

# DIGITAL HISTORY OF CONCEPTS: SENSE CLUSTERING OVER TIME

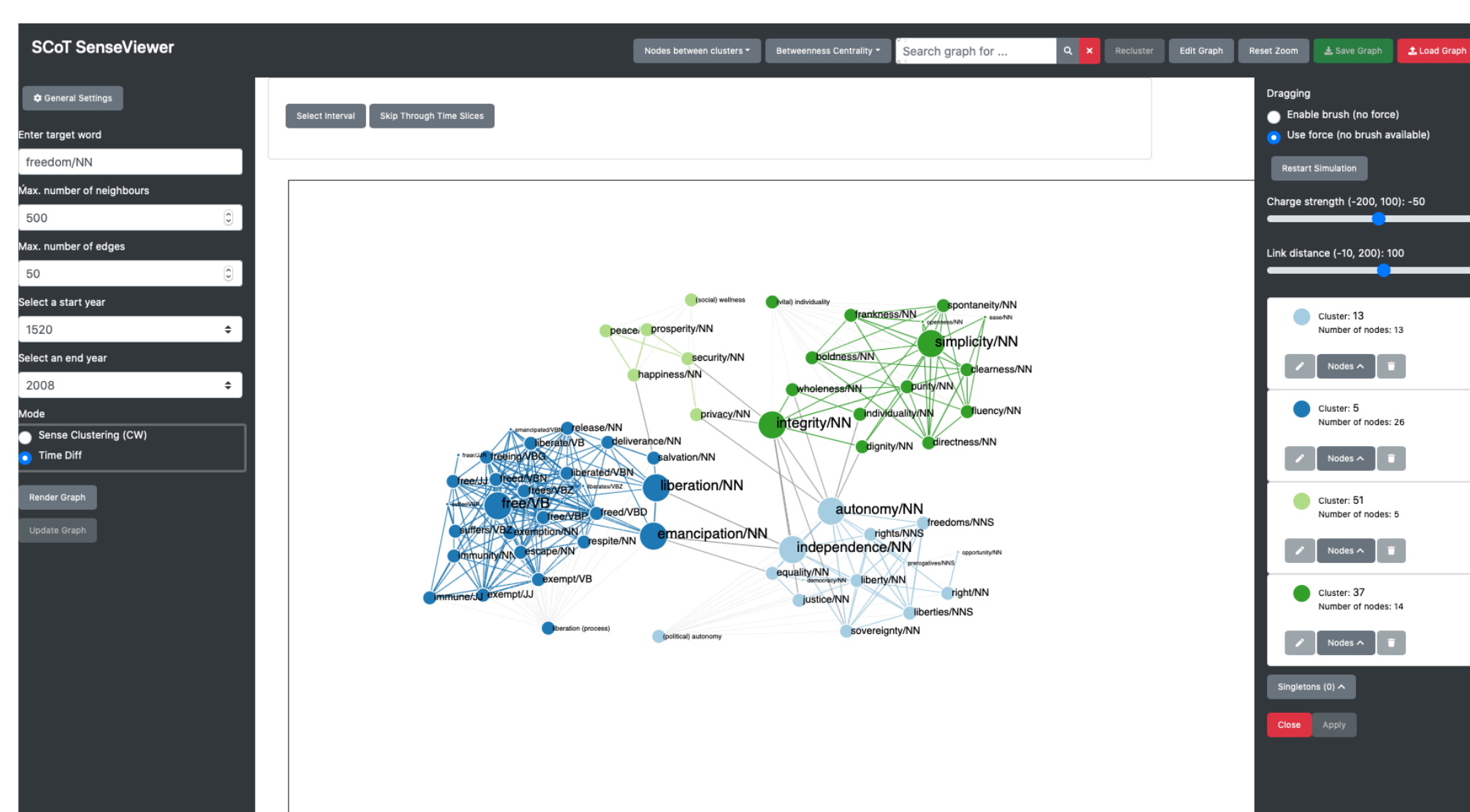
INGA KEMPFFERT<sup>1</sup>, SABA ANWAR<sup>1</sup>, ALEXANDER FRIEDRICH<sup>2</sup>, CHRIS BIEMANN<sup>1</sup>  
<sup>1</sup> U HAMBURG, <sup>2</sup> TU DARMSTADT

## ABSTRACT

We present **SCoT (Sense Clustering over Time)**, a **web interface for tracking and accessing word senses** over time to enable a digital, data-driven history of concepts in an ongoing collaboration between computational linguistics and philosophy [FB 2016].

