

# BIG DATA ANALYSIS



**PRESENTED BY:**

**P. SHARMILA**

**M.VANITHA**

**A.VIMALADEVI**

# Motivation

- The world's technological per-capita capacity to **store information** **doubled** every 40 months
  - As of 2012, 2.5 exabytes ( $2.5 \times 10^{18}$ ) 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
	<p><b>2014/15 Budget Speech</b> <i>"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..."</i></p> <p><b>2015 Policy Address</b> <i>".....to use Kowloon East as a pilot area to explore the feasibility of developing a Smart City."</i></p> <p><b>2015/16 Budget Speech</b> <i>".....free online government information will be released in digital formats to encourage development of more applications by start-ups."</i></p>

# 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: 4<sup>th</sup> Paradigm – Analytics for Science and Engineering**
  - Big Data and Multi-disciplines (Bio, Chemistry, Engineering, Social)

**B**ig data &  
biology





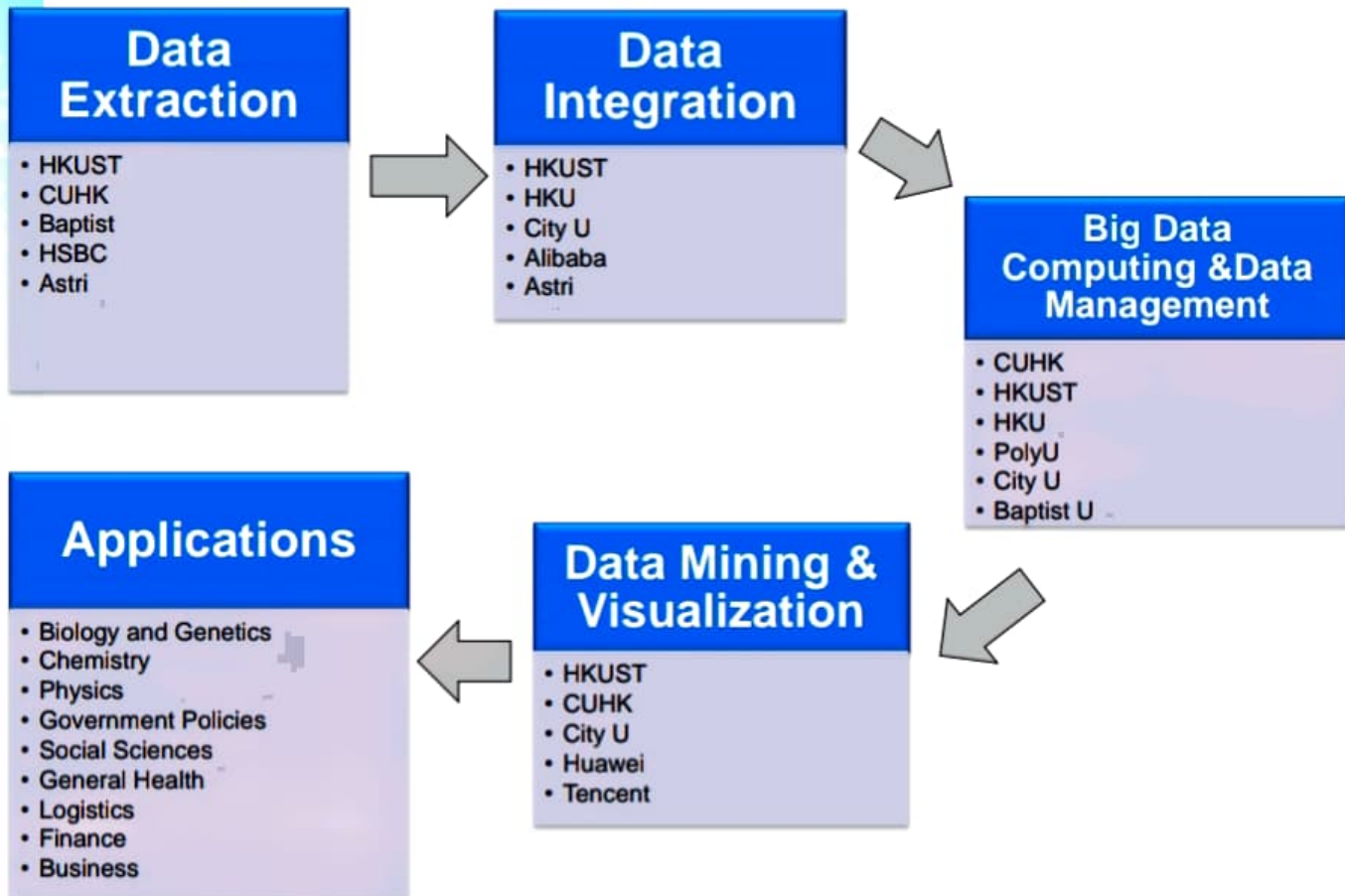
# Why Hong Kong is Ready for the Theme

- 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.





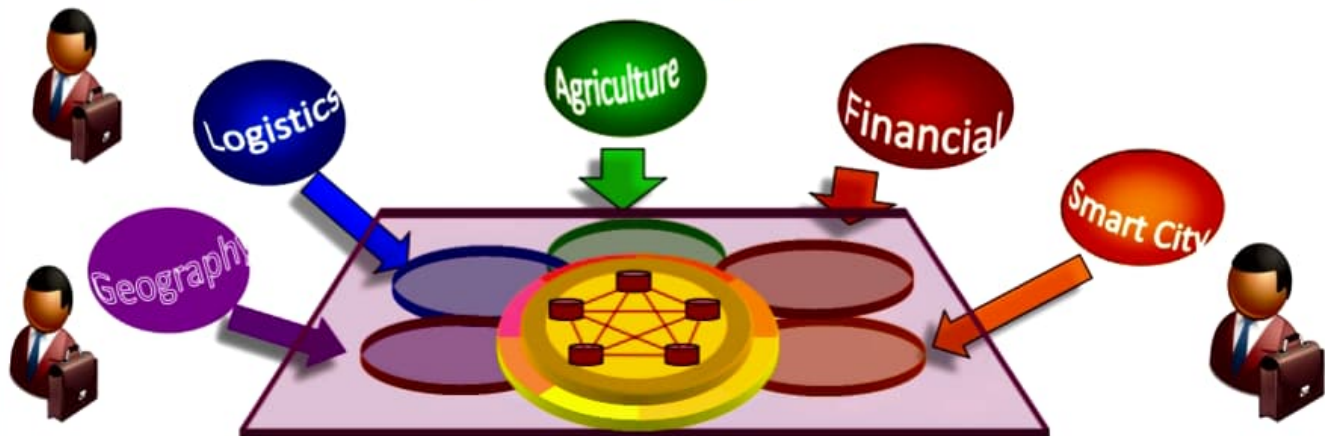
# Big Data Analytics Workflow



# 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.



**THANK YOU**