

CREATING FUTURES THROUGH RESEARCH – MEETING CHALLENGES, EMBRACING
OPPORTUNITIES & DELIVERING IMPACT

**UNIVERSITY DATA ANALYTICS: A CASE OF THE TECHNICAL
UNIVERSITY OF KENYA**

**AUTHORS: OBONYO ISHMAEL, OPIYO ELISHA AND
ODONGO WILLIAM**

PRESENTER: ISHMAEL OBONYO



ORGANIZATION: UNIVERSITY OF NAIROBI

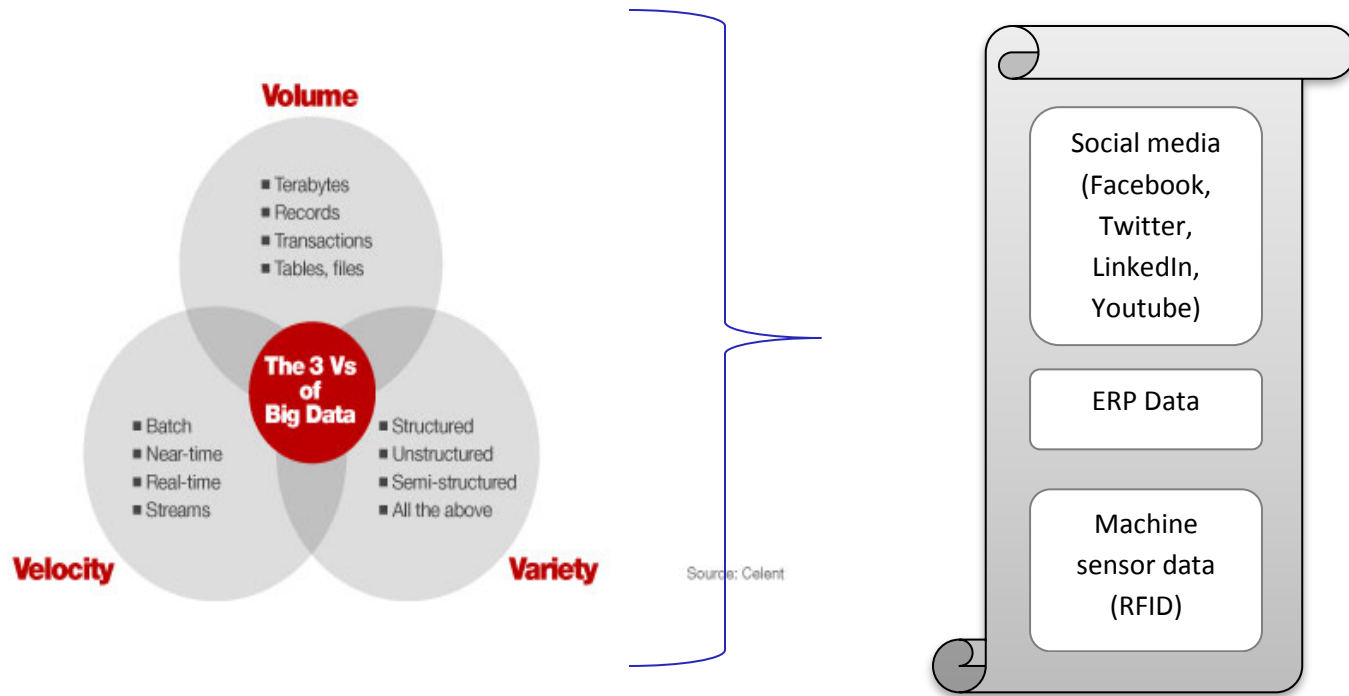
VENUE: INTEL COLLEGE, NAIROBI - KENYA

OVERVIEW

- Definitions of Major Terms
- Background Information
- Problem Statement
- Methodology
- Results
- Conclusion
- References

