### **NFORMATION**. *plus!* Universitätsbibliothek



### **Conquaire: Continuous quality control for research data to ensure reproducibility**

Vidya Ayer, Philipp Cimiano, Fabian Herrmann, Vitali Peil, Christian Pietsch, Andreas Rempel, Jochen Schirrwagen, Johanna Vompras, Cord Wiljes contact: christian.pietsch@uni-bielefeld.de - ORCID: https://orcid.org/0000-0001-8778-1273 - license: CC-BY 4.0 - last modified: June 3, 2019 https://conquaire.uni-bielefeld.de

#### About

We present progress in Conquaire, a DFG project to foster the analytical reproducibility of research results at Bielefeld University, Germany, conducted by CITEC (Center of Excellence in Cognitive Information Technology), Bielefeld University Library and 9 research groups from 2016 to 2019. DFG grant number: 277747081.

### 3. Continuous data quality checks

Inspired by successful software development practices such as versioning and continuous integration (CI), we apply these ideas to research data management (RDM).



## Motivation

**Reproducibility** of research results enables peer validation. However, reproducing research results is a major challenge:

- Nature survey (1576 scientists): 50–70 % reproducibility failures
- Reproducibility Project: 61 % failure rate in Psychology
- pharmaceutical clinical drug trials: 82 % failure rate

Universität Bielefeld

## Conquaire approach to improve data quality

Prime directive: Do not to create any extra effort for researchers, and not to stand in their way. Instead, we try to tap into their existing workflows in 4 easy steps:

- 1. offer an institutional GitLab instance with Large File Storage
- 2. offer training in Git and GitLab as well as for PUB (see below)
- 3. offer data quality checks implemented in GitLab CI (see below)
- 4. connect GitLab to our institutional publication repository PUB for

The GitLab component GitLab Cl can automatically run arbitrary programs whenever files are added or changed. These programs run on a separate server within a temporary Docker container. We provide a .gitlab-ci.yml file that runs them.

It checks for basic FAIR data principles compliance and tries to validate the most popular data file types.

### 4. Publishing datasets

Bielefeld University's institutional publication repository PUB is the place where researchers enter their publications, including data publications. PUB is an in-house development based on the free and open-source LibreCat software.



Ρ	UB front-end:			
a de la compañía de la	PUB - Publications at Bielefeld University - Mozilla Firefox			3 ×
PUB - P	ublications at Bielefe × +			
$\leftrightarrow$	C' 🏠 🕕 https://pub.uni-bielefeld.de 🛛 💀 🏠 👱 III\ 🔍 Search 🥔 😇 💈 🧔	👶 S 🗉 🔬 👫 ?	L 💕 🗟	≡
Pub	lications at Bielefeld University	Deutsch		
Home	Publications Authors Awards About -	+∋ Login		
"PUB depot univer public	– Publications at Bielefeld University <sup>*</sup> is the institutional repository of Bielefeld University. It constitutes the central for publication metadata, publications and research data with the mission to reflect the scientific work of the sity's researchers. Bielefeld scientists of all disciplines can use this service to create and maintain their personal ation lists. The records are freely accessible and partially linked with fulltexts and research data. Learn more	Open Access at Bielefeld University		
0	61357 Publication References Bibliographic data, partially enriched with fulltexts (PDFs etc.)	Research Data Management		
0	4109 Individual Author Pages Documentation of Bielefeld researchers' publishing activities	<u> </u>		
0	252 Data Publications Bibliographic data, partially enriched with research data	PUB Theses		
		1 0 0 1110305		

