



Summary & Future Directions

KDD 2020 Tutorial

Embedding-Driven Multi-Dimensional Topic Mining and Text Analysis

Yu Meng, Jiaxin Huang, Jiawei Han

Computer Science, University of Illinois at Urbana-Champaign

August 23, 2020

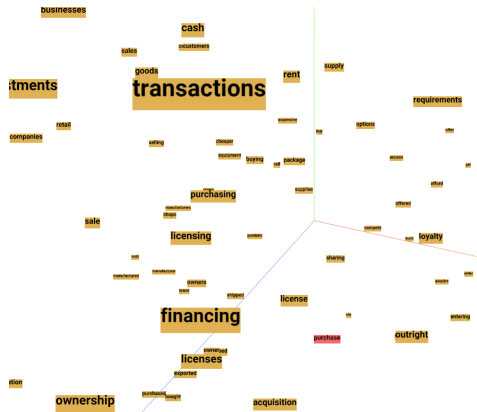


Our Roadmap of This Tutorial

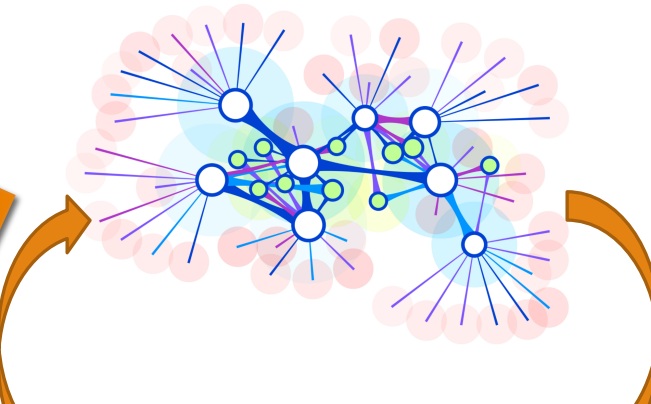
Text Corpus



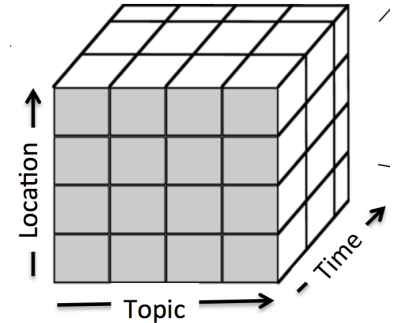
Part I: Text Embedding



Part II: Taxonomy Construction



Part IV: Multi-Dimensional Text Analysis



Part III: User-Guided Topic Mining



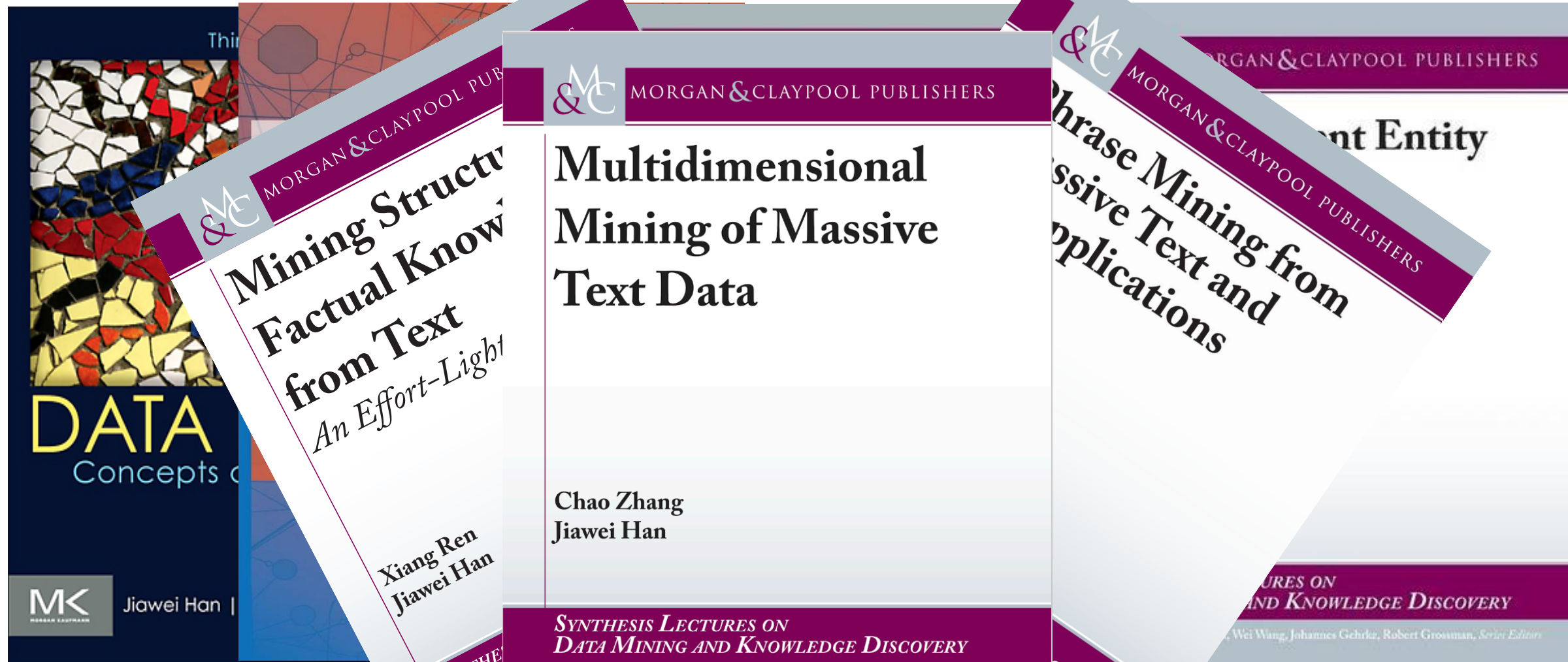
Existing KB

Knowledge

Summary: from Unstructured Text to Knowledge

- ❑ Power of Embedding at Transforming Unstructured Text into Structured Knowledge
- ❑ Mining Structures from Massive Unstructured Text (Texts → Structures)
 - ❑ Automated Text Representation Learning
