

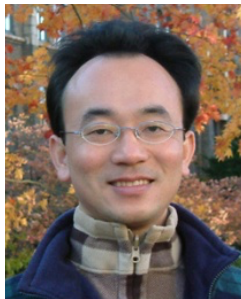
資料探勘 (Data Mining)

ABC：人工智慧，大數據，雲端運算 (ABC: AI, Big Data, Cloud Computing)

1092DM02

MBA, IM, NTPU (M5026) (Spring 2021)

Tue 2, 3, 4 (9:10-12:00) (B8F40)



Min-Yuh Day

戴敏育

Associate Professor

副教授

Institute of Information Management, National Taipei University

國立臺北大學 資訊管理研究所

<https://web.ntpu.edu.tw/~myday>

2021-03-2



課程大綱 (Syllabus)

- | 週次 (Week) | 日期 (Date) | 內容 (Subject/Topics) |
|-----------|------------|--|
| 1 | 2021/02/23 | 資料探勘介紹 (Introduction to data mining) |
| 2 | 2021/03/02 | ABC：人工智慧，大數據，雲端運算
(ABC: AI, Big Data, Cloud Computing) |
| 3 | 2021/03/09 | Python 資料探勘的基礎
(Foundations of Data Mining in Python) |
| 4 | 2021/03/16 | 資料科學與資料探勘：發現，分析，可視化和呈現數據
(Data Science and Data Mining:
Discovering, Analyzing, Visualizing and Presenting Data) |
| 5 | 2021/03/23 | 非監督學習：關聯分析，購物籃分析
(Unsupervised Learning: Association Analysis,
Market Basket Analysis) |
| 6 | 2021/03/30 | 資料探勘個案研究 I
(Case Study on Data Mining I) |

課程大綱 (Syllabus)

- | 週次 (Week) | 日期 (Date) | 內容 (Subject/Topics) |
|-----------|------------|---|
| 7 | 2021/04/06 | 非監督學習：集群分析，行銷市場區隔
(Unsupervised Learning: Cluster Analysis, Market Segmentation) |
| 8 | 2021/04/13 | 監督學習：分類和預測
(Supervised Learning: Classification and Prediction) |
| 9 | 2021/04/20 | 期中報告 (Midterm Project Report) |
| 10 | 2021/04/27 | 監督學習：分類和預測
(Supervised Learning: Classification and Prediction) |
| 11 | 2021/05/04 | 機器學習和深度學習
(Machine Learning and Deep Learning) |
| 12 | 2021/05/11 | 卷積神經網絡
(Convolutional Neural Networks) |

課程大綱 (Syllabus)

週次 (Week) 日期 (Date) 內容 (Subject/Topics)

- 13 2021/05/18 資料探勘個案研究 II
(Case Study on Data Mining II)
- 14 2021/05/25 遞歸神經網絡
(Recurrent Neural Networks)
- 15 2021/06/01 強化學習
(Reinforcement Learning)
- 16 2021/06/08 社交網絡分析
(Social Network Analysis)
- 17 2021/06/15 期末報告 I (Final Project Report I)
- 18 2021/06/22 期末報告 II (Final Project Report II)

**ABC:
AI,
Big Data,
Cloud Computing**

Outline

- **AI**
- **Big Data**
- **Cloud Computing**

FinTech ABCD

AI

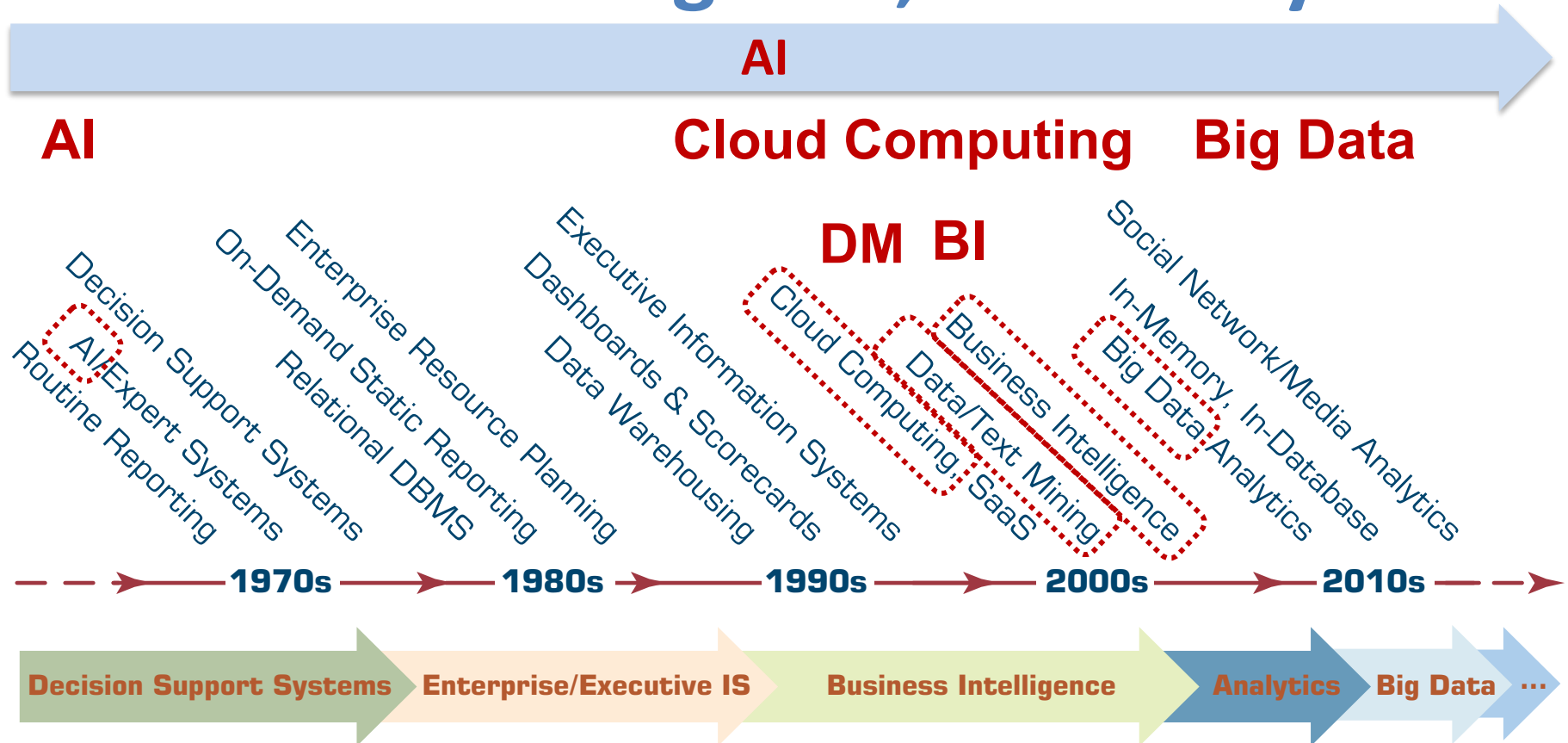
Block Chain

Cloud Computing

Big **D**ata

AI, Big Data, Cloud Computing

Evolution of Decision Support, Business Intelligence, and Analytics



Artificial Intelligence (A.I.) Timeline

A.I. TIMELINE

SYZYG

1950

TURING TEST

Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence

1955

A.I. BORN

Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe "the science and engineering of making intelligent machines"

1961

UNIMATE

First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964

ELIZA

Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966

SHAKY

The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER

Many false starts and dead-ends leave A.I. out in the cold

1997

DEEP BLUE

Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998

KISMET

Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot insofar as it detects and responds to people's feelings

1999

AIBO

Sony launches first consumer robot pet dog AiBO (AI robot) with skills and personality that develop over time

2002

ROOMBA

First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes

2011

SIRI

Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S

2011

WATSON

IBM's question answering computer Watson wins first place on popular \$1M prize television quiz show Jeopardy

2014

EUGENE

Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human

2014

ALEXA

Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks

2016

TAY

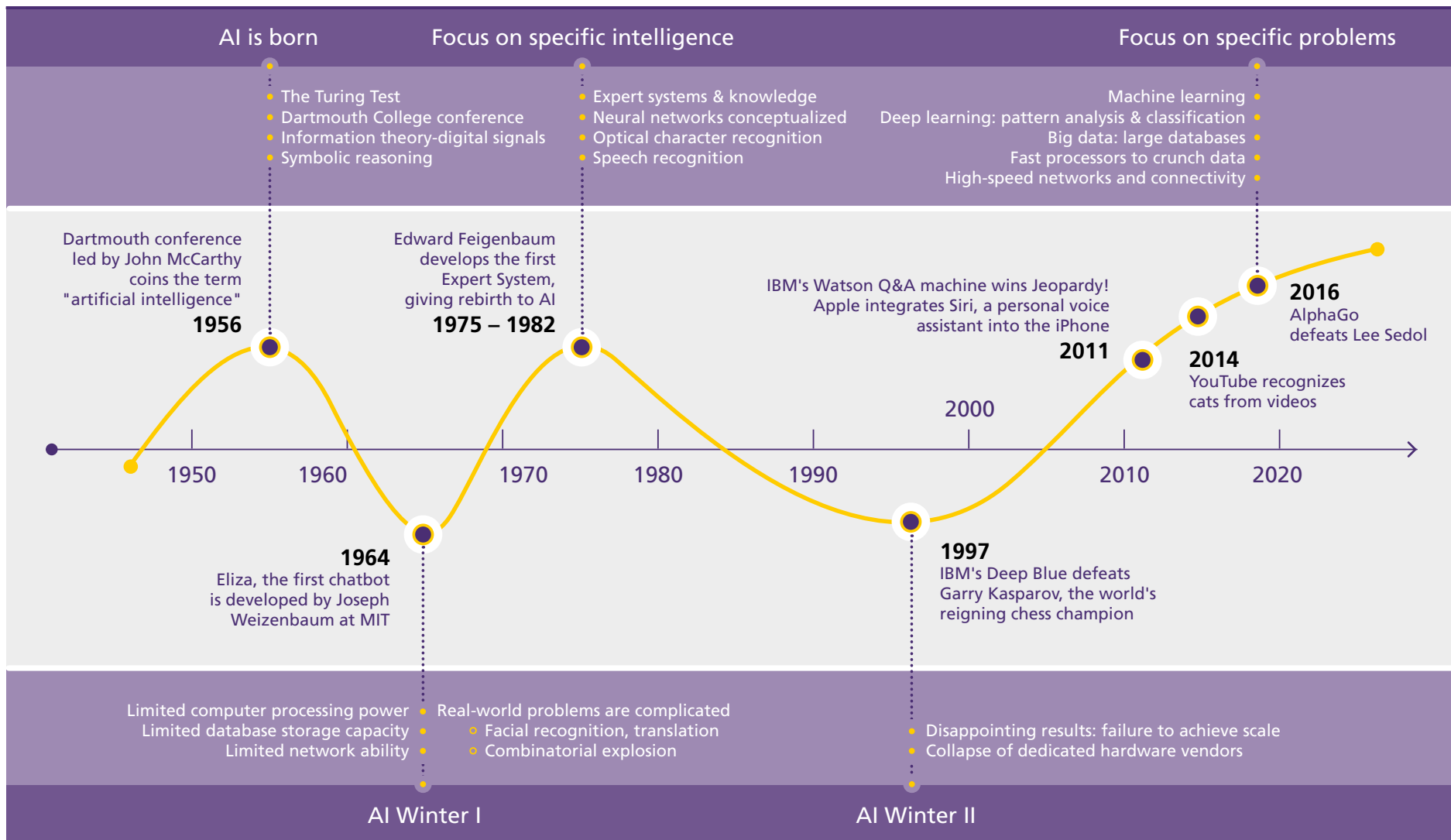
Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments

2017

ALPHAGO

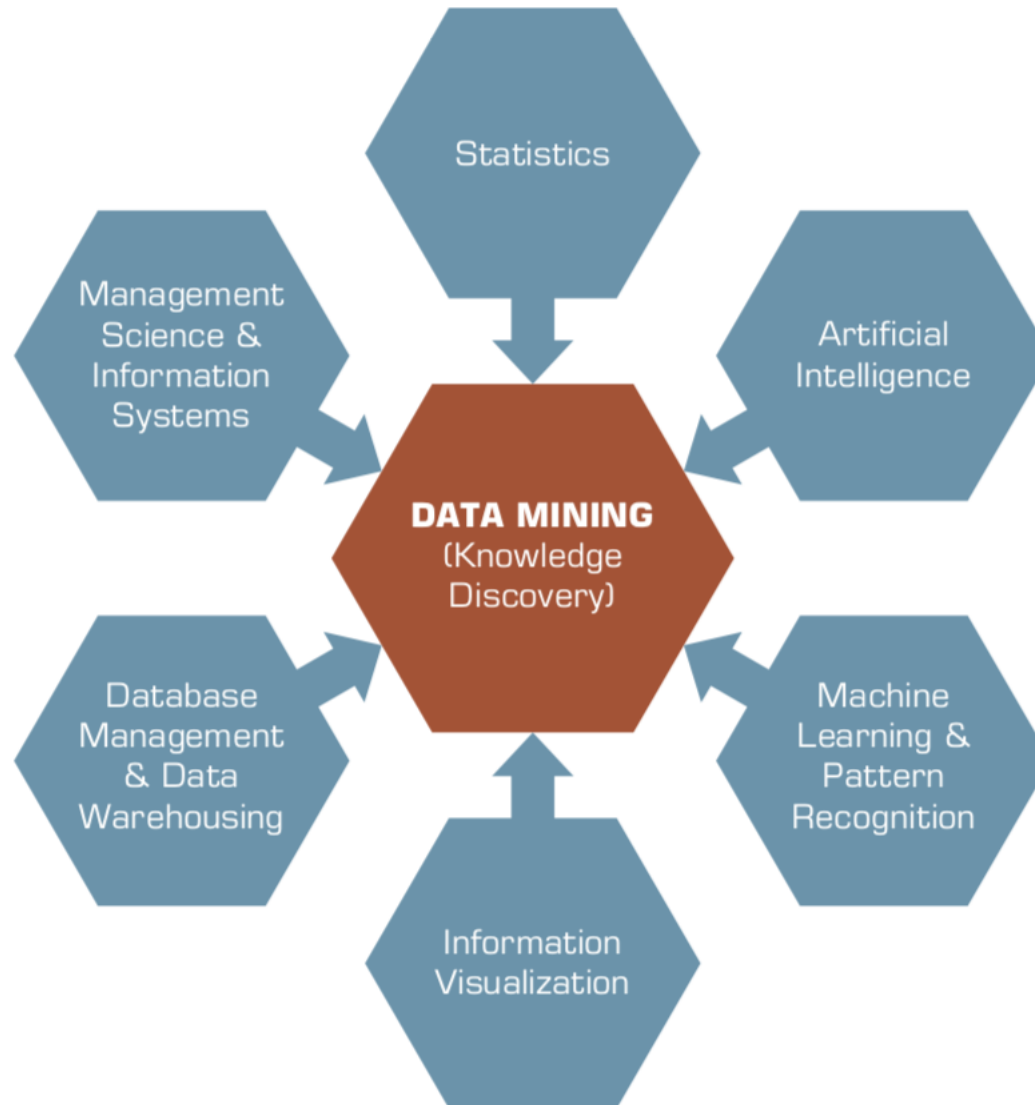
Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2^{170}) of possible positions

The Rise of AI

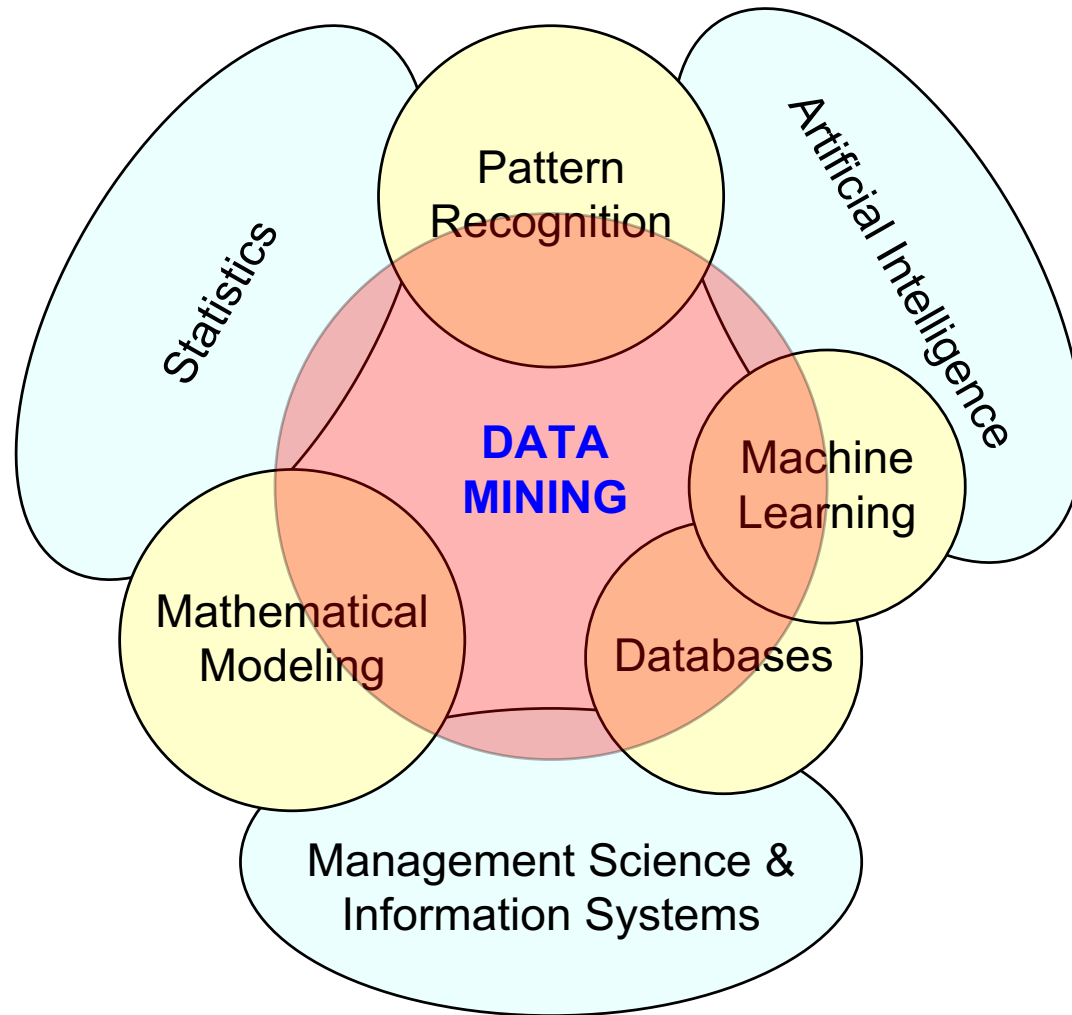


Data Mining

Is a Blend of Multiple Disciplines



Data Mining at the Intersection of Many Disciplines



Data Mining Tasks & Methods

Prediction

Data Mining Tasks & Methods	Data Mining Algorithms	Learning Type
Prediction		
Classification	Decision Trees, Neural Networks, Support Vector Machines, kNN, Naïve Bayes, GA	Supervised
Regression	Linear/Nonlinear Regression, ANN, Regression Trees, SVM, kNN, GA	Supervised
Time series	Autoregressive Methods, Averaging Methods, Exponential Smoothing, ARIMA	Supervised
Association		
Market-basket	Apriori, OneR, ZeroR, Eclat, GA	Unsupervised
Link analysis	Expectation Maximization, Apriori Algorithm, Graph-Based Matching	Unsupervised
Sequence analysis	Apriori Algorithm, FP-Growth, Graph-Based Matching	Unsupervised
Segmentation		
Clustering	k-means, Expectation Maximization (EM)	Unsupervised
Outlier analysis	k-means, Expectation Maximization (EM)	Unsupervised

