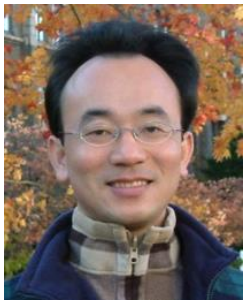
Tamkang University

商業智慧實務
Practices of Business Intelligence

資料科學與巨量資料分析
(Data Science and Big Data Analytics)

1032BI06
MI4
Wed, 9,10 (16:10-18:00) (B130)



Min-Yuh Day
戴敏育
Assistant Professor
專任助理教授
Dept. of Information Management, Tamkang University
淡江大學 資訊管理學系
http://mail. tku.edu.tw/myday/

2015-04-15

課程大綱 (Syllabus)

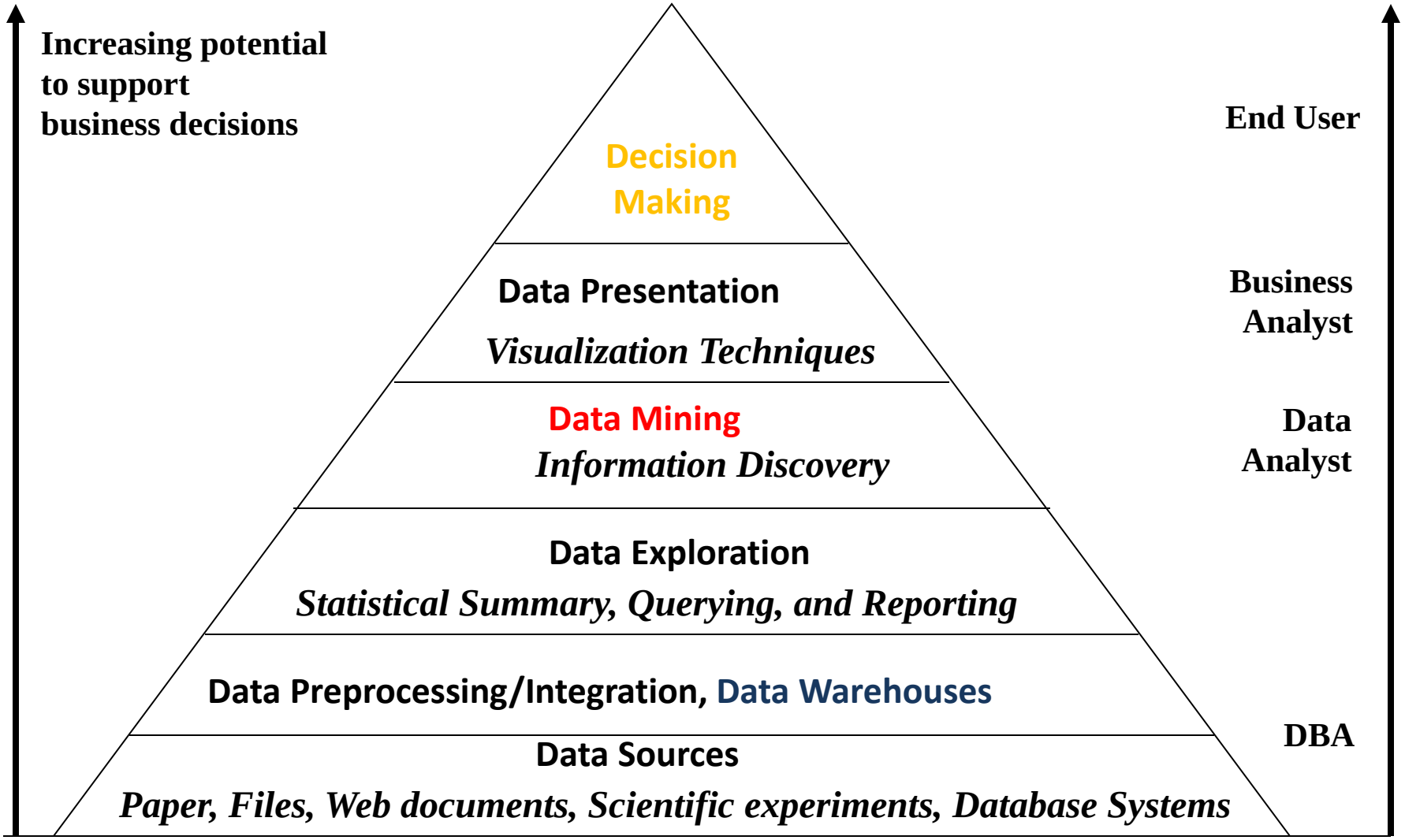
週次 (Week)	日期 (Date)	內容 (Subject/Topics)
1	2015/02/25	商業智慧導論 (Introduction to Business Intelligence)
2	2015/03/04	管理決策支援系統與商業智慧 (Management Decision Support System and Business Intelligence)
3	2015/03/11	企業績效管理 (Business Performance Management)
4	2015/03/18	資料倉儲 (Data Warehousing)
5	2015/03/25	商業智慧的資料探勘 (Data Mining for Business Intelligence)
6	2015/04/01	教學行政觀摩日 (Off-campus study)
7	2015/04/08	商業智慧的資料探勘 (Data Mining for Business Intelligence)
8	2015/04/15	資料科學與巨量資料分析 (Data Science and Big Data Analytics)

課程大綱 (Syllabus)

週次	日期	內容（Subject/Topics）
9	2015/04/22	期中報告 (Midterm Project Presentation)
10	2015/04/29	期中考試週 (Midterm Exam)
11	2015/05/06	文字探勘與網路探勘 (Text and Web Mining)
12	2015/05/13	意見探勘與情感分析 (Opinion Mining and Sentiment Analysis)
13	2015/05/20	社會網路分析 (Social Network Analysis)
14	2015/05/27	期末報告 (Final Project Presentation)
15	2015/06/03	畢業考試週 (Final Exam)

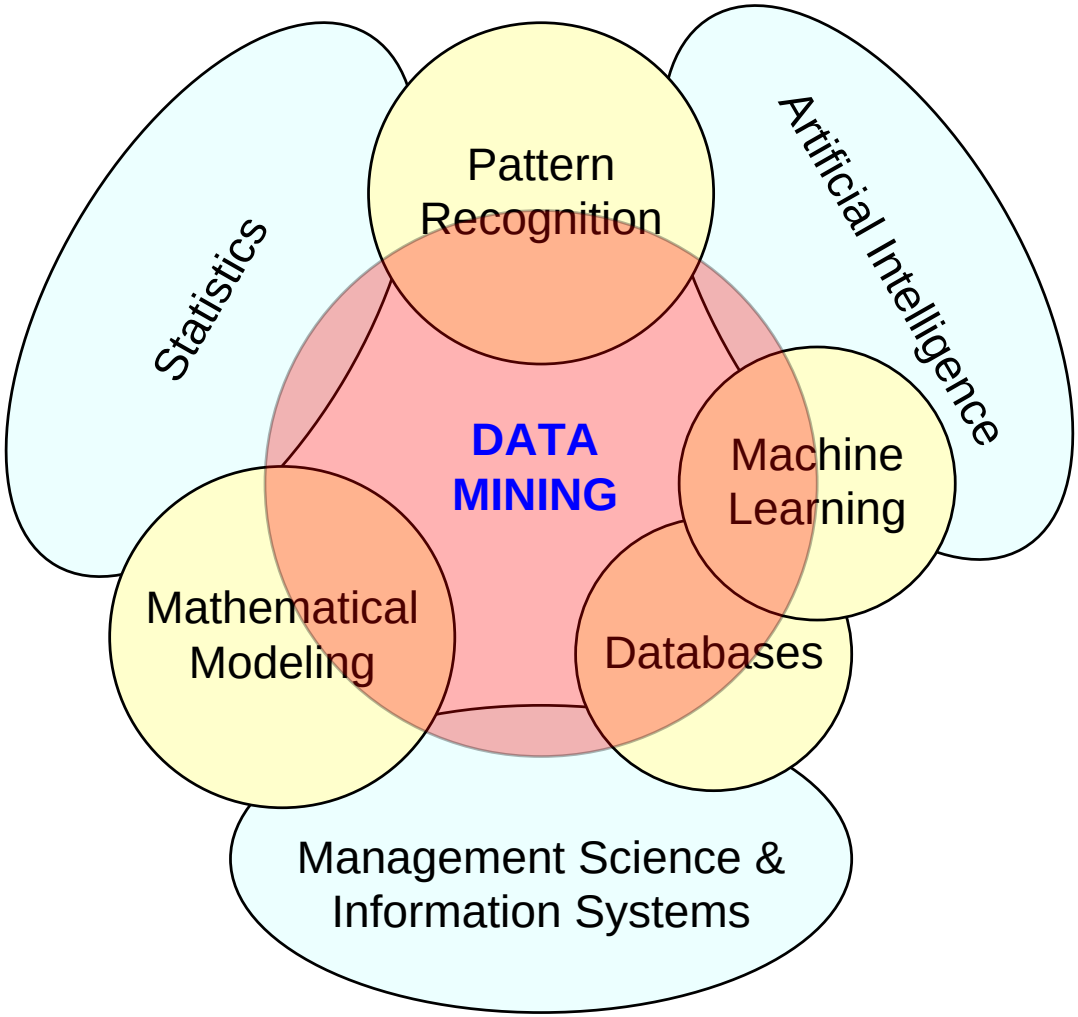
Business Intelligence

Data Mining, Data Warehouses



Source: Han & Kamber (2006)

