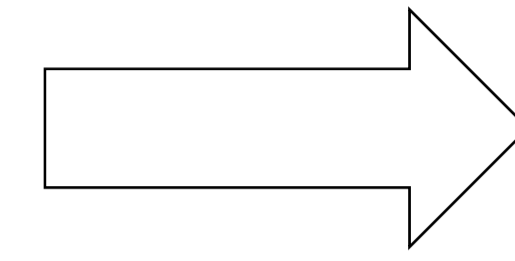
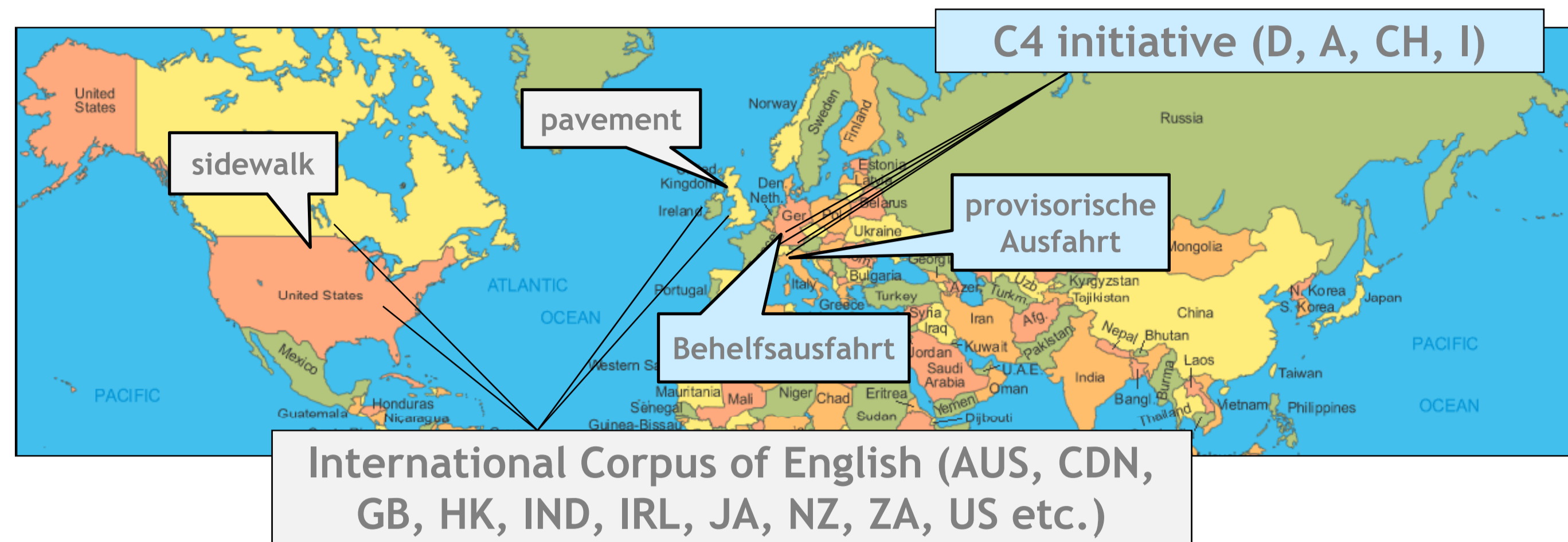


Stefanie Anstein
Institute for Specialised Communication and Multilingualism, EURAC research Bolzano
supervised by Uli Heid, IMS Stuttgart

