

Queenie Luo

✉ queenieluo@g.harvard.edu | 🏠 ealc.fas.harvard.edu/people/queenie-luo | 📄 Google Scholar | 🌐 <https://sites.harvard.edu/queenieluo/>

Education

Harvard University

PH.D. STUDENT IN EAST ASIAN LANGUAGES AND CIVILIZATIONS

Cambridge, MA

Sep. 2020 - PRESENT

- **GPA: 4.0** (up-to-date)

Harvard University

M.S. IN DATA SCIENCE

Cambridge, MA

Sep. 2020 - May. 2023

- **GPA: 4.0**, **MIT cross-registration: 5.0/5.0**
- Harvard John A. Paulson School of Engineering and Applied Sciences
- Thesis Title: “Decoding the Past: Strategic and Historical Insights from Simulating the Battle of Changping”
- Selected Courses: Advances in Computer Vision (MIT), Quantitative Methods for Natural Language Processing (MIT), Sensorimotor Learning (MIT)

Harvard University

M.T.S IN BUDDHIST STUDIES

Cambridge, MA

Sep. 2018 - May 2020

- **GPA: 3.96**
- Harvard Divinity School

Columbia University

B.A. IN RELIGION (MAJOR) AND PHYSICS (CONCENTRATION)

New York, NY

Jan. 2016 - May. 2018

- **GPA: 3.98**
- **Honors:** Summa Cum Laude; Phi Beta Kappa; Departmental Honors; GS Honor Society; Senior Marshal

Parsons School of Design, The New School

B.F.A STUDENT IN FASHION DESIGN

New York, NY

Sep. 2012 - Dec. 2015

- **Awards:** Second Place Honors, Fusion Fashion Competition, Parsons and Fashion Institute of Technology, 2013

Honors & Awards

Liang Qichao Travel Fund, Harvard Fairbank Center

Spring, 2025

Harvard China Fund Research Assistant Grant, Principal Investigator: Michael D. Smith

2024 - 2025

Harvard China Fund Research Assistant Grant, Principal Investigator: Michael Puett

2022 - 2023

Harvard China Fund Faculty Grant, Principal Investigator: Leonard W.J. van der Kuip

2021 - 2023

Harvard University Asia Center Research Grant (2 terms), Harvard University

2019, 2021

Denie S. Weil Fellowship, Harvard Divinity School

2019-2020

Walton Brooks and Helen Celemia McDaniel Fund, Harvard Divinity School

2018-2019

Phi Beta Kappa Outstanding Scholarship Annual Award, Columbia University

May, 2018

William Strong Scholarship (2 terms), Columbia University

2017, 2018

General Studies Scholarship (2 terms), Columbia University

2016, 2017

General Studies Dean's List (4 terms), Columbia University

2016, 2017, 2018

Language Proficiency

Advanced Classical Tibetan, Classical Chinese

Beginning Sanskrit, German

Fluent English

Native Mandarin Chinese

Publications

(† indicates equal contribution.)

[1] **Queenie Luo**, Michael J. Puett, and Michael D. Smith. “A Perspectival Mirror of the Elephant”. *Communications of the ACM*, 67.8. August 2024.

[2] **Queenie Luo** and Michael Puett. “Anglo-American bias could make generative AI an invisible intellectual cage”. *Nature* 629, 1004. May 2024.

[3] **Queenie Luo** and Michael Puett. “We must decolonise AI to overcome cultural bias in the classroom”. *South China Morning Post*. April 24, 2024.

- [4] **Queenie Luo**, Michael J. Puett, and Michael D. Smith. “A Perspectival Mirror of the Elephant: Investigating Language Bias on Google, ChatGPT, YouTube, and Wikipedia”. *ACM Queue* 22, 1. January/February 2024.
- [5] **Queenie Luo**[†] and Yung-Sung Chuang[†] “Cleansing Jewel: A Neural Spelling Correction Model Built On Google OCR-ed Tibetan Manuscripts”. *ACM Transactions on Asian and Low-Resource Language Information Processing*. April 2024.
- [6] **Queenie Luo**, Leonard W. J. van der Kuijp. “Norbu Ketaka: Auto-Correcting BDRC’s E-Text Corpora Using Natural Language Processing and Computer Vision Methods”. *Revue d’Etudes Tibétaines*. *Tibetan digital humanities and natural language processing. Proceedings of the IATS 2022 Panel*
- [7] **Queenie Luo**. “San Cai Fa Mi”. *Harvard-Yenching Institute*. October 26, 2020.

