KoLibri - Leveraging Zurich Zentralbibliothek's Jupyter Notebooks for Metadata Retrieval and Analysis from Alma

Making library data available for researchers.

Bielefeld | 3-7 March 2025



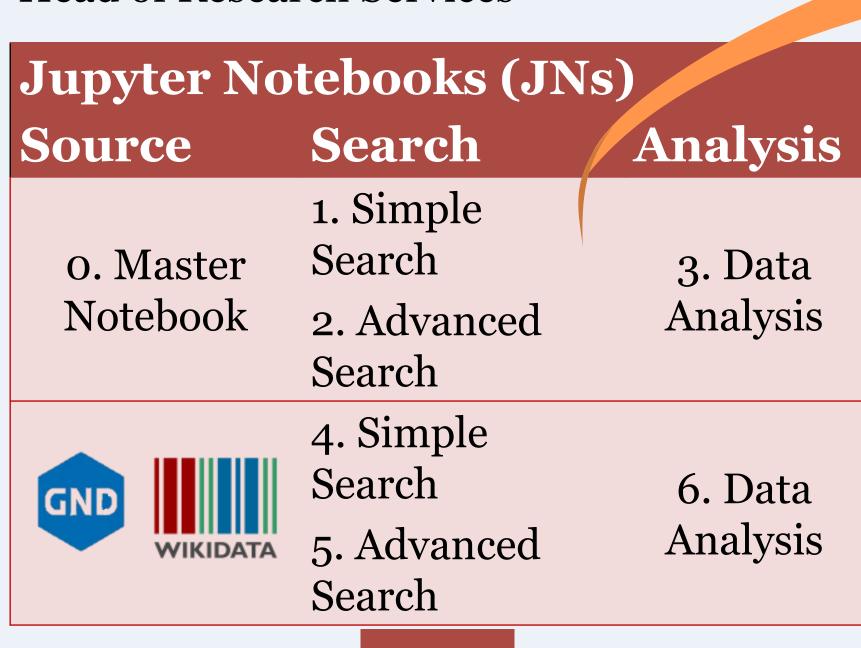




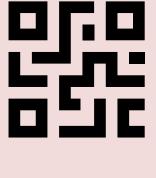
New data-driven tools for data mining: ZB's Jupyter Notebooks

"Following our benchmarking efforts, we are pleased to introduce our own series of Jupyter Notebooks as a new research service, featuring innovative, user-centric functionalities."

- Elisabeth Gamer (ZB), Head of Research Services



o. Master Notebook



- tweak the codebase of ca. 5,000 lines of code
- use advanced Python programming skills



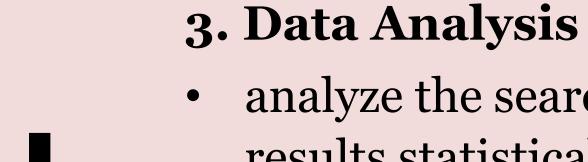
1. Simple Search

- harvest entries of media records
- download their table of contents in bulk

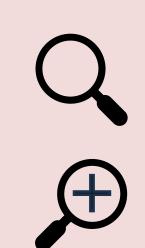


2. Advanced Search

query specific fields in the catalog for more targeted results



- analyze the search results statistically
- create bar and pie charts, word clouds or world maps



