

# Retrieving OJA Data from online platforms

---

**Research Question:** How can we efficiently scrape German online job advertisements (OJAs)?

**Background:** There are several platforms in Germany that provide OJAs. Some of them focus on academic jobs (e.g. academia.de), others more on non-academic jobs (BA Berufesuche). Only BA services provide an API, other platforms require scraping approaches. The first task is to structure the OJA data (e.g. title, occupation), to link them if necessary (to the KldB) and to provide scrapers in order to get a representative view of the OJAs.

**Data:** *German texts!*  
List of chambers

**Literature:**

- 📄 T. Reiser, J. Dörpinghaus, P. Steiner, M. Tiemann. Towards a dataset of digitalized historical German VET and CVET regulations. *Data*, 9(11):128 (2024).
- 📄 T. Reiser, J. Dörpinghaus, P. Steiner. Analyzing Historical Legal Textcorpora: German VET and CVET regulations. *INFORMATIK 2024, Lecture Notes in Informatics (LNI)*, Gesellschaft für Informatik, p. 2001-2012.
- 📄 J. Dörpinghaus, J. Binnewitt, D. Samray, K. Hein. Understanding Informatics in Continuing Vocational Education and Training Data in Germany. *ACM Transactions on Computing Education*, 2024, 24(3), 1-22.
- 📄 J. Dörpinghaus, D. Samray, R. Helmrich. Challenges of Automated Identification of Access to Education and Training in Germany. *Information* 2023, 14(10), 524

**Requirements:** Coding (Python); Databases (SQL); Knowledge of the German Labor Market

---

**Contact:** Dr. Jens Dörpinghaus ([doerpinghaus@uni-koblenz.de](mailto:doerpinghaus@uni-koblenz.de))

---