

# WENHU CHEN

Email: (hustchenwenhu@gmail.com, wenhuchen@uwaterloo.ca), Website: <https://wenhuchen.github.io/>

## ABOUT ME

I'm currently an LM researcher at Perplexity AI. I have been a tenure-track assistant professor at Computer Science of University of Waterloo and Vector Institute since Aug 2022. I became a Canadian CIFAR AI chair since Jan 2023. I mainly work on different aspects of foundation models. My research can be summarized as the following aspects:

- 1) **large language models and agents**: post-training, reasoning, coding, agents, etc.
- 2) **multimodal language models**: post-training, multi-image/video understanding, etc.
- 3) **retrieval augmentation**: information retrieval, retrieval-augmented generation, etc.
- 4) **evaluation**: benchmarks, automatic metrics, reward models, etc.

## WORKING EXPERIENCE

### Perplexity AI

11/03/2025-Now

Location: San Francisco, US

Role: Member of Technical Staff

Manager: Tony Wu

Duties: 1. Building search agents to perform high-accuracy and precision information seeking. 2. Post-training open-source models to match ChatGPT on human preferences (Perplexity Sonar).

### Vector Institute

08/22/2022-Now

Location: Toronto, Canada

Role: Faculty Member (On Leave)

Duties: 1. Conducting research in natural language processing. 2. Service in postdoc and faculty review committee. 3. Hosting research events in Vector Institute.

### University of Waterloo

08/22/2022-Now

Location: Waterloo, Canada

Role: Assistant Professor in Cheriton School of Computer Science (On Leave)

Duties: 1. Leading research projects and supervising graduate students. 2. Teaching fundamental courses. 3. Administrative services (i.e. the graduate admission committee, undergrad hiring committee).

### Google Deepmind (Gemini Team)

08/22/2022-07/28/2025

Location: Remote in Canada

Role: Senior Research Scientist (20% Part-Time)

Manager: William W. Cohen & Ming-Wei Chang

Duties: Contributed to Gemini Multimodal and Evaluation efforts. I mainly participated in the Gemini 2.5 and the Nano-Banana project.

### Google Research

08/16/2021-08/19/2022

Location: Pittsburgh, USA

Role: Research Scientist (Full time)

Manager: William W. Cohen

Duties: Building Next-Generation Retrieval-Augmented Systems to enhance factuality.

## EDUCATION

### University of California, Santa Barbara

09/01/2017-06/15/2021

Location: Santa Barbara, USA

Ph.D in Computer Science

Advisor: William (Yang) Wang and Xifeng Yan

**RWTH Aachen University, Germany**

09/01/2014-02/01/2017

Location: Aachen, Germany

M.Sc. in Electronics Engineering

**Huazhong University of Science and Technology, China**

09/01/2010-06/15/2014

Location: Wuhan, China

B.Sc. in Electronics Engineering

**INTERNSHIP EXPERIENCE****Google Research**

June 15th, 2020 - Dec 18th, 2020

Location: Pittsburgh, PA, USA

Mentors: Ming-Wei Chang, Eva Schlinger, William W. Cohen

Duties: Work as research intern to explore research topics. We Build question answering system that can understand both structured and unstructured information. Our work was published in [ICLR 2021].

**Microsoft AI & Cloud**

June 17th, 2019 - Sep 6th, 2019

Location: Seattle, WA, USA

Mentors: Zhe Gan, Jingjing Liu

Duties: Work as research intern to explore research topics. We build explainable visual question answering system that can ground on images to answer simple questions. Our work was published in [WACV 2021].

**Samsung Research America**

June 18th, 2018 - Sep 14th, 2018

Location: Mountain View, CA, USA

Mentors: Yi-Lin Shen, Hongxia Jin

Duties: Work as research intern to explore research topics. We investigate different compression algorithm in natural language systems to save computation and storage. Our work was published in [NAACL 2019].

**Microsoft Research Asia**

Feb 13th, 2017 - Aug 11th, 2017

Location: Beijing, China

Mentors: Shujie Liu, Ming Zhou

Duties: Work as research intern to explore research topics. We enhance the machine translation systems by utilizing data augmentation strategies. Our work was published in [NAACL 2018].

**eBay Research**

Dec 7th, 2015 - May 27th, 2016

Location: Aachen, Germany

Mentors: Matusov Evgency, Shahram Khadivi

Duties: Work as research intern to explore research topics. We enhance machine translation with statistical alignment algorithm. Our work was published in [AMTA 2016].

