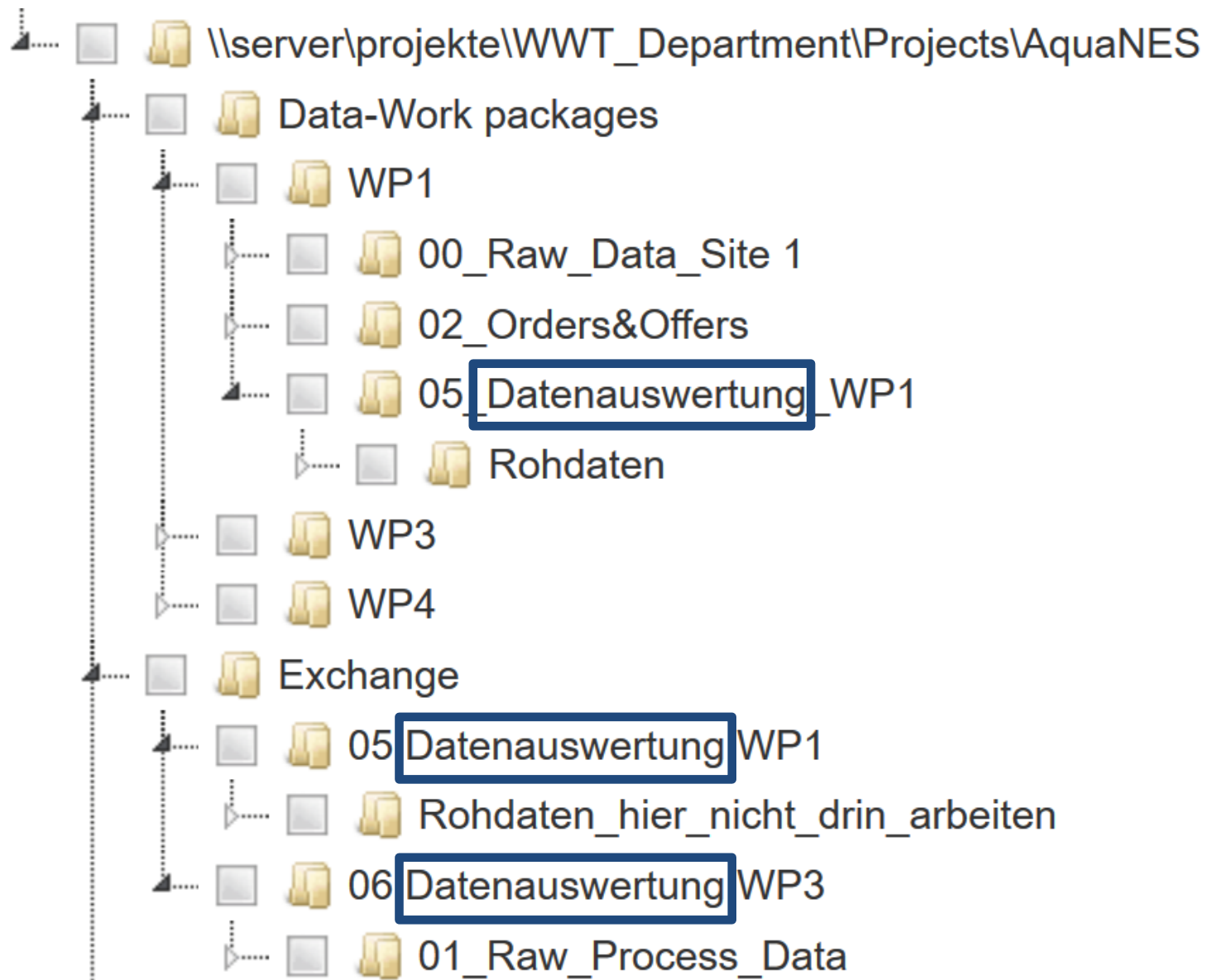


Langer Pfad zum Testprojekt



Langer Pfad zum Testprojekt

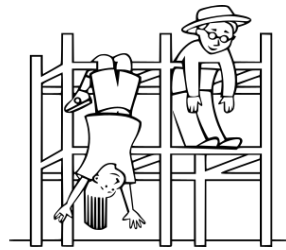
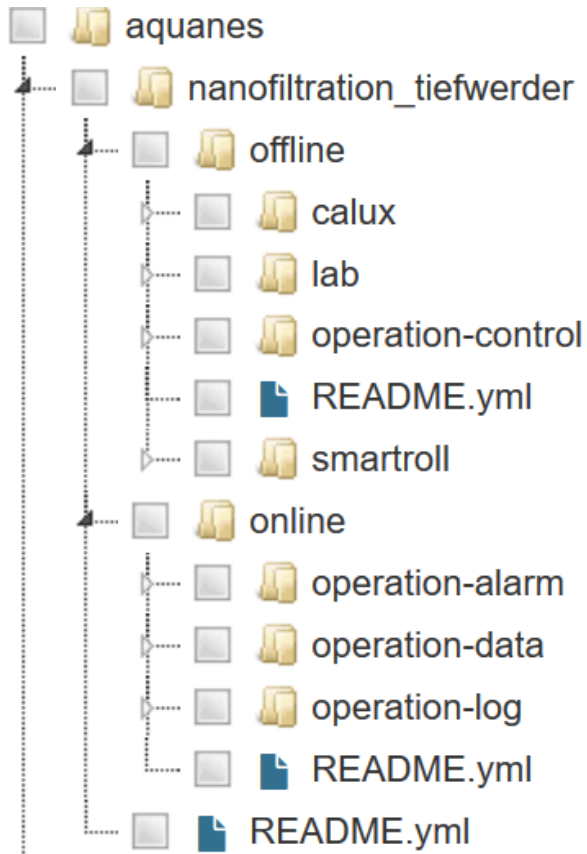






