

# Simon Benigeri

✉ [simon.benigeri@northwestern.edu](mailto:simon.benigeri@northwestern.edu)   [🌐 yourwebsite.com](#)   [in simon-benigeri](#)   [🔊 simon-benigeri](#)

## Summary

---

Ph.D. candidate in AI at Northwestern, researching NLP and conversational AI. 3 years of AI research in cardiology and 3 years of AI industry experience at IBM.

## Education

---

**Northwestern University** *PhD in Artificial Intelligence*

*Sept 2023 – present*

- Advisor: [Larry Birnbaum](#) [🔗](#)
- **Research focus:** Natural Language Processing, Conversational AI

**Northwestern University** *MS in Artificial Intelligence*

*Sept 2020 – Dec 2021*

- GPA: 3.98/4.0

**University of Pennsylvania** *BA in Mathematics*

*Sept 2010 – June 2016*

## Experience

---

**Graduate Research Assistant - PhD in AI**

*Evanston, IL*

*Northwestern University*

*Sep 2023 – present*

- Conducting research to develop practical conversational AI systems, focusing on enhancing reliability in real-world applications and robustness to complex but natural human conversational behavior.
- Submitted a paper to a major NLP conference, focusing on robustness and reasoning in conversational AI.
- Teaching Assistant for CS 337: Natural Language Processing (Fall 2024), guiding students in foundational and advanced concepts of NLP.

**AI Research Engineer - Cardiology**

*Chicago, IL*

*Bluhm Cardiovascular Institute Center for AI, Northwestern University*

*Feb 2022 – present*

*(Transitioned to part-time in Sep 2023)*

- Collaborated with physician-scientists to develop AI/ML solutions for cardiovascular disease, focusing on heart failure and multimodal data integration.
- Led the creation of high-quality cardiac amyloidosis datasets, increasing data quality from 3,000 weakly labeled to 700 gold-standard cases, leading to multiple publications and grant funding.
- Built predictive models for heart failure subtypes and mortality risk using clinical text and structured data, contributing to the team's published research.
- Automated echocardiogram processing with startup Ultromics to enhance diagnostics for heart failure and cardiac amyloidosis, with plans for publication.
- Developed HIPAA-compliant data de-identification tools for DICOM images and clinical notes, ensuring data security and reducing processing time by 75% compared to previous de-identification solutions.

**Senior Technical Consultant - AI**

*Paris, France*

*IBM*

*Feb 2018 – Sep 2020*

- Developed NLP-driven AI systems for enterprise applications (banking compliance, HR, customer service, IT support, construction), integrating machine learning with business logic in collaboration with clients.
- Led the end-to-end model development lifecycle, from data collection and annotation to evaluation, deployment, and monitoring; developed application back-ends.
- Achieved significant impact, including reducing manual effort in analyzing annual reports and adverse news for KYC compliance by 50%, as highlighted in [Global Trade Review](#) [🔗](#).
- Supported customer complaints for the French National Railway Company (SNCF) during strikes via the chatbot [Tout Oui](#) [🔗](#), as referenced in [Ouest-France](#) [🔗](#).

## Publications

---

Automated identification of heart failure with reduced ejection fraction using deep learning-based natural language processing <a href="#">🔗</a> ( <i>JACC: Heart Failure</i> , 2025)	2025
A Flash in the Pan: Better Prompting Strategies to Deploy Out-of-the-Box LLMs as Conversational Recommendation Systems <a href="#">🔗</a> ( <i>ACL: COLING 2025</i> )	2025
Evaluating the performance and potential bias of predictive models for the detection of transthyretin cardiac amyloidosis <a href="#">🔗</a> ( <i>preprint on medRxiv</i> )	2024
Multi-domain Summarization from Leaderboards to Practice: Re-examining Automatic and Human Evaluation <a href="#">🔗</a> ( <i>ACL: GEM 2023</i> , <a href="#">GitHub</a> <a href="#">🔗</a> )	2023

## Projects

---

<b>Probing LLM planning and reasoning by playing 20 questions</b> <a href="#">GitHub</a> <a href="#">🔗</a>	
◦ Built a backend to evaluate LLM reasoning for the game 20 questions; implemented budget forcing (S1: <a href="#">Simple test-time scaling</a> <a href="#">🔗</a> ) to improve reasoning performance.	
<b>CalendarBot</b> <a href="#">GitHub</a> <a href="#">🔗</a>	
◦ Built a Retrieval-Augmented Generation (RAG) chatbot for calendar event retrieval using a hybrid approach: date-based retrieval followed by vector-based retrieval.	
<b>Protecting forest elephants through audio monitoring</b> <a href="#">Blog post</a> <a href="#">🔗</a>	
<b>Deep reinforcement learning agent for stock trading</b> <a href="#">GitHub</a> <a href="#">🔗</a>	
<b>LSTM language model in PyTorch</b> <a href="#">GitHub</a> <a href="#">🔗</a>	

## Skills

---

**Languages:** Python, SQL, Bash

**Technologies:** PyTorch, TensorFlow, Hugging Face Transformers, NLTK, SpaCy, NumPy, pandas, scikit-learn, FastAPI, Docker, Git, AWS, Google Cloud Platform (GCP), Azure, Relational Databases, Vector Databases

**Research and collaboration:** Research writing (ACL, JACC), effective communication with cross-functional teams, presenting research findings, communicating technical concepts to diverse audiences, solving open problems in collaborative environments