Association

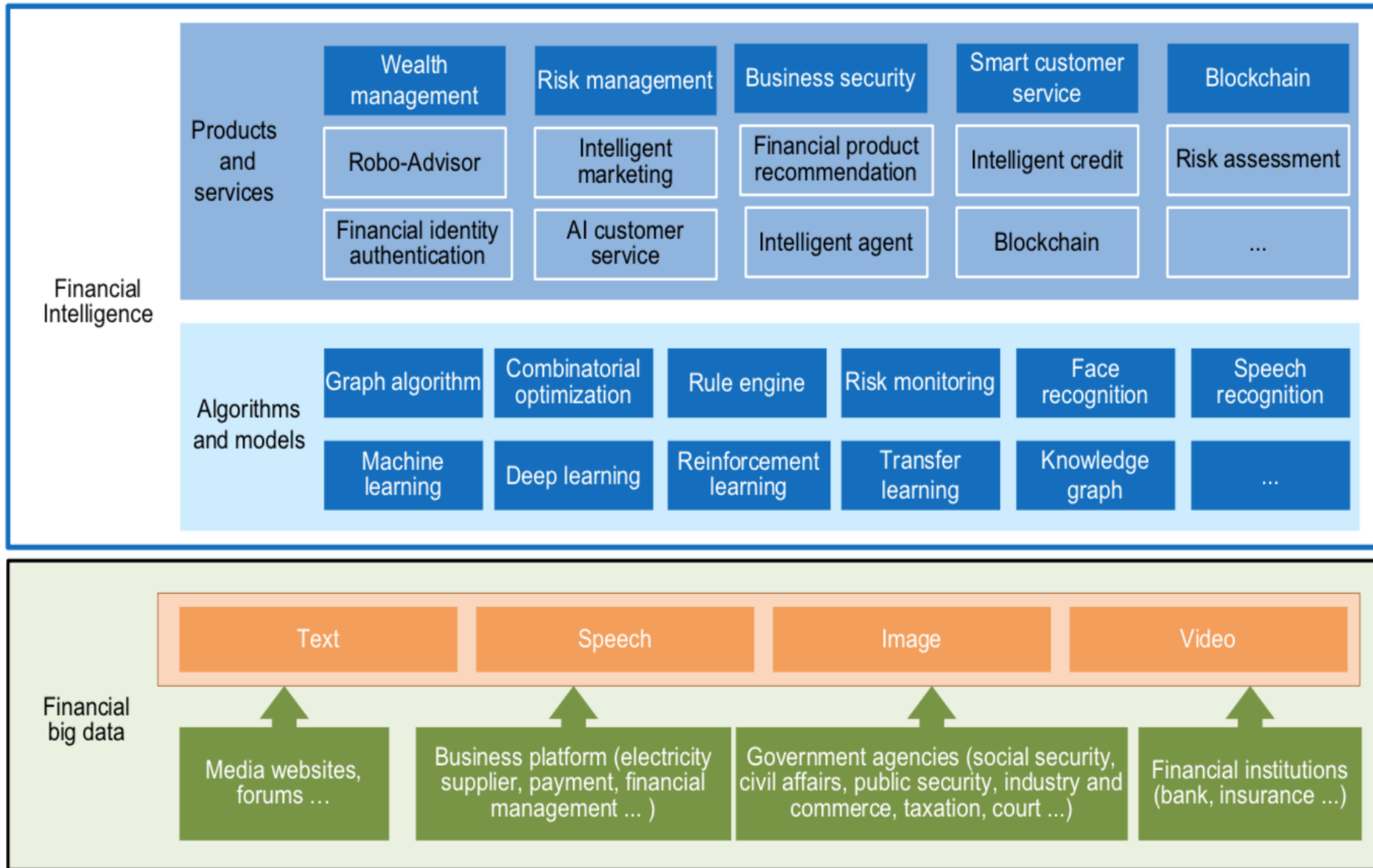
Segmentation

AI 2.0

a new generation of AI
based on the
novel information environment of
major changes and
the development of
new goals.

FinBrain: when Finance meets AI 2.0

(Zheng et al., 2019)

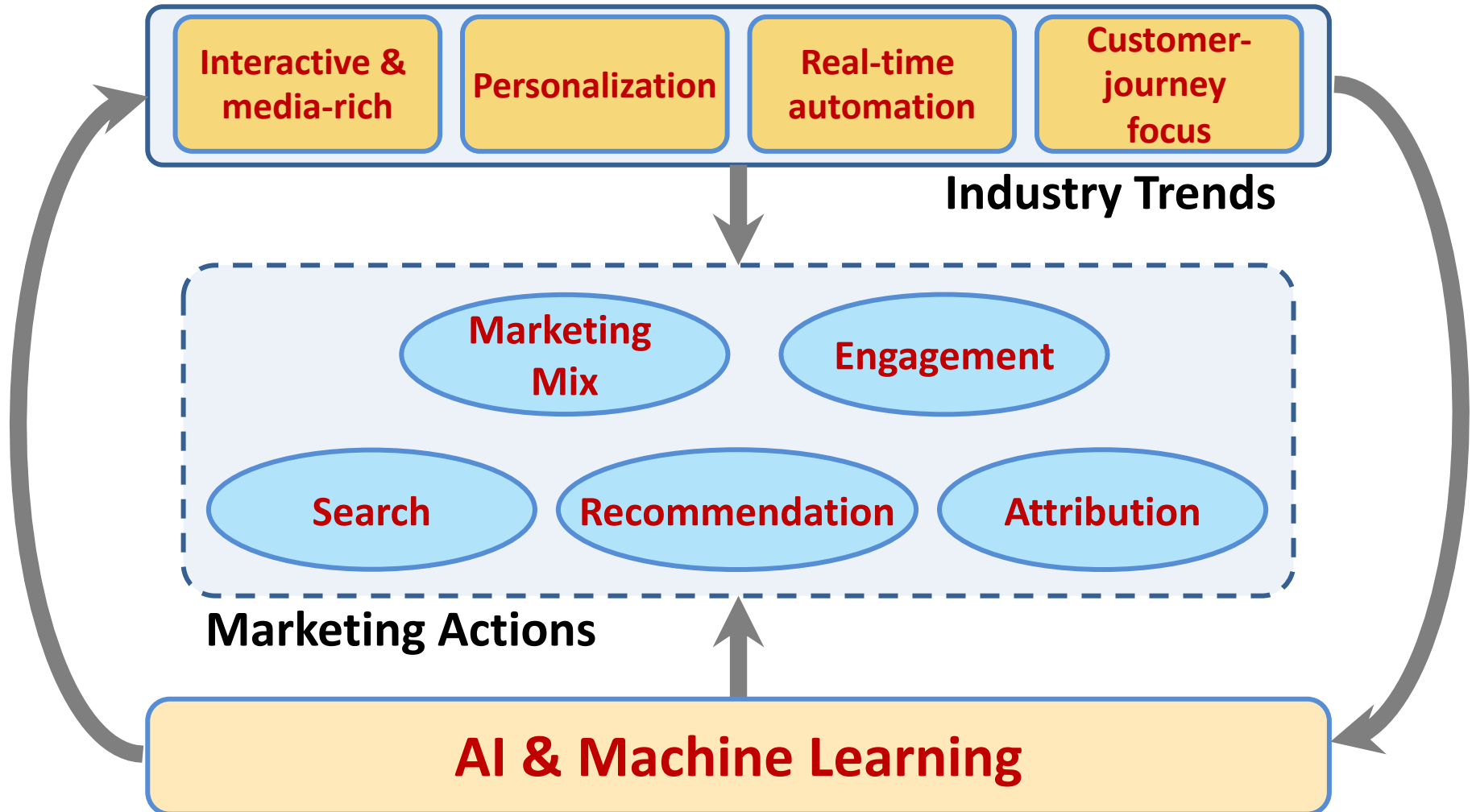


Technology-driven Financial Industry Development

Development stage	Driving technology	Main landscape	Inclusive finance	Relationship between technology and finance
Fintech 1.0 (financial IT)	Computer	Credit card, ATM, and CRMS	Low	Technology as a tool
Fintech 2.0 (Internet finance)	Mobile Internet	Marketplace lending, third-party payment, crowdfunding, and Internet insurance	Medium	Technology-driven change
Fintech 3.0 (financial intelligence)	AI, Big Data, Cloud Computing, Blockchain	Intelligent finance	High	Deep fusion

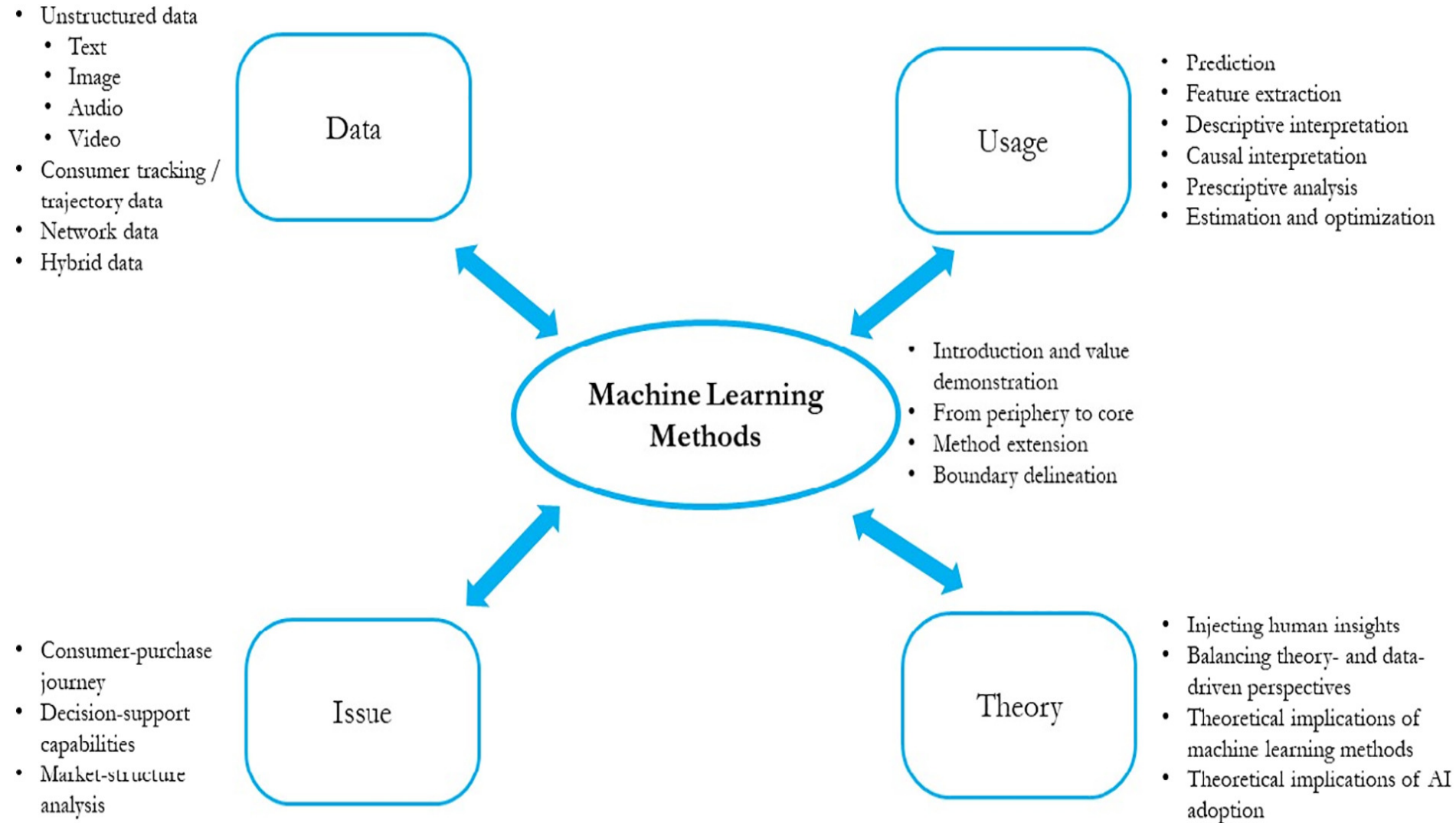
AI-driven Marketing

(Ma and Sun, 2020)



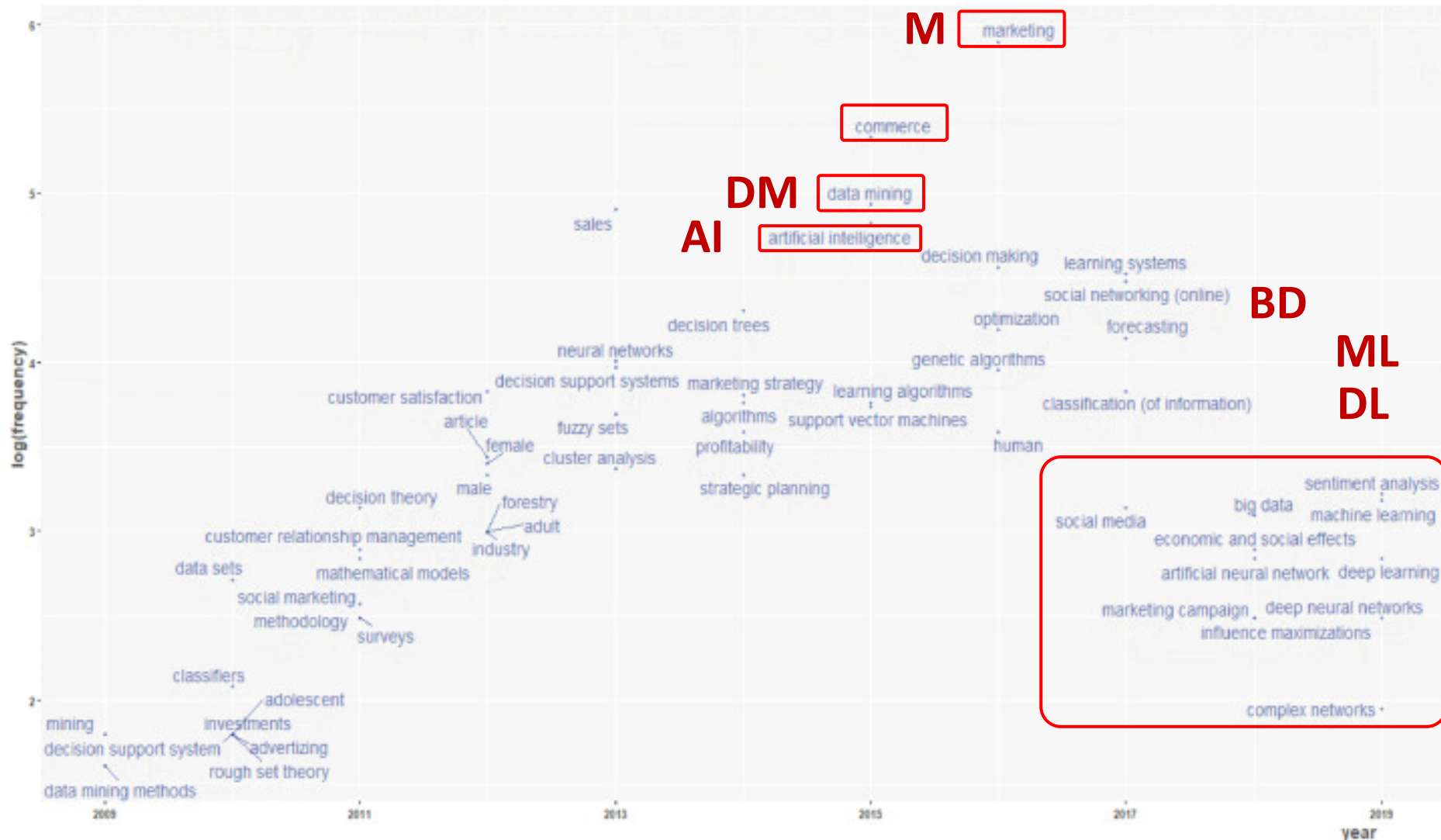
Machine Learning in Marketing Research

(Ma and Sun, 2020)



Artificial Intelligence in Marketing

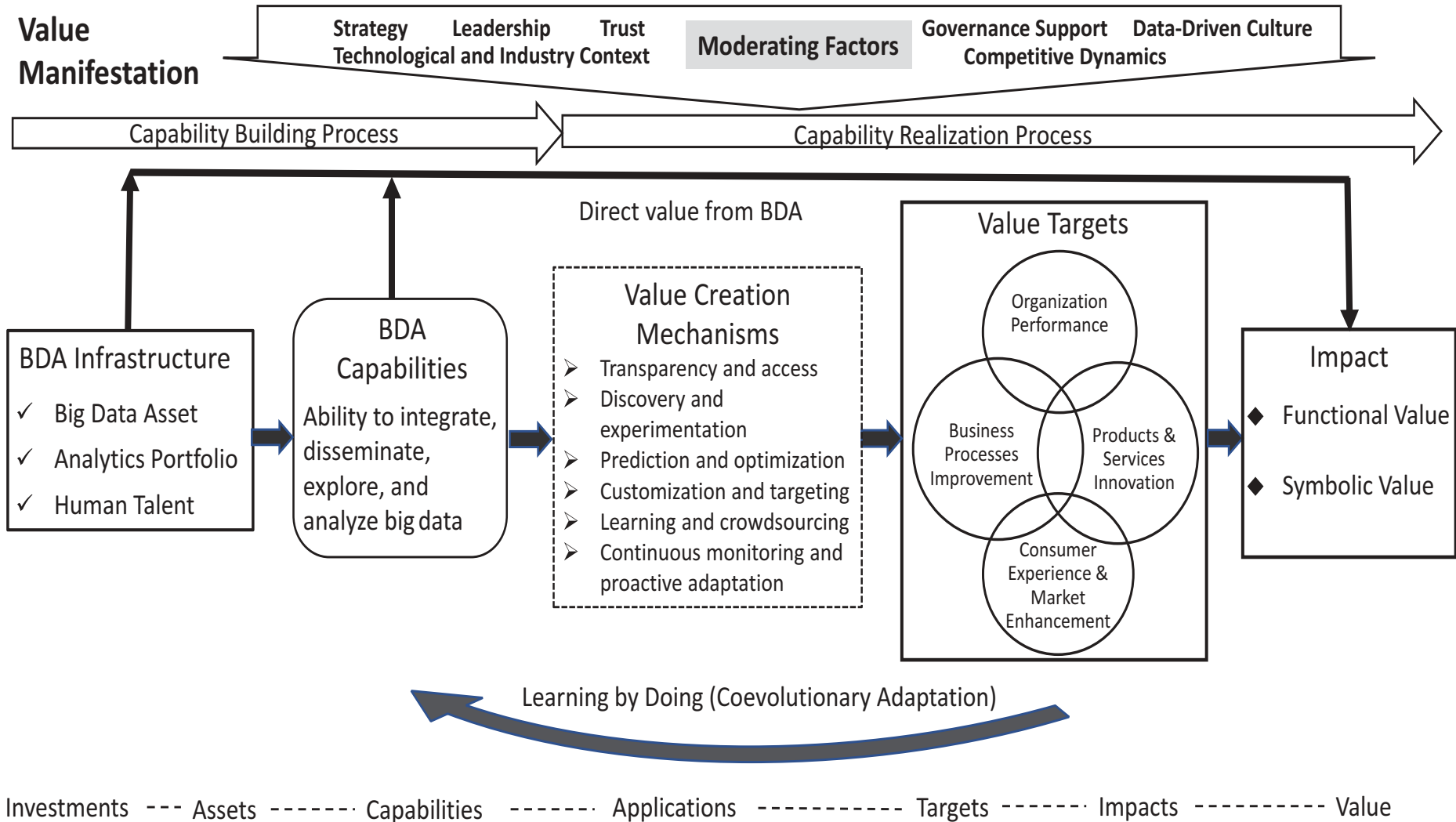
(Verma et al., 2021)



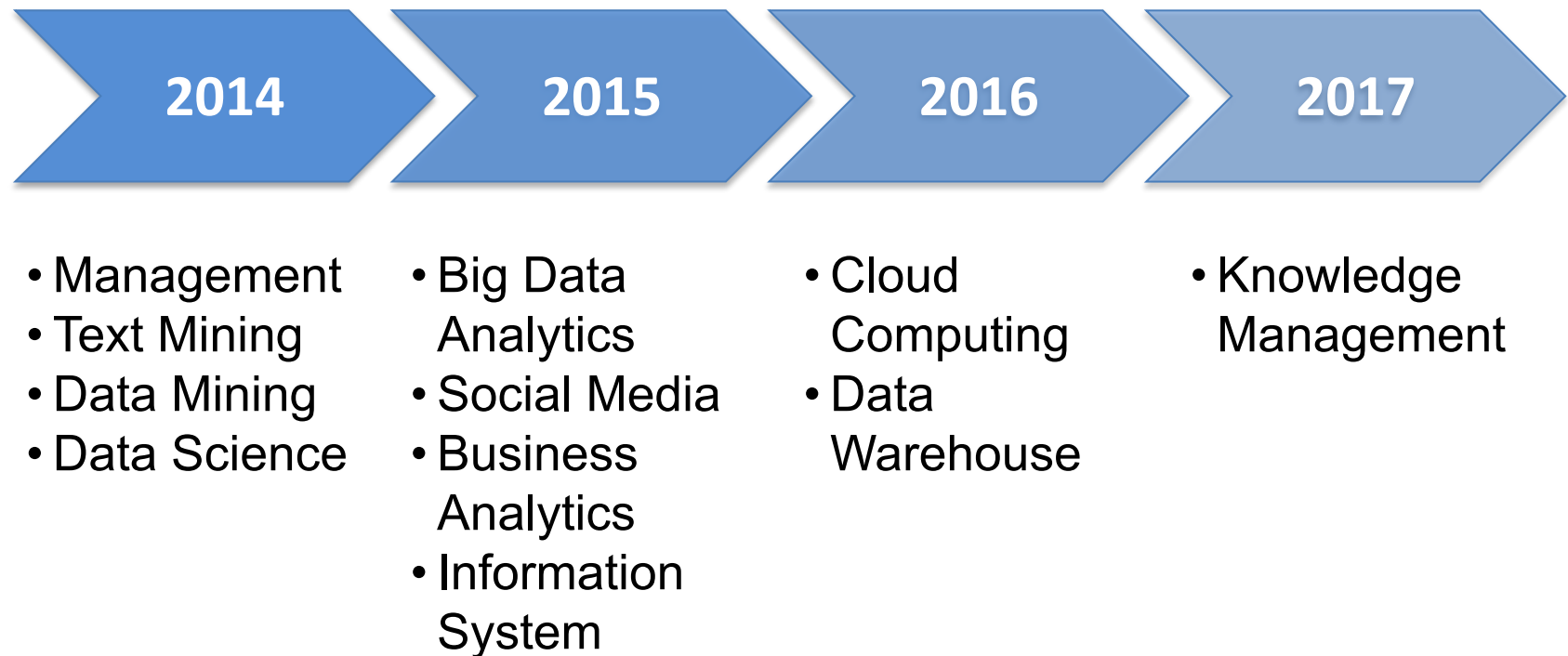
Source: Sanjeev Verma, Rohit Sharma, Subhamay Deb, and Debojit Maitra (2021), "Artificial intelligence in marketing: Systematic review and future research direction." International Journal of Information Management Data Insights (2021): 100002.

