세계 석학과 함께하는 빅데이터 세미나
Big Data Seminar with International Scholars

- Building the Dresden Web Table Corpus

Wolfgang Lehner
Technische Universität Dresden

2:30PM-5:00PM, Oct 12, 2015
Room 309, Building 302

Abstract:
WebTables reflect a rich source of information if properly extracted, cleaned, and enriched. There are many applications which may benefit from the compact and relational-like data set coming out of the web. Within this talk, we will - in a first step - outline potential applications for WebTables and then discuss different WebTables classification approaches to turn billions of raw tables into usable and information rich data sets, which may directly be used in query answering. In a second step, the talk will outline the practical side of building the Dresden Web Table Corpus (DWTC), which is freely and publicly accessible for other research communities. The project represents a nice example of working with big data from an educational perspective and shows that big data research can be done effectively and efficiently within an academic environment.

Bio.

Wolfgang Lehner is full professor and head of the database technology group at the TU Dresden, Germany. His research is dedicated to database system architecture specifically looking at crosscutting aspects from algorithms down to hardware-related aspects in main-memory centric settings. He is part of TU Dresden's excellence cluster with research topics in energy-aware computing, resilient data structures on unreliable hardware, and orchestration of heterogeneous systems; he is also a principal investigator of Germany's national "Competence Center for Scalable Data Services and Solutions" (ScaDS); Wolfgang serves the community in many PCs, is an elected member of the VLDB Endowment, serves on the review board of the German Research Foundation (DFG), and is an appointed member of the Academy of Europe.

Contact Info. :
- Prof. Cha, Sang Kyun (880-4163 / chask@snu.ac.kr)

빅데이터연구원 | 서울대학교

세계 석학과 함께하는 빅데이터 세미나
Big Data Seminar with International Scholars

- On Fusing Heterogeneous Data Sources

Philip S. Yu
University of Ilinois at Chicago

2:30PM-5:00PM, Oct 12, 2015
Room 309, Building 302

Abstract

The problem of big data has become increasingly importance in recent years. On the one hand, the big data is an asset that potentially can offer tremendous value or reward to the data owner. On the other hand, it poses tremendous challenges to distil the value out of the big data. The very nature of the big data poses challenges not only due to its volume, and velocity of being generated, but also its variety, where variety means the data can be collected from various sources with different formats from structured data to text to network/graph data, etc. In this talk, we focus on the variety issue and discuss the recent development in fusing information from multiple data sources, which can be applied to multiple disciplines, including social network analysis, World-Wide Web, database systems, data mining, machine learning, and networked communication and information systems. Fusion of multiple social networks will also be considered.

Bio.

Dr. Philip S. Yu is a Distinguished Professor and the Wexler Chair in Information Technology at the Department of Computer Science, University of Illinois at Chicago. Before joining UIC, he was at the IBM Watson Research Center, where he built a world-renowned data mining and database department. He is a Fellow of ACM and IEEE. Dr. Yu is the recipient of IEEE Computer Society's 2013 Technical Achievement Award for "pioneering and fundamentally innovative contributions to the scalable indexing, querying, searching, mining and anonymization of big data". Dr. Yu has published more than 910 referred conference and journals papers cited more than 63,000 times with an H-index of 117. He has applied for more than 300 patents.

Dr. Yu is the Editor-in-Chief of ACM Transactions on Knowledge Discovery from Data. He is on the steering committee of the IEEE Conference on Data Mining and ACM Conference on Information and Knowledge Management and was a member of the IEEE Data Engineering steering committee. He was the Editor-in-Chief of IEEE Transactions on Knowledge and Data Engineering (2001-2004). He received a Research Contributions Award from IEEE Intl. Conference on Data Mining (ICDM) in 2003, the ICDM 2013 10-year Highest-Impact Paper Award, and the EDBT Test of Time Award (2014). Dr. Yu received his PhD from Stanford University.

Contact Info. :
- Prof. Cha, Sang Kyun (880-4163 / chask@snu.ac.kr)

빅데이터연구원 | 서울대학교