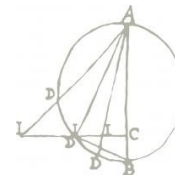# Tool intro: Jupyter Notebooks

Robert Casties
MPI for History of Science, Germany

MAX PLANCK INSTITUTE
FOR THE HISTORY OF SCIENCE

Jupyter(lab)

- https://jupyter.org/
- Web frontend for interactive programming notebooks
- Support for many programming languages, like Python or R
- Can be run locally or on a server

```
    – pip install jupyterlab
    – jupyter-lab
```

Jupyter Notebooks

- Notebooks run a kernel for a specific language, e.g. Python
- Notebooks are made of cells that contain either text (for documentation), lines of the programming language or the output of a programming cell
- You can run notebook cells one-by-one or run the whole notebook as one
- You can interactively change a cell and re-run it when developing your program
- Github can show an uploaded notebook with its output

File   Edit   View   Run   Kernel   Tabs   Settings   Help

Launcher   01_recogito-json-annotatic

Code   Python 3 (ipykernel)

```
[22]:   from rdflib import Graph, Dataset, URIRef
        import re
```

```
[23]:   datapath = "/home/jupyter-mvogl/mydata/"
```

```
[24]:   # the URI for our named graph that will contain our data
        anno_named_graph_uri = URIRef(f'http://example.org/ns/users/malte/genesis-example-ge
```

```
[25]:   # read the JSON-LD file with annotations from Recogito into a Graph
        anno_graph = Graph()
        anno_graph.parse(f'{datapath}gtclbuppjmot3c(1).jsonld',
                         format='json-ld',
                         publicId=anno_named_graph_uri)
```

```
[25]:   <Graph identifier=Na7049bacda2e47a0b3285ff1838b5081 (<class 'rdflib.graph.Graph'>)>
```

```
[26]:   set([x for x in anno_graph.predicates()])
```

```
[26]:   {rdflib.term.URIRef('http://purl.org/dc/terms/creator'),
         rdflib.term.URIRef('http://purl.org/dc/terms/issued'),
         rdflib.term.URIRef('http://purl.org/dc/terms/modified'),
         rdflib.term.URIRef('http://www.w3.org/1999/02/22-rdf-syntax-ns#first'),
         rdflib.term.URIRef('http://www.w3.org/1999/02/22-rdf-syntax-ns#rest'),
         rdflib.term.URIRef('http://www.w3.org/1999/02/22-rdf-syntax-ns#type'),
         rdflib.term.URIRef('http://www.w3.org/1999/02/22-rdf-syntax-ns#value'),
         rdflib.term.URIRef('http://www.w3.org/2000/01/rdf-schema#label'),
         rdflib.term.URIRef('http://www.w3.org/ns/activitystreams#generator'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#end'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#exact'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#hasBody'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#hasPurpose'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#hasSelector'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#hasSource'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#hasTarget'),
         rdflib.term.URIRef('http://www.w3.org/ns/oa#start'),
         rdflib.term.URIRef('http://xmlns.com/foaf/0.1/homepage'),
```

Filter files by name

/ workshop / ingest /

Name	Last Modified
01_recogito-json-annotations-to-blazegraph.ipynb	16 days ago
02_process-person-annotations.ipynb	16 days ago
03_recogito-csv-relations-to-blazegraph.ipynb	16 days ago
recogito-genesis-annotations-nb.nt	16 days ago
recogito-genesis-annotations.nt	16 days ago

Simple   2   1   Python 3 (ipykernel) | Idle   Mem: 183.84 MB   Mode: Command   Ln 1, Col 1   01_recogito-json-annotations-to-blazegraph.ipynb

Jupyter Notebooks

- Can be used for "literate programming" (Don Knuth)
- Can be used for "reproducible science"
- Can be converted into "real" Python programs (may need additional work)
- Can be developed into books using https://jupyterbook.org