

Zwischenvortrag für die Bachelorarbeit am Institut für Informatik
der Freien Universität Berlin

AG Software Engineering

Entwicklung eines Tools zur Ansammlung und Aufbereitung von Datensätzen aus Github für die Forschung im Bereich Technical Debt

Robert Selack
robfu@zedat.fu-berlin.de

Beteuer: Prof. Dr. Lutz Prechelt
Victor Brekenfeld

Gliederung

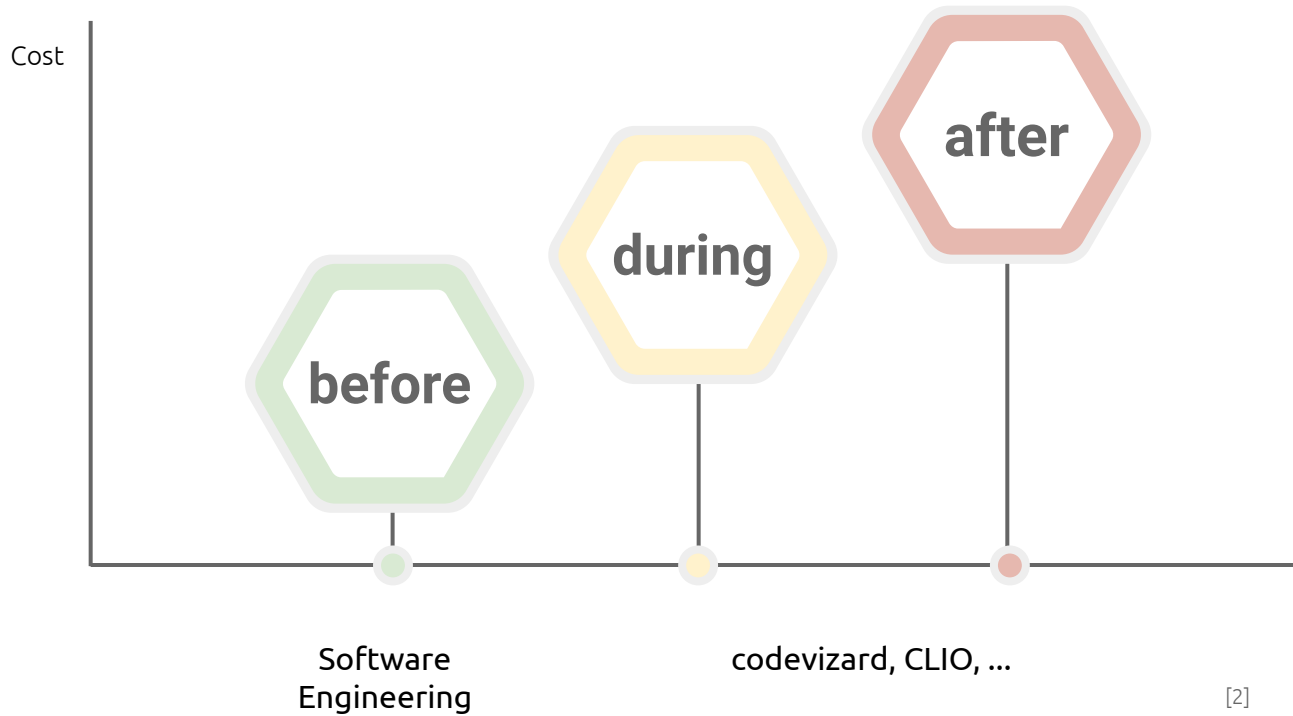
- 1 Einführung**
- 2 Motivation**
- 3 Umsetzung des Tools**

Technical Debt

“A design or construction approach that's expedient in the short term but that creates a technical context in which the same work will cost more to do later than it would cost to do now (including increased cost over time).”