Value Creation by Big Data Analytics

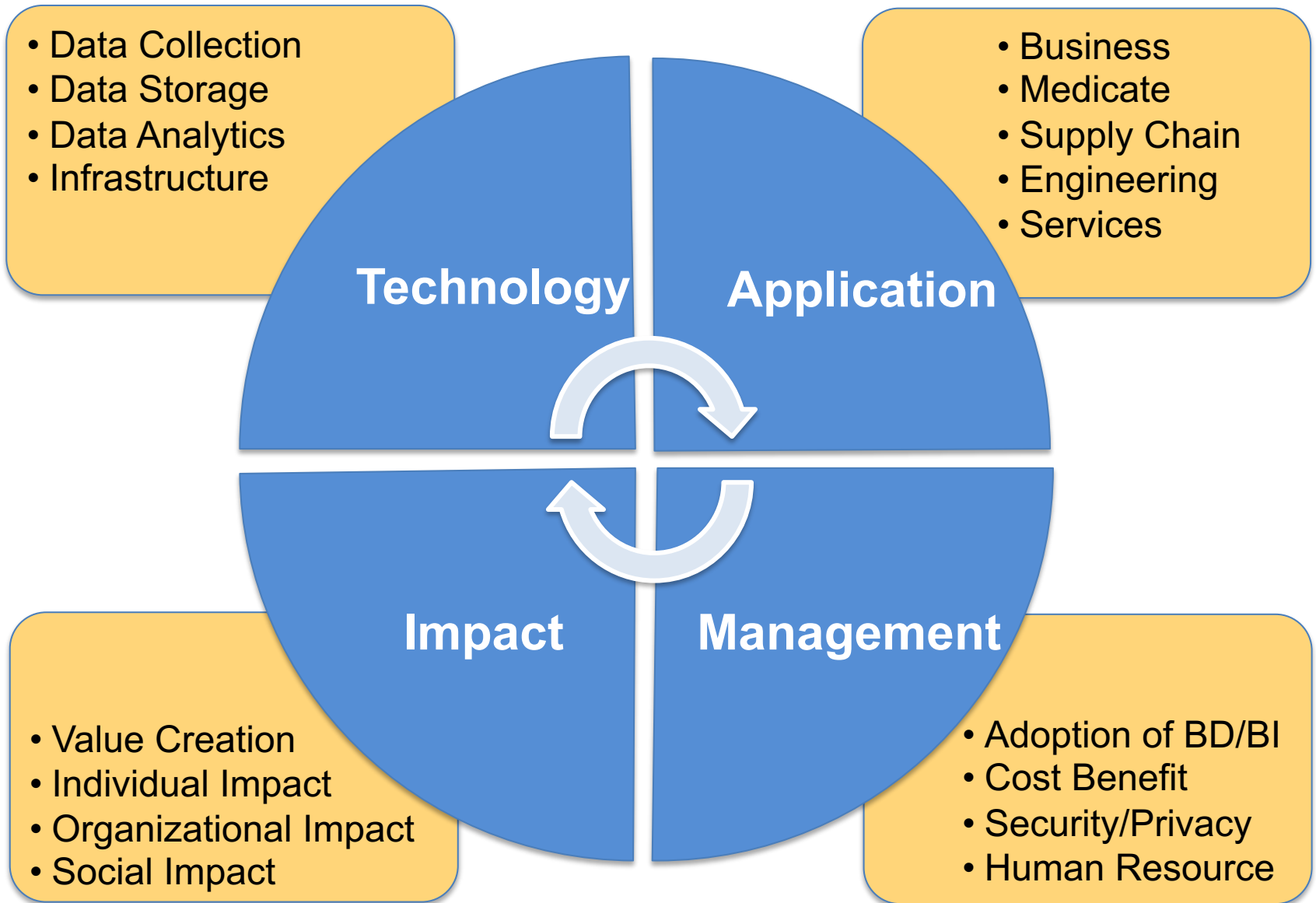
(Grover et al., 2018)



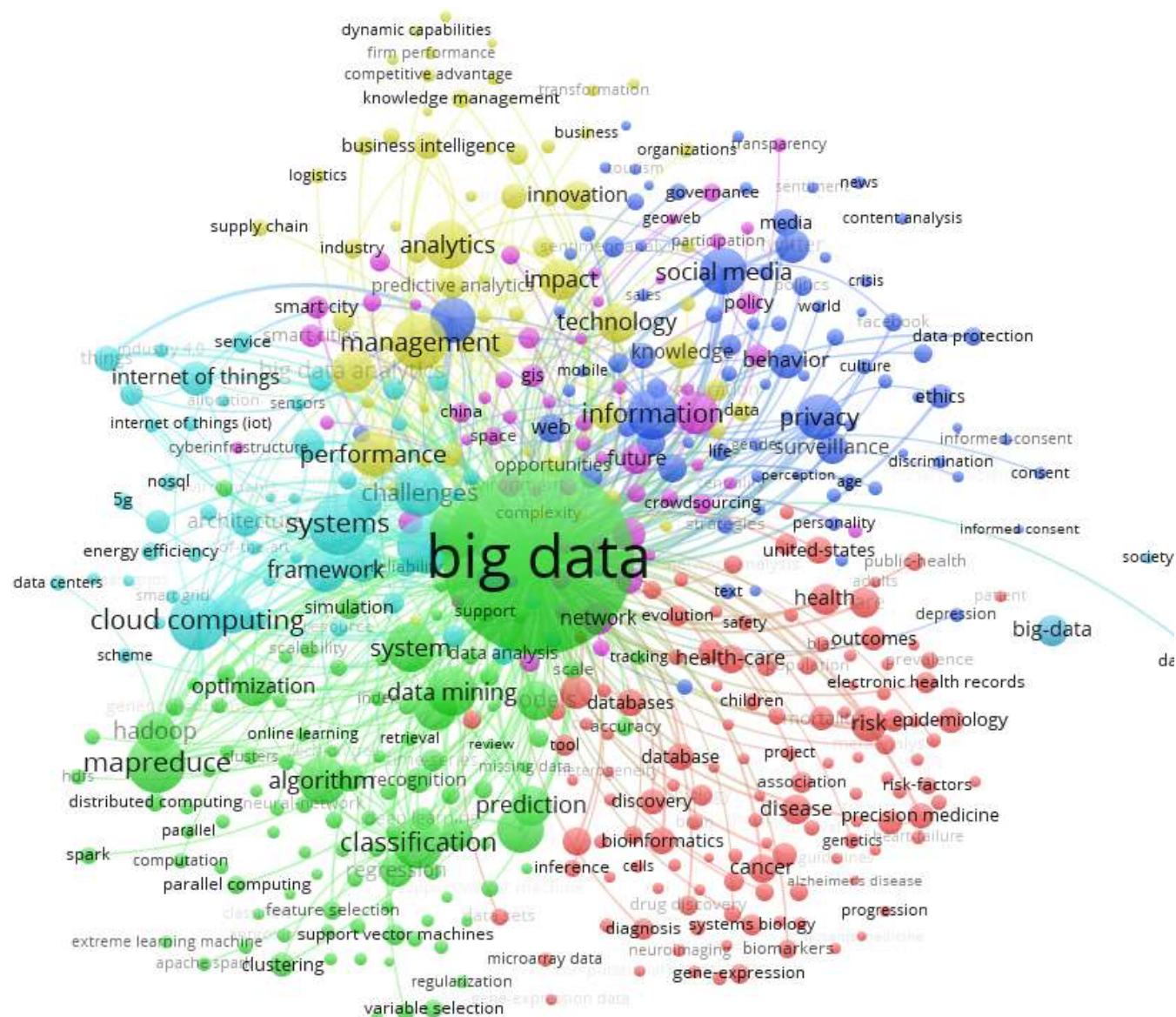
Evolution of top keywords in “BD & BI” publications



Framework for BD and BI Research



Business Intelligence and Big Data analytics



Source: Ting-Peng Liang and Yu-Hsi Liu (2018), "Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study", *Expert Systems with Applications*, Volume 111, 30, 2018, pp. 2-10

Ai

Definition of Artificial Intelligence (A.I.)

Artificial Intelligence

**“... the science and
engineering
of
making
intelligent machines”
(John McCarthy, 1955)**

Artificial Intelligence

**“... technology that
thinks and acts
like humans”**

Artificial Intelligence

**“... intelligence
exhibited by machines
or software”**

4 Approaches of AI

Thinking Humanly	Thinking Rationally
Acting Humanly	Acting Rationally

4 Approaches of AI

2.

**Thinking Humanly:
The Cognitive
Modeling Approach**

3.

**Thinking Rationally:
The “Laws of Thought”
Approach**

1.

**Acting Humanly:
The Turing Test
Approach** (1950)

4.

**Acting Rationally:
The Rational Agent
Approach**

AI Acting Humanly: The Turing Test Approach (Alan Turing, 1950)

- Knowledge Representation
- Automated Reasoning
- Machine Learning (ML)
 - Deep Learning (DL)
- Computer Vision (Image, Video)
- Natural Language Processing (NLP)
- Robotics

Artificial Intelligence: A Modern Approach

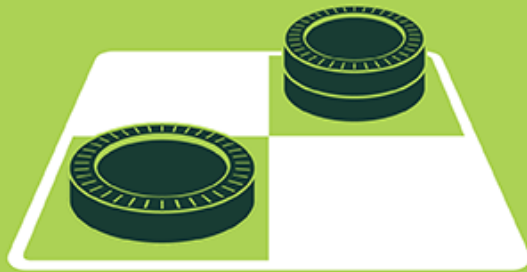
1. Artificial Intelligence
2. Problem Solving
3. Knowledge and Reasoning
4. Uncertain Knowledge and Reasoning
5. Learning
6. Communicating, Perceiving, and Acting
7. Philosophy and Ethics of AI

Artificial Intelligence

Machine Learning & Deep Learning

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

AI, ML, DL

Artificial Intelligence (AI)

Machine Learning (ML)

Supervised
Learning

Unsupervised
Learning

Deep Learning (DL)

