

CoGDat settings and upload

All CoGDat-related scripts/files are deposited under `/home/sysgen/COGDAT_TEST`

- put all data to be uploaded (fastq read1 and 2 and assembly.fa) to `/data` without sub-folders
- required for upload: data in `/data` folder, `cogdat_config.yaml` with correct token and correct dev or data active, `metadata.json` with metadata

Anleitung zu Cogdat gibt es unter <https://docs.cogdat.de/metadata.html>

Die Angaben, was optional ist und was nicht, sind nicht richtig; vermutlich ist alles optional

Set-up config file, json file and data

`cogdat_config.yaml`

- einsehbar mit `less cogdat_config.yaml`
wieder raus gehen mit `Q`
- modifizierbar mit `nano cogdat_config.yaml` (entspricht Notepad+)
wieder raus gehen mit `strg x`; nach Änderungen dann mit `y` bestätigen zum Speichern
- Im config file stehen 2 Möglichkeiten drin: DEV für Testuploads und PROD für echte Uploads. Je nachdem was man machen möchte, muss man das andere inaktivieren mit einem `#` vor der Zeile
Das config file besteht damit aus der URL, wohin die Daten hochgeladen werden und einem token, der User-spezifisch ist (damit wird erkannt, wer die Daten hochgeladen hat)
- Aktuell gespeichert sind die tokens von Michaels Account

Neuen Account mit den entsprechenden tokens erstellen:

- Registrieren bei <https://dev.cogdat.de/> und <https://data.cogdat.de/>
Der Link kommt nach einiger Zeit per Mail; Organisation: Medical Microbiology, University of Tübingen (oder auch was anderes, nicht so wichtig)
- Im Browser von der Workstation/Linux aus anmelden bei <https://dev.cogdat.de/> bzw. <https://data.cogdat.de/>
- Unter "Accounts" - "Manage API Keys" einen Namen vergeben unter "Label", z.B. "Medical Microbiology Dev Key", ein Gültigkeitsdatum eingeben und auf "Generate API Key" klicken.
Es erscheint eine Buchstaben/Zahlenkombination. Das ist der token.
- token mit Rechtsklick - Copy kopieren und nach `nano cogdat_config.yaml` per Pfeiltasten an die richtige Stelle navigieren, den alten token löschen und mit Rechtsklick - Paste den neuen einfügen, dann mit `strg x` und dann `y` und `Enter` raus gehen

`metadata.json`

- Send Caspar the metadata, he will produce the `metadata.json` file and send it back

`data`

- Wenn die fastas nur als multifasta vorliegen, kann man das multifasta trennen mit einem Skript von Ersoy:
`python /home/sysgen/multifasta_fasta/multifasta_to_single.py
<path_to_multi_fasta_file>`
- Copy fasta and fastq files (R1 and R2) to /home/sysgen/COGDAT_TEST/data

Hochladen

```
dmclient --config cogdat_config.yaml shortcuts stage-and-submit  
--label "Microbiology API Test" "test_metadata.json" "data"
```

Vermutlich ist "Microbiology API Test" frei gewählt und kann geändert werden

Danach kann man auf der Seite nachschauen, ob der Upload geklappt hat.

Falsch hochgeladene Daten können nicht gelöscht werden und eine mal vergeben
Proben-ID kann nicht neu vergeben/überschrieben werden.