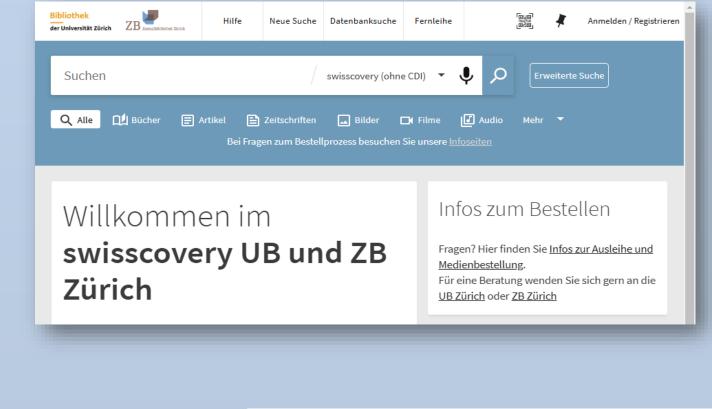
4.-6. Enriched JNs

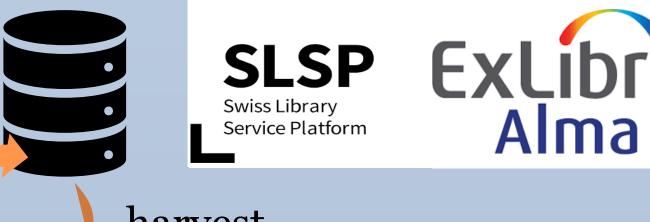
- add out-of-the-box GND/Wikidata data fields
- customize data fields