Data Mining at the Intersection of Many Disciplines



Source: Turban et al. (2011), Decision Support and Business Intelligence Systems

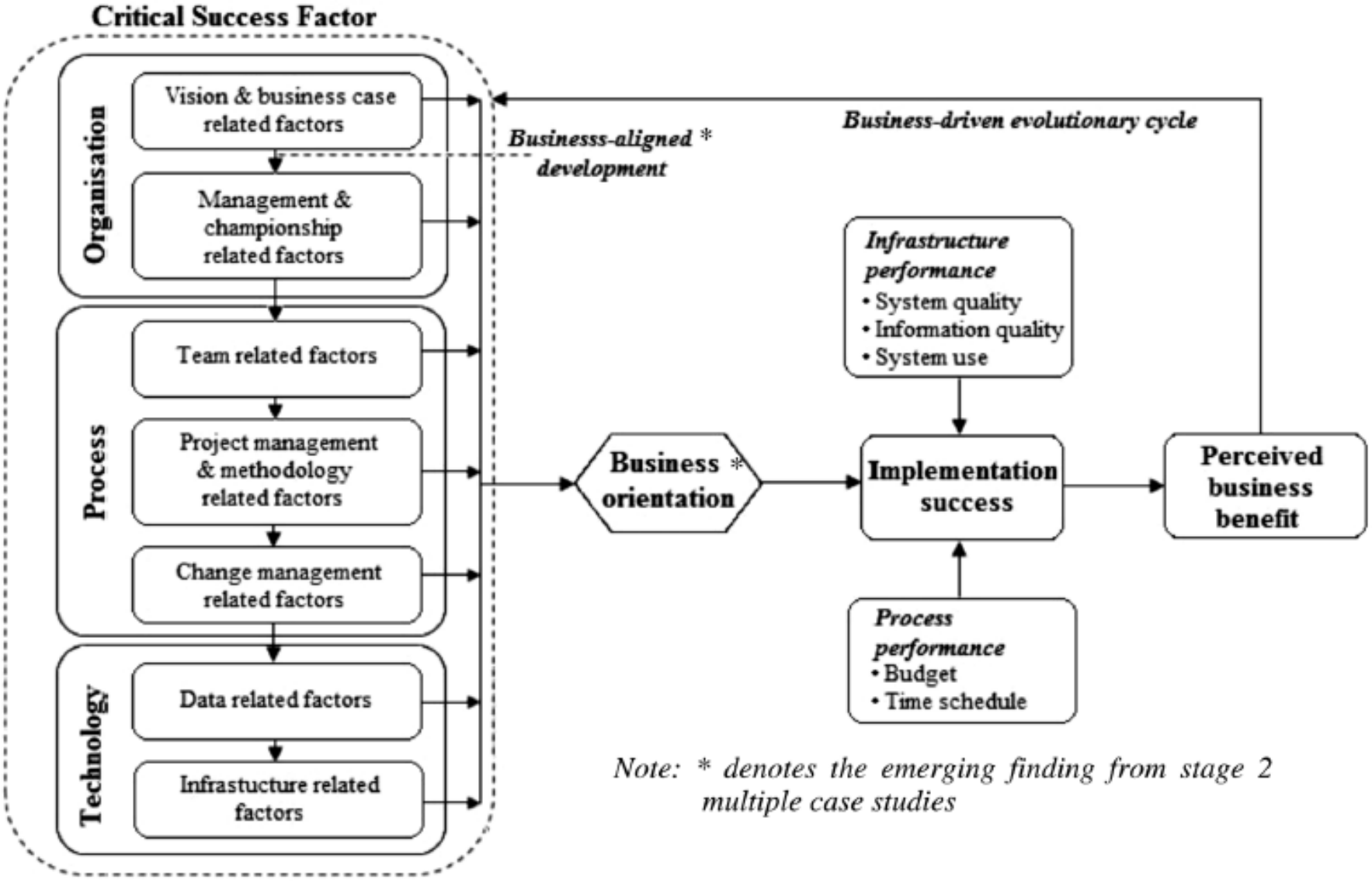
Outline

- Business Intelligence Implementation
- Business Intelligence Trends
- Data Science
- Big Data Analytics
  - Big Data, Big Analytics:
    Emerging Business Intelligence and Analytic Trends for Today's Businesses

Business Intelligence Implementation

Business Intelligence Implementation

CSFs Framework for Implementation of BI Systems



Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence systems. Journal of computer information systems, 50(3), 23.

Critical Success Factors of Business Intelligence Implementation

- Organizational dimension
  - Committed management support and sponsorship
  - Clear vision and well-established business case
- Process dimension
  - Business-centric championship and balanced team composition
  - Business-driven and iterative development approach
  - User-oriented change management.
- Technological dimension
  - Business-driven, scalable and flexible technical framework
  - Sustainable data quality and integrity

Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence systems. Journal of computer information systems, 50(3), 23.

Business Intelligence Trends

Business Intelligence Trends

1. Agile Information Management (IM)
2. Cloud Business Intelligence (BI)
3. Mobile Business Intelligence (BI)
4. Analytics
5. Big Data

Source: http://www.businessspectator.com.au/article/2013/1/22/technology/five-business-intelligence-trends-2013

Business Intelligence Trends: Computing and Service

- Cloud Computing and Service
- Mobile Computing and Service
- Social Computing and Service

Business Intelligence and Analytics

- Business Intelligence 2.0 (BI 2.0)
  - Web Intelligence
  - Web Analytics
  - Web 2.0
  - Social Networking and Microblogging sites
- Data Trends
  - Big Data
- Platform Technology Trends
  - Cloud computing platform

Source: Lim, E. P., Chen, H., & Chen, G. (2013). Business Intelligence and Analytics: Research Directions. *ACM Transactions on Management Information Systems (TMIS)*, *3*(4), 17

Business Intelligence and Analytics: Research Directions

1. Big Data Analytics
   - Data analytics using Hadoop / MapReduce framework
2. Text Analytics
   - From Information Extraction to Question Answering
   - From Sentiment Analysis to Opinion Mining
3. Network Analysis
   - Link mining
   - Community Detection
   - Social Recommendation

Source: Lim, E. P., Chen, H., & Chen, G. (2013). Business Intelligence and Analytics: Research Directions. *ACM Transactions on Management Information Systems (TMIS)*, *3*(4), 17

Data Science

“Data science

is

the study of the generalizable extraction of knowledge from data ”

Source: Dhar, V. (2013). Data science and prediction. Communications of the ACM, 56(12), 64-73.

Data Science

- A common epistemic requirement in assessing whether new knowledge is actionable for decision making is its predictive power, not just its ability to explain the past.

Source: Dhar, V. (2013). Data science and prediction. Communications of the ACM, 56(12), 64-73.

Data Scientist

- A data scientist requires an integrated skill set spanning mathematics, machine learning, artificial intelligence, statistics, databases, and optimization, along with a deep understanding of the craft of problem formulation to engineer effective solutions.

Source: Dhar, V. (2013). Data science and prediction. Communications of the ACM, 56(12), 64-73.

Data Scientist:

The Sexiest Job of the 21st Century

(Davenport & Patil, 2012)(HBR)

Source: Davenport, T. H., & Patil, D. J. (2012). Data Scientist. *Harvard business review*

Data Scientist:

The Sexiest Job of the 21st Century

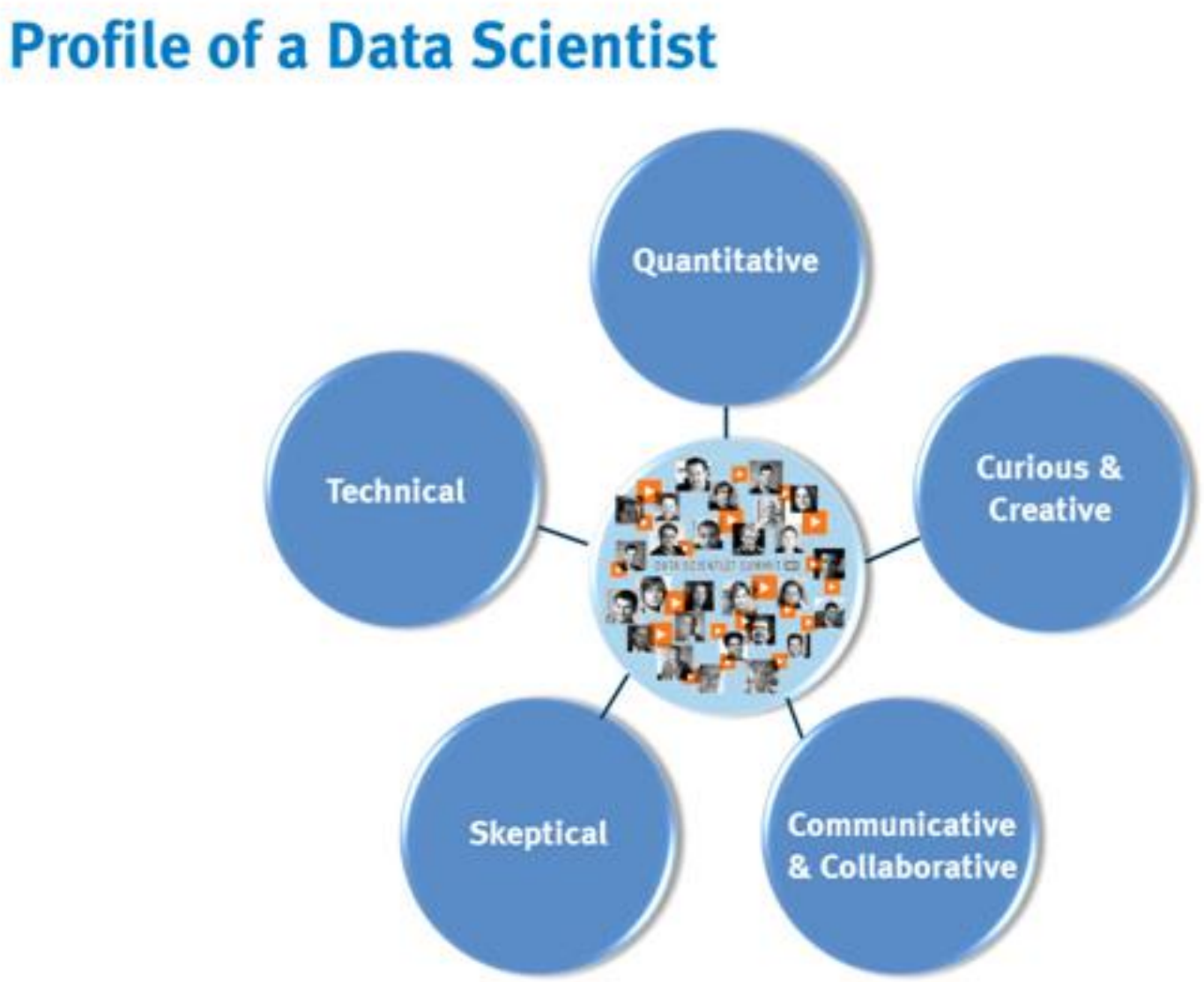
Meet the people who can coax treasure out of messy, unstructured data.

