A distributed system for invoking the BundesAPI with integration to a data warehouse

Research Question:	How can we efficiently build a distributed pipeline system that calls BundesAPI to query data from BA and integrate this data into a data warehouse without quality issues?
Background:	The BundesAPI project (https://github.com/bundesapi) allows you to query several federal agencies, including the Bundesagentur für Arbeit (BA). The goal of this project is to retrieve and store data from several of their systems in order to get a continuous and solid data base for further analysis. This includes: • JOBSUCHE • BERUFENET • BEWERBERBÖRSE • ENTGELDATLAS • NEWPLAN • AUSBILDUNGSSUCHE • STUDIENSUCHE • WEITERBILDUNGSSUCHE One challenge is to obtain interoperable data.
Data:	*German texts!* • BundesAPI Project • Internal Documentation
Literature:	▣ A. Fischer, J. Dörpinghaus. Web Mining of Online Resources for German Labor Market Research and Education: Finding the Ground Truth? *Knowledge* 2024, 4(1), 51–6
Requirements:	Coding (Python); Databases (SQL); Linux; Knowledge of the German Labor Market
Contact:	Dr. Jens Dörpinghaus (doerpinghaus@uni-koblenz.de)