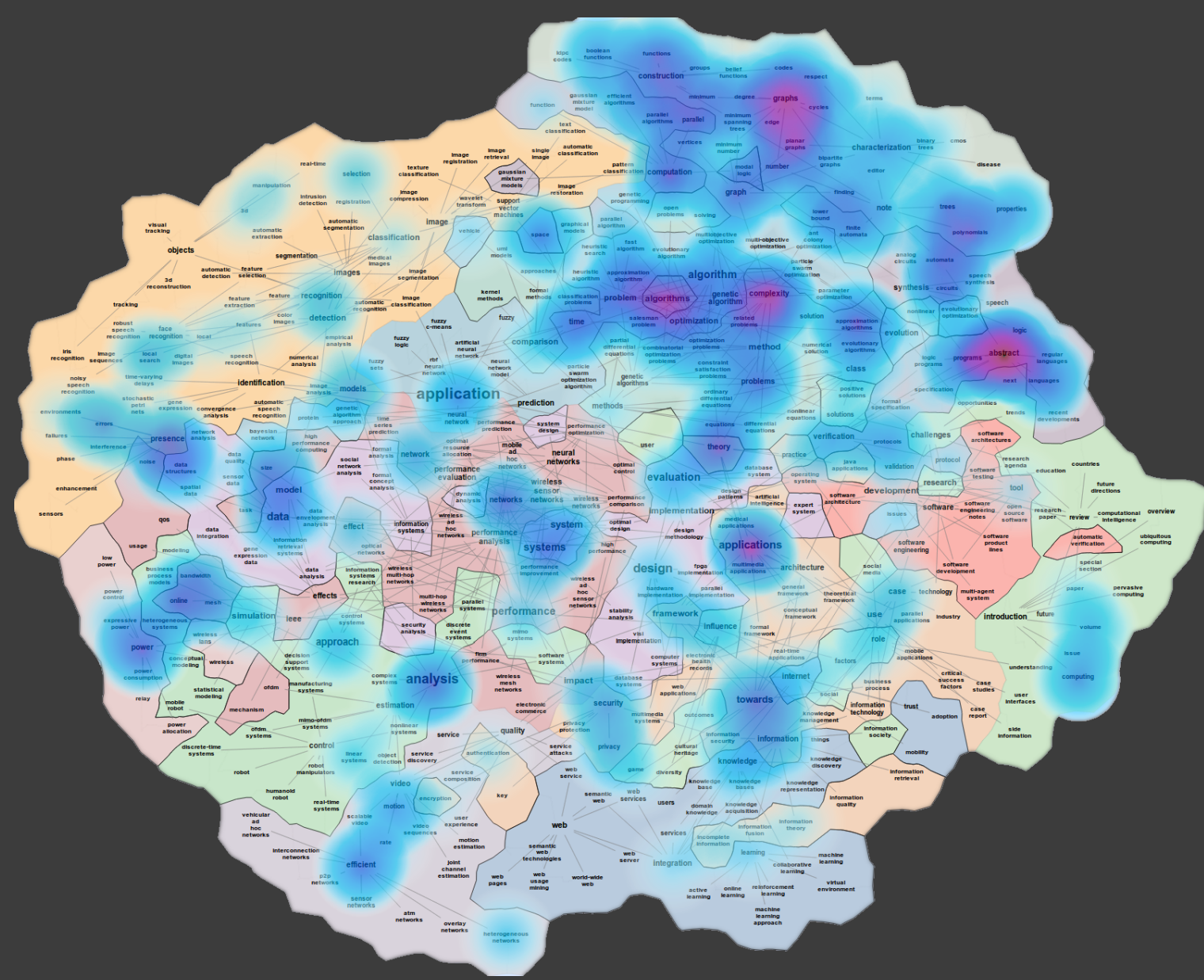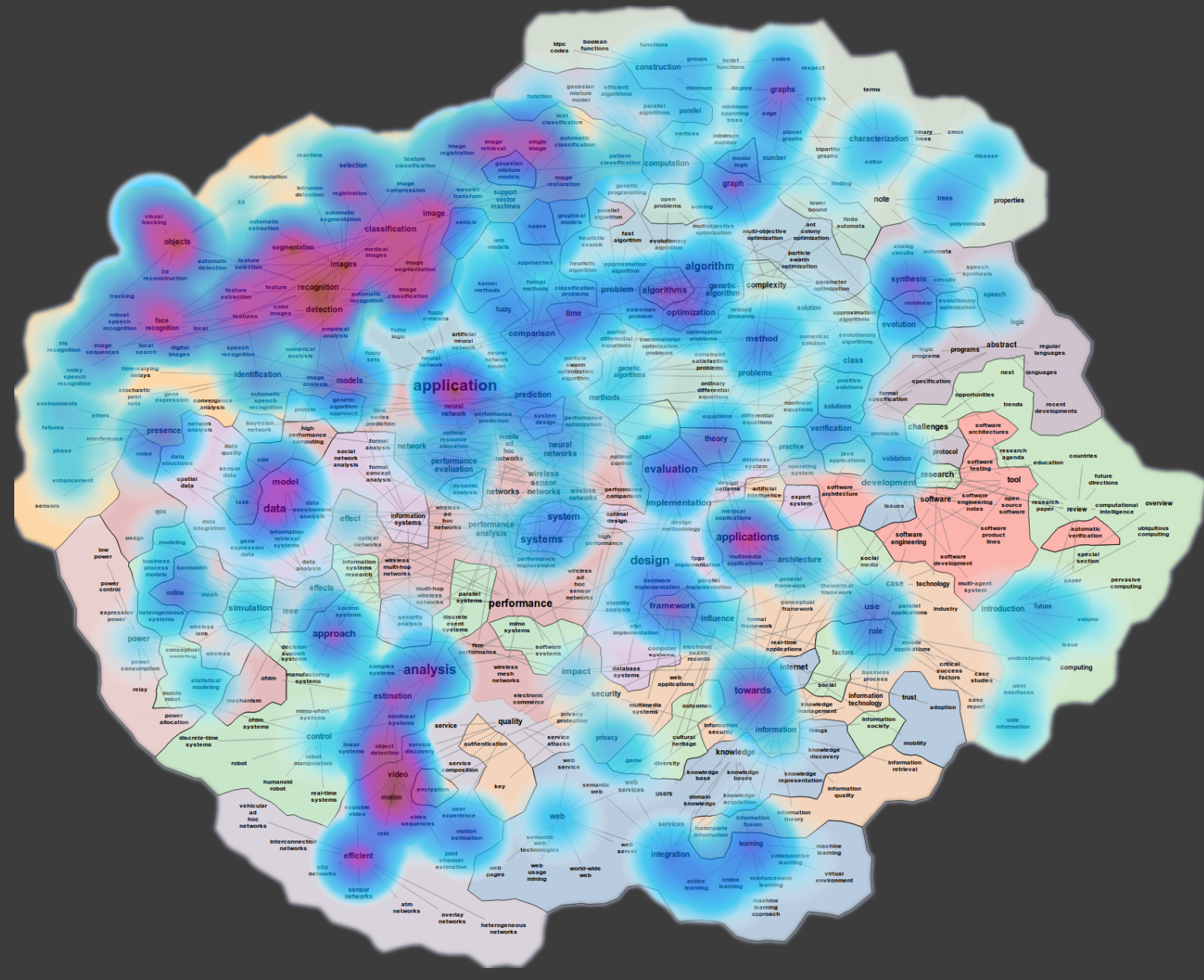