DBMS IMPLEMENTATION - CSE510

K. Selçuk Candan

Name: K. Selçuk Candan

- Professor of computer science and engineering at (SCAI) ASU
- ACM Distinguished Scientist
- Director, ASU Center for Assured and Scalable Data Engineering (CASCADE)
- Director, Enterprise, Media, and Information Technologies Labs (EmitLab)
  - ~10 PhD/MS students (both in US, Italy)

- publications
  - 200+ conference and journal articles + book chapters
  - 9 patents
  - 1 text book

What do I do??

- Executive Committee member, ACM Special Interest Group on Management of Data (SIGMOD)
- Steering Committee member, ACM Int. Conf on Multimedia Retrieval (ICMR)
- Associate editor, IEEE Transactions on Multimedia
- Associate editor, ACM Transactions on Database Systems
- Associate Editor, IEEE Transactions on Knowledge and Data Engineering (2016-2023)
- Associate Editor, IEEE Transactions on Cloud Computing (2016-2021)
- Associate Editor, Proceedings of the VLDB (2019,2021,2022, 2023, 2024)
- Associate Editor, the Very Large Data Bases journal (2005-2012)
- Associate editor, IEEE Transactions on Multimedia (2012-2016)
- Associate editor, Journal of Multimedia (2010-2015)
- Program Chair, ACM SIG Management of Data (SIGMOD) Conference 2023
- General Chair, ACM International Conference on Web Search and Data Mining (WSDM) 2022
- General Chair, IEEE International Conference on Smart Data Services (SMDS) 2020,2021
- Program Chair, ACM Int. Conference on Multimedia Retrieval 2019
- General Chair, IEEE International Conference on Cloud Engineering (IC2E) 2015.
- Publicity Chair, ACM SIG Multimedia Conference 2012
- General Chair, ACM SIGMOD Conference 2012
- General Chair, ACM SIG Multimedia Conference 2011
- PC Chair, the ACM International Conference on Image and Video Retrieval (CIVR) 2010