**TEACHING EXPERIENCE****University of Waterloo, 2023 Winter** : CS 486/686 Introduction to Artificial Intelligence**University of Waterloo, 2024 Winter** : CS 886 Recent Advances on Foundation Models**University of Waterloo, 2024 Spring** : CS 486/686 Introduction to Artificial Intelligence**SELECTED PUBLICATIONS (AS FIRST & LAST AUTHOR)****[TMLR 2026]** VLM2Vec-V2: Advancing Multimodal Embedding for Videos, Images, and Visual Documents  
Rui Meng, Ziyan Jiang, Ye Liu, Mingyi Su, Xinyi Yang, Yuepeng Fu, Can Qin, Zeyuan Chen, Ran Xu, Caiming Xiong, Yingbo Zhou, Wenhua Chen, Semih Yavuz**[TMLR 2025]** BrowserAgent: Building Web Agents with Human-Inspired Web Browsing Actions  
Tao Yu, Zhengbo Zhang, Zhiheng Lyu, Junhao Gong, Hongzhu Yi, Xinming Wang, Yuxuan Zhou, Jiabing Yang, Ping Nie, Yan Huang, Wenhua Chen**[TMLR 2025]** QuickVideo: Real-Time Long Video Understanding with System Algorithm Co-Design  
Benjamin Schneider, Dongfu Jiang, Chao Du, Tianyu Pang, Wenhua Chen

**[TMLR 2025]** StructEval: Benchmarking LLMs' Capabilities to Generate Structural Outputs

Jialin Yang, Dongfu Jiang, Lipeng He, Sherman Siu, Yuxuan Zhang, Disen Liao, Zhuofeng Li, Huaye Zeng, Yiming Jia, Haozhe Wang, Benjamin Schneider, Chi Ruan, Wentao Ma, Zhiheng Lyu, Yifei Wang, Yi Lu, Quy Duc Do, Ziyan Jiang, Ping Nie, Wenhua Chen

**[NeurIPS 2025]** Pixel Reasoner: Incentivizing Pixel-Space Reasoning with Curiosity-Driven Reinforcement Learning  
Alex Su, Haozhe Wang, Weimin Ren, Fangzhen Lin, Wenhua Chen

**[NeurIPS 2025]** General-Reasoner: Advancing LLM Reasoning Across All Domains  
Xueguang Ma, Qian Liu, Dongfu Jiang, Ge Zhang, Zejun Ma, Wenhua Chen

**[NeurIPS 2025]** VL-Rethinker: Incentivizing Self-Reflection of Vision-Language Models with Reinforcement Learning  
Haozhe Wang, Chao Qu, Zuming Huang, Wei Chu, Fangzhen Lin, Wenhua Chen  
**(Spotlight, Top 10% of accepted papers)**

**[NeurIPS 2025]** MoCha: Towards Movie-Grade Talking Character Synthesis

Cong Wei, Bo Sun, Haoyu Ma, Ji Hou, Felix Juefei-Xu, Zecheng He, Xiaoliang Dai, Luxin Zhang, Kunpeng Li, Tingbo Hou, Animesh Sinha, Peter Vajda, Wenhua Chen

**(Spotlight, Top 10% of accepted papers)**

**[TMLR 2025]** PixelWorld: Towards Perceiving Everything as Pixels

Zhiheng Lyu, Xueguang Ma, Wenhua Chen

**[Findings of EMNLP 2025]** VisCoder: Fine-Tuning LLMs for Executable Python Visualization Code Generation  
Yuansheng Ni, Ping Nie, Kai Zou, Xiang Yue, Wenhua Chen

**[EMNLP 2025]** VisualWebInstruct: Scaling up Multimodal Instruction Data through Web Search  
Yiming Jia, Jiachen Li, Xiang Yue, Bo Li, Ping Nie, Kai Zou, Wenhua Chen

**[EMNLP 2025]** Unleashing the Reasoning Potential of Pre-trained LLMs by Critique Fine-Tuning on One Problem  
Yubo Wang, Ping Nie, Kai Zou, Lijun Wu, Wenhua Chen

**[LLM Report 2025]** Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities

Gemini Team

**[TMLR 2025]** ABC: Achieving Better Control of Multimodal Embeddings using VLMs

Benjamin Schneider, Florian Kerschbaum, Wenhua Chen

**[COLM 2025]** ScholarCopilot: Training Large Language Models for Academic Writing with Accurate Citations  
Yubo Wang, Xueguang Ma, Ping Nie, Huaye Zeng, Zhiheng Lyu, Yuxuan Zhang, Benjamin Schneider, Yi Lu, Xiang Yue, Wenhua Chen

