

POLARIS Big Data Panel

Dr. David Belanger
dgb@ieee.org



STEVENS
INSTITUTE *of* TECHNOLOGY
THE INNOVATION UNIVERSITY®

Brief Introduction

- **Currently:**
 - **Senior Research Fellow – Stevens Institute of Technology**
 - **Leader – IEEE Big Data Initiative**
 - **Various: BHEF Big Data Study; NSF Visitor Panel;**
- **History:**
 - **Retired Chief Scientist, and V.P. Information, Software, and Systems Research – AT&T Labs**
 - **Founder (1995) and Leader (until 2012) AT&T Infolab**
 - **33 years at Bell/AT&T Labs**
- **Interests (Technical):**
 - **Data Science/Engineering, Systems, Software**
 - **Big Data Networking, Management, Analysis, Governance**



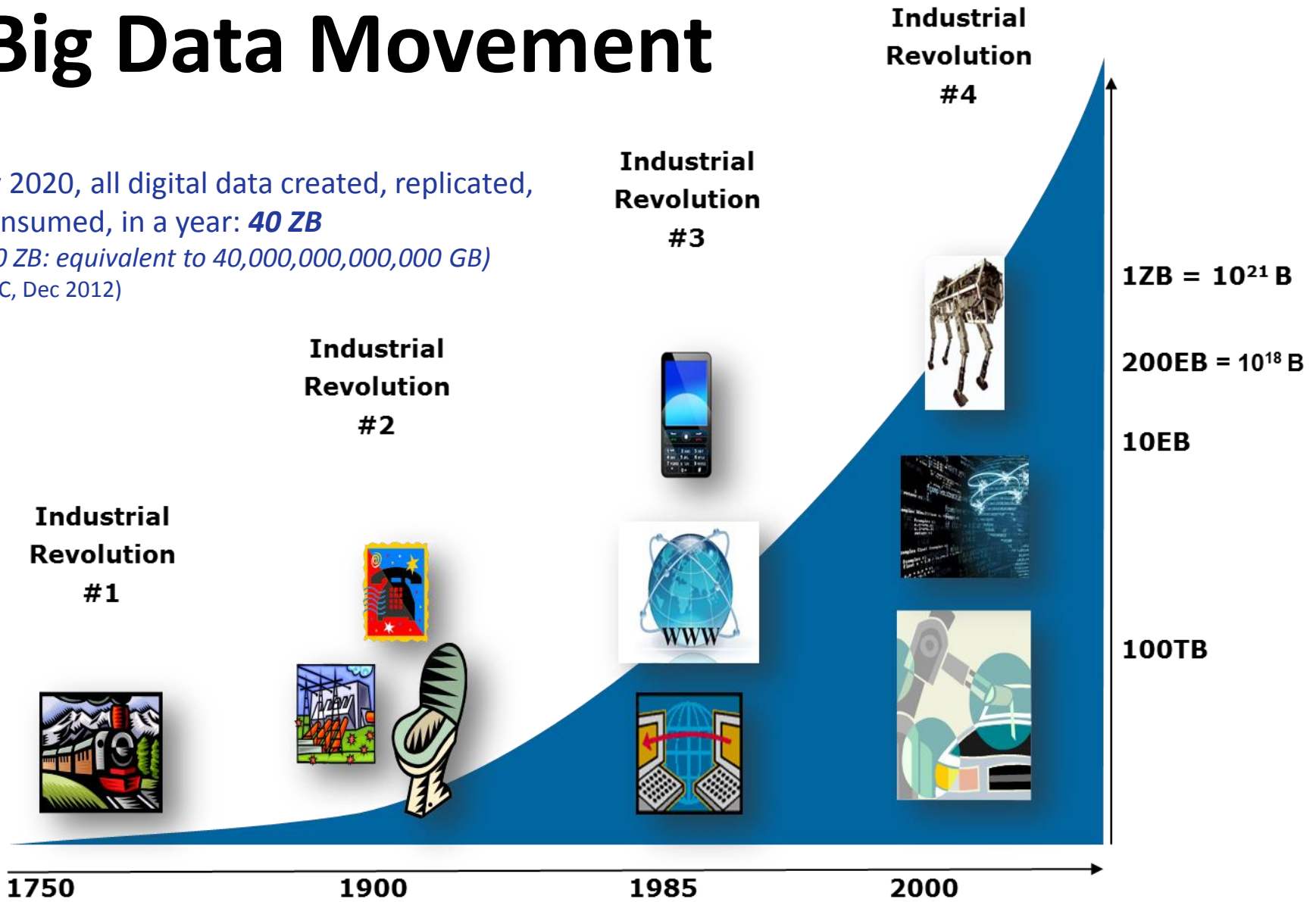
