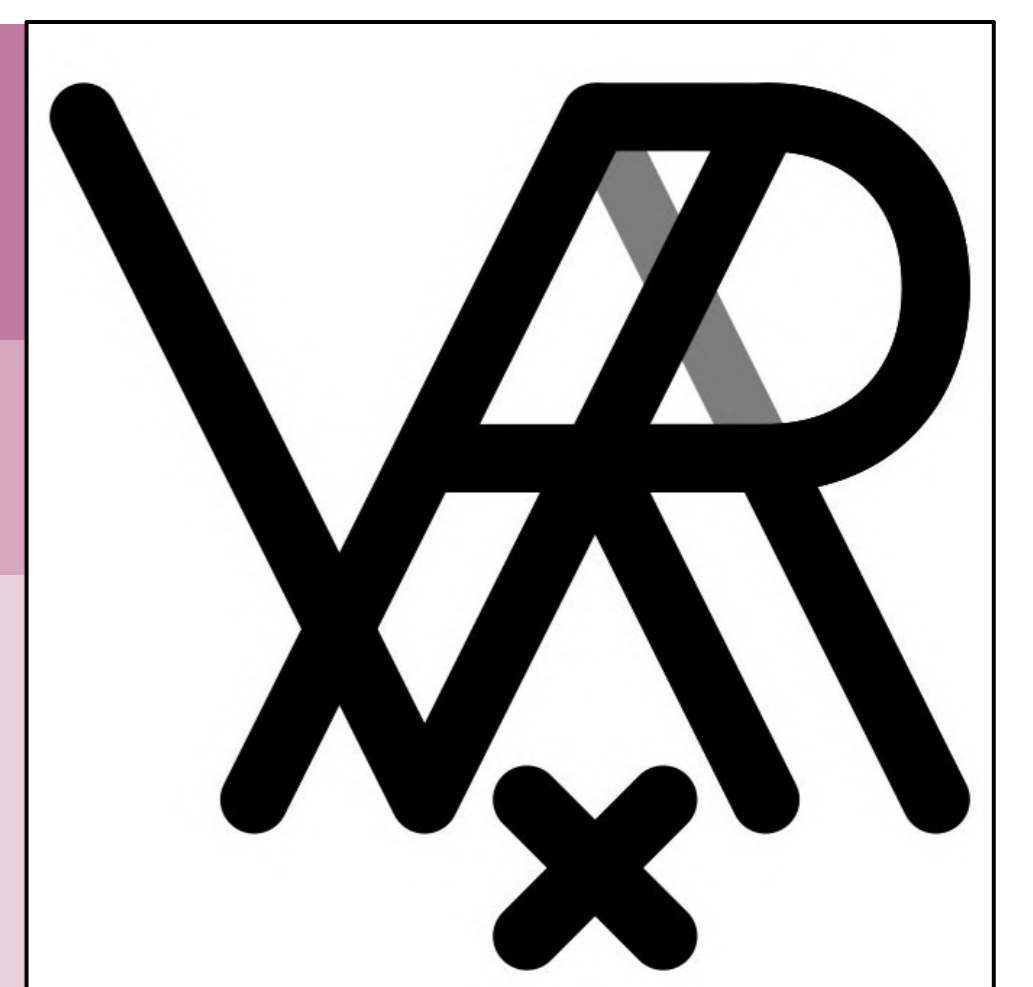


ODXVR x NTS



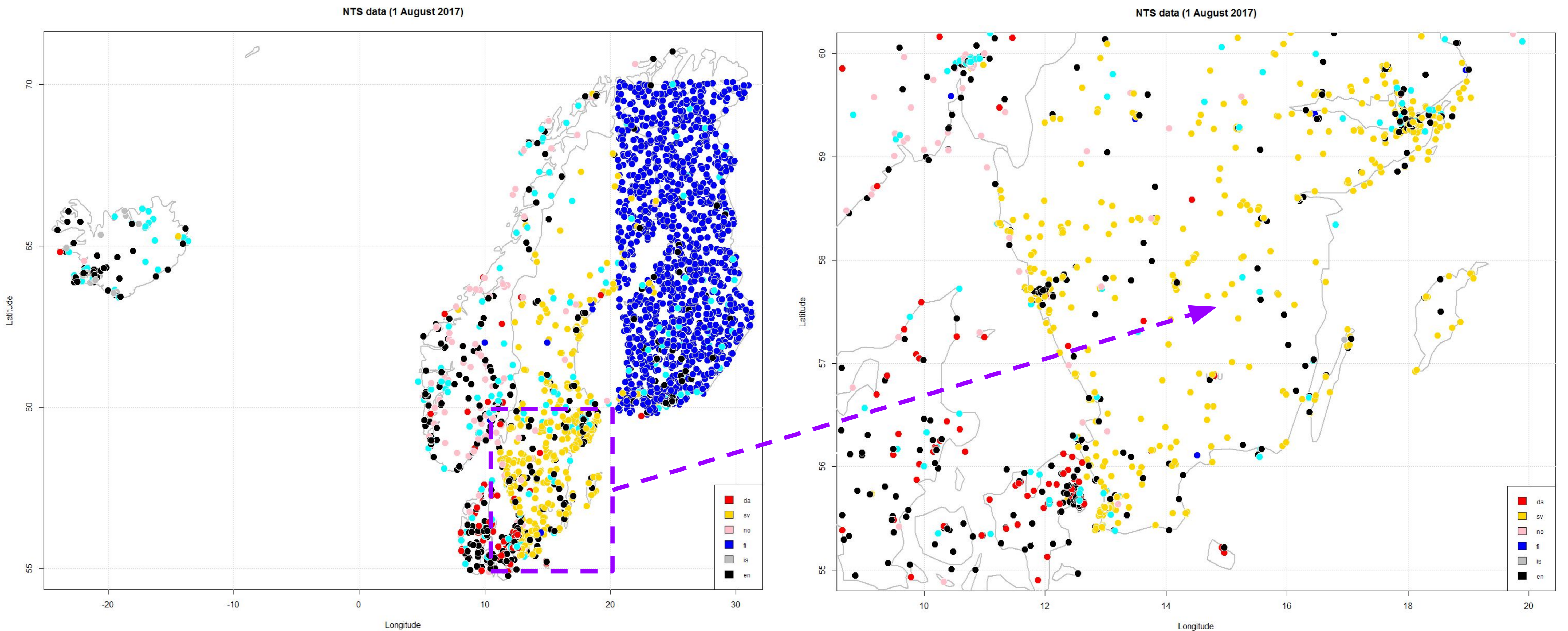
Open Data Exploration in Virtual Reality x Nordic Tweet Stream (a dynamic corpus of tweets)

Aris Alissandrakis, Nico Reski - VRxAR Labs
Mikko Laitinen, Jukka Tyrkkö, Magnus Levin - Department of Languages
Jonas Lundberg - Department of Computer Science

