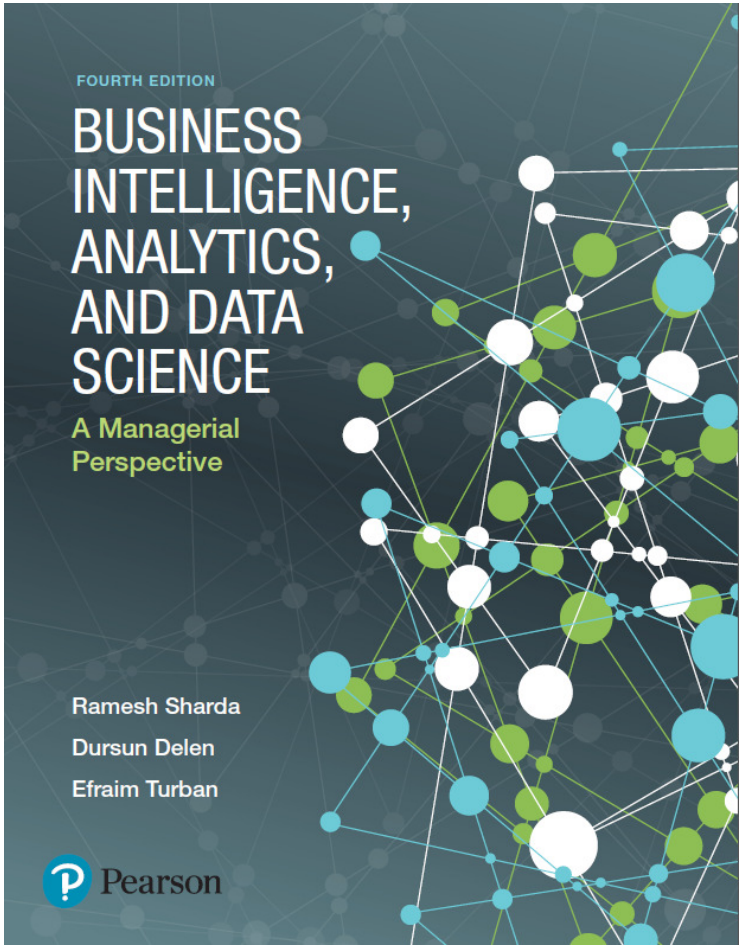
Business Intelligence, Analytics, and Data Science: A Managerial Perspective

Fourth Edition



Chapter 8 – Part A

Future Trends, Privacy and Managerial Considerations in Analytics

Pearson

Learning Objectives (1 of 2)

8.1 Explore some of the emerging technologies that may impact analytics, business intelligence (BI), and decision support

8.2 Describe the emerging Internet of Things (IoT) phenomenon, potential applications, and the IoT ecosystem

8.3 Describe the current and future use of cloud computing in business analytics

8.4 Describe how geospatial and location-based analytics are assisting organizations

Learning Objectives (2 of 2)

8.5 Describe the organizational impacts of analytics applications

8.6 List and describe the major ethical and legal issues of analytics implementation

8.7 Identify key characteristics of a successful data science professional

OPENING VIGNETTE

Analysis of Sensor Data Helps Siemens Avoid Train Failures

Discussion Questions

1. In industrial equipment such as trains, what parameters might one measure on a regular basis to estimate the equipment's current performance and future repair needs?
2. How would weather data be useful in analyzing a train's equipment status?
3. Estimate how much data you might collect in one month using, say, 1,000 sensors on a train. Each sensor might yield 1 KB data per second.
4. How would you propose to store such data sets?

