



### *From words to networks: text-based/ semantic network analysis*

**Jana Diesner** is an Assistant Professor at the iSchool at the University of Illinois at Urbana Champaign. She is also an affiliate at the Department of Computer Science and a Faculty Fellow at the National Center for Supercomputing Applications (NCSA). Jana's research in social computing combines theories and methods from natural language processing, social network analysis and machine learning. In her lab, they develop and advance computational solutions that help people to measure and understand the interplay of information and socio-technical networks. They also bring these solutions into various application context, e.g. in the domain of impact assessment. For more information about Jana's work see <http://people.lis.illinois.edu/~jdiesner/>.

WORKSHOP	FROM WORDS TO NETWORKS: TEXT-BASED/ SEMANTIC NETWORK ANALYSIS
WORKSHOP LENGTH	2 HOURS
INSTRUCTOR	<b>JANA DIESNER</b> , PhD ( <a href="http://people.lis.illinois.edu/~jdiesner/">http://people.lis.illinois.edu/~jdiesner/</a> )

Natural language text data can be a rich source of information about networks. This workshop introduced participants to fundamental theories, concepts and methods for going from words to networks. The focus is on practical, hands-on skills for constructing network data based on text data in an informed, systematic and efficient fashion. Attendants will learn text mining and natural language processing techniques including:

- Entity Extraction, i.e. locating and classifying terms that represent instances of relevant node classes in one-mode and multi-mode networks, e.g. social agents, locations and information.
- Relation Extraction, i.e. linking identified entities into edges based on various criteria, including proximity, syntax and semantics. The extracted networks can be imported into standard SNA tools, incl. R, Gephi, Pajek, UCINET, ORA and visone.
- Analyzing the extracted networks to make meaningful interpretations.
- Text summarization/ distant reading techniques such as topic modeling, corpus statistics, and term weighting techniques.
- Close reading techniques used in the social sciences and humanities, such as text coding.
- Sentiment/ subjectivity analysis, also known as opinion mining.
- Pre-processing techniques such as stemming, n-gram detection and part of speech tagging.
- Measuring the entropy (diversity) of text data.
- Working with various types of text data, e.g. social media data and news data.

### WHO SHOULD ATTEND?

We aim to equip the attendees with the skills and tools needed to apply the covered techniques to their own research questions and text data sets. This is an interdisciplinary and interactive workshop designed to benefit from the participation of attendants from different backgrounds. The material, exercises and mode of delivery are suitable for researchers and practitioners alike. No specific prior knowledge or computational skills are required.

### Why bother?

The functioning and dynamics of real-world networks often involve the production, processing and flow of knowledge and information, which are often represented in the form of text data. In this workshop, participants learn how to a) construct network data based on text data and textual metadata and b) jointly consider text data and network data for analysis. The latter step allows for considering two types

of behavioral information for network analysis, namely social interaction and language use.

Using text data for network analysis is useful for answering questions like: Who is talking to whom, and about what? What social agents or stakeholders are associated with that themes and opinions?

What perception or mental models of certain themes do social agents have? How

can we track the evolution and diffusion of opinions in teams, society and online? Throughout this workshop, we discuss practical applications from various domains for the introduced techniques.

# COINs16

### What to bring to the workshop?

Software: We will use ConText (<http://context.lis.illinois.edu/>). Prior to the workshop, we will send an email to confirmed participants with links and installation instructions for any tools needed. You are invited to bring a laptop to the workshop. If attendants cannot bring a laptop they will still fully benefit from the workshop as we screen-project all live walk-through exercises.

Data: Attendants can work with the sample data that we provide and/ or bring their own data.

### INSTRUCTOR AND CONTACT INFO

Jana Diesner (<http://people.lis.illinois.edu/~jdiesner/>) is an Assistant Professor at the iSchool (a.k.a. Graduate School of Library and Information Science) at the University of Illinois Urbana-Champaign (UIUC). For questions about the workshop email Jana at

✉ [jdiesner@illinois.edu](mailto:jdiesner@illinois.edu)