by Thomas H. Davenport and D.J. Patil

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

Source: Davenport, T. H., & Patil, D. J. (2012). Data Scientist. *Harvard business review*

Data Scientist



Source: https://infocus.emc.com/david_dietrich/what-is-the-profile-of-a-data-scientist/

Data Science

and its Relationship to

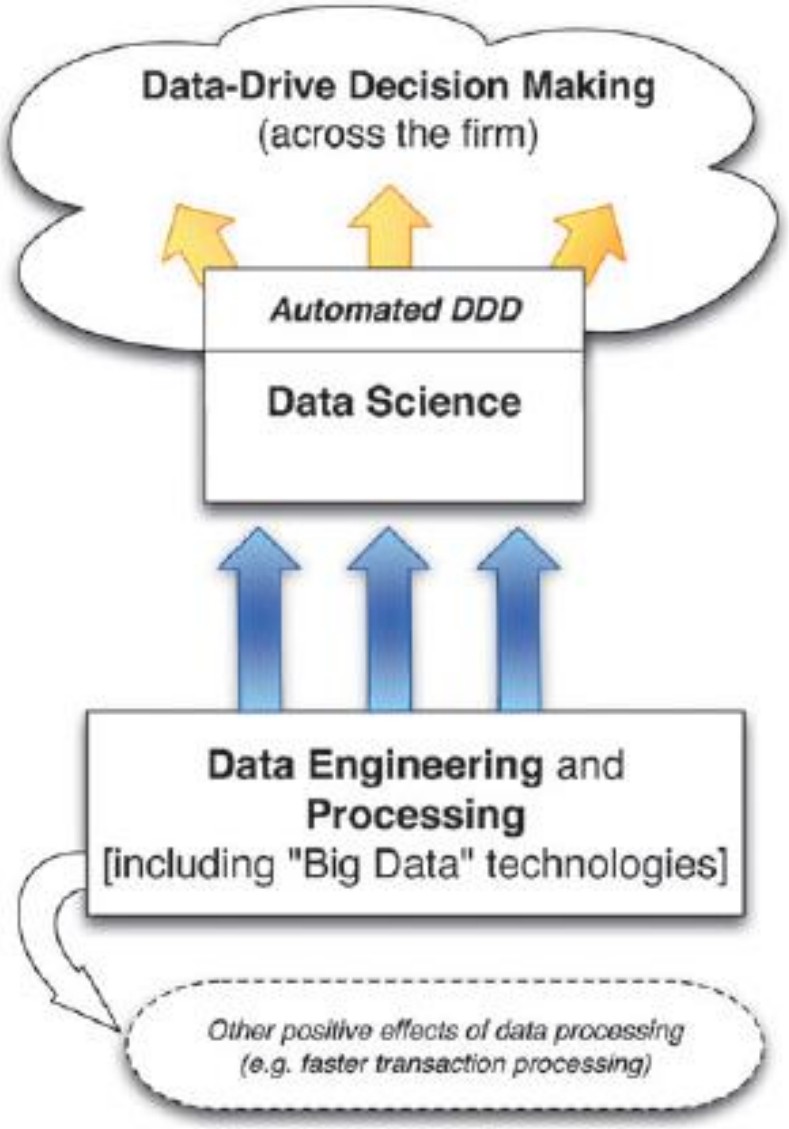
Big Data

and

Data-Driven Decision Making

Source: Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51-59.

Data science in the organization



Source: Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51-59.

Big Data Analytics

Big Data, Big Analytics:

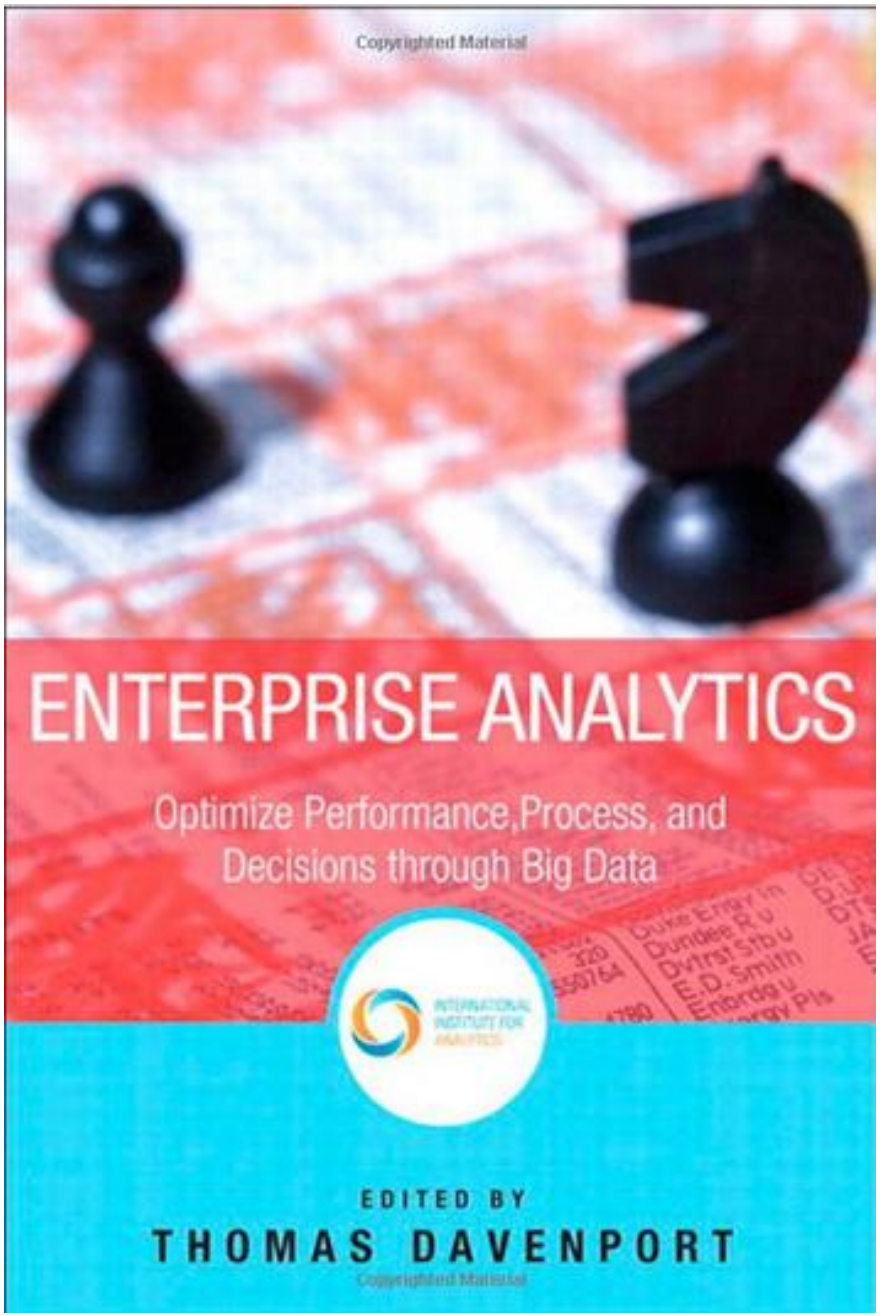
Emerging Business Intelligence and Analytic Trends for Today's Businesses

Big Data:

The Management Revolution

Source: McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution.*Harvard business review*.

Source: McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution.*Harvard business review*.



Source: http://www.amazon.com/Enterprise-Analytics-Performance-Operations-Management/dp/0133039439

Business Intelligence and Enterprise Analytics

- Predictive analytics
- Data mining
- Business analytics
- Web analytics
- Big-data analytics