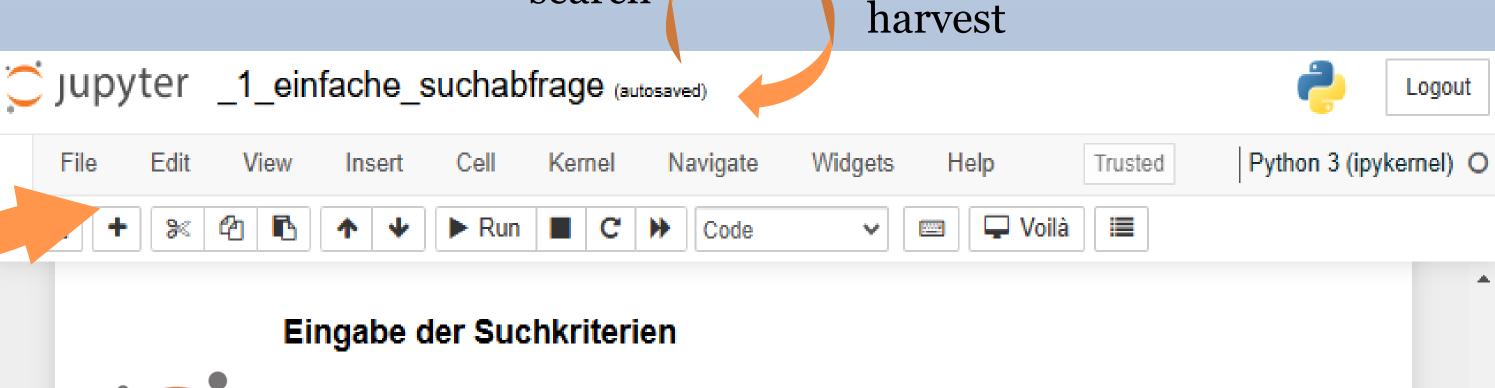
search

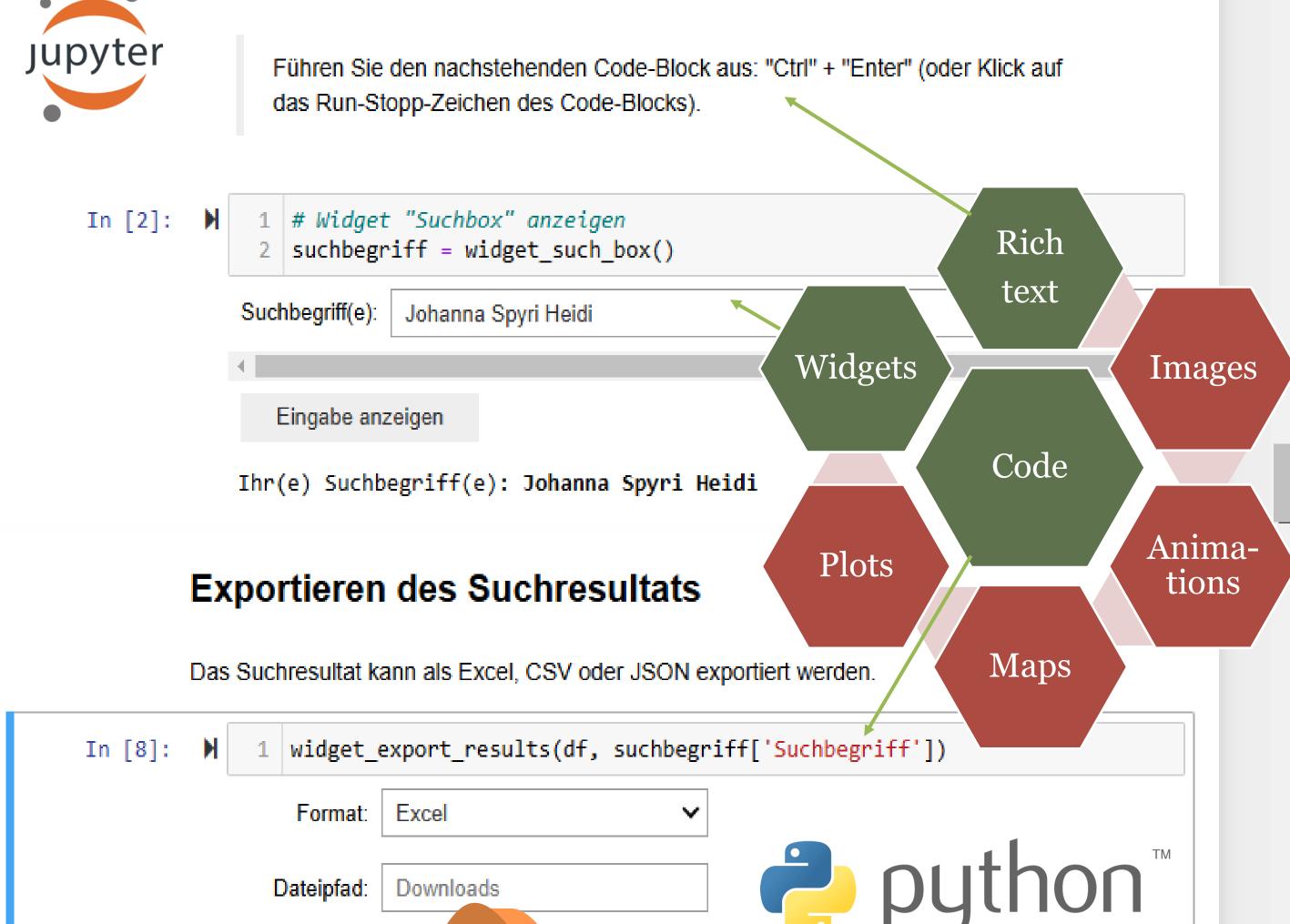
searches the Swisscovery website (> 6 mio. records) for library data. Yet, another SLSP-Alma database with more records (> 25 mio.) in MARC XML format exists. This library catalog can now be searched and harvested using Jupyter Notebooks.





chnis dieses Jupyter Notebooks geöffnet werden.







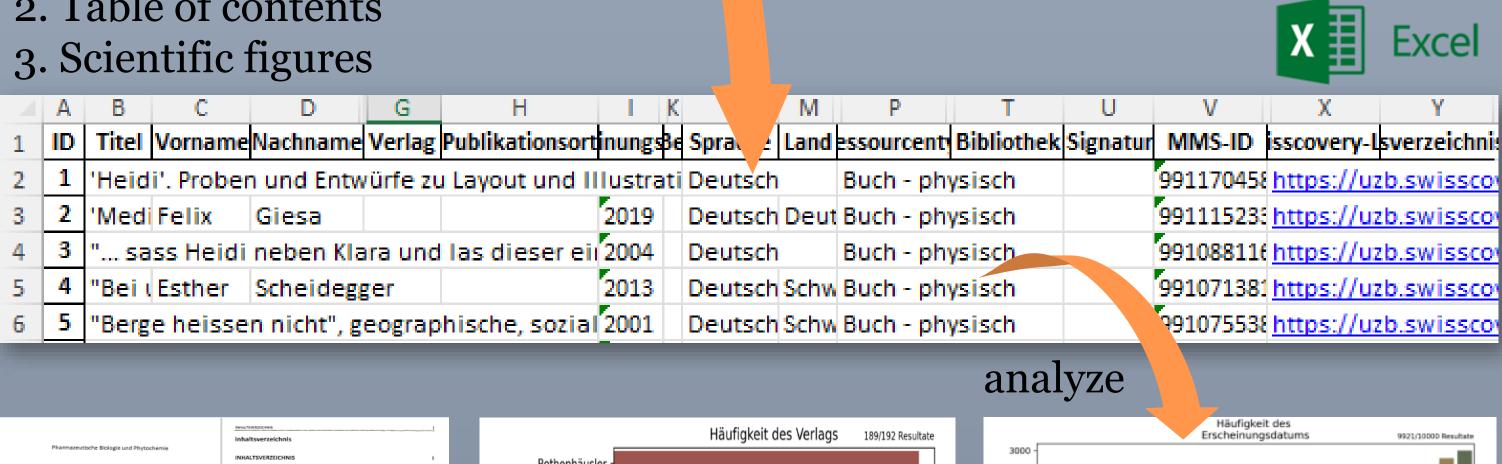
1. Data set (<= 10,000 records)