DEFINITIONS OF MAJOR TERMS

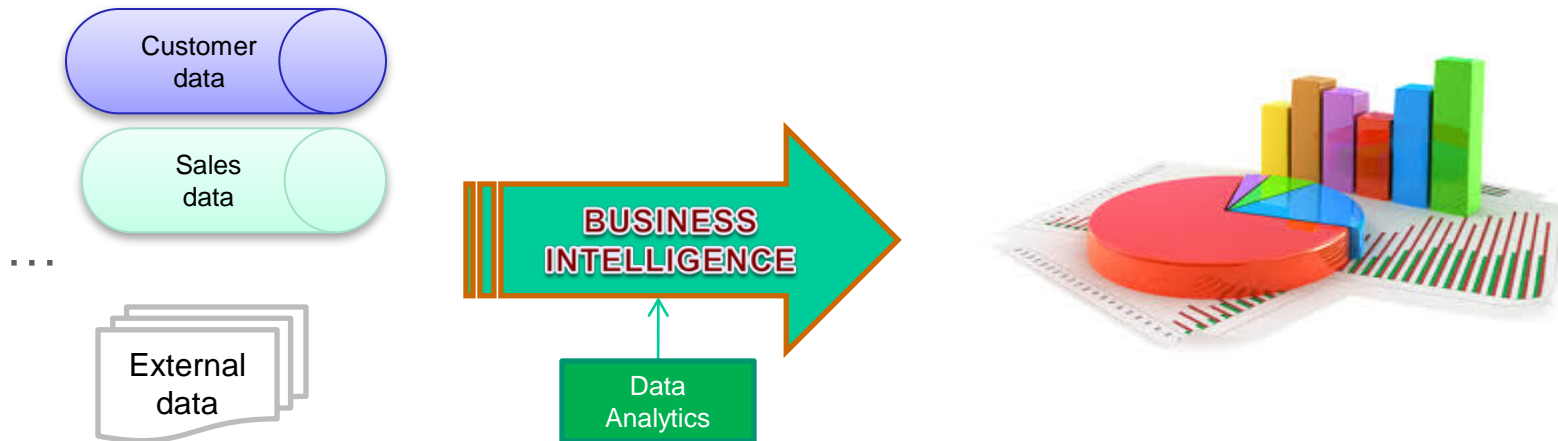
- **Big Data:** described by 3V's - **Volume**, **Velocity** and **Variety** (Chen et al., 2012 & Kwon et al., 2014)



- In addition to the 3 V's, there are other V's: **Veracity**, **Variability**, and **Value** (Gandomi & Haider, 2015)
- It is also a term describing *large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information* (Das & Kalita, 2015).

DEFINITIONS OF MAJOR TERMS

- **Business Intelligence (BI):** set of methodologies, processes, architectures, and technologies used in transforming raw data into meaningful and useful information that is applied in business for competitive advantage (Foley & Guillemette, 2012)



- **Data Analytics:** science that incorporates various disciplines including, but not limited to, data engineering, mathematics, statistics, computing, and domain-specific expertise geared towards transforming data into useful information (Rodgers, Talbut, & Baranovic, 2015, February). It is argued that Business intelligence and analytics (BI&A) has emerged as an important area of study for both practitioners and researchers, reflecting the magnitude and impact of data-related problems to be solved in contemporary business organizations; Data analytics therefore refers to the BI&A technologies that are grounded mostly in data mining and statistical analysis. (Chen, Chiang & Storey, 2012)
- **Hadoop:** open source advanced Data Analytics platform for Big Data adopted in this study for data integration, analytics, and data visualization
- **University Data Analytics:** a process of transforming data pertaining to students, staff, research and other stakeholders into actionable information that can aid a University management during decision making
- **R:** Open source statistical software adopted in this study

BACKGROUND INFORMATION

EFFECTIVE MANAGEMENT OF ORGANIZATION

REQUIRES

Timely access to right **information** pertaining to the organization processes in the right format by the right person

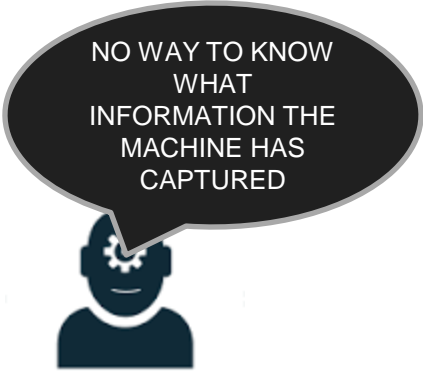
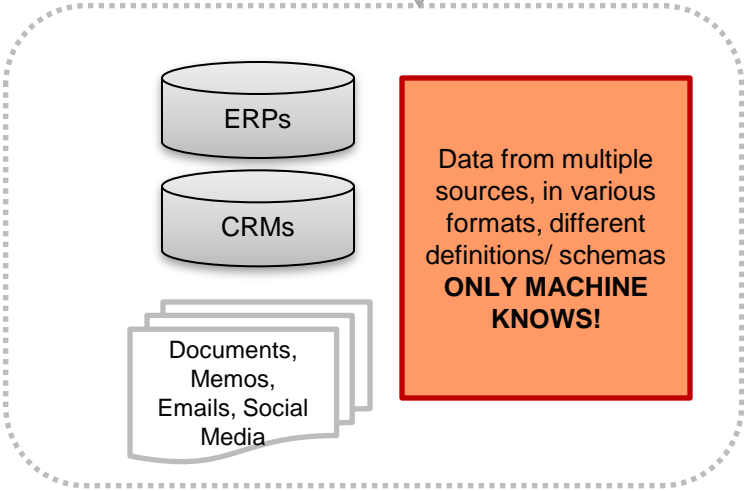
results to

