

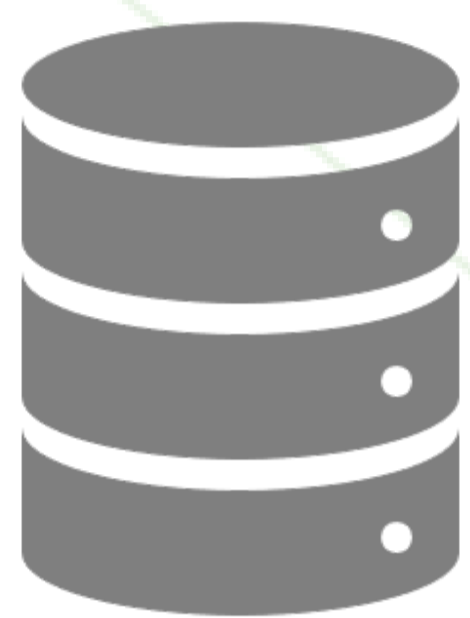
Standardization of HAZOP-studies for application of Artificial Intelligence

Jonas Oeing¹, Prof. Dr.-Ing. Norbert Kockmann¹

1. Laboratory of Equipment Design, TU Dortmund

Motivation

- Safety assessments (e.g. HAZOP) are essential and a very time consuming part in process industry
- A large amount of data is available as a result of detailed safety documentation
- Data offers a lot of unexploited potential for further use

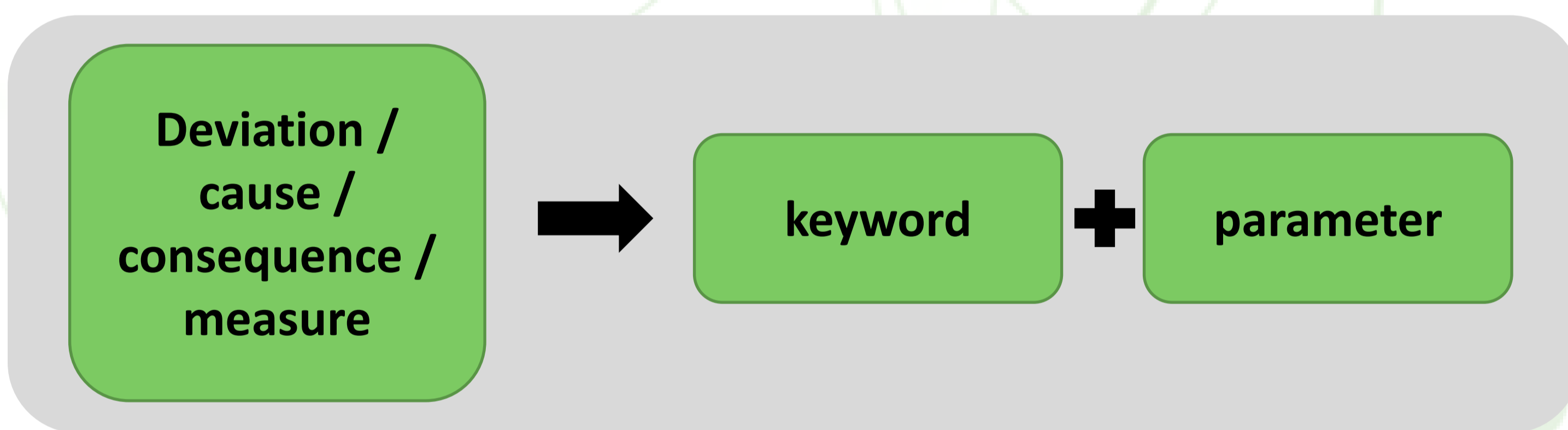


➤ Automated extraction of information from historical HAZOP documents to support future safety assessments