Key Challenges and Opportunities

- **Policy:**
 - **Sharing of Data**
 - **Standards for Data: Usage, Governance, Management**
 - **Ecosystems for Second Generation of Big Data**
 - **Critical application areas – e.g. healthcare, grid, sustainability**
- **Technology:**
 - **Another generation of communications technology: 5G, SDN, NFV, IoT, Fog = Orders of Magnitude more data.**
 - **Data Quality EcoSystem Tools**
- **Directional:**
 - **Big Data In-line to Production Systems – Fully Automated**
 - **Educational infrastructure**



Big Data Movement

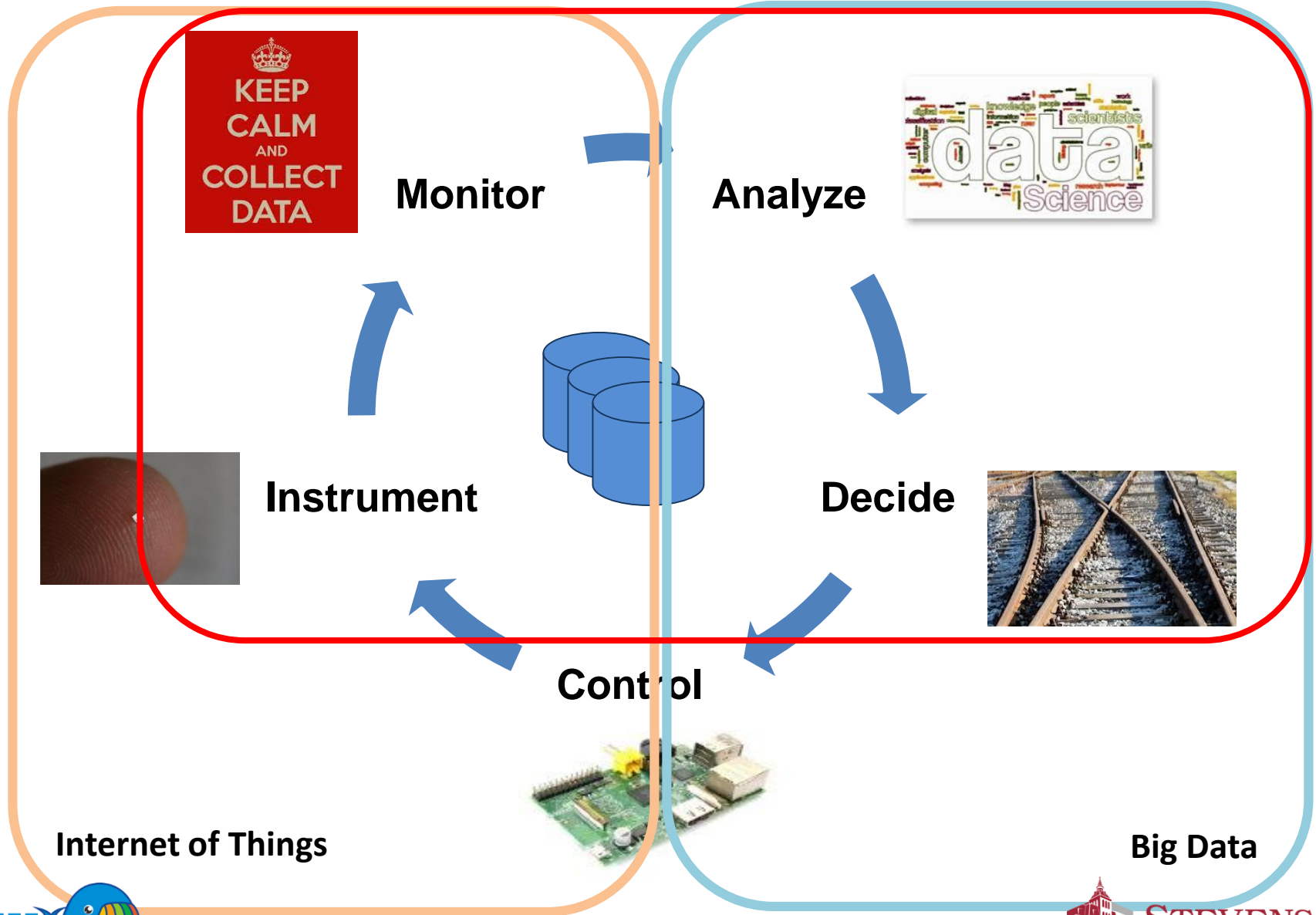
By 2020, all digital data created, replicated, consumed, in a year: **40 ZB**
(40 ZB: equivalent to 40,000,000,000,000 GB)
(IDC, Dec 2012)



