Proposal for the Theme on Big Data Analytics

Qiang Yang, HKUST
May 2015
Jiannong Cao, PolyU
Qi-man Shao, CUHK

Motivation

- The world's technological per-capita capacity to **store information doubled** every 40 months
  - As of 2012, 2.5 exabytes (2.5×1018) of data/day
    - Relational database management systems and desktop statistics and visualization packages often **have difficulty** handling big data.
  - Big Data: new driver for digital economy&society
    - Gartner: hundreds of billions of GDP by 2020.
    - Intangible factor after labor and capital
    - Data Science: The fourth paradigm

The Power of Big Data

- Big Data can bring "**big values**" to our life in almost every aspects.

- Technologically, Big Data is bringing about changes in our lives because it allows **diverse and heterogeneous data to be fully integrated and analyzed to help us make decisions**.

- Today, with the Big Data technology, **thousands of data from seemingly unrelated areas can help support important decisions**. This is the power of Big Data.

- Areas of Applications

  - Health and Well being

  - Policy making and public opinions

  - Smart cities and more efficient society

  - New online educational models: MOOC and Student-Teacher modeling

  - Robotics and human-robot interaction

- Much of this power hinges on Research on Analytics

Hong Kong needs Big Data Research

1. to develop state-of-the-art Big Data platform in research, education and industrial applications, and open it to the Hong Kong society and the world at large, and

2. to make a difference in Smart Cities, Health and Well-being (including supporting aging populations), and modernizing Finance, Education and Logistics in Hong Kong.

Big Data Analytics Objectives
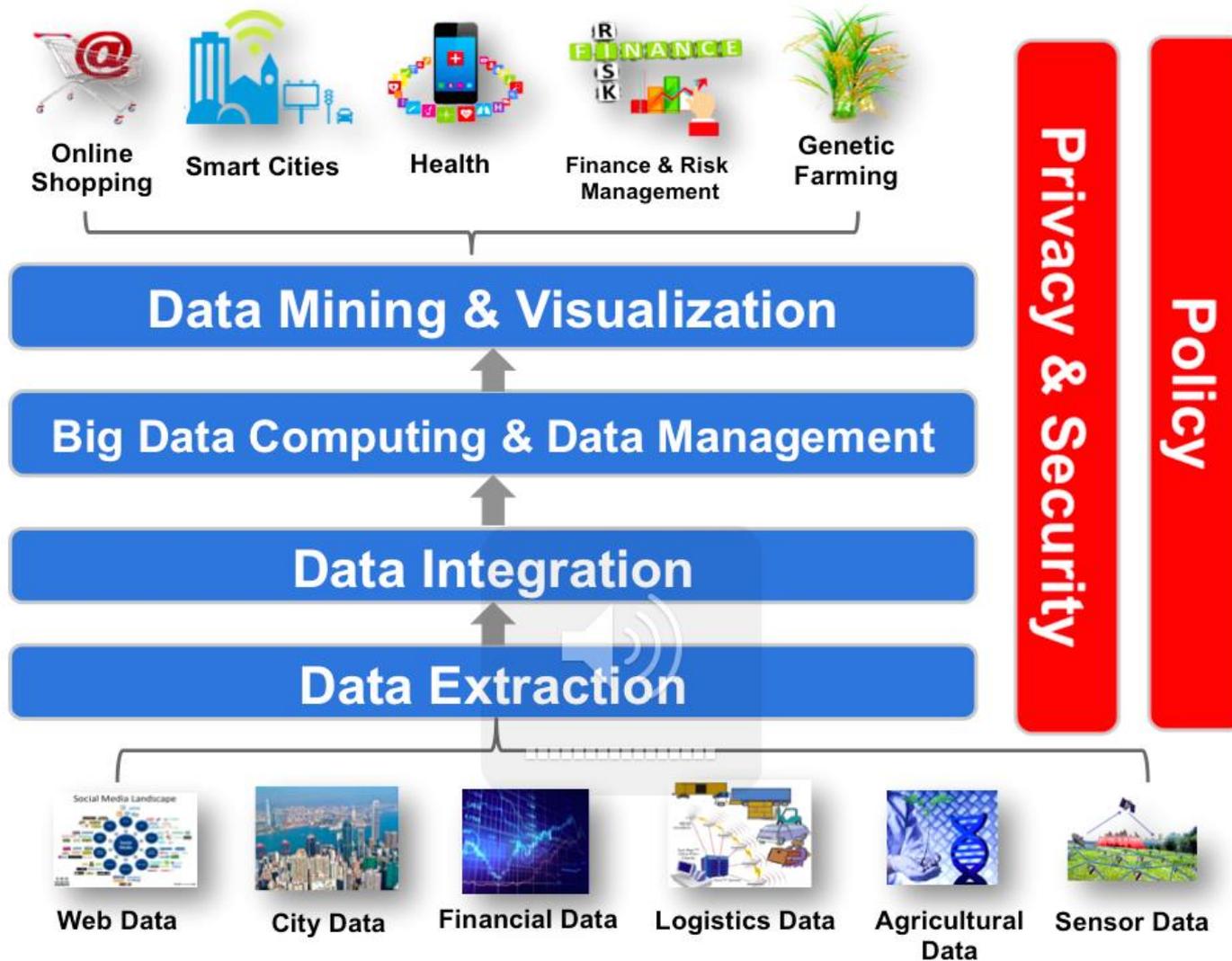
Relation to Smart Cities and IoT

