

RESEARCH INTEREST Computational Social Science, Computational Linguistics, Human-Computer Interaction

EDUCATION **Carnegie Mellon University** 2017 - 2019

Masters in Language Technologies, School of Computer Science

Advisors: [Anthony Tomasic](#), [Eduard Hovy](#), collaborated with [Graham Neubig](#) and [Geoff Kaufman](#)

Relevant courses: Neural Networks for NLP, Computational Semantics, Computational Ethics, Machine Learning, Grammars and Lexicon, Social Web

University of Indonesia 2012 - 2016

Bachelor in Computer Science with the highest distinction (GPA: 3.99 out of 4.00)

Research Assistant at the Information Retrieval Lab, **Advisor:** [Mirna Adriani](#)

EMPLOYMENT **University of Pennsylvania** 2020 - Present

Research Specialist, Advisors: [Lyle Ungar](#) and [Ian Barnett](#)

- World-Well Being Project Lab (measuring psychological well-being based on language)
- Conducted research of NLP for social good and linguistics style analysis
- Developed and maintained data processing pipeline for digital phenotype

University of California, Davis, Natural Language Processing Lab 2019 - 2020

Research Assistant, Advisor: [Zhou Yu](#)

- Built a sociable and persuasive recommendation dialog systems with deep learning model and social science theories (in collaboration with [Facebook AI Research](#))
- Work was published in EMNLP 2020 and submitted to NAACL 2021

PEER-REVIEWED
PUBLICATIONS

[1] ***INSPIRED: Toward Sociable Recommendation Dialog Systems***

[Shirley Anugrah Hayati](#), Dongyeop Kang, Qingxiaoyang Zhu, Weiyang Shi, Zhou Yu

In Proceedings of Conference on Empirical Methods on Natural Language Processing (EMNLP) 2020

[2] ***A Summary of the First Workshop on Language Technology for Language Documentation and Revitalization***

Graham Neubig, Shruti Rijhwani, Alexis Palmer, Jordan MacKenzie, Hilaria Cruz, Xinjian Li, Matthew Lee, Aditi Chaudhary, Luke Gessler, Steven Abney, [Shirley Anugrah Hayati](#), et al.,

LREC 2020 Workshop on Spoken Language Technologies for Under-resourced languages (SLTU)

[3] ***What A Sunny Day ☀️: Towards Emoji-Sensitive Irony Detection***

Aditi Chaudhary*, [Shirley Anugrah Hayati](#)*, Naoki Otani*, Alan W Black (* equal contributions)

EMNLP 2019 Workshop on Noisy User-Generated Text (W-NUT)

[4] ***Analyzing Incorporation of Emotion in Emoji Prediction***

[Shirley Anugrah Hayati](#)* and Aldrian Obaja Muis* (* equal contributions)

North American Association for Computational Linguistics (NAACL) 2019 Workshop on Computational Approaches to Subjectivity, Sentiment, and Social Media Analysis (WASSA)

[5] ***Retrieval-Based Neural Code Generation***

[Shirley Anugrah Hayati](#), Raphael Olivier, Pravalika Avvaru, Anthony Tomasic, Graham Neubig

In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP) 2018

[6] *Short Text Classification on Complaint Documents*

[Shirley Anugrah Hayati](#), Alfian Farizki Wicaksono, Mirna Adriani

CICLing. International Journal of Computational Linguistics and Applications (IJCLA) 2016

[7] *Building Occupation Chronology with Linguistics and Statistical Features*

[Shirley Anugrah Hayati](#), Mirna Adriani

Undergraduate Thesis. 2016

RESEARCH PROJECTS	Sociable and Persuasive Recommendation Dialog Systems (University of California, Davis) 2019 - 2020
	Legal Document Processing (Carnegie Mellon University & Accenture) 2019
	Media “Tagging” as a Lightweight, Informal Implicit Bias Measurement Tool (Human Computer Interaction, Carnegie Mellon University) 2019
	<ul style="list-style-type: none">Developed measurement to reveal users’ implicit biases in the domain of technology and AI.
	Disease Prediction from Medical Records (Carnegie Mellon University & MetLife) 2018-2019
	Bus Fullness Prediction with Participatory Design (Carnegie Mellon University) 2018
	SayHear (Language Technologies Institute, Carnegie Mellon University) 2017 - 2018
	<ul style="list-style-type: none">Building an automatic question-answering system for tabular data;Developed SQL-to-Abstract Syntax Tree neural model for semantic parsing and retrieving answers from the database.
	Automatic Topic Detection on Complaint Documents (University of Indonesia) 2015
	<ul style="list-style-type: none">In collaboration with the Executive Office of the President of the Republic of IndonesiaConducted a text mining research using Latent-Dirichlet Allocation to extract features on short documents.
TEACHING ASSISTANTSHIP	Natural Language Processing (Carnegie Mellon University) 2017
	Natural Language Processing, Artificial Intelligence, Declarative Programming, Calculus, Linear Algebra (University of Indonesia) 2013 - 2017
FELLOWSHIPS, AWARDS	Graduate Research Fellowship 2018 - 2019 Carnegie Mellon University research fellowship including full tuition and monthly stipend
	Top 10 Finalists, Data Mining 2015 Awarded by Gemastik 8, a national IT competition held by Indonesian Ministry of Education
	Silver Medalist (2nd place), User Experience Design 2014 Awarded by Gemastik 7, a national IT competition held by Indonesian Ministry of Education
SERVICE & PROFESSIONAL ACTIVITIES	NAACL Reviewer (2021), EMNLP Volunteer (2020), Women@SCS CMU (2017-2019), ACL, Association of Indonesian Students in the United States (2017 - Present), Reviewer for CMU LTI Student Research Symposium (2018), Competitive Programming Competition Organizer (2014)
SKILLS	
Computer Languages:	Python, Java, HTML, PHP, JavaScript, SQL, CSS, MATLAB, R, F#, Haskell
Natural Languages:	English (Proficient), Indonesian (Native), Chinese (Beginner), Korean (Beginner), French (Beginner)
Tools:	Stanford CoreNLP, NLTK, scikit-learn, SpaCy, Numpy, Pandas, Django, Matplotlib, PyTorch, Dynet, Tensorflow, ParlAI, Mechanical Turk, Figure Eight, Prolific