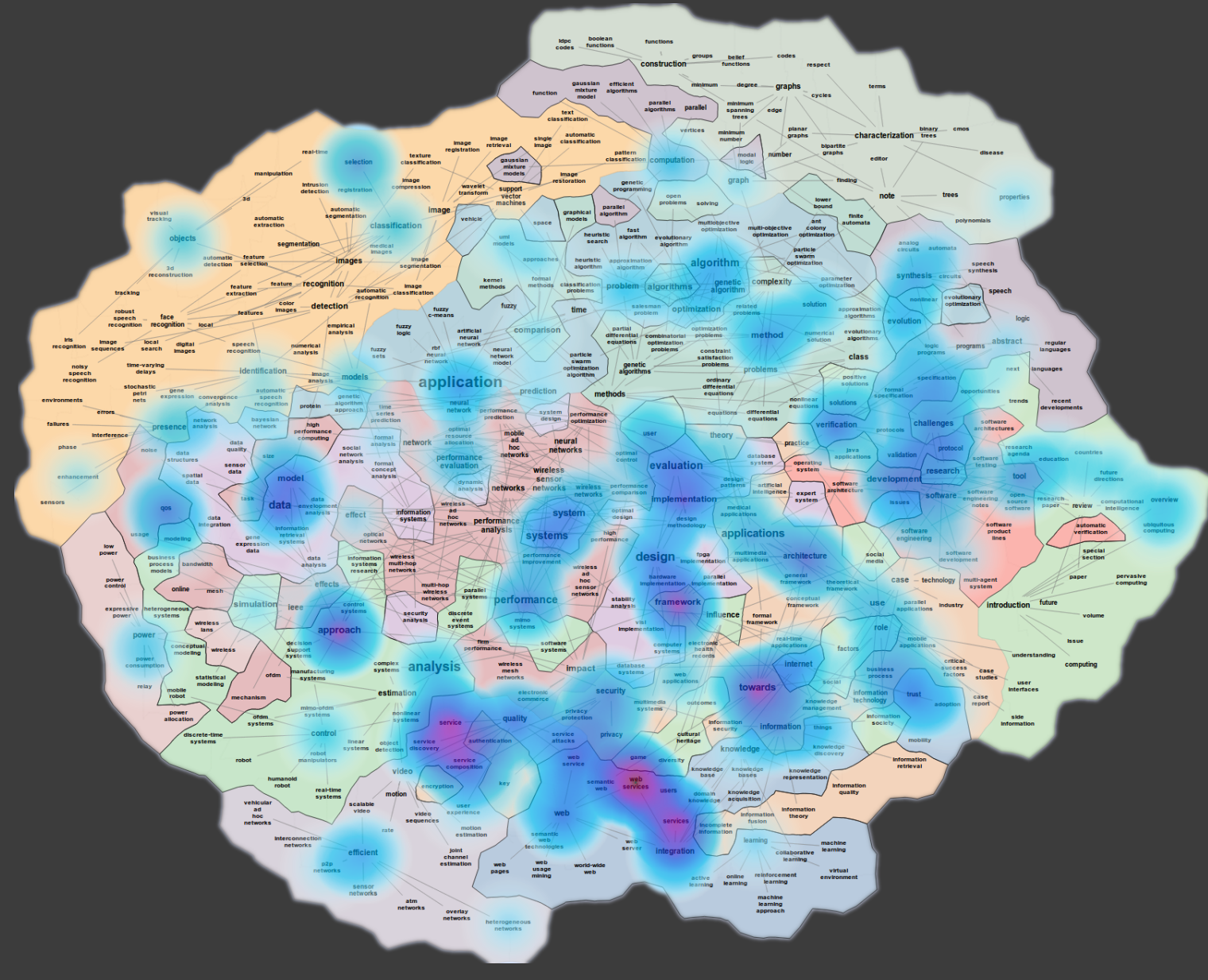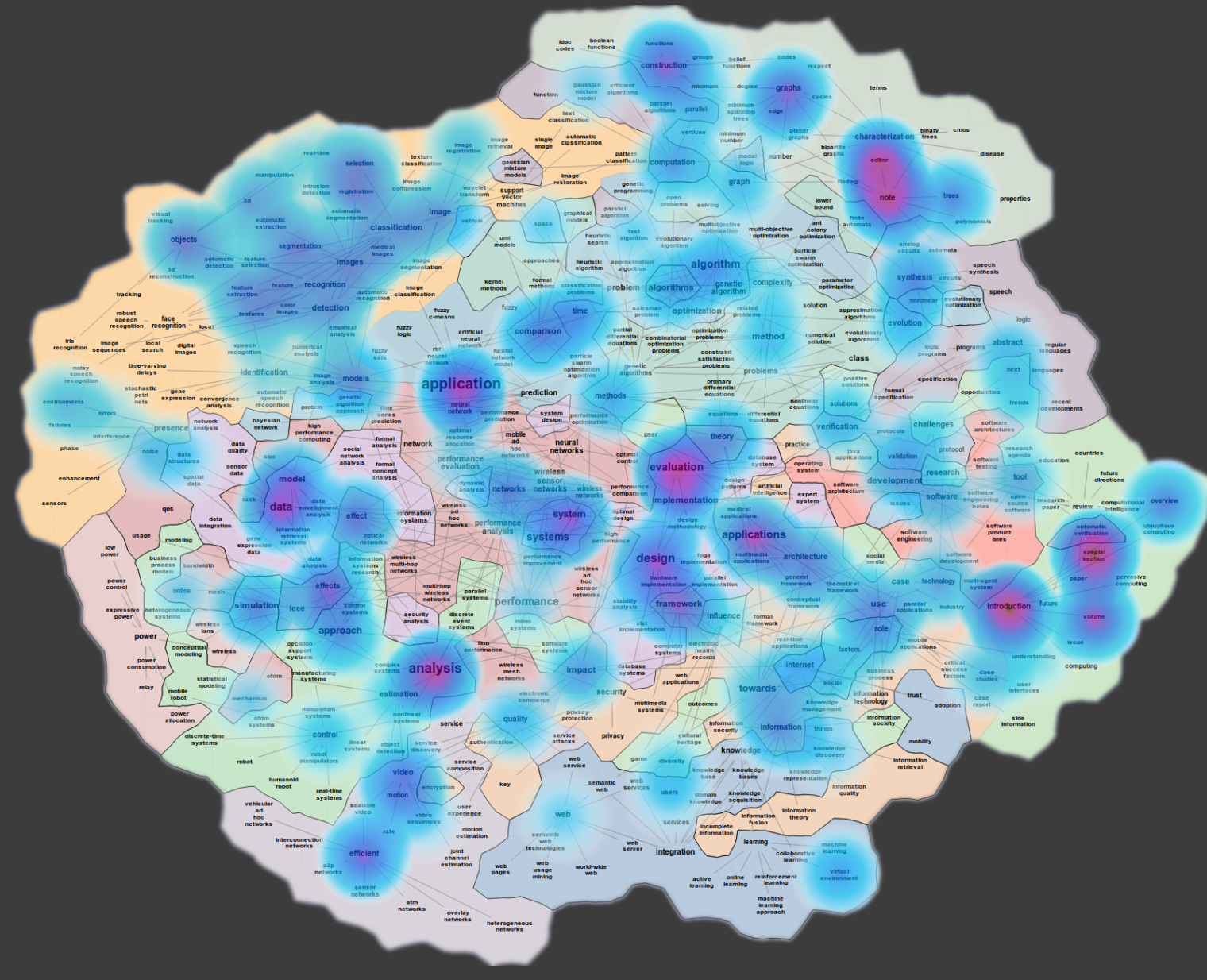# Maps of Computer Science

