

# LỜI NÓI ĐẦU

<b>LỜI NÓI ĐẦU</b> .....	<b>iii</b>
<b>Chương 1: GIỚI THIỆU VỀ BIG DATA</b> .....	<b>1</b>
1.1. Khái niệm Big Data.....	1
1.2. Sự bùng nổ dữ liệu .....	3
1.3. Một số đặc điểm của Big Data.....	5
1.3.1. Dung lượng dữ liệu ( <i>Data volume</i> ).....	5
1.3.2. Tốc độ dữ liệu ( <i>Data velocity</i> ).....	5
1.3.3. Đa dạng dữ liệu ( <i>Data variety</i> ).....	7
1.4. Tầm quan trọng của Big Data trong thời đại ngày nay.....	8
1.5. Cơ hội và thách thức.....	9
<b>Chương 2: NHÂN TỐ THỨC ĐẨY SỰ PHÁT TRIỂN BIG DATA</b> .....	<b>12</b>
2.1. Sự phát triển của thị trường Big Data.....	12
2.2. Sự phát triển liên tục của các nội dung số .....	14
2.3. Sự gia tăng của Internet of Things (IoT) .....	15
2.4. Những sự khởi đầu mạnh mẽ của mã nguồn mở .....	16
2.5. Tăng cường đầu tư cho công nghệ Big Data.....	16
2.6. Nghiên cứu và phát triển liên quan đến tính toán hiệu năng cao ( <i>high-performance computing</i> ) .....	17
2.7. Bức tranh toàn cảnh về Big Data trong tương lai .....	18
2.8. Khoa học dữ liệu và sự gia tăng của các nhà khoa học dữ liệu .....	19
2.8.1. Nguồn dữ liệu và quản lý.....	20
2.8.2. Mô hình hóa phân tích.....	21
2.8.3. Phân tích kinh doanh.....	21
<b>Chương 3: ẢNH HƯỞNG CỦA BIG DATA ĐẾN THIẾT KẾ</b> .....	<b>23</b>

3.1. Big Data được sử dụng trong thiết kế Facebook.....	23
3.2. Apple định hướng thiết kế.....	26
3.3. Big Data trong thiết kế trò chơi.....	27
3.4. Big Data và ô tô.....	29
3.5. Big Data và âm nhạc.....	31
3.6. Big Data và kiến trúc.....	32
3.7. Thiết kế hướng dữ liệu.....	34
3.8. Big Data trong việc thiết kế website tốt hơn.....	35

#### **Chương 4: BIỂU DIỄN THÔNG TIN**

##### **BẢNG HÌNH ẢNH..... 39**

4.1. Xu hướng phân tích dữ liệu.....	39
4.2. Các loại trình bày bằng hình ảnh (visualizations).....	42
4.3. Xây dựng visualizations.....	46
4.4. Sử dụng báo cáo hình ảnh trong đúc kết tri thức.....	48
4.5. Tại sao thông tin dưới dạng hình ảnh mang đến nhiều lợi ích?.....	48
4.6. Hình ảnh và sức mạnh của sự chia sẻ.....	49
4.7. Các nguồn dữ liệu công bố công khai.....	50
4.8. Báo cáo ảnh động thời gian thực (real-time visualization).....	52
4.9. Tại sao hình ảnh dễ hiểu đối với người sử dụng nhưng khó khăn với máy tính?.....	54
4.10. Khía cạnh tâm lý và sinh lý học trong hình ảnh.....	54
4.11. Ảnh hưởng của hình ảnh.....	55

#### **Chương 5: BIG DATA ỨNG DỤNG TRONG**

##### **MARKETING..... 57**

5.1. Ứng dụng Big Data để tăng hiệu quả marketing.....	59
5.2. Big Data và Giám đốc marketing (CMOs).....	60

