

# Long Phan

[lnp26@case.edu](mailto:lnp26@case.edu) • Cleveland, OH • (216) 612 4785  
[linkedin.com/in/Long-Phan-3110](https://www.linkedin.com/in/Long-Phan-3110) • [github.com/justinphan3110](https://github.com/justinphan3110)

---

## EDUCATION

**Case Western Reserve University** Expected MAY 2022  
**Bachelor of Science in Computer Science**  
Concentration: Software Engineering / Deep Learning - Natural Language Processing  
Cumulative GPA: 3.48

---

## Publications

**SciFive: a text-to-text transformer model for biomedical literature**, 2021  
**CoTexT: Multi-task Learning with Code-Text Transformer**, NLP4Prog 2021, a workshop of ACL  
**SPBERT: Pre-training BERT on SPARQLQueries for End-to-end Question Answering over Knowledge Graphs**, in reviewing for ISWC 2021  
**Hierarchical Transformer Encoders for Vietnamese Spelling Correction**, IEA / AIE 2021  
**Leveraging Transfer Learning for Reliable Intelligence Identification on Vietnamese SNSs (ReINTEL)**, 7th International Workshop on Vietnamese Language and Speech Processing VLSP 2020

---

## EXPERIENCE

**National Cancer Institute (NCI):** Deep Learning Research Special Volunteer; October 2020 - present

- Worked in the lab of Dr. Grégoire Altan-Bonnet to develop a new transformer-based pre-trained biomedical language representation model

**Zalo Group:** AI Software Engineer Intern; April - October 2020

- Researched and Developed a scalable system translating domain specific natural language questions into Knowledge Graph's Query in **Java** and **Python**.
- Improved previous techniques and models with more than 0.97 accuracy score
- Created a backend pipeline and an Admin Dashboard to efficiently keep track of new data with **Spark**, **Kafka** and **ReactJs**

**VNG Corporation:** AI Software Engineer Intern; May - August 2019

- Initiated the development of a **first agnostic Named-entity recognition (NER) module** capturing and resolving date/time text into structured data in **Haskell** and **Java** for Ki-Ki, the first Vietnamese intelligent personal assistant
- The module has been deployed for 80 million users on [Zing Mp3](#)

---

## PROJECTS

**Machine Learning Accessibility Initiative:** January 2020 - present

- A web application platform that uses deep learning analysis of input data to provide customized machine learning capabilities for users with no experience in artificial intelligence
- Built with Python Django, MongoDB, ReactJS, Google Cloud Platform

---

## SKILLS

**Programming Language:** Java, JavaScript, Python  
**Frameworks and Libraries:** TensorFlow, Tensor2tensor, ReactJS, React-Native, Flask, Express.js, RestAPI, Elastic Search, Kibana, MongoDB, Docker, SQL, Linux Environment