Venue Heatmaps

Symposium on Theory of Computing

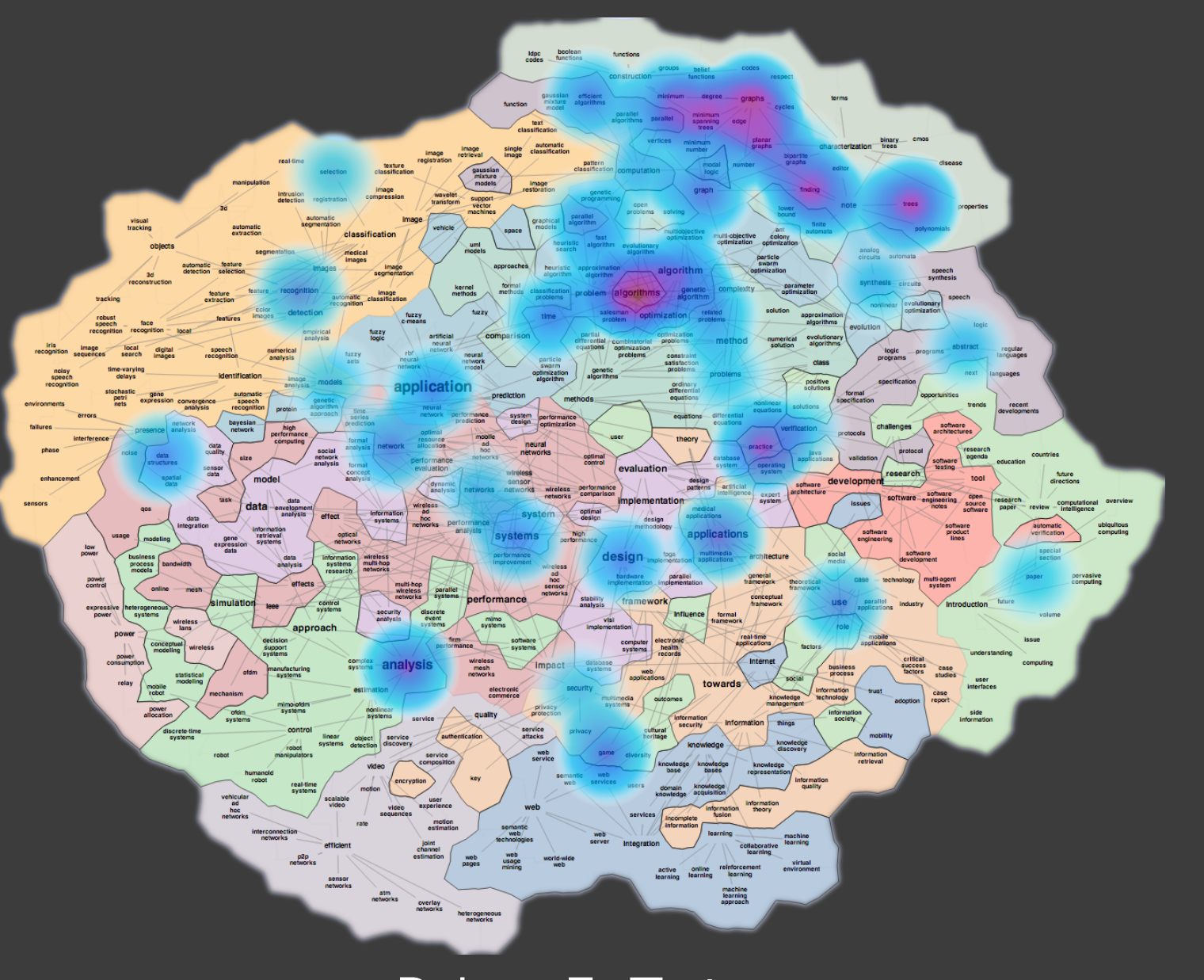Computer Vision and Pattern Recognition
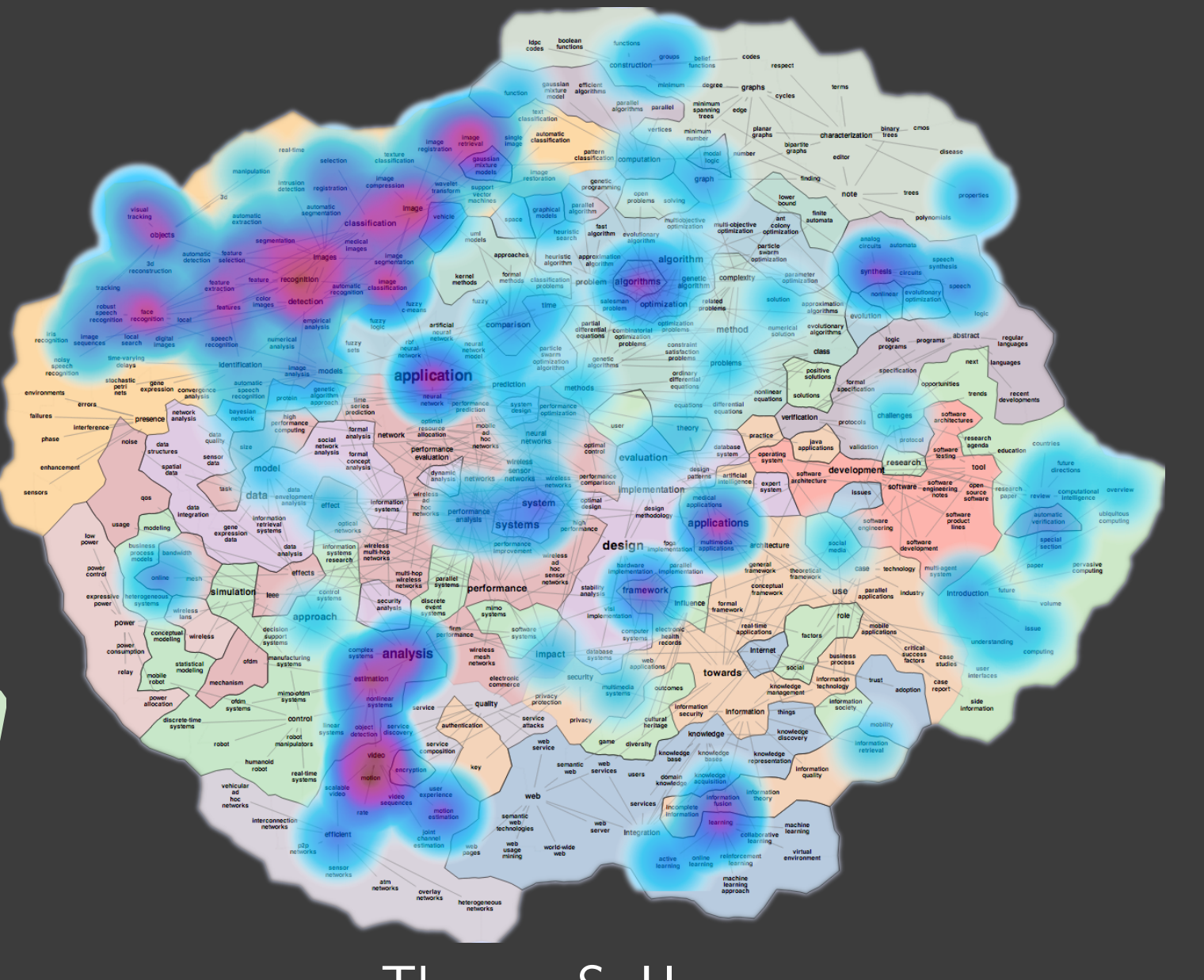