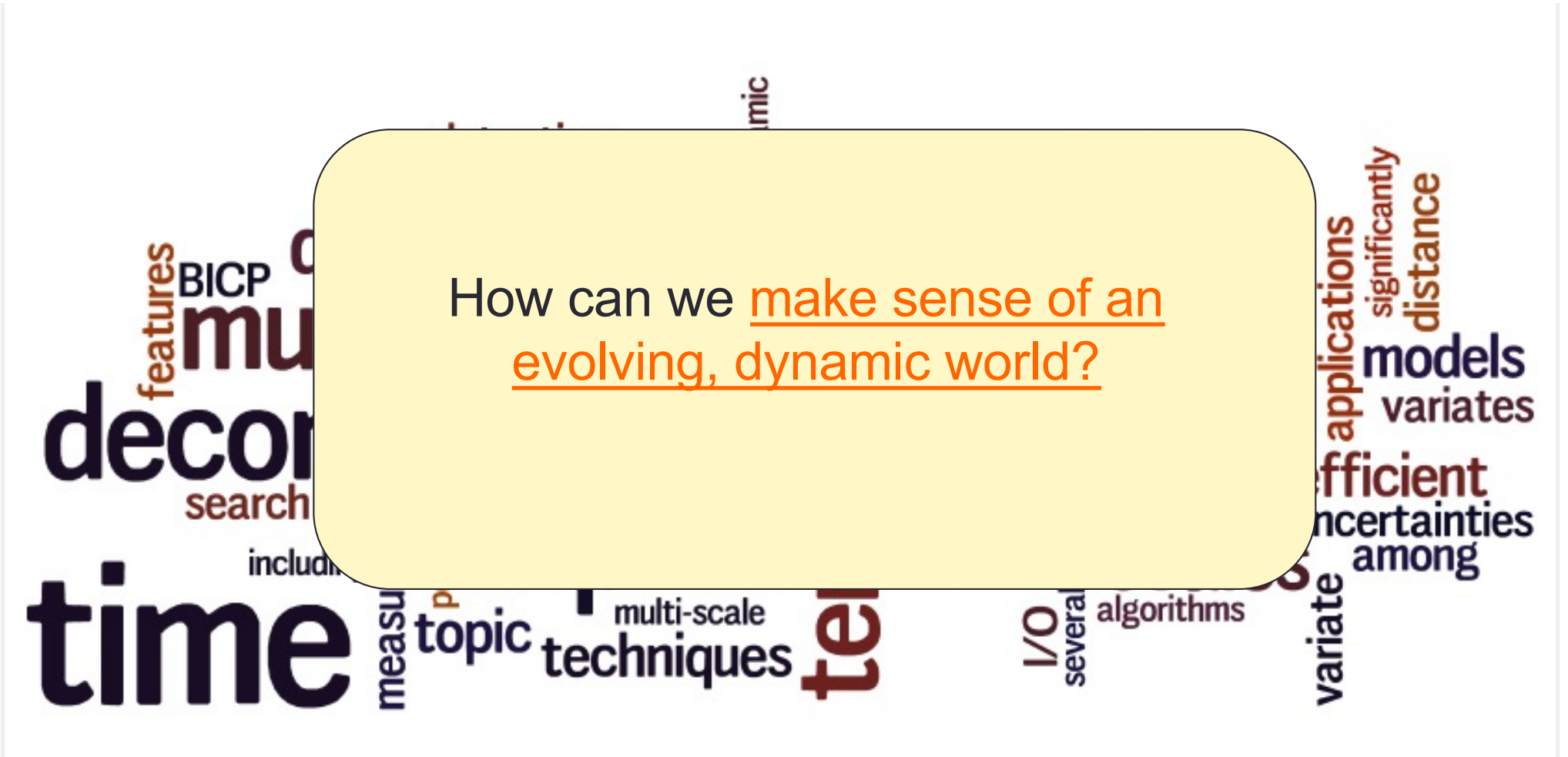
Research Overview…

- **Recent Relevant Grants/Projects:**
- NSF BigData: “Discovering Context-Sensitive Impact in Complex Systems”
- NSF “pCAR: Discovering and Leveraging Plausibly Causal (p-causal) Relationships to Understand Complex Dynamic Systems”
- NSF RAPID ; “RT… …ling”
- NSF: “DataSto… …nse”
- NSF PFI-RP: E… … Performance and Sustainabl…
- NSF II-NEW: “… …rogeneous and Dynamic Data’…
- DOD: “Multivar…
- DOE: “Securin… …silient System (CYDRES)”
- NSF; “Data Ma…
- NSF RAPID ; “U… …ica and Supporting Real-Time Decision Making and Hypothesis Testing through Large Scale Simulations”
- Horizon2020- FourCModelling Conflict, Competition, Cooperation and Complexity: Using Evolutionary Game Theory to Model Realistic Populations

- How can we provide
  - the relevant data/information
  - to the right person/application
  - fast

  to support effective decision making

So what about my team's (recent) work?



We are living in a dynamic, data-rich world…

energy
rehabilitation
finance
security
smart-offices
smart-rooms
production
life-sciences
defense
sports
robotics
elderly-care
retail
supply-chain
child-care
entertainment
personal-data management
transportation
education
space exploration
pet-care
arts
health-care
business/enterprise
sciences
advertisement

Data: Key Enabler

Human-centered IT
Context-aware, location-based services
Accessible work environments
Educational networks

Smart People

Large data in the cloud
Real-time metro-area data analysis
Privacy and access control
Context-aware data warehouses
Public health

Smart Governance

Personalized healthcare
Context-aware, location-based services
Navigational guidance
Ambient intelligence and RFID
Real-time sensor data analysis

Smart Living

DATA

Smart Environment

Context-aware, location-based services
Real-time sensor data analysis
Sustainability
Resilient

Smart Mobility

Real-time sensor data analysis
Large data in the cloud
Real-time metro-area data analysis
Social networks and evolving communities