Internet of Things (IoT)

- IoT is an area with explosive growth
- Connecting physical world to the Internet
- Social Network versus IoT
  - human-to-human vs. machine-to-machine
- Enablers: sensors and sensing devices
- Example
  - Self driving cars
  - Fitness trackers
  - Smartbin – trash detectors detecting fill levels
  - Smart refrigerators, and other appliances

Internet of Things (IoT)

- By 2020, besides computing and communication devices (tablets, phones, and PCs), another 38B things will be connected to the Internet
- Reasons for incredible growth in IoT:
  - Hardware – smaller, affordable, more powerful
  - Availability of BI tools – more capable and cheaper
  - Emergence of new and innovative use cases
- There isn't a universal agreement on the term IoT
  - Web of Things
  - Internet of Systems, …

Application Case 8.1

SilverHook Powerboats Uses Real-Time Data Analysis to Inform Racers and Fans

Questions for Discussion

1. What type of information might the sensors on a race boat generate that would be important for the racers to know? What about for the fans?
2. Which other sports might benefit from similar technologies?
3. What technological challenges might you face in building such systems?

Application Case 8.2

Rockwell Automation Monitors Expensive Oil and Gas Exploration Assets

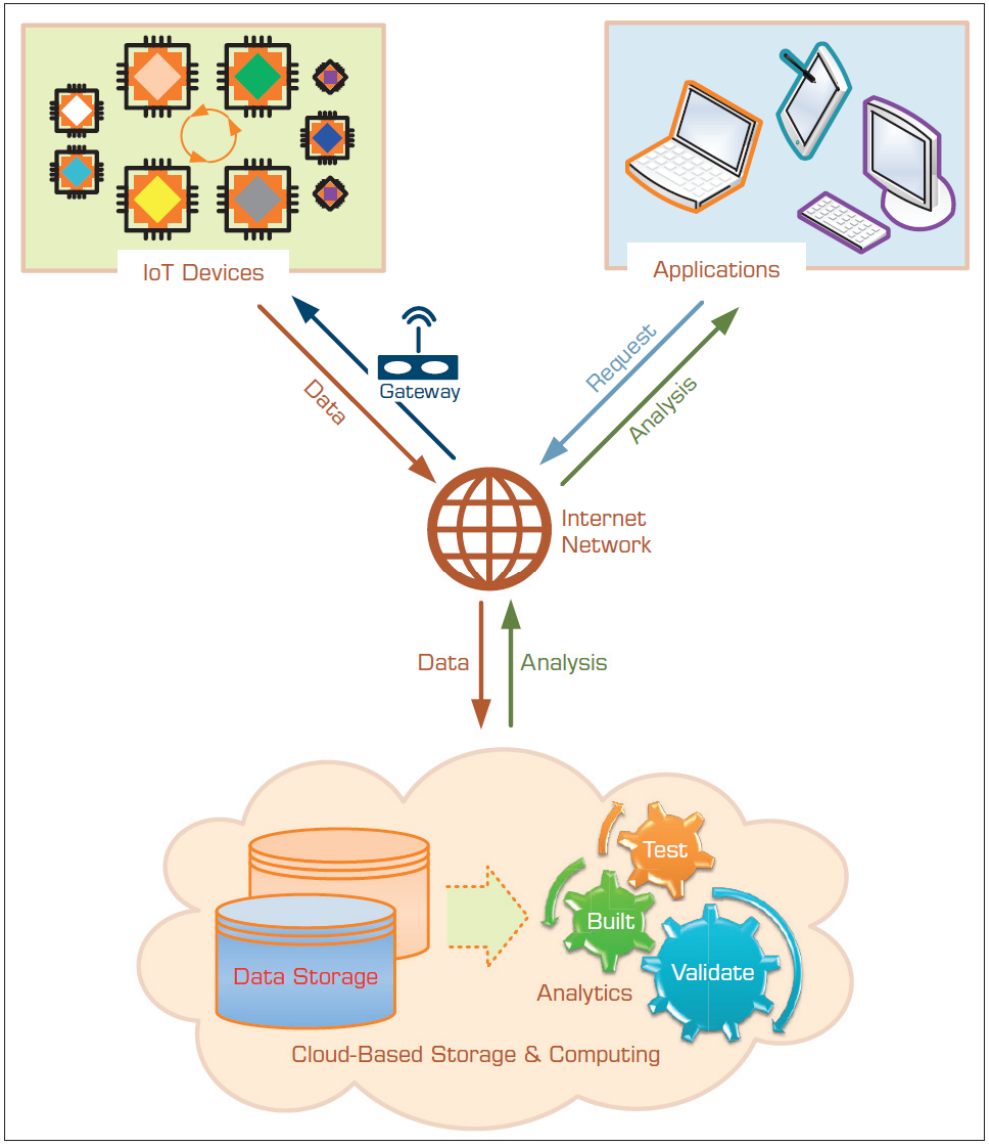
Questions for Discussion

1. What type of information would likely be collected by an oil and gas drilling platform?
2. Does this application fit the three V's of Big Data (volume, variety, velocity)? Why or why not?
3. Which other industries could use similar operational measurements and dashboards?

IoT Technology Infrastructure

- IoT related technology components can be divided into four major blocks:
  1. Hardware
     - physical devices, sensors, and actuators
  2. Connectivity
     - Collecting and sending sensory data to the cloud
  3. Software
     - Integrating, and processing data for patterns
  4. Applications
     - Creating context specific alerts, actionable insight

Building Blocks of IoT Technology Infrastructure



RFID Sensors

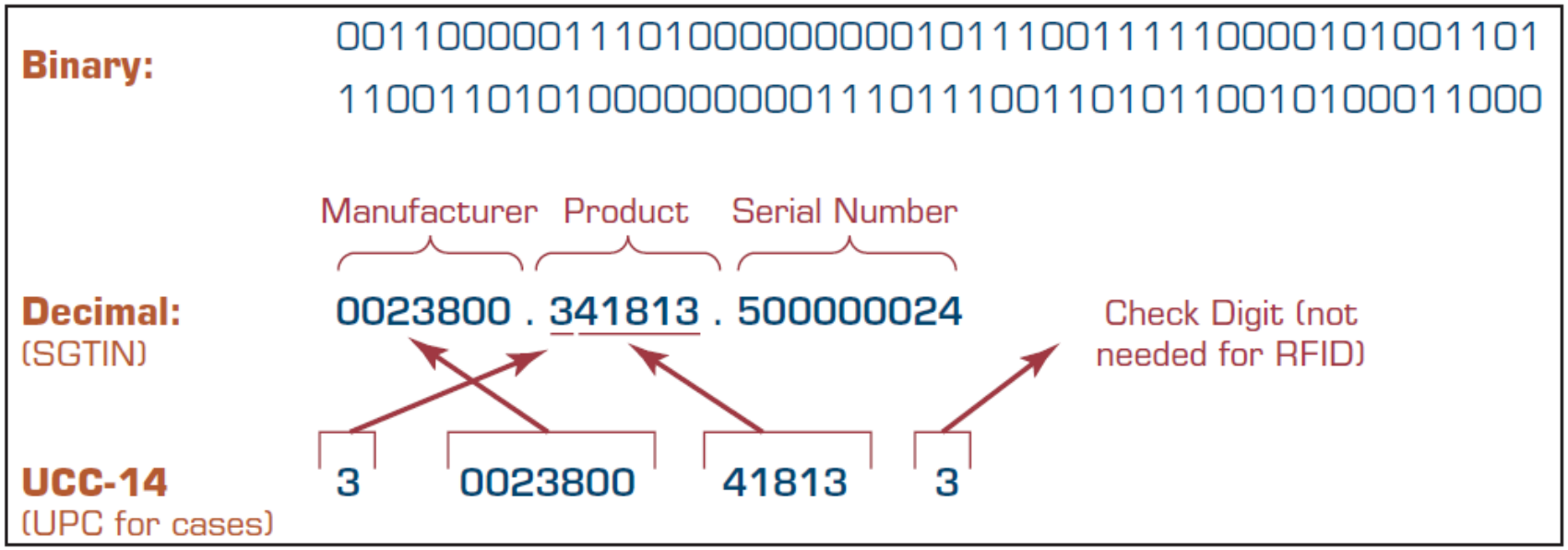
- RFID: radio-frequency identification
- One of the earliest/disruptive sensor technologies
- Part of a family of automatic identification technologies
  - Including ubiquitous barcodes and magnetic strips
- The goal is to use radio-frequency waves to accurately and quickly identify objects
- Use of RFID is led/promoted by large retailers
  - Wal-Mart, Target, Dillard's

