



ISCAM 2018

Rochester Institute of Technology
Rochester, NY, USA

TUTORIAL:

**Location Extraction and
Georeferencing in Social
Media: Challenges,
Techniques, and Applications**

**15th International Conference on
INFORMATION SYSTEMS FOR CRISIS RESPONSE AND
MANAGEMENT**

“Visualizing Crisis”

Workshops and Doctoral Symposium May 20th, 2018

Conference May 21nd-23th, 2018

Rochester New York - USA
Rochester Institute of Technology (RIT)
<https://iscram2018.rit.edu/>

INTRODUCTION TO THE TUTORIAL

Ability to extract or estimate location in social media content, and perform location-centric analyses offer unique and wide-ranging applications. Examples include disaster management, demographic and socio-cultural studies, and spatiotemporal tracking. For instance, location information is critical to reach and rescue disaster-stricken people and dispatch humanitarian assistance. Consequently, there is a pressing need for better understanding of how people express location information explicitly and implicitly on social media, and in general, develop efficient techniques for geospatial computing that spans all information channels. Additionally, location information enables a variety of individual-level and community-level analyses.

Location extraction and georeferencing methods leverage the user-generated content (textual as well as multimedia data, e.g., images, videos), and users' connectivity (social network analysis). The applications of these methods range from detecting communities and localizing individual texts or detecting users' physical locations. However, due to the challenges posed by social media data some techniques did not work reliably on its informal and ill-formed texts or scaled poorly.

In this tutorial, we present the general problem of georeferencing and location extraction, summarizes the state-of-the-art research, discusses challenges, and provides an overview of our recent research accomplishments in the context of disaster management. Furthermore, we provide a hands-on session to get the audience engaged in the process of location retrieval and extraction from social media using data about a recent hurricane event.

Tutorial Website:

<https://sites.google.com/view/isgram2018geotutorial/home>

TUTORIAL TOPICS

- **Geo-context and Geosemantics:**
 - Geospatial Semantics, Technologies, and Data Interoperability.
- **Location Inference and Geocoding:**
 - Top Down Approaches:
 - Employing knowledge-bases (e.g., knowledge graphs, ontologies, gazetteers)
 - Bottom-up Approaches:
 - Natural Language Processing (NLP) and Statistical Models (e.g., language modeling)
 - Deep/Machine Learning technologies (e.g., LSTMs and CRF for sequence labeling)
 - Social Network Analysis
 - Community detection and graph analysis
- **Applications:**
 - Disaster-specific applications (e.g., Storm surge modeling, traffic-related modeling, rapid disaster response, disaster recovery, emergency management, crisis visualization)

- Socio-cultural and demographic studies (e.g., post-disaster depression studies)
- Location-aware services (e.g. recommender systems, opinion analysis)
- **Hands-on Session:**
 - During this session, we will provide the audience a geotagged/annotated Twitter targeted stream of tweets collected from a disaster event. We will build a supervised and unsupervised location taggers and test the performance of each solution on a different unseen disaster dataset. Additionally, we will highlight the differences, advantages, and disadvantages of using the different techniques. This session is going to provide insights while highlighting the challenges of developing an efficient and scalable solution.

PRESENTERS RECRUITMENT

The presenters are from the Ohio Center of Excellence in Knowledge-enabled Computing (<http://www.knoesis.org/>) ([Kno.e.sis](http://www.knoesis.org/)) at Wright State University that performs world-class research in Web 3.0, machine learning, natural language processing, information retrieval, cloud computing, knowledge representation, databases, data mining, bio/health informatics, cognitive science, and visualization, with applications such as disaster management, e-Sciences with a focus on health-care and life sciences, and defense/intelligence. The well-funded center consists of 13 labs and approximately 75 multidisciplinary researchers, including 15 faculty members and more than 40 Ph.D. students from Computer Science, with competitive federal funding (e.g., NIH, NSF, AFOSR, AFRL) and strong industry collaborations.

TUTORIAL STRUCTURE

The tutorial material will be presented in a 3-hour session.