- Steve McConnell [1]

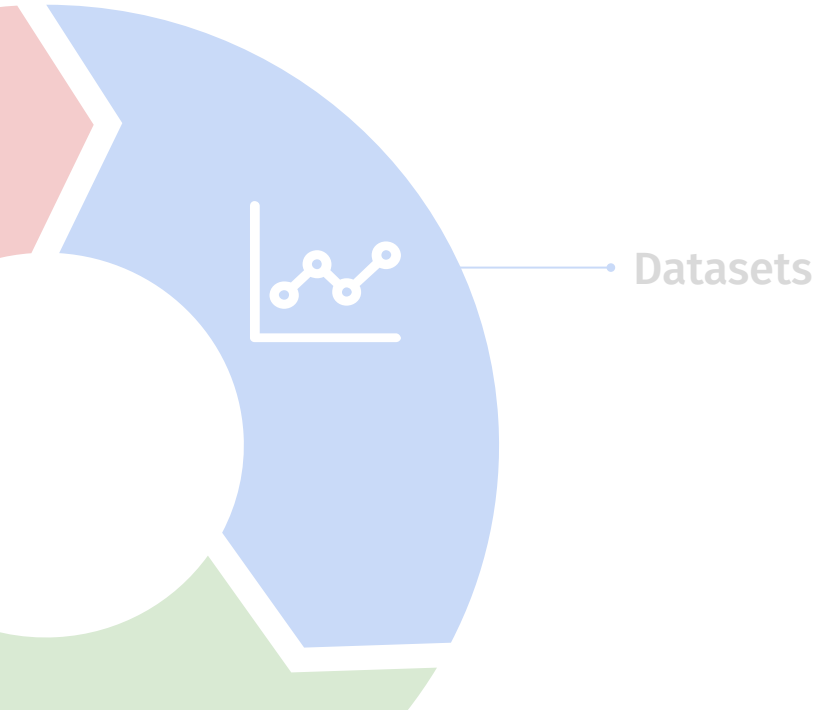
Einführung



Motivation



Motivation



- A. “typical enough to generalize to other projects”
 - B. “suitable for the given invention”
 - C. “has potential for interesting interaction regarding the introduction”
-
- Developer size