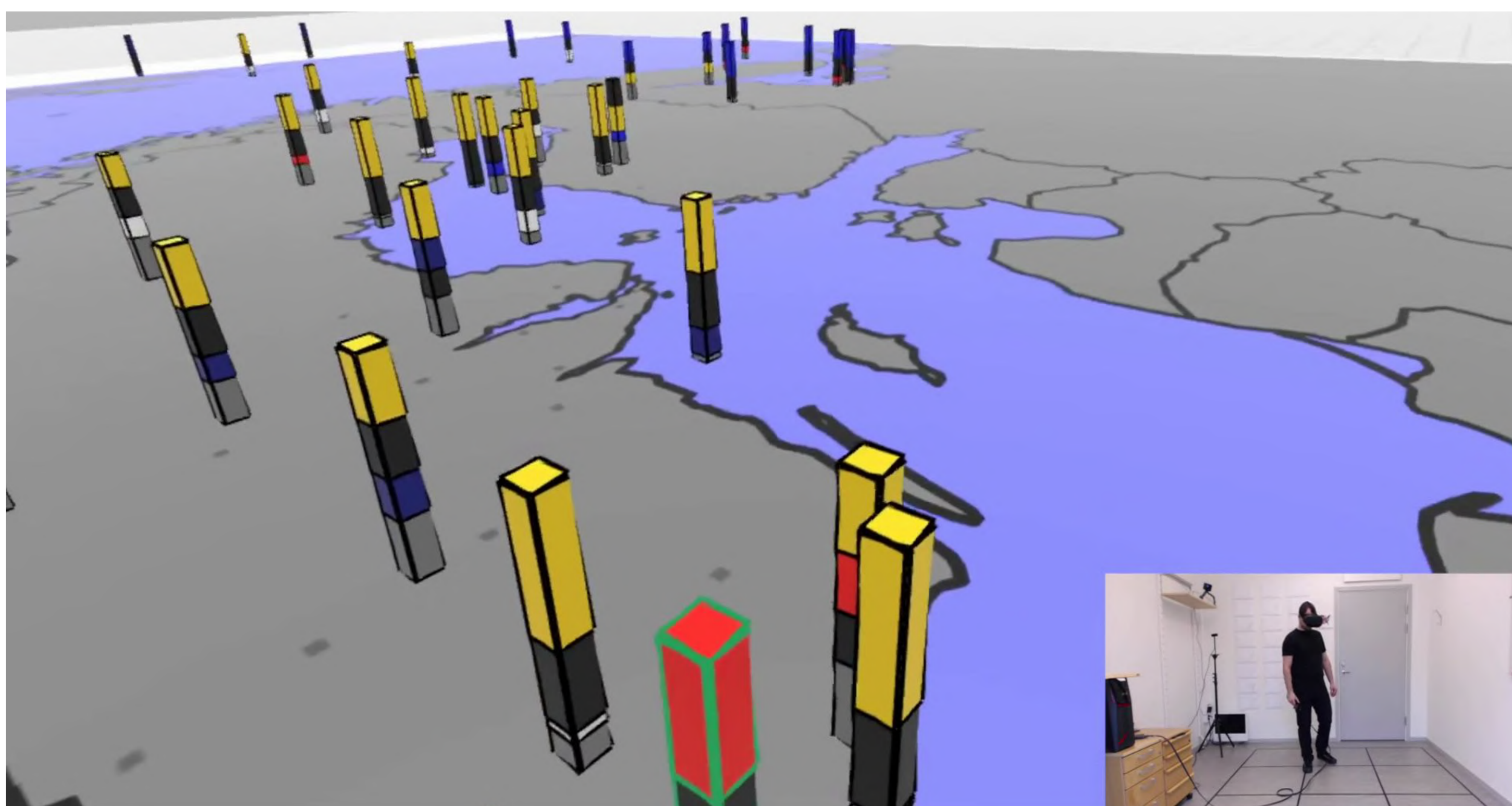
This project within the Data Intensive Digital Humanities group of DISA combines two other projects:

- *Open Data Exploration in Virtual Reality* (ODxVR, Reski and Alissandrakis, 2016) by VRxAR Labs, and
- *Nordic Tweet Stream* (NTS), a cross-disciplinary corpus project of computer scientists and a group of sociolinguists interested in language variability and in the global spread of English (Laitinen et al., forthcoming).

The overall aim of this project is to make the NTS corpus data more accessible to students, and allow them to gain insights on how the distribution of languages changes geographically, and over time.