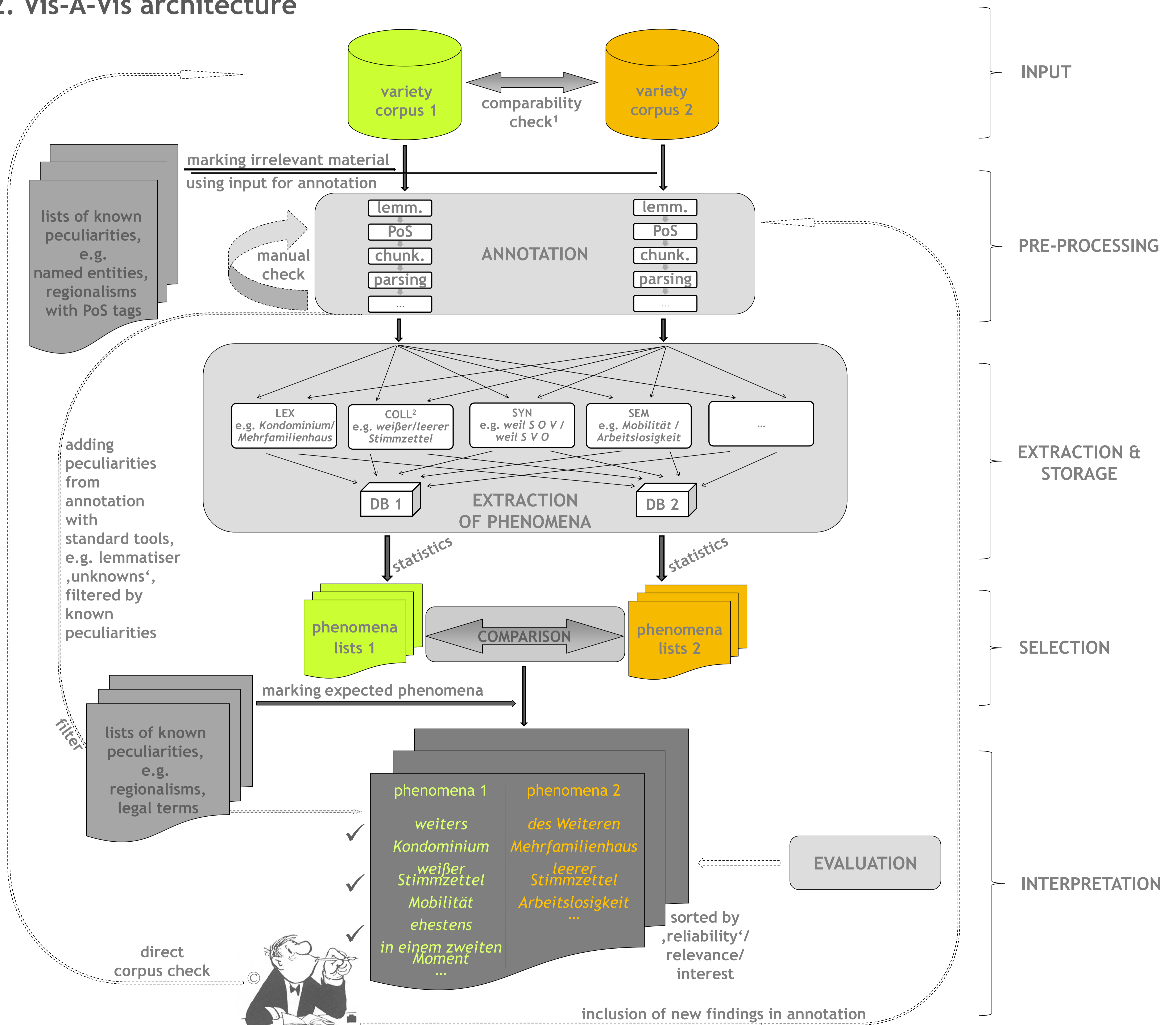
1. Varieties with different NLP coverage



corpus comparison for

- variety documentation <http://www.korpus-suedtirol.it>
- variant lexicography
- standardisation of varieties
- language teaching and learning
- etc.

2. Vis-À-Vis architecture



¹ e.g. Gries, S. T. (2007): Exploring variability within and between corpora: Some methodological considerations. *Corpora*, 109-151.

² e.g. Ritz, J. & Heid, U. (2006): Extraction tools for collocations and their morpho-syntactic specificities. In: *Proceedings of LREC, 1925-1930*.

3. Ongoing work

- implementing detailed workflow by integrating established and original tools
- collecting expected phenomena from studies on South Tyrolean German for their verification
- testing system for German varieties on the basis of C4 corpora - <http://www.korpus-c4.org>