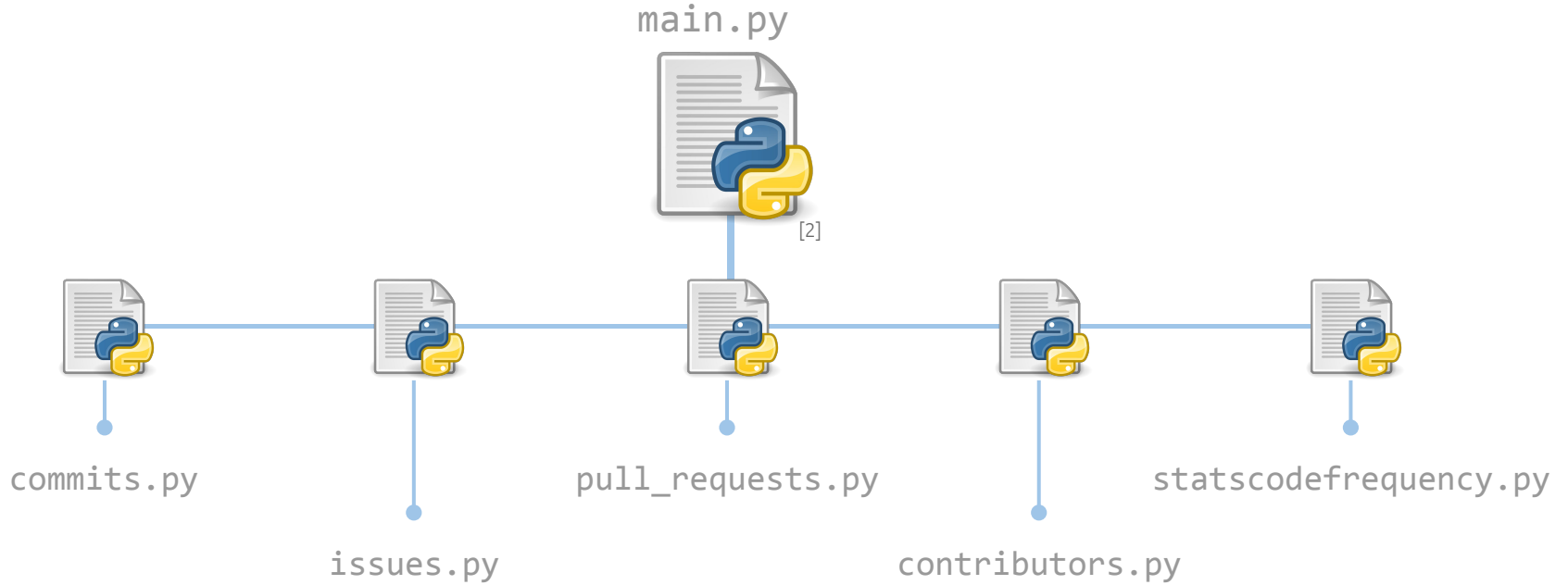
Umsetzung des Tools



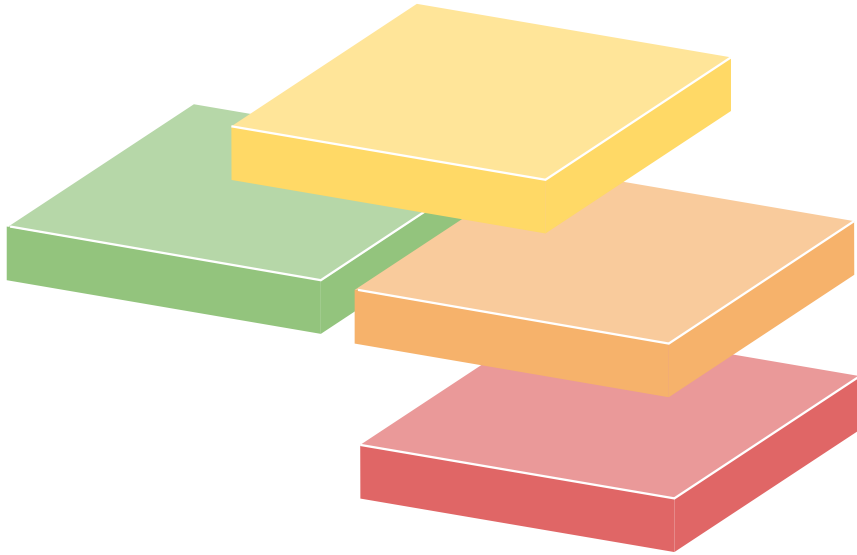
- 1 Python Script
- 2 Configuration File
- 3 Database

