



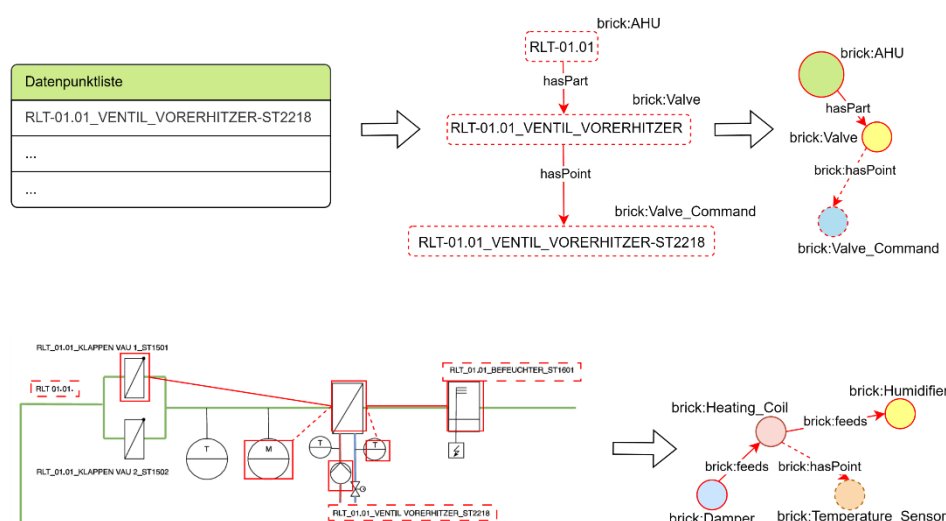
Diplomarbeit Nr. 16/2025

Automatisierte Erstellung von formalen Anlagenbeschreibungen aus Verschiedenen Unstrukturierten Datenquellen

Bearbeitet von: Hüseyin Tokul

Motivation & Problemstellung

Heizungs-, Lüftungs- und Klimasysteme (HLK) gehören zu den größten Energieverbrauchern in Gebäuden. Obwohl große Mengen an Betriebs- und Sensordaten verfügbar sind, werden diese häufig für energetische Analyse nicht genutzt. Grund dafür ist das Fehlen konsistenter, maschinenlesbarer Metadaten. Diese liegen in Bestandsgebäuden meist unstrukturiert und verteilt in verschiedenen Quellen (z. B. Anlagenschemata, Datenpunktbezeichner, Textdokumente) vor. Die manuelle Erstellung strukturierter Modelle in Form von Wissensgraphen ist aufwendig und fehleranfällig. Es existieren bereits automatisierende Ansätze zur **Metadaten-Inferenz**, jedoch liefert keiner dieser Ansätze für sich allein eine vollständige Lösung.



Ziel der Arbeit

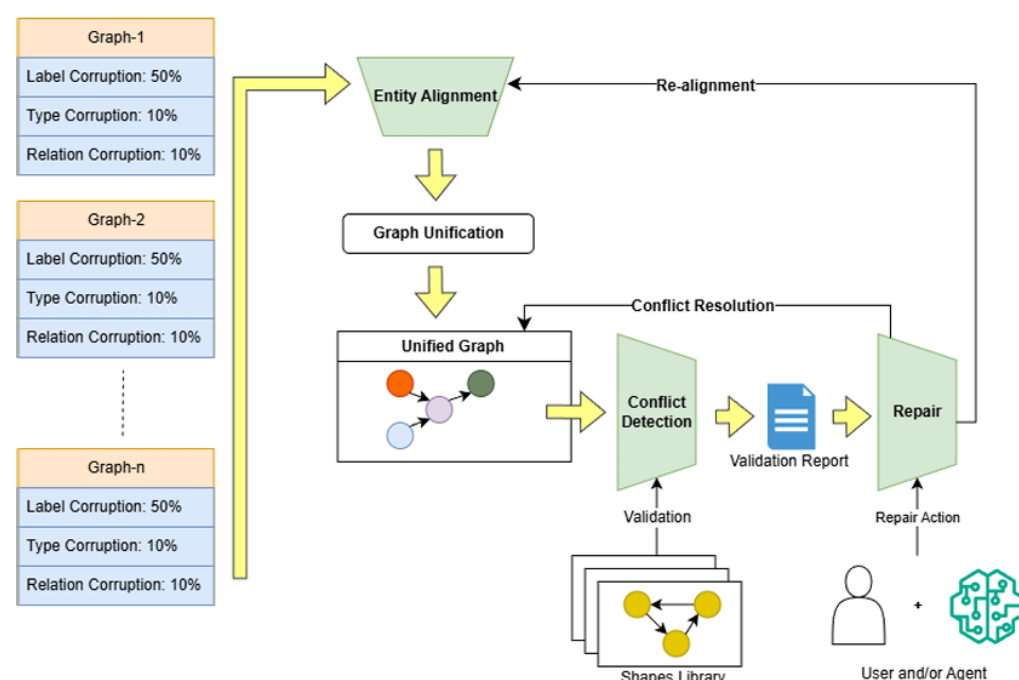
Ziel ist die **automatisierte Zusammenführung mehrerer Wissensgraphen**, die dasselbe HLK-System beschreiben und aus unterschiedlichen automatisierten Inferenzansätzen hervorgehen.

Dabei sollen:

- Informationen aus verschiedenen Quellen kombiniert werden
- Redundanzen genutzt werden
- semantische Konflikte erkannt werden
- Ergebnis: **eine konsistente, integrierte Systembeschreibung**

Methodik

- 1 – Graphenverfälschung
- 2 – Entity Alignment
- 3 – Conflict Detection
- 4 – Repair



Ergebnisse

- Methode auf drei Datensätzen getestet
- Bewertung der Leistung mittels **F1-Score**

Datensatz	Entity Alignment	Conflict Detection-1	Conflict Detection-2
FHM	0.89	0.88	0.91
B1	0.90	0.86	0.89
B2	0.79	0.79	0.88

