Chengkai Li

Associate Professor and Associate Chair, CSE
Director, Innovative Database and Information Systems Research (IDIR) Lab
Ph.D. (University of Illinois at Urbana-Champaign, 2007)

Areas: Big Data Intelligence and Data Science (Database, Data Mining, Web, Natural Language Processing)

Projects:

o data-driven fact-checking, computational journalism **(global leading position, well-known in both computing and professional fact-checking communities)**

o cybersecurity, fake-news detection

o graph database usability, knowledge databases

Funding: $3M (including 4 major NSF grants), $1.6M as Project Lead

Media Coverage: AP Associated Press The Washington Post theguardian Austin American-Statesman 5 and many more
Poynter. WIRED The Dallas Morning News KERA

Achievements: paper awards at top conferences, HP Labs innovation research award, SXSW 2017 panelist, keynote/invited speaker at conferences, general/program chair/organizer of conferences, associate editor of journals

idir.uta.edu

IDIR Students and Collaborators

Current students: **7 Ph.D.**, **3 M.S.**, **2 B.S.**

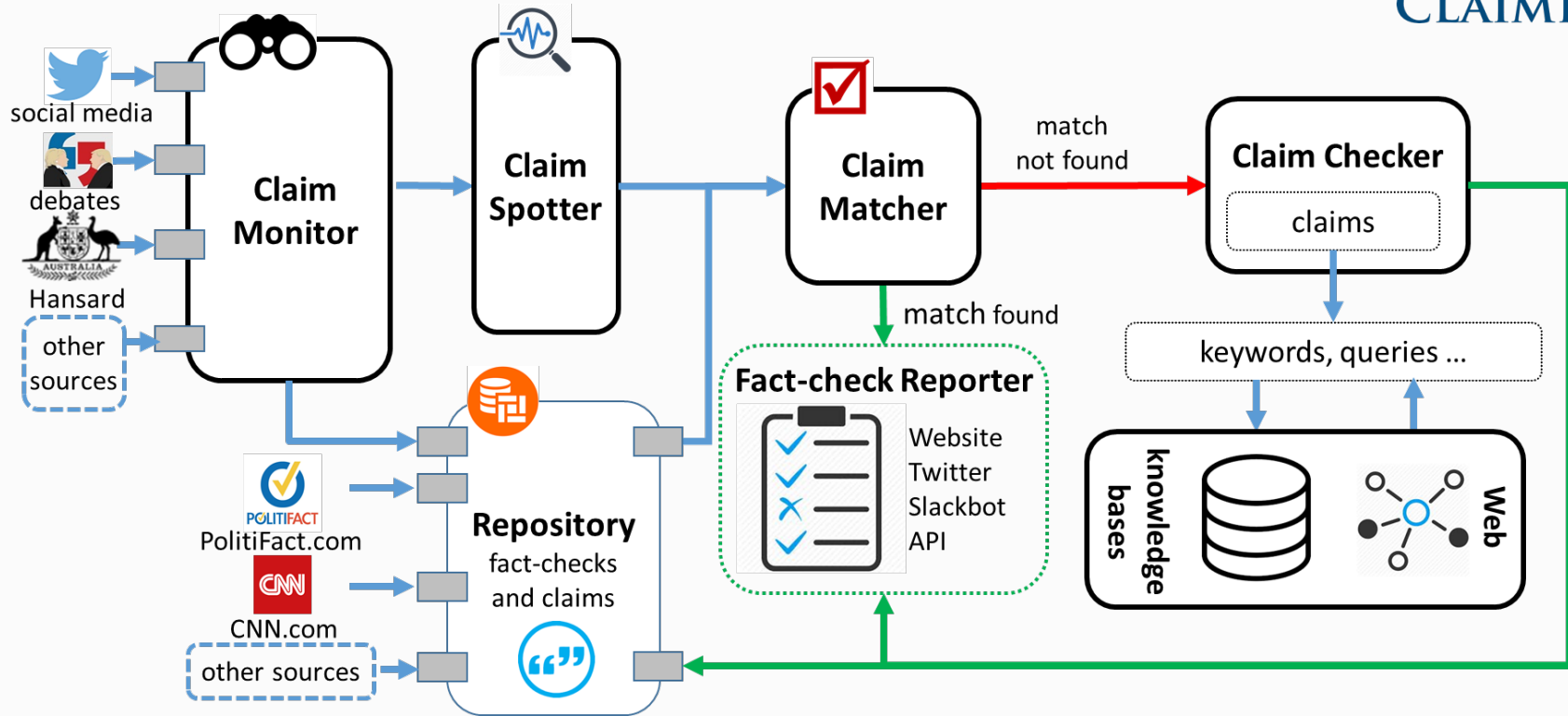Graduated students: **4 Ph.D., 18 M.S.**, **9 B.S.**

Student internships: Amazon A9 Labs, AT&T, Google, HP Labs, Huawei, IBM Research India, NEC Labs, Qatar Computing Research Institute (QCRI), RedHat

First-employment after graduation: **Tenure-track assistant professor :** University of Mississippi; **Researchers:** Huawei US R&D Center, IBM Research India; **Software engineers:** Akamai, Amazon, Ambit Energy, Capital One, Cerner, Copper Labs, Dematic, Electronic Arts (EA) , Google, Microsoft, Pivotal

Current & past collaborators: Duke, George Washington U., Google, HP Labs, Mississippi, Stanford, UCSB, UTD, Chinese Academy of Sciences, Hong Kong Baptist University, Nanjing University, Nanyang Technological University, University of Technology Sydney

The First-ever End-to-end Fact-checking System

idir.uta.edu/claimbuster

Exceptional Fact Finding

FactWatcher (idir.uta.edu/factwatcher): Discovering and monitoring number-based facts pertinent to real-world events (sports, transportation, crime, weather, finance, social media)

- Situational Facts: "No other player scored more pts and reb against DAL than Jordan."
- One-of-the-Few: "Jordan scored 10 pts & 10 reb. Only 3 others have similar performance."
- Prominent Streaks: "The Nikkei 225 closed below 10000 for the 12th consecutive week, the longest such streak since June 2009."

Maverick: Framework and algorithms for discovering non-numeric exceptional facts about entities in knowledge graphs

- Promoting entities: "Denzel Washington followed Sidney Poitier as only the second black to win the Best Actor award."
- Identifying data errors: "Hillary Clinton becomes first female presidential nominee."

Tackling Graph Database Usability Challenges

Challenges

o Large, complex and schema-less graphs capturing millions of entities and billions of relationships between entities.

o Requires substantial understanding of schema and data and complex data cleaning and pre-processing, before one can fetch information or gain insights from data.

Objectives

o Make it easy to understand, query, explore, and clean graph data.

Systems

o GQBE (idir.uta.edu/gqbe): graph query by example

o Orion (idir.uta.edu/orion): auto-suggestion for interactive graph query formulation

o TableView: generating preview tables for knowledge graphs

o Maverick: finding outliers and errors in graphs