5.3. Hoạt động Big Data marketing trong thực tiễn.....	63
5.4. Tự động hóa marketing.....	64
5.5. Big Data – Công cụ tạo nội dung.....	67
5.6. Phương thức PR mới: Big Data và tiếp thị nội dung.....	69
5.7. Sử dụng nguồn lực cộng đồng (Crowdsourcing) .....	71
5.8. Mỗi công ty đều có thể là công ty truyền thông.....	73
5.9. Đo lường kết quả.....	74
<b>Chương 6: BIG DATA VÀ ỨNG DỤNG TRONG</b>	
<b>CUỘC SỐNG .....</b>	<b>76</b>
6.1. Các ứng dụng sức khỏe cá nhân.....	77
6.2. Big Data ứng dụng trong khám chữa bệnh.....	84
6.3. Big Data và việc điều trị bệnh.....	86
6.4. Big Data trong lĩnh vực năng lượng, giải trí và Smart City.....	91
6.5. Big Data trong lĩnh vực bán lẻ .....	93
6.6. Big Data trong lĩnh vực tài chính.....	94
6.7. Điện thoại thông minh trong tương lai .....	95
<b>Chương 7: BIG DATA VÀ CÔNG CỤ PHÂN TÍCH.....</b>	<b>97</b>
7.1. Giới thiệu .....	97
7.2. Phân tích cơ bản (Basic analytics).....	98
7.2.1. Đối chiếu và phân tích khối ( <i>Sliding and dicing</i> ) ....	98
7.2.2. Kiểm soát cơ bản ( <i>Basic Monitoring</i> ) .....	98
7.2.3. Nhận diện vấn đề ( <i>Anomaly identification</i> ).....	99
7.3. Phân tích nâng cao (Advanced analytics) .....	99
7.3.1. Mô hình dự báo ( <i>Predictive modeling</i> ).....	100
7.3.2. Phân tích văn bản ( <i>Text analytics</i> ).....	100
7.3.3. Thống kê và thuật toán khai phá dữ liệu ( <i>Statistical and data-mining algorithms</i> ).....	100

7.4. Phân tích hoạt động (Operationalized analytics).....	101
7.5. Các tình huống trong phân tích Big Data.....	101
7.5.1. <i>Orbitz</i> .....	101
7.5.2. <i>Nokia</i> .....	102
7.5.3. <i>NASA</i> .....	102
7.6. Giải pháp phân tích Big Data.....	102
7.6.1. <i>IBM</i> .....	103
7.6.2. <i>SAS</i> .....	103
7.6.3. <i>Oracle</i> .....	103
7.6.4. <i>Pentaho</i> .....	103
7.7. Phân tích văn bản và phân tích dữ liệu phi cấu trúc.....	104
7.7.1. <i>Khảo sát dữ liệu phi cấu trúc</i> .....	104
7.7.2. <i>Phân tích văn bản</i> .....	105
7.8. Kỹ thuật phân tích và giải mã.....	106
7.9. Các kỹ thuật phân tích khác (Other analytics).....	108
7.9.1. <i>Phân tích dự báo (Predictive analytics)</i> .....	108
7.9.2. <i>Phân tích đồ thị (Graph Analytics)</i> .....	111
7.9.3. <i>Quản lý dữ liệu tổng thể (Master Data Management - MDM)</i> .....	113
7.9.4. <i>Xử lý sự kiện phức tạp (Complex event processing)</i> .....	117
7.9.5. <i>Phân tích kinh doanh di động (Mobile Business Analytics)</i> .....	119
7.9.6. <i>Phân tích âm thanh và phân tích nội dung hình ảnh (Audio analytics and video content analytics)</i> .....	122
7.9.7. <i>Liên hợp dữ liệu (Data Federation)</i> .....	125