Smart Economy

Customer engagement and personalization
Social networks and evolving communities
Recommendation and decision support
Situational-awareness and smart supply-chains

Data: Key Enabler

Human-centered IT
Context-aware, location-based services
Accessible work environments
Educational networks

Large data in the cloud
Real-time metro-area data analysis
Privacy and access control
Context-aware data warehouses
Public health

Smart People

Personalized healthcare
Context-aware, location-based services
Navigational guidance
Ambient intelligence and RFID
Real-time sensor data analysis

Can we leverage "big data" for tools and services to enable communities where inhabitants live and work collectively in harmony, and where the community functions as a system that is resilient, self-sustaining, and livable?

Environment

Smart Mobility

Context-aware, location-based services
Real-time sensor data analysis
Sustainability
Resilient

Smart Economy

Real-time sensor data analysis
Large data in the cloud
Real-time metro-area data analysis
Social networks and evolving communities

Customer engagement and personalization
Social networks and evolving communities
Recommendation and decision support
Situational-awareness and smart supply-chains

“Big Data” Industry Roundtable at ASU

- Co-organized with IBM
- On-site or off-site participation
  - Aerojet,
  - Avnet,
  - Boeing,
  - Facebook
  - Google
  - IBM TJ Watson (Exascale System Software),
  - IBM Smart Analytics
  - IO Data Centers,
  - Johnson Controls,
  - LinkedIn,
  - Lockhed Martin,
  - Mayo Clinic,
  - NEC Labs,
  - Oracle,
  - Salt River Project,
  - SAP

2nd Event…


EARLY WARNING
reengen
RiskSense
CYBERSPONSE
ENABLED SECURITY
SAE INTERNATIONAL
PI CUBE
hp
SCIENCE foundation ARIZONA
Innovation at Work
ORION HEALTH
Johnson
What frustrates you today?
YAHOO!
NIST
National Institute of Standards and Technology
Adobe
Federal Reserve Bank of Kansas City
AMERICAN EXPRESS
yaap
Greater Phoenix
ECONOMIC COUNCIL
Positive Energy

“Big Data” Industry Roundtable at ASU

•Today, the **amount of data** being generated is **massive**.

•This necessitates **engineering of new data architectures** with lots of processing power and tools that can **match the scale** of the data and support **split second decision making**, **through data fusion and integration and analysis and forecasting** algorithms, to help non-data-experts (both government and commercial) make decisions and **generate value**.

"Hunting for the Value Gaps in Data Management, Services, and Analytics”
ACM SIGMOD blog; http://wp.sigmod.org/

Key knowledge gaps..

- Six most critical knowledge competency groups (in terms of the value gap – i.e., the difference between current and desired states of the knowledge area)
  1. temporal and spatial analyses,
  2. summarization, cleaning, visualization, anomaly detection,
  3. real-time processing for streaming data,
     - media analytics
  4. representations and fusion for unstructured/structured data, semantic Web,
     - make unstructured data queriable, prioritize and rank data, correlate and identify the gaps in the data
  5. graph-based models, social networks,
     - entity analytics, (social and other) network analytics
  6. performance and scalability, distributed architectures.

"Hunting for the Value Gaps in Data Management, Services, and Analytics"
ACM SIGMOD blog; http://wp.sigmod.org/

