

ระบบสารสนเทศเพื่อการจัดการ ธุรกิจการค้าสมัยใหม่

INFORMATION SYSTEM FOR MODERN TRADE BUSINESS
Digital Transform for Modern Trade Business



ประจักษ์ เจ็ดโงม
คณะเศรษฐศาสตร์และบริหารธุรกิจ
มหาวิทยาลัยทักษิณ

คำนำ

ในยุคสมัยใหม่นี้ ระบบสารสนเทศสำหรับการบริหารธุรกิจการค้ามีบทบาทสำคัญในการเสริมสร้างประสิทธิภาพและความสามารถในการปรับตัวของธุรกิจ ระบบสารสนเทศส่งผลให้ธุรกิจสามารถบริหารจัดการและวางแผนดำเนินการได้อย่างมีประสิทธิภาพ รวมถึงการช่วยในการแก้ไขปัญหาต่าง ๆ อย่างรวดเร็ว ระบบสารสนเทศนี้มาพร้อมกับฟังก์ชันที่หลากหลายเพื่อช่วยในการจัดการข้อมูลที่สำคัญสำหรับธุรกิจ เช่น ข้อมูลลูกค้า ข้อมูลสินค้า และข้อมูลการขาย แนวทางการเรียนรู้เกี่ยวกับระบบสารสนเทศและสถาปัตยกรรมของระบบสารสนเทศสมัยใหม่จะช่วยให้นักศึกษาและผู้อ่านเข้าใจและมีความรู้เกี่ยวกับการจัดการข้อมูลและสารสนเทศในองค์กร ที่จะทำให้พวกเขาสามารถจัดการข้อมูลได้อย่างมีประสิทธิภาพ นอกจากนี้การศึกษาเกี่ยวกับระบบสารสนเทศยังช่วยให้นักศึกษาเข้าใจการบริหารฐานข้อมูลและการใช้ระบบสารสนเทศในการวิเคราะห์ข้อมูลต่าง ๆ ขององค์กร ตำราฉบับนี้มีเนื้อหาครอบคลุม สำหรับการเตรียมพัฒนาระบบสารสนเทศเพื่อการจัดการธุรกิจการค้าสมัยใหม่ ผู้เขียนคาดหวังว่าตำราจะระบบสารสนเทศเพื่อการจัดการธุรกิจการค้าสมัยใหม่ ฉบับนี้ จะช่วยให้นักศึกษาและผู้อ่านมีความเข้าใจการพัฒนาและออกแบบระบบสารสนเทศเพื่อการจัดการธุรกิจการค้าสมัยใหม่ และสามารถนำไปใช้งานได้จริงในภาคหน้า หากพบข้อผิดพลาดในฉบับตำรา หรือต้องการให้ข้อเสนอแนะใด ๆ สามารถแจ้งข้อคิดเห็นหรือข้อเสนอแนะได้ที่ผู้เขียน เพื่อให้ผู้เขียนได้ปรับปรุงแก้ไขตำราให้ดียิ่งขึ้นต่อไป