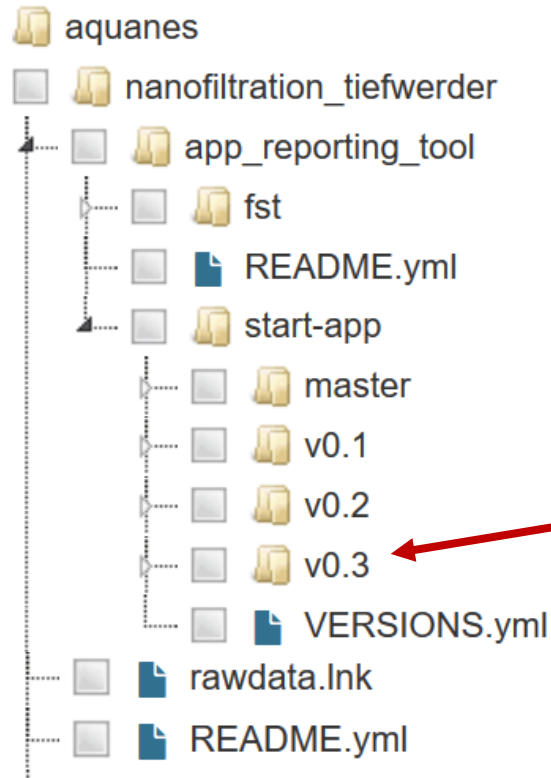
Rohdaten

\\server\rawdata



Datenverarbeitung

\\server\processing



Ergebnisse

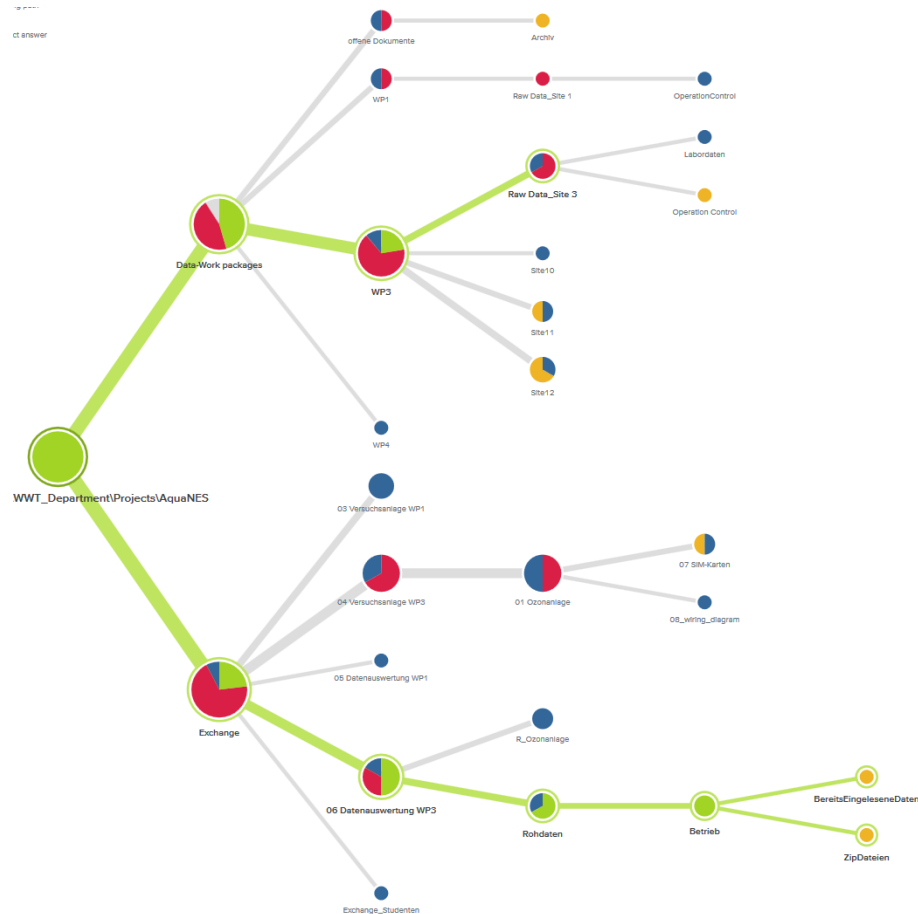
\\server\projekte

Link auf beste verfügbare Version

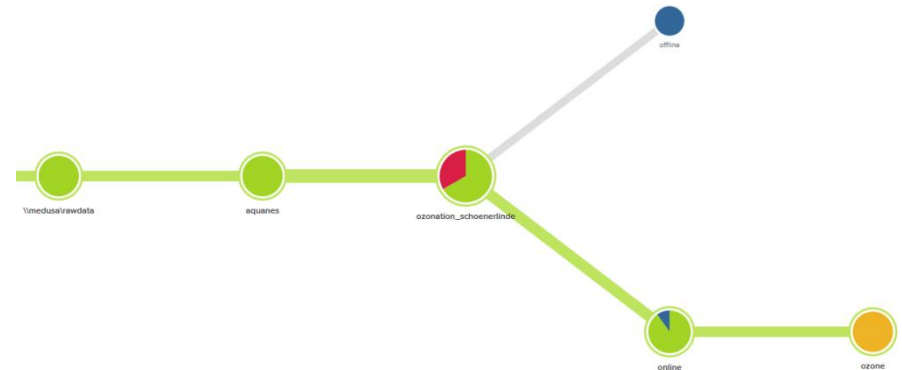


Wo finde ich die Rohdaten des Betriebes der Ozonanlage in Schönerlinde, die im Projekt AquaNES erhoben wurden?

Bisherige Ordnerstruktur



Neue Ordnerstruktur



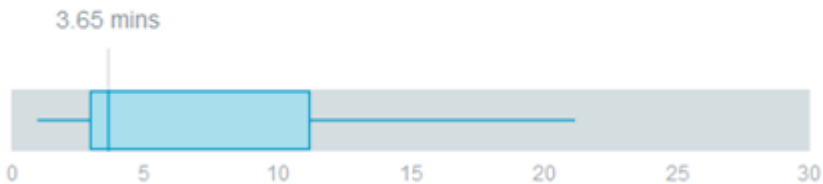
Tool: [Treejack](#)

Wo finde ich die Rohdaten des Betriebes der Ozonanlage in Schönerlinde, die im Projekt AquaNes erhoben wurden?

