INLG 2016

The 9th International
Natural Language Generation conference

Proceedings of the Conference

September 5-8, 2016
Edinburgh, UK
Introduction

Welcome to the 9th International Natural Language Generation Conference (INLG 2016)! INLG is the biennial meeting of the ACL Special Interest Group on Natural Language Generation (SIGGEN). The INLG conference provides the premier forum for the discussion, dissemination, and archiving of research and results in the field of Natural Language Generation (NLG). Previous INLG conferences have been held in Ireland, the USA, Australia, the UK, and Israel. Prior to 2000, INLG meetings were held as international workshops with a history stretching back to 1983. In 2016, INLG is organized in Edinburgh, UK.

The INLG 2016 program includes presentations of substantial, original, and previously unpublished results on all topics related to natural language generation. This year, INLG received 65 submissions (29 full papers, 30 short papers and 6 demo proposals) from 12 different countries. 16 submissions were accepted as full papers (10 presented orally, 6 as posters), 20 as short papers (4 presented orally, and 16 as posters), and 3 demos. In addition, INLG 2016 includes an invited talk by Yejin Choi (University of Washington) and Vera Demberg (Saarland University).

The organizing committee would like to offer their thanks to our invited speakers for agreeing to join us and to the authors of all submitted papers. We have also received sponsorship from Arria and WebNLG, for which we are extremely grateful. Finally, we would like to welcome you to Edinburgh and hope that you have an enjoyable and inspiring visit!

Amy Isard, Verena Rieser and Dimitra Gkatzia
INLG 2016 Co-Chairs
Organizers:
Amy Isard
Verena Rieser
Dimitra Gkatzia
Jekaterina Novikova
Simon Keizer

Program Committee:
Anja Belz
Aoife Cahill
Nina Dethlefs
Katja Filippova
Claire Gardent
Albert Gatt
Pablo Gervás
Alistair Knott
Ioannis Konstas
Emiel Krahmer
Guy Lapalme
David McDonald
Cecile Paris
Paul Piwek
Ehud Reiter
Kristina Striegnitz
Mariët Theune
Ielka van der Sluis
Keith Vander Linden
Leo Wanner
Michael White
Sandra Williams
Helen Hastie
Barbara Di Eugenio
Vera Demberg
Laura Perez-Beltrachini
Saad Mahamood
Andreas Vlachos
Kees van Deemter
David Howcroft
Yejin Choi
Kathy McKeown
Diana Bental
François Portet
Simon Mille
Cyril Labbe
John Kelleher
Sander Wubben
Somayajulu Sripada
Nikolaos Engonopoulos
Amanda Stent
David Schlangen
Sina Zarrieß

Invited Speakers:

Yejin Choi (University of Washington)
Vera Demberg (Saarland University)
# Table of Contents

*Summarising News Stories for Children*  
Iain Macdonald and Advaith Siddharthan ................................................. 1

*Discourse-Driven Narrative Generation With Bipartite Planning*  
David Winer and R. Michael Young .......................................................... 11

*Generating English from Abstract Meaning Representations*  
Nima Pourdamghani, Kevin Knight and Ulf Hermjakob ................................ 21

*Generating summaries of hospitalizations: A new metric to assess the complexity of medical terms and their definitions*  
Sabita Acharya, Barbara Di Eugenio, Andrew D Boyd, Karen Dunn Lopez, Richard Cameron and Gail M Keenan ............................................................... 26

*Designing Algorithms for Referring with Proper Names*  
Kees van Deemter ......................................................................................... 31

*When to Plummet and When to Soar: Corpus Based Verb Selection for Natural Language Generation*  
Charese Smiley, Vassilis Plachouras, Frank Schilder, Hiroko Bretz, Jochen Leidner and Dezhao Song ......................................................................................... 36

*Sketch-to-Text Generation: Toward Contextual, Creative, and Coherent Composition*  
Yejin Choi .................................................................................................... 40

*Abstractive Compression of Captions with Attentive Recurrent Neural Networks*  
Sander Wubben, Emiel Krahmer, Antal van den Bosch and Suzan Verberne ...... 41

*Infusing NLU into Automatic Question Generation*  
Karen Mazidi and Paul Tarau ........................................................................ 51

*Automatic label generation for news comment clusters*  
Ahmet Aker, Monica Paramita, Emina Kurtic, Adam Funk, Emma Barker, Mark Hepple and Rob Gaizauskas ................................................................. 61

*Improving Fluency in Narrative Text Generation With Grammatical Transformations and Probabilistic Parsing*  
Emily Ahn, Fabrizio Morbini and Andrew Gordon ......................................... 70