ประจักษ์ เด็ดโฉม

ผู้เขียนและผู้เรียบเรียง

สารบัญ

	หน้า
สารบัญภาพ	5
สารบัญตาราง.....	8
บทที่ 1 ระบบสารสนเทศเพื่อการจัดการธุรกิจการค้าสมัยใหม่.....	9
(Information System for Modern Business Management)	9
1.1 นิยามเกี่ยวกับธุรกิจการค้าสมัยใหม่.....	9
1.2 ประเภทของระบบสารสนเทศ (Type of Information Systems)	11
1.3 ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูล	20
คำศัพท์ บทที่ 1	23
คำถามและแบบฝึกหัดท้ายบทที่ 1	25
บทที่ 2 การพัฒนาและวิเคราะห์ระบบสารสนเทศสมัยใหม่.....	27
(Modern Information System Development)	27
2.1 ระบบสารสนเทศพื้นฐาน	27
2.2 วงจรการวิเคราะห์ระบบงาน	30
2.3 วงจรการพัฒนากระบวนการสำหรับระบบสารสนเทศ.....	31
2.4 รายละเอียดวงจรการพัฒนากระบวนการ	32
2.5 แนวทางที่ใช้ในการวิเคราะห์และออกแบบระบบสารสนเทศสมัยใหม่.....	36
คำศัพท์ บทที่ 2	42
คำถามและแบบฝึกหัดท้ายบทที่ 2	43
บทที่ 3 ความรู้เบื้องต้นเกี่ยวกับกรอบแนวทางการวางแผนสถาปัตยกรรมระบบ	44
(Introduction to Modern Information System Architecture)	44
3.1 สถาปัตยกรรมองค์กร (Enterprise Architecture: EA) คืออะไร	44
3.2 ประโยชน์ของการใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture: EA)	45
3.3 ตัวอย่างองค์ประกอบและเฟรมเวิร์ค (Framework) ของสถาปัตยกรรมองค์กรสำหรับระบบสารสนเทศสมัยใหม่.....	47
3.4 สรุปประโยชน์ของสถาปัตยกรรมองค์กร TOGAF.....	55
คำศัพท์ บทที่ 3	57
คำถามและแบบฝึกหัดท้ายบทที่ 3	58
บทที่ 4 แนวทางการออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศสมัยใหม่	59
(Guidelines for new Information System Architecture Design)	59

4.1 บทนำแนวทางการออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศ	59
4.2 โมเดลที่ใช้ในการวิเคราะห์และออกแบบระบบ.....	62
4.3 ตัวอย่างเครื่องมือแบบจำลอง	64
คำศัพท์ บทที่ 4	82
คำถามและแบบฝึกหัดท้ายบทที่ 4	83
บทที่ 5 โมเดลการพัฒนาและมาตรฐานสำหรับการสร้างระบบสารสนเทศสมัยใหม่.....	84
(Developing Model and Standard for New Information System Development)	84
5.1 บทนำการสร้างระบบสารสนเทศสมัยใหม่	84
5.2 โมเดลขั้นตอนการพัฒนาซอฟต์แวร์ (Software Development Process Models).....	87
5.3 มาตรฐาน COBIT: Control Objectives for Information and Related Technology (Amorim et al., 2021)	95
คำศัพท์ บทที่ 5	104
คำถามและแบบฝึกหัดท้ายบทที่ 5.....	105
บทที่ 6 การประยุกต์ใช้สถาปัตยกรรมองค์กรเพื่อพัฒนาระบบข้อมูลขนาดใหญ่.....	107
(Applying the Enterprise Architecture for Big Data System Development)	107
6.1 สถาปัตยกรรมองค์กรเพื่อพัฒนาระบบข้อมูลขนาดใหญ่ (Enterprise Architecture for Big Data System).....	107
6.2 สถาปัตยกรรมของข้อมูลขนาดใหญ่ (Architecture of Big Data).....	110
6.3 ประโยชน์ของสถาปัตยกรรมองค์กรสำหรับระบบข้อมูลขนาดใหญ่	118
คำศัพท์ บทที่ 6	121
คำถามและแบบฝึกหัดท้ายบทที่ 6	123
บทที่ 7 กระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่	124
(The Process of Developing Modern Corporate Architecture and Information Systems).....	124
7.1 บทนำกระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่	124
7.2 แนวทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่.....	127
7.3 องค์ประกอบของทีมงานพัฒนาระบบในองค์กรปัจจุบัน	135
7.4 บทบาทของนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile.....	137
คำศัพท์ บทที่ 7	139
คำถามและแบบฝึกหัดท้ายบทที่ 7	140
บทที่ 8 ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูลในธุรกิจการค้าสมัยใหม่	141
(Introduction to Data Mining for Modern Trading Business)	141
8.1 ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูล	141

