

# Ruli Manurung

Staff Computational Linguist

Tokyo, Japan

 [ruli.manurung@gmail.com](mailto:ruli.manurung@gmail.com)

 [rulimanurung.github.io](https://github.com/rulimanurung)

 [in/rulimanurung](https://in.rulimanurung)

 [Google Scholar](#)

 [@rulimanurung](#)

**tl;dr:** I uniquely combine 8 years of experience in the tech industry delivering production-grade AI systems (Google Assistant) with 12 years in academia, bootstrapping an NLP lab from scratch to become one of the premier research labs in the region focusing on diverse topics ranging from Indonesian NLP resources/tools to computational creativity.

## Experience

### Google Japan (Feb. 2016 - present)

Tokyo, Japan

May 2022 - Present

#### Dialog Data & Eval Tech Lead, Google Assistant | Staff Computational Linguist

- Designed and led the implementation of a fully automated pipeline that leverages LLMs to generate high-quality synthetic dialog data with all requisite semantic annotations & context signals for training and evaluating NLP models.
- Established scalable & reliable eval methods of multi-turn dialog parsing.
- Developed fully horizontalized semantic representations for Google Assistant dialog journeys.

Feb. 2020 - Apr. 2022

#### NLP Data Linguist Tech Lead / Manager, Google Assistant | Staff Computational Linguist

- Launched a scaled-up & fully operationalized synthetic data pipeline that delivers high quality human-annotated data for training and evaluating Seq2seq NLP models.
- Co-led a project to leverage rich curated lexicons and linguistically motivated data-augmentation methods to automatically generate machine-synthetic data for training and evaluating NLP models.
- Co-founded and co-led a framework for the development of semantic meaning representations that maximize the development agility and accuracy of Google Assistant NLP models.

Nov. 2017 - Jan. 2020

#### APAC Regional Lead & Tech Lead / Manager, Google Assistant | Computational Linguist

- TL/M of a team (6 full-time engineers) that focuses on developing highly reusable and horizontal annotators for core semantic expressions such as numbers, dates, times, money amounts, etc. The resulting libraries are a fundamental component used by multiple client teams across Google.
- Served as APAC regional lead for a large team of analytical and computational linguists. I oversaw the growth of the team, more than doubling the headcount in 2 years (7→15). This team was mission-critical for launching Google Assistant across all key APAC markets.
- Co-led project to develop high-precision semantic grammars for Google Assistant NLP across multiple locales.

Feb. 2016 - Oct. 2017

#### Lead Linguist, Google Research | Analytical Linguist

- Lead linguist for all aspects specifically related to Indonesian NLP within Google Research, among others: developing detailed annotation guidelines, producing deep linguistic analysis reports, training vendor teams, authoring grammars.
- Co-led project to linguistically analyze and author Natural Language Generation (NLG) modules for Google Assistant text-to-speech messages across all non-English locales. This was critical for enabling all Google Assistant launches in non-English markets, at the time triggering over 1B messages/month.

## Universitas Indonesia (Jan. 2007 - Dec. 2015)

Jakarta, Indonesia

Jan. 2007 -  
Dec. 2015

### IR/NLP lab co-lead & Computer Science Dept. Head | Associate Professor

- Bootstrapped [NLP lab](#) from scratch to become the premier research lab in the region focusing on: (i) Indonesian NLP resources/tools, (ii) computational language creativity.
- Researcher / Co-PI on 11 research grants from bodies including EPSRC(UK), ARC(Australia), British Council (UK), and Ministry of Research & Technology (Indonesia).
- Research publications: [Google Scholar](#) (*h*-index: 22)
- Research supervision: Supervised/co-supervised to completion 61 undergraduate theses, 29 masters theses, and 4 doctoral dissertations.
- Professional service: Co-chair [PACLIC '12](#) (Pacific Asia Conference on Language, Information, and Computation)
- Program committee member:
  - ALR (Workshop On Asian Language Resources) 2011
  - ALTW (Australasian Language Technology Workshop) 2012
  - ICC3 (International Conference on Computational Creativity) 2010-2015
  - MALINDO Workshop 2008-2012
  - PACLIC (Pacific Asia Conference on Language, Information, and Computation) 2010-2013
  - PACLING (Conference of the Pacific Association for Computational Linguistics) 2015
  - LaTeCH (Language Technology for Cultural Heritage, Social Sciences, and Humanities) 2007
  - EvoMUSART (International Conference on Evolutionary and Biologically Inspired Music, Sound, Art and Design) 2007-2015
- Visiting scholar / guest lecturer / invited speaker to multiple research institutes, summer schools: Helsinki University, University of Aberdeen, De La Salle University, Hanoi University of Technology, Bina Nusantara University.

## University of Edinburgh (Jan. 2004 - Nov. 2006)

Edinburgh, UK

Jan. 2004 -  
Nov. 2006

### School of Informatics | Postdoctoral Research Fellow

- Project title: **STANDUP (System To Augment Non-speakers' Dialogue Using Puns)**
- Principal Investigators: Graeme Ritchie (University of Aberdeen), Helen Pain (University of Edinburgh), Annalu Waller (University of Dundee)
- Implemented a novel joke generation system, and designed and implemented a suitable user interface for language impaired children, conducted user studies.
- Press coverage: [BBC](#), [The Guardian](#)

## Education

### University of Edinburgh

Edinburgh, UK

1998-2003

### Ph.D in Artificial Intelligence | School of Informatics

- PhD thesis title: "[An evolutionary algorithm approach to poetry generation](#)"
- Supervisors: Dr. Henry Thompson, Dr. Graeme Ritchie

## Universitas Indonesia

Jakarta, Indonesia

1992-1997

### Bachelor of Computer Science | Faculty of Computer Science

- Undergraduate thesis title: "*Image Compression using the Wavelet Transformation*"
- Graduated *Cum Laude*