International Conference on Web Services

Transactions on Visualization and Computer Graphics

The venue heatmaps above show profiles of particular conferences and journals over the basemap of all of computer science. Each venue's heatmap exhibits a distinct location over the topic basemap: STOC in the northeast over theory topics, CVPR in the northwest over terms related to images and pattern recognition, ICWS in the south over Internet and web topics, and TVCG is associated with visualization topics in all areas of computer science.
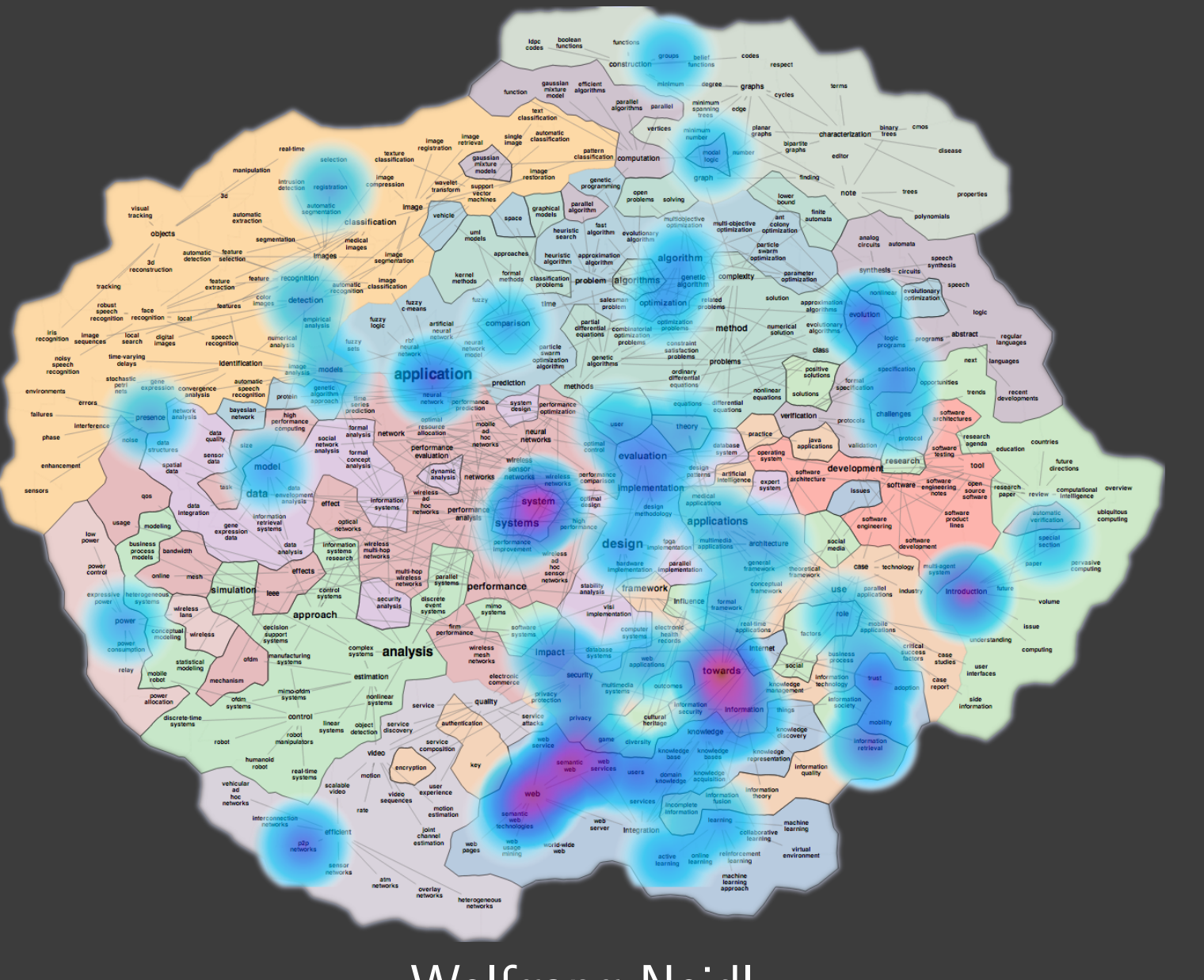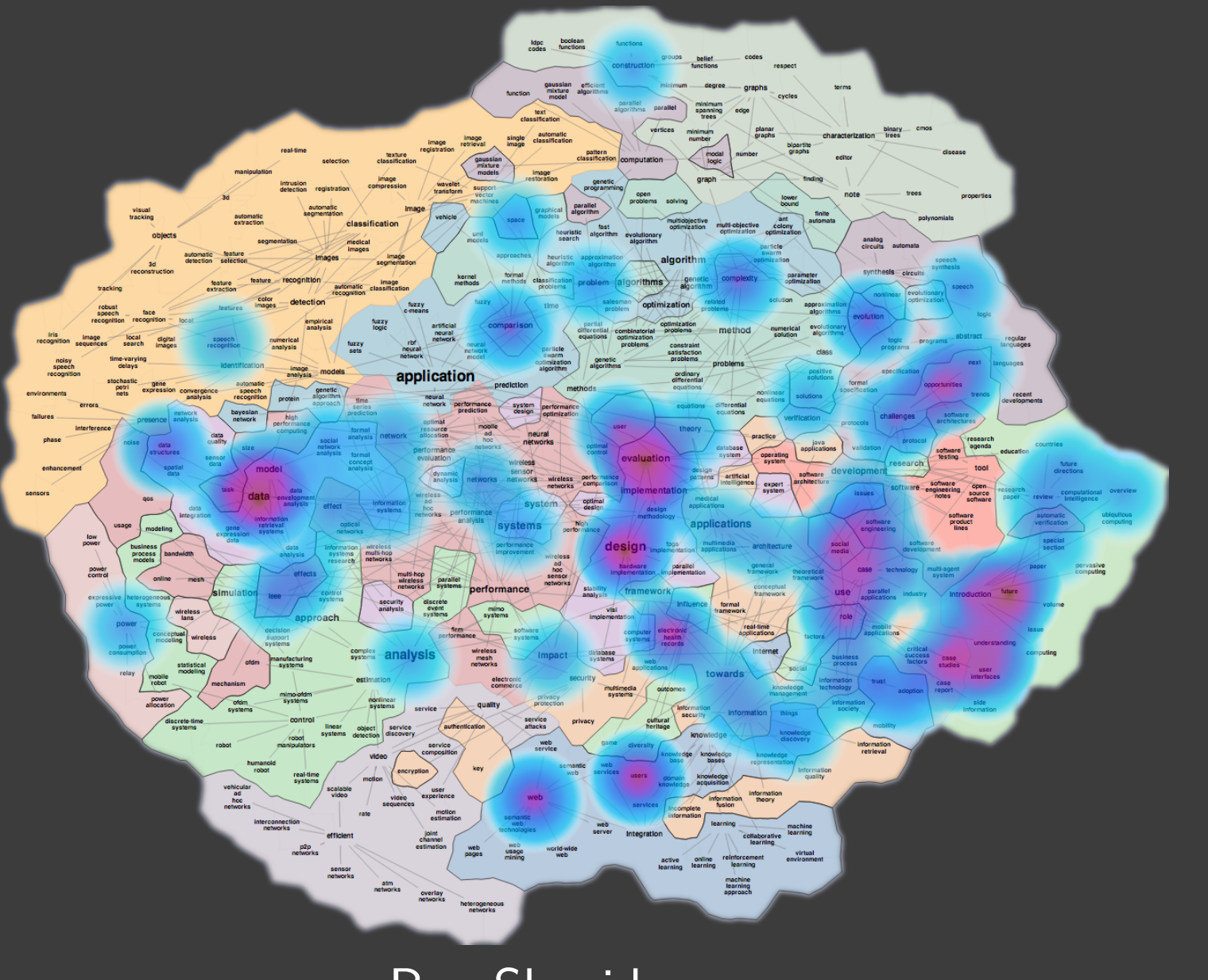
Author Heatmaps
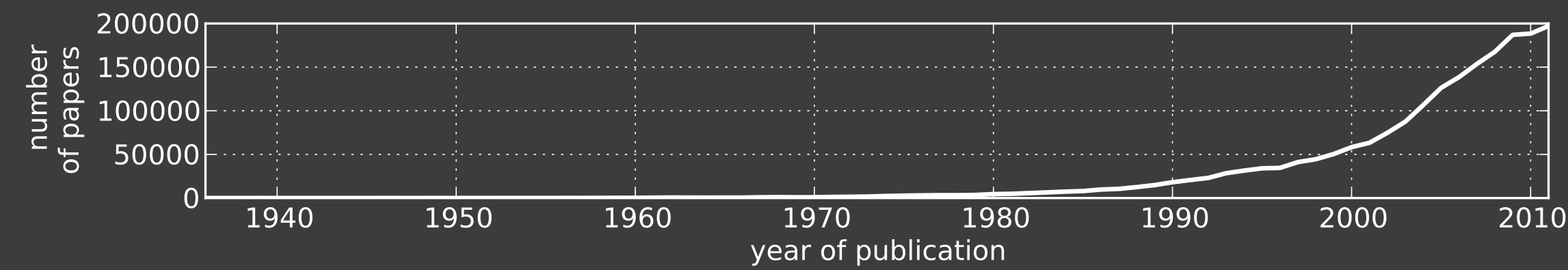

