# Sebastian Gehrmann

✉ gehrmann@seas.harvard.edu
🖥 sebastiangehrmann.com
in sebastiangehrmann
🐦 SebGehr

## Education

2019 **Ph.D. Computer Science**,
*Harvard Paul A. Johnson School of Engineering and Applied Sciences*.
Advisors: Barbara Grosz, Alexander (Sasha) Rush.
Thesis: Human-AI Collaboration for Natural Language Generation with Interpretable Neural Networks. [link]

2018 **S.M. Computer Science**,
*Harvard Paul A. Johnson School of Engineering and Applied Sciences*, GPA: 3.9.

2015 **B.Sc. Computer Science and Business Administration**,
*Nordakademie University of Applied Sciences*, Germany, ECTS Grade A.

2013 **Visiting Undergraduate Student**, *Harvard University*.

## Work Experience

2020-present **Research Scientist**, *Google AI Language*, New York City, NY.
I am working on improving natural language generation systems and make them interpretable, controllable, faithful, fair, and robust.

Summer 2018, **Document Systems and NLP Research Intern**, *Adobe*, San Jose, CA.
2019 Supervised by Franck Dernoncourt, I studied how generated text can help humans understand and share information in long documents.

Fall 2018 **Graduate Research Intern**, *IBM Research*, Cambridge, MA.
In collaboration with Hendrik Strobelt, I worked on interactive inference for text generation models and analysis tools for controllable generation of protein sequences.

Summer 2016 **Collaborator**, *Fondazione Bruno Kessler*, Trento, Italy.
Supervised by Massimo Zancanaro and Oliviero Stock, I developed an automatic facilitator to support teenagers with conflicting backgrounds to create shared narratives. We evaluated this facilitator in a human-subject study between Arabic and Israeli native speakers.

2014–2017 **Co-Founder**, *Digital Natives*, Frankfurt, Germany.
I was the technical leader of a small media agency that developed web sites, improved SEO and handled social media for small and medium sized businesses. Notable clients include Girardet Verlag and No-Stop Consulting.

2011–2015 **Dual Student**, *Eppendorf AG*, Hamburg, Germany.
To finance my undergraduate studies, I held a full time position in cooperation with Nordakademie. My responsibilities included the project management for a new customer relationship management system. Among other tasks I created and led training sessions for service technicians across Europe and designed algorithms for prioritizing service tickets and deploying technicians.

# Conference and Journal Publications

**Note** A more up-to-date list of my publications can be found on my [google scholar profile].

2020    **The GEM Benchmark: Natural Language Generation, its Evaluation and Metrics**, *Sebastian Gehrmann, et al.*.
arXiv preprint 2102.01672

2021    **Accelerating Antimicrobial Discovery with Controllable Deep Generative Models and Molecular Dynamics**, *Payel Das, Tom Sercu, Kahini Wadhawan\*, Inkit Padhi\*; Sebastian Gehrmann\*, Flaviu Cipcigan, Vijil Chenthamarakshan, Hendrik Strobelt, Cicero dos Santos, Pin-Yu Chen, Yi Yan Yang, Jeremy Tan, James Hedrick, Jason Crain, Aleksandra Mojsilovic.*
Nature Biomedical Engineering

2020    **Learning to Evaluate Translation Beyond English: BLEURT Submissions to the WMT Metrics 2020 Shared Task**, *Thibault Sellam, Amy Pu, Hyung Won Chung, Sebastian Gehrmann, Qijun Tan, Markus Freitag, Dipanjan Das, Ankur P Parikh.*
WMT

2020    **The Language Interpretability Tool: Extensible, Interactive Visualizations and Analysis for NLP Models**, *Ian Tenney\*, James Wexler\*, Jasmijn Bastings, Tolga Bolukbasi, Andy Coenen, Sebastian Gehrmann\*, Ellen Jiang, Mahima Pushkarna, Carey Radebaugh, Emily Reif, Ann Yuan .*
EMNLP Demo

2020    **ToTTo: A controlled table-to-text generation dataset**, *Ankur P Parikh, Xuezhi Wang, Sebastian Gehrmann\*, Manaal Faruqui, Bhuwan Dhingra, Diyi Yang, Dipanjan Das.*
EMNLP

2020    **Investigating gender bias in language models using causal mediation analysis**, *Jesse Vig\*, Sebastian Gehrmann\*, Yonatan Belinkov\*, Sharon Qian, Daniel Nevo, Yaron Singer, Stuart Shieber.*
NeurIPS

2020    **A Corpus for Detecting High-Context Clinical Indications in Intensive Care Patient Notes Focusing on Frequently Readmitted Patients**, *Edward Moseley, Joy Wu, Jonathan Welt, John Foote Jr., Patrick Tyler, David Grant, Eric Carlson, Sebastian Gehrmann, Franck Dernoncourt, Leo Celi.*
Proceedings of The 12th Language Resources and Evaluation Conference.