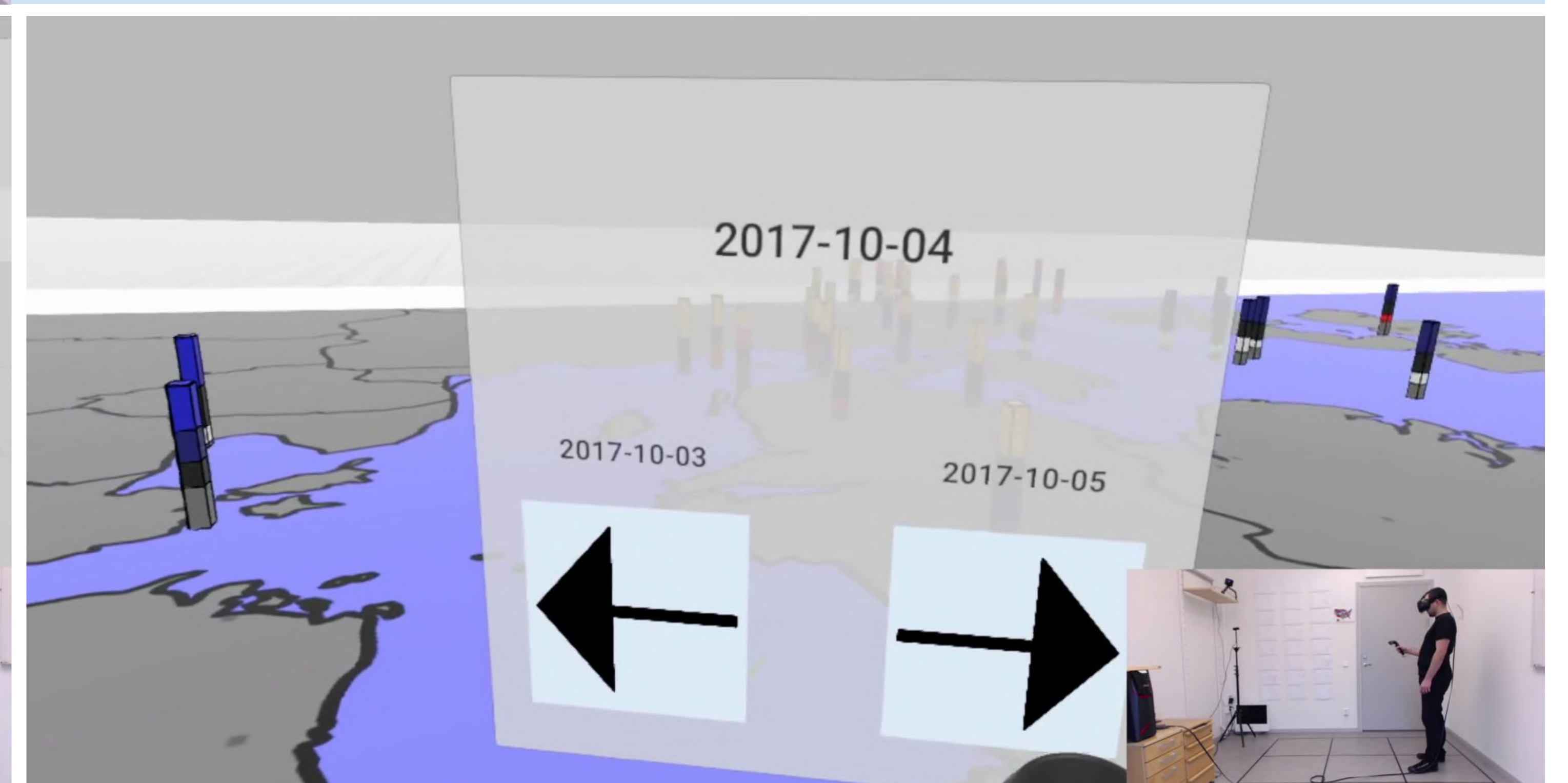