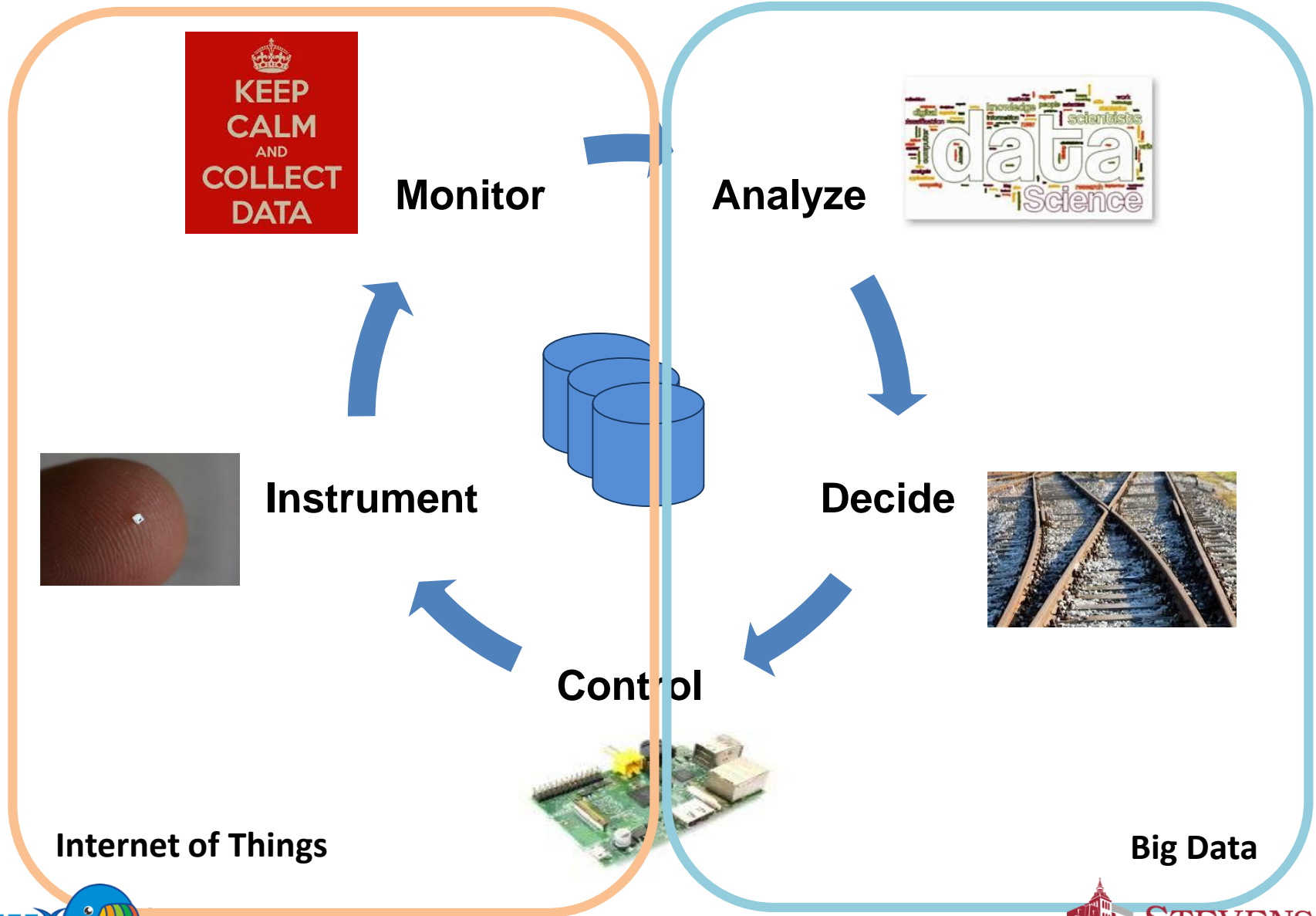
R. J. Gordon: Is US economic growth over? Faltering innovation confronts the six headwinds. CEPR Policy Insight No 63