 - ❑ Automated Multi-Faceted Taxonomy Construction
 - ❑ Automated Topic Mining
 - ❑ Automated Document Assignment in Multidimensional Text Cube
 - ❑ Automated Comparative Summarization in Multidimensional Text Cube
- ❑ Still a lot of work to do from unstructured text to structured knowledge

Our Journey: From Big Data to Big Structures & Knowledge



Han, Kamber and Pei,
Data Mining, 3rd ed. 2011

Yu, Han
Link

*SYNTHESIS LECTURES ON
DATA MINING AND KNOWLEDGE DISCOVERY*

Jiawei Han, Lise Getoor, Wei Wang, Johannes Gehrke, Robert Grossman, *Series Editors*

Information Networks, 2012

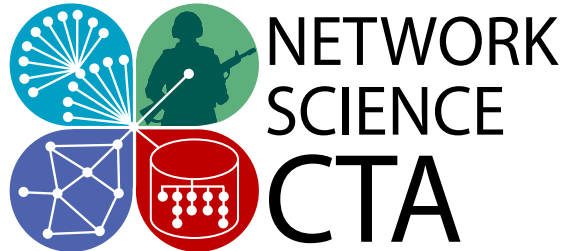
Y. Sun: SIGKDD'13 Dissertation Award

Jiawei Han and Han, Mining Latent Entity
Structures, 2015

Wang: SIGKDD'15 Dissertation Award

Acknowledgement

- Thanks for the research support from: ARL/NSCTA, NIH, NSF, DHS, ARO, DTRA





Thank you !
Q&A