RFID Sensors

- How does RFID work?
  - Tag – a circuit attached to the product to be identified
  - Interrogator (i.e., reader) – with antennas and a computer to detect objects, store the data, and take due actions
- Tags can be passive or active
  - Passive tag – small, inexpensive, no power source
  - Active tag – larger, more expensive, has power source
- Which one is better?
  - Retail uses passive tags, others may use active tags
- RFID + Sensors can be used for perishable goods

Data Representation in RFID

- Data representation for a given application domain
- For Retail: Electronic Product Code (EPC)
- RFID tags contain 96 bits of data



Fog Computing

- Data produced by IoT is huge in size (problem)
- Fog computing is to address the issue by
  - Proposing fog nodes to process the data close to IoT
  - Fog nodes – any device including routers or switches

TABLE 8.1 Difference between Fog Nodes and a Cloud Platform

Fog Nodes	Cloud Platform
Receive data from IoT devices	Receives and aggregates data from fog nodes
Run IoT real-time analytics in millisecond response time	Analysis is performed on huge amounts of business data and can take hours or weeks

  - Data Center/Cloud → Fog Device → Physical Device (i.e., sensors that generate the data)

Other IoT Considerations

- IoT Platforms – many large companies are in it
  - Amazon AWS IoT, Microsoft Azure IoT Suite, Predix IoT, Platform by General Electric (GE), IBM Watson IoT solutions, and Teradata Unified Data Architecture
- IoT Start-up Ecosystem
  - Many start-up companies are emerging in the field of IoT
  - Examples include Sigfox, 3D Robotics, Canary, Athos, Greenwave, Jawbone, FreedomPop, Razer, and Ring
  - Fitbit – one of the most successful IoT startups
- See Figure 8.3 for a pictorial representation of IoT Ecosystem

Internet of Things (IoT) Ecosystem

Managerial Considerations in the Internet of Things

1. Organizational Alignment
2. Interoperability Challenges
3. Security

- Emerging growth of IoT and its potential to help us achieve the vision of *smart cities*, *smart grid*, smart *anything*

Cloud Computing and Business Analytics

- A style of computing in which dynamically scalable and often virtualized resources are provided over the Internet.
- Users need not have knowledge of, experience in, or control over the technology infrastructures in the cloud that supports them.
- Cloud computing = utility computing, application service provider grid computing, on-demand computing, software-as-a-service (SaaS), …
  - Cloud = Internet
  - Related “-as-a-services”: infrastructure-as-a-service (IaaS), platforms-as-a-service (PaaS)

Cloud Computing Example

- Web-based e-mail → cloud computing application
  - Stores the data (e-mail messages)
  - Stores the software (e-mail programs)
  - Centralized hardware/software/infrastructure
  - Centralized updates/upgrades
  - Access from anywhere via a Web browser
  - e.g., Gmail
- Web-based general application = cloud application
  - Google Docs, Google Spreadsheets, Google Drive,…
  - Amazon.com's Web Services