CNN

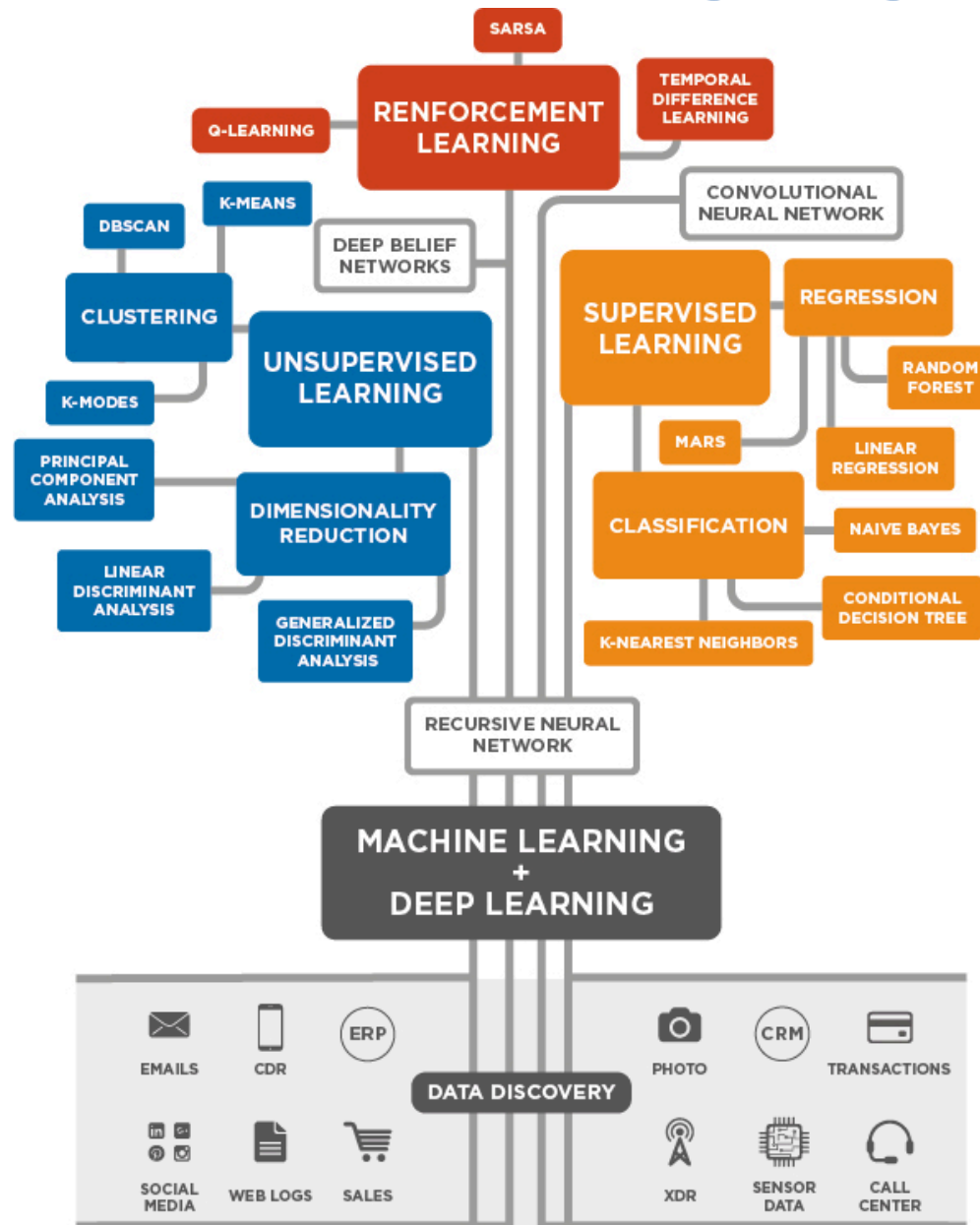
RNN LSTM GRU

GAN

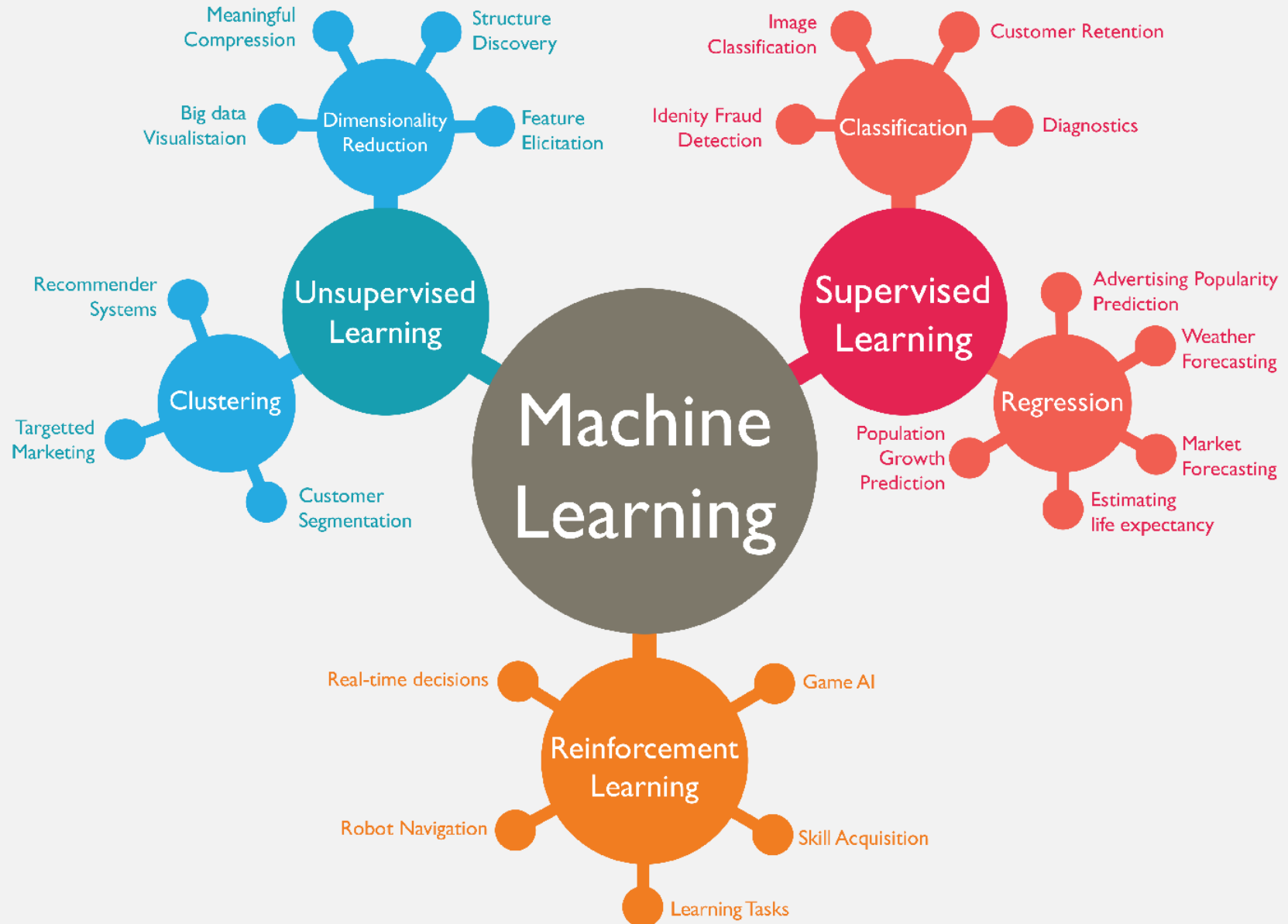
Semi-supervised
Learning

Reinforcement
Learning

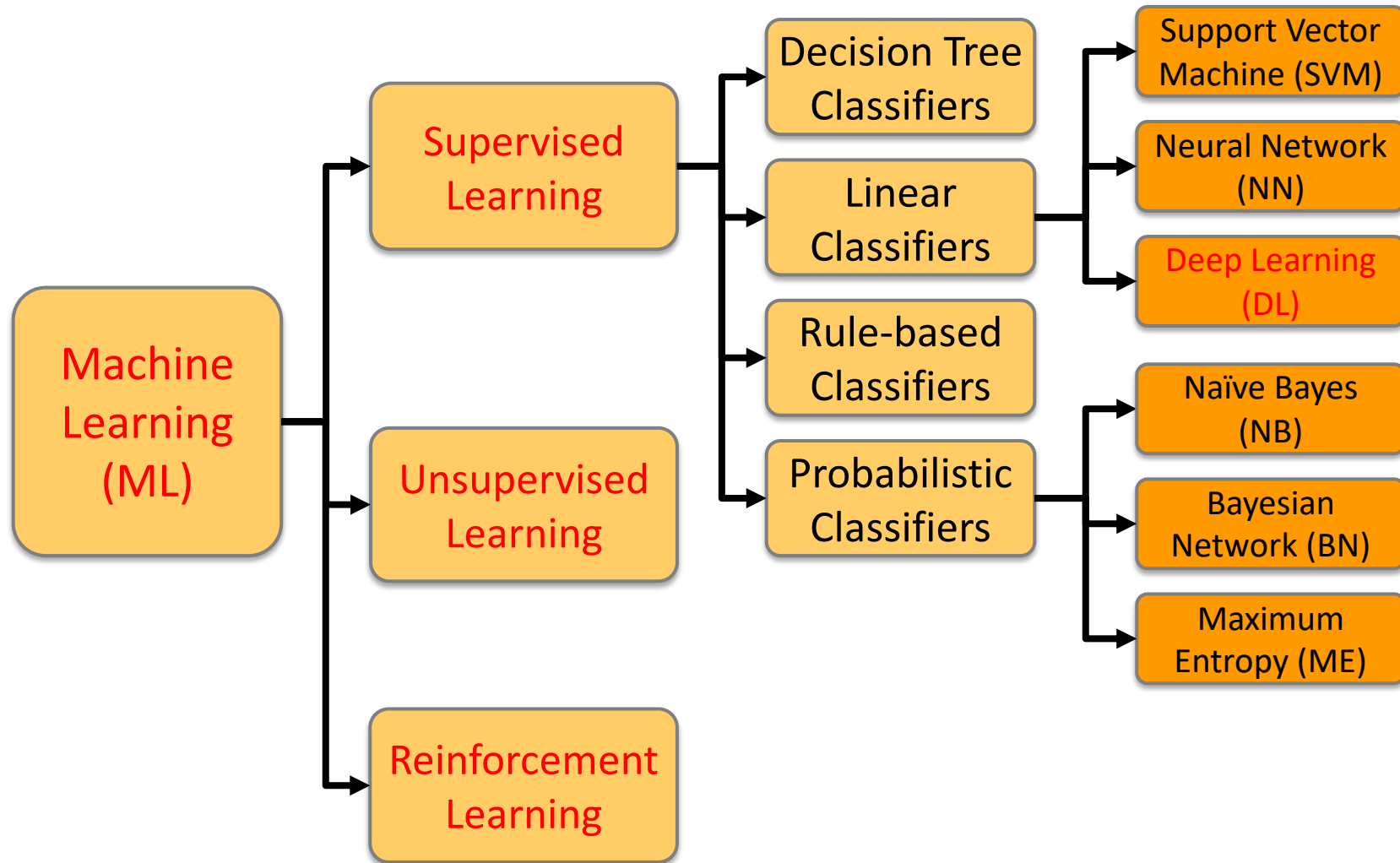
3 Machine Learning Algorithms



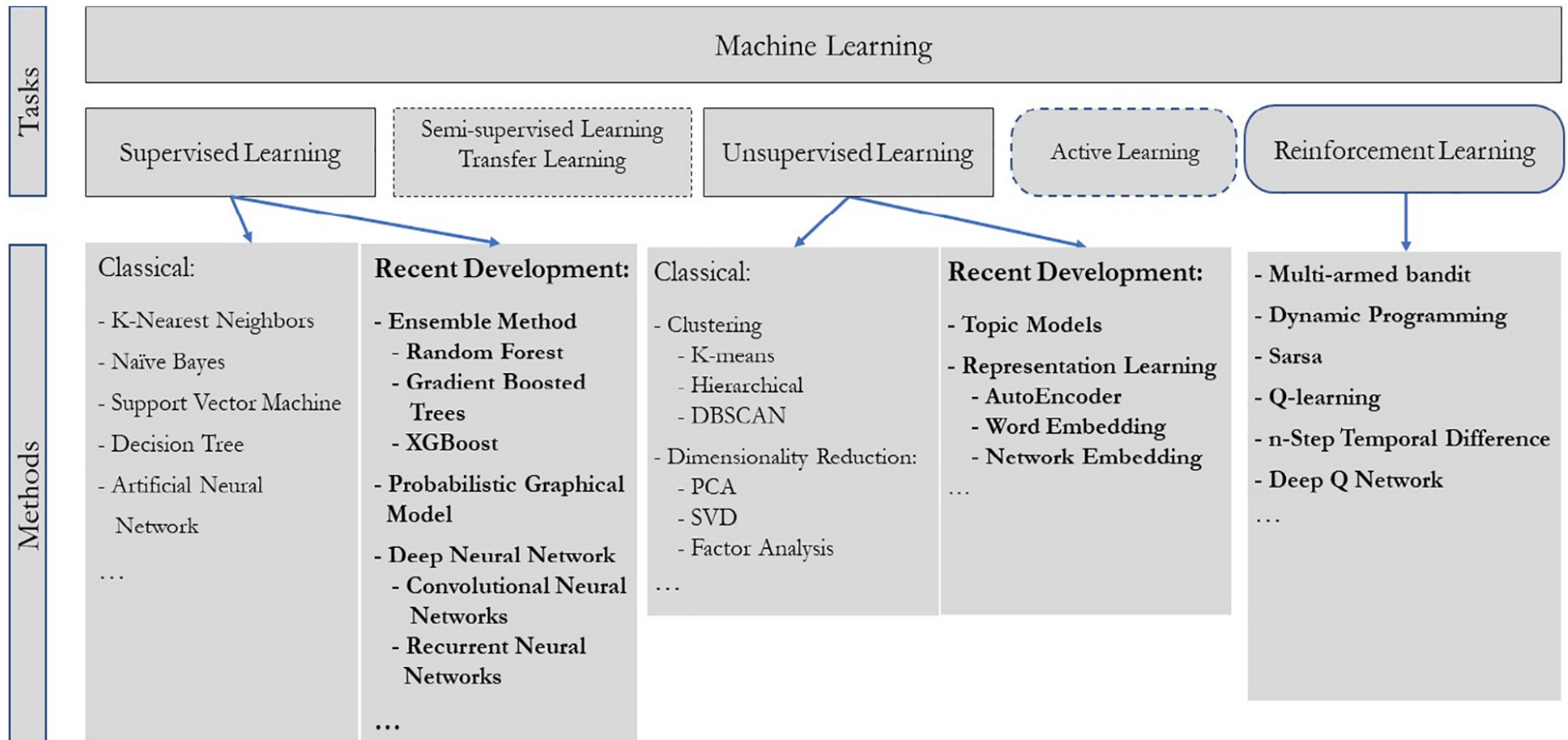
Machine Learning (ML)



Machine Learning (ML) / Deep Learning (DL)



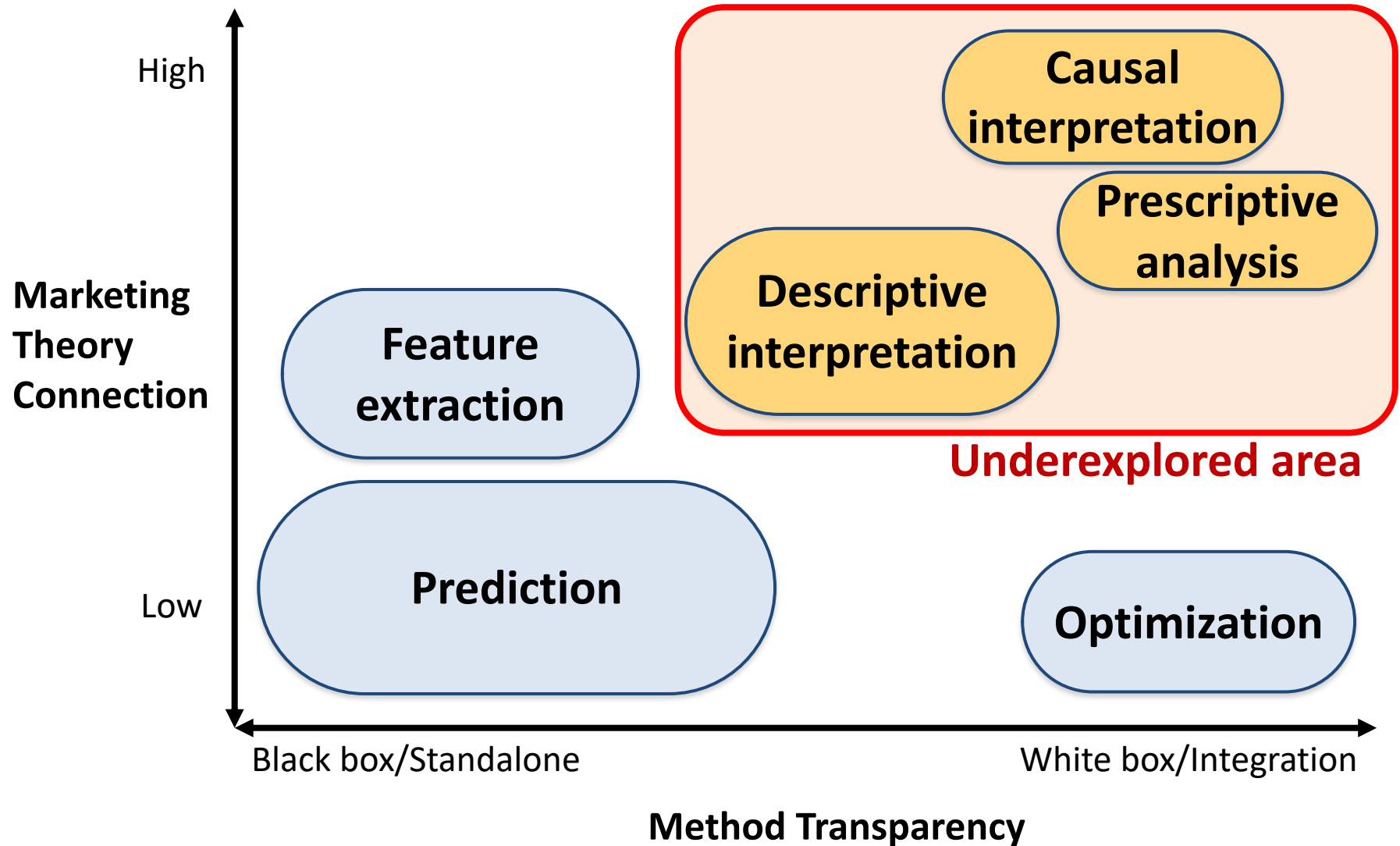
Machine Learning Tasks and Methods



Note: Several entries in the diagram, e.g. word embedding or multi-armed bandit, refer to specific problem formulations for which a collection of methods exist.

: Tasks that take input data as given
 : Tasks that involve interactive data acquisition
 Dashed border: methods not elaborated in paper text
 Bold type: highlights recent developments

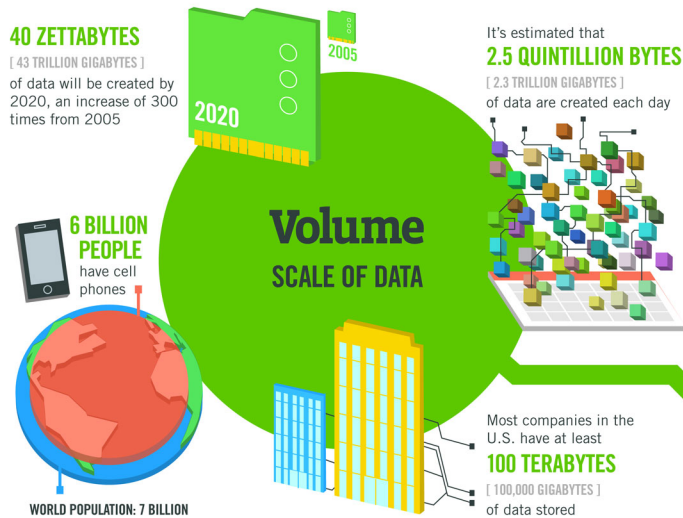
Machine Learning Methods in Marketing



Big Data

Big Data Analytics and Data Mining

Big Data 4 V



The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015
4.4 MILLION IT JOBS
will be created globally to support big data, with 1.9 million in the United States



As of 2011, the global size of data in healthcare was estimated to be

150 EXABYTES
[161 BILLION GIGABYTES]



30 BILLION PIECES OF CONTENT
are shared on Facebook every month



Variety

DIFFERENT FORMS OF DATA

By 2014, it's anticipated there will be **420 MILLION WEARABLE, WIRELESS HEALTH MONITORS**

4 BILLION+ HOURS OF VIDEO
are watched on YouTube each month



400 MILLION TWEETS
are sent per day by about 200 million monthly active users



The New York Stock Exchange captures **1 TB OF TRADE INFORMATION** during each trading session



Modern cars have close to **100 SENSORS** that monitor items such as fuel level and tire pressure

Velocity

ANALYSIS OF STREAMING DATA

By 2016, it is projected there will be **18.9 BILLION NETWORK CONNECTIONS** – almost 2.5 connections per person on earth



1 IN 3 BUSINESS LEADERS

don't trust the information they use to make decisions



in one survey were unsure of how much of their data was inaccurate

Veracity

UNCERTAINTY OF DATA

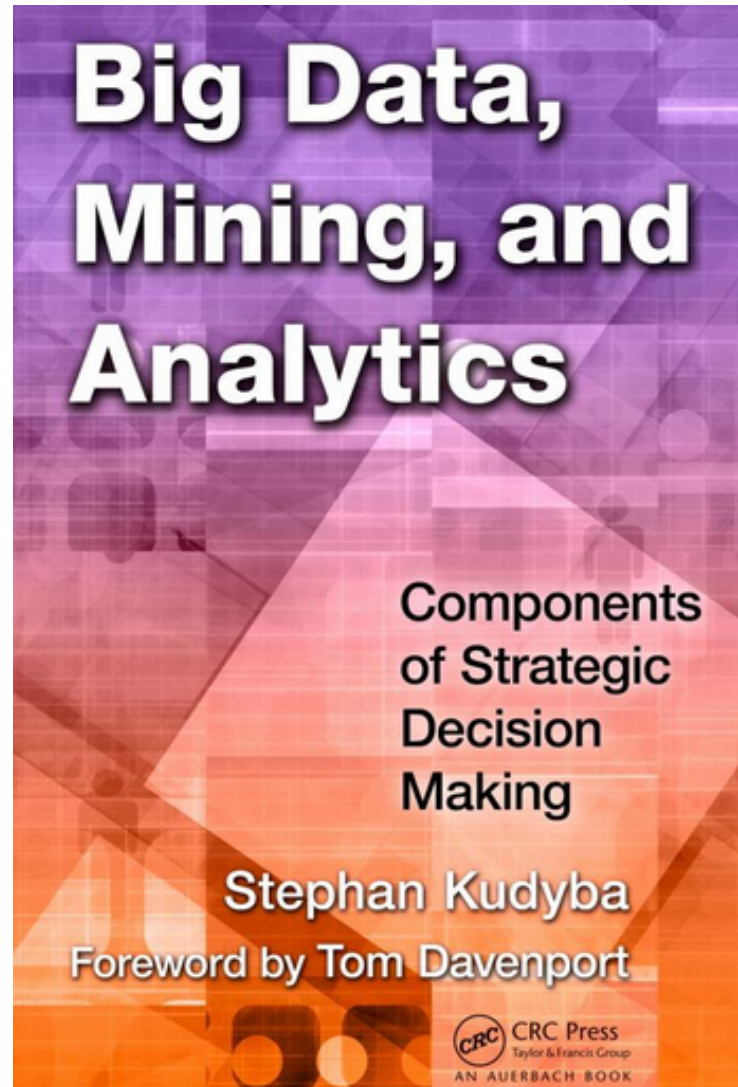
Poor data quality costs the US economy around

\$3.1 TRILLION A YEAR

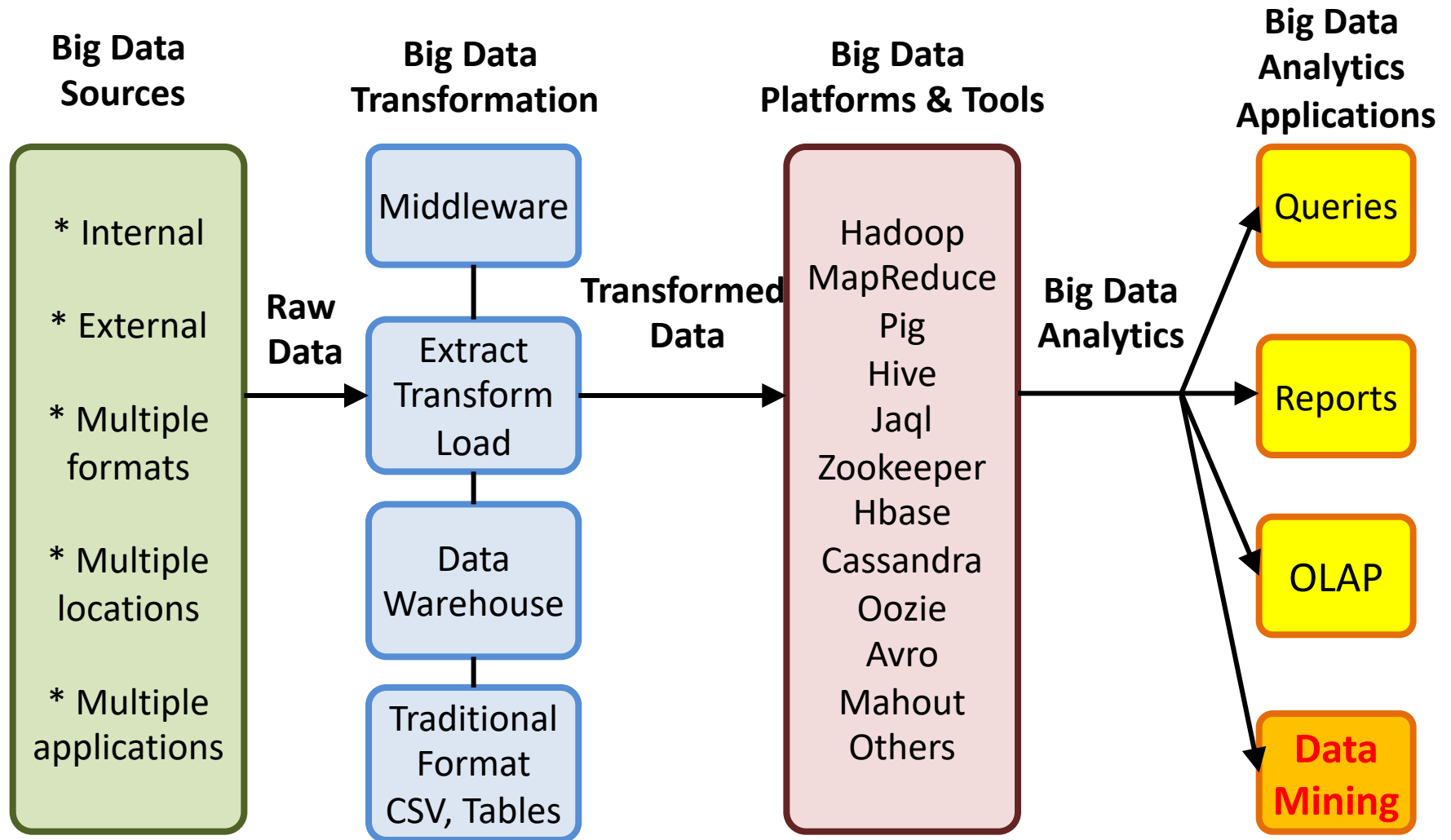


Value

Stephan Kudyba (2014),
Big Data, Mining, and Analytics:
Components of Strategic Decision Making, Auerbach Publications