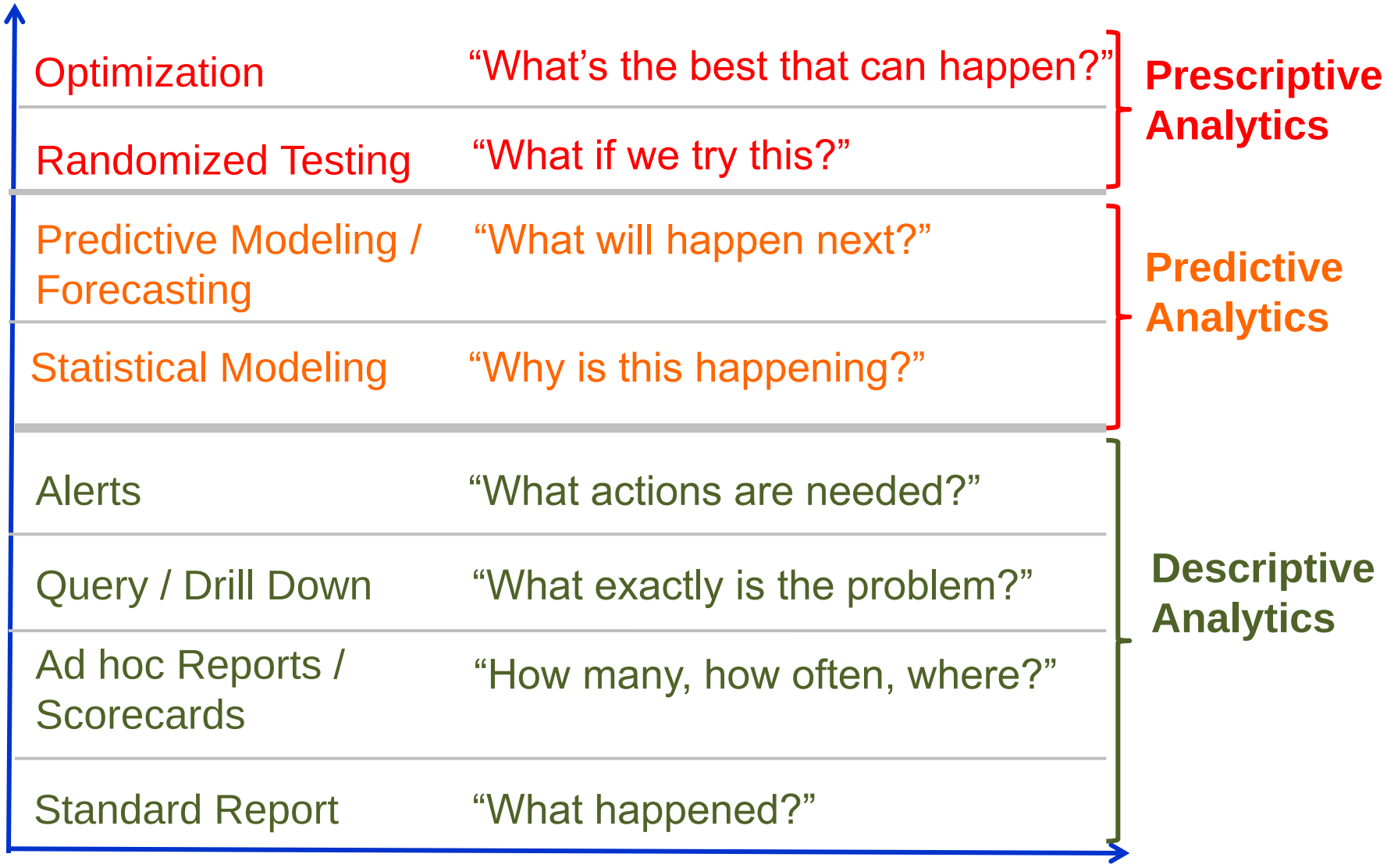
Source: Thomas H. Davenport, "Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data", FT Press, 2012

Three Types of Business Analytics

- Prescriptive Analytics
- Predictive Analytics
- Descriptive Analytics

Source: Thomas H. Davenport, "Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data", FT Press, 2012

Three Types of Business Analytics



Source: Thomas H. Davenport, "Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data", FT Press, 2012

Big-Data Analysis

- **Too Big,**
  too Unstructured,
  too many different source
  to be manageable through traditional databases

The Rise of "Big Data"

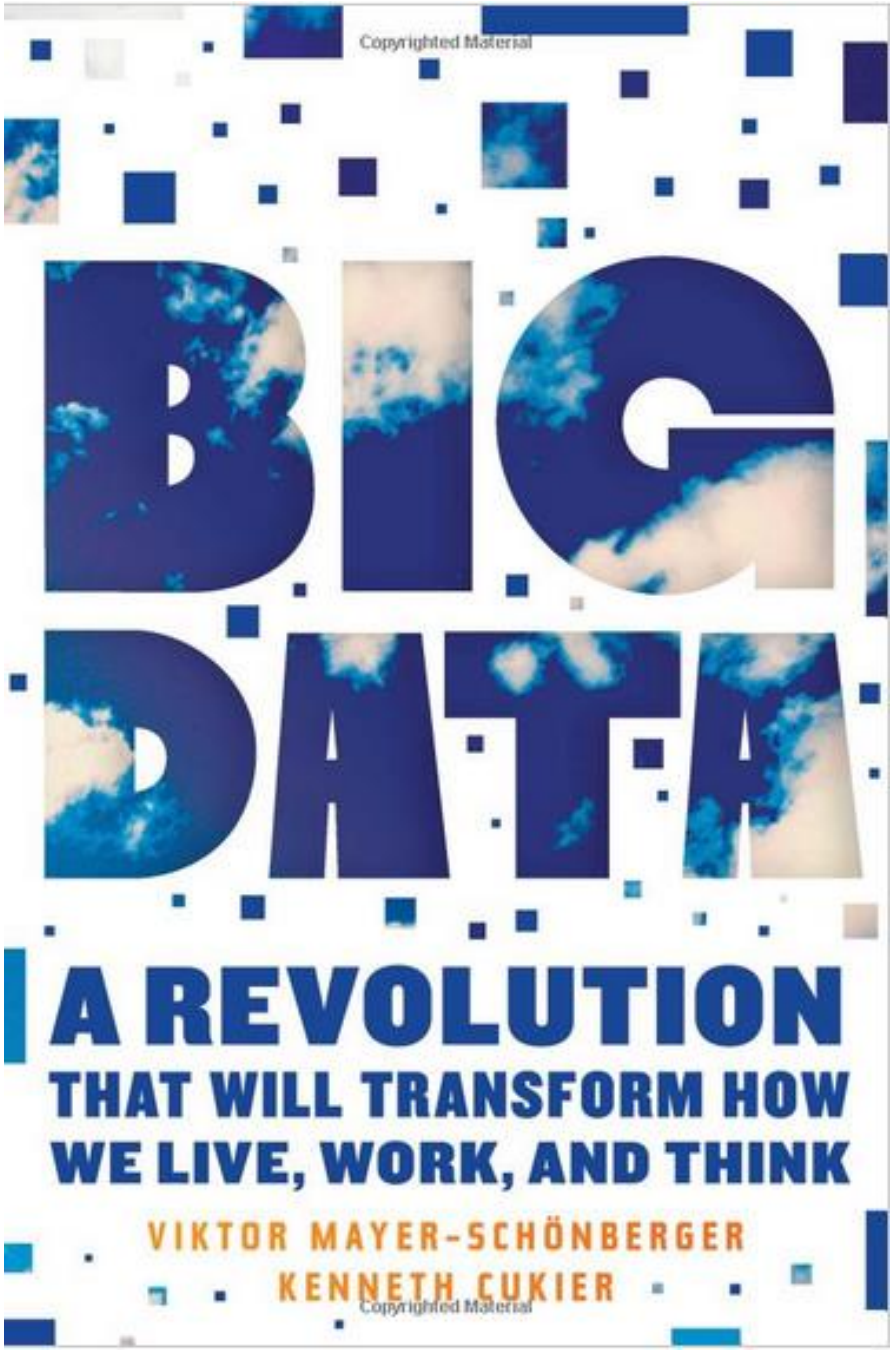
- "Too Big" means databases or data flows in **petabytes (1,000 terabytes)**
  - Google processes about 24 petabytes of data per day
- "Too unstructured" means that the data isn't easily put into the traditional rows and columns of conventional databases

Source: Thomas H. Davenport, "Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data", FT Press, 2012

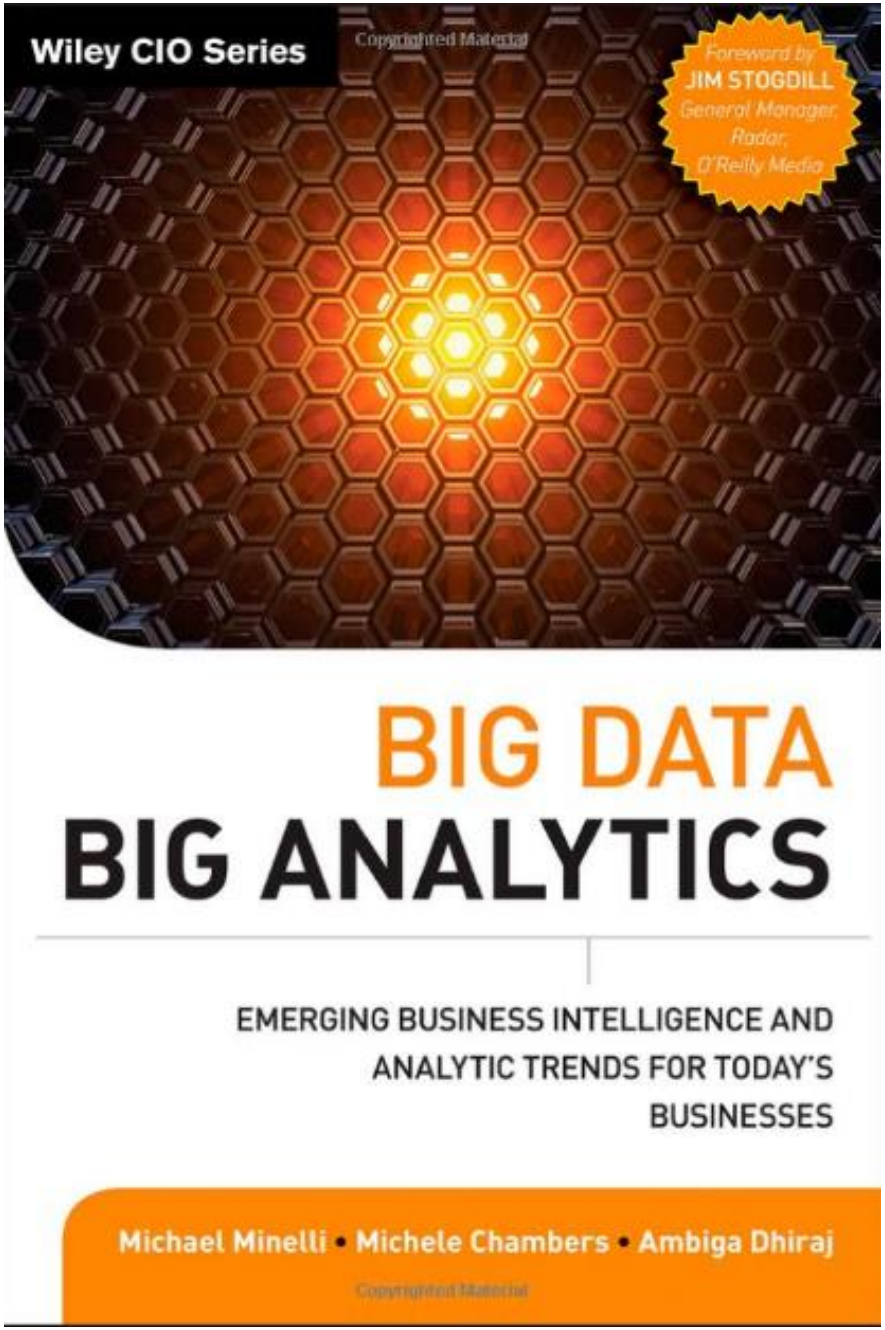
Examples of Big Data

- Online information
  - Clickstream data from Web and social media content
    - Tweets
    - Blogs
    - Wall postings
- Video data
  - Retail and crime/intelligence environments
  - Rendering of video entertainment
- Voice data
  - call centers and intelligence intervention
- Life sciences
  - Genomic and proteomic data from biological research and medicine

