Master thesis

Circuit diagram extraction: Extraction of technically correct logical data models (topology) from circuit diagram images

Initial situation and motivation

Many transmission system operators have circuit diagrams in a wide variety of formats: Scans, PDFs, images, CAD files.

These documents contain valuable information on primary technology and single-line diagrams, but cannot usually be analysed automatically.

At the same time, there are numerous norms and standards (e.g. in the DACH region, EU and internationally) as well as specifications from outside the industry that can be used to create standardised logical data models.

Problem definition

To optimise internal processes, circuit diagram images should be automatically converted into structured, technically correct data models (e.g. topology information). Methods exist, particularly in the context of machine learning, which could support this process (semi-)automatically.

The challenge is to correctly recognise the relevant switching elements, connections and topologies from a wide variety of file formats - often without explicit metadata.

Research question(s)

1. Which technologies and algorithms (e.g. image recognition, OCR, deep learning) are best suited to recognising electrical symbols, lines and topology in circuit diagram images?
2. What standards and methods are there for the structured storage of extracted data?
3. Which procedure is best suited to process a wide variety of file formats (incl. PDF, PNG, CAD) and implement them prototypically via Python?

Desired methodology

- Literature research
- Programming
- Benchmarking and comparison of relevant methods
- Prototype with basic functionality and
- Preparation of results in a structured form