Umsetzung des Tools

1 Python Script



Umsetzung des Tools



1 Python Script

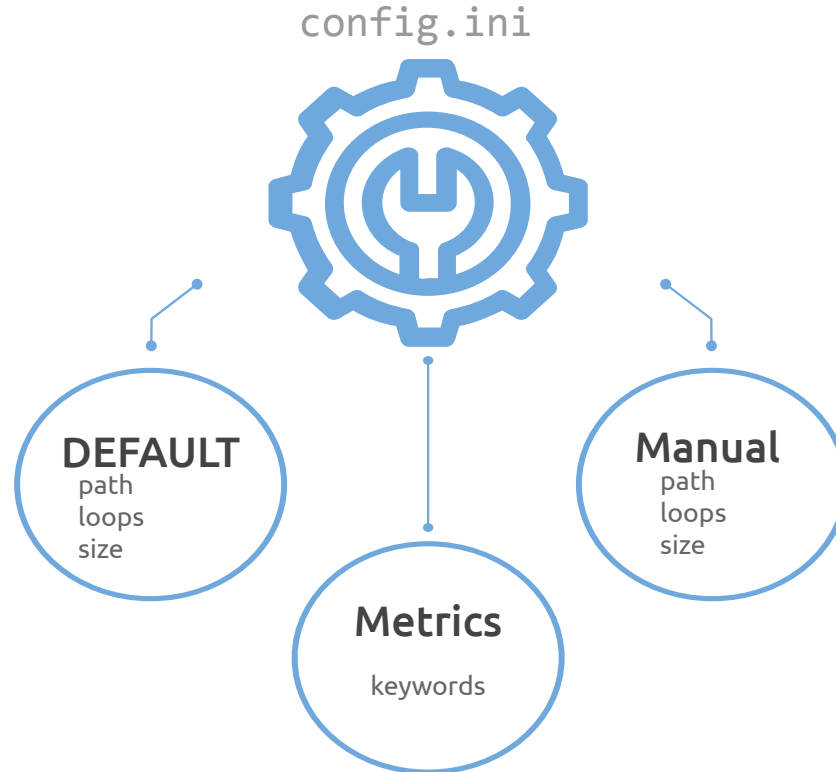
1 **Repository** 1/30

2 **Metric** 1/5000
Issues

2 **Metric** 1/5000
Contributors

3 **Issue** 1/5000
Comments

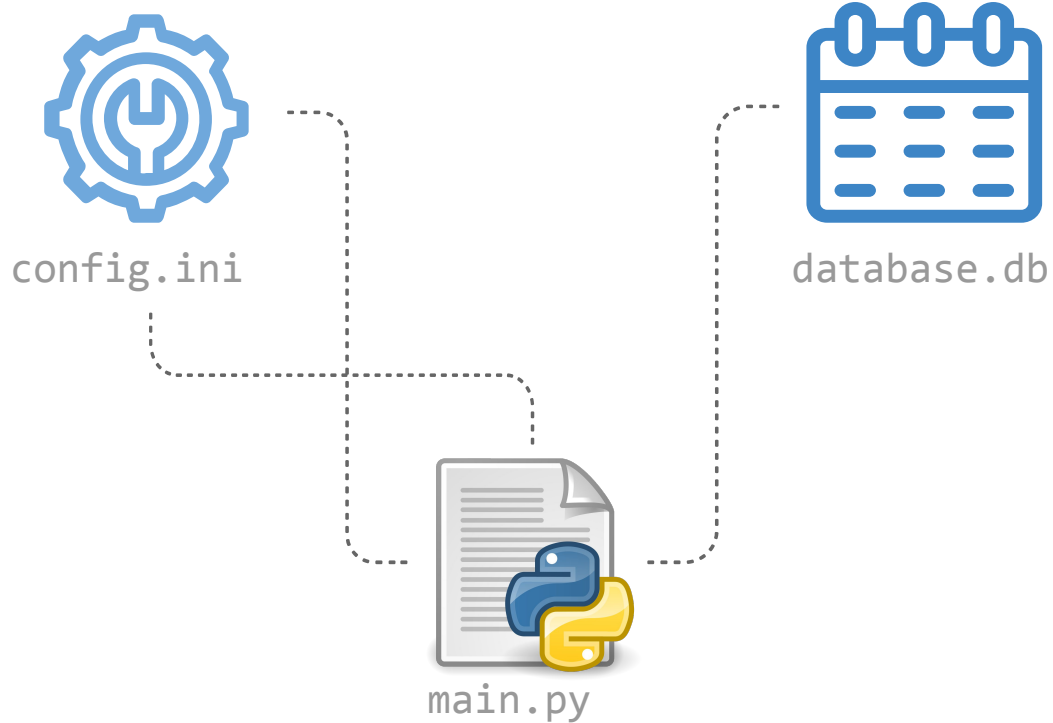
Umsetzung des Tools



Umsetzung des Tools

3

Database

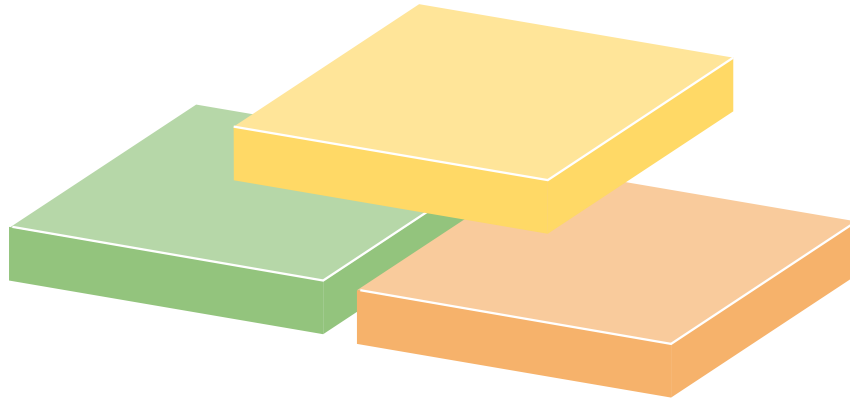
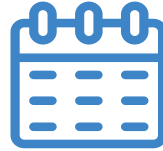