Source: Thomas H. Davenport, "Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data", FT Press, 2012



Source: http://www.amazon.com/Big-Data-Analytics-Intelligence-Businesses/dp/111814760X



Source: http://www.amazon.com/Big-Data-Analytics-Intelligence-Businesses/dp/111814760X

Big Data, Big Analytics:

Emerging Business Intelligence and Analytic Trends for Today's Businesses

- What Big Data is and why it's important
- Industry examples (Financial Services, Healthcare, etc.)
- Big Data and the New School of Marketing
- Fraud, risk, and Big Data
- Big Data technology
- Old versus new approaches
- Open source technology for Big Data analytics
- The Cloud and Big Data

Source: http://www.amazon.com/Big-Data-Analytics-Intelligence-Businesses/dp/111814760X

Big Data, Big Analytics:
Emerging Business Intelligence and Analytic Trends for Today's Businesses

- Predictive analytics
- Crowdsourcing analytics
- Computing platforms, limitations, and emerging technologies
- Consumption of analytics
- Data visualization as a way to take immediate action
- Moving from beyond the tools to analytic applications
- Creating a culture that nurtures decision science talent
- A thorough summary of ethical and privacy issues

Source: http://www.amazon.com/Big-Data-Analytics-Intelligence-Businesses/dp/111814760X

What is BIG Data?

Volume

Large amount of data

Velocity

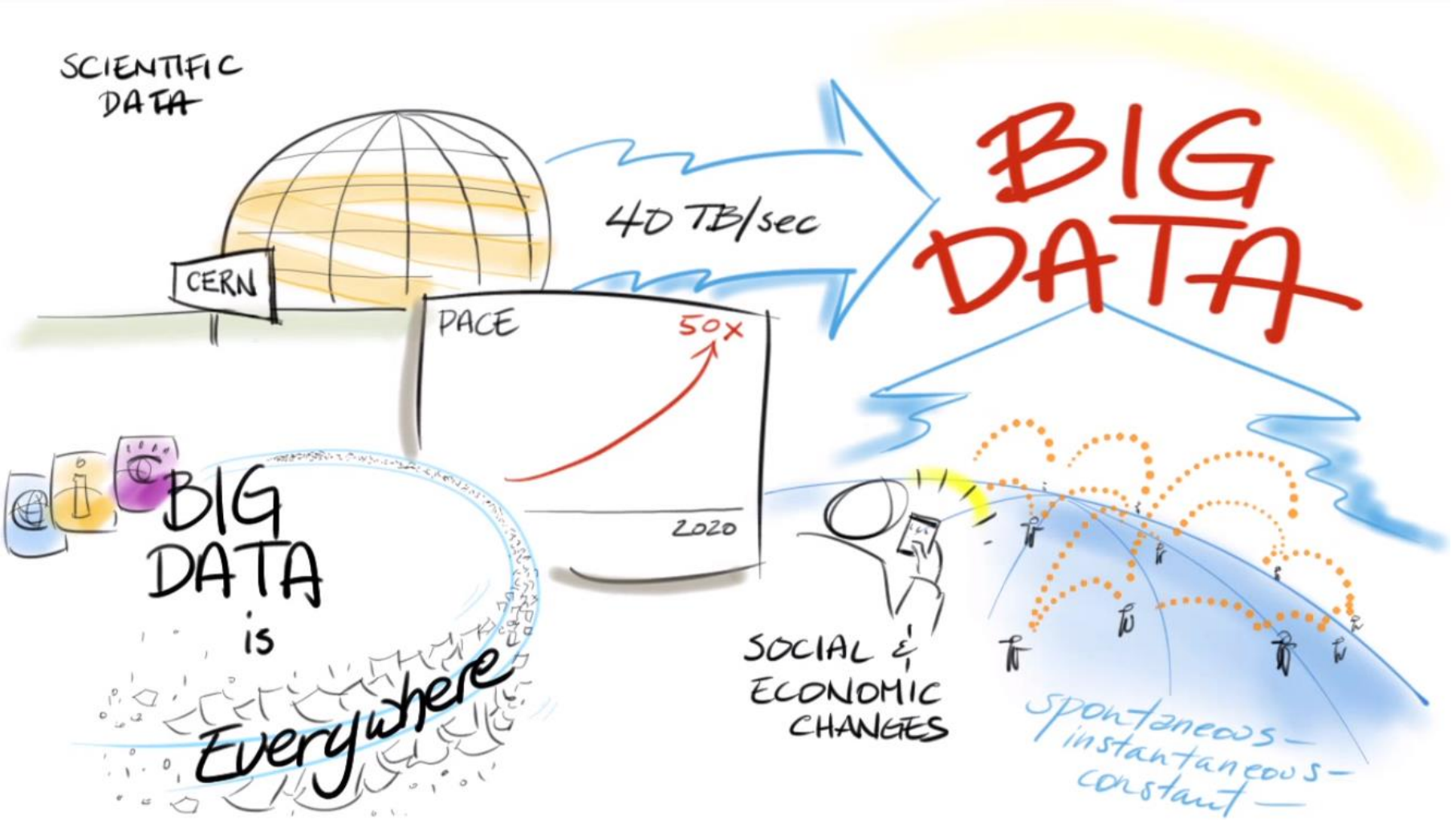
Needs to be analyzed **quickly**

Variety

Different types of structured and unstructured data

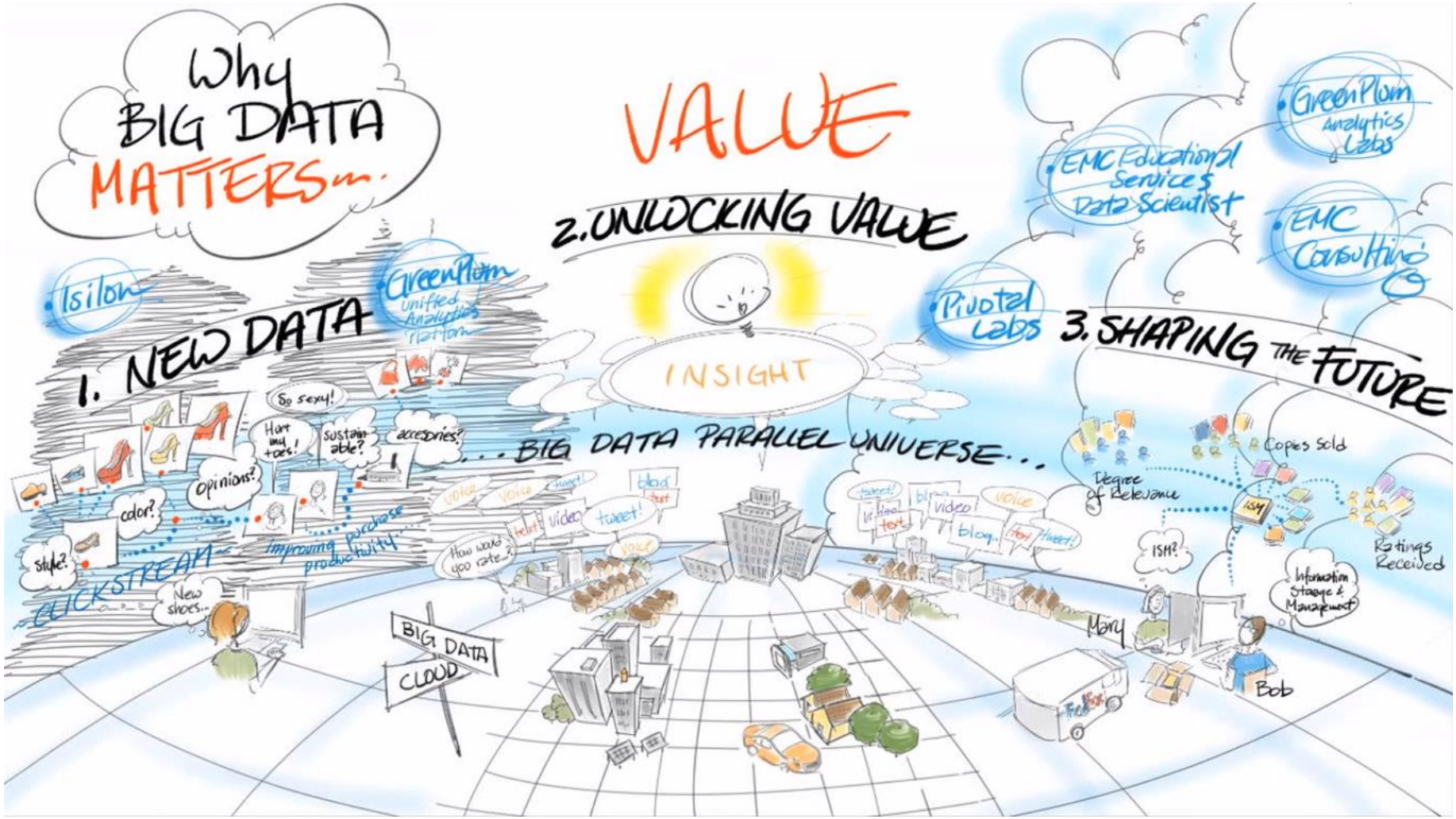
Source: http://visual.ly/what-big-data

Big Ideas: How Big is Big Data?