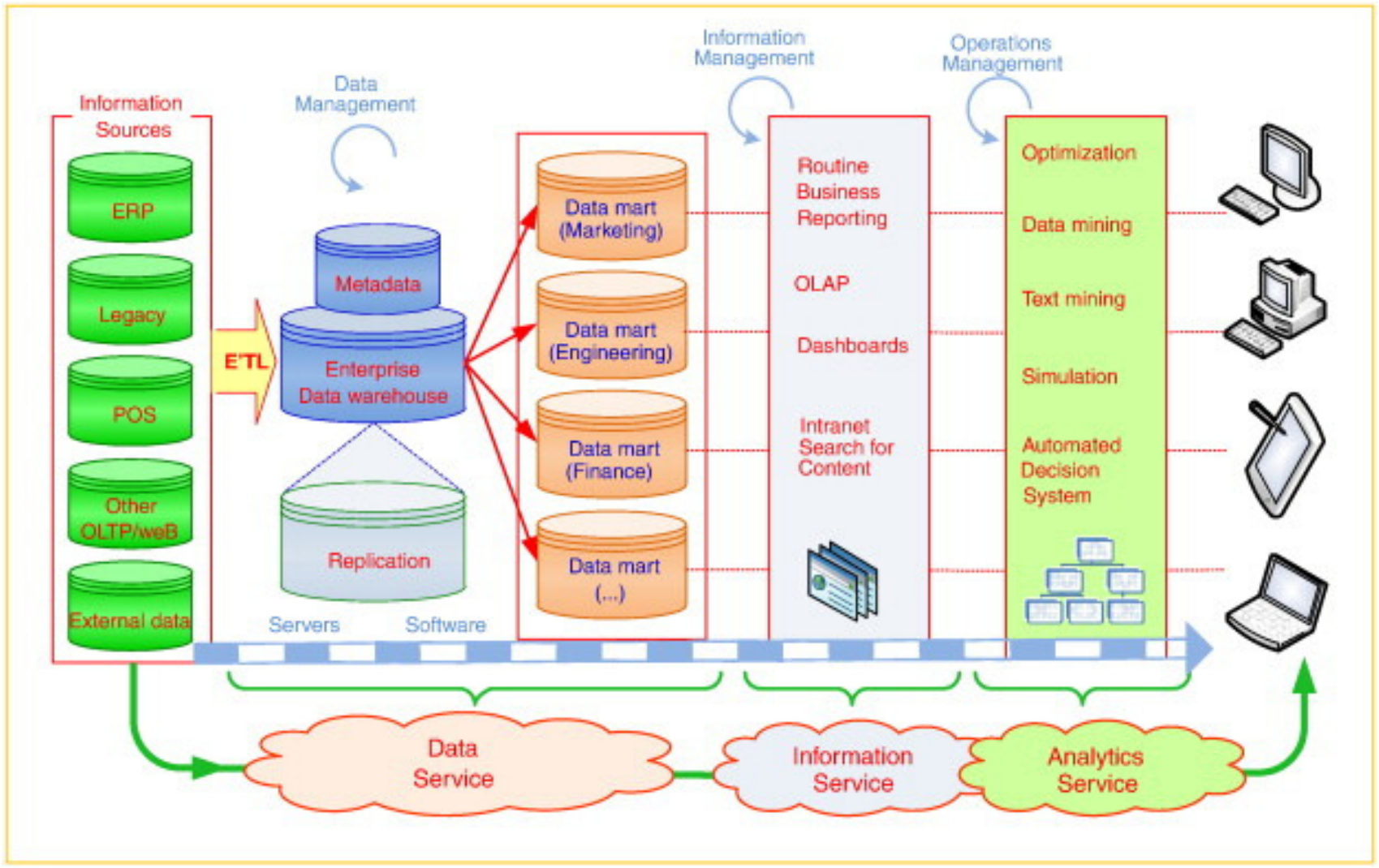
Cloud Computing Example

- Cloud computing is used in
  - e-commerce, BI, CRM, SCM, …
- Business model
  - Pay-per-use
  - Subscribe/pay-as-you-go
- Companies that offer cloud-computing services
  - Google, Yahoo!, Salesforce.com
  - IBM, Microsoft (Azure)
  - Sun Microsystems/Oracle

Cloud Computing and Service-Oriented Thinking

- Service-oriented thinking is one of the fastest-growing paradigms today
- Toward building agile data, information, and analytics capabilities as services
- Service orientation + DSS/BI
- Component-based service orientation fosters
  - Reusability, Substitutability, Extensibility, Scalability, Customizability, Reliability, Low Cost of Ownership, Economy of Scale,…

