Saad Mahamood

Nationality: British 🇬🇧
LinkedIn: https://de.linkedin.com/in/dr-saad-mahamood-0073745
Homepage: https://www.saad.me.uk

PROFILE

An experienced computational linguist with extensive academic and commercial experience in building and researching AI-based Natural Language Generation systems that convert large datasets into textual summaries. I have extensive experience of working alongside domain experts. I'm also capable of leading or working as part of a team, and able to be inventive to research and solve real-world problems. I am looking to broaden my experiences by tackling new challenges.

EDUCATION

2006-2010	**University of Aberdeen, Scotland 🏴󠁧󠁢󠁳󠁣󠁴󠁿 – PhD in Computing Science**
2005-2006	**University of Sussex, England 🏴󠁧󠁢󠁥󠁮󠁧󠁿 – MSc in Informatics**
2001-2005	**Aberystwyth University, Wales 🏴󠁧󠁢󠁷󠁬󠁳󠁿 – BSc in Computing Science**

PHD DISSERTATION

Thesis Title: "Generating affective Natural Language for Parents of Neonatal Infants."

My PhD focused on the development of a Natural Language Generation (NLG) based system to convert clinical neonatal data into automatic daily summaries for parents of babies in Neonatal Intensive Care. This was done by summarising the information in a baby's electronic patient record to produce reports focusing on matters of most interest to the parent. My thesis also focused on the informational and emotional needs of parents and tried to understand how these needs could be met by generating tailored automatic reports. During my PhD I constructed a prototype NLG system to generate such reports for parents and evaluated this system with parents that previously had a baby in neonatal care.

WORK EXPERIENCE

NLG Expert & Lead Data Scientist, trivago N.V., Düsseldorf, Germany 🇩🇪   March 2018 - Present

Employed by trivago as an expert to help build NLG data-to-text solutions. In particular, helping to generate accommodation descriptions automatically with a high degree of content and linguistic variation for the purposes of search engine optimisation, updatability, coverage, and to inform users. I have also worked along side a colleague to help built a solution that could generate a unique selling point text from user reviews.

Responsibilities also include NLP and Data Science related tasks as well as leading a team of Data Scientists and Analysts. I work on both leading data science solutions within trivago and working on projects to solve problems such as scoring user reviews, computing star ratings for alternative accommodations, data validation/verification, geocode consensus modelling, and calculating objective image quality scores for accommodations images.

I have also worked to help review research papers for academic conferences such as ACL and INLG and writing research papers for academic conferences. In my spare time, I help to maintain the open source *SimpleNLG* surface realiser project (https://github.com/simplenlg/simplenlg) and maintain an active interest in keeping up with the latest research developments in NLG and NLP.

Senior Natural Language Generation Engineer, Arria NLG, Aberdeen, Scotland 🏴󠁧󠁢󠁳󠁣󠁴󠁿   August 2012 - February 2018

Employed by Arria NLG to research and develop the creation of new generic tools that can enable the quicker/easier production of NLG systems for commercial clients. In the past most NLG systems contained many bespoke software components specific to a particular big data client, thus making reuse inherently difficult. Recently, I have been involved in helping to build a new NLG microplanner and multilingual (English and German) realiser that allows for the generation of texts in more than one

language from a given syntactical sentence representation. This has involved dealing with aspects such as aggregation, referring expression generation, syntax, morphology, morphophonology, and orthography.

I have direct experience working and engaging with commercial clients to deliver NLG based solutions. I have also taken the lead to author an academic research paper within a commercial environment and helped to review a patent application based directly on the intellectual property that I've worked on. Additionally, in my spare time, I helped to maintain the open source *SimpleNLG* surface realiser project and also worked on a research paper with an external academic collaborator.

Honorary Research Fellow, Computing Science, University of Aberdeen, Scotland 🏴   **July 2012 - February 2018**

Honorary appointment to continue the work I had done during my previous Research Fellow tenure on a part-time basis. This involved working on research papers, reviewing research papers for conferences, and attending the NLG research group to discuss relevant papers.

Research Fellow, Computing Science, University of Aberdeen, Scotland 🏴   **July 2010 - June 2012**

Appointed as a research fellow to work on a follow-up project on research work done during my PhD. I was responsible for the development, deployment, refinement, and evaluation of an NLG system to convert neonatal clinical data into reports for parents of pre-term neonates. I worked closely with clinical domain experts and neonatal clinical data to help refine the quality of texts generated and bring about the integration of our NLG system with the existing electronic health record software. I also developed the web-based interface for parents to access the automatically generated reports from a public computer. Additionally, I was also responsible for project managing a research assistant, given the responsibility to apply for NHS ethics approval for onward evaluation, and also performed tutorial responsibilities for student demonstration classes.