Robert E. Tarjan

Thomas S. Huang

Wolfgang Nejdl

Ben Shneiderman

The author heatmaps above show profiles of particular researchers over the basemap of all of computer science. Researchers were selected with research interests corresponding to the venue heatmaps at left: Tarjan paired with STOC, Huang with CVPR, Nejdl with ICWS, and Shneiderman with TVCG. Each author heatmap and the corresponding venue heatmap cover similar areas of the underlying basemap, illustrating an alignment of the researcher's interests and the topics covered in the conference or journal.

These topical maps show popular research areas in computer science, generated from the titles of computer science research papers in the DBLP computer science bibliography. Larger terms appear more frequently in paper titles. Terms that frequently appear together in titles are drawn close to one another in the map, and clustered to create "countries" of related topics. Terms are extracted by matching part-of-speech tags and ranked using the *C-value* algorithm. Term co-occurrence similarity is calculated using a modified Jaccard Coefficient function. 70,000 papers were sampled from DBLP to create the map above.

DBLP Papers

Maps are made by extracting terms from paper titles in the DBLP bibliography of computer science papers. DBLP contains

- – 2,184,720 papers
- – 1,324 journals
- – 6,904 conferences
- – 1,237,445 authors
- – A date range of 1936–2012
- – An average of 10.3 words per title
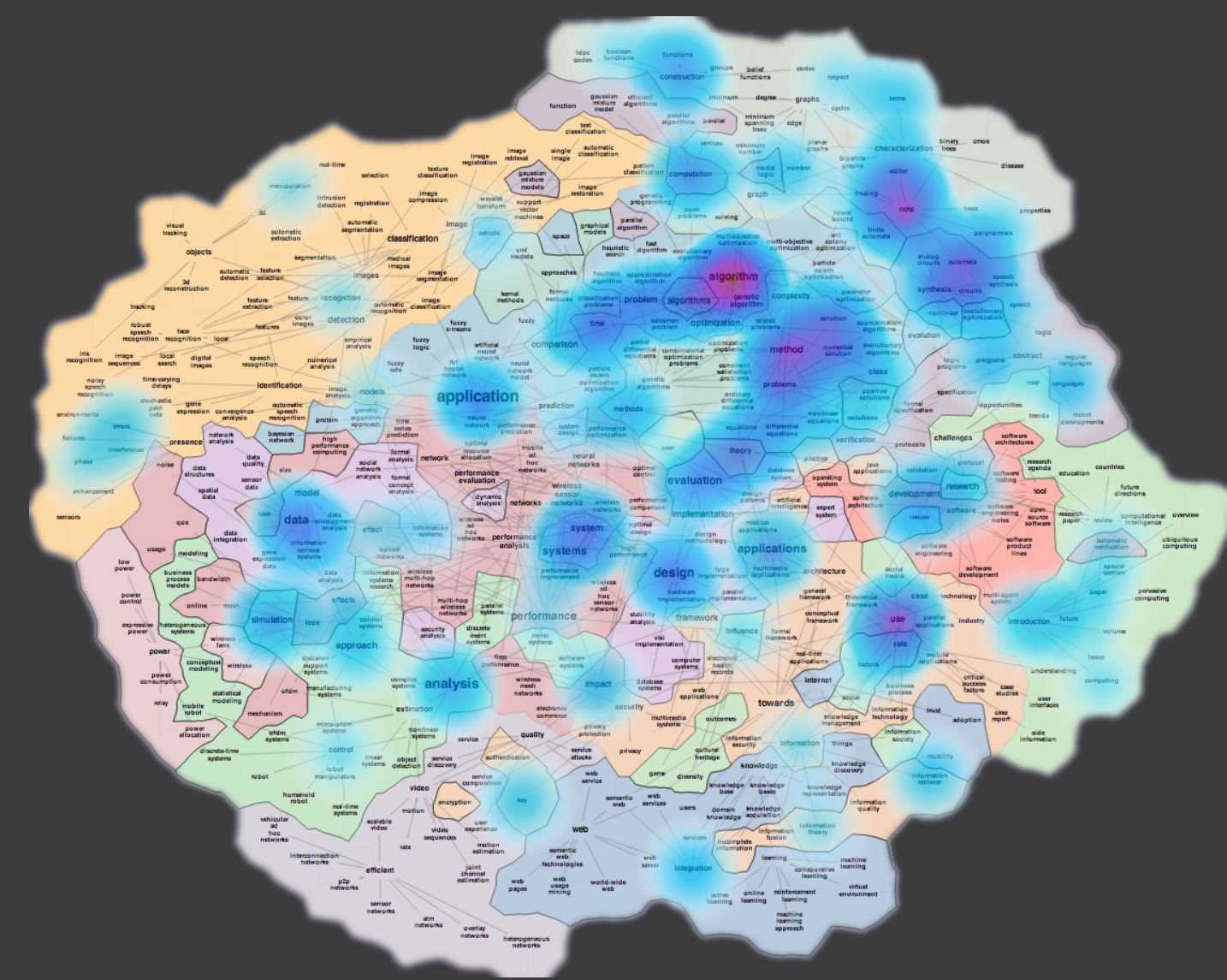- – An average of 2.9 terms per title
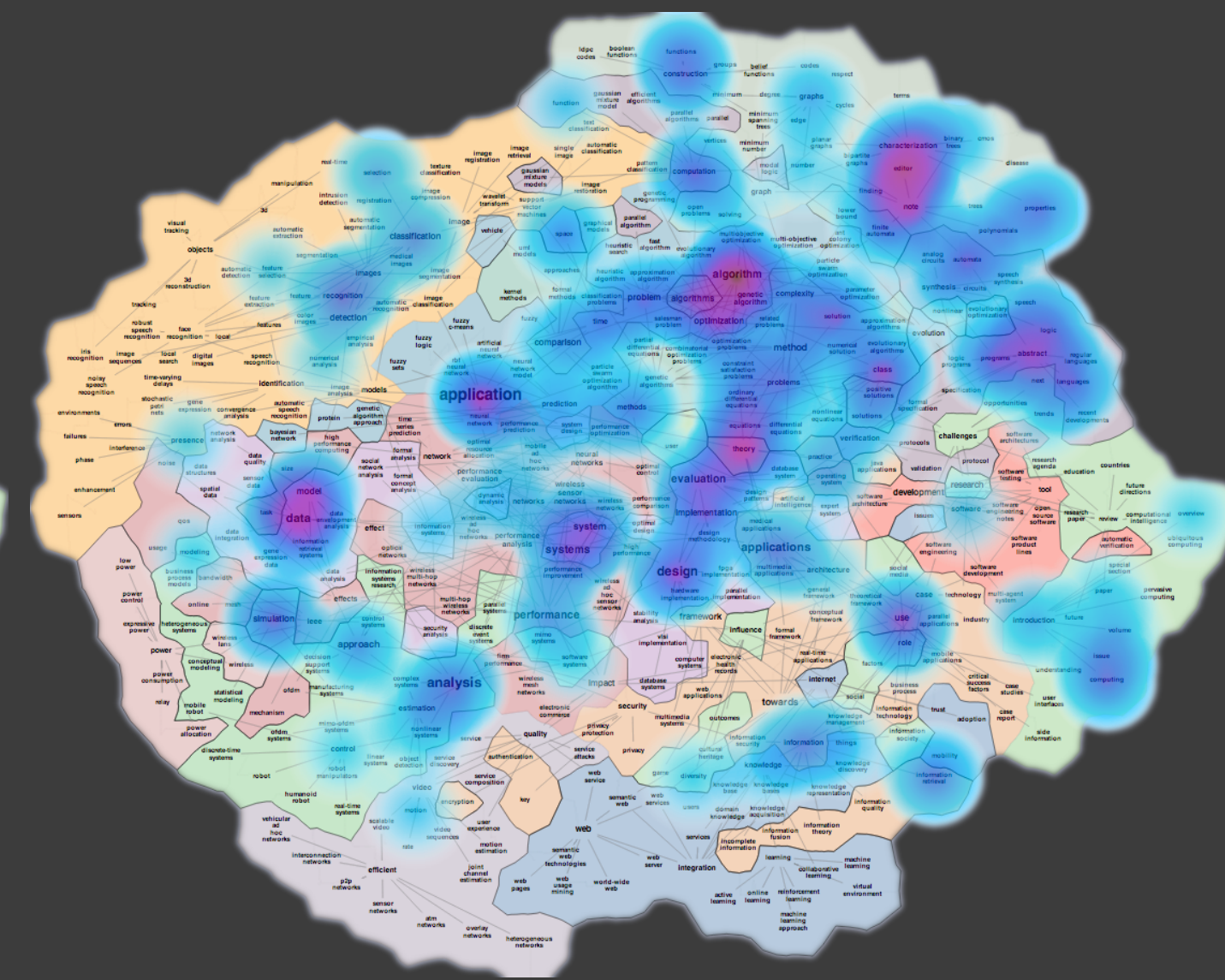

Heatmap Construction

Heatmaps show term usage within a target set of papers, from a particular author, venue, or time period. The color intensity of the heatmap is determined by how frequently the terms underneath were used in the target set of papers, showing the importance of these topics within the documents.