Source: http://www.youtube.com/watch?v=eEpxN0htRKI

Big Ideas: Why Big Data Matters



Source: http://www.youtube.com/watch?v=eEpxN0htRKI

Key questions enterprises are asking about Big Data

- How to store and protect big data?
- How to backup and restore big data?
- How to organize and catalog the data that you have backed up?
- How to keep costs low while ensuring that all the critical data is available when you need it?

Source: http://visual.ly/what-big-data

Volumes of Data

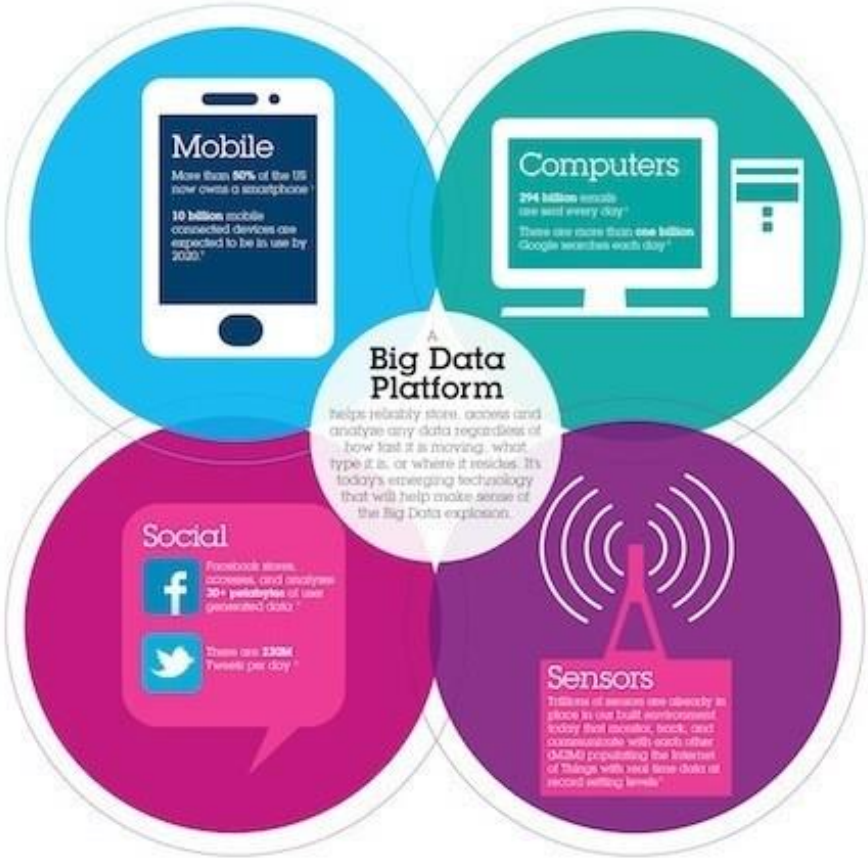
- Facebook
  - **30 billion pieces of content** were added to Facebook this past month by 600 million plus users
- Youtube
  - **More than 2 billion videos** were watch on YouTube yesterday
- Twitter
  - **32 billion searches** were performed last month on Twitter

Source: http://visual.ly/what-big-data

Big Data:

Making the World go Round

Big Data is growing and moving fast from a variety of sources; are you keeping up?



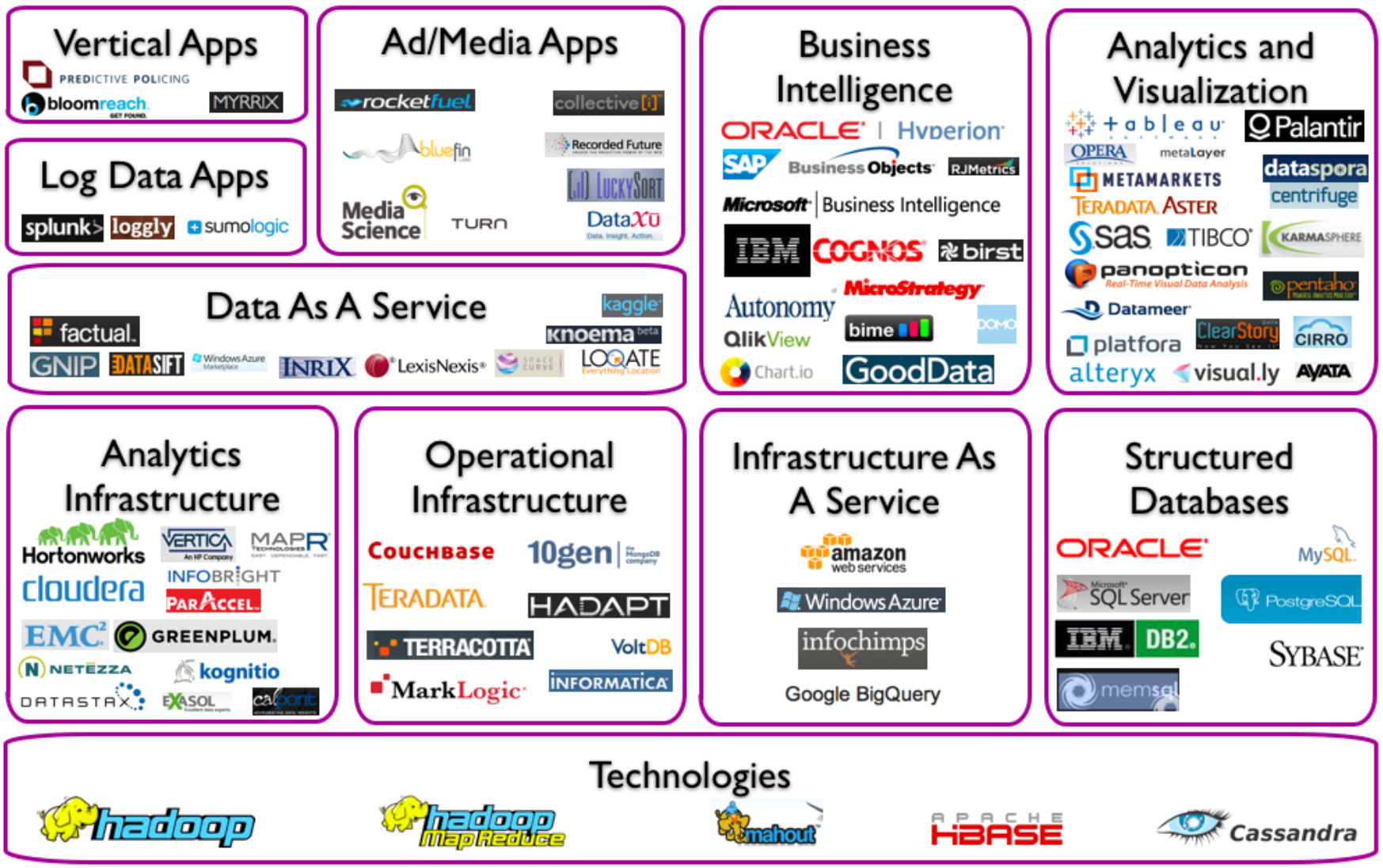
Information gathered by IBM:

1. Clarkin Sharma Consulting - US Mobile Data Market Update Q3 2012
2. 2011 Economist
3. IBM - Managing the Big Flood of Big Data in Digital Marketing
4. Google - How Google Search Works
5. Wikibon - Taming Big Data
6. IBM - Managing the Big Flood of Big Data in Digital Marketing
7. IBM

IBM

Source: http://www.business2community.com/big-data/big-data-big-insights-for-social-media-with-ibm-0501158

