

ACL 2014

**Workshop on
Interactive Language Learning, Visualization, and Interfaces**

Proceedings of the Workshop

June 27, 2014
Baltimore, Maryland, USA

©2014 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-941643-15-0

Title Sponsor: Idibon

idibon

language technologies for
a connected world

Introduction

People acquire language through social interaction. Computers learn linguistic models from data, and increasingly, from language-based exchange with people. How do computational linguistic techniques and interactive visualizations work in concert to improve linguistic data processing for humans and computers? How can statistical learning models be best paired with interactive interfaces? How can the increasing quantity of linguistic data be better explored and analyzed? These questions span statistical natural language processing (NLP), human-computer interaction (HCI), and information visualization (Vis), three fields with natural connections but infrequent meetings. Vis and HCI are niches in NLP; Vis and HCI have not fully utilized the statistical techniques developed in NLP. This workshop aims to assemble an interdisciplinary community that promotes collaboration across these fields.

Three themes define this first workshop:

Active, Online, and Interactive Machine Learning Statistical machine learning (ML) has yielded tremendous gains in coverage and robustness for many tasks, but there is a growing sense that additional error reduction might require a fresh look at the human role. Presently, human inputs are often restricted to passive annotation in ML research. However, the fields of ML and HCI are both developing new techniques—such as active learning, incremental/online learning, and crowdsourcing—that attempt to engage people in novel and productive ways. How do we jointly solve the learning questions that have been the domain of NLP and address research topics in HCI such as managing human workers and increasing the quality of their responses?

Language-based user interfaces NLP techniques have entered mainstream use, but the field currently focuses more on building and improving systems and less on understanding how users interact with them in real-world environments. User interface (UI) design decisions can affect the perceived or actual performance of a system. For example, while machine translation (MT) quality improved considerably over the last decade, studies found that human translators disliked MT output for reasons unrelated to translation quality. Many existing systems present sentence-level translations in the absence of relevant context, and disrupt rather than contribute to a translator’s workflow. How do we best integrate learning methods, user behavior understanding, and human-centered design methodology?

Text Visualization and Analysis The quantity and diversity of linguistic corpora is swelling. Recent work on visualizing text data annotated with linguistic structures (e.g., syntactic trees, hypergraphs, and sequences) has produced tools that enable exploration of thematic and recurrence patterns in text. Visual representations built on the outputs of word-level models (e.g., sentiment classifiers, topic models, and continuous word embedding models) now power exploratory analysis of legal documents, political text, and social media content. Beyond adding analytic value, interactive visualization can also reduce the upfront effort needed to set up, configure, and learn a tool, as well as promote adoption. How do we pair appropriate NLP techniques and visualizations to assist both expert and non-technical users, who encounter a growing amount of linguistic data in their professional and everyday lives?

Organizers

Jason Chuang	University of Washington (USA)
Spence Green	Stanford University (USA)
Marti Hearst	UC Berkeley (USA)
Jeffrey Heer	University of Washington (USA)
Philipp Koehn	Johns Hopkins University (USA)

Program Committee

Vicente Alabau	Robin Hill
Cecilia Aragon	Eser Kandogan
Chris Callison-Burch	Frank Keller
Francisco Casacuberta	Katie Kuksenok
Allison Chaney	Laurens van der Maaten
Christopher Collins	Christopher D. Manning
John DeNero	Aditi Muralidharan
Marian Dörk	Burr Settles
Jacob Eisenstein	John Stasko
Jim Herbsleb	Fernanda Viégas
Martin Wattenberg	

Invited Speakers

Chris Culy	Universität Tübingen
Marti Hearst	UC Berkeley
Jimmy Lin	University of Maryland, College Park
Noah Smith	Carnegie Mellon University
Krist Wongsuphasawat	Twitter

