# Chirag Nagpal, PhD Al Research Scientist, Meta Superintelligence Labs

- Carnegie Mellon University, Pittsburgh, USA	
Doctor of Philosophy (Language and Information Technologies)	Fall 18 - 23
Survival and Time-to-Event Analysis, Graphical Models, Mixed-Integer Non-Linear Programming	
Committee : Artur Dubrawski (Chair), Bhiksha Raj, Louis-Philippe Morency, Russell Greiner (University of Alberta) and Katherine Heller (Google and Duke University)	
Master of Science (Language Technologies)	Fall 16 - 18
Coursework : Advanced Natural Language Processing, Advanced Multimodal Machine Learning, Neural Networks for Natural Language Processing, Machine Learning for Text Mining	
- Army Institute of Technology, University of Pune, India	Fall 12 - 16
Bachelor of Computer Engineering	
First Position in Class, Chief of Army Staff Gold Medal in Academics	
- Al Research Scientist, Meta GenAl, Menlo Park	Spring 25 -
	Spring 23 - 25
	3pring 23 - 25
Built on-policy RL recipes for Gemma 2 alignment research used by 100+ FTEs.	
- Research Intern Responsible Al Google Research (Remote due to COVID-19)	Spring 22
- ,	<i>5pmg 22</i>
	Summer 20
- ,	Julillier 20
	Summer 19
Bayesian methods to mitigate systemic analyst bias and error in equities forecasts.	
- Science for Social Good Fellow, IBM TJ Watson Research Center, New York	Summer 18
Causal neural networks to recover heterogeneous treatment effects.	
auton-survival: an Open-Source Package for Regression. Counterfactual Estimation.	
Evaluation and Phenotyping with Censored Time-to-Event Data. [link] [github] [docs] [blog]	
Chirag Nagpal, Willa Potosnak, and Artur Dubrawski	
MLHC - Machine Learning for Healthcare Conference '22	
CMU 10-708, Probabilistic Graphical Models	Fall 20
Teaching Assistant for Prof. Pradeep Ravikumar. [webpage]	
CMU 11-761, Language and Statistics	Fall 19
Teaching Assistant for Prof. Bhiksha Raj. [webpage]	
Selected Peer Reviewed Journal, Conference and Symposium Papers	
"Estimation of Language Model Generation Length under a Context Window of Fixed Size" [link] Chirag Nagpal - Unpublished Monograph	
"Preference Models assume Proportional Hazards of Utilities" [link] Chirag Nagpal - Unpublished Monograph	
"Speech Recognition with LLMs Adapted to Disordered Speech using Reinforcement Learning" [1506]	
Chirag Nagpal, Subhashini Venugopalan, Katrin Tomanek and others.	
ICASSP - International Conference on Acoustics, Speech, and Signal Processing '25	
"Helping or Herding? Reward Model Ensembles Mitigate but do not Eliminate Reward Hacking" <a href="Link">[Link</a> Jacob Eisentstein, Chirag Nagpal, Alekh Agarwal, Jonathan Berant and others.  COLM - Conference on Language Modelling '24	<u>1</u>
	Doctor of Philosophy (Language and Information Technologies) Survival and Time-to-Event Analysis, Graphical Models, Mixed-Integer Non-Linear Programming Committee: Artur Dubrawski (Chair), Bhiksha Raj. Louis-Philippe Morency, Russell Greiner (University of Alberta) and Katherine Heller (Google and Duke University)  Master of Science (Language Technologies) Coursework: Advanced Natural Language Processing, Advanced Multimodal Machine Learning, Neural Networks for Natural Language Processing, Machine Learning for Text Mining  - Army Institute of Technology, University of Pune, India Bachelor of Computer Engineering First Position in Class, Chief of Army Staff Gold Medal in Academics  - Al Research Scientist, Meta GenAl, Menlo Park Designed on-policy Rt. recipes and lead safety post-training for Llama 4 Scout, Maverick and other internal product lines. Launches: Off-policy Grounding, Generative Reward Models, False Refusal Suppression, Multilingual and Multimodal Safety.  - Research Scientist, Google Research, San Francisco Foundational Research in LLM Alignment. Contributor to Gemini 2.5. Built on-policy Rt. recipes for Gemma 2 alignment research used by 100+ FTEs.  - Research Intern, Responsible Al, Google Research (Remote due to COVID-19) Algorithmic Fairness in Integer Risk Scoring Systems.  - Research Intern, Google Brain (Remote due to COVID-19) Deep Semi-Parametric Mixtures for calibrated estimation of Time-to-Event.  - Summer Associate, JP Morgan Al Research, New York City Bayasian methods to mitigate systemic analyst bias and error in equities forecasts.  - Science for Social Good Fellow, IBM TJ Watson Research Center, New York Manager: Dr. Kush R. Varshney  Causal neural networks to recover heterogeneous treatment effects.  - Summer Associate, JP Morgan Al Research, New York City Bayasian methods to mitigate systemic analyst bias and error in equities forecasts.  - Science for Social Good Fellow, IBM TJ Watson Research Center, New York Manager: Dr. Kush R. Varshney  Chirag Nagpal, Willia Potosnak, and Ar

"Transforming and Combining Rewards for Aligning Large Language Models" [link] Zihao Wang, Chirag Nagpal, Jonathan Berant, Sanmi Koyejo, Victor Veitch and others.