- Part 1
 - Duration: 60 minutes
 - Description: This is going to motivate the location inference problem and introduce the concepts behind geospatial context and location-based techniques and applications.
 - Topics:
 - Geo-context and Geolocations
 - Employing knowledge-bases
 - Natural Language Processing (NLP) and Statistical Models
- Break for 10 minutes
- Part 2
 - Duration: 60 minutes
 - Description: More detailed technical and state-of-the-art technology for georeferencing and geocoding from social media texts and networks. We will discuss specific applications and future directions in the field.
 - Topics:
 - Deep/Machine Learning technologies

- Community detection and graph analysis
 - Applications
- Break for 10 minutes
- Part 3
 - Duration: 40 minutes
 - Description: Hands-on Session
 - Structure: We will design a hands-on exercise for attendees to experience the power of location-based applications and visualization. This exercise is going to use disaster data from Twitter and would involve the use of knowledge bases from the web.

TUTORIAL PRESENTERS AND COLLABORATORS

	<p>Hussein S. Al-Olimat <i>hussein@knoesis.org</i> Kno.e.sis, Wright State University</p> <p>Hussein is the lead researcher on our NSF funded project, “Social and Physical Sensing Enabled Decision Support for Disaster Management and Response” (HazardSEES). Currently, his research is focused on Geosemantics and Georeferencing. He is the lead on the Location-related work and has authored the research paper titled “Location Name Extraction from Targeted Text Streams using Gazetteer-based Statistical Language Models.”</p>
	<p>Amir Hossein Yazdavar <i>amir@knoesis.org</i> Kno.e.sis, Wright State University</p> <p>Amir is a researcher at Kno.e.sis Center Wright State University. He is broadly interested in applying machine learning (incl. deep learning) and semantic web (incl. creation and use of knowledge graphs for social media analytics). He has a particular interest in studying people’s behavior, psychological status, social connectivity, and community demographics in social media.</p> <p>Related projects:</p> <ul style="list-style-type: none"> ● Modeling Social Behavior Depression ● Twitris 3.0 – Sentiment Analysis for Analyzing Presidential Election ● Gender-Based Violence in 140 Characters or Fewer: A BigData Case Study of Twitter. <p>Recent publication focusing on studying geographical analysis of large-scale user-generated content on social media:</p> <ul style="list-style-type: none"> ● On the Challenges of Sentiment Analysis for Dynamic Events ● Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media



Krishnaprasad Thirunarayan

tkprasad@knoesis.org
Kno.e.sis, Wright State University

[Prof. Thirunarayan](#) has offered several tutorials on “Trust Networks” and “Semantics-empowered Big Data Processing”, and has many publications that span the areas to be covered in this tutorial including “[Knowledge enabled approach to predict the location of twitter users](#)”, “[Understanding City Traffic Dynamics Utilizing Sensor and Textual Observations](#)”, and “[Extracting city traffic events from social streams](#)”. He is also a collaborator on several NSH and NIH funded projects covering social data analytics and applications spanning crisis management to health informatics.



Amit Sheth

amit@knoesis.org
Kno.e.sis, Wright State University

[Prof. Sheth](#) has offered 30+ usually well-attended tutorials at some of the top conferences (e.g., WWW, VLDB, SIGMOD, ICDE, ICWSM). His tutorial with his former PhD student at ICWSM 2013 on a topic relevant to ISCRAM “[Crisis Mapping, Citizen Sensing and Social Media Analytics for Response Coordination](#)” has over 52,000 views. He has led or is leading several significant NSF and NIH funded projects of relevance to this topic area (involving social media and/or crisis management/coordination: eg., “[Social Media Enhanced Organizational Sensemaking in Emergency Response](#),” and “[Hazards SEES: Social and Physical Sensing Enabled Decision Support for Disaster Management and Response](#),” with well over 50 publications). Technology from these projects have been used for coordination during real crisis (e.g., 2014 Kashmir Floods, 2014 Uttarakhand Floods) that have received major international media coverage, for comprehensive simulation for first responders in Dayton area, and for commercialization leading to Cognovi Labs started with the technology licensed from the university research he led. Some of his 55 keynotes covered topics relevant to this tutorial. He is one of the top authors in Computer Science, WWW and Semantic Web.





ISCRAM 2018

Rochester Institute of Technology
Rochester, NY, USA