instant data publishing incl. DOI minting, linking, and archiving

### 1. GitLab.UB

Sign in · GitLab - Mozilla F	Firefox _ =		Projects · Dashboard · GitLab -	• Mozilla Firefox	
Sign in · GitLab × +					
) → C û https://gitlab.ub.uni-bielefeld.de/user ···· 🕸 ★ 🛓	🛝 🔍 Search 🥔 🗟 😨 🔞 S 🗉 » 🗄	← → C l l l https://gitlab.u	b.uni-bielefeld.de/dasi 🚥 🕸 🛧 🛓	III\ Q Search	🧈 🖯 🧕 🛡 👶 S 🗉 🗆
Universität Bielefeld		Bolversitik Eduted Projects ✓ Groups ✓ More ✓	∕ 🖿 🖌 🖬 🗸 s	earch or jump to	Q ();24 (1) (25 Q ~ 🍕
		Drojecto			Newpro
		Projects			New pro
GitLab.UB: Bielefeld University's	Sign in	Your projects 209 Starred projects 9 E	Explore projects	Filter by name	Last updated
collaborative development	Username or email	All Personal			
environment		Kompetenzzentrum Forschungsda	aten / Organisation A		
	Password	Dies ist das zentrale Repositorium des	Kompetenzzentrums Forschungsdaten für di	na te inte na te inte	Updated 4 hour
Universität Bielefeld Universitätsbibliothek		, PUB / librecat-core A Maintainer	r)		
This is a GitLah instance bosted at Rielefeld University and run by <b>Bielefeld</b>	Remember me Forgot your password?	Mirror of the original LibreCat applicat	ion from https://github.com/LibreCat/LibreCa	nat 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Updated 6 hour
University Library's LibTec department.	Sign in	D BASE / Protokalla A Maintainer		<b>≜</b> 0 ೪0	Lindated 1 da
Members of Bielefeld University please use the "BITS single sign-on"				NO 80	opulled i da
In case of difficulties, please contact <b>Christian Pietsch</b> .	Signin with	BASE / frontend	6 / / / · / · · · · · · ·	<b>★</b> 0 ¥0	Updated 1 da
······································		Hier gibt es das Beste aus /usr/local/vu	ufind/ alias /base/ vom BASE-Webserver		
		F Christian Pietsch / frontend A	Maintainer	<b>★</b> 0 Ÿ0	Updated 2 day
		N oa2020-de / <b>nikola-website</b> O ( Forget Jekyll, here comes Nikola!	Maintainer	🛊 0 🖞 0	Updated 2 day
Explore Help About GitLab					
		C Conquaire / conquaire-book Q ( The Conquaire Book	Owner	<b>★</b> 0 ¥2	Updated 2 day
.gitlab-ci.yml · master · Conquaire / quality_chee	cks · GitLab - Mozilla Firefox - 🔹		Admin Area · GitLab - Moz	illa Firefox	
itlab-ci.yml · master · Cor × +	IN O Search 🦧 🖂 🖪 🐽 🚜 S 🗐 🚿 🖛	$ \begin{array}{c c} & & & \\ & & & \\ & & & \\ \hline & & & \\ \hline & & & \\ \hline & & & \\ & & & \\ \hline & & & \\ \hline & & & \\ & & & \\ \hline \\ \hline$	h uni-bielefeld de/adm 🛛 💀 🕭 🗸	III\ Q Search	
Projects × Groups × More × 🐚 🗡 🕂 Sear	rch or iump to 9 D) 24 D1 (25 Q ~ (A) ~	Projects × Groups × More ×		earch or jump to	Q []]24 []] [25 Q ~ (
		Admin Area > Dashboard			
B.gitlab-ci.yml 1.33 KB <sup>C</sup> <sup>™</sup>	Image: Web IDE     Replace     Delete				
1 quality-check:		Brojects: 975	llsors: 5	51	Groups: 103
<pre>image: python:3.6-alpine 5</pre>					010405.100
<pre>6 before_script: 7 # copy python scripts on Conquaire server to execute the</pre>	m from the server	(°1) New project	New user		New group
<pre>8 - rm -rf /opt/conquaire/quality_checks/src 9 - cp -R src/ /opt/conquaire/quality_checks/ " Drift CT Variables form sitlab</pre>		<u>ل</u>			
<pre>10 # Print CI Variables from gittab. 11 - echo "USER_NAME "\${GITLAB_USER_NAME} 12 - echo "USER_EMATL "\${GITLAB_USER_EMATL}</pre>		Statistics	Features		Components
<pre>13 - echo "PROJECT_URL "\${CI_PROJECT_URL} 14 - echo "PROJECT_PATH "\${CI_PROJECT_PATH}</pre>		© Forks	111 Sign up	ധ	GitLab 11.8.1 (657d50
<pre>15 - echo "COMMIT_SHA "\${CI_COMMIT_SHA} 16 # Create temporary mail configuration files.</pre>		lssues 1	4,939 LDAP	Ċ	GitLab Shell 8.4
<pre>17 - mkdir /etc/ssmtp 18 - echo "root=\${GITLAB_USER_EMAIL}" &gt; /etc/ssmtp/ssmtp.co 10 - echo "root=\$</pre>	nf	Merge Requests	273 Gravatar	Ċ	GitLab Workhorse v8.
<ul> <li>ecno "mailhub=conquaire.uni-bielefeld.de" &gt;&gt; /etc/ssmt</li> <li>echo "hostname=gitlab-runner.conquaire.uni-bielefeld.d</li> <li># Install pythop lyml and ssmtp package for conding food</li> </ul>	<pre>p/ssmcp.cont e" &gt;&gt; /etc/ssmtp/ssmtp.conf back mail in the docker container</pre>	0) Notes 4	42,947 OmniAuth	•	GitLab API
<ul> <li>22 * Instatt python txmt and ssmtp package for sending feed</li> <li>22 - apk add py3-lxml ssmtp</li> <li>23</li> </ul>	Dack mait in the oocker container.	© Snippets	45 Reply by email	U	Ruby 2.5.3p1
24 script:		SSH Keys	263 Container Registry	•	Kails 5.0.
<pre>25 # - ts -ta /var/www/ntmt/feedback 26 - /usr/bin/python3 /opt/conquaire/quality_checks/src/mai 27</pre>	<pre>n.py -f /var/www/html/feedback/ -l "https://conquaire.</pre>	Milestones	195 GitLab Pages	C	postgresql 9.6
<pre>27 28 # Choose docker CI runner on gitlab server. 20 tags:</pre>		Active Users	547 Shared Runners	•	Gitaly Servers
30 - dockerexec					
		Latest projects	Latest users		Latest groups
		>>	urs ano	3 hours ago	3 days a