Smart data challenges in a dynamic world

energy

rehabilitation

finance

security

smart-offices

SINT

(S)parsity

(I)mprecision

(N)oise

Variable (T)rust

3Vs

(V)olume

(V)elocity

(V)ariety

HMLE

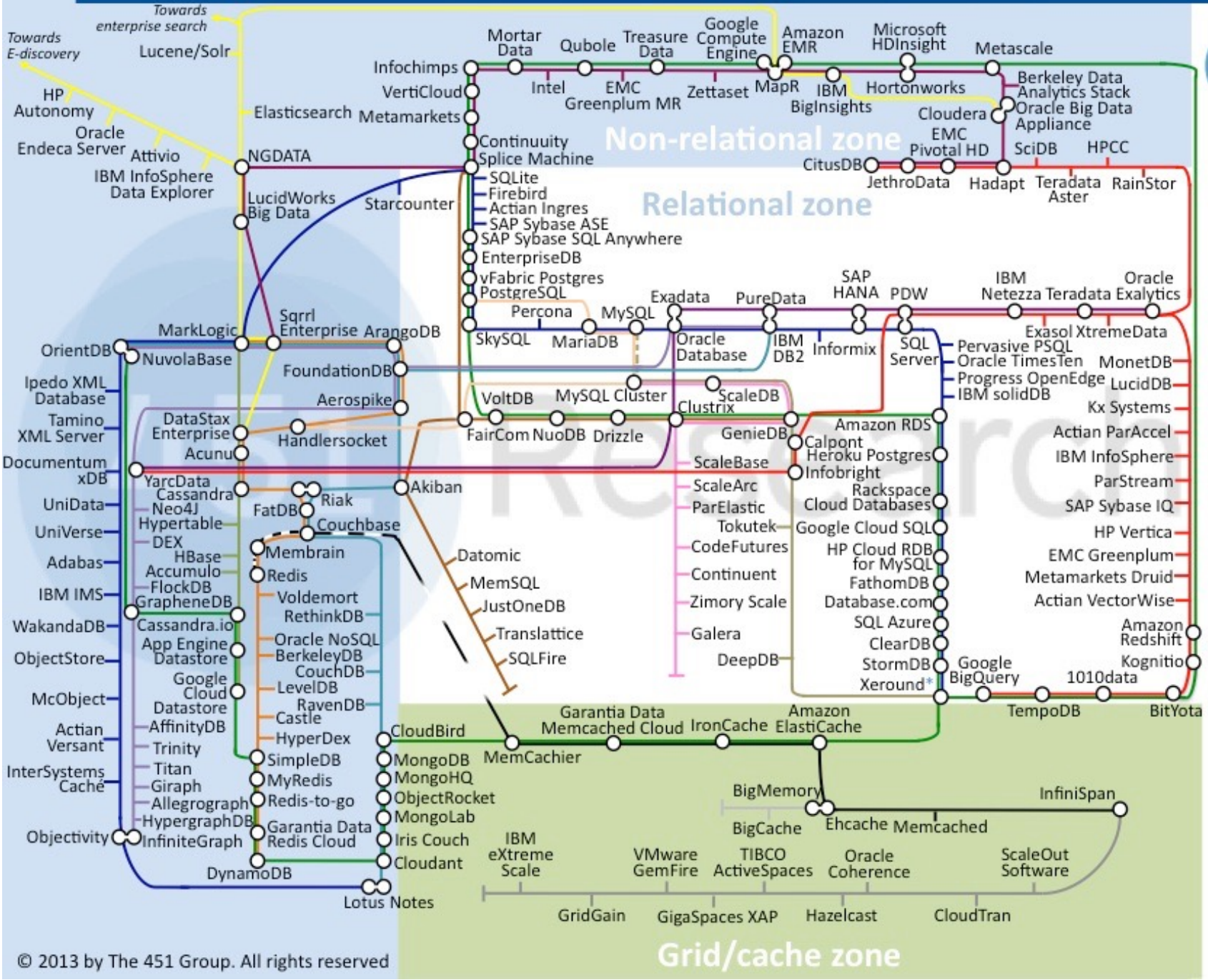
(H)igh-dimensional

(M)ulti-modal

Inter-(L)inked

(E)volving

Database Landscape Map – June 2013



451 Research

Key:

- General purpose
- Specialist analytic
- -as-a-Service
- NoSQL extension
- BigTables
- Graph
- Document
- Key value stores
- Key value direct access
- Hadoop
- NewSQL extension
- MySQL storage engines
- Advanced clustering/sharding
- New SQL databases
- Data caching extension
- Data caching
- Data grid
- Search
- Appliances
- Off-heap memory

*Xeround closed May 2013 until further notice