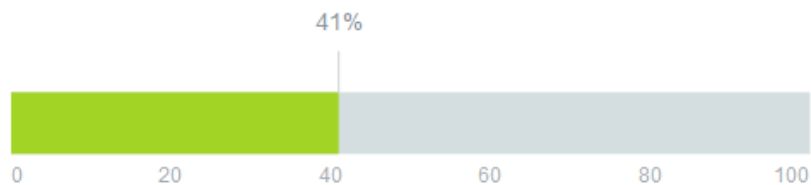
Bisherige Ordnerstruktur

n = 9

Time taken ?



Success ?



Directness ?



Neue Ordnerstruktur

n = 9

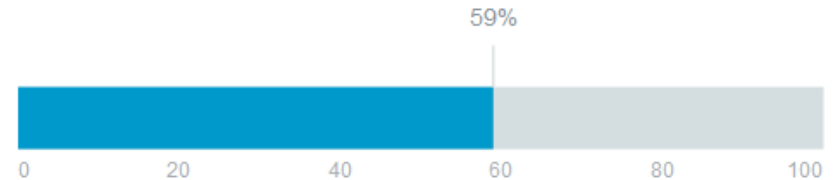
Time taken ?



Success ?



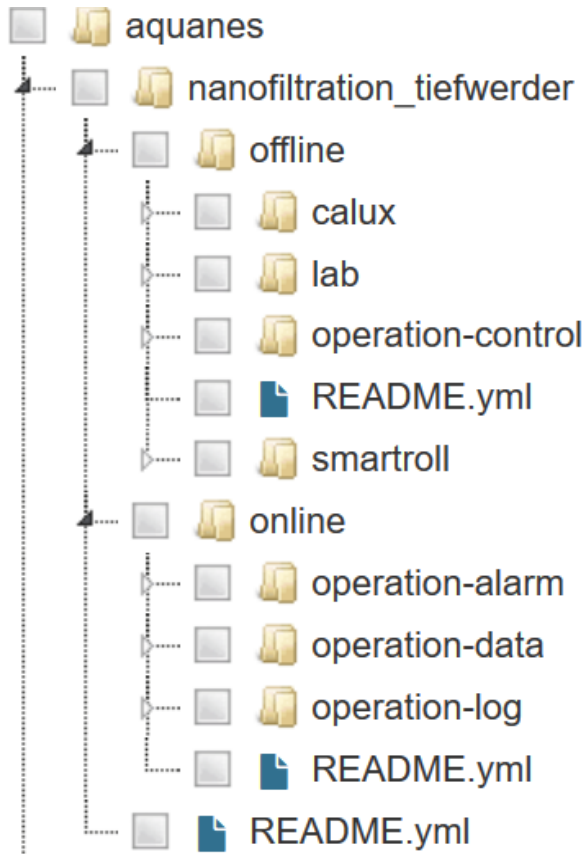
Directness ?