As an alternative to the existing upload button, researchers can now select one of their GitLab.UB projects to publish any version and get a DOI for it. PUB talks to the GitLab API.

Data publications can be linked to conventional publications. Metadata can be re-used in the Linked Open Data spirit via an open API in formats such as MODS and DataCite XML.

Soon, PUB will display a badge indicating the overall result of the data quality checks from step 3.

Ð			PUB Theses
·	11777 Open Access Public Freely accessible documents	cations	
0	2073 Theses		i
	Bielefeld dissertations, habilitat	ion, and selected master and bachelor theses	ORCID
DUI			ORCID ORCID
PUB	B Training für Einsteiger Teil 1: Persönliche I	Publikationslisten erstellen Di., 07.05.2019 09:15 Uhr (ca. 90	Min.)
PUB	für Einsteiger Teil 2: Dateien ergä	nzen, Daten exportieren Di., 26.03.2019 09:15 Uhr (ca. 90 Min	AUTHINICATE DISPLAY 2018
Welt	weit eindeutige Autorenidentifikat	tion mit ORCID Mi., 27.03.2019 09:15 Uhr (45 Min.)	
Inter	ested in (more) information about	PUB, but no time to join the scheduled events?	
We w	vill gladly advise you on how to best u tes or 2 hours - single, individually, a	use PUB and answer the questions you might have. No matter if t your workplace. Please contact us!	it's 10
	and a subsection of the subsec		
Contact	Policy Legal Notice		Support Publication Services
vrivacy s	Statement	Connecting Research and Researchers	Susanne Riedel Phone: 0521/106-4058
Powered	by LibreCat	Further Memberships & Certificates	publikationsdienste.ub@uni-blelefeld.de
ø 2010-2	ors bieleteta university Library		
IKATIONE ERSITAT EFELD	N # Home I My Publicatio	on List @ 🖓 PUB How-to	🏴 Deutsch 🍰 cpietsch 🗸 🕞 Lo
	Deserve b De	4-	
	Research Da	ita	
	Research Da	IIA	
	Research Da	✓ Save × Cancel ♠ Submit	
	Research Da	Save X Cancel & Submit	
	Research Da	✓ Save × Cancel & Submit	
	Details	Save × Cancel Submit	
	Details Publication Type	✓ Save × Cancel	-
	Details Publication Type DOI	Save × Cancel Submit   Research Data   e.g. 10.4119/unibl/2695161	I want to register a new DOI
	Creator *	✓ Save × Cancel	□ I want to register a new DOI ○ extern ● ••••• •
	Research Da Details Publication Type DOI Creator *	Save     Cancel     Submit       Research Data       e.g. 10.4119/unibl/2695161       First Name     Chris       Last Name     Pietsch	□ I want to register a new DOI ○ extern ● □unBit + -
	Research Da Details Publication Type DOI Creator * Department	Save × Cancel Submit   Research Data   e.g. 10.4119/unibl/2695161   First Name   Chrite   Last Name   Pietsch   Universitätsbibliothek	I want to register a new DOI  extern  t t t t t t t t t t t t t t t t t t
	Research Da Details Publication Type DOI Creator * Department Project	Save       Cancel       Submit         Research Data       e.g. 10.4119/unibi/2695161         First Name       Chris       Last Name         Universitätsbibliothek       Start typing, or use (arrow down) to choose a project from the second secon	I want to register a new DOI C extern C Guilet + - T the list.
	Research Da Details Publication Type DOI Creator * Department Project Research Group	Save       Cancel       Submit         Research Data       e.g. 10.4119/unibi/2695161         First Name       Chris       Last Name         Universitätsbibliothek       Start typing, or use (arrow down) to choose a project fro         Start typing, or use (arrow down) to choose a research g	I want to register a new DOI C extern C Tongit + - m the list. + -
	Research Da Details Publication Type DOI Creator * Department Project Research Group Titie *	Save × Cancel Submit   Research Data   e.g. 10.4119/unibl/2695161   First Name   Chris   Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g   e.g. Handbook of eBook Publishing	I want to register a new DOI         ● extern       ● □         ● extern       ● □         + -         m the list.       + -         proup from the list.       + -
	Research Da Details Publication Type DOI Creator * Department Project Research Group Titte * Publisher *	Save × Cancel Submit   Research Data   e.g. 10.4119/unibi/2695161   First Name   Chris   Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g   e.g. Handbook of eBook Publishing   Bielefeld University	I want to register a new DOI C extern  C Total + - m the list. + - group from the list. + -