Least Frequently Used Terms ⸺ Most Frequently Used Terms
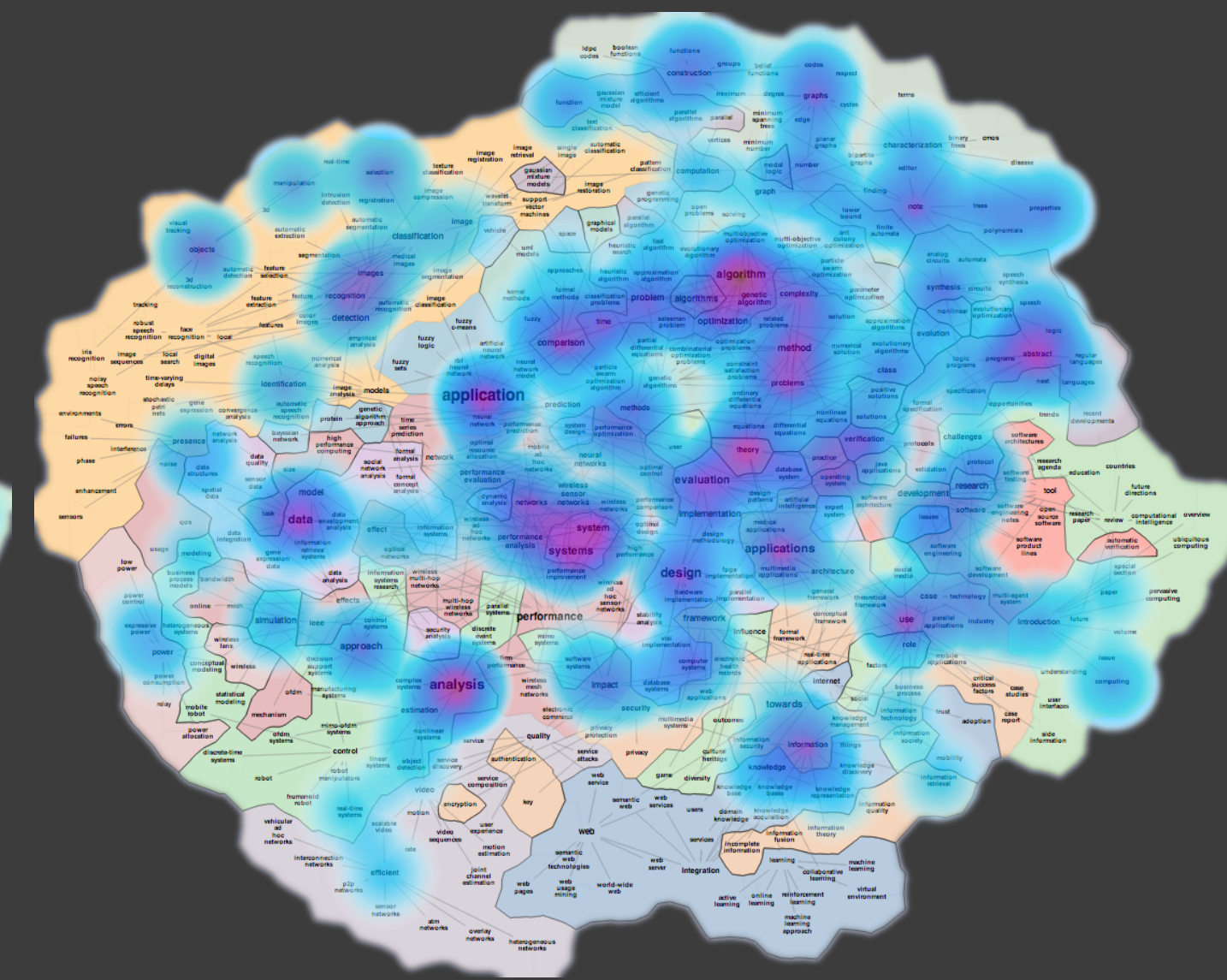
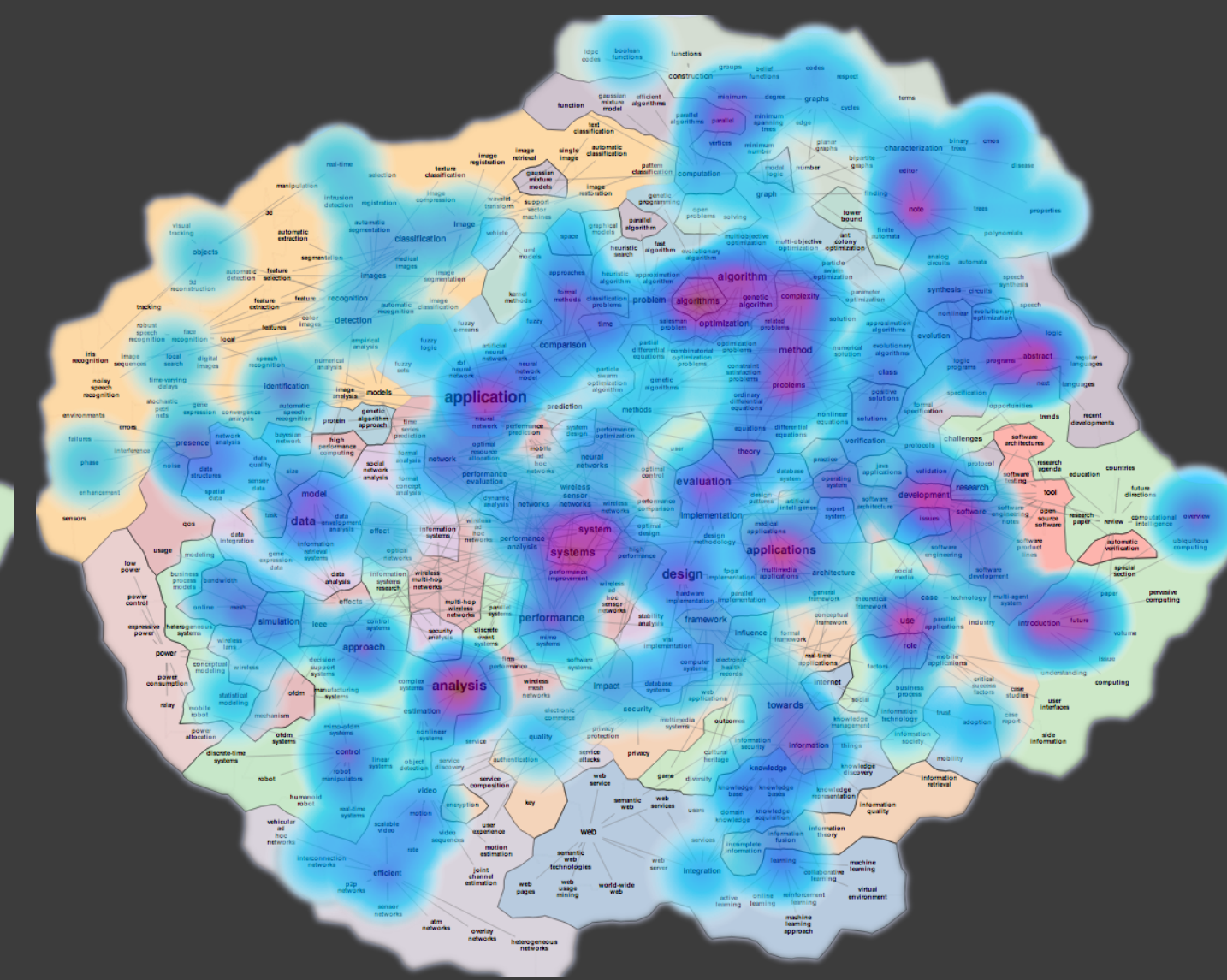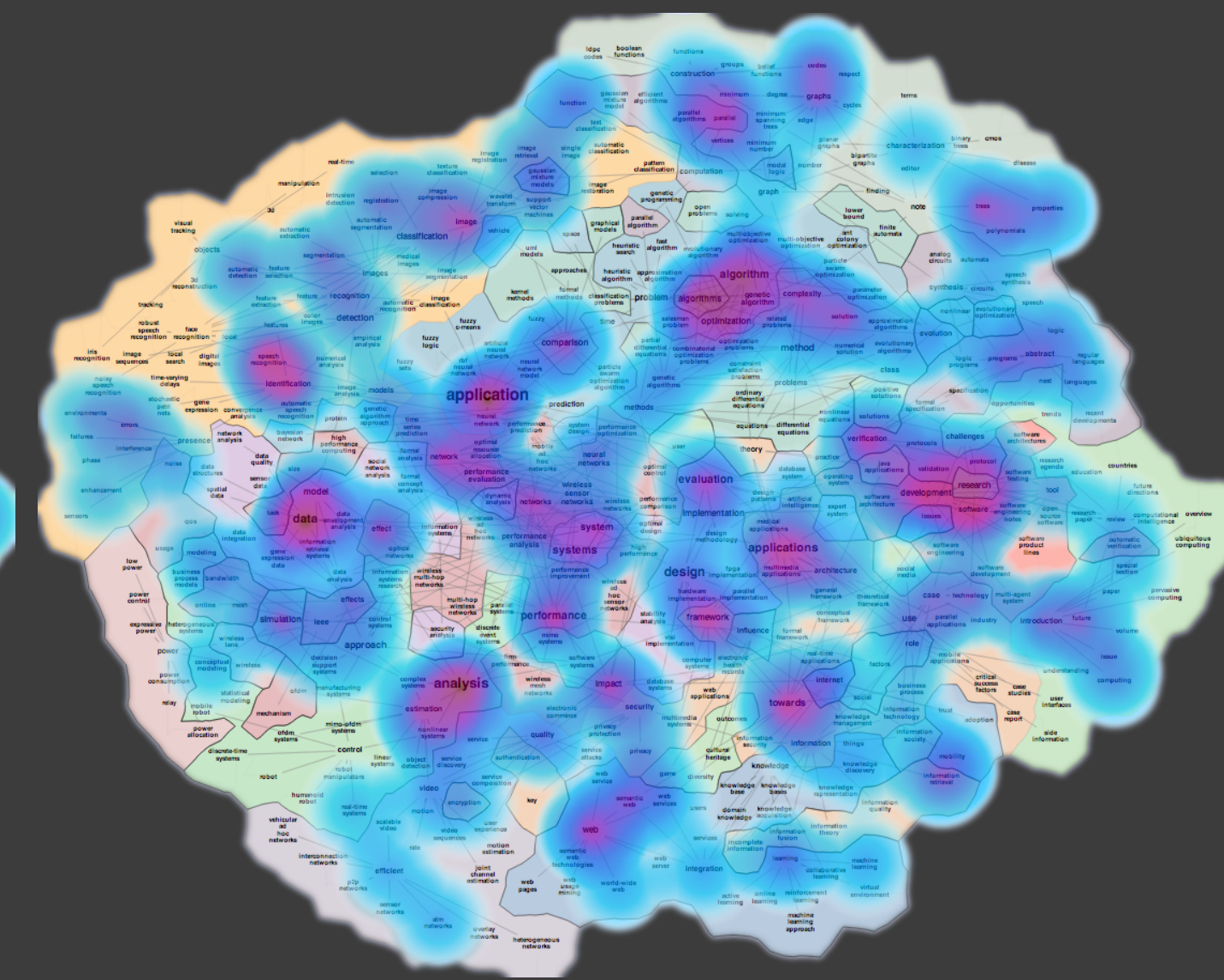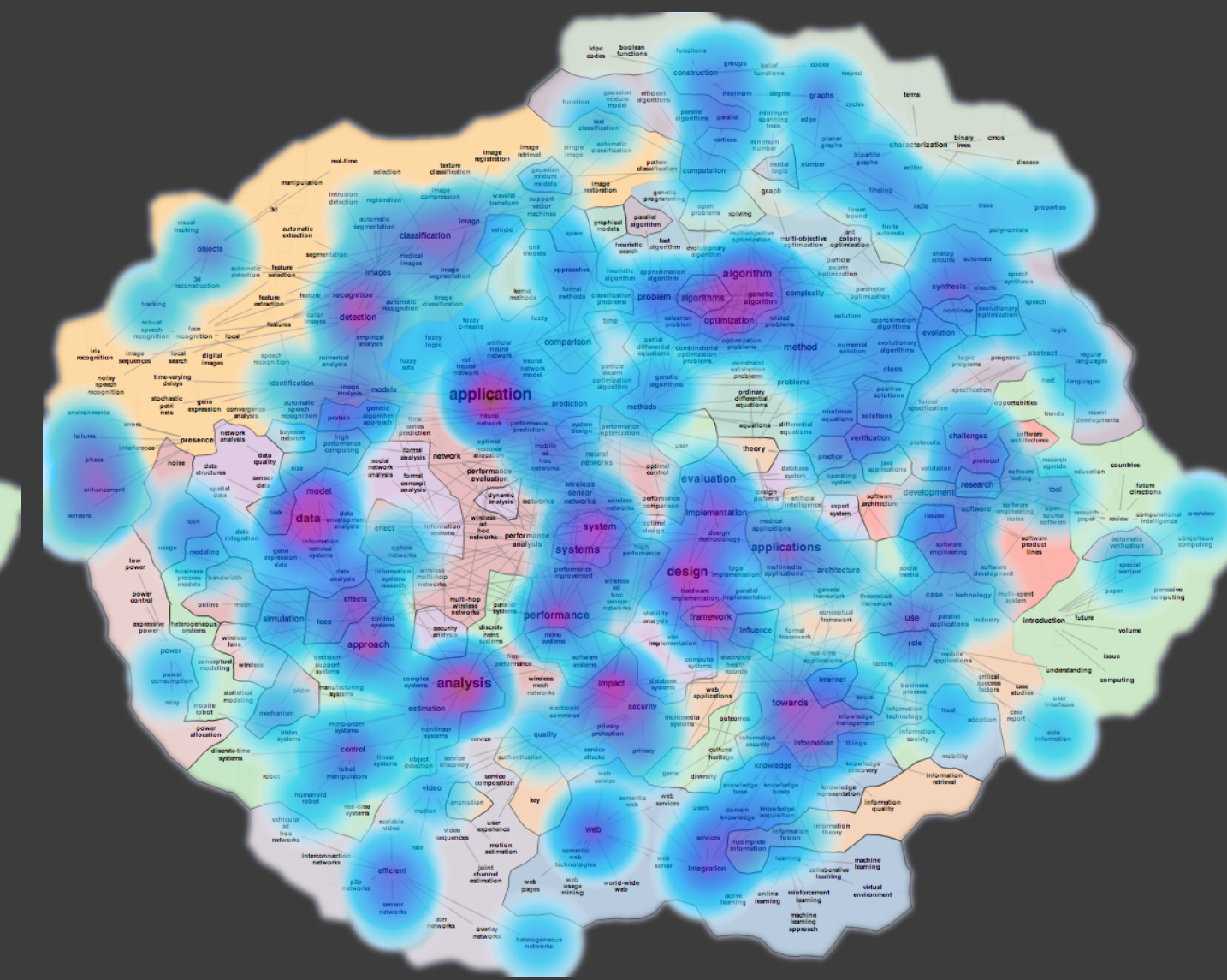## Temporal Heatmaps

1954 - 1963

1964 - 1973

1974 - 1983

1984 - 1993

1994 - 2003

2004 - 2013

The temporal heatmaps above show samples of papers in ten year periods from 1954 to present, providing profiles of computer science research within each time period. In the first few decades, the heatmaps have the highest intensity in the northeast, over algorithmic topics and terms from theoretical computer science. As the field matures and applications develop throughout the next few decades, we see a migration and diffusion over the rest of the map, with growing intensity over applications and systems. In the past two decades, terms in the southwest related to the Internet, data, and performance gain intensity as research in these topics becomes widespread.

Authors: Daniel Fried and Stephen Kobourov. Department of Computer Science, University of Arizona