7.10. Mô hình trong phân tích dữ liệu Big Data ..... 125

7.11. Một số vấn đề tồn tại của Big Data ..... 127

**Chương 8: NỀN TẢNG CÔNG NGHỆ CHO  
BIG DATA ..... 129**

8.1. Cơ sở dữ liệu (CSDL) quan hệ SQL và cơ sở dữ  
liệu NoSQL..... 129

    8.1.1. Cơ sở dữ liệu quan hệ SQL ..... 129

    8.1.2. Cơ sở dữ liệu NoSQL..... 130

    8.1.3. Một số đặc điểm của CSDL NoSQL ..... 131

    8.1.4. Ứng dụng CSDL NoSQL ..... 132

8.2. Cơ sở dữ liệu quan hệ và dữ liệu lớn..... 134

    8.2.1. Cơ sở dữ liệu quan hệ ..... 134

    8.2.2. Cơ sở dữ liệu hướng không quan hệ  
(Nonrelational Databases) ..... 136

    8.2.3. CSDL hướng cặp khóa-giá trị (Key-Value Pair  
Databases) ..... 137

    8.2.4. Cơ sở dữ liệu hướng văn bản (Document  
Databases) ..... 137

    8.2.5. Cơ sở dữ liệu hướng đồ thị (Graph Databases).... 137

8.3. Nền tảng MapReduce..... 139

8.4. Khám phá thế giới của Hadoop và sơ lược về cài  
đặt hệ thống ..... 141

    8.4.1. Khái niệm Hadoop..... 141

    8.4.2. Hiểu về hệ thống tập tin phân tán Hadoop ..... 144

    8.4.3. Xây dựng Big Data dựa trên nền tảng Hadoop  
và MapReduce ..... 145

    8.4.4. Sơ lược về cài đặt Hadoop ..... 145

    8.4.5. Cài đặt VMware Player..... 146

8.4.6. Thiết lập môi trường ảo hóa.....	146
8.4.7. Cài đặt Plugin cho Hadoop MapReduce.....	150
8.5. Một số thao tác cơ bản trong HDFS và MapReduce Eclipse.....	154
8.6. Tương tác với HDFS .....	159
8.7. Sự so sánh giữa Oracle NoSQL DataBase, HBase và MongoDB.....	166
8.8. Kết luận chương.....	169
<b>Chương 9: CÁC GIẢI PHÁP BIG DATA TRONG THỰC TẾ.....</b>	<b>170</b>
9.1. Tác động kinh tế của Big Data .....	170
9.1.1. Xác định các loại dữ liệu và các nguồn dữ liệu.....	170
9.1.2. Hiệu chỉnh về quy trình nghiệp vụ hay tạo ra quy trình mới.....	174
9.1.3. Ảnh hưởng kỹ thuật của các quy trình làm việc Big Data.....	174
9.1.4. Tìm nhân tài để hỗ trợ các dự án Big Data.....	175
9.1.5. Tính toán ROI từ các đầu tư Big Data .....	176
9.2. Việc quản lý dữ liệu doanh nghiệp và Big Data.....	177
9.2.1. Định nghĩa quản lý dữ liệu doanh nghiệp.....	177
9.3. Tạo một lộ trình triển khai Big Data .....	178
9.3.1. Hiểu được tính cấp thiết của kinh doanh (business urgency).....	178
9.3.2. Dự đoán đúng mức công suất .....	179
9.3.3. Lựa chọn phương pháp phát triển phần mềm.....	179
9.3.4. Tập hợp các kỹ năng và cân bằng ngân sách .....	180
9.4. Khởi động lộ trình Big Data của doanh nghiệp.....	180
9.5. Big Data như một công cụ lên kế hoạch kinh doanh ....	183

9.5.1. Giai đoạn 1: Lập kế hoạch với dữ liệu .....	184
9.5.2. Giai đoạn 2: Thực hiện các phân tích .....	185
9.5.3. Giai đoạn 3: Kiểm tra các kết quả.....	185
9.5.4. Giai đoạn 4: Điều chỉnh dựa trên kế hoạch.....	186
9.6. Thêm vào các chiều mới cho chu trình lên kế hoạch ....	186
9.6.1. Giai đoạn 5: Kiểm soát trong thời gian thực .....	187
9.6.2. Giai đoạn 6: Điều chỉnh các tác động.....	187
9.6.3. Giai đoạn 7: Cho phép thử nghiệm .....	187
9.7. Bảo mật các phân tích dữ liệu trong tương lai.....	188
9.8. Bắt đầu chiến lược Big Data.....	188
9.9. Lên kế hoạch cho Big Data.....	190
9.10. Chuyển đổi các quy trình kinh doanh với Big Data....	190
<b>DANH MỤC TỪ VIẾT TẮT .....</b>	<b>191</b>
<b>GIẢI THÍCH THUẬT NGỮ .....</b>	<b>193</b>
<b>TÀI LIỆU THAM KHẢO.....</b>	<b>194</b>