

The APIS Project

matthias.schloegl@oeaw.ac.at

- Based on the "Austrian Biographical Dictionary" (ÖBL)
- ~18.500 biographies of people who "had an impact" on - back then - Austrian soil and died between 1815 and 1955
- The APIS project intends to semantically enrich these biographies (relations to people, places etc.)
- ..and use that data to research migration patterns.

“Stud. in Wien und Linz, legte 1795 in St. Florian die Ordensgelübde ab, 1797 Priesterweihe, arbeitete zuerst in der Seelsorge, dann als Prof. für Bibelstudium des Neuen und Alten Testamentes am Lyzeum in Linz, Dir. des Gymnasiums in Linz und Mitarbeiter der „Linzer Monatsschrift“. 1814–23 Kanzleidir. des Stiftes, 1818 Leiter des Stiftsdekanates, 1823 Propst, 1823–48 Generaldir. der o.ö. Gymnasien...”

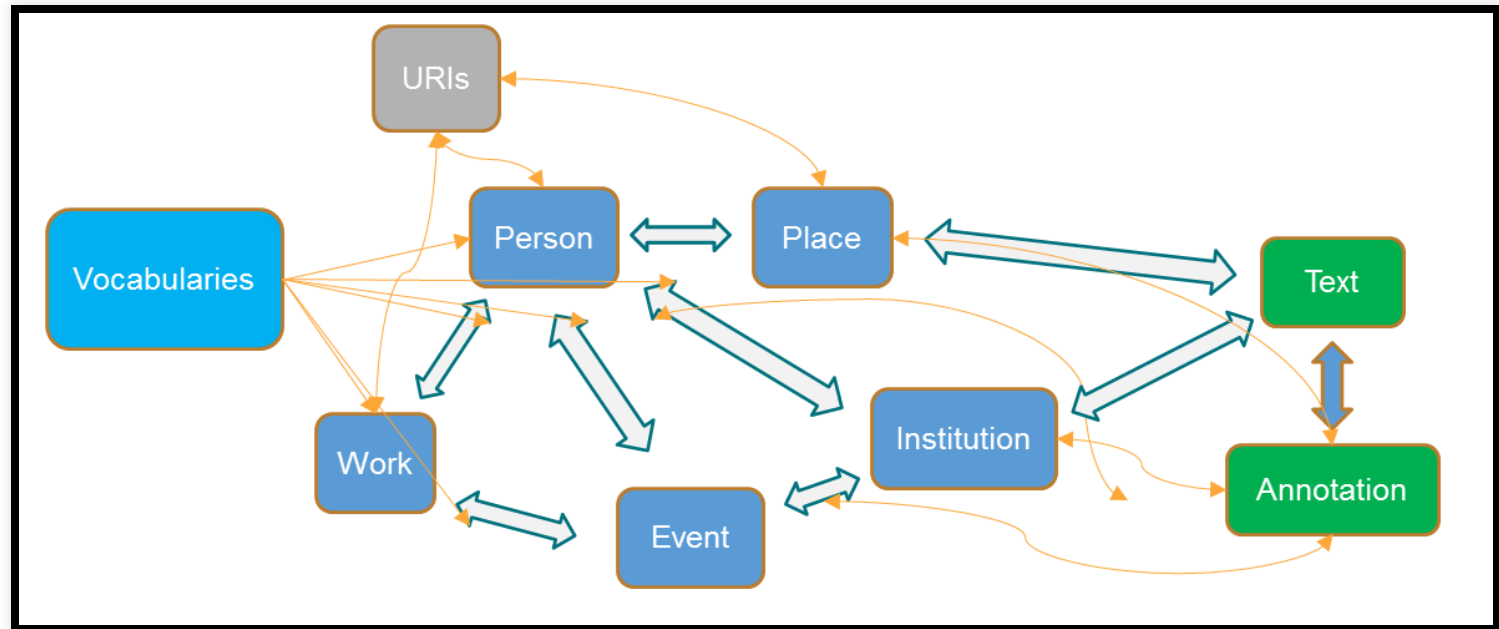
*“Stud. in **Wien** und **Linz**, legte 1795 in St. Florian die Ordensgelübde ab, 1797 Priesterweihe, arbeitete zuerst in der Seelsorge, dann als Prof. für Bibelstudium des Neuen und Alten Testamentes am Lyzeum in Linz, Dir. des Gymnasiums in Linz und Mitarbeiter der „Linzer Monatsschrift“. 1814–23 Kanzleidir. des Stiftes, 1818 Leiter des Stiftsdekanates, 1823 Propst, 1823–48 Generaldir. der o.ö. Gymnasien...”*

ideas

- simple data model >> serialize later in various formats/data models (karma, ontop)
- rock solid and widely used software stack >> less bugs, better documented, easier to develop
- hybrid approach >> researchers and tools use the same webtool/database