Big Data Landscape

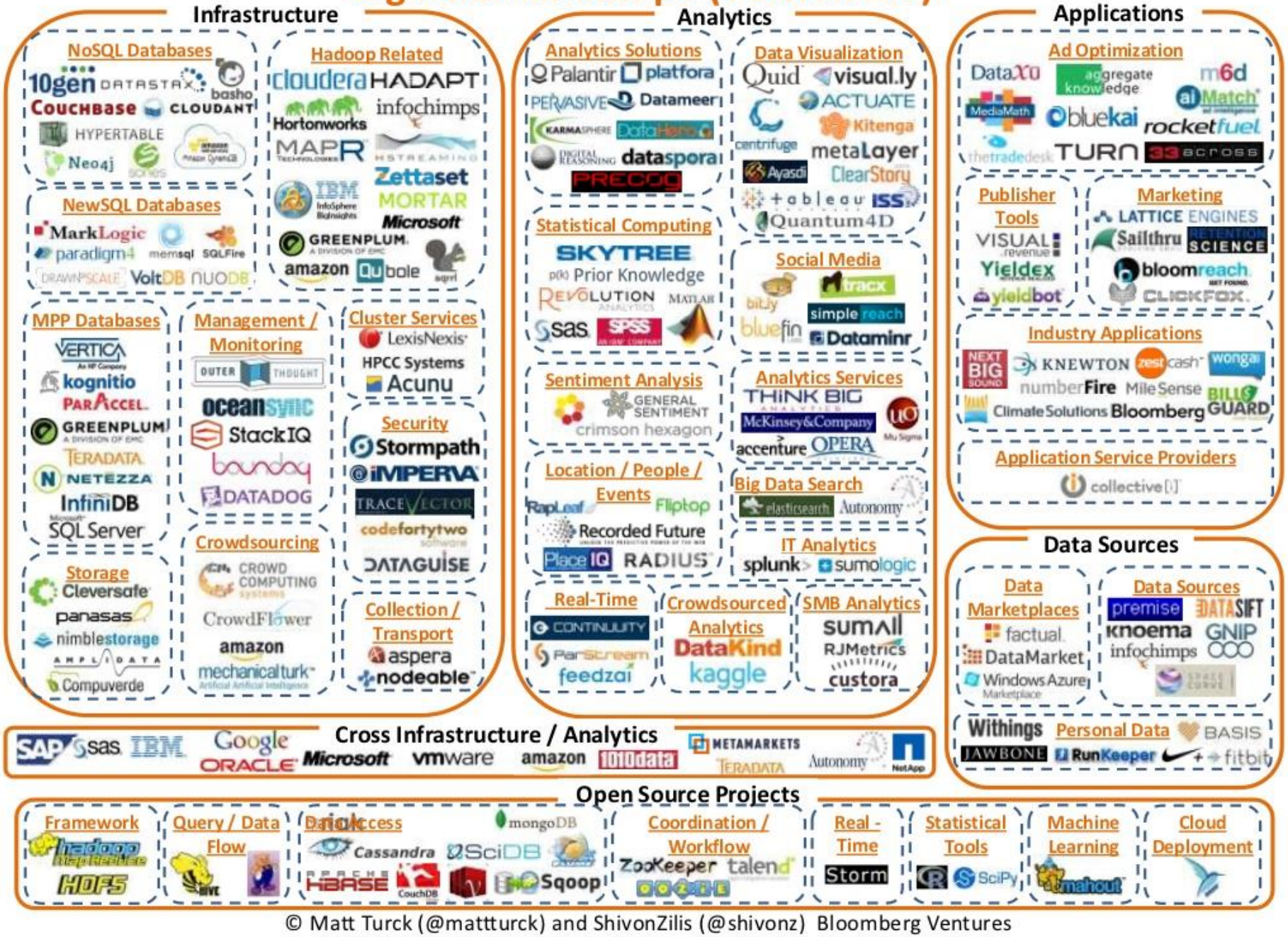


dave@vcdave.com

blogs.forbes.com/davefeinleib

Source: http://www.forbes.com/sites/davefeinleib/2012/06/19/the-big-data-landscape/

Big Data Landscape (Version 2.0)

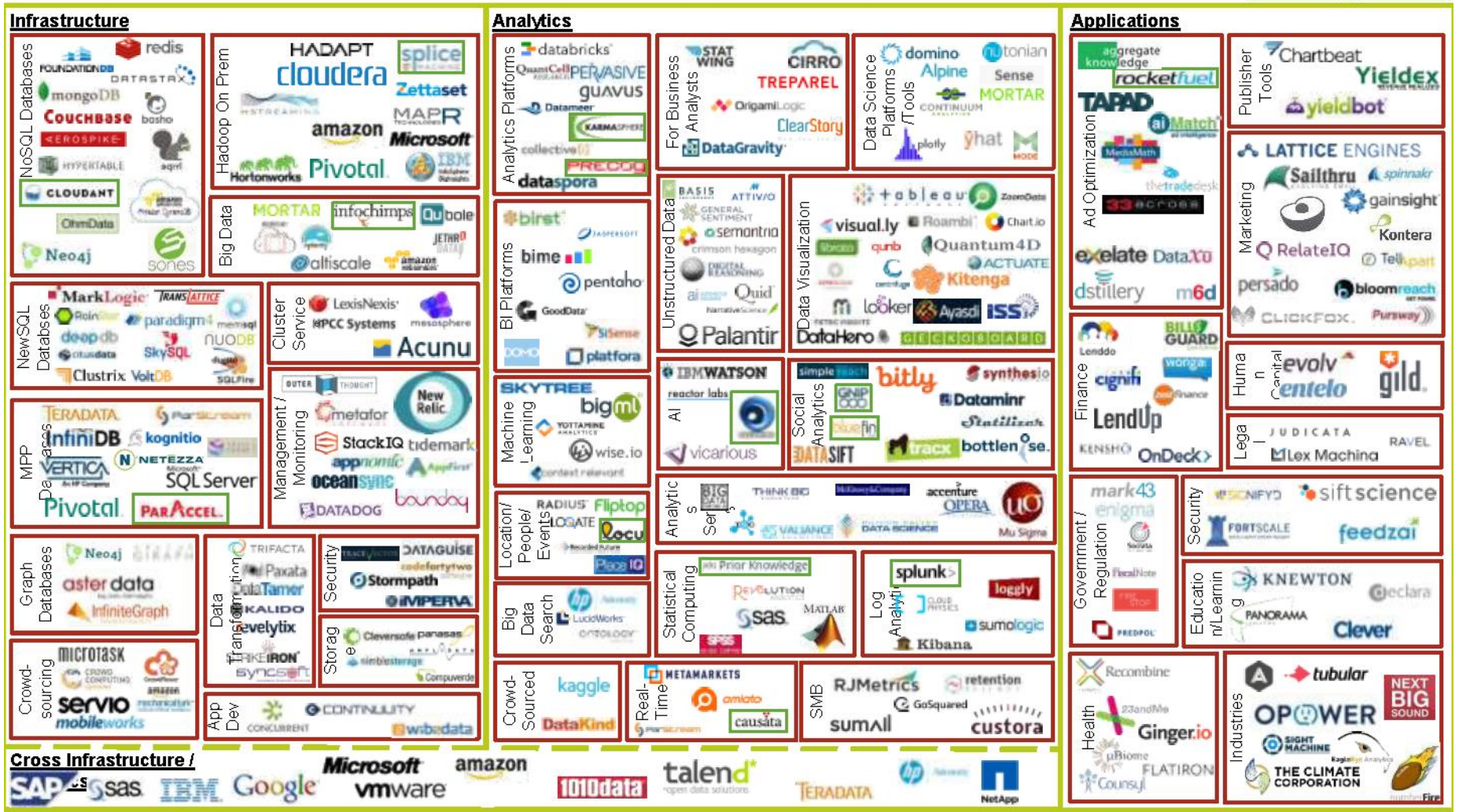


© Matt Turck (@mattturck) and ShivonZilis (@shivonz) Bloomberg Ventures

Source: http://mattturck.com/2012/10/15/a-chart-of-the-big-data-ecosystem-take-2/