2020    **ExBERT: A Visual Analysis Tool to Explore Learned Representations in Transformers Models**, *B Hoover, H Strobelt, S Gehrmann.*
ACL 2020 Demo, NeurIPS 2019 Demo

2019    **Memory-Augmented Recurrent Neural Networks Can Learn Generalized Dyck Languages**, *M Süzgün, S Gehrmann, Y Belinkov, S Shieber.*
arXiv preprint 1911.03329

2019 **Encoder-Agnostic Adaptation for Conditional Language Generation**, *Z Ziegler, L Melas-Kyriazi, S Gehrmann, A Rush*.
arXiv preprint 1908.06938

2019 **Generating Abstractive Summaries with Finetuned Language Models**, *S Gehrmann\*, Z Ziegler\*, A Rush*.
INLG 2019 Generation Challenge Track

2019 **Margin Call: an Accessible Web-based Text Viewer with Generated Paragraph Summaries in the Margin**, *N Rizvi, S Gehrmann, F Dernoncourt*.
INLG Demos 2019

2019 **Visual Interaction with Deep Learning Models through Collaborative Semantic Inference**, *S Gehrmann\*, H Strobelt\*, R Krüger, HP Pfister, A Rush*.
VAST 2019

2019 **Evaluating an Automated Mediator for Joint Narratives in a Conflict Situation**, *M Zancanaro, O Stock, G Schiavo, A Cappelletti, S Gehrmann, D Canetti, O Shaked, S Fachter, R Yifat, R Mimran, T Weiss*.
Behaviour & Information Technology

2019 **GLTR: Statistical Detection of Fake Text**, *S Gehrmann, H Strobelt, A Rush*.
ACL Demos 2019, **Nominated for Best Demo Paper**

2019 **Improving Human Text Comprehension through Semi-Markov CRF-based Neural Section Title Generation**, *S Gehrmann, S Layne, F Dernoncourt*.
NAACL 2019

2018 **Seq2Seq-Vis: A Visual Debugging Tool for Sequence-to-Sequence Models**, *H Strobelt\*, S Gehrmann\*, M Behrisch, A Perer, HP Pfister, A Rush*.
VAST 2018, **Honorable Mention**

2018 **Bottom-Up Abstractive Summarization**, *S Gehrmann, Y Deng, A Rush*.
EMNLP 2018

2018 **End-to-End Content and Plan Selection for Natural Language Generation**, *S Gehrmann, F Dai, H Elder, A Rush*.
INLG 2018

2018 **Towards Controllable Generation of Diverse Natural Language**, *H Elder, S Gehrmann, A O'Connor, Q Liu*.
INLG 2018 - E2E NLG challenge special session

2018 **Comparing Rule-Based and Deep Learning Models for Patient Phenotyping**, *S Gehrmann, F Dernoncourt, Y Li, E Carlson, J Wu, J Welt, J Foote, E Moseley, D Grant, P Tyler, L Celi*.
PloS One 13 (2)

2017 **Behind the Scenes: A Medical Natural Language Processing Project**, *J Wu, F Dernoncourt, S Gehrmann, P Tyler, E Moseley, E Carlson, D Grant, Y Li, J Welt, L Celi*.
International Journal of Medical Informatics

2016 **Visual Analysis of Hidden State Dynamics in Recurrent Neural Networks**,
*H Strobelt, S Gehrmann, HP Pfister, A Rush*.
IEEE Transactions on Visualization and Computer Graphics 2017

2015 **Deploying AI Methods to Support Collaborative Writing: a Preliminary Investigation**, *S Gehrmann, L Urke, O Amir, B Grosz*.
CHI 2015, Extended Abstracts on Human Factors in Computing Systems

## Workshop Publications

2019 **Interactive White-Box Models through Collaborative Semantic Inference**,
*S Gehrmann\*, H Strobelt\*, R Krüger, HP Pfister, A Rush*.
ACL 2019 Workshop on BlackboxNLP

2019 **LSTMs can perform Dynamic Counting**, *M Süzgün, S Gehrmann, Y Belinkov, S Shieber*.
ACL 2019 Workshop on Deep Learning and Formal Languages

2019 **Interactive Visual Exploration of Latent Space (IVELS) for peptide auto-encoder model selection**, *T Sercu, S Gehrmann, H Strobelt, P Das, I Padhi, C Dos Santos, K Wadhawan, V Chenthamarakshan*.
ICLR 2019 Workshop on Deep Generative Models for Highly Structured Data

2018 **Bottom-Up Abstractive Summarization**, *S Gehrmann, Y Deng, A Rush*.
Amazon Graduate Student Symposium 2019