Umsetzung des Tools

3

Database

repositories



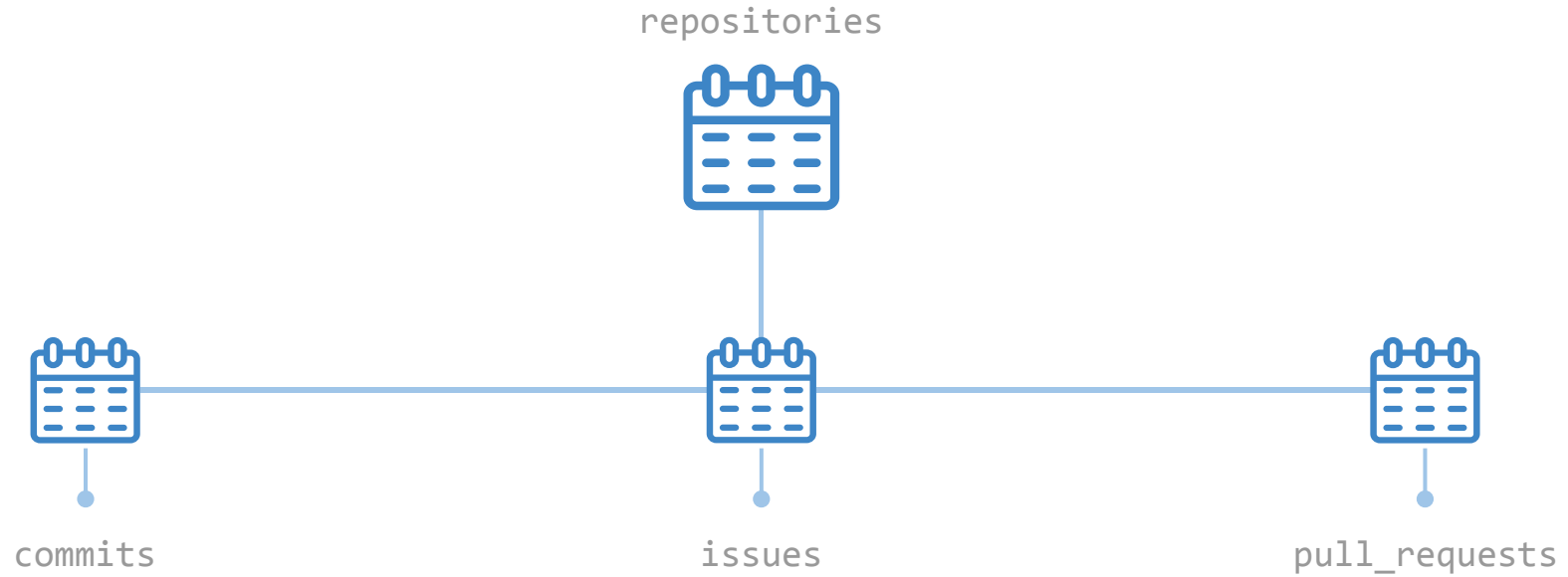
1

2

Umsetzung des Tools

3

Database



Umsetzung des Tools

3

Database



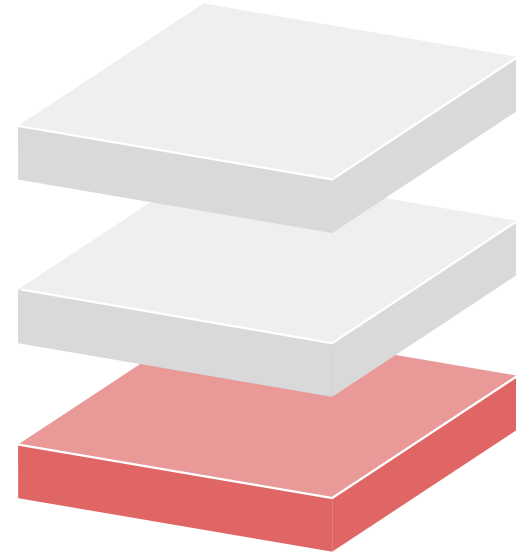
commits



issues



pull_requests



1

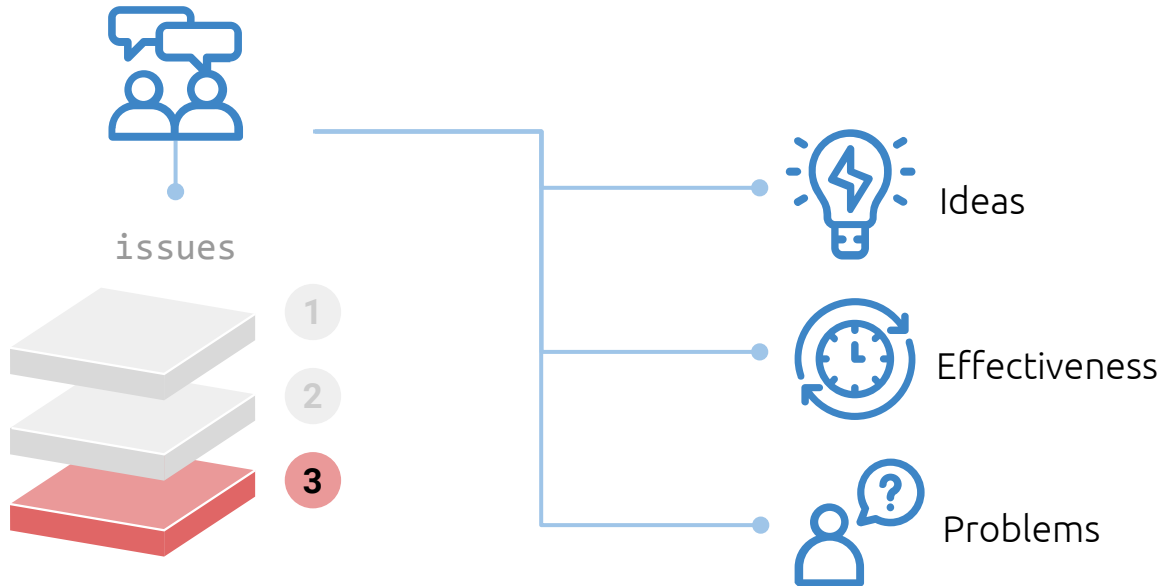
2

3

Umsetzung des Tools

3

Database



Umsetzung des Tools

3

Database

repositories



```
robert@may20: ~/OSCTool/Code
```

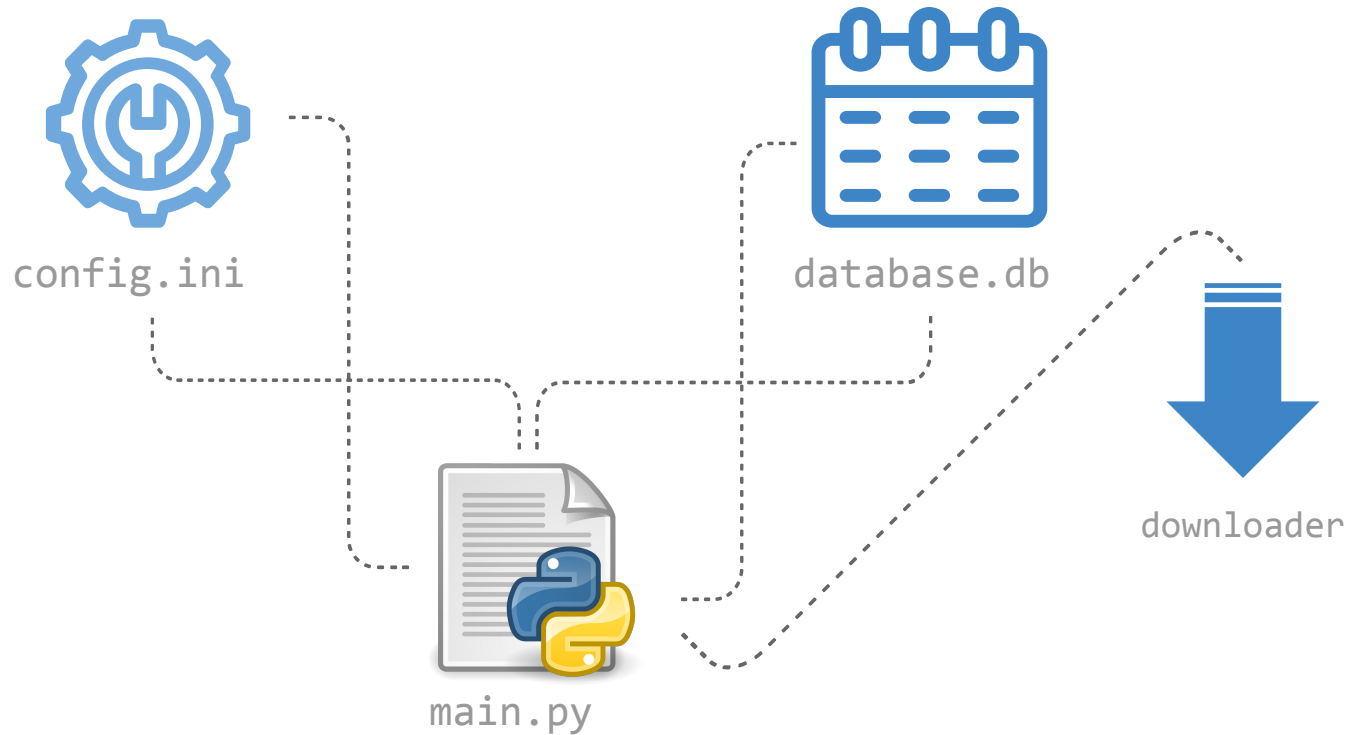
repo_id	repo_creator	repo_name	repo_size	downloaded	last_access	code_frequency_additions	code_frequency_deletions	code_frequency_difference	contributors	commits_amount	commits_keywords	issues_amount	issues_keywords	issues_labels	pull_requests_amount	pull_requests_keywords	pull_requests_labels
528	DocSavage	rails-authorization-plugin	651	0	2020-09-11	16789	14937	0.893949368683747	5	120							
16898	protocool	ack-tmbundle	501	0	2020-09-11	3085	306	0.12511555915721	4	27	old, new, refactor	5					
18257	jamesu	rallscollab	2868	0	2020-09-11	134465	112403	0.835927564793812	11	864	refactor, new, old	7	old, new				