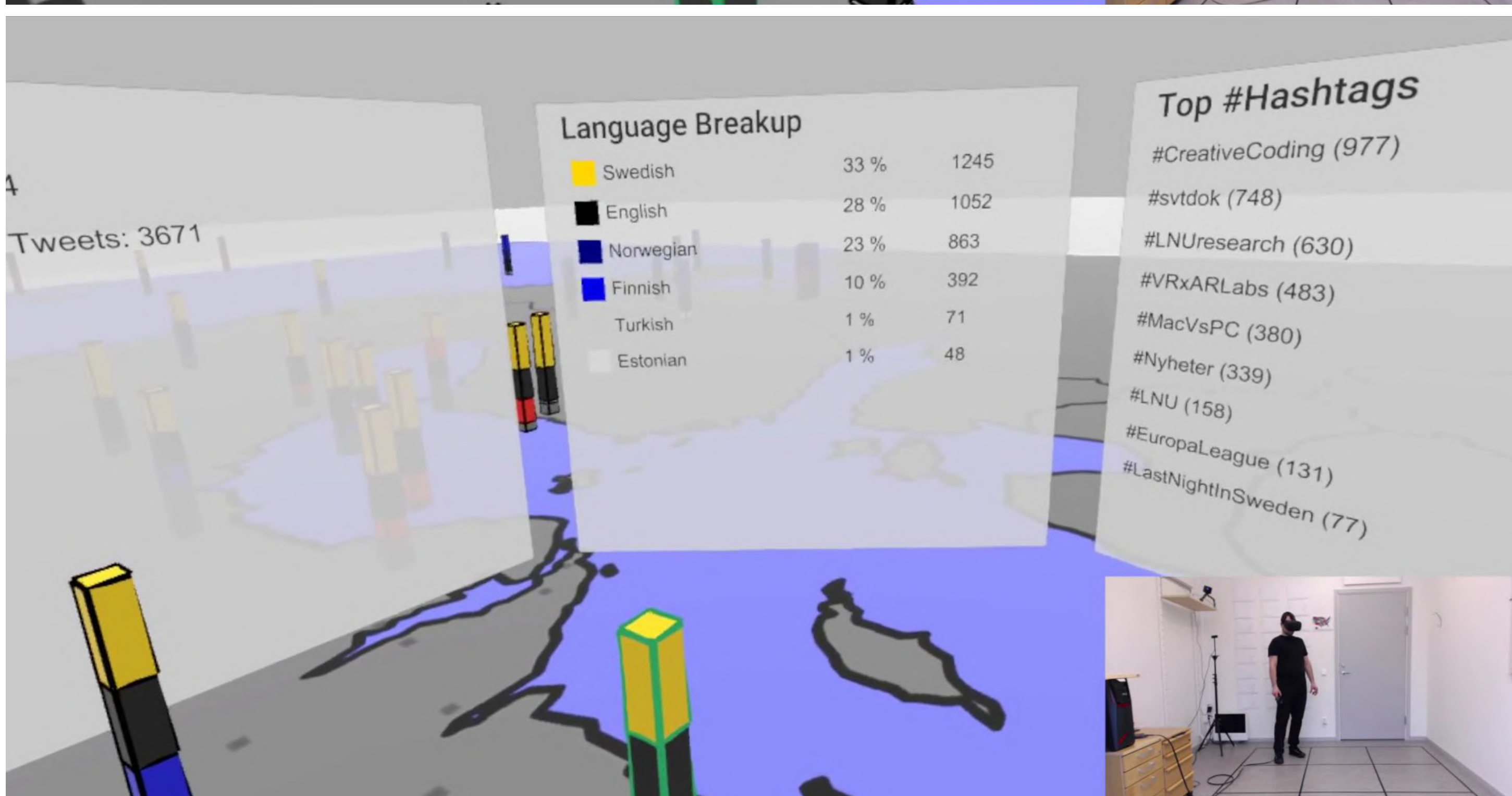