ICML - International Conference on Machine Learning '24

"Risk-Aware Framework Development for Disruption Prediction : Alcator C-Mod and DIII-D Survival Analysis" [link] Zander Keith, Chirag Nagpal, Cristina Rea and Alex Tinguely.

JFE - Journal of Fusion Energy '24

"Participatory Systems for Personalized Prediction" [link] Hailey James, Chirag Nagpal, Katherine Heller, and Berk Ustun.

NeurIPS - Neural Information Processing Systems '23 (Spotlight Paper)

"Counterfactual Phenotyping with Censored Time-to-Events"  $\underline{{\tt [arXiv]}}$   $\underline{{\tt [code]}}$ 

Chirag Nagpal, Mononito Goswami, Keith Dufendach, and Artur Dubrawski KDD - ACM Conference on Knowledge Discovery and Data Mining '22

"Deep Cox Mixtures for Survival Regression" [arXiv] [code]

Chirag Nagpal, Steve Yadlowsky, Negar Rostamzadeh, and Katherine Heller

MLHC - Machine Learning for Healthcare Conference '21

Taught in Prof. David Sontag's Machine Learning for Health course at MIT and Harvard. [link]

"Deep Survival Machines: Fully Parametric Survival Regression and Representation Learning

for Censored Data with Competing Risks" [arXiv] [code]

Chirag Nagpal, Xinyu (Rachel) Li, and Artur Dubrawski

JBHI - IEEE Journal of Biomedical and Health Informatics '21

Spotlight Presentation at NeurIPS ML for Health Workshop '19, (Top 3% out of over 300 submissions.)

Chirag Nagpal\*, Vincent Jeanselme\*, and Artur Dubrawski

AAAI Spring Symposium - Survival Prediction: Algorithms, Challenges and Application '21

"Interpretable subgroup discovery in treatment effect estimation with application to opioid prescribing guidelines"

Chirag Nagpal, Dennis Wei, Bhanukiran Vinzamuri, Monica Shekhar, Sara E. Berger, Subhro Das, Kush R. Varshney

CHIL - Conference on Health, Inference and Learning '20 [arXiv] [code]

### **Abstracts and Posters at Medical Conferences**

ICCAI '22, "Identification of patients with stable coronary artery disease who benefit from ACE inhibitors using Cox mixture model for heterogeneous treatment effects"

STS Coronary '22, "Novel Machine Learning Technique Defines Patients Who Benefit from Off-Pump CABG"

ISICEM '22, "Phenogrouping of hemorrhagic trauma patients using latent variable machine learning."

CCM '18, "Accuracy of identifying venous thromboembolism by administrative coding compared to manual review."

#### Mentoring

- Fall '22 : Shakirah Cooper, Biomedical Engineering, Carnegie Mellon
- Summer '22 : Mingzhu Liu, BS Michigan at Ann Arbor → MS Robotics, Carnegie Mellon
- Summer '22 : Van H. Le, BS Math and Economics, Hollins University, Virginia
- Fall '21 : Willa Potosnak, BS Duquesne University, PA → Robotics PhD, Carnegie Mellon
- Fall '19 : **Xinyu (Rachel) Li**, MS Information Systems, Heinz College → Robotics PhD, Carnegie Mellon

# SERVICE

## Organization

Co-organizer for the AAAI Spring Symposium on Survival Prediction 2021, 2023

Area Chair

Conferences: MLHC 2024, ML4H 2025

Reviewer

Journals: IEEE Journal of Biomedical and Health Informatics, Journal of Forecasting, Frontiers in Immunology

Conferences: NeurIPS, ICML, ICLR, MLHC, CHIL, ML4H

**Departmental Service** 

Member, School of Computer Sciences Dean's PhD Students Advisory Committee [webpage]

Member, Admissions Committee, Robotics Institute Summer Scholar's Program. [webpage]

Chair, SCS DEC/5, CMU Computer Science Graduate Students Social Organization. [webpage]

Member, International Student's Association, Carnegie Mellon. [webpage]

## PERSONAL

Citizenship: Indian, Languages: English and Hindi

Interests: Equitation, Trivia Quizzing, Squash, Making and DIY, Amateur Radio ( Callsign: VU2CND )