- World economic forum ranking HK's infrastructure: #1
  - Maintain the lead in IT Infrastructure
- East Kowloon Project: Energizing Hong Kong via Smart Cities
- Big Data:
  - IoT provides the infrastructure for collecting the data
  - Smart Cities as important application goal

Smart City | Background

2014/15 Budget Speech
"In the recent fourth update of the Digital 21 Strategy, we have proposed a series of initiatives under the theme of Smarter Hong Kong, Smarter Living…"

2015 Policy Address
"……to use Kowloon East as a pilot area to explore the feasibility of developing a Smart City."

2015/16 Budget Speech
"……free online government information will be released in digital formats to encourage development of more applications by start-ups.

6

Research Objectives

- **Big Data Analytics: data mining and machine learning**
  - Large-scale machine learning, data mining and data visualization
- **Big Data Computing: data center support for Analytics**
  - Big data collection and transformation, integration and distributed data management and computing
- **Big Data Theory, Privacy&Security issues on Analytics**
  - Big data sampling and statistical theory, Big data security and privacy
- **Big Data Science: 4th Paradigm – Analytics for Science and Engineering**
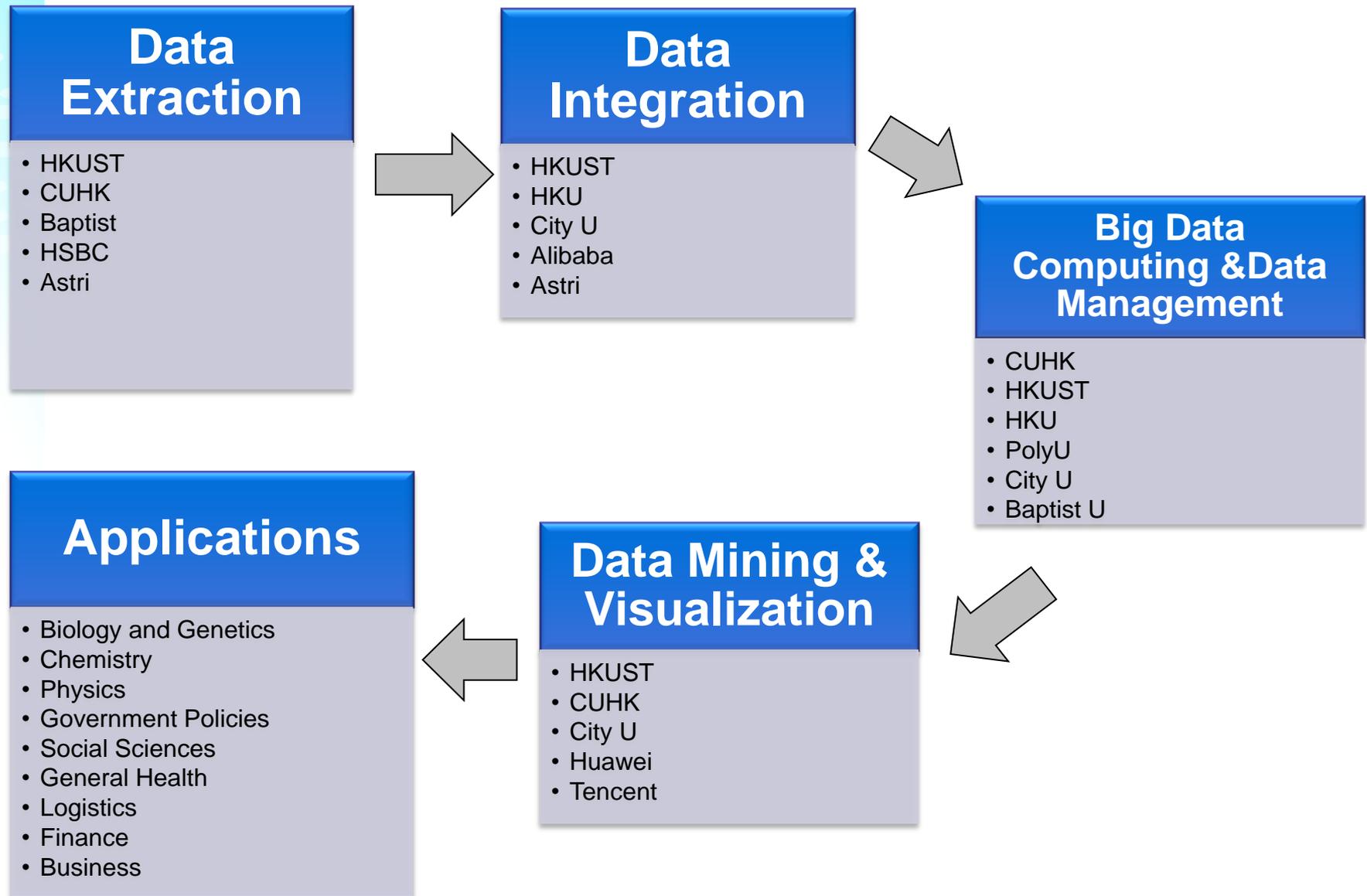 - Big Data and Multi-disciplines (Bio, Chemistry, Engineering, Social)

Why Hong Kong is Ready for the Theme

- We have the best researchers in machine learning, data mining, data management, sensor networks, statistics, and multidisciplinary research such as bioinformatics
  - China National 973 Projects on Big Data
  - IEEE Transactions on Big Data: EiC
  - ACM KDD Conferences: PC and Conference Chairs
  - Winner of Big Data related international competitions
- New industries based on lots of data
  - Financial industry, logistics industry, education sector, government services, etc.
- We have many potential collaborators and partners
 - Huawei, Tencent, Baidu, Alibaba, Google, Microsoft, etc.

IEEE TRANSACTIONS ON BIG DATA

Big Data Analytics Workflow

Data Extraction

- HKUST
- CUHK
- Baptist
- HSBC
- Astri

Data Integration

- HKUST
- HKU
- City U
- Alibaba
- Astri

Big Data Computing &Data Management

- CUHK
- HKUST
- HKU
- PolyU
- City U
- Baptist U

Applications

- Biology and Genetics
- Chemistry
- Physics
- Government Policies
- Social Sciences
- General Health
- Logistics
- Finance
- Business