The two figures above show 2D visualizations of the NTS data for just a single day (entire region, left; close up, right). Although aesthetically pleasing, it is arguably more difficult for them to be used by non-experts to gain insights from such big datasets.



We developed an application that allows a user to explore and interact with tweets and associated metadata downloaded by the NTS project from Denmark, Finland, Iceland, Norway and Sweden, within a virtual reality environment.

The data are visualised as nodes floating above a map of the Nordic region. Each node is a stacked cuboid representing the percentages of the top three languages within aggregated tweets for a specific area (town, county, region, etc).



Each individual node can be further explored, providing a more detailed breakup of all the represented languages, along with popular hashtags, overall number of tweets for that area, and other demographic information.

The NTS corpus currently consists of more than one year's worth of Twitter data. Within the virtual environment, besides moving spatially exploring different Nordic areas, it is also possible for a user to move through time to different time periods (days, weeks, months, etc) and compare how the language distributions varied.

- Reski, N. & Alissandrakis, A. (2016). **Change your Perspective: Exploration of a 3D Network created from Open Data in an Immersive Virtual Reality Environment.** In Proceedings of the 9th International Conference on Advances in Computer-Human Interactions (ACHI 2016), April 24-28, Venice, Italy. pages 403-410
- Laitinen, M., Lundberg, J., Levin, M. & Lakaw, A. Forthcoming. **Creating the Nordic Tweet Stream: A real-time monitor corpus of rich and big data.** Journal of Universal Computer Science.