Analyzing the Language of Food on Social Media
Talk by Stephen Kobourov, University of Arizona

Úterý 1. prosince 13.30 v místnosti 5.066

Abstract: We investigate the predictive power behind the language of food on social media. We collect a corpus of over three million food-related posts from Twitter and demonstrate that many latent population characteristics can be directly predicted from this data: overweight rate, diabetes rate, political leaning, and home geographical location of authors. For all tasks, our language-based models significantly outperform the majority-class baselines. Performance is further improved with more complex natural language processing, such as topic modeling. We analyze which textual features have most predictive power for these datasets, providing insight into the connections between the language of food, geographic locale, and community characteristics. Lastly, we design and implement an online system for real-time query and visualization of the dataset. Visualization tools, such as geo-referenced heatmaps, semantics-preserving word-clouds and temporal histograms, allow us to discover more complex, global patterns mirrored in the language of food. The results of this work were covered by media outlets, including the Washington Post and the Guardian.

Short Bio: Stephen Kobourov is a Professor of Computer Science at the University of Arizona. He completed BS degrees in Mathematics and Computer Science at Dartmouth College in 1995, and a PhD in Computer Science at Johns Hopkins University in 2000. He was a Research Scientist at AT&T Research Labs, a Humboldt Fellow at the University of Tübingen in Germany, and in 2015-16 is working as a Distinguished Fulbright Chair at Charles University in Prague. His main research area is computational geometry and theoretical computer science where he published over 150 papers in scientific journals and conferences.