Good Management Decisions



(Davenport & Prusak, 1998)

What brings about the challenge?



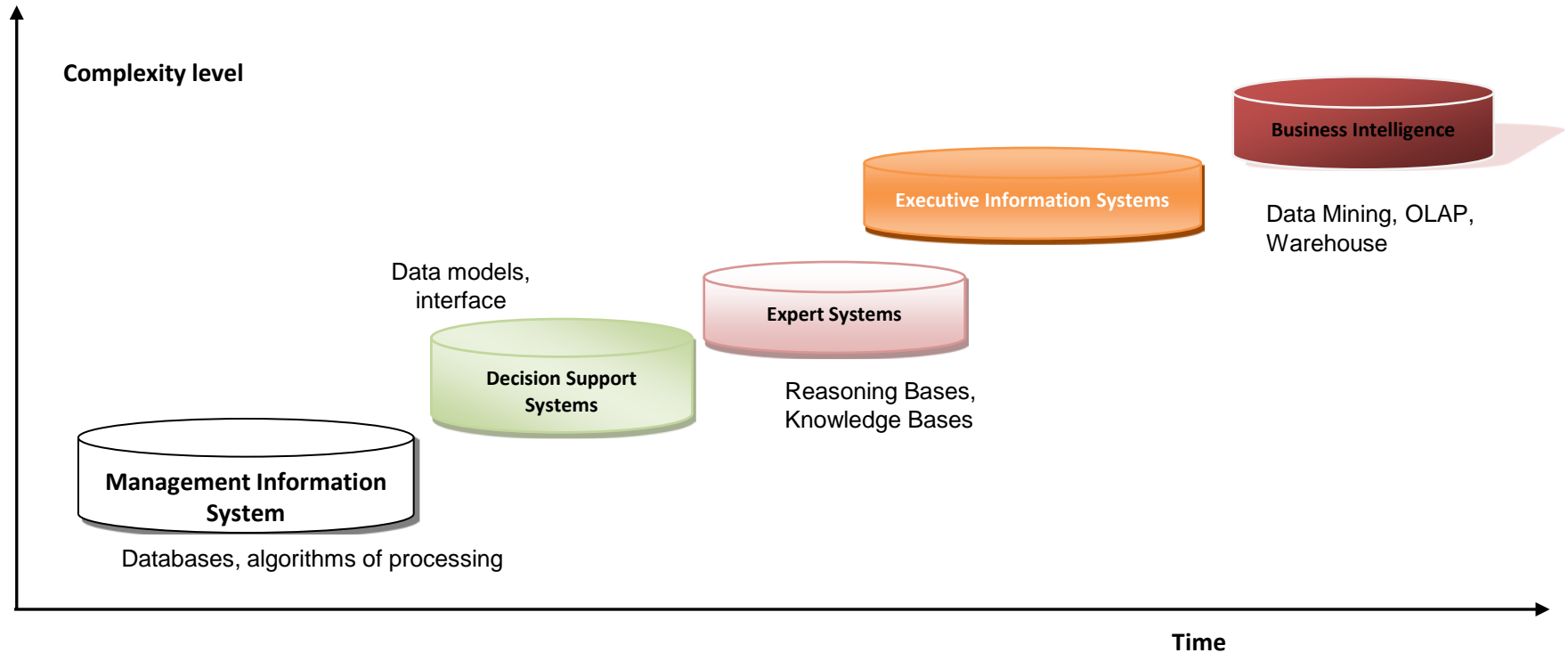
DECISION MAKER – HUMAN BEING

(Herring, 1992 ; Malik, 2005)