*The Multilingual Affective Soccer Corpus (MASC): Compiling a biased parallel corpus on soccer reportage in English, German and Dutch*  
Nadine Braun, Martijn Goudbeek and Emiel Krahmer .................................. 74

*Challenges of Argument Mining: Generating an Argument Synthesis based on the Qualia Structure*  
Patrick Saint-Dizier ...................................................................................... 79
Tense and Aspect in Runyankore Using a Context-Free Grammar
Joan Byamugisha, C. Maria Keet and Brian DeRenzi ........................................ 84

Task demands and individual variation in referring expressions
Adriana Baltaretu and Thiago Castro Ferreira .................................................. 89

Category-Driven Content Selection
Rania Mohammed, Laura Perez-Beltrachini and Claire Gardent ........................ 94

Evaluative Pattern Extraction for Automated Text Generation
Chia-Chen Lee and Shu-Kai HSIEH ................................................................. 99

Statistics-Based Lexical Choice for NLG from Quantitative Information
Xiao Li, Kees van Deemter and Chenghua Lin .................................................. 104

Incremental Generation of Visually Grounded Language in Situated Dialogue (demonstration system)
Yanchao Yu, Arash Eshghi and Oliver Lemon ................................................. 109

Unsupervised Sentence Simplification Using Deep Semantics
Shashi Narayan and Claire Gardent ............................................................... 111

Enabling text readability awareness during the micro planning phase of NLG applications
Priscilla Moraes, Kathleen Mccoy and Sandra Carberry .................................. 121

How can we adapt generation to the user’s cognitive load?
Vera Demberg ............................................................................................ 132

Selecting Domain-Specific Concepts for Question Generation With Lightly-Supervised Methods
Yiping Jin and Phu Le .................................................................................. 133

Statistical Natural Language Generation from Tabular Non-textual Data
Joy Mahapatra, Sudip Kumar Naskar and Sivaji Bandyopadhyay ..................... 143

Paraphrase Generation from Latent-Variable PCFGs for Semantic Parsing
Shashi Narayan, Siva Reddy and Shay B. Cohen .............................................. 153

The WebNLG Challenge: Generating Text from DBPedia Data
Emilie Colin, Claire Gardent, Yassine Mrabet, Shashi Narayan and Laura Perez-Beltrachini .. 163

The aNALoGuE Challenge: Non Aligned Language GEneration
Jekaterina Novikova and Verena Rieser ......................................................... 168

A Challenge Proposal for Narrative Generation Using CNLs
Eugenio Concepción, Gonzalo Méndez, Pablo Gervás and Carlos León ............. 171

On the verbalization patterns of part-whole relations in isiZulu
C. Maria Keet and Langa Khumalo ............................................................... 174

SimpleNLG-IT: adapting SimpleNLG to Italian
Alessandro Mazzei, Cristina Battaglini and Cristina Bosco ................................ 184
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Don’t Mention the Shoe! A Learning to Rank Approach to Content Selection for Image Description Generation</td>
<td>Josiah Wang and Robert Gaizauskas</td>
<td>193</td>
</tr>
<tr>
<td>Good Automatic Authentication Question Generation</td>
<td>Simon Woo, Zuyao Li and Jelena Mirkovic</td>
<td>203</td>
</tr>
<tr>
<td>Automatic Generation of Student Report Cards</td>
<td>Amy Isard and Jeremy Knox</td>
<td>207</td>
</tr>
<tr>
<td>Collecting Reliable Human Judgements on Machine-Generated Language: The Case of the QG-STEC Data</td>
<td>Keith Godwin and Paul Piwek</td>
<td>212</td>
</tr>
<tr>
<td>Ranking Automatically Generated Questions Using Common Human Queries</td>
<td>Yllias Chali and Sina Golestanirad</td>
<td>217</td>
</tr>
<tr>
<td>Towards proper name generation: a corpus analysis</td>
<td>Thiago Castro Ferreira, Sander Wubben and Emiel Krahmer</td>
<td>222</td>
</tr>
<tr>
<td>An Analysis of the Ability of Statistical Language Models to Capture the Structural Properties of Language</td>
<td>Aneiss Ghodsi and John DeNerano</td>
<td>227</td>
</tr>
<tr>
<td>Enhancing PTB Universal Dependencies for Grammar-Based Surface Realization</td>
<td>David L. King and Michael White</td>
<td>232</td>
</tr>
<tr>
<td>Effect of Data Annotation, Feature Selection and Model Choice on Spatial Description Generation in French</td>
<td>Anja Belz, Adrian Muscat, Brandon Birmingham, Jessie Levacher, Julie Pain and Adam Quinquenel</td>
<td>237</td>
</tr>
<tr>
<td>QGASP: a Framework for Question Generation Based on Different Levels of Linguistic Information</td>
<td>Hugo Patinho Rodrigues, Luisa Coheur and Eric Nyberg</td>
<td>242</td>
</tr>
<tr>
<td>Automatic Reports from Spreadsheets: Data Analysis for the Rest of Us</td>
<td>Pablo Duboue</td>
<td>244</td>
</tr>
<tr>
<td>Towards Generating Colour Terms for Referents in Photographs: Prefer the Expected or the Unexpected?</td>
<td>Sina Zarrieß and David Schlangen</td>
<td>246</td>
</tr>
<tr>
<td>Absolute and Relative Properties in Geographic Referring Expressions</td>
<td>Rodrigo de Oliveira, Somayajulu Sripada and Ehud Reiter</td>
<td>256</td>
</tr>
<tr>
<td>Crowd-sourcing NLG Data: Pictures Elicit Better Data</td>
<td>Jekaterina Novikova, Oliver Lemon and Verena Rieser</td>
<td>265</td>
</tr>
</tbody>
</table>
Conference Program