Courtesy of Mark Davis

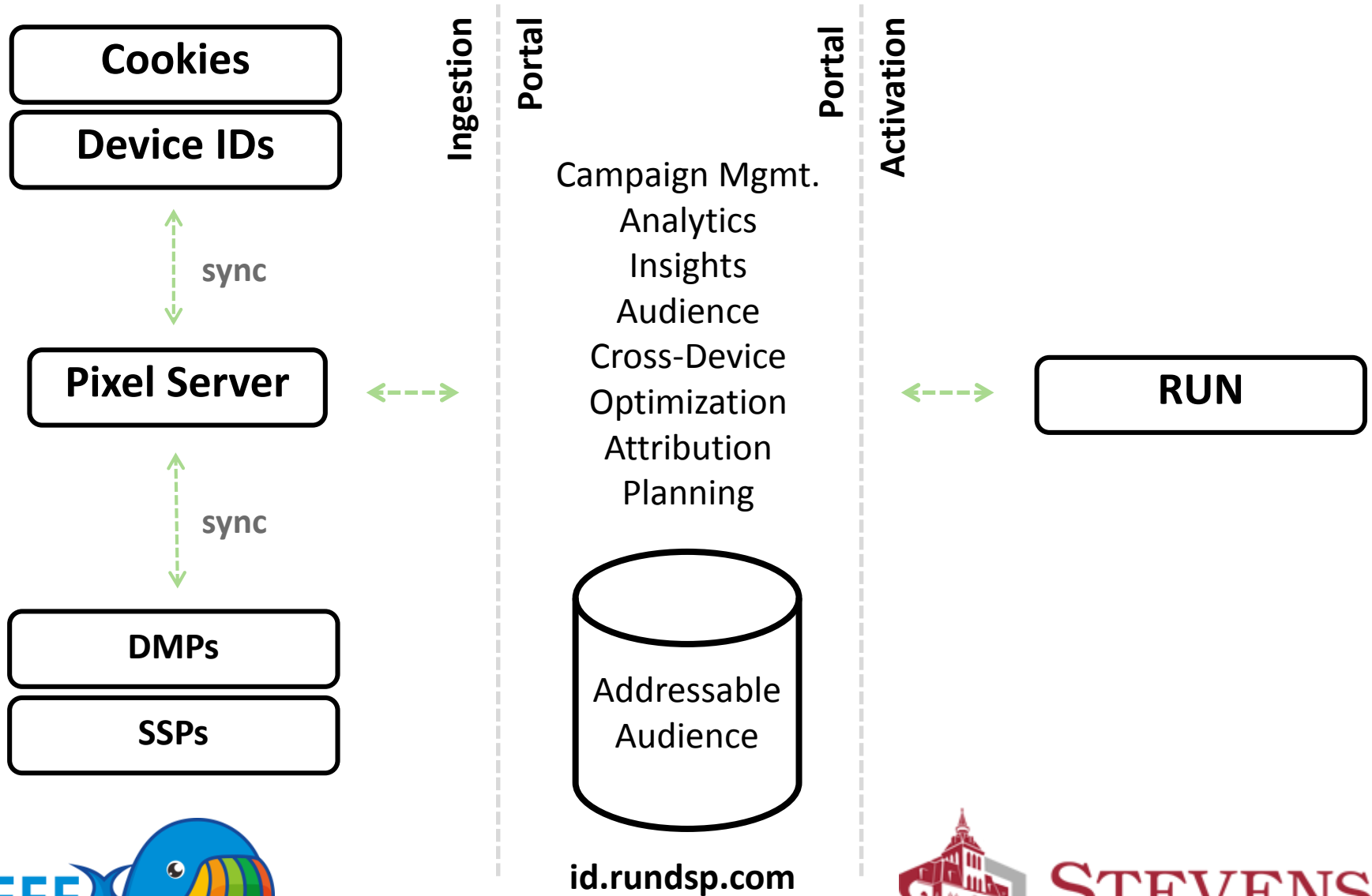
Data Analysis Lifecycle: IoT



Data Analysis Lifecycle: IoT



RUN* Today



*Source: RUN Inc.



STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

A Few Infrastructure Issues

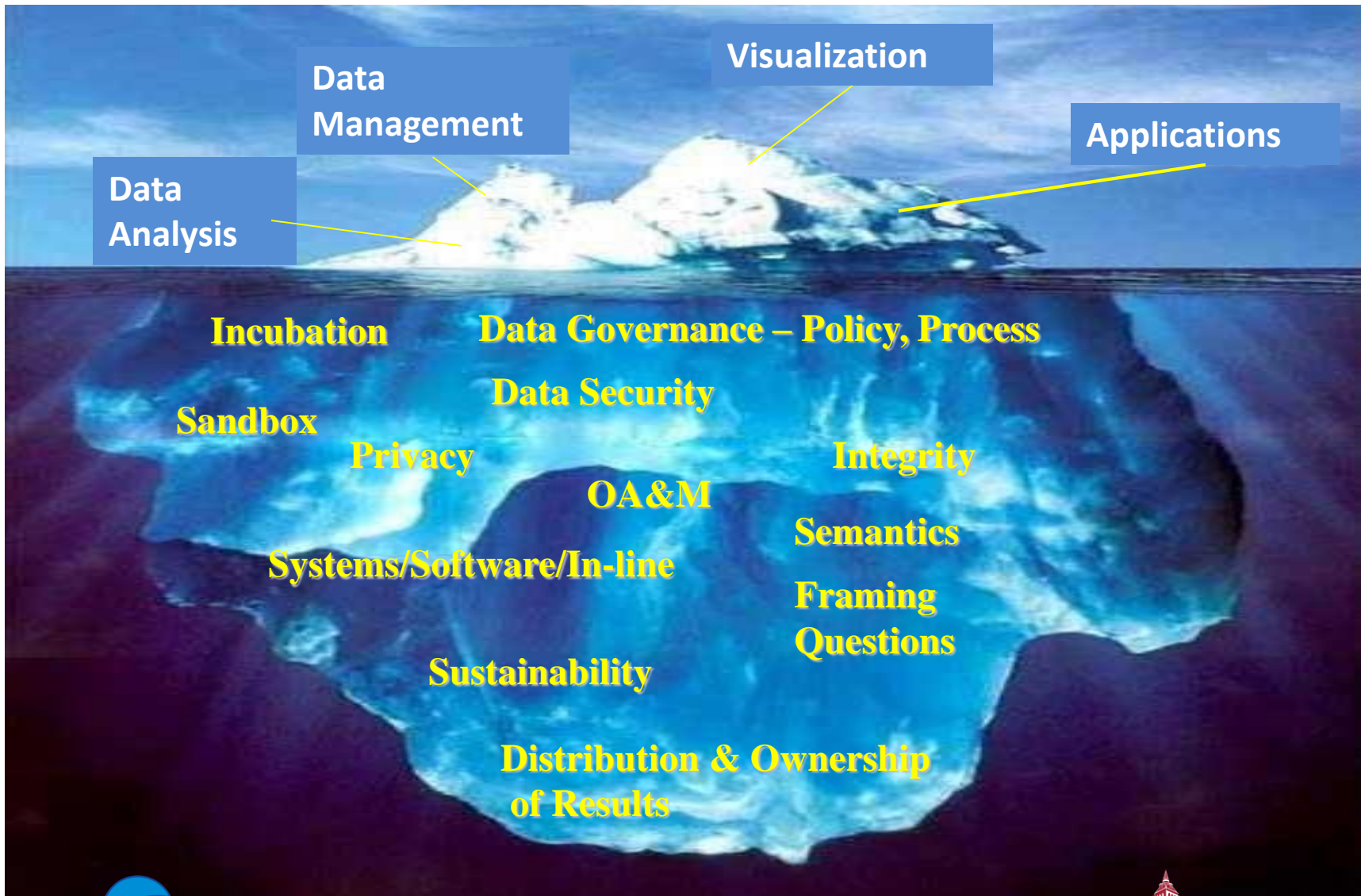
- **Data Science Related Issues**
 - **Drift**
 - **Tools**
 - **Scale**
 - **Verification**
 - **Automatic Data Preparation**
- **Data Engineering Related Issues**
 - **Distribution & Scale**
 - **NoSQL? CAP?**
 - **Hitting the Wall**
- **Software Engineering Related Issues**
 - **Configuration Management**
 - **Testing and Performance**
- **Systems Related Issues**
 - **Standards – e.g. data transfer**
 - **Fault Tolerance**
 - **Change Management and Propagation of Changes.**

