Honglei Liu

✉ liuhonglei@gmail.com
🌐 https://honglei.ai/
in https://www.linkedin.com/in/honglei-liu-a53a435a/

Education

2012 – 2017	**Ph.D., Computer Science**
	University of California, Santa Barbara (UCSB)
	Advisor: *Prof. Xifeng Yan*
	Thesis: *Mining Patterns and Networks from Sequence Data*

2010 – 2012	**M.S., Computer Science** *(Graduate with Honors, top 1%)*
	Northeastern University (NEU), China
	Advisors: *Prof. Xiaochun Yang and Prof. Bin Wang*
	Thesis: *Approximate Text Alignment Techniques and Optimization Approaches*
	(Outstanding Graduate Dissertation Award)

2006 – 2010	**B.S., Computer Science** *(Honors Class, top 3%)*
	Northeastern University (NEU), China
	Thesis: *Design and Implementation of an Approximate Querying Approach Based on Berkeley DB*
	(Outstanding Undergraduate Dissertation Award)

Experience

04/2020 – Present	**Research Scientist Manager**, Facebook
02/2020 – 04/2020	**Staff Research Scientist**, Facebook
01/2019 – 02/2020	**Senior Research Scientist**, Facebook
07/2017 – 12/2018	**Research Scientist**, Facebook
06/2016 – 09/2016	**Intern**, Facebook
	Topic: *Indexing and Mining Billions of Time Series*
07/2015 – 09/2015	**Bioinformatics Intern**, Illumina
	Topic: *Fast Specificity Checking for Multiplex PCR Primer Design*

Publications

1. **User Memory Reasoning for Conversational Recommendation**

   Hu Xu, Seungwhan Moon, Honglei Liu, Bing Liu, Pararth Shah, Bing Liu, Philip S. Yu. arXiv preprint arXiv:2006.00184 (2020).

2. **Federated User Representation Learning**

   Duc Bui, Kshitiz Malik, Jack Goetz, Honglei Liu, Seungwhan Moon, Anuj Kumar, Kang G. Shin. arXiv preprint arXiv:1909.12535 (2019).

3. **Active Federated Learning**

   Jack Goetz, Kshitiz Malik, Duc Bui, Seungwhan Moon, Honglei Liu, Anuj Kumar. Workshop on Federated Learning for Data Privacy and Confidentiality at Neural Information Processing Systems (NeurIPS 2019).

4. **Global Textual Relation Embedding for Relational Understanding**

Zhiyu Chen, Hanwen Zha, Honglei Liu, Wenhu Chen, Xifeng Yan, Yu Su. Proc. of the Annual Meeting of the Association for Computational Linguistics (ACL 2019). (Short Paper)

5. **Interpretability of Deep Reinforcement Learning Models in a Conversational System**

   Honglei Liu, Parath Shah, Wenxuan Li, Wenhai Yang, Anuj Kumar. in submission

6. **Explore-Exploit: A Framework for Interactive and Online Learning**

   Honglei Liu, Anuj Kumar, Wenhai Yang, Benoit Dumoulin. Systems for Machine Learning Workshop at Neural Information Processing Systems (NeurIPS 2018).

   Open-sourced as **ReAgent**, a modular, end-to-end platform for building reasoning systems

7. **PoQaa: Text Mining and Knowledge Sharing for Scientific Publications**

   Keqian Li, Ping Zhang, Honglei Liu, Hanwen Zha, Xifeng Yan. Proc. of Int. Conf. on Knowledge Discovery and Data Mining (KDD 2018). (demo)

8. **In Vitro Validation of in Silico Identified Inhibitory Interactions**

   Honglei Liu, Daniel Bridges, Connor Randall, Sara A. Solla, Bian Wu, Paul Hansma, Xifeng Yan, Kenneth S. Kosik, Kristofer Bouchard. Journal of Neuroscience Methods 321 (2019): 39-48.

9. **Global Relation Embedding for Relation Extraction**

   Yu Su*, Honglei Liu*, Semih Yavuz, Izzeddin Gur, Huan Sun, Xifeng Yan. Proc. of the Annual Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2018). (*: Equal Contribution)

10. **Active Learning of Functional Networks from Spike Trains**

    Honglei Liu, Bian Wu. SIAM Int. Conf. on Data Mining (SDM 2017).

11. **Fast Motif Discovery in Short Sequences**

    Honglei Liu, Fangqiu Han, Hongjun Zhou, Xifeng Yan, Kenneth S. Kosik. Proc. of Int. Conf. on Data Engineering (ICDE 2016).

    Software licensed to **SerImmune Inc.** to produce real world value

12. **ALAE: Accelerating Local Alignment with Affine Gap Exactly in Biosequence Databases**

    Xiaochun Yang, Honglei Liu, Bin Wang. Proc. of Int. Conf. on Very Large Data Bases (VLDB 2012).

13. **Approximate Substring Query Algorithms Supporting Local Optimal Matching**

    Honglei Liu, Xiaochun Yang, Bin Wang, Rong Jin. Journal of Frontiers of Computer Science and Technology, 2011.

Patents

1. **Personalized Federated Learning for Assistant Systems**

   US Patent Pending 16/815,990

2. **Active Federated Learning for Assistant Systems**

   US Patent Pending 16/815,960

3. **Methods, Mediums, and Systems for Training a Model**

   US Patent Pending 16/731,321

4. **Methods, Mediums, and Systems for Representing a Model in a Memory of Device**

   US Patent Pending 16/731,345

5. **Methods, Mediums, and Systems for Providing a Model for an End-User Device**

   US Patent Pending 16/731,304

6. **Suppressing Reminders for Assistant Systems**

   US Patent Pending 16/733,044

7. **Smart Assistant Systems**

   US Patent Pending 62/923,342

8. **Interpretability of Deep Reinforcement Learning Models in Assistant Systems**

   US Patent Pending 16/389,769

9. **Systems and Methods for Off-Target Sequence Detection**

   US Patent Pending 20180075186

10. **Biological Sequence Local Comparison Method Capable of Obtaining Complete Solution**

    CN Patent Granted 102750461

11. **An Electric Automobile Battery Replacing Device**

    CN Patent Granted 202089042

Open Source

ReAgent   A modular, end-to-end platform for building reasoning systems. It closes the loop of turning actions into feedback, and feedback into training data for RL and online learning. ReAgent is used at Facebook to drive tens of billions of decisions per day.

PyText   A deep-learning based NLP modeling framework built on PyTorch. PyText is used at Facebook to iterate quickly on new modeling ideas and then seamlessly ship them at scale.

ASC   A fast motif discovery tool that is 10,000 times faster than MEME while preserving the same accuracy. ASC+MEME reduces the running time of MEME from weeks to a few minutes with even better accuracy. ASC was licensed to SerImmune Inc. funded by NIH, illumina, Merck, etc. to find motifs from massive protein sequences generated by modern sequencing techniques.