

CNS 2016

**2nd Workshop on  
Computing News  
Storylines**

**Proceedings of the Workshop**

November 5, 2016  
Austin, Texas, USA

©2016 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Curran Associates  
57 Morehouse Lane  
Red Hook, New York 12571  
USA  
Tel: +1-845-758-0400  
Fax: +1-845-758-2633  
[curran@proceedings.com](mailto:curran@proceedings.com)

ISBN 978-1-945626-27-2

## Introduction

This volume contains the proceedings of the 2nd Workshop on Computing News Storylines (CNewsStory 2016) held in conjunction with the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016) in Austin, Texas, USA, on 5 November 2016. Narratives are at the heart of information sharing. Ever since people began to share their experiences, they have connected them to form narratives. The study of storytelling and the field of literary theory called narratology have developed complex frameworks and models related to various aspects of narrative such as plots structures, narrative embeddings, characters’ perspectives, reader response, point of view, narrative voice, narrative goals, and many others. These notions from narratology have been applied mainly in Artificial Intelligence and to model formal semantic approaches to narratives (e.g. Plot Units developed by Lehnert (1981)). In recent years, computational narratology has qualified as an autonomous field of study and research. Narrative has been the focus of a number of workshops and conferences (AAAI Symposia, Interactive Storytelling Conference (ICIDS), Computational Models of Narrative). Furthermore, reference annotation schemes for narratives have been proposed (NarrativeML by Mani (2013)).

The workshop aims to bring together researchers from different communities working on representing and extracting narrative structures in news, a text genre popular in NLP research but which has received little attention in research into narrative structure, representation and analysis.

Current advances in NLP technology have made it possible to look beyond scenario-driven, atomic extraction of events from single documents and work towards extracting story structures from multiple documents, while these documents are published over time as news streams. Policy makers and information specialists are increasingly in need of tools that support them in finding salient stories in large amounts of information to more effectively implement policies, monitor actions of “big players” in society and check facts. Their tasks often revolve around reconstructing cases either with respect to specific entities (e.g. person or organisations) or events (e.g. the 2016 presidential elections). Storylines represent explanatory schemas that enable us to make better selections of relevant information but also projections for the future. They constitute a huge potential for exploiting news data in an innovative way.

Of the 14 submissions we received, 8 were accepted that touch upon different aspects of narrative research in news. Three contributions describe approaches to detect storylines either from news (Brüggermann et al.), from Tweets (Krishnan and Eisenstein), or from news but with metadata added via Twitter (Poghosyan and Ifrim). Besides detecting storylines, different aspects of storylines such as diegesis and point of view are also addressed (Eisenberg and Finlayson). The second topic that is addressed is annotation and representation of storylines (Caselli and Vossen and O’Gorman et al.). Related to this is the analysis of the distribution of narrative schemas in a corpus, which may help further the discussion on corpus creation (Simonson and Davis). Finally, ideas on how to put storylines to use in a newsroom are discussed in Caswell.

We would like to thank the members of the Program Committee for their timely reviews and the authors for their contributions.



**Organizers:**

Tommaso Caselli, Vrije Universiteit Amsterdam  
Ben Miller, Georgia State University  
Marieke van Erp, Vrije Universiteit Amsterdam  
Piek Vossen, Vrije Universiteit Amsterdam  
David Caswell, Reynolds Journalism Institute & University of Missouri

**Program Committee:**

Alexandra Balahur, European Commission Joint Research Centre, Ispra, Italy  
Sabine Bergler, Computer Science, Columbia University, Canada  
Matje van de Camp, De Taalmonsters, The Netherlands  
Reginald Chua, Thomson Reuters, USA  
Leon Derczynski, University of Sheffield, UK  
Mark Finlayson, Florida International University, USA  
Martijn Kleppe, Koninklijke Bibliotheek, Den Haag, The Netherlands  
Bernardo Magnini, HLT-FBK, Italy  
Roser Morante, Vrije Universiteit Amsterdam, The Netherlands  
Nasrin Mostafazadeh, University of Rochester, USA  
Vivi Nastase, Institut fur Computerlinguistik, University of Heidelberg, Germany  
Silvia Pareti, Google Inc. & University of Edinburgh  
Octavian Popescu, IBM Watson Research Center, USA  
Ellen Riloff, University of Utah, USA  
Jonathan Stray, Columbia University, USA  
Xavier Tannier, LIMSI-CNRS, France  
Marc Verhagen, Brandeis University, USA

**Invited Speaker:**

Eduard Hovy - Bridging the Gap between Event Macro-structures and Event Micro-structures  
Carnegie Mellon University



## Table of Contents

<i>Computable News Ecosystems: Roles for Humans and Machines</i> David Caswell .....	1
<i>Storyline detection and tracking using Dynamic Latent Dirichlet Allocation</i> Daniel Bruggemann, Yannik Hermey, Carsten Orth, Darius Schneider, Stefan Selzer and Gerasi- mos Spanakis .....	9
<i>Real-time News Story Detection and Tracking with Hashtags</i> Gevorg Poghosyan and Georgiana Ifrim .....	20
<i>Nonparametric Bayesian Storyline Detection from Microtexts</i> Vinodh Krishnan and Jacob Eisenstein .....	30
<i>Automatic Identification of Narrative Diegesis and Point of View</i> Joshua Eisenberg and Mark Finlayson .....	36
<i>Richer Event Description: Integrating event coreference with temporal, causal and bridging annotation</i> Tim O’Gorman, Kristin Wright-Bettner and Martha Palmer .....	47
<i>NASTEAs: Investigating Narrative Schemas through Annotated Entities</i> Dan Simonson and Anthony Davis .....	57
<i>The Storyline Annotation and Representation Scheme (StaR): A Proposal</i> Tommaso Caselli and Piek Vossen .....	67





# Workshop Program

**Saturday, November 5, 2016**

**09:00–10:30 Session 1: Opening Session**

**09:00–09:10** *Welcome and Opening Remarks*

09:10–10:10 *Bridging the Gap between Event Macro-structures and Event Micro-structures*  
Ed Hovy, Language Technologies Institute, CMU

**10:30–11:00** *Coffee Break*

**11:00–12:30 Session 2: Morning Session**

11:00–11:25 *Computable News Ecosystems: Roles for Humans and Machines*  
David Caswell

11:25–11:50 *Storyline detection and tracking using Dynamic Latent Dirichlet Allocation*  
Daniel Bruggemann, Yannik Hermeij, Carsten Orth, Darius Schneider, Stefan Selzer and Gerasimos Spanakis

11:50–12:15 *Real-time News Story Detection and Tracking with Hashtags*  
Gevorg Poghosyan and Georgiana Ifrim

12:15–12:30 *Nonparametric Bayesian Storyline Detection from Microtexts*  
Vinodh Krishnan and Jacob Eisenstein

**12:30–14:00** *Lunch*

**Saturday, November 5, 2016 (continued)**

**14:00–15:30 Session 3: Afternoon Session**

14:00–14:25 *Automatic Identification of Narrative Diegesis and Point of View*  
Joshua Eisenberg and Mark Finlayson

14:25–14:50 *Richer Event Description: Integrating event coreference with temporal, causal and bridging annotation*  
Tim O’Gorman, Kristin Wright-Bettner and Martha Palmer

14:50–15:15 *NASTEIA: Investigating Narrative Schemas through Annotated Entities*  
Dan Simonson and Anthony Davis

15:15–15:30 *The Storyline Annotation and Representation Scheme (StaR): A Proposal*  
Tommaso Caselli and Piek Vossen

**16:00–17:00 Session 4: Discussion and Closing Remarks**