23664	dima	restfolx_framework	8433	0	2020-09-11	155364	124576	0.79874826212467	4	584	refactor, new, old	17	old, refactor	Unassigned, 1	2	new, old, refactor	dependencies
24558	tincharper	role_requirement	549	0	2020-09-11	2752	1610	0.65020809767442	3	51	new	4					
26479	pele	oauth	397	0	2020-09-11	12588	6467	0.513743247537337	29	366	old, new, refactor	15	new		5	new, old	dependencies
158501	eznobus	chef-deploy	130	0	2020-09-11	1724	649	0.370480110809281	4	52	refactor, old	0					
201169	acastipera	statistics	186	0	2020-09-11	762	188	0.25330922237197	3	30	new	4	new				
318338	flyerhzn	css_sprite	164	0	2020-09-11	2991	1937	0.647609495152123	7	201	new, refactor	4	new, old		3		
331903	voloko	twitter-stream	369	0	2020-09-11	1667	739	0.443311337732454	17	105	new	19	new, old		6	new	
326807	creattionx	do	97	0	2020-09-11	736	194	0.263506956521739	2	19		0					
814266	nryan43	maglc_encoding	150	0	2020-09-11	240	80	0.333333333333333	5	13		3	old, new		1		
977871	geokit	geokit-rails3	449	0	2020-09-11	17870	15936	0.89177392275602	15	280	old, new	15	new, old		4	new	
1025642	sergeche	eclipse-rencoding	1552	0	2020-09-11	84393	65459	0.775644899430486	1	86	refactor, old, new	4	new, old		0		
1034856	nyrodev	nyroModul	419	0	2020-09-11	7129	4676	0.655912470192173	5	91	refactor, new, old	16	new, old	tips, Todo	1	new	
1212224	andrewdavey	vogue	263	0	2020-09-11	1579	829	0.520515832805573	8	61	refactor, new	17	old, new, refac		1		
1433223	techtierous	tabulous	738	0	2020-09-11	278030	219935	0.791047728662375	5	212	old, new, refactor	5	old, new		1		
1575368	anthonyshort	stitch-css	695	0	2020-09-11	7152	5627	0.78677293064877	1	114	new, old	4	new, old		0		
2060544	JSFoundation	standards	55	0	2020-09-11	1324	78	0.0589123867069486	6	19		23	new, refactor, TC39, W3C, In	1			
2057431	Graphite-Tat	tattle	2288	0	2020-09-11	129400	25361	0.195060887735550	15	228	new, old, refactor	4	new, old, feature, proc	1			
2856799	njonuschat	bootstrap-sass-rails	1999	0	2020-09-11	42432	41354	0.974594645559528	14	315	old, new	1	old, new, refac	feature	0		