8.2 นิยามและความหมายนักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist) ..	147
8.3 กรณีศึกษาเกี่ยวกับนักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist).....	150
8.4 ความรู้เบื้องต้นเกี่ยวกับ Data Preprocessing	151
8.5 กรณีศึกษาเกี่ยวกับ Data Preprocessing.....	152
คำศัพท์ บทที่ 8	156
คำถามและแบบฝึกหัดท้ายบทที่ 8	157
บทที่ 9 เทคนิคและ อัลกอริทึม สำหรับการทำเหมืองข้อมูล	159
(Techniques and Algorithms for Data Mining)	159
9.1 ความรู้เบื้องต้นเกี่ยวกับเทคนิคและ อัลกอริทึม สำหรับการทำเหมืองข้อมูล	159
9.2 ความรู้เบื้องต้นเกี่ยวกับเทคนิคการจำแนกประเภท (Classification)	162
9.3 ความรู้เบื้องต้นเกี่ยวกับเทคนิคการแบ่งกลุ่ม (Clustering).....	168
9.4 ความรู้เบื้องต้นเกี่ยวกับเทคนิคกฎความสัมพันธ์ (Association Rules).....	173
9.5 ความรู้เบื้องต้นเกี่ยวกับเทคนิคระบบแนะนำ (Recommendation System).....	179
9.6 ความรู้เบื้องต้นเกี่ยวกับเทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network: ANN).....	186
คำศัพท์ บทที่ 9	190
คำถามและแบบฝึกหัดท้ายบทที่ 9	191
บทที่ 10 เครื่องมือและเทคนิคที่สำคัญสำหรับการทำเหมืองข้อมูล	193
(Important Tools and Techniques for Data Mining)	193
10.1 การประยุกต์ใช้โมเดลกระบวนการวิเคราะห์ข้อมูล (The CRISP-DM model).....	195
10.2 การทำความเข้าใจข้อมูล (Data Understanding) การเตรียมข้อมูล (Data Preparation).....	196
10.3 ความท้าทายในการเตรียมข้อมูล (Data Preparation).....	200
10.4 การนำไปใช้ (Consumption).....	201
10.5 แนวทางในการเล่าเรื่องด้วยข้อมูล (Data Storytelling).....	212
คำศัพท์ บทที่ 10	215
คำถามและแบบฝึกหัดท้ายบทที่ 10.....	216
บรรณานุกรม.....	217
ภาคผนวก	223
ดัชนี	258

สารบัญภาพ

รูปที่ 1. 1 ประเภทของระบบสารสนเทศจำแนกตามระดับชั้นองค์กร	17
รูปที่ 1. 2 ส่วนผสมของข้อมูลในแต่ละระดับชั้นองค์กร.....	17
รูปที่ 1. 3 Power BI dashboard and Report Example (ภาพจากการเรียนการสอนของผู้เขียน).....	19
รูปที่ 1. 4 แผนภาพแสดงกระบวนการทำเหมืองข้อมูล (Data Mining Process) ปรับปรุงจาก (IBM, 2011).....	21
รูปที่ 1. 1 ประเภทของระบบสารสนเทศจำแนกตามระดับชั้นองค์กร	18
รูปที่ 1. 2 ส่วนผสมของข้อมูลในแต่ละระดับชั้นองค์กร.....	18
รูปที่ 1. 3 Power BI dashboard and Report Example (ภาพจากการเรียนการสอนของผู้เขียน).....	20
รูปที่ 1. 4 แผนภาพแสดงกระบวนการทำเหมืองข้อมูล (Data Mining Process) ปรับปรุงจาก (IBM, 2011).....	22
รูปที่ 2.1 องค์ประกอบของระบบสารสนเทศพื้นฐาน.....	29
รูปที่ 2.2 วงจรการพัฒนากระบวน SDLC.....	32
รูปที่ 2.3 ขั้นตอนการวิเคราะห์ความต้องการของระบบ.....	34
รูปที่ 2.4 องค์ประกอบระบบที่ต้องออกแบบ (Systems' Components for Designing).....	35
รูปที่ 2.5 ระบบฐานข้อมูลขนาดใหญ่ของระบบบันทึกข้อมูลสุขภาพผู้ป่วย.....	38
รูปที่ 2.6 รูปเทคโนโลยีที่สำคัญ สำหรับการออกแบบสารสนเทศสมัยใหม่.....	39
รูปที่ 3.1 กรอบแนวคิดและหลักการในการพัฒนาสถาปัตยกรรมองค์กรสากล	47
รูปที่ 3.2 ประวัติโดยย่อของสถาปัตยกรรมองค์กร.....	50
รูปที่ 3.3 องค์ประกอบของสถาปัตยกรรมองค์กร TOGAF.....	51
รูปที่ 3.4 แนวทางการใช้งานตามสถาปัตยกรรมองค์กร TOGAF.....	53
รูปที่ 3.5 ขั้นตอนการออกแบบสถาปัตยกรรมองค์กรตามแนวทาง TOGAF.....	54
รูปที่ 4.1 ปัจจัยเพื่อพิจารณาสำหรับการบริหารโครงการสารสนเทศสมัยใหม่.....	60
รูปที่ 4.2 กรอบการบริหารโครงการสารสนเทศสมัยใหม่.....	61
รูปที่ 4.3 ผังงานเรียงลำดับ (Sequence).....	66
รูปที่ 4.4 ผังงานแบบมีเงื่อนไข SELECTION.....	67
รูปที่ 4.5 ผังงานแบบทำซ้ำ ITERATION.....	68
รูปที่ 4.6 ตัวอย่างผังงานพื้นฐานทั่วไป (ไม่มีการใช้ IT).....	68
รูปที่ 4.7 สัญลักษณ์ของ (Data Flow Diagram: DFD).....	69
รูปที่ 4.8 ระดับของ แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level)	70