Research Experience

Institute for Quantitative Social Science, Harvard University

Sep. 2021 - PRESENT

DATA SCIENTIST

- Principal Investigator: Peter Bol
- Built real-time **deep learning/machine learning** models as web applications on the Amazon AWS platform with S3, Redshift, Data Pipeline, DynamoDB, Lambda and SageMaker service.
- Developed Lepton: Natural Language Processing (NLP) Applications in China Historical Records Data Mining. [[Lepton Model](#)]
- Developed Kraft: Korean Romanization from **Transformer** [[Kraft Model](#)]

OCR for Chinese Paleographic Materials Project, Fairbank Center for Chinese Studies, Harvard University

Aug. 2022 - Aug. 2023

LEAD RESEARCH SCIENTIST

- Principal Investigator: Michael Puett
- Developed a customized Optical Character Recognition (OCR) pipeline for Chinese paleographic materials using classical **Computer Vision** and **Deep Learning** techniques to digitize Bronze inscriptions from the Shang and Zhou Dynasties, sources from the Two Han Dynasties, and Bamboo slips.

Norbu Ketaka Project, Fairbank Center for Chinese Studies, Harvard University

Jan. 2022 - Aug. 2023

LEAD RESEARCH SCIENTIST

- Principal Investigator: Leonard W.J. van der Kuijp
- Led *Norbu Ketaka* project to produce **one million** pages of highly accurate dataset for Classical Tibetan for machine translation usage. [[Access our collection](#)]
- Developed a customized **Computer Vision** model to pre-process the Classic Tibetan manuscript images to reduce the noise produced by Google OCR API.
- Implemented the Tibetan NLP **BERT model** to auto-detect and labeled the textual noise in the OCR-ed Classic Tibetan texts.
- Developed a personnel project management tool to facilitate the distribution of raw data to offshore team members based on their Tibetan proficiency and time availability.

Conference Presentations

Greater Boston Digital Research and Pedagogy Symposium

Boston, USA

THE BOSTON DIGITAL HUMANITIES CONSORTIUM IS COMMITTED TO THE COLLABORATIVE DEVELOPMENT IN THE DIGITAL HUMANITIES AND COMPUTATIONAL SOCIAL SCIENCES.

Apr. 2023

- China Biographical Database (CBDB) Kinship Visualization

Tools of the Trade: the Way Forward

Cambridge, USA

CONFERENCE ON THE TRANSITION FROM PRINT TO DIGITAL TOOLS, DATABASES, AND PLATFORMS IN CHINESE, JAPANESE, KOREAN, VIETNAMESE AND BUDDHIST STUDIES

Mar. 2023

- Lepton: Natural Language Processing Applications in China Historical Records Data Mining
- Kraft (Korean Romanization from Transformer)

16th International Association for Tibetan Studies (IATS) Seminar

Prague, Czech Republic

THE PANEL OF TIBETAN DIGITAL HUMANITIES AND NATURAL LANGUAGE PROCESSING

Jul. 2022

- Norbu ketaka: Auto-Correcting BDRC’s E-Text Corpus Using Natural Language Processing and Computer Vision Methods

Fourth International Conference on the Wutai Culture

Shanxi, China

IDENTITY AND NETWORKS IN BUDDHISM AND EAST ASIAN RELIGIONS

Jul. 2018

- The Clothing Culture in Tang Dynasty

Web Design Project

WEBSITE & GRAPHIC DESIGNER, DIGITAL MEDIA FELLOW

- Advisor: Michael Puett [[Gened1091 Website](#)]
- Developed student project gallery's web service, and departmental data visualization platform with WordPress and Javascript.
- Implemented multiple high-volume data sets from RESTful APIs and Data Transfer files into Google Cloud Platform by Google BigQuery and Dataflow. Designed the automation process to QA and monitor the data ETL process with Google Data Studio.
- Designed a collection of icons and illustrations specifically tailored for course materials.

Scholarly Website, *Department of East Asian Languages and Civilizations, Harvard University*

WEBSITE & GRAPHIC DESIGNER

- Advisor: Melissa McCormick [[Scholarly Website](#)]
- Designed and developed scholarly website.

Teachings

CS109A: Introduction to Data Science, Teaching Fellow

- Supported Lab 5 which covered topics of cross-validation, confidence intervals & bootstrap, statsmodels, Ridge and Lasso regularization.
- Supported Lab 6 which covered topics of missingness, logistic regression, and polynomial logistic regression.
- Supported Lab 7 which covered topics of Decision Trees, Bagging, and Random Forest

AC221: Critical Thinking in Data Science, Teaching Fellow

- Lead a lecture human language bias (Apr. 6, 2024).
- Prepared critical thinking and writing tips for students.
- Facilitated in-class discussion.

Skills

Programming Languages	Python, SQL, JavaScript, Mathematica, Shell Scripting
Python Libraries	PyTorch, HuggingFace Transformer series, Pygame, OpenAI Gym Tensorflow, Keras, NumPy, Scikit-Learn, pandas, OpenCV, Matplotlib
Toolkits	LaTex, Docker, Git/GitHub, Linux, WordPress, Postman Photoshop, Clip Studio Paint, Stable Diffusion, Illustrator Premiere Pro, After Effect, Animate, Character Animator, Audition
Cloud Computing	AWS, Google Cloud Platform