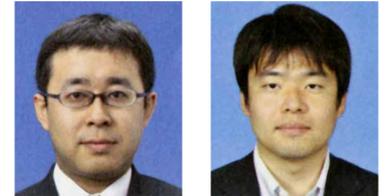# M.TOYODA, NEMOTO, YOSHINAGA, GODA LAB.

Platform for Processing and Visualizing Real World and Social Big Data

### Department of Informatics and Electronics

Interactive Data Analysis, Earth Observation Data Engineering, Natural Language Processing, Data Platform Engineering
Department of Information and Communication Engineering, Graduate School of Information Science and Technology   https://www.tkl.iis.u-tokyo.ac.jp/

Our Lab has focused on completely novel and surprising research on system software, advanced applications, hardware and algorithms, which is the basis of the technology for handling large quantities of data. We have been developing high performance database engine based on a novel out-of-order execution principle, ultra-large-scale cyber space mining systems, 100 petabyte-scale global environment information system based on database engineering, and natural language processing for understanding society trend and our thought.

Development of High Performance Database Engine for Big Data Era

Our research group has been developing Ultrafast Database Engine based on a novel out-of-order execution principle. This new database engine has the capability to achieve significant performance boosting for analytical queries in the Big Data era. The group has also been constructing an experimental system of next-generation strategic social services (cyber physical services) to clarify effectiveness of the new database engine.
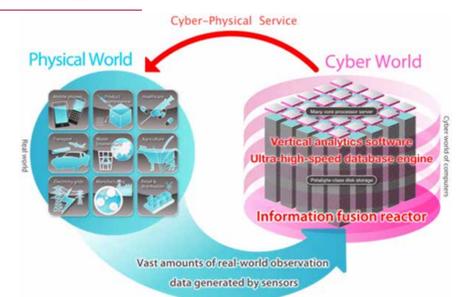
Ultra-large-scale Cyber-Physical Space Analysis System

Our lab has continuously collected Japanese Web pages since 1999 and has constructed a Web archive system including about tens of billions of URLs, billions of blog articles, and tens of billions of tweets. Based on this peta scale archive, we are developing structural, contents, and temporal analysis systems including information diffusion extraction, inter-media comparison, and real-time deep text analysis. The cyber space information is integrated with physical space information such as mobility data for traffic analysis. Results of analysis can be interactively visualized on a large-scale high-resolution display wall.

Huge-scale spatio-temporal visualization system on the display wall

## 100 Petabyte Scale Global Environmental Information Fusion

We have been developing a large scale global environmental information fusion system for data integration and analysis that includes the supporting functions of life cycle data management, data search, information exploration, scientific analysis, and partial data down-loading.

Global environmental information storage

Natural Language Processing for Reading Social Bigdata

Aiming to understand human mind and the real world from text and aid our communication and language understanding, we study on natural language processing that process text efficiently and accurately with computers. The pursuit of such technologies leads to computational linguistics that reveals the mechanism of languages and human intelligence. We have so far realized trend analysis in accumulating social big data. We are now implementing our chatbot on smart speakers for supporting healthcare of the elderly.
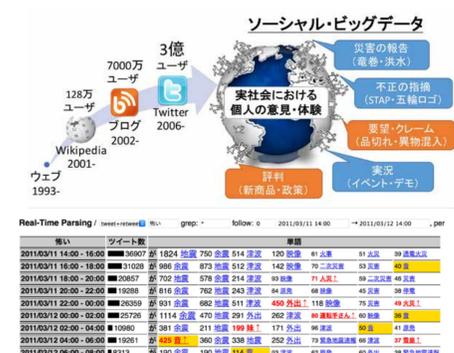
Stream analysis during the 2011 Tohoku earthquake and tsunami