Service-Oriented DSS/BI

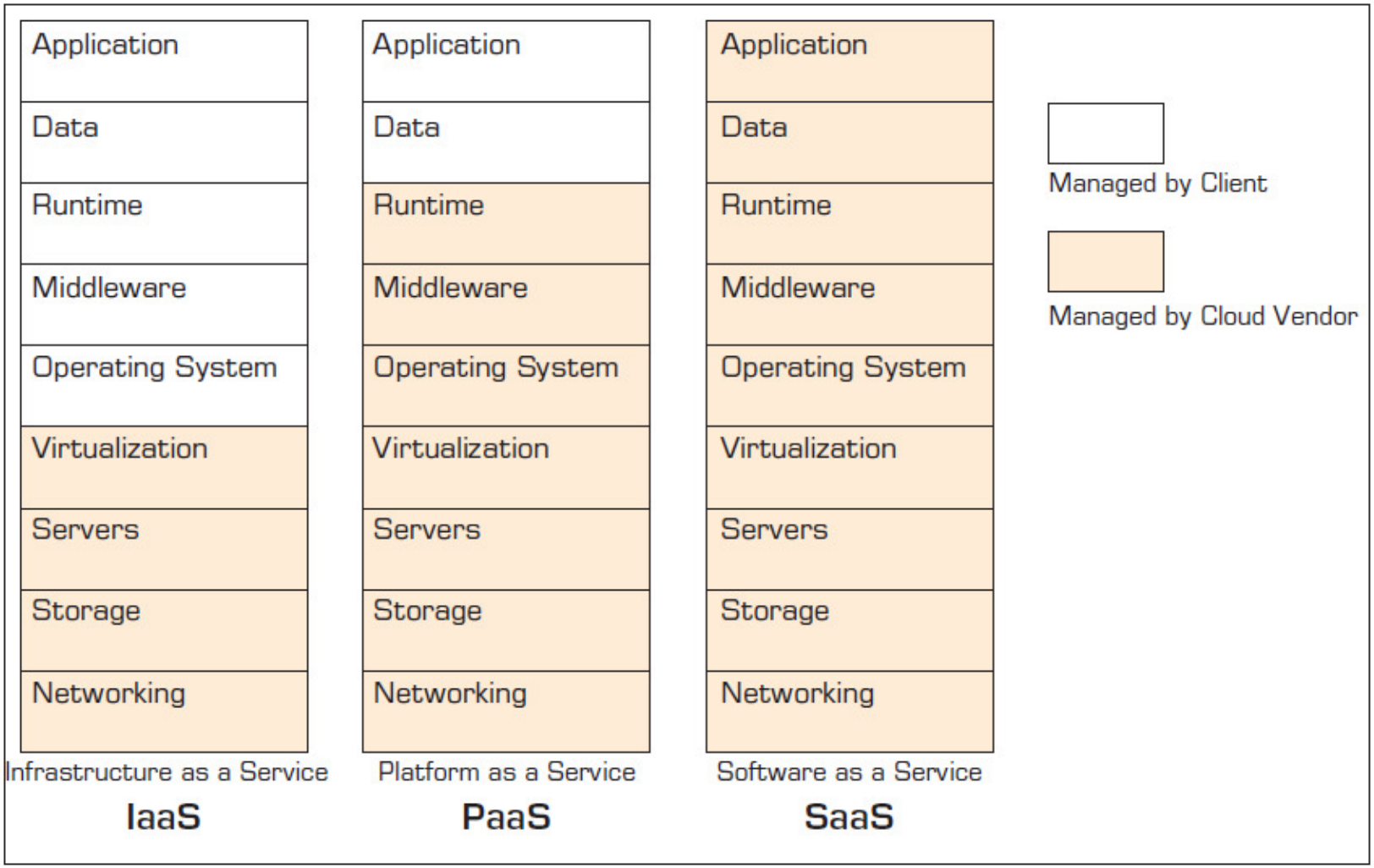


Variations of Service-Oriented Architecture and the Cloud

- Data as a Service (DaaS)
- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)
- …

- Why so many .aaS is emerging?