Exportieren

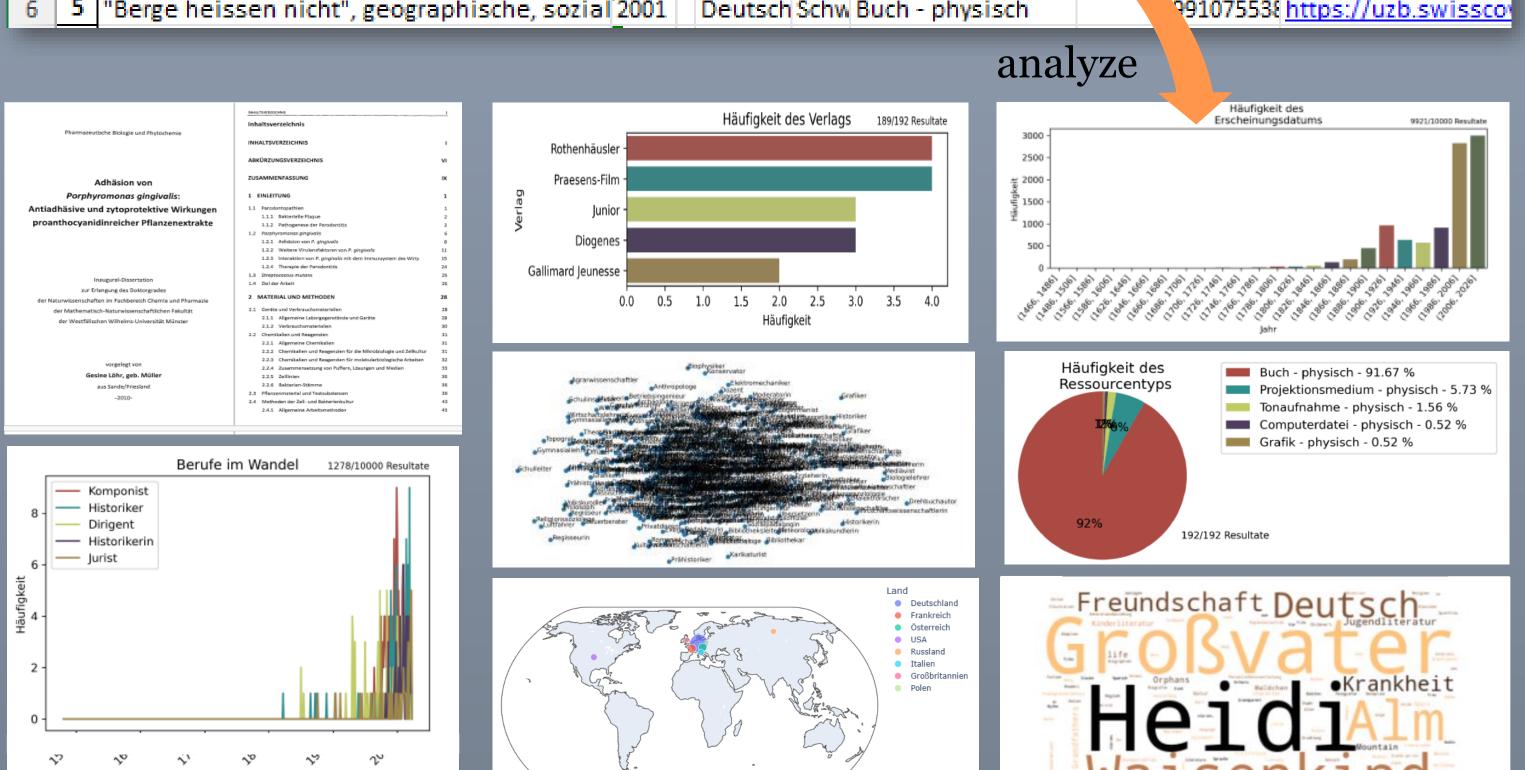
Jetzt kann die Ergebnisdatei im Vel

- 2. Table of contents

PDF



retrieve



KoLibri is feeding on book flower nectar!