Table of Contents

<i>MiTextExplorer: Linked brushing and mutual information for exploratory text data analysis</i> Brendan O'Connor	1
<i>Interactive Learning of Spatial Knowledge for Text to 3D Scene Generation</i> Angel Chang, Manolis Savva and Christopher Manning	14
<i>Dynamic Wordclouds and Vennclouds for Exploratory Data Analysis</i> Glen Coppersmith and Erin Kelly	22
<i>Active Learning with Constrained Topic Model</i> Yi Yang, Shimei Pan, Doug Downey and Kunpeng Zhang	30
<i>GLANCE Visualizes Lexical Phenomena for Language Learning</i> MeiHua Chen, Shih-Ting Huang, Ting-Hui Kao, Hsun-wen Chiu and Tzu-Hsi Yen	34
<i>SPIED: Stanford Pattern based Information Extraction and Diagnostics</i> Sonal Gupta and Christopher Manning	38
<i>Interactive Exploration of Asynchronous Conversations: Applying a User-centered Approach to Design a Visual Text Analytic System</i> Enamul Hoque, Giuseppe Carenini and Shafiq Joty	45
<i>MUCK: A toolkit for extracting and visualizing semantic dimensions of large text collections</i> Rebecca Weiss	53
<i>Design of an Active Learning System with Human Correction for Content Analysis</i> Nancy McCracken, Jasy Suet Yan Liew and Kevin Crowston	59
<i>LDavis: A method for visualizing and interpreting topics</i> Carson Sievert and Kenneth Shirley	63
<i>Hierarchy: Visualization for Hierarchical Topic Models</i> Alison Smith, Timothy Hawes and Meredith Myers	71
<i>Concurrent Visualization of Relationships between Words and Topics in Topic Models</i> Alison Smith, Jason Chuang, Yuening Hu, Jordan Boyd-Graber and Leah Findlater	79

Conference Program

Friday June 27, 2014

8:30 **Opening Remarks**

8:45 **Invited Talk**
Jimmy Lin

Research Papers

9:30 *MiTextExplorer: Linked brushing and mutual information for exploratory text data analysis*
Brendan O'Connor

9:50 *Interactive Learning of Spatial Knowledge for Text to 3D Scene Generation*
Angel Chang, Manolis Savva and Christopher Manning

10:10 *Dynamic Wordclouds and Vennclouds for Exploratory Data Analysis*
Glen Coppersmith and Erin Kelly

10:30 **Coffee Break**

11:00 **Invited Talk**
Noah Smith

11:45 **Invited Talk**
Marti Hearst

12:30 **Lunch Break**

2:00 **Invited Talk**
Chris Culy

2:45 **Interactive Demo Session**

Active Learning with Constrained Topic Model
Yi Yang, Shimei Pan, Doug Downey and Kunpeng Zhang

GLANCE Visualizes Lexical Phenomena for Language Learning
MeiHua Chen, Shih-Ting Huang, Ting-Hui Kao, Hsun-wen Chiu and Tzu-Hsi Yen

SPIED: Stanford Pattern based Information Extraction and Diagnostics
Sonal Gupta and Christopher Manning

Interactive Exploration of Asynchronous Conversations: Applying a User-centered Approach to Design a Visual Text Analytic System
Enamul Hoque, Giuseppe Carenini and Shafiq Joty

Friday June 27, 2014 (continued)

Interactive Demo Session (continued)

MUCK: A toolkit for extracting and visualizing semantic dimensions of large text collections

Rebecca Weiss

Design of an Active Learning System with Human Correction for Content Analysis

Nancy McCracken, Jasy Suet Yan Liew and Kevin Crowston

LDavis: A method for visualizing and interpreting topics

Carson Sievert and Kenneth Shirley

Hiearchie: Visualization for Hierarchical Topic Models

Alison Smith, Timothy Hawes and Meredith Myers

Concurrent Visualization of Relationships between Words and Topics in Topic Models

Alison Smith, Jason Chuang, Yuening Hu, Jordan Boyd-Graber and Leah Findlater

4:00

Invited Talk

Krist Wongsuphasawat

4:45

Discussion and Closing Remarks