

## CONTACT

⚡ <https://xudongolivershen.github.io>  
 ⚡ <https://www.linkedin.com/in/oliverxudongshen>  
 ✉ [xudong.shen@u.nus.edu](mailto:xudong.shen@u.nus.edu)  
 📞 (+65) 88296866 or (+86) 18658143615

## EDUCATION

Ph.D. in Integrative Sci. & Eng. (Comput. Sci. Focus), *National Univ. of Singapore* **2019-Present**  
 B.A. in Naval. Arch. & Ocean Eng., *Zhejiang Univ.*, China **2015-2019**

## TECHNICAL SKILLS

### Machine Learning (ML):

Experienced with (i) the complete ML pipeline in Python (numpy, scikit-learn, pandas, PyTorch, HuggingFace,...); (ii) with image (face), sensor (satellite remote sensing), sequential (skeleton time sequence), language, and FinTech (P2P lending) data; and (iii) with vision, NLP, multimodal, and generative models.

### Statistics:

Big data analysis and structural modelling using frequentist and Bayesian approaches using the R (survival, rms,...) and Python (PyMC3,...) stack. Experienced with survival modelling, financial default analysis, and social and network data analysis.

### Development:

Experienced with dev tools Git, Docker, Kubernetes, AWS.

## EXPERIENCE

*Doctoral Researcher, N-CRIPT Centre, NUS, Singapore* **2019-Present**

- (a) Designed and executed research in ML and finance: (i) delivered solutions in Python to mitigate unfairness in large ML systems, e.g., face recognition, (ii) theoretical and empirical modeling of million P2P lending data using Python (PyMC3) & R (survival, rms,...) to estimate default risk, profitability, and discrimination.
- (b) Actively contributed (60+ commits) to community-driven projects on biases and robustness in NLP and AI, such as [BIG-Bench](#), [BigScience](#), and [NL-Augmenter](#).
- (c) Communicated scientific insights to both academic and industry audience in 6 papers (2 first-authored, 3 published in leading journals and conferences) and 3+ talks.
- (d) Assisted teaching the course *Computational Methods for Business Analytics* twice, totally 150+ hours to 200+ undergraduate students. Prepared weekly recap materials and videos, and delivered weekly Q&A sessions.

*Research Intern, Sea AI Lab, Sea Limited, Singapore*

**2022-Present**

- (a) Drive research projects involving 5+ researchers from 2 institutes (NUS & Sea) on bias and safety issues of diffusion-based image generation models, e.g., fair language representation and its impact on image generation.

*Undergraduate Researcher, Second Institute of Oceanography, China*

**Jul. 2017-Jul. 2019**

- (a) Handled the full stack of satellite remote sensing image segmentation using Python & Matlab, incl. data preprocessing, labeling, neural net training and validating.

*Undergraduate Researcher, Waseda University, Japan*

**Sep. 2018-Feb. 2019**

- (a) Analyzed the effectiveness of Japan's policy toward the elimination of Persistent Organic Pollutants (POPs), and proposed policy recommendations in China.

*Undergraduate Researcher, York University, Canada*

**May-Aug. 2018**

- (a) Worked on fluid dynamics simulation (based on the Navier-Stokes PDE) of a hydrodynamically focused printing process for printed electronics, using COMSOL and FLUENT, in collaboration with 5 other researchers.

PREPRINTS	1.Kaustubh D Dhole, Varun Gangal, Sebastian Gehrmann, Aadesh Gupta, ..., <u>Xudong Shen</u> , ...: "NL-Augmenter: A Framework for Task-Sensitive Natural Language Augmentation". In <i>arXiv preprint, arXiv:2112.02721</i> (2021).	
	2.Aarohi Srivastava, Abhinav Rastogi, Abhishek Rao, Abu Awal Md Shoeb, ..., <u>Xudong Shen</u> , ...: "Beyond the Imitation Game: Quantifying and Extrapolating the Capabilities of Language Models". In <i>arXiv preprint, arXiv:2206.04615</i> (2022).	
	3. <u>Xudong Shen</u> , Tianhui Tan, Tuan Q. Phan, Jussi Keppo: "Unwarranted Gender Disparity in Online P2P Lending: Evidence of Affirmative Action". In <i>arXiv preprint, arXiv:2210.07864</i> (2022).	
PUBLICATIONS	1.Ziwei Xu, <u>Xudong Shen</u> , Yongkang Wong, Mohan Kankanhalli: "Unsupervised Motion Representation Learning with Capsule Autoencoders". In <i>NeurIPS</i> (2021).	
	2. <u>Xudong Shen</u> , Yongkang Wong, Mohan Kankanhalli: "Fair Representation: Guaranteeing Approximate Multiple Group Fairness for Unknown Tasks". In <i>IEEE Trans. PAMI</i> (2022).	
	3.Yizhong Wang, Swaroop Mishra, Pegah Alipoormolabashi, Yeganeh Kordi, ..., <u>Xudong Shen</u> , ...: "Benchmarking Generalization via In-Context Instructions on 1,600+ Language Tasks". In <i>EMNLP</i> (2022).	
INVITED TALKS	Workshop on Enormous Language Models, <i>ICLR 2021</i> , Online	2021
	N-CRiPT Seminar, <i>National University of Singapore</i> , Singapore	2022
	<i>International Conference on Smart Finance 2022</i> , Online (Hong Kong)	2022
SERVICE	<i>Referee</i> : EMNLP 2022, ICLR 2023, WACV 2023.	
ADDITIONAL TRAINING	Master Kong Dream Scholarship Program, <i>Waseda University</i> , Japan	Sep. 2018-Feb. 2019
	Globalink Research Internship, <i>York University</i> , Canada	May.-Aug. 2018
	Erasmus+ Student Mobility, <i>Università di Trento</i> , Italy	Feb.-Jun. 2017