Different Types of Cloud Offerings



Essential Technologies for Cloud Computing

- Virtualization
  - Creation of a virtual version of something like an operating system or server
  - Example: logical division of a hard drive to create two separate hard drives in a computer
- Levels of virtualization
  - Network virtualization
  - Storage virtualization
  - Server virtualization
- Relates to which cloud service is employed

Cloud Deployment Models

- Private cloud
- Public cloud
- Hybrid cloud
- Which cloud model is good for you?
- Major cloud platform providers in analytics:
  - Amazon Elastic Beanstalk
  - IBM Bluemix
  - Microsoft Azure
  - Google App Engine
  - OpenShift

Representative Analytics as a Service Offering

- Teradata - Aster Analytics as a Service
- IBM Watson Analytics
- MineMyText.com
- SAS Visual Analytic and Visual Statistics
- Tableau
- Showflake
- Predix by General Electric

→ Most of these have free/restricted/trial offerings

Illustrative Analytics Applications Employing the Cloud Infrastructure

- MD Anderson Cancer Center Utilizes Cognitive Computing Capabilities of IBM Watson to Give Better Treatment to Cancer Patients
- Public School Education in Tacoma, Washington, Uses Microsoft Azure Machine Learning to Predict School Dropouts
- Dartmouth-Hitchcock Medical Center Provides Personalized Proactive Healthcare Using Microsoft Cortana Analytics Suite

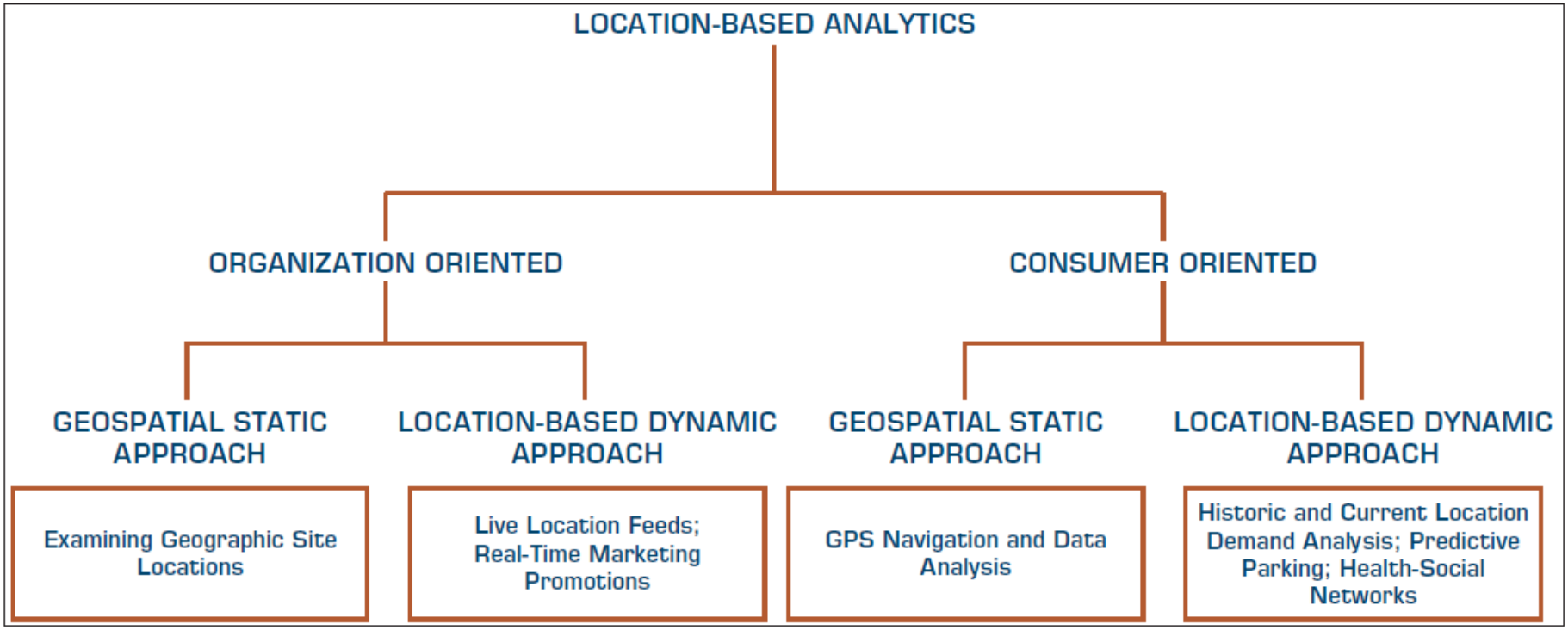
Illustrative Analytics Applications Employing the Cloud Infrastructure