Architecture of Big Data Analytics

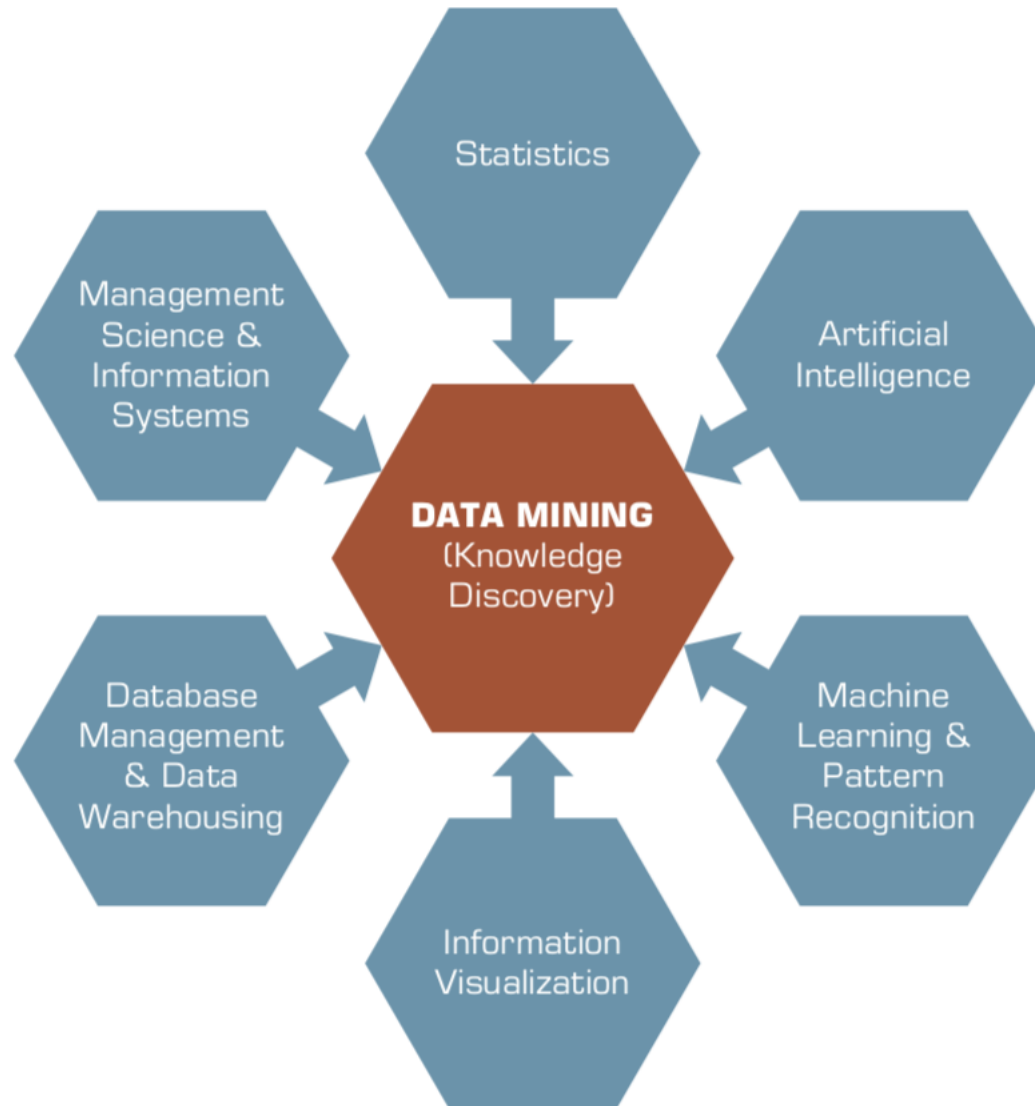


Architecture of Big Data Analytics

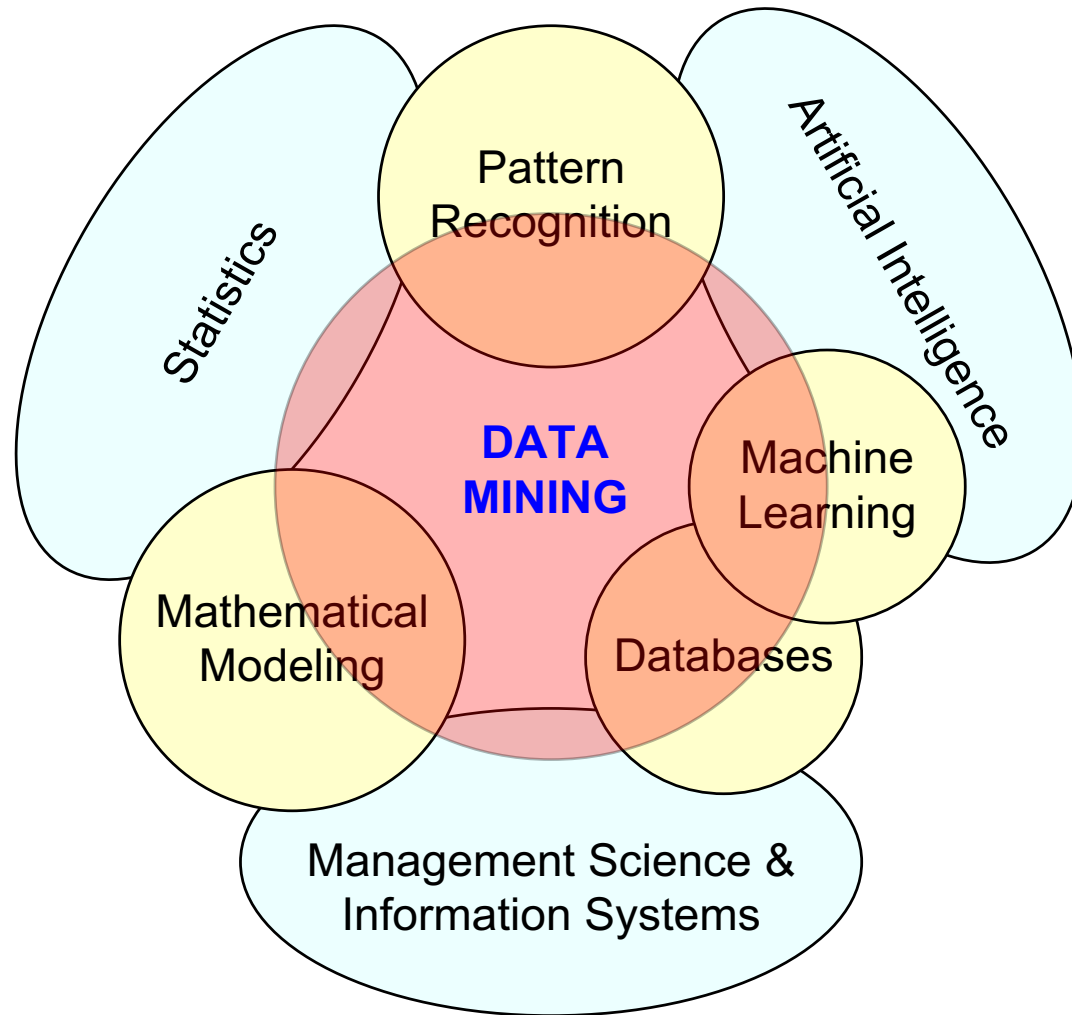


Data Mining

Is a Blend of Multiple Disciplines



Data Mining at the Intersection of Many Disciplines



Data Mining Tasks & Methods

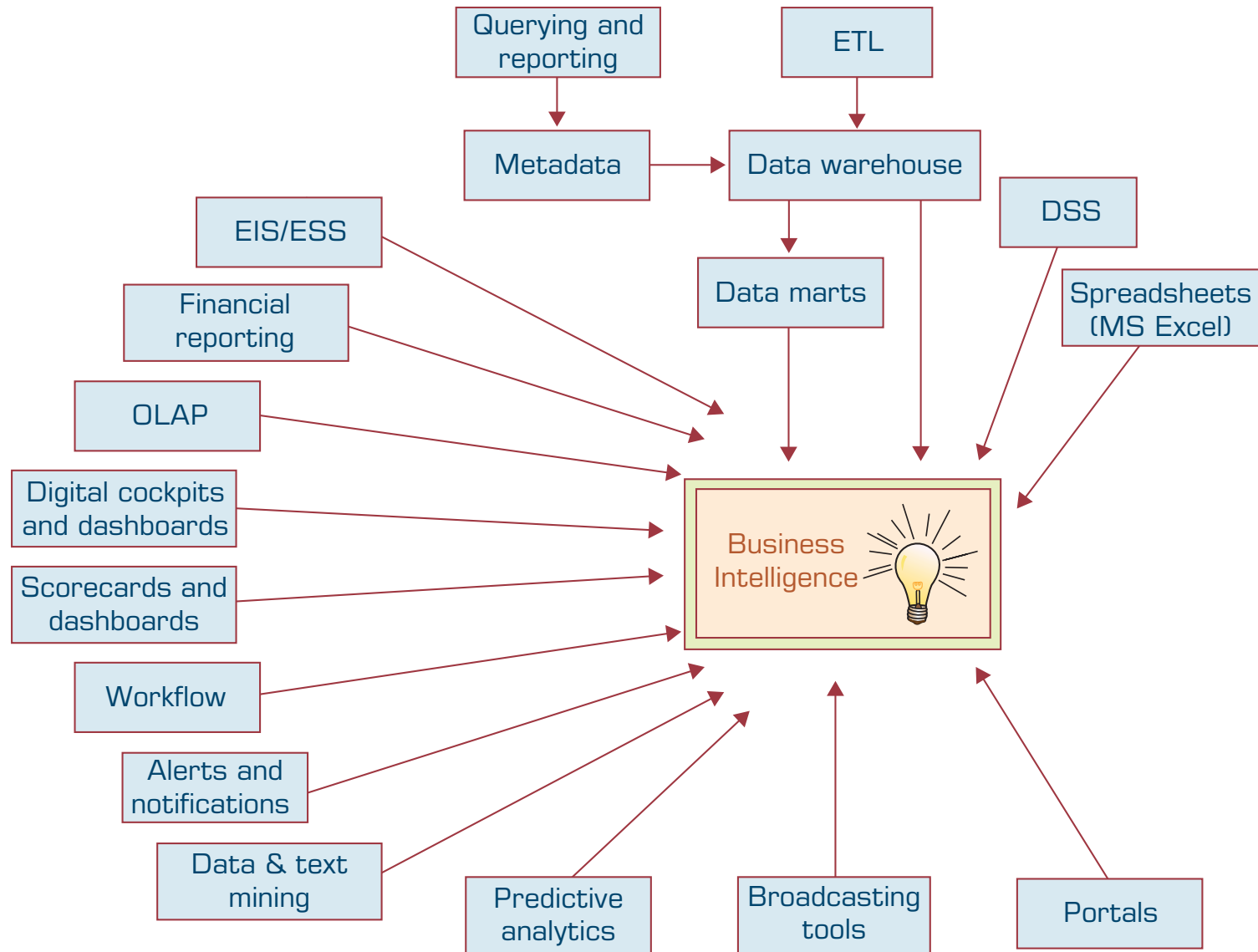
Prediction

Data Mining Tasks & Methods	Data Mining Algorithms	Learning Type
Prediction		
Classification	Decision Trees, Neural Networks, Support Vector Machines, kNN, Naïve Bayes, GA	Supervised
Regression	Linear/Nonlinear Regression, ANN, Regression Trees, SVM, kNN, GA	Supervised
Time series	Autoregressive Methods, Averaging Methods, Exponential Smoothing, ARIMA	Supervised
Association		
Market-basket	Apriori, OneR, ZeroR, Eclat, GA	Unsupervised
Link analysis	Expectation Maximization, Apriori Algorithm, Graph-Based Matching	Unsupervised
Sequence analysis	Apriori Algorithm, FP-Growth, Graph-Based Matching	Unsupervised
Segmentation		
Clustering	k-means, Expectation Maximization (EM)	Unsupervised
Outlier analysis	k-means, Expectation Maximization (EM)	Unsupervised

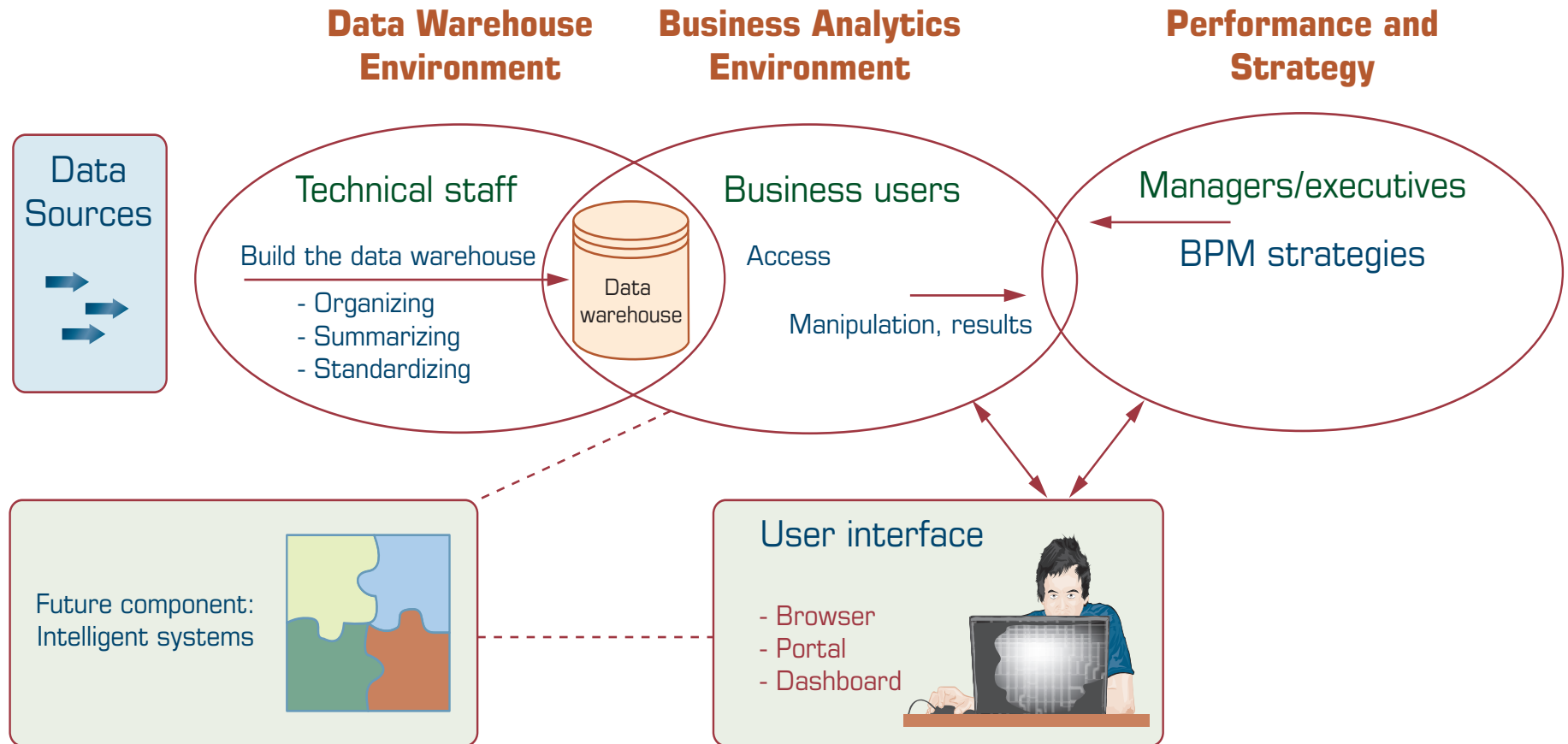
Association

Segmentation

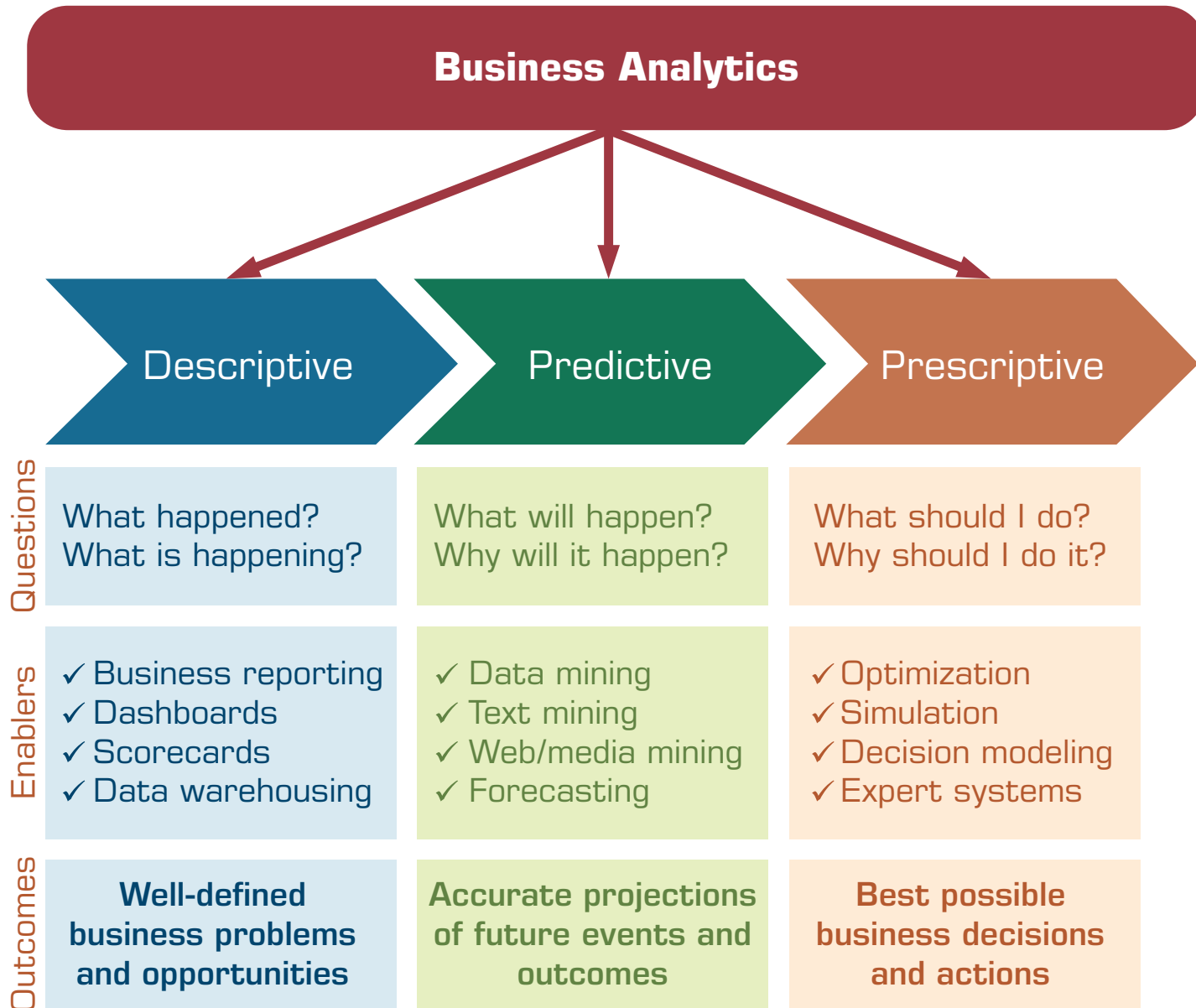
Evolution of Business Intelligence (BI)



