

Yonghui Wu, Ph.D.

Curriculum Vitae

**The University of Texas
Health Science Center at Houston
School of Biomedical Informatics**

Office Address: 7000 Fannin St, Suite 800
Houston, TX 77030
Office Phone: 1-713-500-3903
E-mail: yonghui.wu@uth.tmc.edu

EDUCATION

- 2010 **Ph.D. in Computer Application Technology** *Harbin Institute of Technology, Harbin, China.*
- 2005 **M.S. in Computer Science and Technology** *Harbin Institute of Technology, Harbin, China.*
- 2003 **B.S. in Computer Science and Technology** *Harbin University of Science and Technology, Harbin China.*

ACADEMIC APPOINTMENTS

- 2016–now **Research Assistant Professor** *School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston TX.*
- 2014–2015 **Research Scientist** *School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston TX.*
- 2012–2014 **Post-doc Research Fellow** *School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston TX.*
- 2010–2012 **Post-doc Research Fellow** *Department of Biomedical Informatics, Vanderbilt University, Nashville TN.*

PROFESSIONAL ACTIVITIES

Award Committee

- 2013 *ICIBM* International Conference on Intelligent Biology and Medicine, Nashville TN, US.

Program Committee

- 2014 *ICBO*, International Conference on Biomedical Ontology, Workshop on Vaccine and Drug Ontology Studies, Houston TX, US.
- 2012–2014 *ICIBM*, International Conference on Intelligent Biology and Medicine, Nashville TN, US.

Journal Reviewer

Journal of the American Medical Informatics Association (JAMIA), Journal of Medical Internet Research (JMIR), Journal of Biomedical Informatics (JBI), Pattern Recognition (PR), Neurocomputing, Neural Computing and Applications, PLoS ONE, BMC Medical Informatics and Decision Making, *DATABASE*, International Journal of Medical Informatics, International Journal of Data Mining and Bioinformatics, Computers in Biology and Medicine.

Conference Reviewer

- 2011–2014 AMIA Fall Symposium.
- 2013–2014 World Congress on Medical and Health Informatics (*MedInfo*).
- 2012–2014 International Conference on Intelligent Biology and Medicine (*ICIBM*).
- 2011 International Conference on Intelligent Computing (*ICIC*).

HONORS/AWARDS

- 2014 **Top-ranked (#1) disorder mention recognition system** *2014 Semantic Evaluation (SemEval) challenge, Task 7 - Analysis of Clinical Text.*
- 2013 **Top-ranked (#1) disorder mention and clinical abbreviation disambiguation system** *2013 ShARe/CLEF eHealth Shared Tasks in Clinical NLP.*
- 2012 **Top-ranked (#1) “temporal relation extraction” system** *2012 i2b2 Clinical NLP challenge.*
- 2012 **Finalist of best paper award (Co-first author), AMIA Translational Bioinformatics Summit.**
 - Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *JAMIA.*
- 2008 **Prominent research award for national 863 program of China, ICRC Lab, Harbin Institute of Technology, China.**
- 2003 **Excellent student leader of Heilongjiang province, China.**

TEACHING EXPERIENCES

- 2015 **Guest Lecture, HI6300 - Advanced Health Information Systems.** UTHealth, School of Biomedical Informatics.
- 2014 **Guest Lecture, HI6330 - BioMedical Natural Language Processing.** UTHealth, School of Biomedical Informatics.
- 2013 **Guest Lecture, HI6330 - BioMedical Natural Language Processing.** UTHealth, School of Biomedical Informatics.

INVITED PRESENTATIONS

- 2015 **AMIA annual symposium, A Study of Neural Word Embeddings for Named Entity Recognition in Clinical Text** San Francisco, US.

- 2013 **AMIA annual symposium**, *Building a Large Clinical Abbreviation Sense Inventory from Discharge Summaries* Washington DC, US.
- 2013 **7th Workshop on Data and text mining in biomedical informatics**, *A prototype application for real-time recognition and disambiguation of clinical abbreviations* San Francisco, US.
- 2012 **AMIA annual symposium** *A comparative study of current clinical natural language processing systems on handling abbreviations in discharge summaries* Chicago Illinois, US.
- 2012 **ICIBM** *Identifying the status of genetic lesions in cancer clinical trial documents using machine learning* Nashville TN, US.
- 2011 **AMIA annual symposium** *Detecting Abbreviations in Discharge Summaries using Machine Learning Methods* Washington DC, US.