Data Mining & Visualization

- HKUST
- CUHK
- City U
- Huawei
- Tencent

Multi-disciplinary Big-data Analytics

- **Objectives:**
  - Interdisciplinary, multi-university, multi-team research on heterogeneous scientific and technological big data analytics
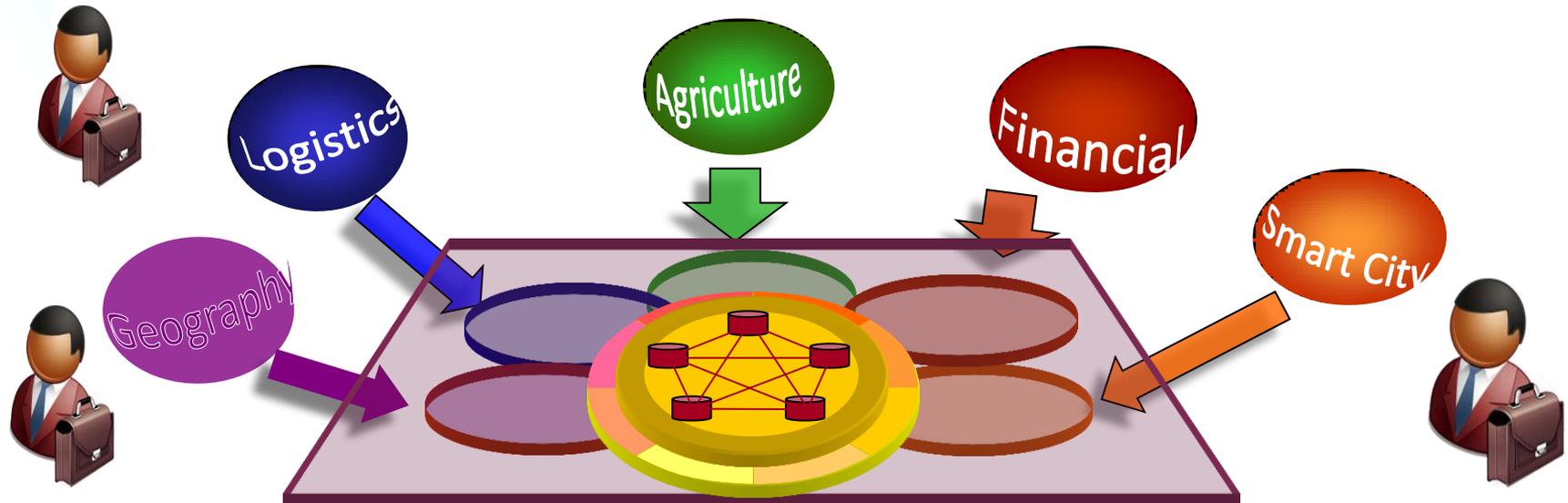
Why Big Data needs Team Work?

- Big data analytics is necessarily a joint effort by researchers from academic institutions, government and society and industry.

  - The government and industry are sources of Big Data, and providers of problems and challenges,

  - The academic researchers are solution providers.

  - When it comes to package the solutions from university labs to transfer to the real world, universities and industry must work together to build scalable and robust solutions.

Expected Outcomes

- New methodologies and solutions for Big Data research

- New applications that impact the society and industry in Hong Kong and beyond, and

- New digital economies created based on big data

- New educational programs for students; cultivating leaders for the Big Data society and industry.

Breakthroughs Expected

- New algorithms, methodologies, systems and applications in Big Data

- New knowledge from Big data applications in Science, Engineering, and Societal Problems

- New insights into Big data practices in real world

- New ways to protect security and privacy of big data relevant to individuals and organizations

Call for Proposals in Big Data Analytics

- **Foundations in Big Data Analytics Research**:
  - developing and studying fundamental theories, algorithms, techniques, methodologies, technologies to address the effectiveness and efficiency issues to enable the applicability of Big Data problems;

- **Innovative Applications in Big Data Analytics**:
  - developing techniques, methodologies and technologies of key importance to a Big Data problem that requires the seamless cooperation of domain scientists with big data researchers.

Benefit for the Hong Kong Society

- Can Hong Kong transform into a more data-driven society and maintain its leadership globally?

- How can companies in Hong Kong become more competitive with Big Data technology?

- Can Data Science in Hong Kong's research fields benefit from strong foundations on big data?

- Can the government become more efficient with big data driven methodologies in decision making?

- Can education, finance, logistics and health benefit from the ever increasing data?

Projected impact of climate change on agricultural yields