Wednesday, September 7, 2016

8:00–9:15  Registration

9:15–9:30  Welcome / Opening remarks

9:30–10:30  Oral Session: Summarisation / Narrative

09:30–10:00  Summarising News Stories for Children
Iain Macdonald and Advaith Siddharthan

10:00–10:30  Discourse-Driven Narrative Generation With Bipartite Planning
David Winer and R. Michael Young

10:30–11:00  Coffee break

11:00–12:20  Oral Session: short papers

11:00–11:20  Generating English from Abstract Meaning Representations
Nima Pourdamghani, Kevin Knight and Ulf Hermjakob

11:20–11:40  Generating summaries of hospitalizations: A new metric to assess the complexity of medical terms and their definitions
Sabita Acharya, Barbara Di Eugenio, Andrew D Boyd, Karen Dunn Lopez, Richard Cameron and Gail M Keenan

11:40–12:00  Designing Algorithms for Referring with Proper Names
Kees van Deemter

12:00–12:20  When to Plummet and When to Soar: Corpus Based Verb Selection for Natural Language Generation
Charase Smiley, Vassilis Plachouras, Frank Schilder, Hiroko Bretz, Jochen Leidner and Dezhao Song
Wednesday, September 7, 2016 (continued)

12:20–13:30  Lunch

13:30–14:30  Invited Talk
Yejin Choi, session chair: Amy Isard

–13:30  Sketch-to-Text Generation: Toward Contextual, Creative, and Coherent Composition
Yejin Choi

14:30–16:00  Posters and Demos with coffee

15:00–15:30  Abstractive Compression of Captions with Attentive Recurrent Neural Networks
Sander Wubben, Emiel Krahmer, Antal van den Bosch and Suzan Verberne

15:00–15:30  Infusing NLU into Automatic Question Generation
Karen Mazidi and Paul Tarau

15:00–15:30  Automatic label generation for news comment clusters
Ahmet Aker, Monica Paramita, Emina Kurtic, Adam Funk, Emma Barker, Mark Hepple and Rob Gaizauskas

15:00–15:30  Improving Fluency in Narrative Text Generation With Grammatical Transformations and Probabilistic Parsing
Emily Ahn, Fabrizio Morbini and Andrew Gordon

15:00–15:30  The Multilingual Affective Soccer Corpus (MASC): Compiling a biased parallel corpus on soccer reportage in English, German and Dutch
Nadine Braun, Martijn Goudbeek and Emiel Krahmer

15:00–15:30  Challenges of Argument Mining: Generating an Argument Synthesis based on the Qualia Structure
Patrick Saint-Dizier

15:00–15:30  Tense and Aspect in Runyankore Using a Context-Free Grammar
Joan Byamugisha, C. Maria Keet and Brian DeRenzi

15:00–15:30  Task demands and individual variation in referring expressions
Adriana Baltaretu and Thiago Castro Ferreira

15:00–15:30  Category-Driven Content Selection
Rania Mohammed, Laura Perez-Beltrachini and Claire Gardent
Wednesday, September 7, 2016 (continued)

15:00–15:30  Evaluative Pattern Extraction for Automated Text Generation  
Chia-Chen Lee and Shu-Kai HSIEH

15:00–15:30  Statistics-Based Lexical Choice for NLG from Quantitative Information  
Xiao Li, Kees van Deemter and Chenghua Lin

15:00–15:30  Incremental Generation of Visually Grounded Language in Situated Dialogue  
(demonstration system)  
Yanchao Yu, Arash Eshghi and Oliver Lemon