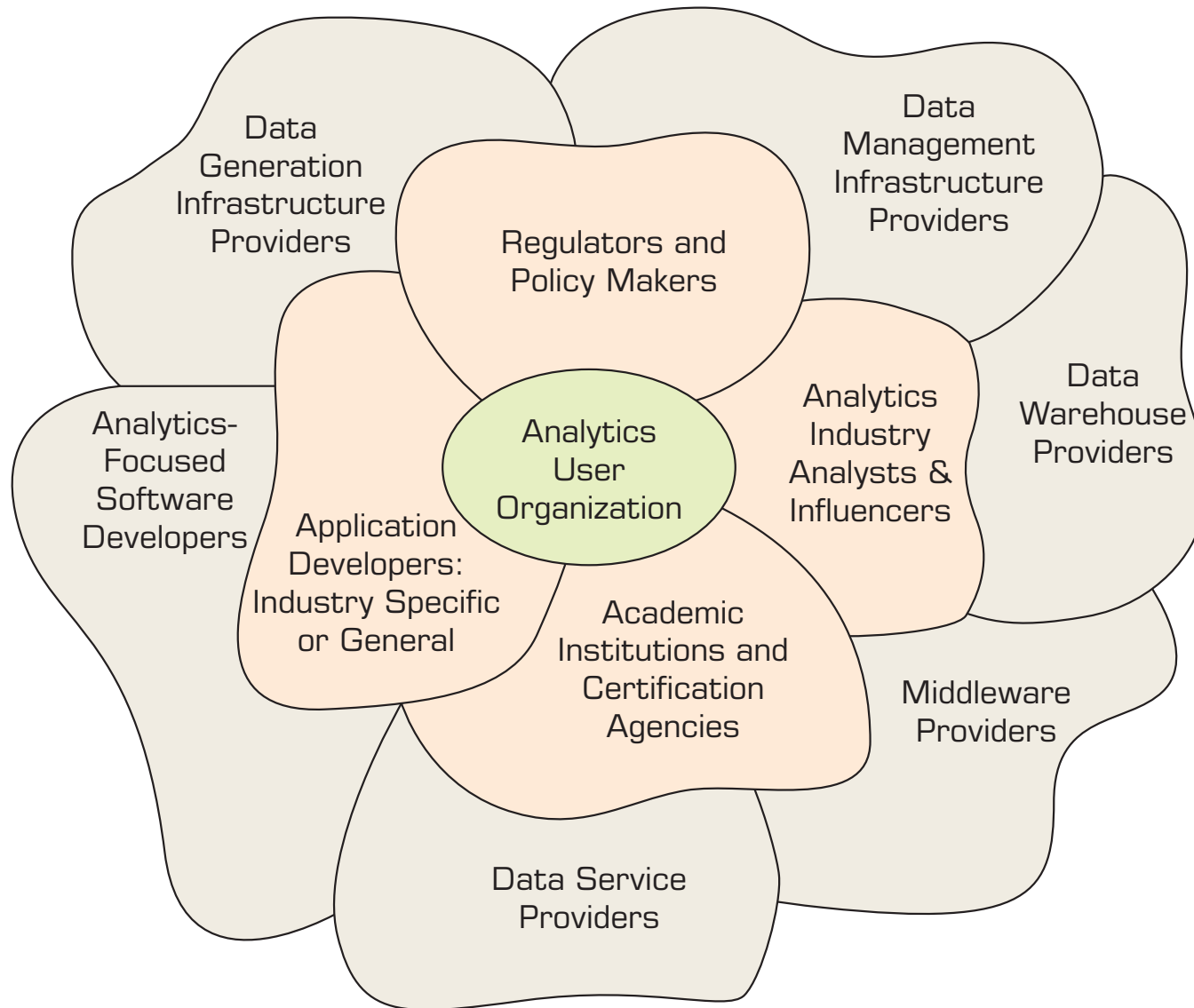
A High-Level Architecture of BI



Three Types of Analytics



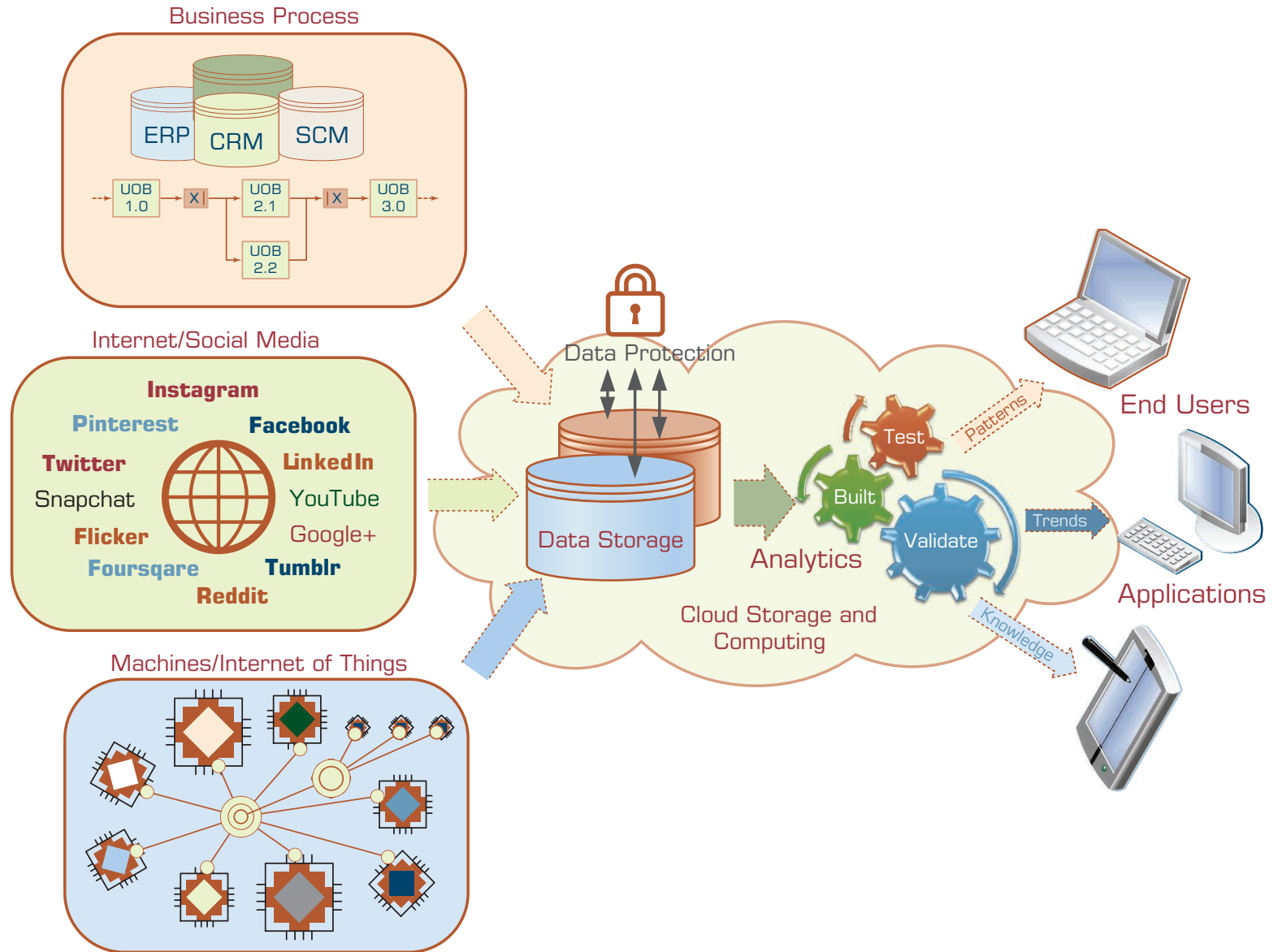
Analytics Ecosystem



A word cloud visualization featuring various job titles and roles. The most prominent words are "Marketing", "Developer", "Consultant", "Analytics", "Business", "Senior", "Risk", "Scientist", "Information", "Intern", "Engineer", "Associate", "Manager", "Specialist", "Programmer", "Security", "Database", "Administrator", "Health", "Revenue", "Science", "Quantitative", "Mining", "Technology", "Systems", "Professional", "Applications", "Modeling", "Technologies", "Care", "Director", "Global", "SAP", "Architect", "Hadoop", "Market", "Application", "Statistical", "System", "Programming/Support", "Drug", "Assurance", "Center", "Big Data", "Innovation", "Technical", "Fraud", "Planning", "Demand", "Strategic", "Applied", "Mining", "Business", "Analyst", "Suite", "Pricing", "Lead", "Statistics", "Evaluation", "Basis", "Computer", "Effectiveness", "Healthcare", "Communications", "Assistant", "Analysis", "Mobile", "Services".

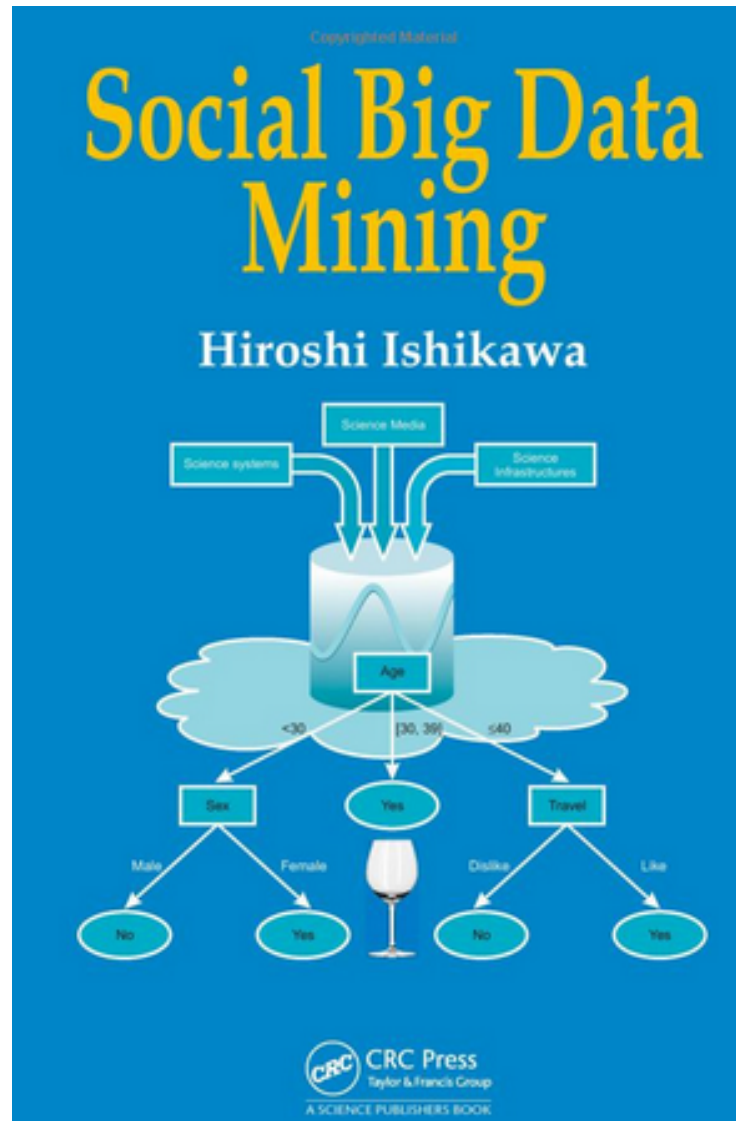
54

A Data to Knowledge Continuum



Social Big Data Mining

(Hiroshi Ishikawa, 2015)

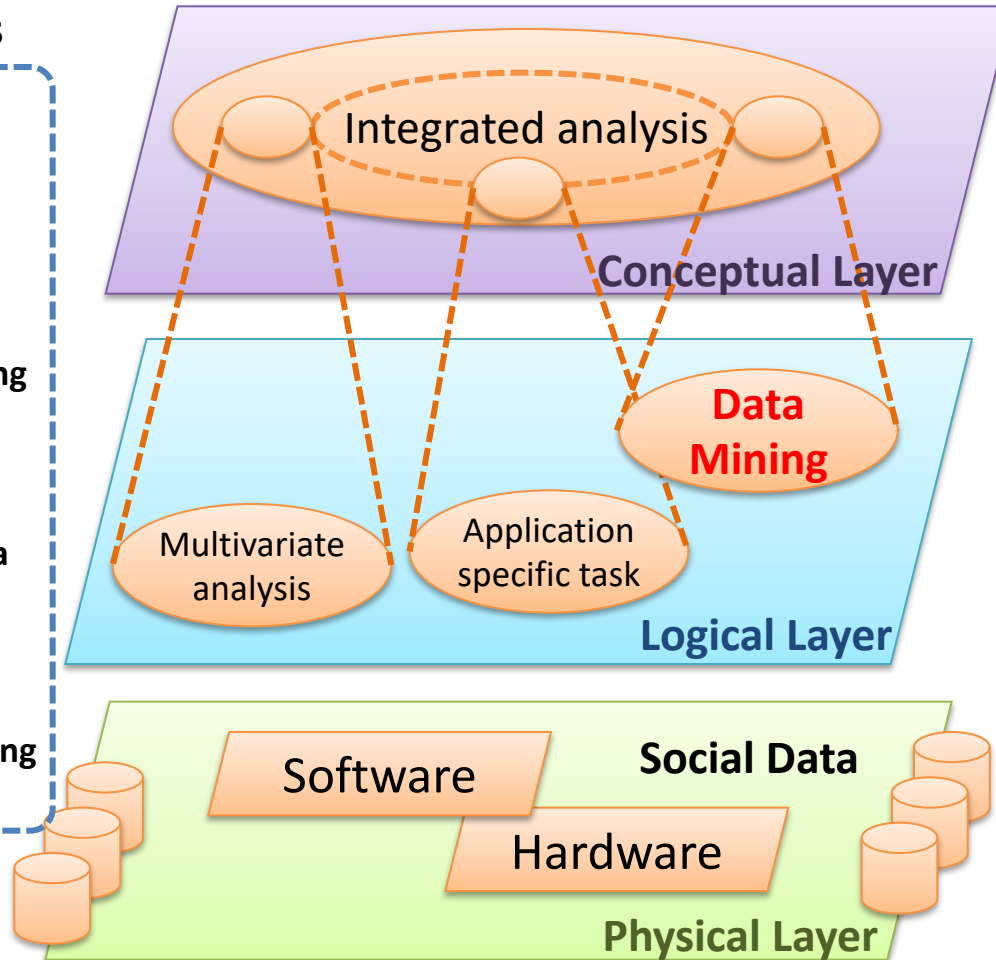


Architecture for Social Big Data Mining

(Hiroshi Ishikawa, 2015)

Enabling Technologies

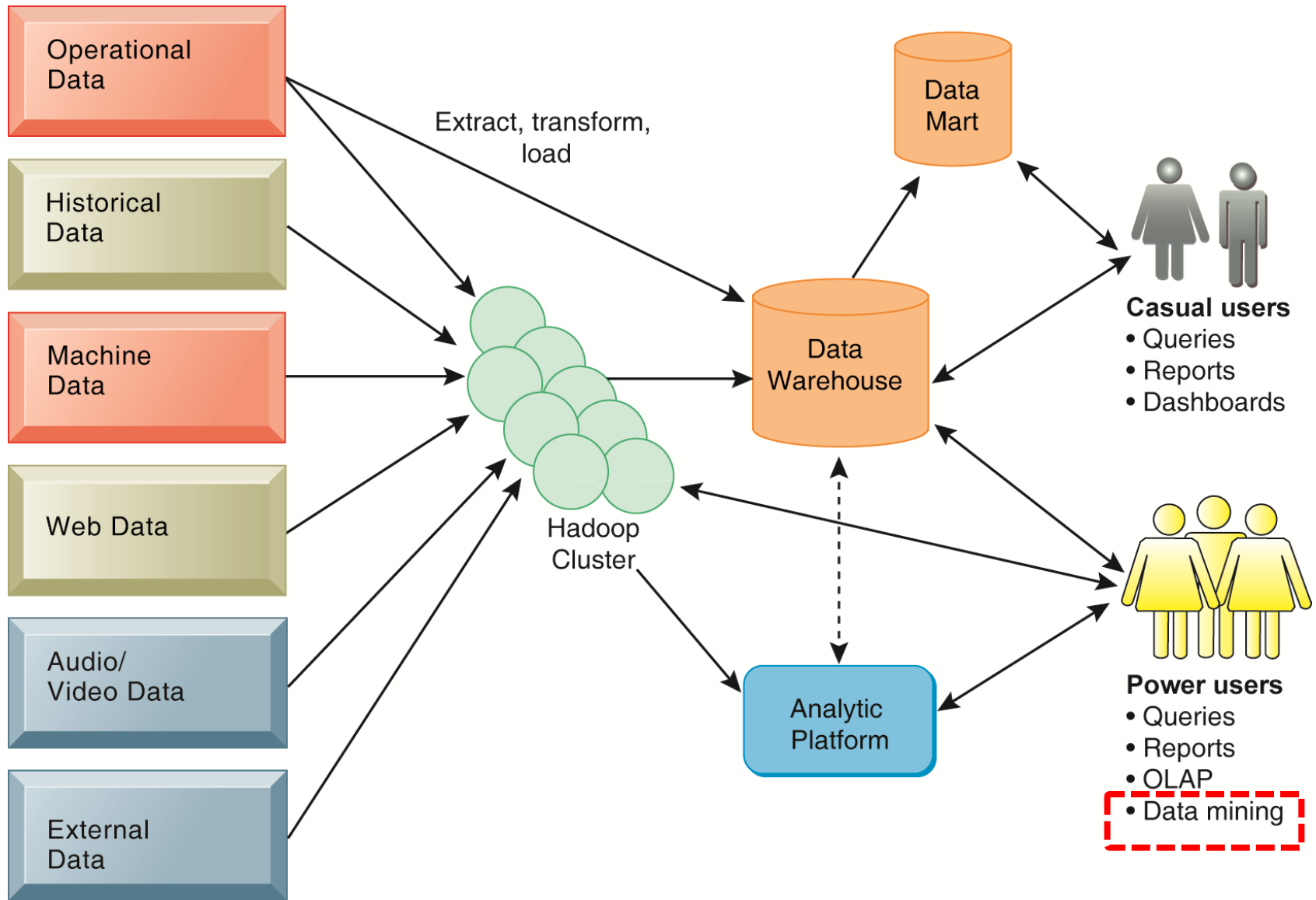
- Integrated analysis model
- Natural Language Processing
- Information Extraction
- Anomaly Detection
- Discovery of relationships among heterogeneous data
- Large-scale visualization
- Parallel distrusted processing



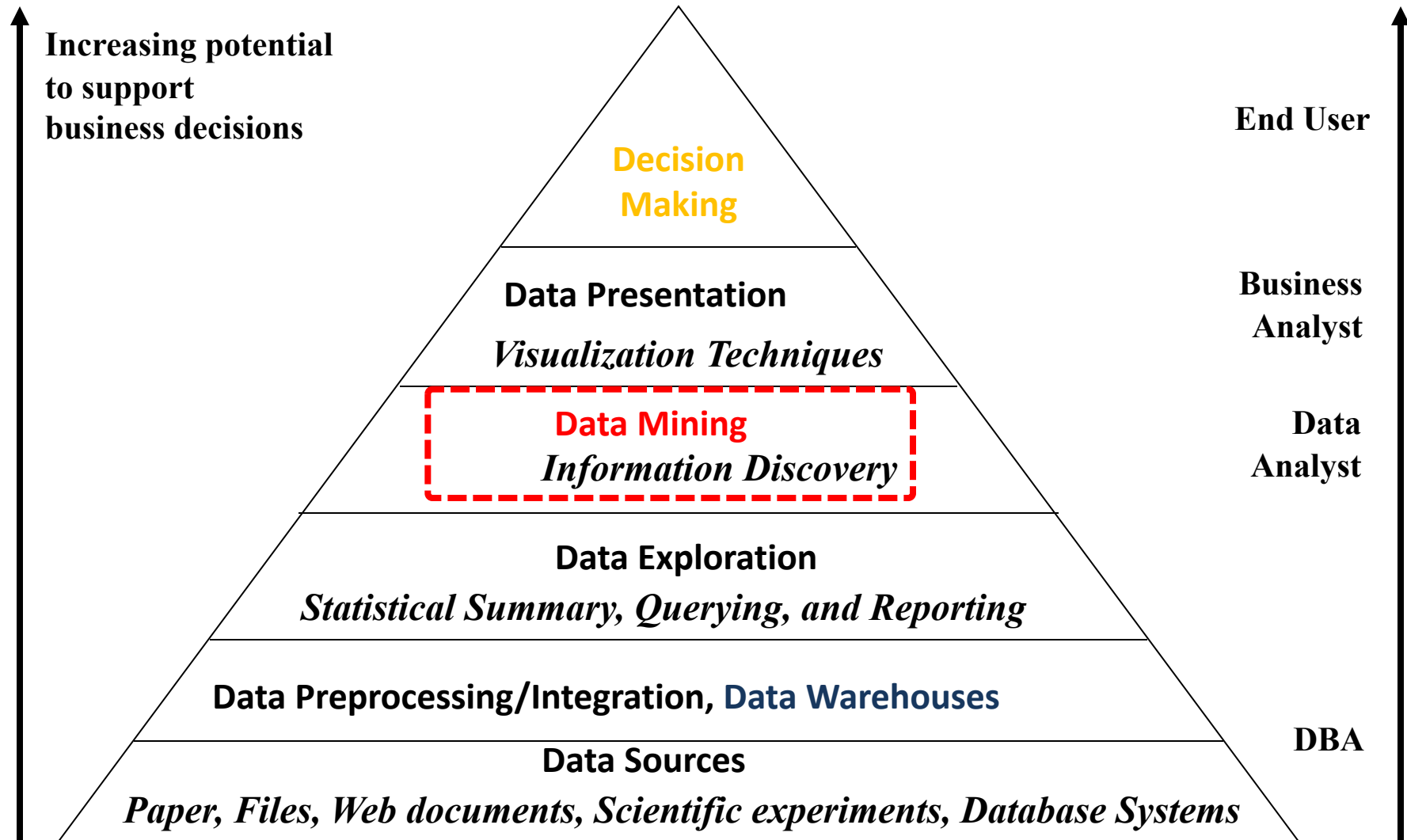
Analysts

- Model Construction
- Explanation by Model
- Construction and confirmation of individual hypothesis
- Description and execution of application-specific task

Business Intelligence (BI) Infrastructure



Business Intelligence and Data Mining





Data Mining:

Core **Analytics** Process

The **KDD** Process for
Extracting Useful **Knowledge**
from Volumes of **Data**

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996).