BACKGROUND INFORMATION

THE MAJOR CHALLENGE! DATA => INFORMATION => KNOWLEDGE. *THE MORE THE DATA THE GREATER THE CHALLENGE!*

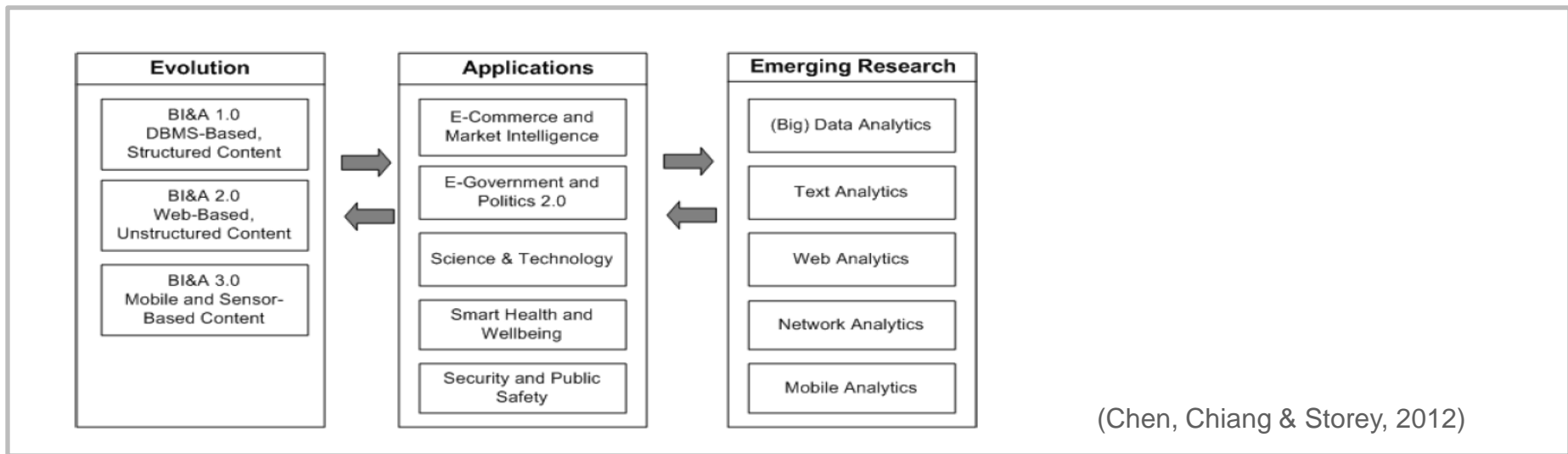
WHAT HAS BEEN DONE TO ADDRESS THE CHALLENGE



(Olszak & Ziemia ,2007)

BACKGROUND INFORMATION

EMERGENCE OF DATA ANALYTICS

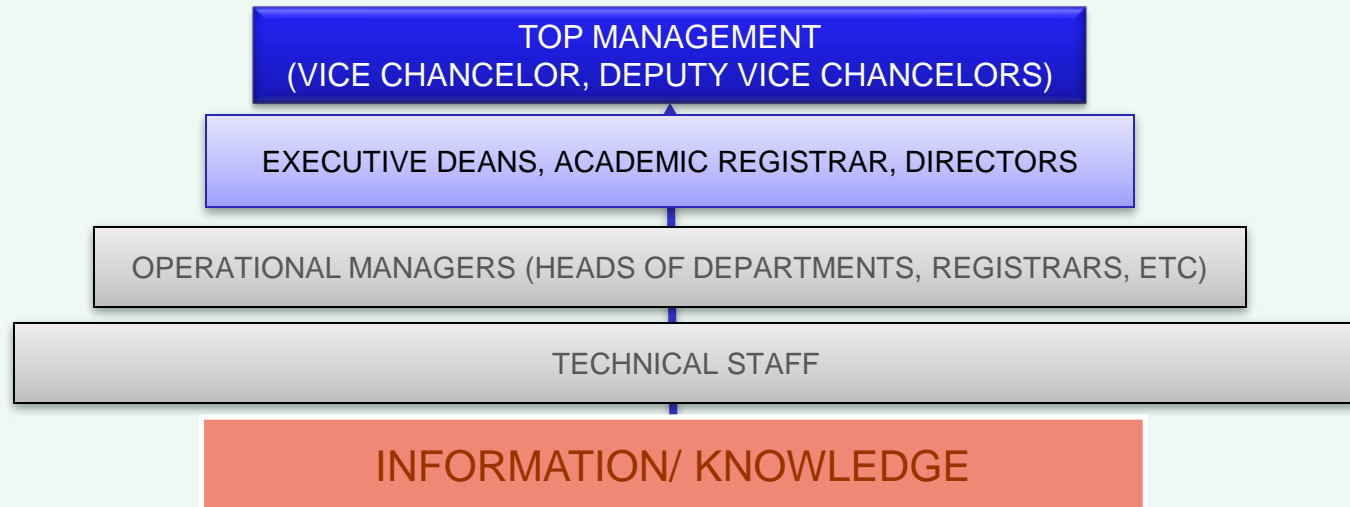


OBJECTIVES

The main objective: to perform data analytics pertaining to university operations using open source tools like Hadoop framework and R, taking a case of the Technical University of Kenya