References: https://data.zb.uzh.ch



search prompt example: Johanna Spyri's Heidi for data analysis

Johanna Spyri Heidi Suchbegriff(e): -- Freundschaft Deutsch

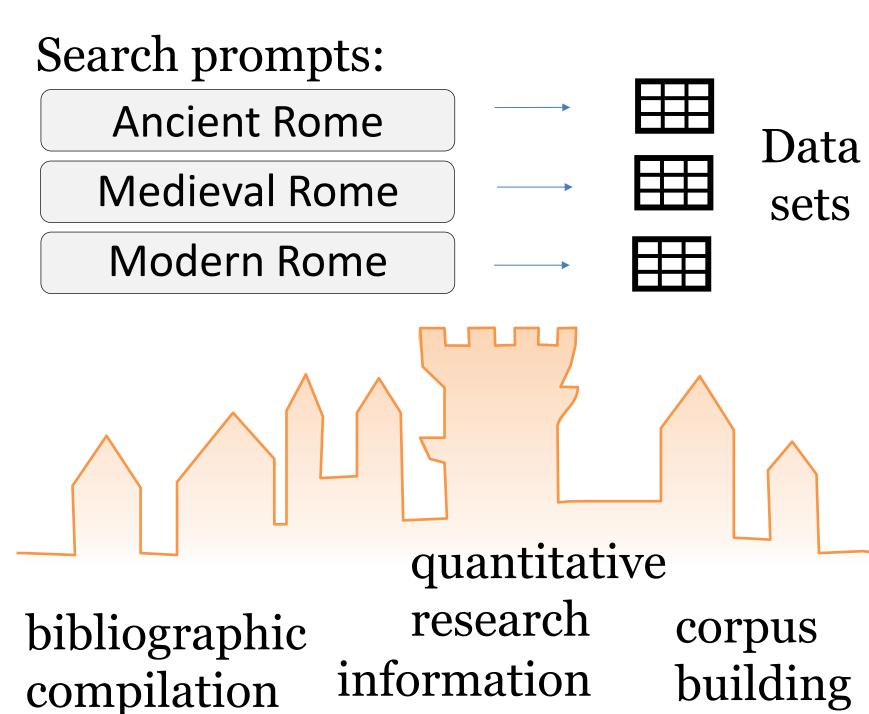
Top-5 related words:

'Grandfather', 'orphan', 'alpine pasture', 'German' and 'friendship'.

- Publications: Mostly, physical books written in German between 2010 and 2020.
- Special resource types: A small fraction (1.68%) of the media items in the library holdings consists of and electronic audio physical recordings and computer files.
- Geographical focus:

Switzerland and the Alps, however, there are also references to Japan and the USA.

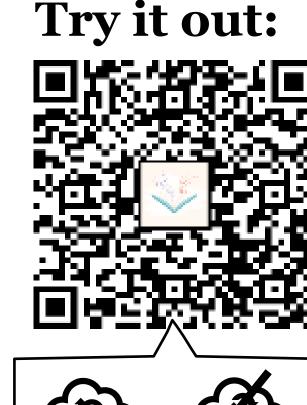
(4) Pooling various data sets for building castles in the sky

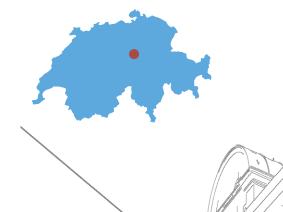


visualization

Key benefits

- ✓ No login
- ✓ Open-source
- ✓ Transparency
- ✓ Usability vs. logic





Contact Us:

Bibliotheksinformatik Zurich Zentralbibliothek



bibliotheksinformatik @zb.uzh.ch