Our exploration tool allows to **visualize the semantics of conceptual terms and their change over time**, based on digital text analysis.

The visualization shall help conceptual historians to identify and track significant changes of conceptual terms **within and across different periods of time in different text corpora** as well as to narrow down relevant contexts and source documents for further study.



The paradigms of the target word are displayed as a **graph, where word nodes are connected with edges indicating their similarity**. For the formation of concepts, **graph clustering [B06]** provides an automatic initialization of **colour-coded sense clusters**, which can be **labelled and post-edited**.

We employ colour coding to visualize differences and to show, for selected time intervals, which paradigm's respective senses are stable, are added, or fall out of use. This mode also allows for stepping through the time intervals for visually analysing the dynamics of change.

We use **distributional semantics and word sense induction** to enable this data-driven approach to tracking word senses over time.

## HISTORY OF CONCEPTS

Research in the history of concepts deals with the historical semantics of terms, with a special focus on **the changes of scientific, cultural or political basic concepts** related with them [MS 2016]. Basic concepts – such as „freedom, „power“, „life,“ or „time“ – are regarded as indispensable for the conception of objects or frames of reference in which we align our cultural, scientific or political understanding of the world and ourselves. Significant shifts of meaning in the semantics of basic concepts can therefore point to changes in our understanding of the world and ourselves.

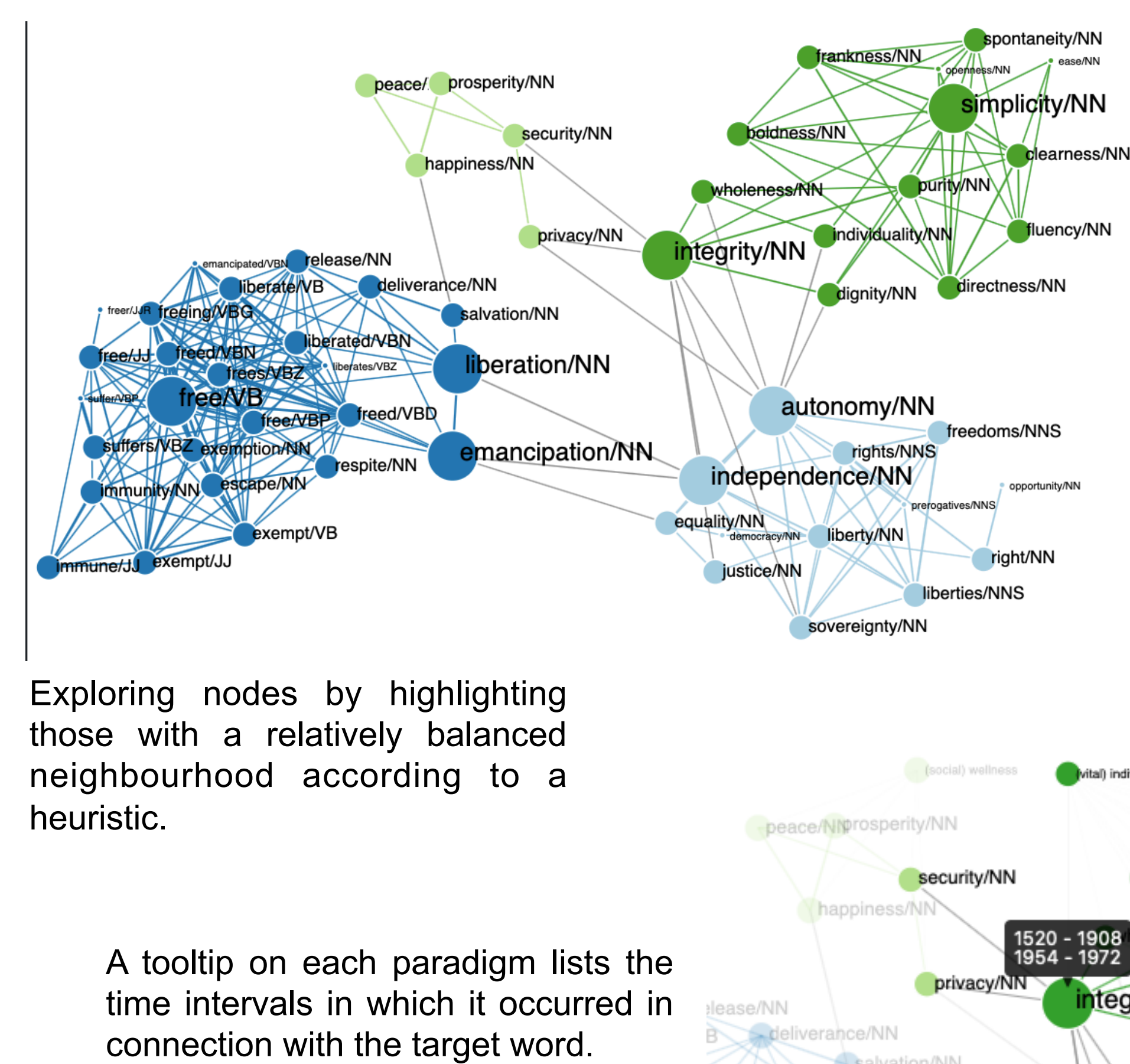
A crucial question in the history of concepts is how relevant contexts and periods of time of such changes can be identified. With our tool, we are able to **detect such changes over time**. The diachronic analysis of the term ‚freedom‘ for example indicates a semantic reduction of the term to its political meaning (in the sense of ‚liberty‘). Finding reason for this reduction would be a possible task of a conceptual history of ‚freedom‘ based on this tool.