รูปที่ 4.9 ตัวอย่างแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level 0) ของระบบทะเบียน นักศึกษา	70
รูปที่ 4.10 ตัวอย่างแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level 1) ของระบบทะเบียน นักศึกษา	71
รูปที่ 4.11 ตัวอย่างแผนภาพกระแสข้อมูลสัญลักษณ์ของ ER-Diagram	72
รูปที่ 4.12 ตัวอย่างแผนภาพ E-R Diagram	73
รูปที่ 4.13 ตัวอย่างแผนภาพ E-R Diagram ระบบ e-Learning	75
รูปที่ 4.14 ตัวอย่างแผนภาพ Class diagram	79
รูปที่ 4.15 ตัวอย่างแผนภาพ Class diagram จำลองความสัมพันธ์.....	80
รูปที่ 4.16 ตัวอย่างโมเดลแบบจำลองเชิงวัตถุ (Modelling object).....	80
รูปที่ 5.1 องค์ประกอบของระบบสารสนเทศ.....	85
รูปที่ 5.2 โมเดลน้ำตก (Waterfall Model) (Laudon & Laudon, 2013).....	88
รูปที่ 5.3 แบบ Iterative Model & Incremental Model (Laudon & Laudon, 2013).....	89
รูปที่ 5.4 แบบ Evolutionary (Laudon & Laudon, 2013).....	90
รูปที่ 5.5 โมเดลแบบ Incremental Model (Laudon & Laudon, 2013).....	91
รูปที่ 5.6 กระบวนการทำงานของ “Scrum” (ผู้เขียน).....	93
รูป 6.1 แบบจำลองสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ (DGA, 2022).....	108
รูปที่ 7.1 ความคาดหวังกับกระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่.....	125
รูปที่ 7.2 แนวทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่.....	128
รูปที่ 7.3 มุมการวิเคราะห์ขอบเขตของสถาปัตยกรรมองค์กร (Scope).....	129
รูปที่ 7.4 มุมการศึกษาความเป็นไปได้ (Feasibility Study).....	130
รูปที่ 7.5 ผังโครงสร้างองค์กรส่วนงานสารสนเทศยุคใหม่.....	135
รูปที่ 7.6 องค์ประกอบทีมงานสารสนเทศยุคใหม่.....	136
รูปที่ 8.1 ภาพแสดงกระบวนการทำเหมืองข้อมูล (Data Mining).....	143
รูปที่ 8.2 กรณีศึกษา “เดอะมอลล์” เจาะพฤติกรรมลูกค้าด้วย Data Mining ทำให้ลูกค้าประทับใจ	