Challenges

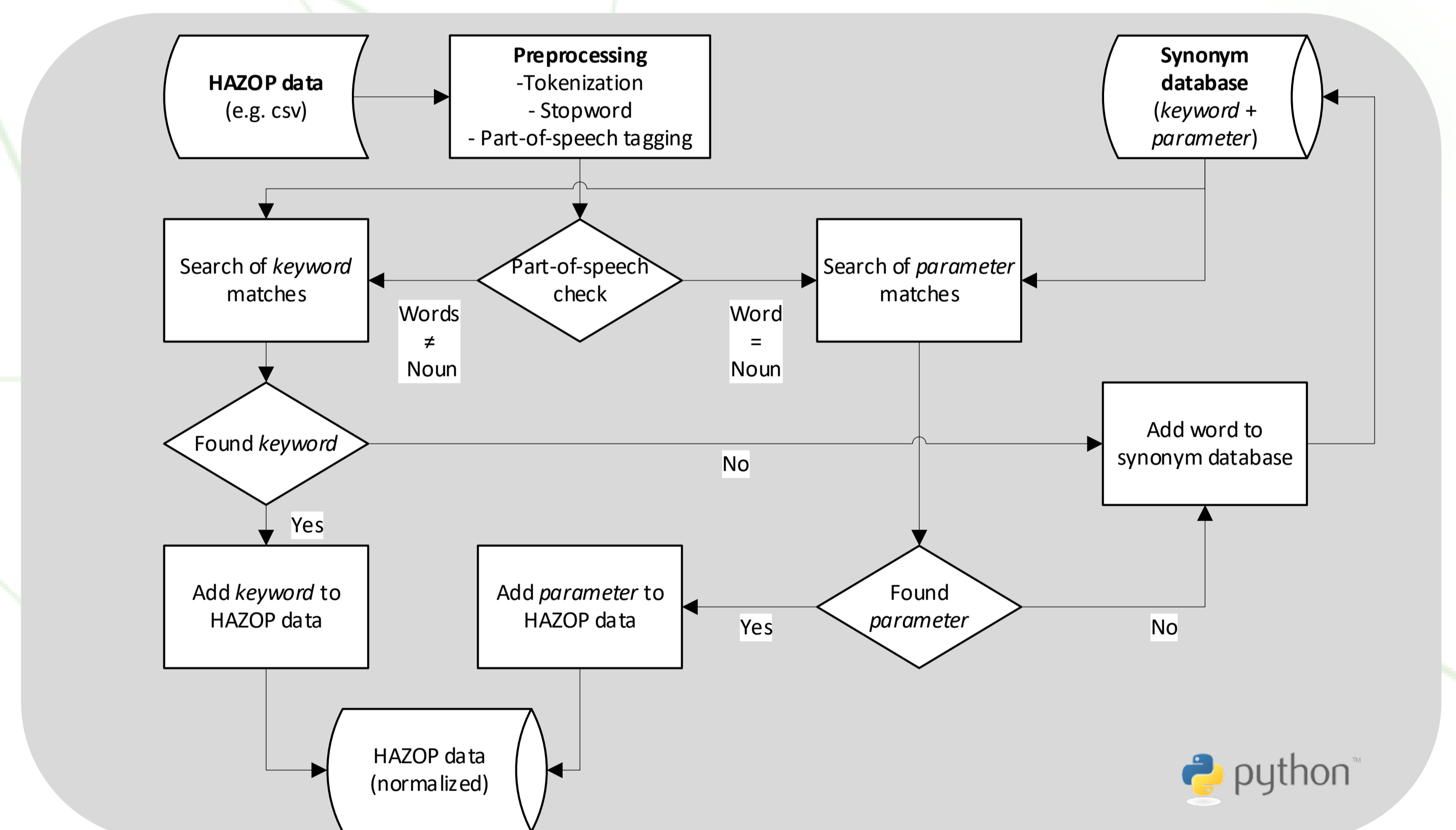
- Formats of safety considerations are often not machine-readable
- Fokus therefore concentrates on HAZOP analyses, which are based on .csv files
- High variation of entries
 - Individual formulations and language of the persons carrying out the HAZOP
 - Table structures and formatting show deviations
- Using natural language processing NLP, extensive detection and normalization of entries must be performed to make the data available for machine learning / artificial intelligence

Normalization of entries



- Entries / descriptions are transferred into **keyword / parameter / location** combinations
 - Detection of relevant information by part-of-speech tagging
 - Standardization by synonym databases
 - Implementation in Python via the packages nltk^[1] and spacy^[2]