## WORKING WITH SCOT

The user can **edit** the clusters and correct the system's initial hypothesis. Amongst other things, the cluster names and colours can be modified. The user can also label the sense clusters and change the cluster assignment of any word. The user can also add new clusters if the paradigms are too coarse and delete whole clusters if they are not of interest. Adding more nodes to the graph and re-clustering is also possible.

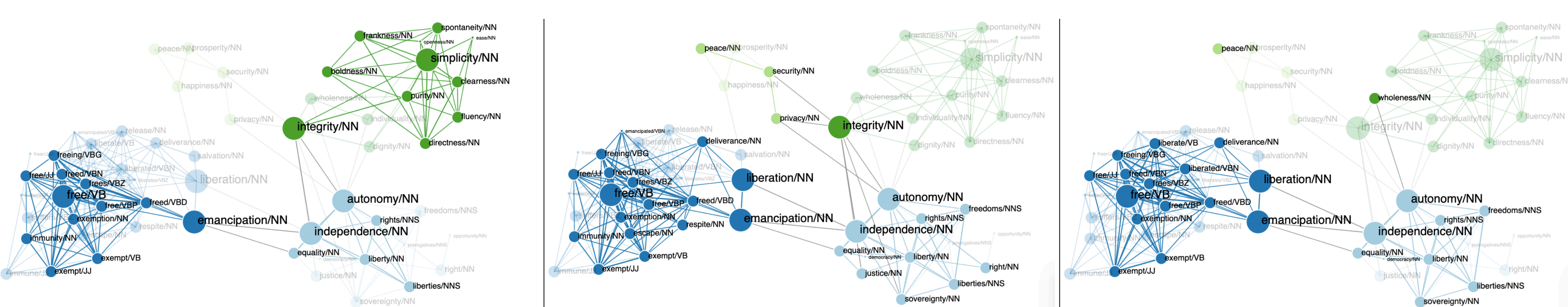
The user can **explore** the graph via searching for certain paradigms, applying different centrality measures, or stepping through the different time slices of a concept. The user can also explore which senses of a paradigm are stable or fall out of use or are added to it at a certain point in time.

The user can save sessions in a file and continue working on a previously saved graph, which can be shared with others.



## VISUALIZING CHANGE IN CONCEPTS

A user can step through the individual time slices of a concept. Here, the change in the concept „freedom“ is visualized as an example:



1520 - 1908

1954 - 1972

2006-2008

## AVAILABLE CORPORA

- **Google Books 1520 – 2008 (English)**  
<http://ltdemos.informatik.uni-hamburg.de/scot/>
- **FinnNews 1999 – 2015 (Finnish)**  
<http://ltdemos.informatik.uni-hamburg.de/finnews/>

## REFERENCES

- Biemann, C. (2006): Chinese whispers: an efficient graph clustering algorithm and its application to natural language processing problems. Proceedings of TextGraphs-1, pp. 73–80, New York City, NY, USA
- Friedrich, A. & C. Biemann (2016). Digitale Begriffsgeschichte? Methodologische Überlegungen und exemplarische Versuche am Beispiel moderner Netzsemantik, *Forum Interdisziplinäre Begriffsgeschichte* 5(2):78-96
- Müller, E. & F. Schmieder (2016): *Begriffsgeschichte und historische Semantik: ein kritisches Kompendium*. Berlin: Suhrkamp.