16:00–17:00  Oral Session: Simplification and Readability

16:00–16:30  Unsupervised Sentence Simplification Using Deep Semantics  
Shashi Narayan and Claire Gardent

16:30–17:00  Enabling text readability awareness during the micro planning phase of NLG applications  
Priscilla Moraes, Kathleen Mccoy and Sandra Carberry

19:00  Drinks, dinner, ceilidh at Ghillie Dhu - Announcement of best paper

Thursday, September 8, 2016

09:00–10:00  Invited Talk  
Vera Demberg, session chair: Verena Rieser

9:00–10:00  How can we adapt generation to the user’s cognitive load?  
Vera Demberg
Thursday, September 8, 2016 (continued)

10:00–10:30  Oral Session, Statistical NLG 1

10:00–10:30  Selecting Domain-Specific Concepts for Question Generation With Lightly-Supervised Methods
Yiping Jin and Phu Le

10:30–11:00  Coffee break

11:00–12:00  Oral Session, Statistical NLG 2

11:00–11:30  Statistical Natural Language Generation from Tabular Non-textual Data
Joy Mahapatra, Sudip Kumar Naskar and Sivaji Bandyopadhyay

11:30–12:00  Paraphrase Generation from Latent-Variable PCFGs for Semantic Parsing
Shashi Narayan, Siva Reddy and Shay B. Cohen

12:00–12:45  GenChal

12:00–12:15  The WebNLG Challenge: Generating Text from DBPedia Data
Emilie Colin, Claire Gardent, Yassine Mrabet, Shashi Narayan and Laura Perez-Beltrachini

12:15–12:30  The aNALoGuE Challenge: Non Aligned Language GEneration
Jekaterina Novikova and Verena Rieser

12:30–12:45  A Challenge Proposal for Narrative Generation Using CNLs
Eugenio Concepción, Gonzalo Méndez, Pablo Gervás and Carlos León
Thursday, September 8, 2016 (continued)

12:45–13:45 Lunch

13:45–15:15 Posters and demos with coffee

14:45–15:15 On the verbalization patterns of part-whole relations in isiZulu
C. Maria Keet and Langa Khumalo

14:45–15:15 SimpleNLG-IT: adapting SimpleNLG to Italian
Alessandro Mazzei, Cristina Battaglino and Cristina Bosco

14:45–15:15 Don’t Mention the Shoe! A Learning to Rank Approach to Content Selection for Image Description Generation
Josiah Wang and Robert Gaizauskas

14:45–15:15 Good Automatic Authentication Question Generation
Simon Woo, Zuyao Li and Jelena Mirkovic

14:45–15:15 Automatic Generation of Student Report Cards
Amy Isard and Jeremy Knox

14:45–15:15 Collecting Reliable Human Judgements on Machine-Generated Language: The Case of the QG-STEC Data
Keith Godwin and Paul Piwek

14:45–15:15 Ranking Automatically Generated Questions Using Common Human Queries
Yllias Chali and Sina Golestanirad

14:45–15:15 Towards proper name generation: a corpus analysis
Thiago Castro Ferreira, Sander Wubben and Emiel Krahmer

14:45–15:15 An Analysis of the Ability of Statistical Language Models to Capture the Structural Properties of Language
Aneiss Ghodsi and John DeNero

14:45–15:15 Enhancing PTB Universal Dependencies for Grammar-Based Surface Realization
David L. King and Michael White

14:45–15:15 Effect of Data Annotation, Feature Selection and Model Choice on Spatial Description Generation in French
Anja Belz, Adrian Muscat, Brandon Birmingham, Jessie Levacher, Julie Pain and Adam Quinquenel
Thursday, September 8, 2016 (continued)

14:45–15:15  **QGASP: a Framework for Question Generation Based on Different Levels of Linguistic Information**  
Hugo Patinho Rodrigues, Luisa Coheur and Eric Nyberg

14:45–15:15  **Automatic Reports from Spreadsheets: Data Analysis for the Rest of Us**  
Pablo Duboue

**15:15–16:45 Oral Session, REG and Vision**

16:15–16:45  **Towards Generating Colour Terms for Referents in Photographs: Prefer the Expected or the Unexpected?**  
Sina Zarrieß and David Schlangen

15:15–15:45  **Absolute and Relative Properties in Geographic Referring Expressions**  
Rodrigo de Oliveira, Somayajulu Sripada and Ehud Reiter

15:45–16:15  **Crowd-sourcing NLG Data: Pictures Elicit Better Data.**  
Jekaterina Novikova, Oliver Lemon and Verena Rieser

**Closing Remarks**