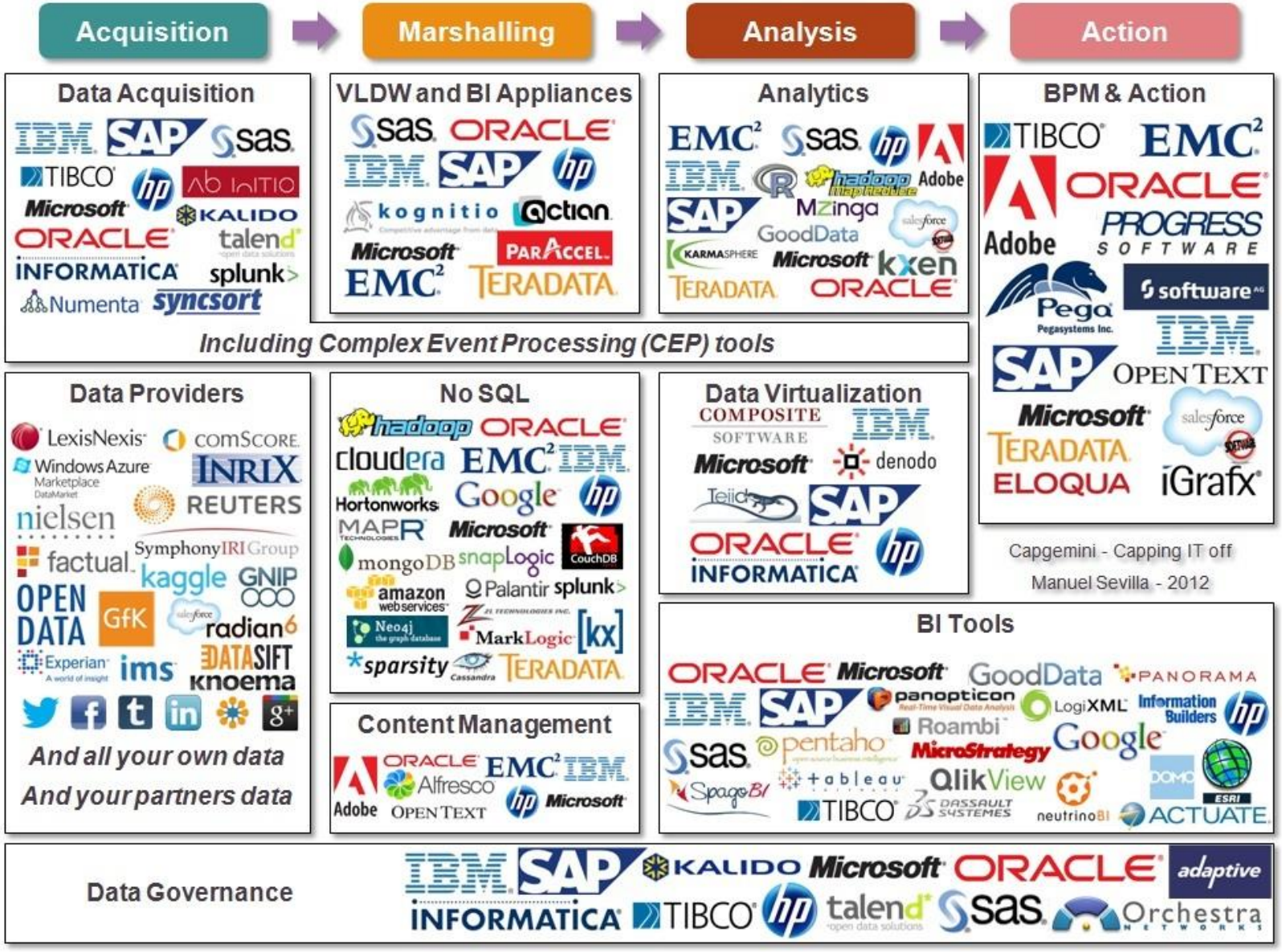
BIG DATA LANDSCAPE, VERSION 3.0

Exited: Acquisition or IPO



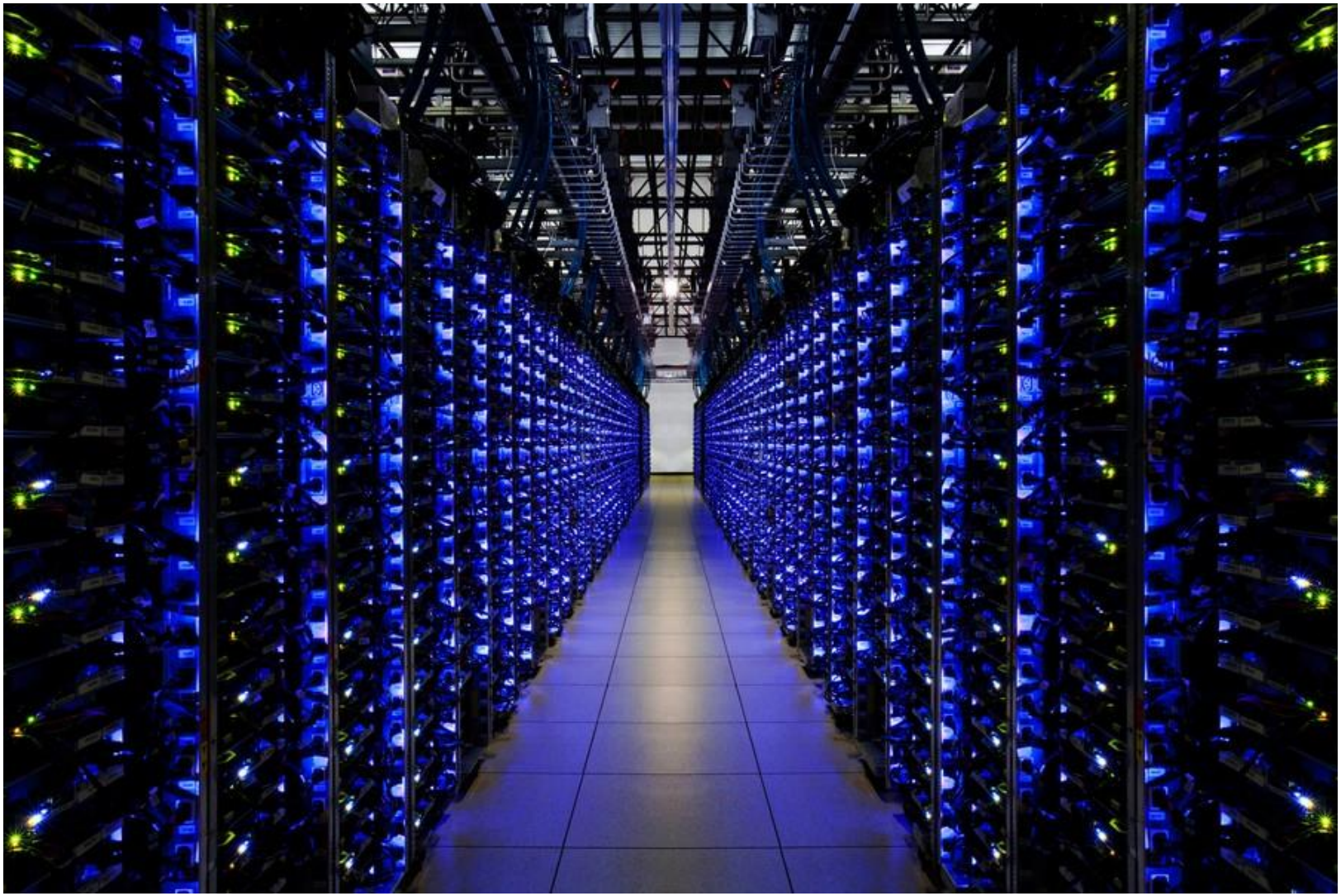
Source: http://www.slideshare.net/mjft01/big-data-landscape-matt-turck-may-2014

Big Data Vendors and Technologies



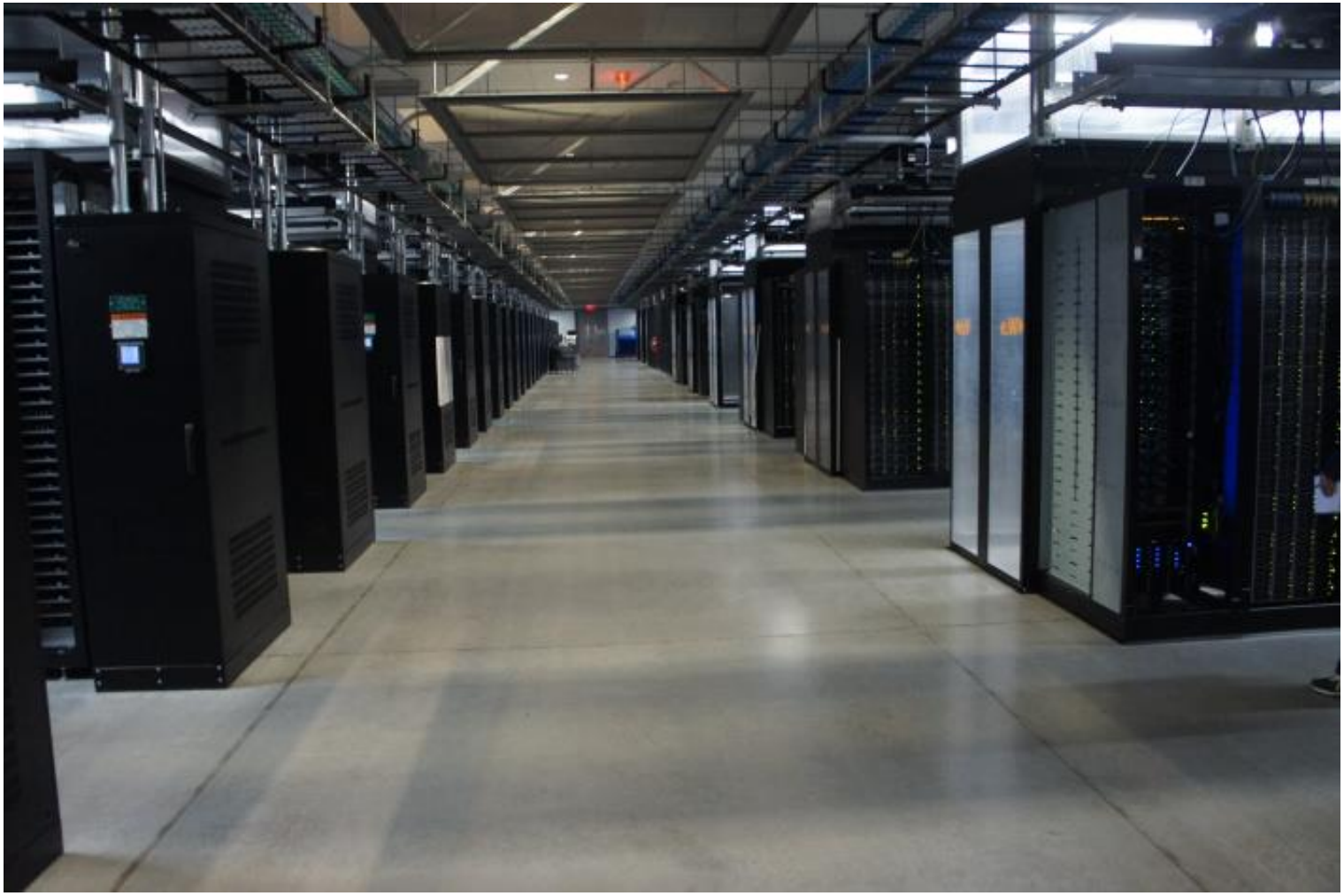
Source: http://www.capgemini.com/blog/capping-it-off/2012/09/big-data-vendors-and-technologies-the-list

Processing Big Data
Google



Source: http://whatsthebigdata.files.wordpress.com/2013/03/google_datacenter.jpg

Processing Big Data, Facebook



http://gigaom.com/2012/08/17/a-rare-look-inside-facebooks-oregon-data-center-photos-video/

Summary

- Business Intelligence Implementation
- Business Intelligence Trends
- Data Science
- Big Data Analytics
  - Big Data, Big Analytics:
    Emerging Business Intelligence and Analytic Trends for Today's Businesses

References

- Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence systems. Journal of computer information systems, 50(3), 23.
- Lim, E. P., Chen, H., & Chen, G. (2013). Business Intelligence and Analytics: Research Directions. *ACM Transactions on Management Information Systems (TMIS)*, *3*(4), 17
- McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard business review*.
- Davenport, T. H., & Patil, D. J. (2012). Data Scientist. *Harvard business review*.
- Provost, F., & Fawcett, T. (2013). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51-59.
- Dhar, V. (2013). Data science and prediction. Communications of the ACM, 56(12), 64-73.
- Thomas H. Davenport,
  Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data,
  FT Press, 2012
- Michael Minelli, Michele Chambers, Ambiga Dhiraj,
  Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses,
  Wiley, 2013
- Viktor Mayer-Schonberger, Kenneth Cukier,
  Big Data: A Revolution That Will Transform How We Live, Work, and Think,
  Eamon Dolan/Houghton Mifflin Harcourt, 2013