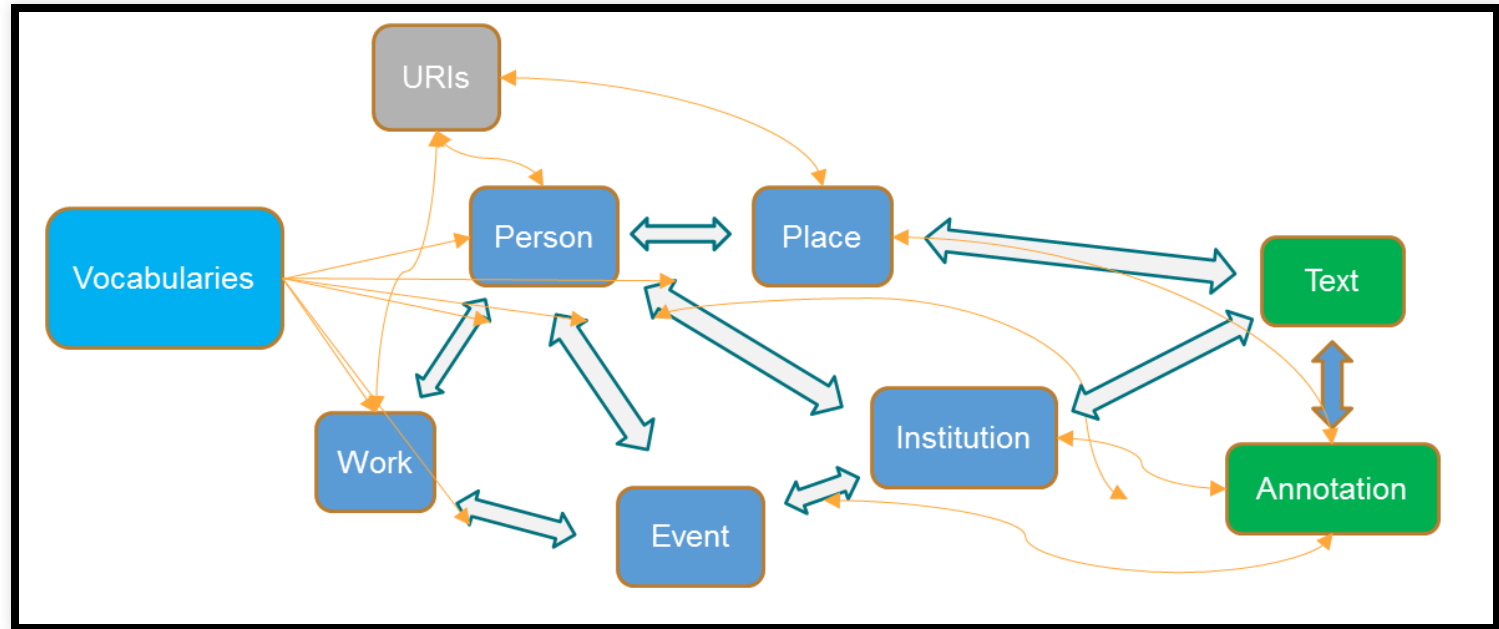
datamodel #1

5 core entities: people, places, institutions, events and works



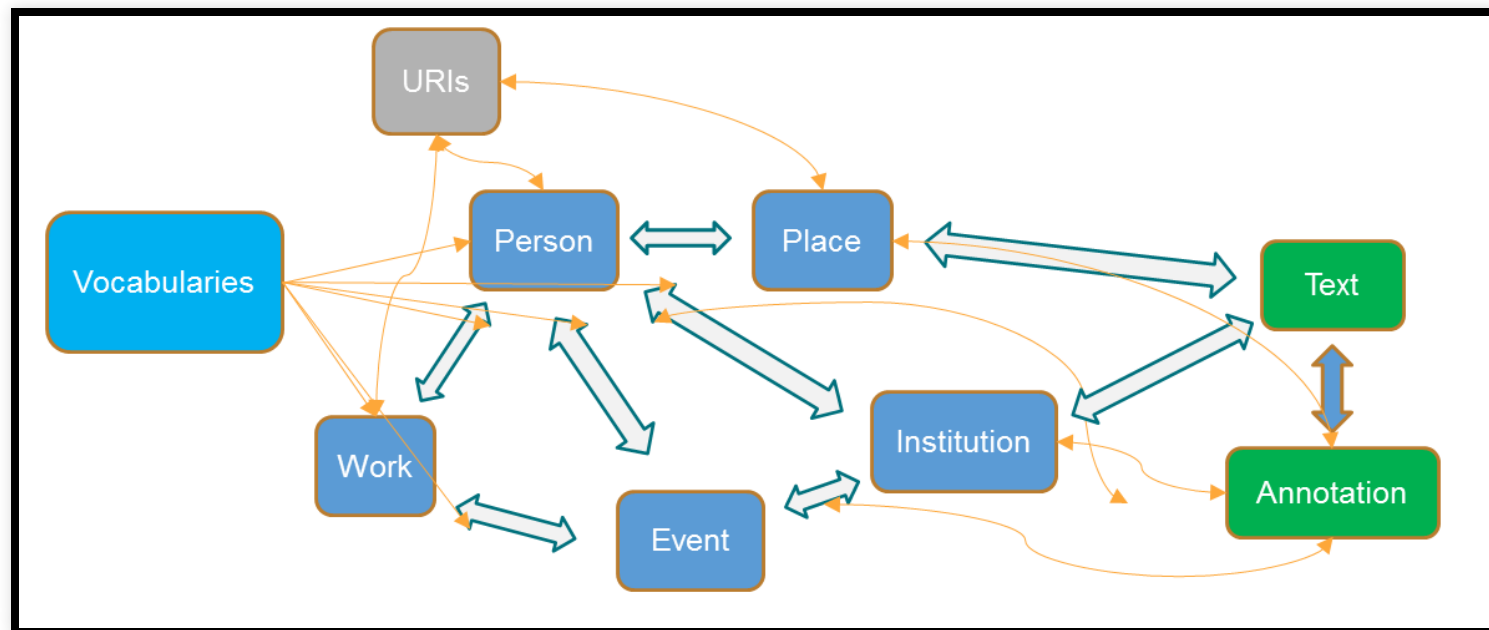
datamodel #2

all interrelated >>
relations contain a type
(SKOS), a time span and
some free text fields