SKILLS

Computer Programming Languages: Java and Python.
Cloud Technologies: AWS and GCP.

Languages: English 🇬🇧 (native), German 🇩🇪 (B1.1)

PUBLICATIONS

- Simon Millie, Kaustubh D. Dhole, **Saad Mahamood**, Laura Perez-Beltrachini, Varun Gangal, Mihir Kale, Emiel van Miltenburg (2021). Automatic Construction of Evaluation Suites for Natural Language Generation Datasets. *To appear at the thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021)*.
- Emiel van Miltenberg, Miruna Clinciu, Ondrej Dusek, Dimitra Gkatzia, Stephanie Inglis, Leo Lappänen, **Saad Mahamood**, Emma Manning, Stephanie Schoch, Craig Thompson, and Lulu Wen (2021). Underreporting of errors in NLG output, and what to do about it. *To appear at the 14th International Natural Language Generation Conference (INLG 2021)*. Aberdeen, Scotland 🏴.
- Miruna Clinciu, Dimitra Gkatzia, and **Saad Mahamood** (2021). It's common sense, isn't it? Demystifying Human Evaluations in Commonsense-enhanced NLG systems. *Proceedings of the w*orkshop on Human Evaluation of NLP systems (HumEval) at the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL). Kyiv, Ukraine 🇺🇦.
- David M. Howcroft, Anya Belz, Miruna-Adriana Clinciu, Dimitra Gkatzia, Sadid A. Hasan, **Saad Mahamood**, Simon Mille, Emiel van Miltenburg, Sashank Santhanam, Verena Rieser (2020). Twenty Years of Confusion in Human Evaluation: NLG Needs Evaluation Sheets and Standardised Definitions. *Proceedings of the 13th International Natural Language Generation Conference (INLG 2020)*. Dublin, Ireland 🇮🇪.
- **Saad Mahamood** and Maciej Zembrzuski (2019). Hotel Scribe: Generating High Variation Hotel Descriptions. *Proceedings of the 12th International Natural Language Generation Conference (INLG 2019)*. Tokyo, Japan 🇯🇵.
- Dimitra Gkatzia and **Saad Mahamood** (2015). A Snapshot of NLG Evaluation Practices 2005-2014. *Proceedings of the 15th European Workshop on Natural Language Generation — ENLG-15*. Brighton, United Kingdom 🇬🇧.
- **Saad Mahamood**, William Bradshaw, and Ehud Reiter (2014). Generating Annotated Graphs using the NLG Pipeline Architecture. *Proceedings of the 8th International Natural Language Generation 2014 (INLG 2014)*. Philadelphia, PA, USA 🇺🇸.
- **Saad Mahamood** and Ehud Reiter (2012). Working with Clinicians to Improve a Patient-Information NLG system. *Proceedings of the International Natural Language Generation 2012 — INLG 2012*. Starved Rock, IL, USA 🇺🇸.
- **Saad Mahamood** and Ehud Reiter (2011). Generating Affective Natural Language for Parents of Neonatal Infants. *13th European Workshop on Natural Language Generation — ENLG-11*. Nancy, France 🇫🇷.
- Wendy Moncur, **Saad Mahamood**, Ehud Reiter, Yvonne Freer (2009). Involving healthcare consumers in knowledge acquisition for virtual healthcare. Virtual Healthcare Interaction 2009 (AAAI Fall Symposium Series).
- Albert Gatt, François Portet, Ehud Reiter, Jim Hunter, **Saad Mahamood**, Wendy Moncur, Somayajulu Sripada (2009). From Data to Text in the Neonatal Intensive Care Unit: Using NLG Technology for Decision Support and Information Management. AI Communications 22:153-186.

- **Saad Mahamood**, Ehud Reiter, Chris Mellish (2008). Neonatal Intensive Care Information for Parents — An Affective Approach. *Proceedings of the Twenty-First IEEE International Symposium on Computer-Based Medical Systems*. Jyväskylä, Finland 🇫🇮.
- **Saad Mahamood**, Ehud Reiter, Chris Mellish (2007). A Comparison of Hedged and Non-hedged NLG Texts. Proceedings of the 11th European Workshop on Natural Language Generation — ENLG-07. Schloss Dagstul, Germany 🇩🇪.

PATENTS

- Method and apparatus for document planning — *Inventor*: Saad Mahamood (2014), *Patent Number*: US20160232152A1.