RESEARCH PROGRAM

Current Projects:

- 03/01/2013 - **Cancer Prevention & Research Institute of Texas R1307**, *PI – Hua Xu*,
02/28/2018 CPRIT Rising Stars Award - Repurposing Existing Drugs for Cancer Treatment using Electronic Health Records, Role:Co-Investigator .
- 09/29/2014 - **NLM 2R01LM010681-05**, *PI – Hua Xu*, Interactive machine learning methods
09/28/2018 for clinical natural language processing, Role: Co-Investigator.

Completed Projects:

- 05/31/2010 - **NLM R01LM010681**, *PI - Hua Xu*, Real-time Disambiguation of Abbreviations in
5/30/2014 Clinical Notes, Role: Postdoc Research Fellow.
- 04/10/2010 - **SHARP ONC 90TR000401**, *PI - Jiajie Zhang*, National Center for Cognitive
03/31/2014 Informatics and Decision Making in Healthcare SHARP, Role: Postdoc Research Fellow.

RESEARCH INTERESTS

Clinical Natural Language Processing (NLP), Machine Learning, Text/Data Mining, Pharmacovigilance.

PUBLICATIONS

Peer Reviewed Articles

- [1] Jun Xu, Hee-Jin Lee, Jia Zeng, Yonghui Wu, Yaoyun Zhang, Liang-Chin Huang, Amber Johnson, Vijaykumar Holla, Ann M. Bailey, Trevor Cohen, Funda Meric-Bernstam, Elmer Bernstam, Hua Xu. Extracting genetic alteration information for personalized cancer therapy from ClinicalTrials.gov. *J Am Med Inform Assoc* 2016, accepted.

- [2] Wu, Y., Xu, J., Jiang, M., Zhang, Y., Xu, H. A Study of Neural Word Embeddings for Named Entity Recognition in Clinical Text. *AMIA Annu Symp Proc*, 2015, Accepted.
- [3] Yonghui Wu, Jun Xu, Yaoyun Zhang, and Hua Xu. Clinical Abbreviation Disambiguation Using Neural Word Embeddings. *ACL-IJCNLP 2015*, 2015:171.
- [4] Yonghui Wu, Min Jiang, Jianbo Lei, Hua Xu. Named Entity Recognition in Chinese Clinical Text Using Deep Neural Network. *Studies in health technology and informatics*, 2015, 216:624-8.
- [5] Zhang Y, Xu J, Wang J, Wu Y, Parkasam M and Xu H. UTH-CCB@BioCreative V Track 2: Recognizing Chemical Entities in Patents. *Proceedings of the Fifth BioCreative Challenge Evaluation Workshop* 2015:147-148.
- [6] Xu J, Wu Y, Zhang Y, Wang J, Liu R, Wei Q, and Xu H. UTH-CCB@BioCreative V CDR Task: Identifying Chemical-induced Disease Relations in Biomedical Text. *Proceedings of the Fifth BioCreative Challenge Evaluation Workshop* 2015:254-259.
- [7] Xu J, Zhang Y, Wu Y, Wang J, Dong X, and Xu H. Citation Sentiment Analysis in Clinical Trial Papers. *AMIA Annu Symp Proc* 2015, Accepted.
- [8] Xu J, Zhang Y, Wang J, Wu Y, Jiang M, Soysal E, and Xu H. UTH-CCB: The Participation of the SemEval 2015 Challenge Task 14. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)* 2015:311-314.
- [9] Sun J, Zhao M, Jia P, Wang L, Wu Y, Iverson C, Zhou Y, Bowton E, Roden D, Denny J, Aldrich M, Xu H, Zhao Z. Deciphering signaling pathway networks to understand the molecular mechanisms of metformin. *PLoS Computational Biology* 2015, 11(6):e1004202.
- [10] Yonghui Wu, Joshua C. Denny, S. Trent Rosenbloom, Randolph A. Miller, Dario A. Giuse, Min Song, Hua Xu. A Preliminary Study of Clinical Abbreviation Disambiguation in Real Time. *Appl Clin Inform* 2015, 6(2), 364-374.
- [11] Yonghui Wu, Joshua Denny, S. Trent Rosenbloom, Randolph A. Miller, Dario A. Giuse, Min Song, Hua Xu. A prototype application for real-time recognition and disambiguation of clinical abbreviations. In *CIKM, Proceedings of the 7th international workshop on Data and text mining in biomedical informatics (DTMBIO '13)* 2013.
- [12] Buzhou Tang, Yudong Feng, Xiaolong Wang, Yonghui Wu, Yaoyun Zhang, Min Jiang, Jingqi Wang, Hua Xu. A comparison of conditional random fields and structured support vector machines for chemical entity recognition in biomedical literature. *J Cheminform* 2015, 7(Suppl 1):S8.
- [13] Min Jiang, Yonghui Wu, Anushi Shah, Priyanka Priyanka, Joshua C. Denny, Hua Xu. Extracting and standardizing medication information in clinical text - the MedEx-UIMA system. *AMIA Summit on Clinical Research Informatics (CRI)* 2014.