The **KDD Process** for
Extracting Useful **Knowledge**
from Volumes of **Data**.

Communications of the ACM, 39(11), 27-34.

Knowledge Discovery in Databases creates the context for developing the tools needed to control the flood of data facing organizations that depend on ever-growing databases of business, manufacturing, scientific, and personal information.

The KDD Process for Extracting Useful Knowledge from Volumes of Data

AS WE MARCH INTO THE AGE of digital information, the problem of data overload looms ominously ahead. Our ability to analyze and understand massive datasets lags far behind our ability to gather and store the data. A new generation of computational techniques and tools is required to support the extraction of useful knowledge from the rapidly growing volumes of data. These techniques and tools are the subject of the emerging field of knowledge discovery in databases (KDD) and data mining.

Large databases of digital information are ubiquitous. Data from the neighborhood store's checkout register, your bank's credit card authorization device, records in your doctor's office, patterns in your telephone calls,

Usama Fayyad,

Gregory Piatetsky-Shapiro,

and Padhraic Smyth

and many more applications generate streams of digital records archived in huge databases, sometimes in so-called data warehouses.

Current hardware and database technology allow efficient and inexpensive reliable data storage and access. However, whether the context is business, medicine, science, or government, the datasets themselves (in raw form) are of little direct value. What is of value is the knowledge that can be inferred from the data and put to use. For example, the marketing database of a consumer

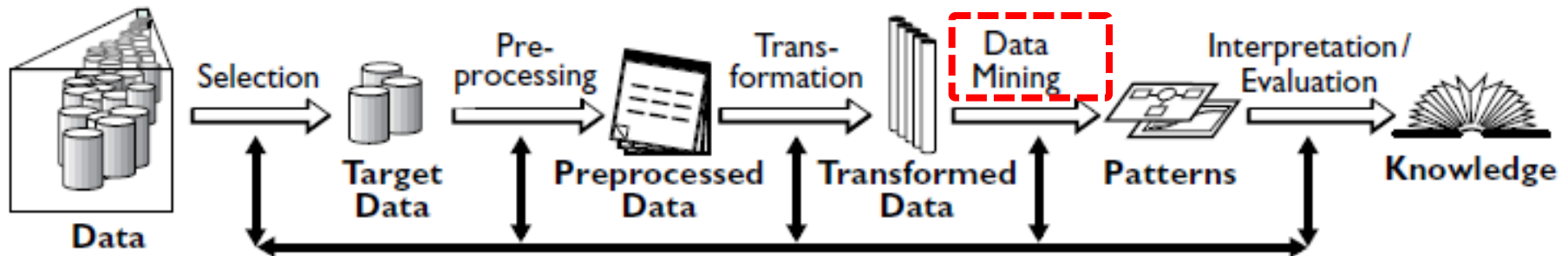


REPLY HERE

Data Mining

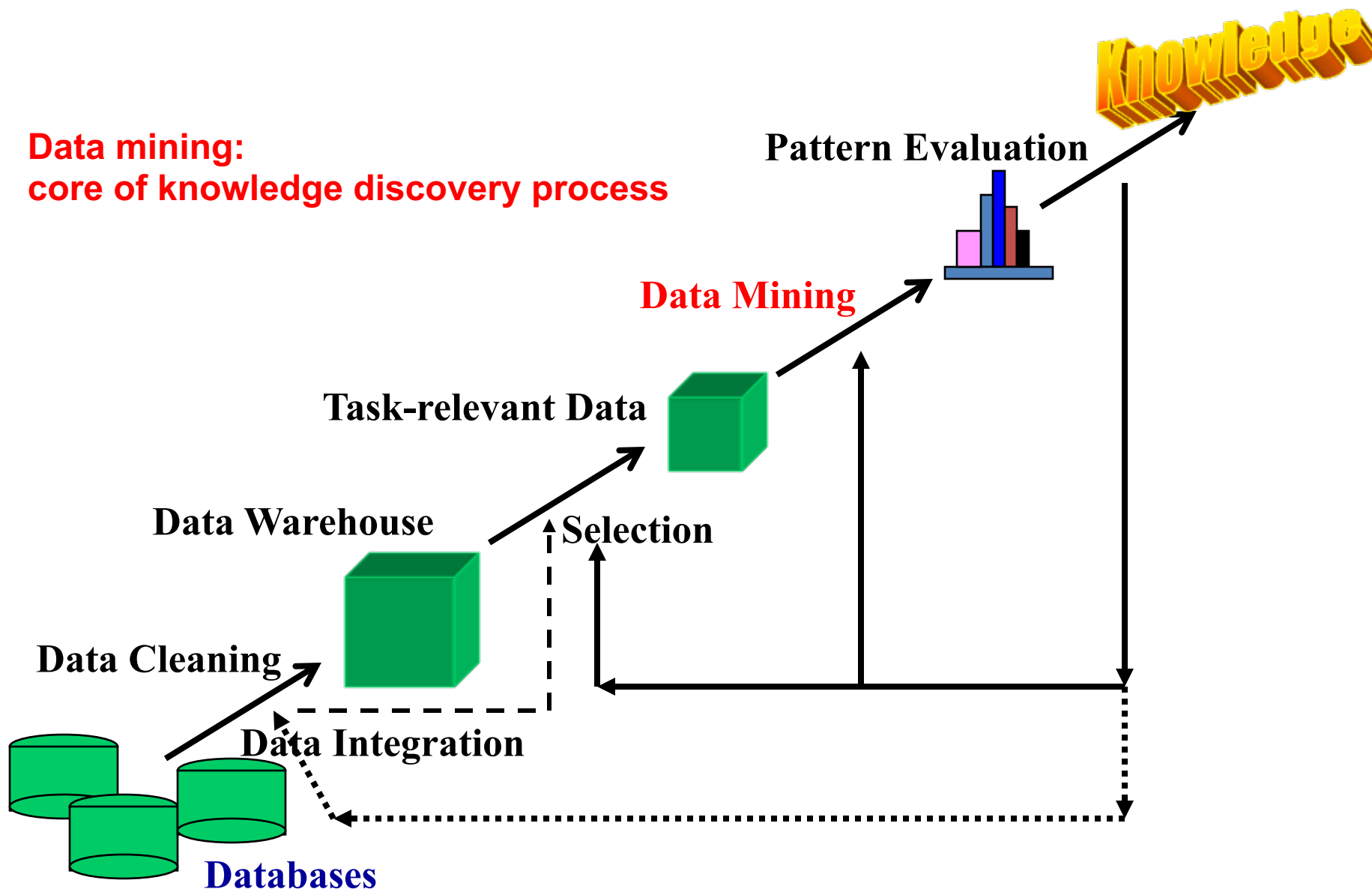
Knowledge Discovery in Databases (KDD) Process

(Fayyad et al., 1996)



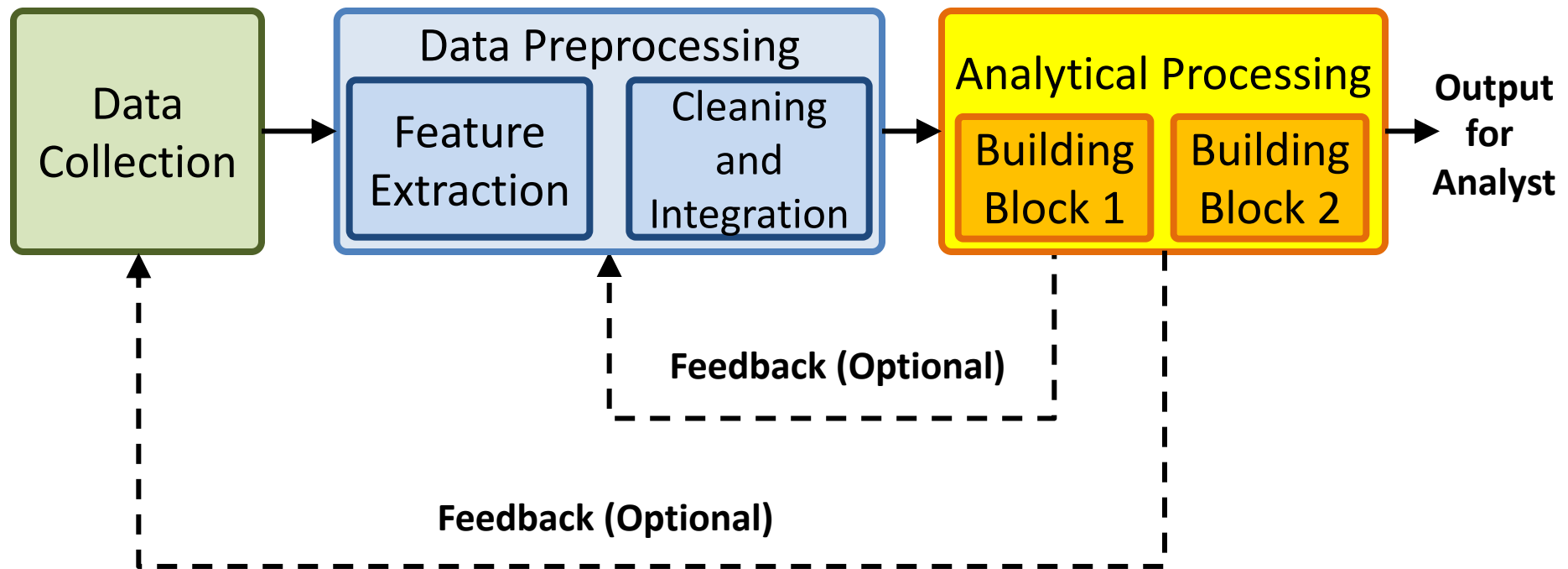
Knowledge Discovery (KDD) Process

Data mining:
core of knowledge discovery process

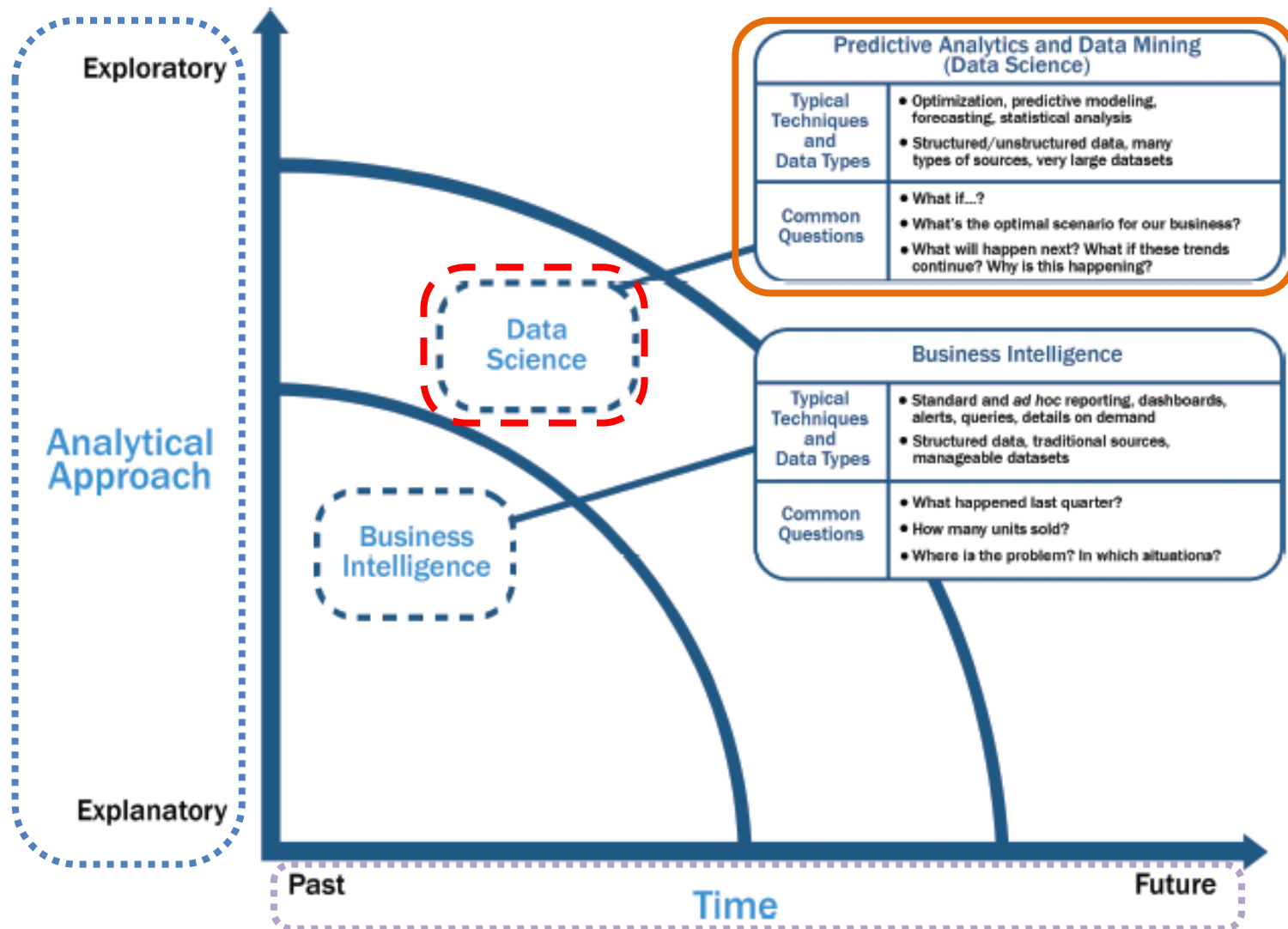


Data Mining Processing Pipeline

(Charu Aggarwal, 2015)



Data Science and Business Intelligence



Data Science and Business Intelligence



Predictive Analytics and Data Mining (Data Science)

Past

Time

Future

Predictive Analytics and Data Mining (Data Science)

Structured/unstructured data, many types of sources,
very large datasets

Optimization, predictive modeling, forecasting statistical analysis

What if...?
What's the optimal scenario for our business?
What will happen next?
What if these trends continue?
Why is this happening?

Cloud Computing



Cloud Computing

AWS

Amazon Web Services



Analytics



Application Integration



AR & VR



AWS Cost Management



Blockchain



Business Applications



Compute



Customer Engagement



Database



Developer Tools



End User Computing



Game Tech



Internet of Things



Machine Learning



Management & Governance



Media Services



Migration & Transfer



Mobile



Networking & Content
Delivery



Quantum Technologies



Robotics



Satellite



Security, Identity &
Compliance

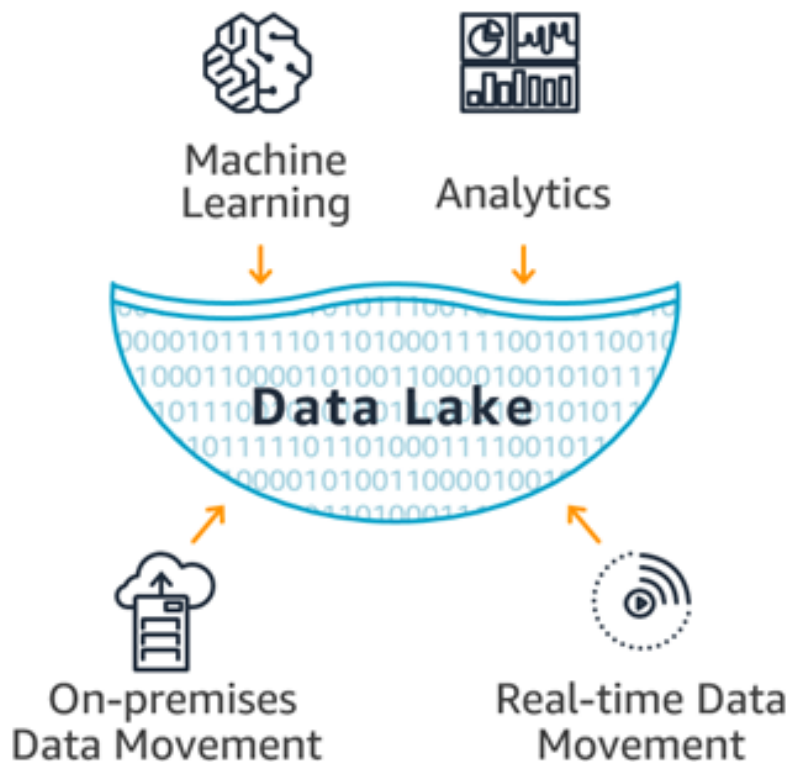


Storage

Data Lakes and Analytics on AWS

Data Movement

Import your data from on-premises, and in real-time.



Data Lake

Store any type of data securely, from gigabytes to exabytes.

Analytics

Analyze your data with a broad selection of analytic tools and engines.

Machine Learning

Forecast future outcomes, and prescribe actions.



