

# XI CHEN

xchen4@umass.edu (citations 200+)  $\diamond$  <https://www.linkedin.com/in/xi-chen-91a5ab221/>

## EDUCATION

<b>University of Massachusetts Amherst</b>	<b>Amherst, MA</b>
• PhD in Computer Science (major GPA: 3.9/4)	2019-present
<b>Central South University</b>	<b>Changsha, China</b>
• B.E. in Computer Science (GPA 3.9/4)	2014-2018

## EXPERIENCE

<b>University of Massachusetts Amherst</b>	<b>Amherst, MA</b>
<b>Research Assistant</b>	2019.9 - present
<ul style="list-style-type: none"><li>• Refined a state-of-art transformer model to analyze global news graph with trillions of edges.</li><li>• Quantified country effects on news similarity with greedy factor selection, event clustering, and regression.</li><li>• Developed a large dataset of multilingual news article pairs with human-in-the-loop active learning framework.</li></ul>	
<b>Zhipu AI</b>	<b>Beijing, China</b>
<b>Research Intern</b>	2018.3 - 2018.9
<ul style="list-style-type: none"><li>• Disambiguated scholars with heterogeneous random forest (deployed on prestigious AI platform Aminer).</li><li>• Predicted social links by probabilistic graph time series with psychological latent variable.</li></ul>	
<b>Briup Technology corporation</b>	<b>Jiangsu, China</b>
<b>Software Engineer Intern</b>	2017.6 - 2017.8
<ul style="list-style-type: none"><li>• Developed a server-client smart traffic system on sand table as to route optimization.</li></ul>	

## PROJECT

<b>LLM learning for global disaster news</b>
<ul style="list-style-type: none"><li>• Optimize prompts for GPT4's precision and recall on identifying global news covering disasters.</li><li>• Built regression models to quantify country factors with greedy feature selection and event clustering.</li></ul>
<b>Large-scale Multilingual News Graph Learning</b>
<ul style="list-style-type: none"><li>• Extract cross-lingual name entities with Spacy, Polyglot, and Wikipedia concepts.</li><li>• Recall, pre-rank, and rank similar news article pairs based on rules and active learning classifiers.</li><li>• Predict global news network for trillions of pairs with refined multilingual transformer and bi-directional index.</li></ul>
<b>Personalized Opinion Graph Time Series Learning</b>
<ul style="list-style-type: none"><li>• Model the dynamics of opinions and personalities with probabilistic graph time series.</li><li>• Devise an extended EM algorithm for learning the graph time series and get competitive performance.</li></ul>
<b>Symbolic Math Expression Learning</b>
<ul style="list-style-type: none"><li>• Built a sequence-to-sequence model to compute partial differential of formula.</li></ul>

## SELECTED PUBLICATION

<b>International News Synchronization during Covid-19 Pandemic</b>
<b>Xi Chen</b> , Scott Hale, David Jurgens, Mattia Samory, Przemyslaw Grabowicz. International World Wide Web Conference (The WebConf), 2023.
<b>A Multilingual News Similarity Dataset for Media Network and Bias Research</b>
<b>Xi Chen</b> , Scott Hale, David Jurgens, Mattia Samory, Przemyslaw Grabowicz. International AAAI Conference on Web and Social Media (ICWSM), 2023.
<b>Multilingual Document-level Similarity</b>
<b>Xi Chen</b> , Ali Zeynali, Chico Camargo, Fabian Flock, Devin Gaffney, Przemyslaw Grabowicz, Scott Hale, David Jurgens, Mattia Samory. International Workshop on Semantic Evaluation at NAACL (SEMEVAL), 2022.

## SKILL

**PhD Courses:** Machine Learning, Optimization, Natural Language Processing, Database, Algorithm  
**Programming Language and framework:** Python, C/C++, R, Pytorch, Spark, Java, SQL, Linux, Matlab

## HONOR

Orals at top-tier computer science conferences (WebConf 2023, NAACL 2022, TADA 2023, IC2S2 2022&2023).  
Nomination for Chinese Exceptional Student (1/6500 ISchool undergrads and grads), 2018.  
Ranked global Top50 of Autochess players (50/millions).