อยู่นานมากขึ้น.....	146
รูปที่ 8.3 กรณีศึกษาการออกแบบการจัดเก็บข้อมูลเพื่อการวิเคราะห์พฤติกรรมผู้บริโภคบน	
นาฬิกาอัจฉริยะ (Wearable Smart Watch).....	147
รูปที่ 8.4 กรณีศึกษากระบวนการจัดเตรียมข้อมูลผู้ป่วยโรคหัวใจล้มเหลว (Congestive Heart Failure : CHF).....	153
รูปที่ 8.5 กรณีศึกษากระบวนการจัดเตรียมข้อมูลชีวิตประจำวัน “LIFELOG”	154
รูปที่ 9.1 กรณีศึกษาเทคนิคการจำแนกประเภท (Classification) แบบ Decision Tree เพื่อจำแนกผู้มีสุขภาพดี.....	167
รูปที่ 9.2 กรณีศึกษาเทคนิคการจำแนกประเภท (Classification) แบบ Decision Tree เพื่อจำแนกหาเหตุการณ์เลือกเรียนออนไลน์.....	167
รูปที่ 9.3 วิธีการแบบลำดับชั้น (Hierarchical Algorithms).....	170
รูปที่ 9.4 วิธีการแบบไม่เป็นลำดับชั้น (Non-Hierarchical Algorithms) หรือ Partitional clustering.....	171
รูปที่ 9.5 กรณีศึกษาเทคนิคการจัดกลุ่ม (Clustering) เพื่อแบ่งกลุ่มลูกค้าและประเมิน Credit Score ของบริษัท Ricult.....	172
รูปที่ 9.6 Explicit Rating และ Implicit Rating.....	181
รูปที่ 9.7 การทำงานของระบบแนะนำ (Recommendation System).....	182
รูปที่ 9.8 อัลกอริทึม (Algorithm) ในการคัดกรองข้อมูลของ Recommendation System.....	183
รูปที่ 9.9 Content-Based Filtering (CBF).....	184
รูปที่ 9.10 Collaborative Filtering (CF).....	185
รูปที่ 9.11 องค์ประกอบของโครงข่ายประสาทเทียม (Artificial Neural Network).....	187
รูปที่ 9.12 โครงข่ายประสาทเทียม (Artificial Neural Network) แบบป้อนไปข้างหน้า (Feed Forward).....	188
รูปที่ 9.13 โครงข่ายประสาทเทียม (Artificial Neural Network) แบบป้อนไปข้างหน้า	

(Feed Forward).....	188
รูปที่ 10.1 โมเดลกระบวนการวิเคราะห์ข้อมูล (The CRISP-DM model).....	194
รูปที่ 10.2 มุมมองสถาปัตยกรรมฐานข้อมูลขนาดใหญ่.....	197
รูปที่ 10.3 4 เทคนิคการแปลงข้อมูล (Data Transformation).....	198
รูปที่ 10.4 Confusion Matrix.....	203
รูปที่ 10.5 ROC Curve.....	204
รูปที่ 10.6 ค่าความผิดพลาด (Error) ที่ใช้วัดประเมินผลโมเดล (Model Evaluation).....	205
รูปที่ 10.7 ส่วนประกอบของ Power BI.....	207
รูปที่ 10.8 ห้องการเรียนรู้การสอนด้วย Tableau ด้านการวิเคราะห์และ	
นำเสนอข้อมูลของผู้เขียน.....	208
รูปที่ 10.9 ประเภทของสิทธิ์การใช้งาน Tableau 3 ประเภท.....	208

สารบัญตาราง

ตารางที่ 2.1 ประเด็นภาระงานสำหรับนักวิเคราะห์และออกแบบระบบ.....	34
ตารางที่ 4.1 ตัวอย่างสัญลักษณ์ของผังงาน (Flow Chart).....	65
ตารางที่ 5.1 มาตรฐาน COBIT และวัตถุประสงค์ย่อย.....	96
ตารางที่ 6.1 เครื่องมือสำหรับการจัดการข้อมูลขนาดใหญ่.....	116
ตารางที่ 8.1 กระบวนการทำเหมืองข้อมูล “CRISP-DM” Cross-Industry Standard Process for Data Mining.....	144
ตารางที่ 8.2 การเปรียบเทียบความแตกต่างระหว่าง นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist).....	149
ตารางที่ 9.1 การเปรียบเทียบเทคนิคการจำแนกประเภท (Classification) ขั้นพื้นฐาน.....	163
ตารางที่ 9.2 ชุดข้อมูลคุณลักษณะสำหรับการตัดสินใจเล่นเบสบอล.....	164
ตารางที่ 9.3 สูตรพื้นฐานสำหรับการเรียนรู้ต้นไม้การตัดสินใจ.....	165
ตารางที่ 9.4 การแบ่งกลุ่มวิเคราะห์ การทำนายราคาและจำนวนวัน และความนิยมของกรู๊ปทัวร์ ไปญี่ปุ่นของนักท่องเที่ยวไทย.....	173