SELECT repo_id FROM repositories WHERE contributors >= 5 AND commits_keywords LIKE '%refactor%';

Umsetzung des Tools

3

Database



Sources

[1] Kruchten, P., Nord, R. L., Ozkaya, I., & Falessi, D. (2013). [Technical debt: towards a crisper definition report on the 4th international workshop on managing technical debt](#). *ACM SIGSOFT Software Engineering Notes*, 38(5), 51-54.

[2] Zazworka, N., Izurieta, C., Wong, S., Cai, Y., Seaman, C., & Shull, F. (2014). [Comparing four approaches for technical debt identification](#). *Software Quality Journal*, 22(3), 403-426.

[3] Christopher Oezbek, Lutz Prechelt. [On Understanding How to Introduce an Innovation to an Open Source Project](#). Proceedings of the 29th International Conference on Software Engineering Workshops (ICSEW '07), IEEE Computer Society, Minneapolis, MN, May 20-26, 2007. Presented at 1st International Workshop on Emerging Trends in FLOSS Research and Development, Minneapolis, May 20-26, 2007.

freepik - <https://www.freepik.com/>

slidesgo - <https://slidesgo.com/>

Image sources

[1] https://www.fu-berlin.de/sites/corporate-design/grundlagen/_medien/fu_logo.png?html=1&locale=de&pageref=91083503&ref=91084456

[2] python icon done in the Tango! style. Part of the CodeTangoSet. From Vinicius Depizzol <https://commons.wikimedia.org/wiki/File:Text-x-python.svg> license: free software



Free

Free