EPSTEIN INSTITUTE SEMINAR ▪ ISE 651

Building Domain-Specific Search Engines
for Investigative Decision Support

ABSTRACT – With the advent of the Web, the barrier-to-entry for illicit activities like human trafficking, securities fraud and narcotics has been lowered. Building advanced decision support systems that help resource-strapped investigative agencies, such as law enforcement, but also organizations like the Securities and Exchange Commission, investigate and generate promising leads is an important problem with potential for social impact. In this talk, I will discuss an architecture called the Domain-specific Insight Graphs (DIG) that allows non-technical domain experts like police and fraud officers to construct personalized fine-grained search engines, with no programming, over large raw corpora comprising webpages, tables and documents. DIG uses advanced Artificial Intelligence techniques to semi-automatically structure the corpus into a knowledge graph, where nodes are entities and attributes, and labeled edges express entity-entity and entity-attribute relationships. The DIG search engine supports rich capabilities, including facets, maps, timelines, images, and even dossier generation. DIG is completely open-source and can be set up on a laptop for arbitrary domains. Currently, it is being used by over 200 law enforcement to fight human trafficking in the US, and has successfully been used in human trafficking prosecutions in the last year. The architecture has since been used to support other decision-centric applications such as geopolitical forecasting, event mining, academic literature exploration and social media analysis.

Dr. Mayank Kejriwal
Research Scientist
USC Information Sciences Institute

SPEAKER BIO – Mayank Kejriwal is a Research Scientist at the USC Information Sciences Institute. He received his Masters and PhD from the University of Texas at Austin in 2016 under the supervision of Daniel P. Miranker, and holds three undergraduate degrees in Engineering Physics, Computer Engineering, and Banking and Finance. His dissertation, titled 'Populating a Linked Data Entity Name System', was awarded the Best Dissertation Award in 2017 by the Semantic Web Science Association and was also published as a book in the 'Studies in the Semantic Web' series. His work on large-scale information integration has been published in multiple IEEE, ACM, Springer and Elsevier academic venues, including IEEE Intelligent Systems, IEEE Transactions on Big Data, and Journal of Web Semantics. He has delivered tutorials and demonstrations at KDD, AAAI and ISWC, and was recently a winning recipient of the Allen Institute for Artificial Intelligence's Key Scientific Challenges Proposal. Currently, he is co-authoring a textbook on knowledge graphs (MIT Press, 2018) with Pedro Szekely and Craig Knoblock.

USCViterbi
School of Engineering
Daniel J. Epstein Department of
Industrial and Systems Engineering

TUESDAY, JANUARY 16, 2018
3:30PM – 4:50PM
USC ANDRUS GERONTOLOGY CENTER (GER), ROOM 206