

GitMA: CATMA-Daten via Git abrufen und mittels Python-Bibliothek weiterverarbeiten

Unsere Struktur

CATMA nutzt eine GitLab-Instanz als Backend. Alle Projektdaten werden in mehreren Git-Repositories gespeichert. Es gibt eine sogenannte "root-repository", welche alle Daten in folgender Ordnerstruktur bereitstellt:



Weitere Informationen finden Sie hier: catma.de/documentation/git-access/

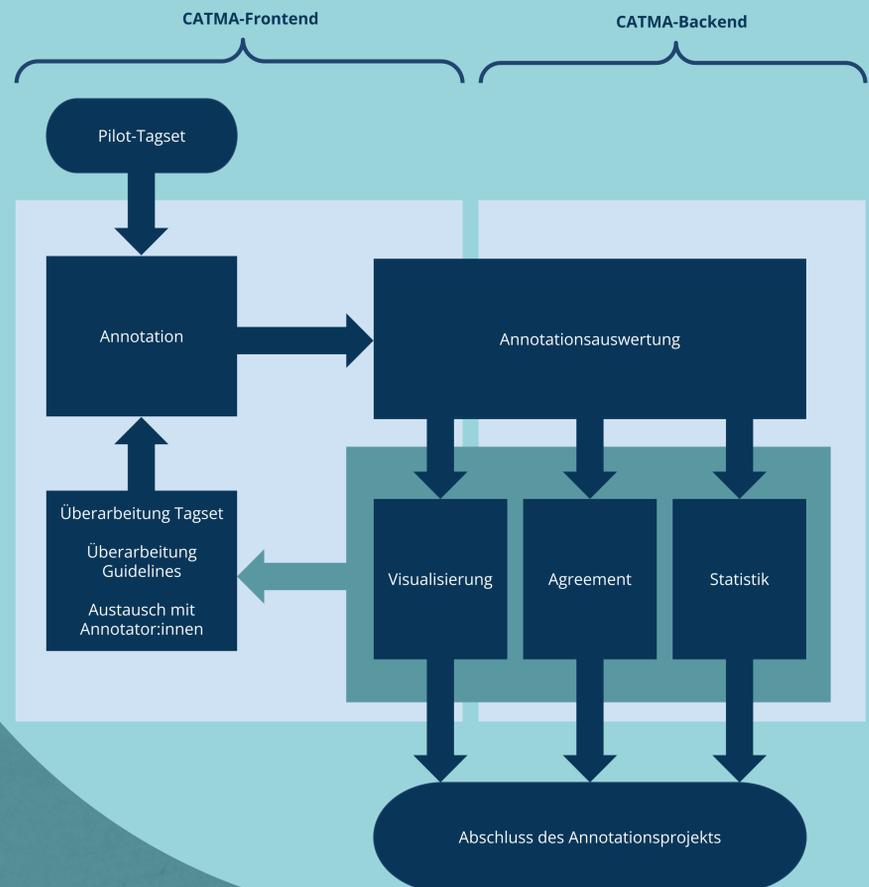
'git clone' eines CATMA Projekts

Voraussetzungen:

- Sie haben bereits einen CATMA-Account und Zugriff auf ein Projekt.
- Sie haben die Logindaten für Ihren Account, bzw. ein Access-Token (falls Sie sich per Google in CATMA einloggen), zur Hand. Ein Access-Token können Sie sich direkt über CATMA oder über das GitLab-Backend holen.

```
git clone --recurse-submodules <catma root repository url>
```

Unsere Idee für Annotationsprojekt-Management



Unser Toolkit
github.com/forTEXT/gitma

Installation

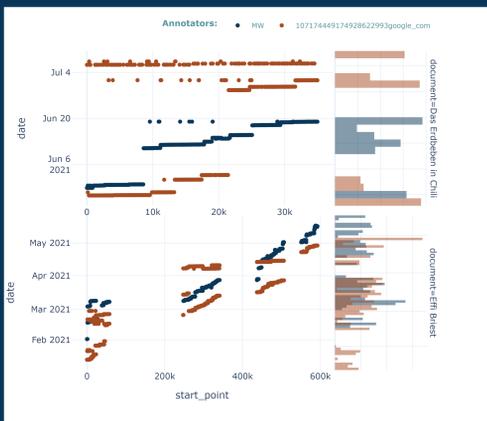
```
pip install git+https://github.com/forTEXT/gitma
```

oder via **docker**

catma.de/documentation/git-access/gitma-docker-image/

Annotationsfortschritt

```
my_project = CatmaProject(
    project_directory='../test/demo_project/',
    project_name='EVENT'
)
my_project.annotation_progression()
```



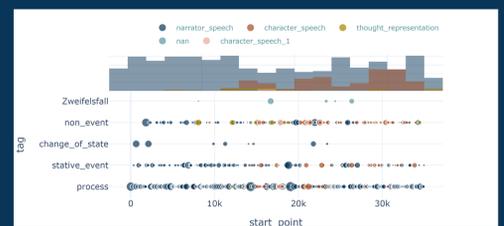
Interaktive Projektübersicht

```
my_project.plot_interactive(color_col='tag')
```



Annotation Collections Visualisieren

```
my_project.plot_interactive(color_col='tag')
```



Annotation Collection

pandas DataFrame

```
my_project.ac_dict['Verwandlung_MV'].df
```

tag	annotation	start_point	end_point	prop:representationtype	
0	change_of_state	Als Gregor Samsa ein[...]	34	95	['narrator_speech']
1	change_of_state	fand er sich in sei[...]	96	167	['narrator_speech']
2	process	Er lag auf seinem po[...]	189	212	['narrator_speech']
3	process	und sah seinen gew[...]	213	328	['narrator_speech']
4	process	wenn er den Kopf ei[...]	222	252	['narrator_speech']
5	stative_event	auf dessen Höhe sich[...]	330	428	['narrator_speech']
6	process	Seine vielen, im Ve[...]	439	543	['narrator_speech']
7	non_event	was ist mit mir gesc[...]	546	572	['thought_representation']
8	process	dachte erf[...]	574	583	['narrator_speech']
9	non_event	Es war kein Traum[...]	585	662	['narrator_speech']
10	stative_event	sein Zimmer, ein ri[...]	684	717	['narrator_speech']
11	stative_event	über dem Tisch hing[...]	719	854	['narrator_speech']
12	stative_event	auf dem eine ausein[...]	735	815	['narrator_speech']
25	non_event	und alle Narrheiten[...]	1393	1420	['thought_representation']
26	process	dachte er[...]	1423	1452	['narrator_speech']
27	stative_event	aber das war gänzli[...]	1434	1470	['narrator_speech']
28	stative_event	denn er war gewöhnt[...]	1472	1526	['narrator_speech']
29	non_event	konnte sich aber in[...]	1528	1664	['narrator_speech']

Inter Annotator Agreement

```
project.gamma_agreement(
    annotation_collections=[
        'Effi_Briest_GS',
        'Effi_Briest_MW',
    ])
```

Gold Annotation

```
project.create_gold_annotations(
    ac_1_name='Effi_Briest_MW',
    ac_2_name='Effi_Briest_GS',
    gold_ac_name='gold_annotation_krambambuli',
    excluded_tags=['Zweifelsfall'],
    min_overlap=0.95,
    push_to_gitlab=True,
)
```