**[COLM 2025]** Critique Fine-Tuning: Learning to Critique is More Effective than Learning to Imitate  
Yubo Wang, Xiang Yue, Wenhua Chen

**[ICCV 2025]** Vamba: Understanding Hour-Long Videos with Hybrid Mamba-Transformers  
Weiming Ren, Wentao Ma, Huan Yang, Cong Wei, Ge Zhang, Wenhua Chen

**[ACL 2025]** AceCoder: Acing Coder RL via Automated Test-Case Synthesis  
Huaye Zeng, Dongfu Jiang, Haozhe Wang, Ping Nie, Xiaotong Chen, Wenhua Chen

**[ACL 2025]** TheoremExplainAgent: Towards Video-based Multimodal Explanations for LLM Theorem Understanding  
Max Ku, Thomas Chong, Jonathan Leung, Krish Shah, Alvin Yu, Wenhua Chen

**[CVPR 2025]** Enhancing Long-Duration and High-Resolution Video Understanding by Video SpatioTemporal Augmentation  
Weiming Ren, Huan Yang, Jie Min, Cong Wei, Wenhua Chen

**[TMLR 2025]** Long-context LLMs Struggle with Long In-context Learning  
Tianle Li, Ge Zhang, Quy Duc Do, Xiang Yue, Wenhua Chen

**[ICLR 2025]** OmniEdit: Building Image Editing Generalist Models Through Specialist Supervision  
Cong Wei, Zheyang Xiong, Weiming Ren, Xinrun Du, Ge Zhang, Wenhua Chen

**[ICLR 2025]** VLM2Vec: Training Vision-Language Models for Massive Multimodal Embedding Tasks  
Ziyan Jiang, Rui Meng, Xinyi Yang, Semih Yavuz, Yingbo Zhou, Wenhua Chen

**[ICLR 2025]** MEGA-Bench: Scaling Multimodal Evaluation to over 500 Real-World Tasks  
Jiacheng Chen, Tianhao Liang, Sherman Siu, Zhengqing Wang, Kai Wang, Yubo Wang, Yuansheng Ni, Wang Zhu, Ziyan Jiang, Bohan Lyu, Dongfu Jiang, Xuan He, Yuan Liu, Hexiang Hu, Xiang Yue, Wenhua Chen

**[TMLR 2024]** MANTIS: Interleaved Multi-Image Instruction Tuning  
Dongfu Jiang, Xuan He, Huaye Zeng, Cong Wei, Max Ku, Qian Liu, Wenhua Chen  
**(Best Paper Award)**

**[TMLR 2024]** AnyV2V: A Plug-and-Play Framework For Any Video-to-Video Editing Tasks  
Max Ku, Cong Wei, Weiming Ren, Huan Yang, Wenhua Chen  
**(Reproducibility Certification)**

**[NeurIPS 2024]** MMLU-Pro: A More Robust and Challenging Multi-Task Language Understanding Benchmark  
Yubo Wang, Xueguang Ma, Ge Zhang, Yuansheng Ni, Abhranil Chandra, Shiguang Guo, Weiming Ren, Aaran Arulraj, Xuan He, Ziyan Jiang, Tianle Li, Max Ku, Kai Wang, Alex Zhuang, Rongqi Fan, Xiang Yue, Wenhua Chen  
**(Spotlight, Top 10% of accepted papers)**

**[NeurIPS 2024]** GenAI Arena: An Open Evaluation Platform for Generative Models  
Dongfu Jiang, Max Ku, Tianle Li, Yuansheng Ni, Shizhuo Sun, Rongqi Fan, Wenhua Chen

**[NeurIPS 2024]** MAmmoTH2: Scaling Instructions from the Web  
Xiang Yue, Tuney Zheng, Ge Zhang, Wenhua Chen

**[EMNLP 2024]** Unifying Multimodal Retrieval via Document Screenshot Embedding  
Xueguang Ma, Sheng-Chieh Lin, Minghan Li, Wenhua Chen, Jimmy Lin

**[EMNLP 2024]** VideoScore: Building Automatic Metrics to Simulate Fine-grained Human Feedback for Video Generation  
Xuan He, Dongfu Jiang, Ge Zhang, Max Ku, Achint Soni, Sherman Siu, Haonan Chen, Abhranil Chandra, Ziyan Jiang, Aaran Arulraj, Kai Wang, Quy Duc Do, Yuansheng Ni, Bohan Lyu, Yaswanth Narsupalli, Rongqi Fan, Zhiheng Lyu, Yuchen Lin, Wenhua Chen

**[Findings of EMNLP 2024]** Augmenting Black-box LLMs with Medical Textbooks for Clinical Question Answering  
Yubo Wang, Xueguang Ma, Wenhua Chen