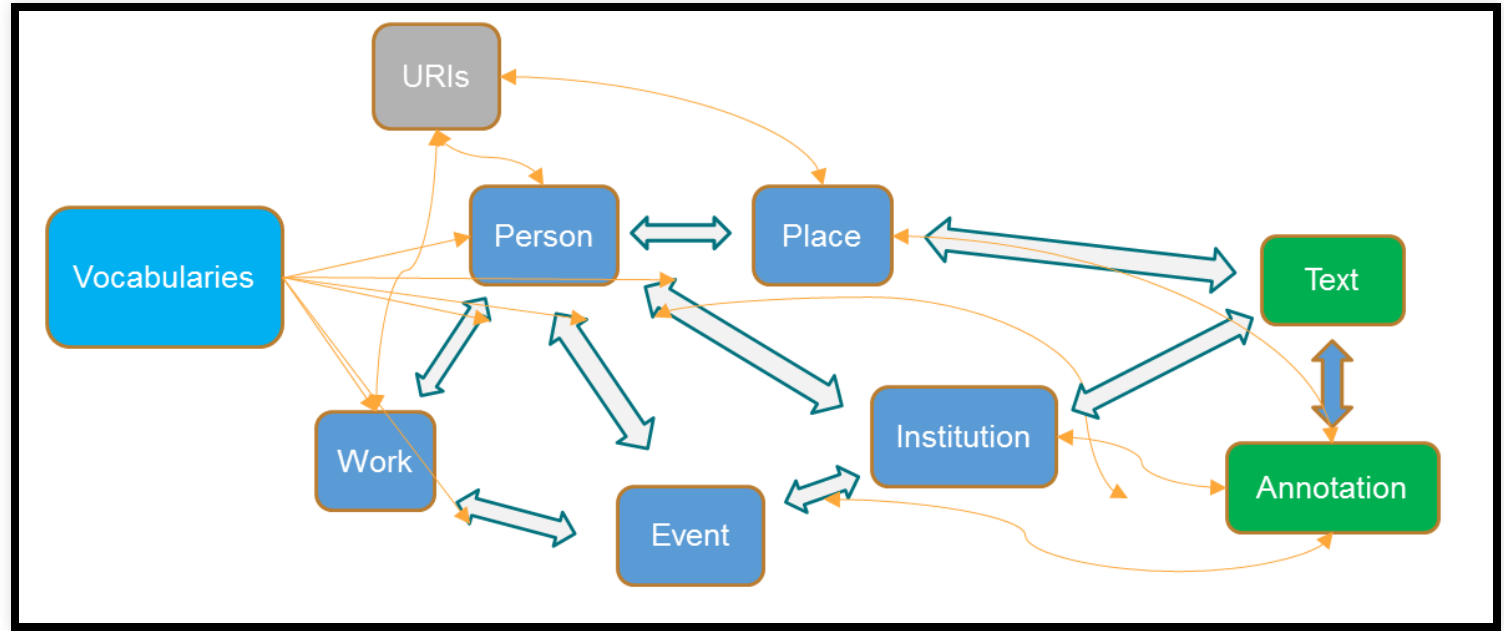
datamodel #3

every entity type can have
several text types
attached



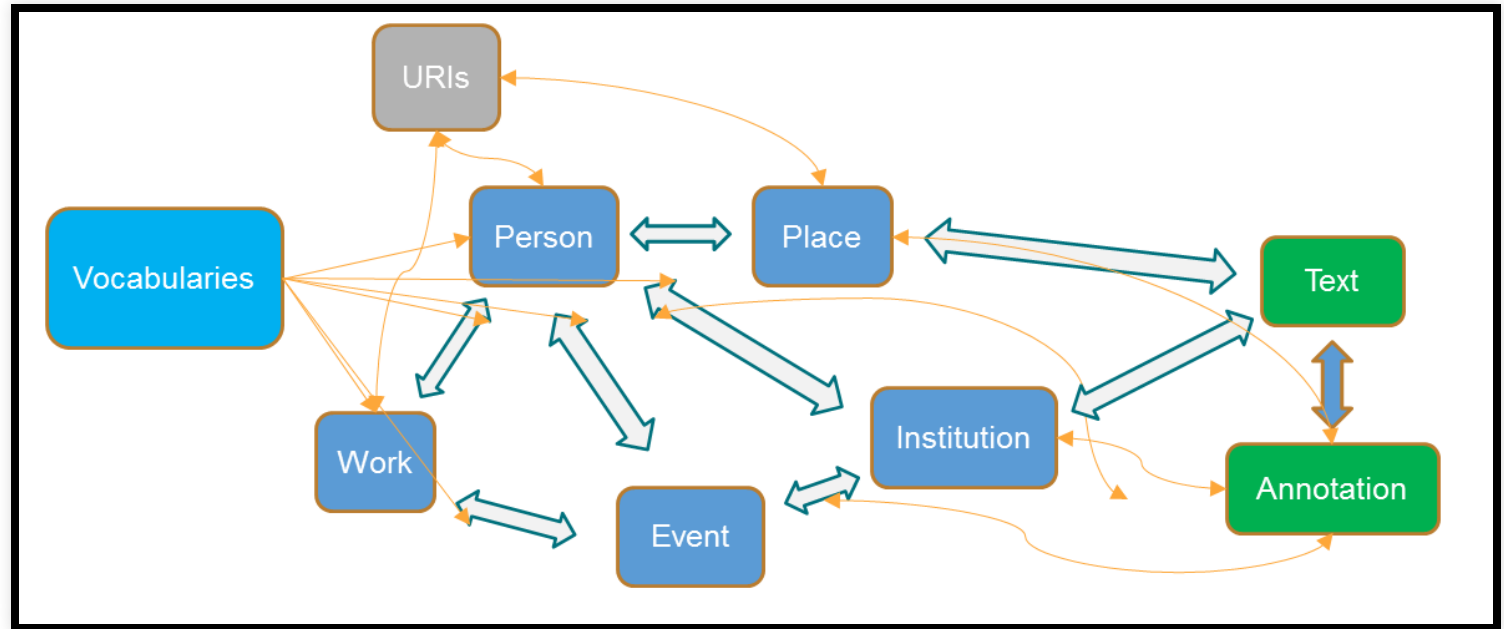
datamodel #4

every annotation is linked to an annotation project, a user and a entity or relation >> allows for multiple annotations (e.g. different tools) of one and the same text



datamodel #5

every change of an object is serialized to json and stored with user and timestamp (revisioning)



software

- Django (Python Web Development Framework)
- MySQL (Database)
- Stanbol (Reference resources)
- spacy and Gate for NLP

webapp #1

- customizable filters and lists
- download of filtered lists in various formats
- creating new entities from customizable reference resources

webapp #2

- metadata of entity on the left
- link to other entities on the right
- auto-import of entities from reference resources

webapp #3

- annotation of text
- annotation projects
- filter annotations to show

webapp #4

- iteratively create network
- add whole collections
- export to graphml/json

Thanks!

matthias.schloegl@oeaw.ac.at / <https://github.com/acdh-oeaw/apis-core> / <https://apisdev.acdh.oeaw.ac.at>