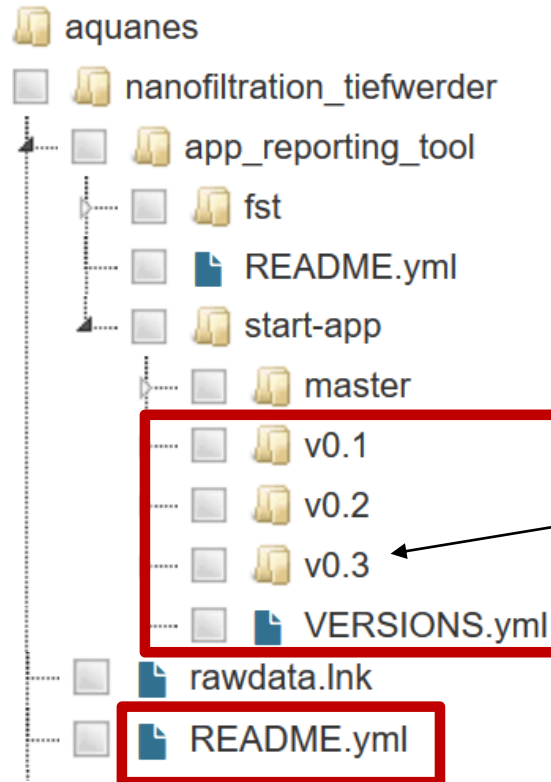
Rohdaten

\\server\rawdata



Datenverarbeitung

\\server\processing



Ergebnisse

\\server\projekte

Link auf beste verfügbare Version

Metadaten im YAML-Format

```
*\\medusa\\processing\\aquanes\\nanofiltration_tiefwerder\\README.yml - Notepad++
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
---
1  ---
2  README.yml:
3      created_by: "Michael Rustler"
4      last_modified: 2019-03-06
5      last_modified_by: "Michael Rustler"
6
7  app_reporting-tool:
8      description: "App for nanofiltration pilot plant Tiefwerder"
9      documented_at: "https://github.com/KWB-R/aquanes.report"
10     documented_by: "Michael Rustler"
11
```

Tool:



Integration des FDM in unser QMS

1. Berücksichtigung bereits bei der **Projektentwicklung** (Daten-/Softwaremanagementpläne)
2. Regelmäßige Überprüfung der Umsetzung unserer “Best practices” im Rahmen der **Project Reviews**
3. **Publikationen** (in öffentlichen Repositorien)

„mindestens 90% aller Projektberichte werden auf Plattform [Zenodo](#) hochgeladen“

FDM Tools made by FAKIN

Tool 1: Pfadanalyse-App



Analyse Paths

Table Statistics Duplicates Sankey Treemap Scatter Frequencies Explore

Load saved paths from

20190809_WWT_Department

Paths are relative to: //server/projekte/WWT_Department

Search:

	toplevel	folder	filename	extension	type	size	modified	depth
	["Pro	aquan	All	All	All	All	All	All
46970	Projects	AquaNES	Administration		directory		2017-09-04	3
46971	Projects	AquaNES	Closure & Evaluation		directory		2012-03-06	3
46972	Projects	AquaNES	Communication		directory		2019-03-19	3
46973	Projects	AquaNES	Data-Work packages		directory		2019-06-05	3
46974	Projects	AquaNES	Exchange		directory		2019-07-24	3
46975	Projects	AquaNES	Reports		directory		2019-07-24	3
46976	Projects	AquaNES	SyncToy_145f67da-70a3-4771-be57-4a798af2dfa7.dat	.dat	file	0.000	2019-03-12	3
46977	Projects	AquaNES	SyncToy_4b98f4aa-04f5-4115-80e8-a7b209eae704.dat	.dat	file	0.000	2018-12-14	3
46978	Projects	AquaNES/Administration	Budget		directory		2019-07-25	4
46979	Projects	AquaNES/Administration	Contracts		directory		2018-08-01	4

Showing 1 to 10 of 11,198 entries (filtered from 101,876 total entries)

Previous 1 2 3 4 5 ... 1120 Next

<https://kwb-r.gitlab.io/fakin/app/fakin.path.app>

Tool 1: Pfadanalyse-App



Analyse Paths

Table Statistics Duplicates Sankey Treemap Scatter Frequencies Explore

Indicator	Value
File/folder name quality	40.8 %
Number of files/folders matching '^Dok1'	0
Number of files/folders matching '^Mappe1'	0
Number of files/folders matching '^Pr.sentation1'	0
Number of files/folders matching '^Kopie von'	10
Number of files/folders matching 'gdalwmscache'	0
Number of files/folders matching 'win-library'	0
Number of files/folders matching '^README\.txt\$'	1
Number of files/folders matching '^README\.yaml\$'	0

Longest paths

//server/projekte/WWT_Department/Projects/AquaNES/Exchange/06 Datenauswertung WP3/01_Raw_Process_Data/00_Site_Operation/01_Operation_Control/02_Operation_Control_CW/03 Akkumulierter Durchfluss CW digitalisiert/Archiv/2018_11_27_Bedvolumina_CW_VolPerSquareMeter.xlsx	length 265
---	---------------

Longest filenames

Rahmenbedingungen für die umweltgerechte Nutzung von behandeltem Abwasser zur landwirtschaftlichen Bewässerung.pdf	length 118
---	---------------