Normalization workflow

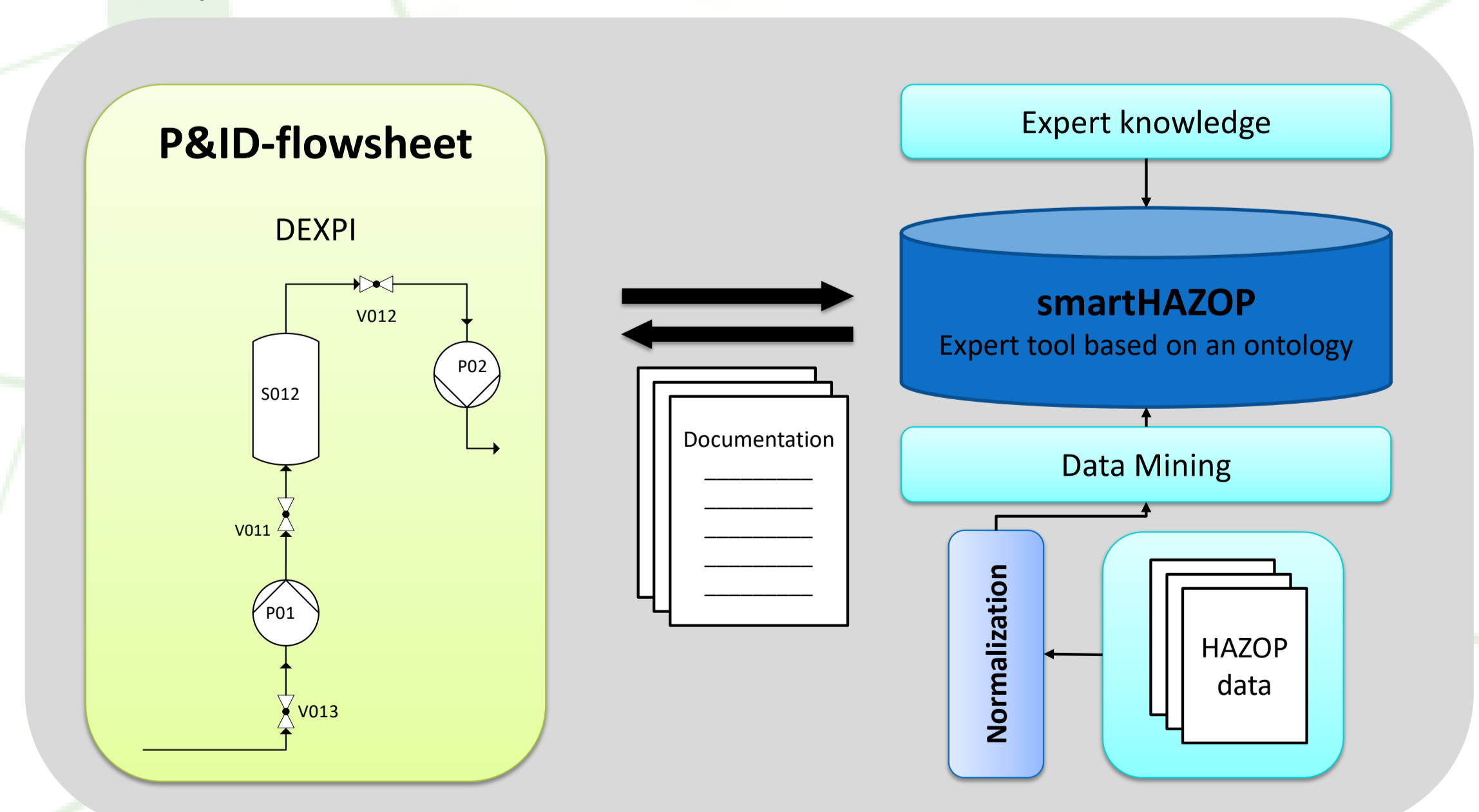


Conclusion

- To obtain information from HAZOP data, they must be transformed into a uniform format and entries must be normalized
- Normalization of entries can be done by analysis of part of speech in linkage of synonym databases (dictionaries)
 - First feasibilities could be proven by normalization of HAZOP data from industry (Bayer AG)
 - Reliable application requires higher robustness via targeted expansion of synonym databases

Future application

- Map Safety relations to DEXPI based P&IDs by using an ontology based HAZOP-expert tool
 - Expert tool can be derived from historical HAZOP data



Outlook

- Machine learning such as automated multi-labelling or word embeddings could be used for an extension of the synonym databases

[1] NLTK project, LTK 3.6 documentation, visited: 16.04.2021, <https://www.nltk.org/#>

[2] spaCy project, paCy v.3.0 API, visited: 16.04.2021, <https://spacy.io/api>

Contact