- [14] Yonghui Wu, Buzhou Tang, Min Jiang, Sungrim Moon, Joshua C. Denny, Hua Xu. Clinical Acronym/Abbreviation Normalization using a Hybrid Approach. *Proceedings of CLEF 2013 Evaluation Labs and Workshop* 2013.
- [15] Yaoyun Zhang, Jingqi Wang, Buzhou Tang, Yonghui Wu, Min Jiang, Yukun Chen, Hua Xu UTH_CCB: A Report for SemEval 2014-Task 7 Analysis of Clinical Text. *SemEval Proc* 2014.
- [16] Buzhou Tang, Xiaolong Wang, Yonghui Wu, Min Jiang, Jingqi Wang, Hua Xu. Recognizing Chemical Entities in Biomedical Literature using Conditional Random Fields and Structured Support Vector Machines. *BioCreative Challenge Evaluation Workshop* 2013; 2:70-4.
- [17] Buzhou Tang, Yonghui Wu, Min Jiang, and Hua Xu. Recognizing and Encoding Disorder Concepts in Clinical Text using Machine Learning and Vector Space Model. *Proceedings of CLEF 2013 Evaluation Labs and Workshop* 2013.
- [18] Yonghui Wu, Jianbo Lei, Wei-Qi Wei, Buzhou Tang, Joshua C. Denny, S. Trent Rosenbloom, Ran-dolph A. Miller, Dario A. Giuse, Kai Zheng, Hua Xu. Analyzing Differences between Chinese and English Clinical Text: A Cross-Institution Comparison of Discharge Summaries in Two Languages. *Stud Health Technol Inform* 2013; 192:662-6.
- [19] Buzhou Tang, Hongxin Cao, Yonghui Wu, Min Jiang and Hua Xu. Recognizing clinical entities in hospital discharge summaries using Structural Support Vector Machines with word representation features. *BMC Medical Informatics and Decision Making* 2013; 13(Suppl 1):S1.
- [20] Buzhou Tang, Yonghui Wu, Min Jiang, Yukun Chen, Joshua C Denny, Hua Xu. A hybrid system for temporal information extraction from clinical text. *J Am Med Inform Assoc* 2013 Sep-Oct;20(5):828-35.
- [21] Buzhou Tang, Hongxin Cao, Yonghui Wu, Min Jiang, Hua Xu. Clinical Entity Recognition using Structural Support Vector Machines with Rich Features. *Proceedings of the ACM sixth international workshop on Data and text mining in biomedical informatics* 2012:13-20.
- [22] Yonghui Wu, Joshua C. Denny, S. Trent Rosenbloom, Ran-dolph A. Miller, Dario A. Giuse, Hua Xu. A comparative study of current clinical natural language processing systems on handling abbreviations in discharge summaries. *AMIA Annu Symp Proc* 2012:997-1003.
- [23] Yonghui Wu, Mia A Levy, Christine M Micheel, Paul Yeh, Buzhou Tang, Michael J Cantrell, Stacy M Cooreman and Hua Xu. Identifying the status of genetic lesions in cancer clinical trial documents using machine learning. *BMC Genomics* 2012;13(Suppl 8):S21.
- [24] Mei Liu, Yonghui Wu, Yukun Chen, Jingchun Sun, Zhongming Zhao, Xue-wen Chen, Michael Edwin Matheny, Hua Xu. Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. *J Am Med Inform Assoc* 2012 Jun 1;19(e1):e28-e35, **co-first author**.