AWS Products

Analytics

- **Amazon Athena**
 - Query data in S3 using SQL
- **Amazon CloudSearch**
 - Managed search service
- **Amazon EMR**
 - Hosted Hadoop framework
- **Amazon Elasticsearch Service**
 - Run and scale Elasticsearch clusters
- **Amazon Kinesis**
 - Analyze real-time video and data streams
- **Amazon Redshift**
 - Fast, simple, cost-effective data warehousing
- **Amazon QuickSight**
 - Fast business analytics service
- **AWS Data Pipeline**
 - Orchestration service for periodic, data-driven workflows
- **AWS Glue**
 - Prepare and load data



Machine Learning on AWS

Machine learning in the hands of every developer and data scientist



Build

Connect to other AWS services and transform data in SageMaker notebooks



Train

Use SageMaker's algorithms and frameworks, or bring your own, for distributed training



Tune

SageMaker automatically tunes your model by adjusting multiple combinations of algorithm parameters



Deploy

Once training is completed, models can be deployed to SageMaker endpoints, for real-time predictions



AWS Certified **Cloud Practitioner**

AWS Certified **Solutions Architect**

AWS Certified **Big Data Specialty**

AWS Certified **Machine Learning Specialty**

Available AWS Certifications

aws certified
Updated May 2019

Professional

Two years of comprehensive experience designing, operating, and troubleshooting solutions using the AWS Cloud



Specialty

Technical AWS Cloud experience in the Specialty domain as specified in the exam guide

Associate

One year of experience solving problems and implementing solutions using the AWS Cloud



BD

Foundational

Six months of fundamental AWS Cloud and industry knowledge



ML

SAA

CLF



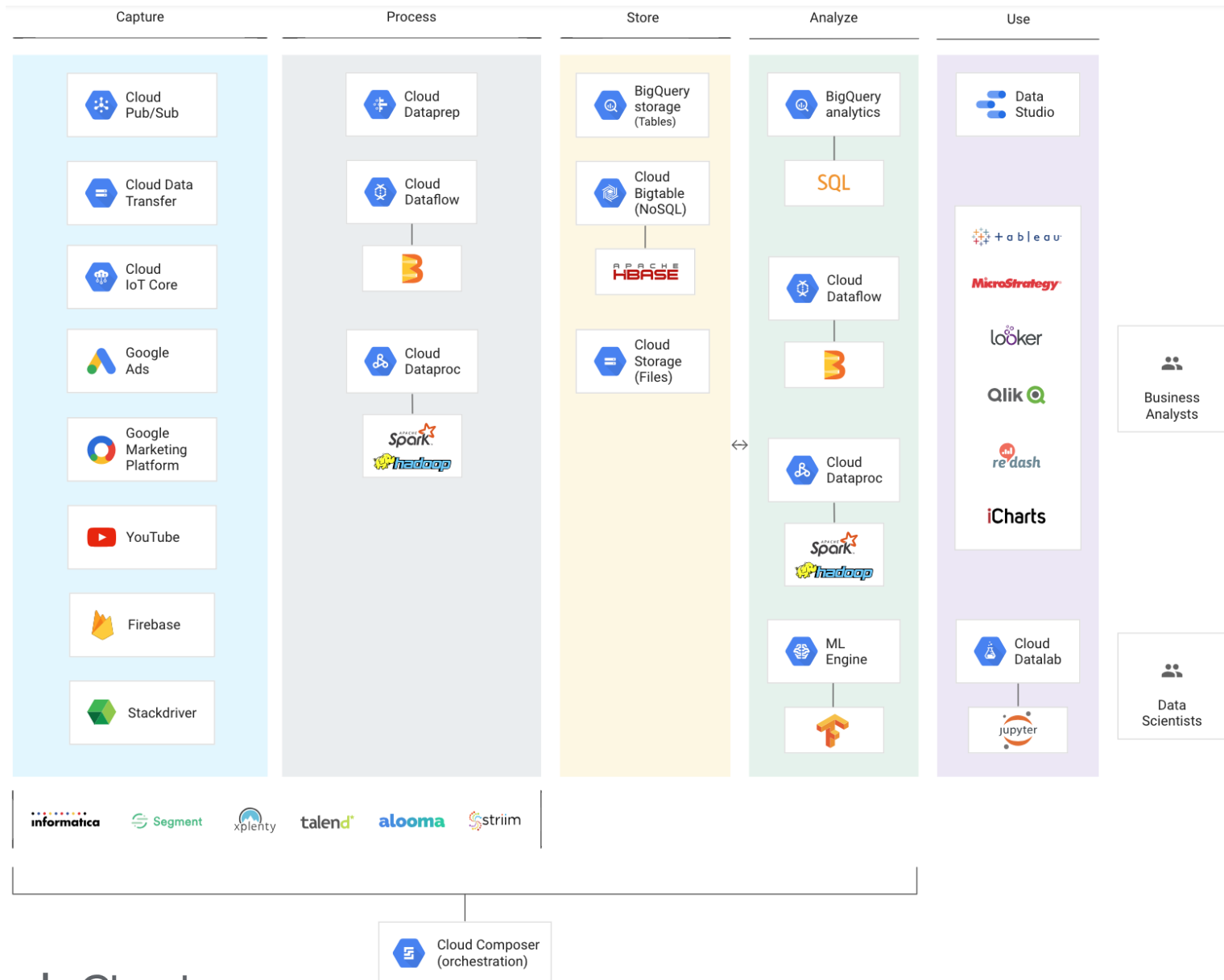
國立臺北大學
National Taipei University





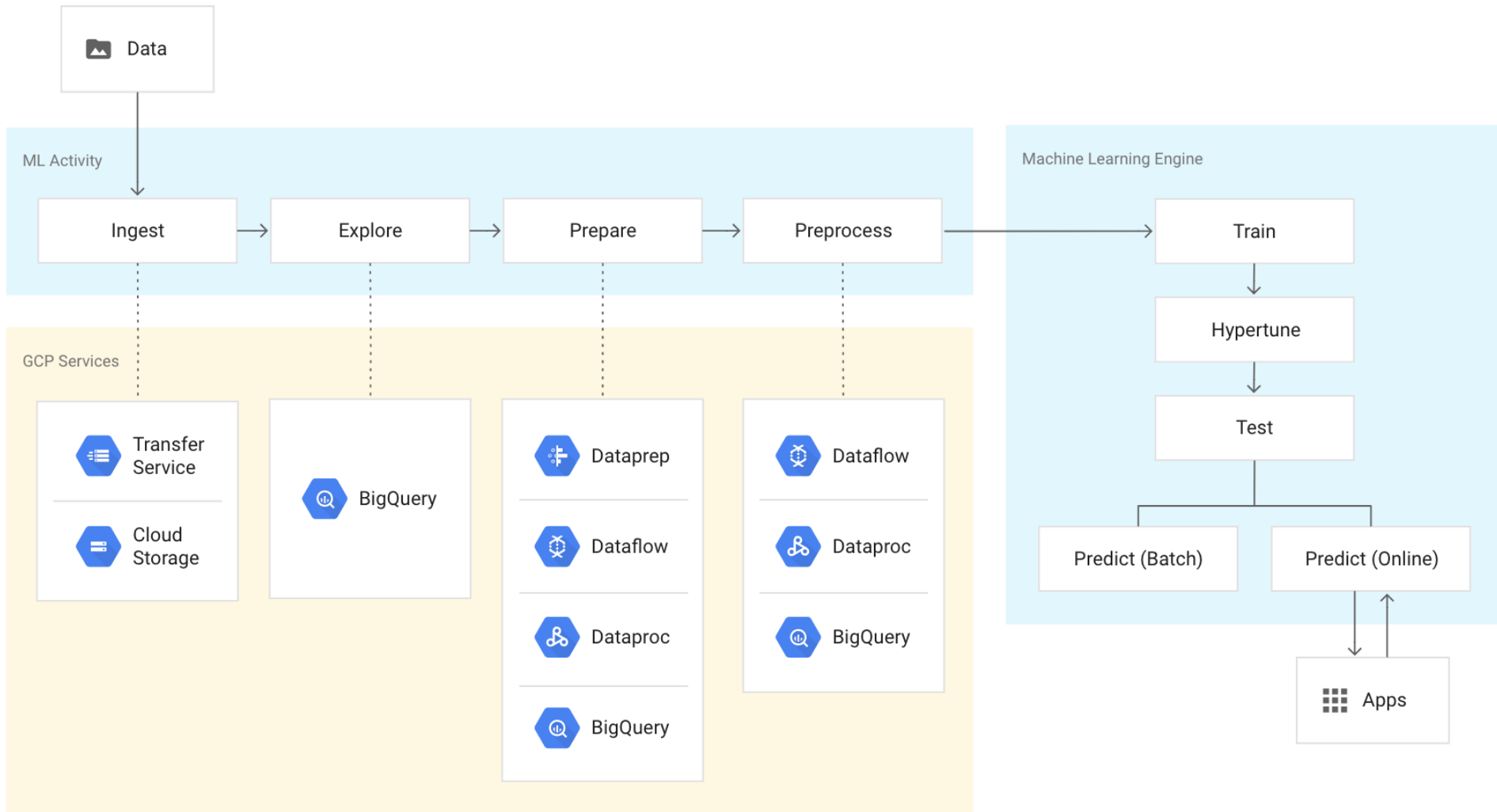
Google Cloud

Google Cloud Big Data Analytics

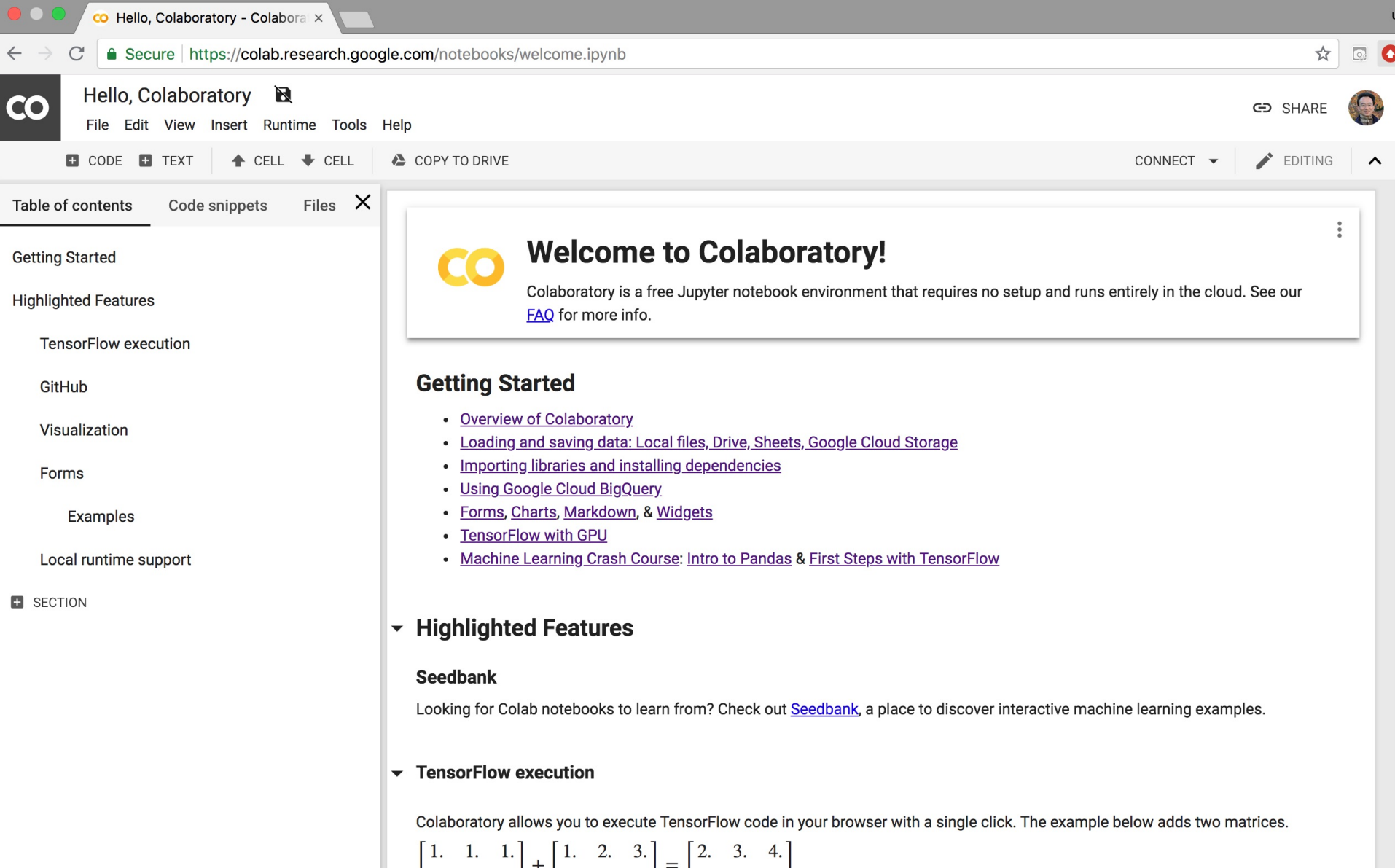


Google Cloud

Machine learning and Cloud AI



Google Colab



The screenshot shows the Google Colaboratory interface in a web browser. The browser's address bar displays the URL <https://colab.research.google.com/notebooks/welcome.ipynb>. The page header includes the Colab logo, the text "Hello, Colaboratory", and a menu with options: File, Edit, View, Insert, Runtime, Tools, and Help. On the right side of the header, there is a "SHARE" button and a user profile icon. Below the header, a toolbar contains icons for "CODE", "TEXT", "CELL", and "COPY TO DRIVE". A sidebar on the left lists navigation options: "Table of contents", "Code snippets", and "Files". The main content area features a "Welcome to Colaboratory!" message with the Colab logo and a brief description of the environment. Below this, a "Getting Started" section lists several links for further exploration. A "Highlighted Features" section is partially visible, showing a "Seedbank" link and a "TensorFlow execution" section.

Table of contents

- Getting Started
- Highlighted Features
 - TensorFlow execution
- GitHub
- Visualization
- Forms
- Examples
- Local runtime support

SECTION

Welcome to Colaboratory!

Colaboratory is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud. See our [FAQ](#) for more info.

Getting Started

- [Overview of Colaboratory](#)
- [Loading and saving data: Local files, Drive, Sheets, Google Cloud Storage](#)
- [Importing libraries and installing dependencies](#)
- [Using Google Cloud BigQuery](#)
- [Forms, Charts, Markdown, & Widgets](#)
- [TensorFlow with GPU](#)
- [Machine Learning Crash Course: Intro to Pandas & First Steps with TensorFlow](#)

Highlighted Features

Seedbank

Looking for Colab notebooks to learn from? Check out [Seedbank](#), a place to discover interactive machine learning examples.

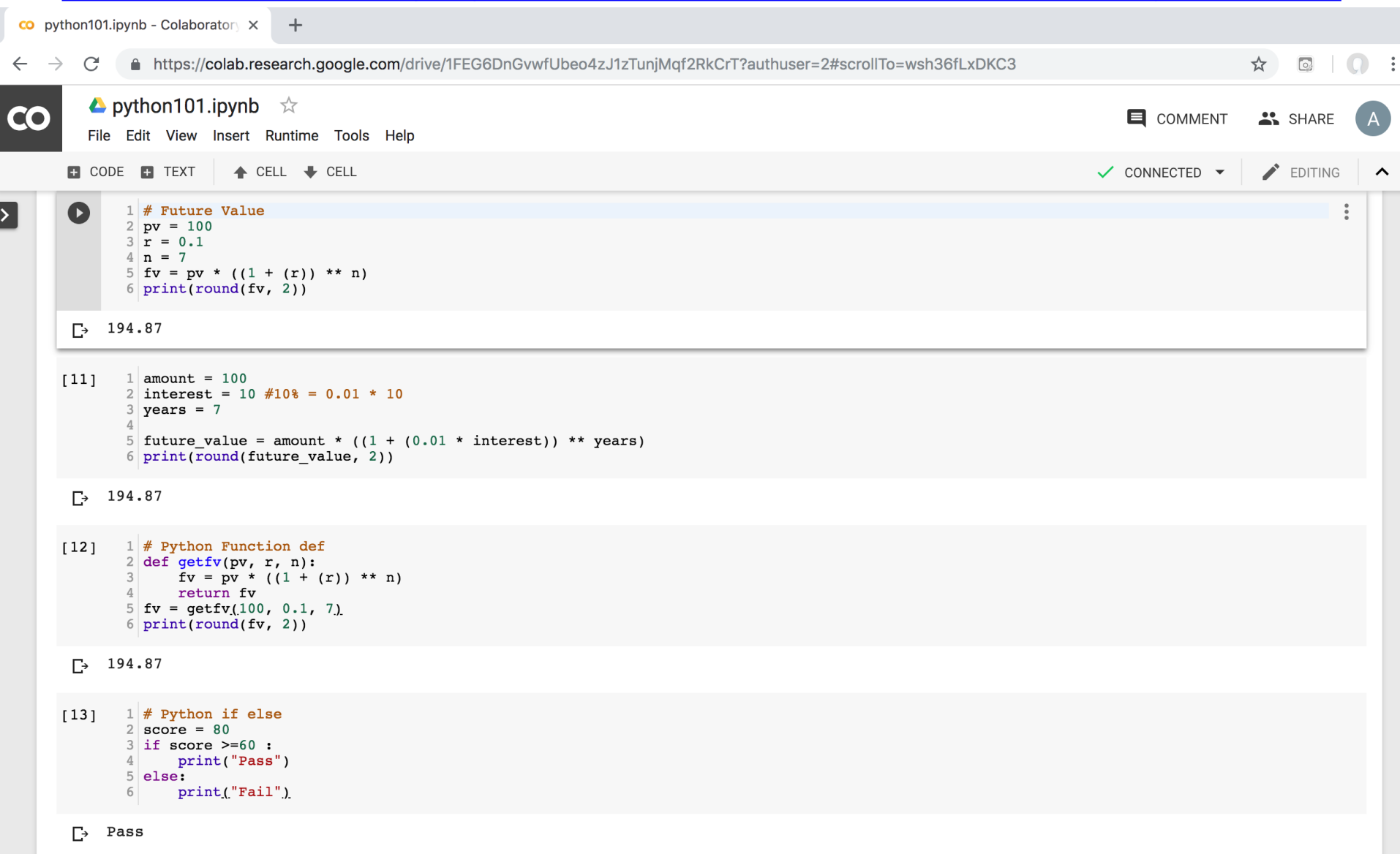
TensorFlow execution

Colaboratory allows you to execute TensorFlow code in your browser with a single click. The example below adds two matrices.

$$\begin{bmatrix} 1. & 1. & 1. \end{bmatrix} + \begin{bmatrix} 1. & 2. & 3. \end{bmatrix} = \begin{bmatrix} 2. & 3. & 4. \end{bmatrix}$$

Python in Google Colab (Python101)

<https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT>



python101.ipynb - Colaboratory

File Edit View Insert Runtime Tools Help

CODE TEXT CELL CELL

CONNECTED EDITING

```
1 # Future Value
2 pv = 100
3 r = 0.1
4 n = 7
5 fv = pv * ((1 + (r)) ** n)
6 print(round(fv, 2))
```

194.87

```
[11] 1 amount = 100
2 interest = 10 #10% = 0.01 * 10
3 years = 7
4
5 future_value = amount * ((1 + (0.01 * interest)) ** years)
6 print(round(future_value, 2))
```

194.87

```
[12] 1 # Python Function def
2 def getfv(pv, r, n):
3     fv = pv * ((1 + (r)) ** n)
4     return fv
5 fv = getfv(100, 0.1, 7)
6 print(round(fv, 2))
```

194.87

```
[13] 1 # Python if else
2 score = 80
3 if score >=60 :
4     print("Pass")
5 else:
6     print("Fail").
```

Pass

<https://tinyurl.com/aintpuppython101>

Summary

- **AI**
- **Big Data**
- **Cloud Computing**

References

- Stuart Russell and Peter Norvig (2020), Artificial Intelligence: A Modern Approach, 4th Edition, Pearson.
- Aurélien Géron (2019), Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition, O'Reilly Media.
- Robert Layton (2017), Learning Data Mining with Python - Second Edition, Packt Publishing.
- Jiawei Han, Micheline Kamber and Jian Pei (2011), Data Mining: Concepts and Techniques, Third Edition, Morgan Kaufmann.
- Sanjeev Verma, Rohit Sharma, Subhamay Deb, and Debojit Maitra (2021), "Artificial intelligence in marketing: Systematic review and future research direction." International Journal of Information Management Data Insights, 100002.
- Liye Ma and Baohong Sun (2020), "Machine learning and AI in marketing – Connecting computing power to human insights." International Journal of Research in Marketing, 37, no. 3, 481-504.
- Ramesh Sharda, Dursun Delen, and Efraim Turban (2017), Business Intelligence, Analytics, and Data Science: A Managerial Perspective, 4th Edition, Pearson.
- Xiao-lin Zheng, Meng-ying Zhu, Qi-bing Li, Chao-chao Chen, and Yan-chao Tan (2019), "Finbrain: When finance meets AI 2.0." Frontiers of Information Technology & Electronic Engineering 20, no. 7, pp. 914-924.
- Ting-Peng Liang and Yu-Hsi Liu (2018), "Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study", Expert Systems with Applications, 111, no. 30, pp. 2-10.
- Jared Dean (2014), Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners, Wiley.
- Mehmet Kaya, Jalal Kawash, Suheil Khoury, and Min-Yuh Day (2018), Social Network Based Big Data Analysis and Applications, Lecture Notes in Social Networks, Springer International Publishing.
- Varun Grover, Roger HL Chiang, Ting-Peng Liang, and Dongsong Zhang (2018), "Creating Strategic Business Value from Big Data Analytics: A Research Framework", Journal of Management Information Systems, 35, no. 2, pp. 388-423.
- Stephan Kudyba (2014), Big Data, Mining, and Analytics: Components of Strategic Decision Making, Auerbach Publications.