Tool 1: Pfadanalyse-App

← →

http://127.0.0.1:3530/

🔍 ↻

Analyse Paths

✕

🏠 ★ ⚙️

Analyse Paths

Table

Statistics

Duplicates

Sankey

Treemap

Scatter

Frequencies

Explore

Min. File Size in MiB

10

1,000

1 201 401 601 801 1,000

Show

10

entries

Search:

	filename	size	count	potential
1	MA_SW.pdf	67.439	2	67.439
2	NewOzoneLst.Rdata	51.483	2	51.483
3	OzoneNew2_2019_KW30.csv	44.196	2	44.196
4	Bedienungsanleitung_Aqua-TROLL-100-200.pdf	34.423	2	34.423
5	Handbuch AQASYS 8.0 - A2013.05.pdf	27.495	2	27.495
6	Exo_KW29_KW38.csv	19.705	2	19.705
7	water-10-01599-2.pdf	15.947	2	15.947
8	Demoware D3.2- Show-case-of-the-environmental-benefits-and-risk-assessment-of-reuse-schemes(1).pdf	11.927	2	11.927

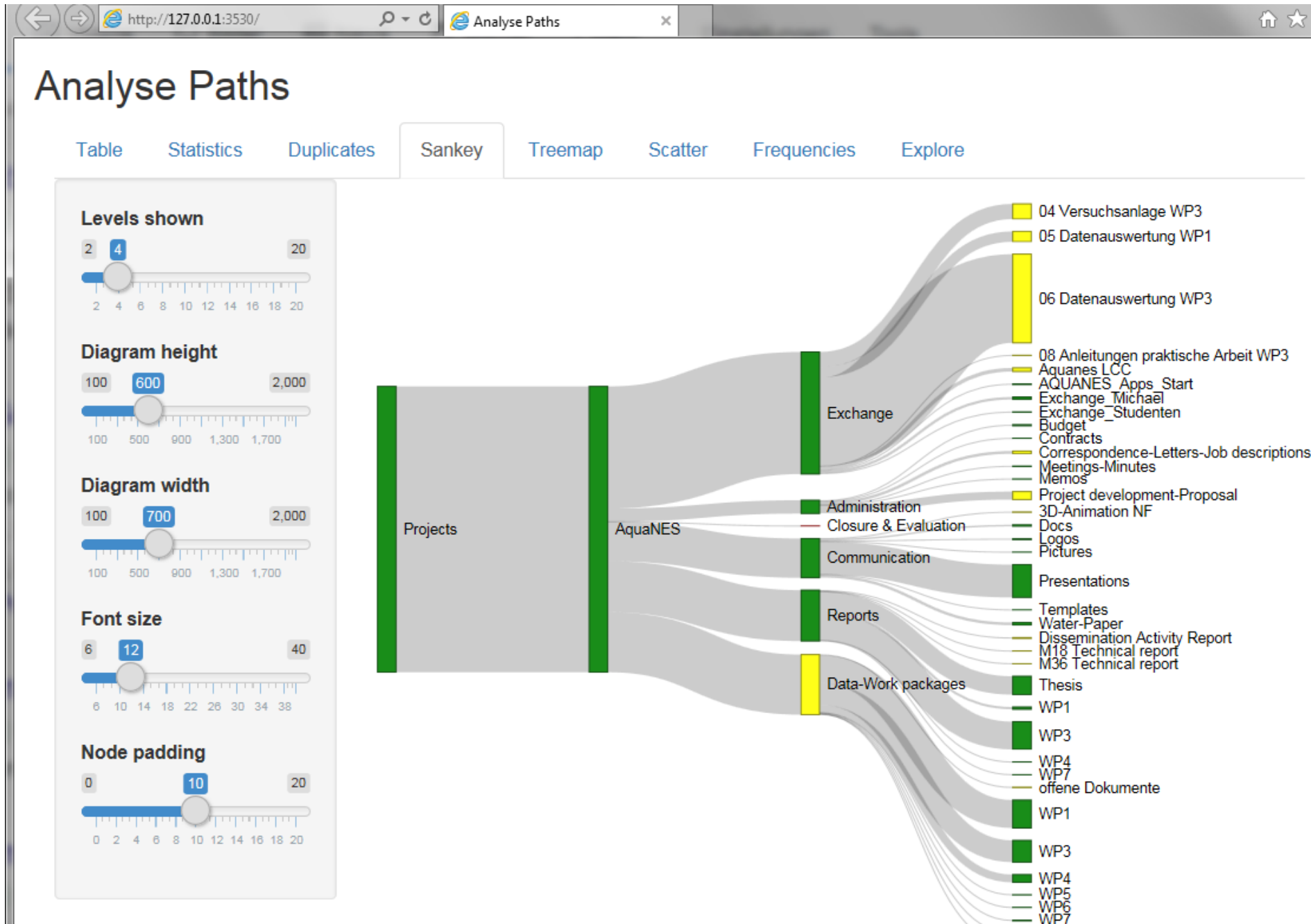
Showing 1 to 8 of 8 entries

Previous

1

Next

Tool 1: Pfadanalyse-App



Tool 1: Pfadanalyse-App



Analyse Paths

Table

Statistics

Duplicates

Sankey

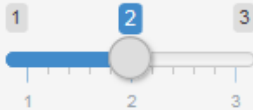
Treemap

Scatter

Frequencies

Explore

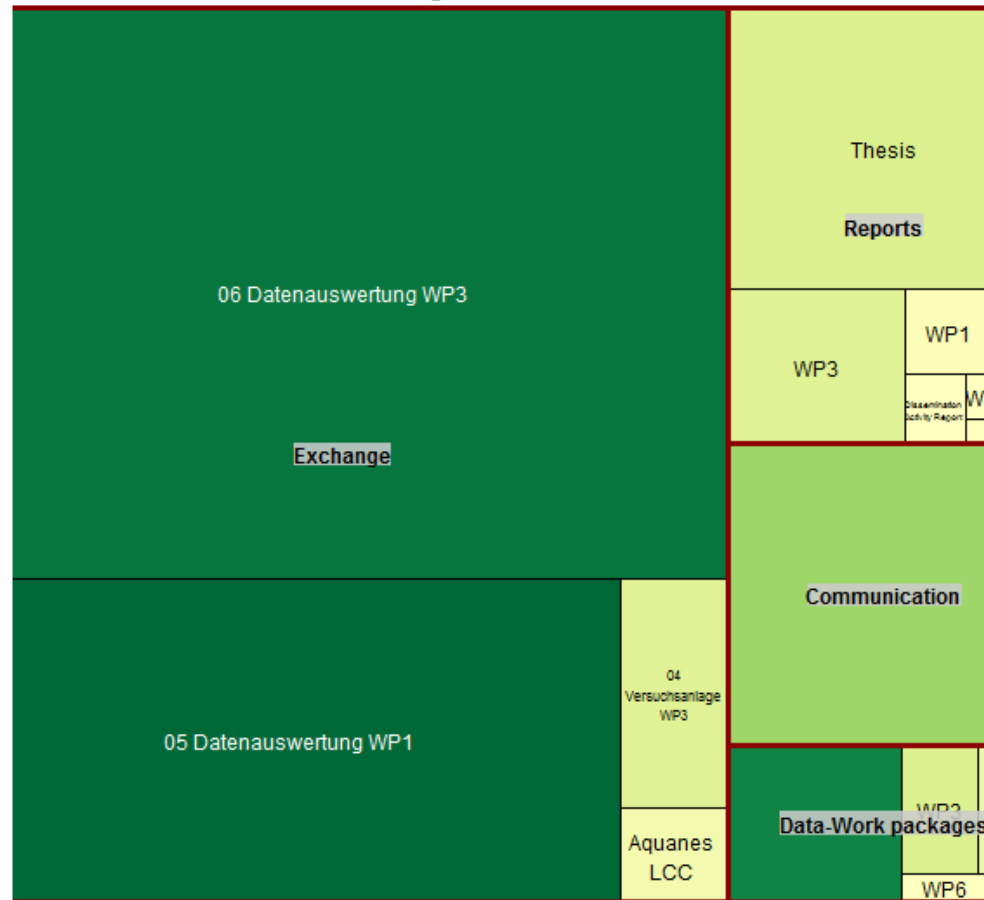
Levels shown



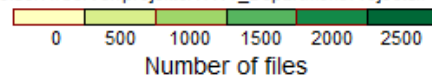
Rectangle area represents

- ☒ total size
- ☐ total number of files

Rectangle size = total size



Root folder: //server/projekte/WWT_Department/Projects/AquaNES

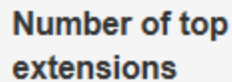