**[COLM 2024]** StructLM: Towards Building Generalist Models for Structured Knowledge Grounding  
Alex Zhuang, Ge Zhang, Tianyu Zheng, Xinrun Du, Junjie Wang, Weiming Ren, Stephen W. Huang, Jie Fu, Xiang Yue, Wenhua Chen

**[ACL 2024]** VIEScore: Towards Explainable Metrics for Conditional Image Synthesis Evaluation  
Max Ku, Dongfu Jiang, Cong Wei, Xiang Yue, Wenhua Chen

**[Findings of ACL 2024]** OpenCodeInterpreter: Integrating Code Generation with Execution and Refinement  
Tianyu Zheng, Ge Zhang, Tianhao Shen, Xueling Liu, Bill Yuchen Lin, Jie Fu, Wenhua Chen, Xiang Yue

**[TMLR 2024]** TIGERScore: Towards Building Explainable Metric for All Text Generation Tasks  
Dongfu Jiang\*, Yishan Li\*, Ge Zhang, Wenhai Huang, Bill Yuchen Lin, Wenhua Chen

**[ECCV 2024]** UniIR: Training and Benchmarking Universal Multimodal Information Retrievers  
Cong Wei, Yang Chen, Haonan Chen, Hexiang Hu, Ge Zhang, Jie Fu, Alan Ritter, Wenhua Chen  
**(Oral, Top 5% of accepted papers)**

**[CVPR 2024]** Instruct-Imagen: Image Generation with Multi-modal Instruction  
Hexiang Hu\*, Kelvin C.K. Chan\*, Yu-Chuan Su\*, Wenhua Chen\*, Yandong Li, Kihyuk Sohn, Yang Zhao, Xue Ben, Boqing Gong, William Cohen, Ming-Wei Chang, Xuhui Jia  
**(Oral, Top 3% of accepted papers)**

**[CVPR 2024]** MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI  
Xiang Yue, Yuansheng Ni, Kai Zhang, Tianyu Zheng, Ruqi Liu, Ge Zhang, Samuel Stevens, Dongfu Jiang, Weiming Ren, Yuxuan Sun, Cong Wei, Botao Yu, Ruibin Yuan, Renliang Sun, Ming Yin, Boyuan Zheng, Zhenzhu Yang, Yibo Liu, Wenhai Huang, Huan Sun, Yu Su, Wenhua Chen  
**(Best Paper Finalist, Top 1% of accepted papers)**

**[ICLR 2024]** ImagenHub: Standardizing the evaluation of conditional image generation models  
Max Ku, Tianle Li, Kai Zhang, Yujie Lu, Xingyu Fu, Wenwen Zhuang, Wenhua Chen

**[ICLR 2024]** MAmmoTH: Building Math Generalist Models through Hybrid Instruction Tuning  
Xiang Yue\*, Xingwei Qu, Ge Zhang, Yao Fu, Wenhao Huang, Huan Sun, Yu Su, Wenhua Chen\*  
(Spotlight, Top 20% of accepted papers)

**[WACV 2024]** Synthesizing Coherent Story with Auto-Regressive Latent Diffusion Models  
Xichen Pan, Pengda Qin, Yuhong Li, Hui Xue, Wenhua Chen  
(Oral, Top 5% of accepted papers)

**[TMLR 2023]** DreamEdit: Subject-driven Image Editing  
Tianle Li, Max Ku, Cong Wei, Wenhua Chen

**[TMLR 2023]** Program of Thoughts Prompting: Disentangling Computation from Reasoning for Numerical Reasoning Tasks  
Wenhua Chen\*, Xueguang Ma\*, Xinyi Wang, William W. Cohen

**[EMNLP 2023]** TheoremQA: A Theorem-driven Question Answering dataset  
Wenhua Chen, Ming Yin, Max Ku, Elaine Wan, Xueguang Ma, Jianyu Xu, Tony Xia, Xinyi Wang, Pan Lu

**[NeurIPS 2023]** Subject-driven Text-to-Image Generation via Apprenticeship Learning  
Wenhua Chen, Hexiang Hu, Yandong Li, Nataniel Ruiz, Xuhui Jia, Ming-Wei Chang, William W. Cohen

**[ACL 2023]** Few-shot In-context Learning on Knowledge Base Question Answering  
Tianle Li, Xueguang Ma, Alex Zhuang, Yu Gu, Yu Su and Wenhua Chen