PROBLEM STATEMENT

THE TECHNICAL UNIVERSITY OF KENYA

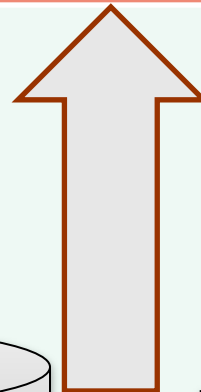
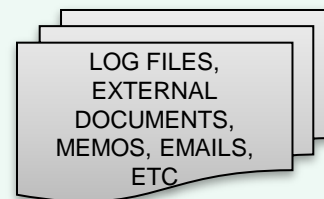
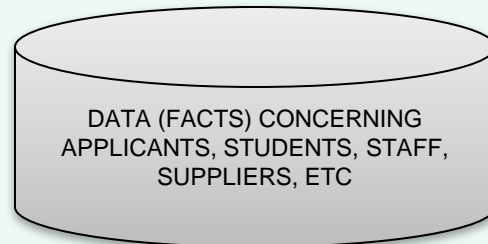


INFORMATION REQUIRED:

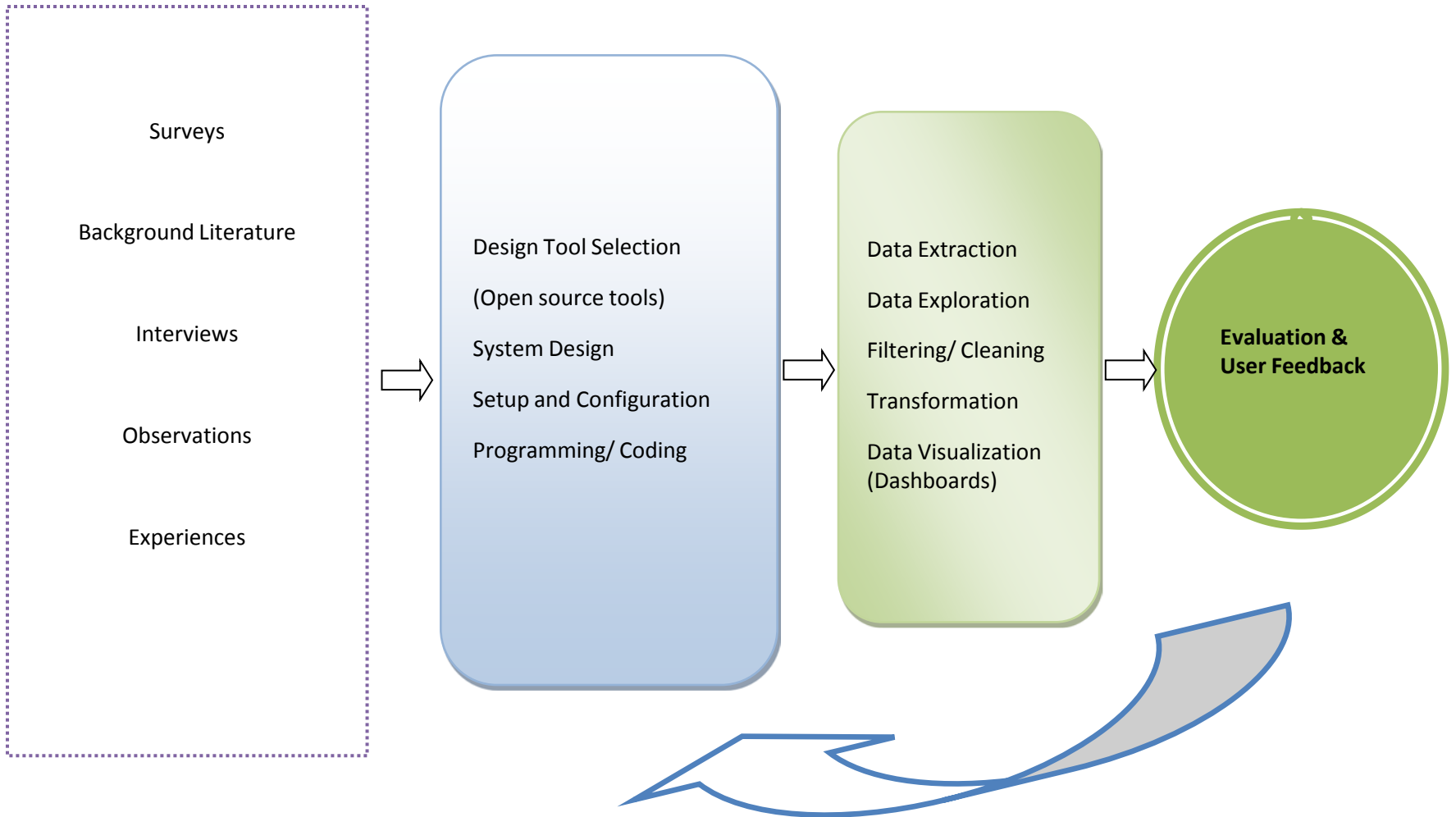
- ❑ ACCURATE NUMBER OF STUDENTS & STAFF
- ❑ STUDENT ADMISSION TREND
- ❑ STAFF COMPOSITION
- ❑ PUBLICATION IMPACT
- ❑ PROGRAMME SUCCESS FACTOR
- ❑ WEBOMETRIC RANKING TREND
- ❑ ETC