Tool 1: Pfadanalyse-App



Analyse Paths

Table	Statistics	Duplicates	Sankey	Treemap	Scatter	Frequencies	Explore
-------	------------	------------	--------	---------	---------	-------------	---------



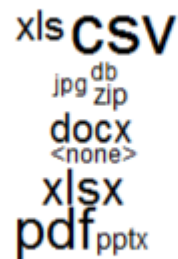
Min. cex



Max. cex Multiple



Filename Extensions



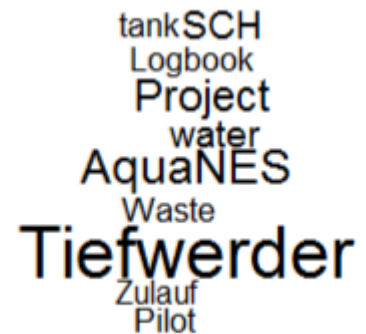
Filename extension	Count
csv	2224
pdf	1706
xlsx	1303
xls	990
docx	959

Words in Folder Names



Word in folder name	Count
erledigt	49
Archiv	34
cache	31
viewer	21
Data	19

Words in Filenames



Word in filename	Count
Tiefwerder	1268
AquaNES	715
Project	626
SCH	476
water	457

Tool 2: FAKIN Wissensspeicher



<https://kwb-r.gitlab.io/fakin/#tool>



FAKIN Latest **Tools** Projects Apps Code People Publications Tags Talks



Tools

used within FAKIN

Collaboration Data Analysis Data Management Plan DOI Knowledge Repo
Learning Metadata Programming Publication Repository Reproducibility
Survey Search Testing Version Control Web Hosting All

TREEJACK

Information
architecture validation
software

LIMESURVEY

Professional online
surveys with
LimeSurvey. The
worldwide leading open
source survey software
as a professional SaaS
solution or as a self-
hosted Community
Edition.

People

involved in the project

Team FAKIN



Michael Rustler