**[Findings of EACL 2023]** Large Language Models are few(1)-shot Table Reasoners  
Wenhua Chen

**[EACL 2023]** Augmenting Pre-trained Language Models with QA-Memory for Open-Domain Question Answering  
Wenhua Chen, Pat Verga, Michiel de Jong, John Wieting, William Cohen

**[ICLR 2023]** Re-Imagen: Retrieval-Augmented Text-to-Image Generator  
Wenhua Chen, Hexiang Hu, Chitwan Saharia, William W. Cohen

**[EMNLP 2022]** MuRAG: Multimodal Retrieval-Augmented Generator for Open Question Answering over Images and Text  
Wenhua Chen, Hexiang Hu, Xi Chen, Pat Verga, William W. Cohen

**[NeurIPS 2021]** A Dataset for Answering Time-Sensitive Questions  
Wenhua Chen, Xinyi Wang, William Yang Wang

**[EMNLP 2020]** KGPT: Knowledge-Grounded Pre-Training for Data-to-Text Generation  
Wenhua Chen, Yu Su, Xifeng Yan, William Yang Wang.

**[Findings of EMNLP 2020]** HybridQA: A Dataset of Multi-Hop Question Answering over Tabular and Textual Data  
Wenhua Chen, Hanwen Zha, Zhiyu Chen, Wenhan Xiong, Hong Wang, William Wang.

**[WACV 2021]** Meta Module Network for Compositional Visual Reasoning  
Wenhua Chen, Zhe Gan, Linjie Li, Yu Cheng, William Wang, Jingjing Liu.  
(Best Student Paper Honorable Mention, Top 1% of accepted papers)

**[Findings of EMNLP 2020]** HybridQA: A Dataset of Multi-Hop Question Answering over Tabular and Textual Data  
Wenhua Chen, Hanwen Zha, Zhiyu Chen, Wenhan Xiong, Hong Wang, William Wang.

**[ACL 2020]** Logical Natural Language Generation from Open-Domain Tables  
Wenhua Chen, Jianshu Chen, Yu Su, Zhiyu Chen and William Yang Wang.

**[ICLR 2020]** TabFact: A Large-scale Dataset for Table-based Fact Verification  
Wenhua Chen, Hongmin Wang, Jianshu Chen, Yunkai Zhang, Hong Wang, Shiyang Li, Xiyou Zhou and William Yang Wang.

**[ACL 2019]** Semantically Conditioned Dialog Response Generation via Hierarchical Disentangled Self-Attention  
Wenhua Chen, Jianshu Chen, Pengda Qin, Xifeng Yan and William Yang Wang.

**[NAACL 2019]** How Large A Vocabulary Does Text Classification Need? A Variational Approach on Vocabulary Selection  
Wenhua Chen, Yu Su, Yilin Shen, Zhiyu Chen, Xifeng Yan and William Yang Wang.

[EMNLP 2018] XL-NBT: A Cross-lingual Neural Belief Tracking Framework  
Wenhu Chen, Jianshu Chen, Yu Su, Xin Wang, Dong Yu, Xifeng Yan and William Yang Wang.

[ACL 2018] No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling  
Wenhu Chen\*, Xin Wang\*, Yuan-Fang Wang and William Yang Wang.

[NAACL 2018] Variational Knowledge Graph Reasoning  
Wenhu Chen, Wenhan Xiong, William Yang Wang, Xifeng Yan.

[NAACL 2018] Generative Bridging Network in Neural Sequence Prediction  
Wenhu Chen, Guanlin Li, Shuo Ren, Shujie Liu, Zhirui Zhang, Mu Li, Ming Zhou.

## ACADEMIC AWARD

---

2025 TMLR Outstanding Paper Award.  
2025 Math Golden Jubilee Award.  
2024 CVPR Best Paper Finalist.  
2023 ACL-IJCNLP Area Chair Award.  
2022 CIFAR AI Chair.  
2021 WACV Best Student Paper Honorable Mention.  
2019 NeurIPS Outstanding Reviewer Award.

## FUNDINGS

---

CIFAR Solution Network: 2026 - 2028  
JELF-CFI Award: 2025 - 2030  
CIFAR AI Chair Fund: 2022 - 2027  
NSERC Discovery Fund: 2023 - 2028  
Mitacs Accelerate Fund: 2023 - 2025  
CIFAR AI Catalyst Fund: 2024 - 2026  
Canada National Research Council (AI4D): 2024 - 2026  
Canada National Research Council (New Beginning): 2025 - 2026

## SERVICE

---

Program Committee: NAACL, ACL, EMNLP, ICML, NeurIPS, ICLR  
Area Chair: AAAI, EMNLP, ACL, NAACL