บทที่ 1 ระบบสารสนเทศเพื่อการจัดการธุรกิจการค้าสมัยใหม่

(Information System for Modern Business Management)

ในปัจจุบันธุรกิจการค้าสมัยใหม่ได้เปลี่ยนแปลงไปหลายมิติเช่น ลักษณะของธุรกิจที่มีแพลตฟอร์ม (Platform) ทั้งบนโลกอินเทอร์เน็ตและการค้าแบบดั้งเดิม พฤติกรรมผู้บริโภค ลักษณะของตลาด สภาพการแข่งขัน และสมรรถนะทางเทคโนโลยีขององค์กร ดังนั้นเพื่อเตรียมความพร้อมผ่านการพัฒนาสารสนเทศและเทคโนโลยีที่สามารถตอบสนองความต้องการที่หลากหลาย และสร้างประสบการณ์ที่ดีแก่ผู้บริโภคได้ จำเป็นที่ต้องเข้าใจกระบวนการทำงานในปัจจุบันและสามารถระบุจุดที่ควรปรับปรุง เพื่อให้ธุรกิจสามารถดำเนินต่อไปได้

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจการเปลี่ยนแปลงของธุรกิจการค้าสมัยใหม่ และความสำคัญของระบบสารสนเทศในการรองรับความเปลี่ยนแปลงนี้
2. สามารถอธิบายประเภทของระบบสารสนเทศแบบดั้งเดิมและสารสนเทศสมัยใหม่ที่สำคัญต่อธุรกิจการค้าสมัยใหม่
3. สามารถอธิบายความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูล

1.1 นิยามเกี่ยวกับธุรกิจการค้าสมัยใหม่

ในการศึกษาคำนิยามเกี่ยวกับธุรกิจการค้าสมัยใหม่ (Modern Business Management) หรือ ธุรกิจค้าปลีกสมัยใหม่ (Modern Trade) ในช่วงปี พ.ศ. 2560-2565 มีนักวิชาการทั้งไทยและต่างประเทศได้แก่ คณิศร นิธิสัจชัย หนูเช็กและคณะ (2561) ธาณิก กสิวิทย์อำนวย (2561) และ Brunet-Thornton ผู้เขียนได้สรุปความหมายเบื้องต้นไว้ดังนี้ ธุรกิจการค้าสมัยใหม่คือรูปแบบการทำธุรกิจการค้าที่ใช้เทคโนโลยีเข้ามาช่วยจัดการและดำเนินการภายในองค์กรเพื่อให้สามารถนำเสนอสินค้าและบริการให้แก่ผู้บริโภคได้ มีกระบวนการจัดจำหน่าย ในการนำเสนอสินค้าที่มีความหลากหลาย มีระบบการจัดการ โดยอาศัยความรู้ความชำนาญและเทคโนโลยีมาช่วยในการตอบสนองความต้องการของลูกค้า นำไปสู่ความสัมพันธ์ที่เชื่อมโยงกันระหว่าง ผู้ผลิต ผู้ค้าปลีกและลูกค้า ธุรกิจยุคใหม่ต้องพึ่งพาระบบสารสนเทศ ผู้จัดการทั่วไปและผู้จัดการสายงาน รวมถึงผู้ปฏิบัติงาน จะต้องเกี่ยวข้องกับระบบสารสนเทศและการตัดสินใจ โดยใช้เทคโนโลยีเป็นสิ่งสำคัญ เหตุใดการมีส่วนร่วมของพวกเขาจึงมีความสำคัญอย่างยิ่งต่อความสำเร็จของธุรกิจ