Project Manager
FAKIN, Data scientist

Reproducible research,
Data management,
Programming (R &
Python)



Hella
Schwarzmüller

Head of Department
"Groundwater"



Hauke
Sonnenberg

Researcher



Christoph
Sprenger

Head of Department
"Groundwater"

Test Projects



Lea
Conzelmann

Researcher



Michael Stapf

Researcher



Andreas
Matzinger

Researcher



Nicolas Caradot

Researcher



Jeannette
Jährg

Researcher



Christian Menz

Researcher

Communication (internal)



Lea
Conzelmann

Researcher



Anne
Kleyböcker

Researcher



Ulf Mieke

Deputy Director, Head
of Department
"Process Innovation"



Christian Remy

Researcher



Pascale Rouault

Deputy Director, Head
of Department "Urban
Systems"



Bodo Weigert

Deputy Director, Head
of Finances,
Administration,
Communication



Marcus
Weispfennig

IT Administrator



Daniel Wicke

Researcher



Malte Zamzow

Researcher



Michael Stapf

Researcher



Wolfgang Seis

Researcher



Andreas
Matzinger

Researcher



Mathias Riechel

Researcher



Nicolas Caradot

Researcher



Jeannette
Jährg

Researcher



Christian Menz

Researcher



Fabian Kraus

Researcher



Christian
Loderer

Researcher



Rabea-Luisa
Schubert

Researcher



Roberto Tatis
Muvdi

Researcher



Tobias Evel

Commercial project



Kuangxin Zhou

Researcher