CHALLENGES/ ISSUES:

- TOO MUCH TIME REQUIRED FOR REPORT GENERATION
- STAFF BUSY WITH REPORT GENERATION INSTEAD OF WORKING
- INCOMPLETE DATA FOR IMPORTANT DECISION MAKING
- RIGID AND STALE REPORTS
- PERIODICAL REPORTS

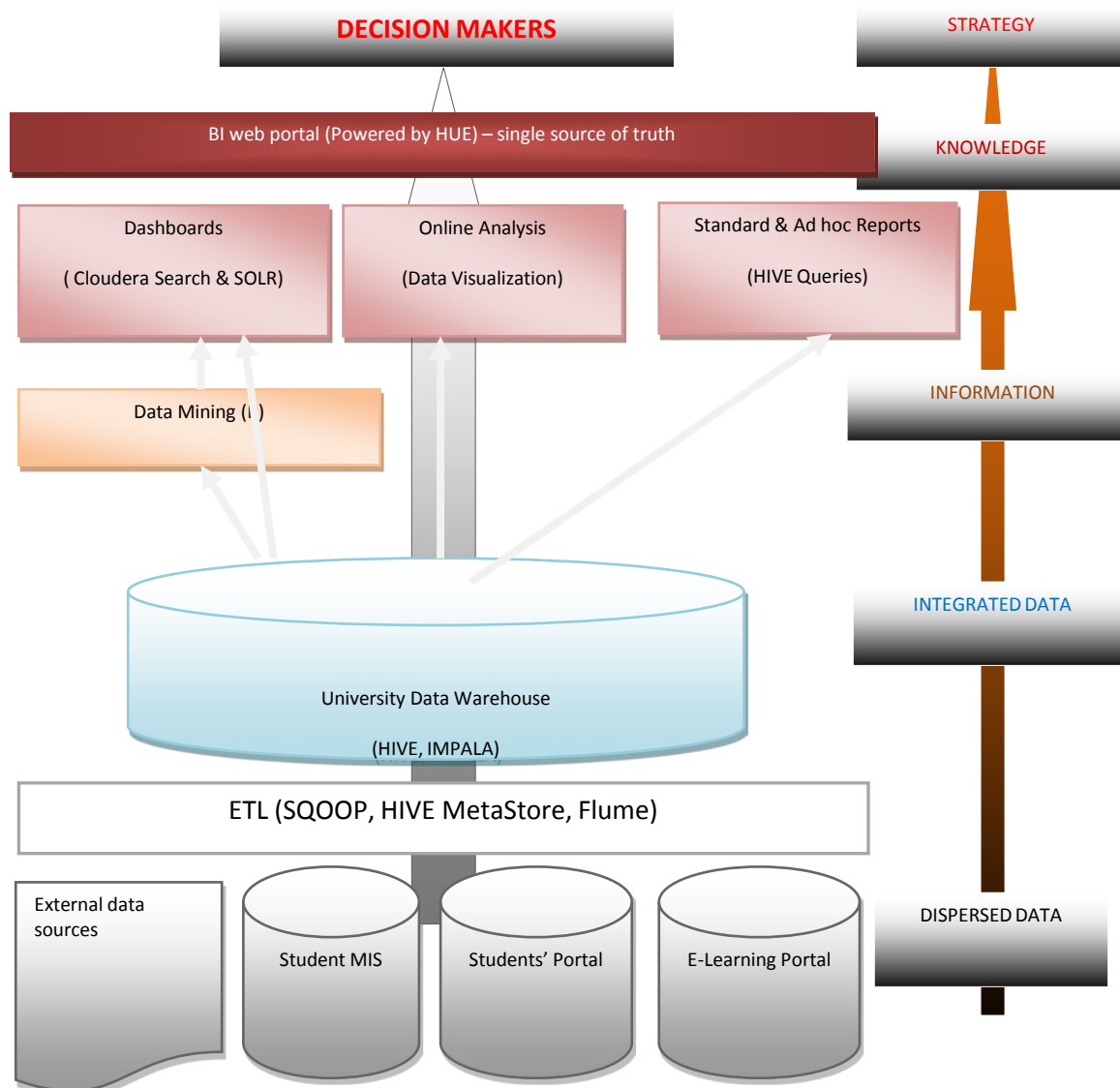


METHODOLOGY



METHODOLOGY

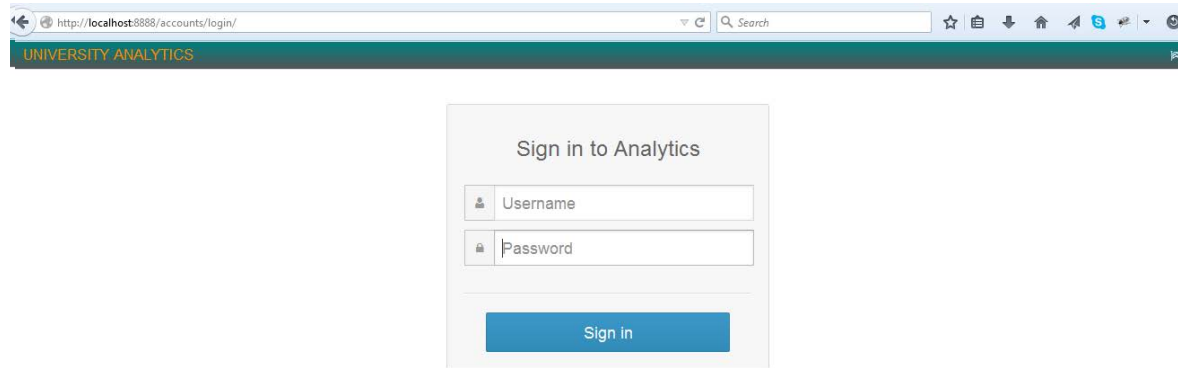
ANALYTICAL SYSTEM ARCHITECTURE



METHODOLOGY

SYSTEM FEATURES

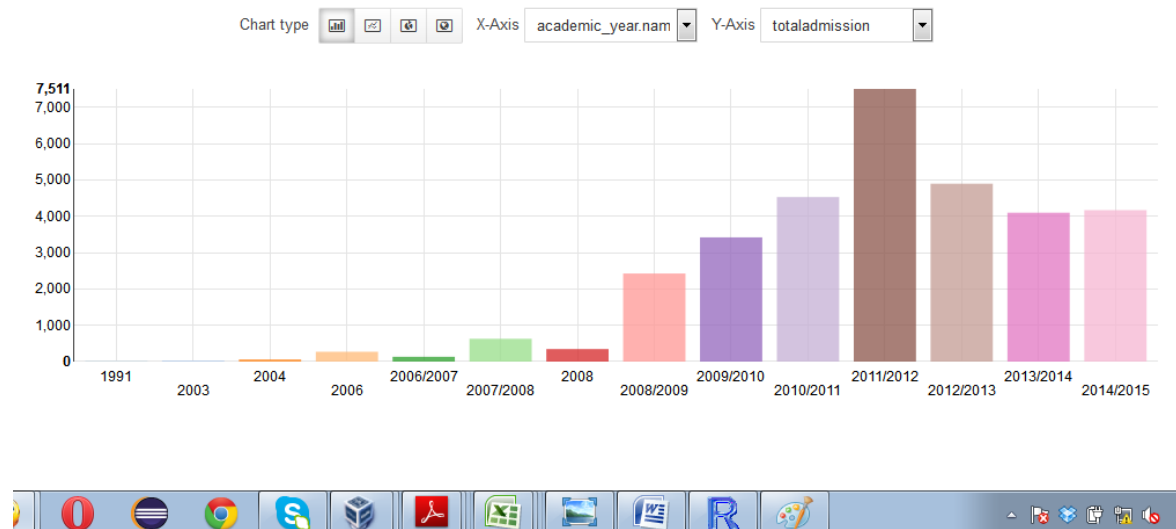
BI PORTAL LOGIN



METHODOLOGY

SYSTEM FEATURES

SAMPLE DASHBOARD



Student Admission Pattern

RESULTS

- ❑ THERE WAS CONSISTENT AND RELIABLE INFORMATION (USERS MORE CONFIDENT ON THE INFORMATION GENERATED)
- ❑ THERE WAS FASTER ACCESS TO REQUIRED INFORMATION (EXAMPLE STAFF COMPOSITION REPORT THAT WAS REQUIRED BY THE AUDITOR GENERAL COULD BE GENERATED WITHIN HOURS INSTEAD OF WEEKS)
- ❑ THERE WAS A STUDENT PROFILE REPORT THAT WAS USED TO DETERMINE THE NEEDINESS LEVELS OF STUDENTS IN ALLOCATION OF ROOMS FOR ACCOMMODATION
- ❑ STAFF ETHNIC BALANCE WAS USED IN INFLUENCING THE HIRING OF NEW STAFF MEMBERS