- [25] Yonghui Wu, Mei Liu, W. Jim Zheng, Zhongming Zhao, and Hua Xu. Ranking gene-drug relationships in biomedical literature using latent dirichlet allocation. *Pac Symp Biocomput* 2012.
- [26] Hua Xu, Yonghui Wu, Noémie Elhadad, Peter D. Stetson, Carol Friedman. A new clustering method for detecting rare senses of abbreviations in clinical notes. *J Biomed Inform* 2012 Dec;45(6):1075-83.
- [27] Mei Liu, Michael E Matheny, Yonghui Wu, ERM Hinz, Joshua C Denny, Jonathan S Schildcrout, Randolph A Miller, Hua Xu. Detecting Adverse Drug Reactions Using Inpatient Medication Orders and Laboratory Tests Data, *IEEE Second International Conference on Healthcare Informatics, Imaging and Systems Biology (HISB)* 2012.
- [28] Jingchun Sun, Yonghui Wu, Hua Xu, and Zhongming Zhao. DTome: a web-based tool for drug-target interactome construction *BMC Bioinformatics* 2012;13(Suppl 9):S7.
- [29] Yonghui Wu, S. Trent Rosenbloom, Joshua C. Denny, Randolph A. Miller, Subramani Mani, Dario A. Giuse, Hua Xu. Detecting Abbreviations in Discharge Summaries using Machine Learning Methods. *AMIA Annu Symp Proc* 2011:1541-9.
- [30] Yonghui Wu, Xiaolong Wang, Yuxin Ding, Jun Xu. Adaptive On-line Web Topic Detection Method for Web News Recommendation System. *Acta Sinica Electronica* 2010;38(11):2620-24.
- [31] Yonghui Wu, Xiaolong Wang, Yuxin Ding, Jun Xu. Topic based Automatic News Recommendation Using Topic Model and Affinity Propagation. *IEEE ICMLC* 2010:1299–1304.
- [32] Yonghui Wu, Yuxin Ding, Xiaolong Wang, and Jun Xu. A Comparative Study of Topic Models for Topic Clustering of Chinese Web News. *IEEE, International Conference on Computer Science and Information Technology* 2010.
- [33] Yonghui Wu, Yuxin Ding, Xiaolong Wang, and Jun Xu. Topic Detection by Topic Model Induced Distance Using Biased Initiation. *Lecture Notes in Computer Science/AST* 2010:310–323.
- [34] Yonghui Wu, Yuxin Ding, Xiaolong Wang, and Jun Xu. On-line Hot Topic Recommendation Using Tolerance Rough Set Based Topic Clustering. *Journal of Computers* 2010;5(4):549-56.
- [35] Yonghui Wu, Yuxin Ding, Xiaolong Wang, and Jun Xu. Design and Implementation of a C/S Structured Distributed Multi-thread Crawler Using ACE and PVFS. *Journal of Computational Information System* 2008;4(3):883–90.
- [36] Hongzhi Guo, Qingcai Chen, Xiaolong Wang, Zhiyong Wang, Yonghui Wu. STRank: A SiteRank algorithm using semantic relevance and time frequency. *IEEE Systems, Man and Cybernetics* 2009:4876-4881.
- [37] Jun Xu, Yuxin Ding, Xiaolong Wang, and Yonghui Wu. Genre identification of Chinese finance text using machine learning method. *Proceedings of the 2008 IEEE International Conference on Systems, Man, and Cybernetics, Singapore* October 2008.

- [38] Jun Xu, Yuxin Ding, Xiaolong Wang, and Yonghui Wu. Bayesian Approaches to Genre Identification of Chinese Finance Text. *Journal of Computational Information System*
- [39] Yuxin Ding, Xiaolong Wang, Lebin Lin, Qi Zhang, Yonghui Wu. The Design and Implementation of The Crawler-Inar. *IEEE ICMLC 2006*.

Abstracts/Posters:

- [1] Yonghui Wu, Adam Wright, Hua Xu, Allison B. McCoy, Dean F. Sittig. A Computable Problem-Medication Knowledgebase based on Electronic Health Records. *AMIA Annu Symp Proc 2014*, Poster paper.
- [2] Yaoyun Zhang, Buzhou Tang, Min Jiang, Jingqi Wang, Yonghui Wu, Hua Xu. Domain Adaptation for Semantic Role Labeling of Clinical Text. *AMIA Annu Symp Proc 2014*, Poster paper.
- [3] Yonghui Wu, S Trent Rosenbloom, Joshua C. Denny, Randolph Miller, Dario Giuse, Hua Xu. Building a Large Clinical Abbreviation Sense Inventory from Discharge Summaries. *AMIA Annu Symp Proc 2013*, Podium paper with oral presentation.
- [4] Buzhou Tang, Yonghui Wu, Min Jiang, Hua Xu. Clinical Entity Recognition Using Structural Support Vector Machines. *AMIA Annu Symp Proc 2012*, Poster paper.
- [5] Anushi Shah, Min Jiang, Yonghui Wu, Joshua C. Denny, Hua Xu. MedEx-UIMA - An Open-Source System for Medication Information Extraction from Clinical Text. *AMIA Annu Symp Proc 2012*, Poster paper.