- Mankind Pharma Uses IBM Cloud Infrastructure to Reduce Application Implementation Time by 98%
- Gulf Air Uses Big Data to Get Deeper Customer Insight
- Chime Enhances Customer Experience Using Snowflake

Location-Based Analytics

- Geospatial Analytics
- Geocoding
  - Visual maps
  - Postal codes
  - Latitude & Longitude
- Enables aggregate view of a large geographic area
- Integrate “where” into customer view

Location-Based Analytics



Location-Based Analytics

- Location-based databases
- Geographic Information System (GIS)
  - Used to capture, store, analyze, and manage the data linked to a location
  - Combined with integrated sensor technologies and global positioning systems (GPS)
- Location Intelligence (LI)?
  - Interactive maps that further drill down to data/information details about any location

Use of Location-Based Analytics

- Retailers – location + demographic details combined with other transactional data can help …
  - determine how sales vary by population level
  - assess locational proximity to other competitors and their offerings
  - assess the demand variations and efficiency of supply chain operations
  - analyze customer needs and complaints
  - better target different customer segments
  - …

GIS Applications

- In addition to business/retail applications, GIS based analytics are being used in
  - Agricultural applications
  - Crime analysis
  - Disease spread prediction
- For more applications, look at
  - esri.com (producer of ArcGIS)
  - grindgis.com
- LI can be combined with weather and environmental data to create a richer data/information infrastructure

Application Case 8.4

Great Clips Employs Spatial Analytics to Shave Time in Location Decisions

Questions for Discussion

1. How is geospatial analytics employed at Great Clips?
2. What criteria should a company consider in evaluating sites for future locations?
3. Can you think of other applications where such geospatial data might be useful?

Application Case 8.5

Starbucks Exploits GIS and Analytics to Grow Worldwide

Questions for Discussion

1. What type of demographics and GIS information would be relevant for deciding on a store location?
2. It has been mentioned that Starbucks encourages its customers to use its mobile app. What type of information might the company gather from the app to help it better plan operations?
3. Will the availability of free Wi-Fi at Starbucks' stores provide any information to Starbucks for better analytics?

A Multimedia Exercise in Analytics Employing Geospatial Analytics

- Go To Teradata University Network (TUN)
- Find the BSI Case video on "The Case of the Dropped Mobile Calls"
- Watch the video via TUN or at YouTube youtube.com/watch?v=4WJR_Z3exw4
- Also, look at the slides at slideshare.net/teradata/bsi-teradata-the-case-of-the-dropped-mobile-calls
- Discuss the case

Real-Time Location Intelligence

- Many devices are constantly sending out their location information
  - Cars, airplanes, ships, mobile phones, cameras, navigation systems, …
    - GPS, Wi-Fi, RFID, cell tower triangulation
- Reality mining?
  - Real-time location information = real-time insight
  - Path Intelligence (pathintelligence.com)
    - Footpath – movement patterns within a city or store
    - How to use such movement information

Application Case 8.6
Quiznos Targets Customers for Its Sandwiches

Questions for Discussion

1. How can location-based analytics help retailers in targeting customers?
2. Research similar applications of location-based analytics in the retail domain.

Analytics Applications for Consumers

- Explosive growth of the apps industry
  - iOS, Android, Windows, Blackberry, Amazon, …
  - Directly used by consumers (not businesses)
  - Enabling consumers to become more efficient
  - Interesting Examples
    - CabSense – finding a taxi in New York City
      - Rating of street corners; interactive maps, …
    - ParkPGH – finding a parking spot
      - Downtown Pittsburgh, Pennsylvania
  - App industry is already 25B in size and growing (wsj.com/apps)