- ❑ USERS WERE ABLE TO IDENTIFY NEW INFORMATION THAT WAS NOT PREVIOUSLY AVAILABLE: FOR EXAMPLE STAFF DISTRIBUTION BASED ON JOB GROUPS EXPOSED A JOB GROUP THAT WAS CROWDED

CONCLUSION



ALTHOUGH NOT ALL DATA SOURCES WERE INCLUDED, THERE WERE TANGIBLE BENEFITS IN THE TECHNICAL UNIVERSITY

HENCE, REPLICATION CAN BE DONE IN OTHER UNIVERSITIES



REPLICATION TO OTHER SECTORS OF THE ECONOMY

CHALLENGES FACED

- ❑ DATA CLEANING AND TRANSFORMATION DIFFICULT
- ❑ COMPLETELY UNRETRIVABLE DATA
- ❑ MANAGEMENT STILL NOT CONFIDENT ON EVIDENT-BASED DECISION MAKING
- ❑ LACK OF DATA ANALYTICAL SKILLS IN UNIVERSITY COMMUNITY
- ❑ DATA FROM EXTERNAL SOURCES LIKE EMAILS, SOCIAL MEDIA STILL DO NOT HAVE A LOT OF RELEVANT INFORMATION REQUIRED BY MANAGEMENT
- ❑ MODELING FOR PREDICTIVE ANALYTICS DIFFICULT DUE TO MISSING/ LESS DATA

- ❑ PUBLIC SERVICE (PUBLIC PERCEPTION ON GOVERNMENT)
- ❑ HEALTH SECTOR (DISEASE DIAGNOSIS, POSSIBILITY OF OUTBREAKS, FAKE DRUGS)
- ❑ BANKING SECTOR (FRAUD DETECTION, POSSIBILITY OF LOAN DEFAULTING)
- ❑ INSURANCE (FRAUDULENT CLAIMS)

SECURITY

REFERENCES

- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS quarterly*, 36(4), 1165-1188.
- Das, S., & Kalita, H. K. (2015). Advanced Dimensionality Reduction Method for Big Data. *Research Advances in the Integration of Big Data and Smart Computing*, 198.
- Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Harvard Business Press.
- Evelson, B., & Norman, N. (2008). Topic overview: Business intelligence. *Forrester research*, 61.
- Foley, É., & Guillemette, M. G. (2012). What is business intelligence?. *Organizational Applications of Business Intelligence Management: Emerging Trends: Emerging Trends*, 52.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Herring, J. P. (1992). The role of intelligence in formulating strategy. *Journal of Business Strategy*, 13(5), 54-60.
- Malik, S. (2005). *Enterprise dashboards: design and best practices for IT*. John Wiley & Sons.
- Olszak, C. M., & Ziemba, E. (2007). Approach to building and implementing business intelligence systems. *Interdisciplinary Journal of Information, Knowledge, and Management*, 2(2007), 134-148.
- Rodgers, M. L., Talbut, M. H., & Baranovic, K. (2015, February). Data Analytics for Beginners. In *The National Teaching & Learning Forum* (Vol. 24, No. 2, pp. 10-11).