Status of Research

- **Healthy and Dynamic:**
 - **Networking**
 - **New Analytic Techniques, Including Information Visualization**
 - **Tools for Data Management, Analytics, Data Streams, etc.**
 - **Physical – e.g. Quantum Computing**
 - **Security**
- **Emerging:**
 - **In-line Systems, Systems/Software Engineering**
- **Some:**
 - **Second Generation – Broad Usage**
 - **Governance**
 - **Data and System Quality**



A More Complete Picture

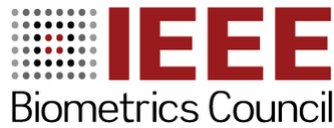


Areas for EU/US Collaboration

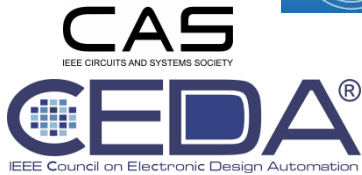
- **A few clear areas:**
 - **Standards**
 - **Education**
 - **Big Data Research Infrastructure – Grand Challenge**
 - **Critical Applications**
- **IEEE Big Data Initiative (bigdata.ieee.org):**
 - **Conferences**
 - **Publications**
 - **Data Repositories**
 - **Data Analytics Competitions**
 - **Standards**
 - **Education**
 - **Applications: e.g. EMBS, smartcities.ieee.org**
 - **Infrastructure: e.g. iot.ieee.org, cybersecurity.ieee.org**
 - **Membership from global corporations**

25 OUs Supporting BDI

➤ **David Belanger** (ComSoc) leads the initiative, with program management by Kathy Grise



IEEE Sensors Council



IEEE Educational Activities



IEEE Technical Activities



IEEE STANDARDS ASSOCIATION



Big Data in a Programmable World