	Research Da Details Publication Type DOI Creator * Department Project Research Group Title * Publisher * Year of Publication *	Save x Cancel Submit   Research Data   e.g. 10.4119/unibi/2695161   First Name   Chris   Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g   e.g. Handbook of eBook Publishing   Bielefeld University   e.g. 2015	I want to register a new DOI C extern  C Total + - m the list. + - group from the list. + -
	Research Da Details Publication Type DOI Creator * Department Project Research Group Title * Publisher * Year of Publication *	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chris     Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro     Start typing, or use (arrow down) to choose a research g     e.g. Handbook of eBook Publishing     Bielefeld University     e.g. 2015	I want to register a new DOI  C extern C Total T
	Research data un	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chris   Last Name   Pietsch   Universitätsbibliothek     Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g   e.g. Handbook of eBook Publishing   Bielefeld University   e.g. 2015   Published while	I want to register a new DOI  C extern  C C T is i + −  m the list. + −  proup from the list. + −  c c on temployed at UniBi
	Research data up	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chric     Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g      e.g. Handbook of eBook Publishing     Bielefeld University   e.g. 2015   Published while	I want to register a new DOI  C extern C G G G G G G G G G G G G G G G G G G
	Research data up Upload File(s) (max. 500 Dot Creator * Department Project Research Group Title * Year of Publication *	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chris     Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro     Start typing, or use (arrow down) to choose a research generative down) to choose a research generative down) to choose a research generative down to c	a not employed at UniBi
	Research data up Upload File(s) (max. 500 Drop files here to upload	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chris     Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g      e.g. Handbook of eBook Publishing     Bielefeld University   e.g. 2015   Published while   DOad	I want to register a new DOI   extern   Image: I want to register a new DOI   Image: I want to register a new DOI <td< td=""></td<>
	Research data up Upload File(s) (max. 500 Drop files here to upload V Copyright and Licen	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chris     Last Name   Pietsch   Universitätsbibliothek   Start typing, or use (arrow down) to choose a project fro     Start typing, or use (arrow down) to choose a research generation of eBook Publishing     Bielefeld University   e.g. 2015    Published while Doad XMB per file) or click!	e not employed at UniBi
	Research data up Upload File(s) (max. 500 Drop files here to upload Ve Copyright and Licen GitLab Project URL	Save Cancel     Research Data     e.g. 10.4119/unibi/2695161     First Name   Chrit   Last Name   Pietsch   Universitätsbibliothek     Start typing, or use (arrow down) to choose a project fro   Start typing, or use (arrow down) to choose a research g   e.g. Handbook of eBook Publishing   Bielefeld University   e.g. 2015   Published while    Noad   MB per file)   or click!   Situab Project URL e.g. https://gittab.ub.uni-bielefeld.dd	I want to register a new DOI   extern   Image: state in the list.   Image: state in the list. </td

#### **Related work**

This project was initiated by Najko Jahn's observation that researchers started to use GitHub for tasks other than software development (see the project proposal on the Conquaire website for quantitative details).

# 2. Training

- In 2016, delegates from 8 research groups (project partners) attended several hands-on Git training sessions based on the Software Carpentry course https://swcarpentry.github.io/git-novice/.
- Once in each semester since summer 2018, we have offered a Git and GitLab training session to Bielefeld University staff. They were almost fully booked.

Later, GitHub started to co-operate with Zenodo to offer a data publication service not unlike the one we provide for Bielefeld University but without pre-configured quality checks.

More recently, GitLab added a feature called Auto DevOps that provides pre-configured quality and security checks for software development.

Ongoing work		
How to display the badge?	Universität Bieleferdi       A. C.++ Implementation of the reversed Attentional Vector Sum (rAVS) model         Download	