

# Daniel Deutsch

ddeutsch@seas.upenn.edu

<https://danieldeutsch.github.io/>

## Education

- 2017 – Present    **University of Pennsylvania**  
PhD in Computer and Information Science  
Advisor: Dan Roth  
Thesis Title: *Question Answering-Based Methods for Developing and Evaluating Summarization Systems*
- 2013 – 2014    **Johns Hopkins University**  
MSE in Computer Science  
Advisor: David Yarowsky
- 2010 – 2013    **Johns Hopkins University**  
BS in Computer Science

## Work Experience

- 2014 – 2017    **Microsoft, Bing (Applied Scientist II)**  
*Entity Linking Team*  
Worked on the core algorithm for Bing's entity linker. Developed new features based on the web index, query logs, and other sources. Redesigned pipeline to use a distributed architecture.
- Question-Answering Team*  
Improved the coverage of Bing's Factoid Answer via mining high-precision query patterns.

## Research Interests

Text generation, automatic summarization, summarization evaluation, constrained inference, reproducibility

## Publications & Pre-Prints

- arXiv *Question-Based Salient Span Selection for More Controllable Text Summarization.* **Daniel Deutsch** and Dan Roth. [Link](#).
- TACL 2021 *A Statistical Analysis of Summarization Evaluation Metrics Using Resampling Methods.* **Daniel Deutsch**, Rotem Dror, and Dan Roth. [Link](#).
- CoNLL 2021 *Understanding the Extent to which Content Quality Metrics Measure the Information Quality of Summaries.* **Daniel Deutsch** and Dan Roth. [Link](#).
- TACL 2021 *Towards Question-Answering as an Automatic Metric for Evaluating the Content Quality of a Summary.* **Daniel Deutsch**, Tania Bedrax-Weiss, and Dan Roth. [Link](#).
- NLP-OSS 2020 *SacreROUGE: An Open-Source Library for Using and Developing Summarization Evaluation Metrics.* **Daniel Deutsch** and Dan Roth. [Link](#).
- COLING 2020 *Is Killed More Significant than Fled? A Contextual Model for Salient Event Detection.* Disha Jindal, **Daniel Deutsch**, and Dan Roth. [Link](#).
- CoNLL 2019 *A General-Purpose Algorithm for Constrained Sequential Inference.* **Daniel Deutsch**,\* Shyam Upadhyay,\* and Dan Roth. \*Equal contribution. [Link](#).
- EMNLP 2019 *Summary Cloze: A New Task for Content Selection in Topic-Focused Summarization.* **Daniel Deutsch** and Dan Roth. [Link](#).
- ACL 2018 *A Distributional and Orthographic Aggregation Model for English Derivational Morphology.* **Daniel Deutsch**,\* John Hewitt,\* and Dan Roth. \*Equal contribution. [Link](#).

## Software Packages

**Creator of Repro**, a library for improving the usability and reproducibility of research code via Docker. <https://github.com/danieldeutsch/repro>

**Creator of SacreROUGE**, a library for using and developing summarization evaluation metrics. <https://github.com/danieldeutsch/sacrerouge>

## Professional Activities

- NAACL 2022 Reproducibility Track Co-Chair
- GEM 2021 – Present Contributor to the human evaluation (2021, 2022) and automatic evaluation (2022) teams for the Generation, Evaluation, and Metrics (GEM) workshop

- MASC 2020      Program Committee Member for the Mid-Atlantic Student Colloquium (MASC)
- 2018 – 2019      Organizer of the Computational Linguistics seminar at the University of Pennsylvania
- 2018 – Present      Maintainer of the NLP website at the University of Pennsylvania

## Internships

- Summer 2020      **Dataminr (Research Scientist Intern)**  
*Hosts: Joel Tetreault and Ke Zhang*
- Summer 2019      **Amazon (Applied Scientist Intern)**  
*Hosts: Lluís Màrquez and Hugo Zaragoza*