2018 **Debugging Sequence-to-Sequence with Seq2Seq-Vis**, *H Strobelt\*, S Gehrmann\*, M Behrisch, A Perer, HP Pfister, A Rush*.
EMNLP 2018 Workshop on BlackboxNLP

2018 **End-to-End Content and Plan Selection for Natural Language Generation**,
*S Gehrmann, F Dai, H Elder, A Rush*.
OpenNMT Workshop 2018

## Leadership

| | |
|---|---|
| 2021 | **GEM Workshop: Natural Language Generation, Evaluation, and Metrics**.<br>I co-organized this workshop, co-located with ACL 2021. |
| 2020 | **1st Workshop on Evaluating NLG Evaluation**.<br>I co-organized this workshop, co-located with INLG 2020. |
| 2019 | **International Natural Language Generation Conference**, *Tokio, Japan*.<br>I was on the organizing committee of the INLG Conference and acted as publicity chair. |
| 2016–2019 | **Computer Science Graduate Council**, *Harvard University*.<br>As president, I lead the communication of students with department administrators and planned biweekly social events for over 100 students and faculty |
| 2012–2015 | **Naktienclub**, *Nordakademie University of Applied Sciences*.<br>As president, I educated students about personal finance and investing through lecture series and acted as managing director of the club's own investment fund of over $10,000. |

## Teaching

| | |
|---|---|
| ACL 2020 | **Interpretability and analysis in neural NLP** , *Co-Organizer with Yonatan Belinkov and Ellie Pavlick*, Tutorial. |
| Winter 2020 | **Structure and Interpretation of Deep Networks**, *Co-Organizer*, MIT. |
| Fall 2017 | **CS182: Artificial Intelligence**, *Teaching Fellow*, Harvard SEAS. |
| Fall 2016 | **CS108: Intelligent Systems: Design and Ethical Challenges**, *Teaching Fellow*, Harvard SEAS. |
| Fall 2018 | **HST.953: Collaborative Data Science in Medicine**, *Mentor*, MIT, Cambridge, MA. |
| 2015 & 2016 | **Data Analysis with Python**, *Intructor*, Nordakademie University of Applied Science, Elmshorn, Germany. |
| 2013 & 2014 | **Introduction to Finance**, *Instructor, Organizer*, Nordakademie University of Applied Science, Elmshorn, Germany. |

## Awards and Scholarships

| | |
|---|---|
| 2019 | Selected as a member of the Siebel Scholar class of 2020 |
| 2019 | One of ten recipients of the Adobe Research Fellowship<br>(selected from over 1,000 applicants) |
| 2019 | 2 x Adobe Research Gift for Harvard NLP ($5,000 each) |
| 2019 | Best automated scores at the TL;DR challenge out of 12 submitted systems |
| 2018 | Harvard Open-Access Publishing Equity award ($1,500) |
| 2017 | Best automated scores at the E2E NLG challenge out of 60 submitted systems |
| 2016/17 | Karina A. Chen Graduate Student Research Fellowship in Engineering and Applied Sciences, Harvard University |

| | |
|---|---|
| 2013/14 | Award for outstanding service to the community of Nordakademie |
| 2006–2010 | First place in state championship in chess (first board in team of eight) |
| 2008 | Runner up in Mathematics Olympiad on state level |
| 2006, 2007, 2009 | Third place in Mathematics Olympiad on state level |

## Open Source

| | |
|---|---|
| OpenNMT-py | Top 5 contributor of one of the largest sequence-to-sequence modeling toolkits |
| Bottom-Up | Main author of models for state-of-the-art summarization modeling |
| GLTR | Backend author for the tool to detect generated text from language models |
| LSTMVis | Backend author for visual analysis of hidden states in recurrent neural networks |
| Seq2Seq-Vis | Backend author for interactive debugging of sequence-to-sequence models |

## Reviewing and Program Committees

| | |
|---|---|
| Area Chair | ACL 2020 (summarization), EACL 2021 (generation), NAACL 2021 (summarization), ACL 2021 (interpretability) |
| Conferences | ACL 2018, 2019, CoNLL 2019, EMNLP 2019 (outstanding reviewer), 2020 (outstanding reviewer) IEEE VIS 2019, NAACL 2019, NeurIPS 2020 (top 10% reviewer), INLG 2018, IUI 2018, 2019 |
| Workshops | BlackboxNLP 2020, SustaiNLP 2020, VisXAI 2018, 2019, 2020, NeuralGen 2019, New Frontiers in Summarization 2019, HAI-GEN 2020, 2021, EACL SRW 2020, WebNLG+ 2020, |
| Journals | IJMI, JAIR, JMIR, JMLR, PLoS One |