แนวโน้มของธุรกิจค้าในปัจจุบันและอนาคต สรุปได้ดังนี้

1. การแข่งขันของธุรกิจมีแนวโน้มรุนแรงขึ้นจากคู่แข่งรายใหม่ทั้งในและต่างประเทศ รวมถึงคู่แข่งจากร้านค้าออนไลน์ (E-commerce)
2. อำนาจซื้อของผู้บริโภคทำให้ ห้างสรรพสินค้า ร้านค้า ต้องให้บริการในรูปแบบการเลือกซื้อลักษณะ One Stop Shopping ในขั้นตอนเดียว เพื่อให้สอดคล้องกับพฤติกรรมบริโภคที่เปลี่ยนแปลงไปในปัจจุบัน
3. ร้านค้าปลีกรูปแบบใหม่ที่เชื่อมโยงธุรกิจค้าปลีกหลายประเภทเข้าด้วยกัน มีการรวมตัวกันระหว่างผู้ค้าปลีก ผู้ค้าส่ง และผู้ผลิต โดยผู้เล่นรายเดียว หรือบริษัทเดียว และมีการขายร้านค้าขายปลีกเอง เช่น ธุรกิจแฟรนไชส์ (Franchise Business)
4. ให้บริการชำระเงินผ่านระบบออนไลน์ ทุกประเภทในจุดรับชำระเงิน เช่น บัตรเครดิต (Credit Card) บริการกระเป๋าเงิน (Money Wallet) พร้อมเพย์ (Prompt-Pay) บริการโอนเงิน หรือ บัตรเดบิต เป็นต้น
5. พฤติกรรมของคนหรือลูกค้าในยุคปัจจุบันที่ไม่ชอบบรอมมากขึ้น บริษัทต่างๆ จึงให้ความสำคัญกับการให้ลูกค้าบริการตนเอง (Self Service) เช่น การสามารถค้นหาคำตอบเอง หรือ แก้ปัญหานั้นๆได้ด้วยตัวเองในช่องทาง ต่างๆ โดย การนำเทคโนโลยีมาช่วยในการบริการ
6. การนำเทคโนโลยีสมัยใหม่ เช่น ระบบการวางแผนทรัพยากรองค์กร ERP (Enterprise Resource Planning) มาช่วยในการบริหาร เพื่อ การเชื่อมโยงระบบทั้งองค์กร ในการจัดการห่วงโซ่อุปทานหลัก การผลิต การบริการ การเงิน และกระบวนการอื่นๆ ขององค์กร รวมถึงการคาดคะเนยอดขาย (Sale Forecast) การควบคุมสต็อก (Stock Management) รวมถึง การนำเทคโนโลยีใหม่ๆ เช่น บิ๊กดาต้า (Big data) มาใช้วิเคราะห์ข้อมูลเชิงลึกของลูกค้าและธุรกิจ และเพิ่มประสิทธิภาพการบริหารจัดการสินค้าและโลจิสติกส์เพื่อกระจายสินค้าสู่ผู้บริโภคอย่างรวดเร็ว ซึ่งจะเพิ่มความได้เปรียบในการแข่งขัน
7. มีการเพิ่มขึ้นของการค้าปลีกแบบไม่มีร้านค้า หรือ การทำการตลาดในสังคมออนไลน์ (Social Media Marketing) เช่น ขายทางคอมพิวเตอร์ ทางโทรศัพท์ โดยใช้กลยุทธ์ต่างๆ ให้เหมาะสมหรือสอดคล้องกับแพลตฟอร์ม (Platform) ไม่ว่าจะเป็นการสร้างคอนเทนต์ การโพสต์ข้อความ รูปภาพ หรือวิดีโอ หลักๆ แล้วก็เพื่อโฆษณาสินค้าให้เข้าถึงกลุ่มลูกค้าเป้าหมายได้มากยิ่งขึ้น ตัวอย่างของแพลตฟอร์ม (Platform) เช่น Shopee LAZADA TikTok YouTube Instagram Facebook Twitter เป็นต้น

ในส่วนของ ระบบสารสนเทศ เพื่อให้เข้าใจความหมาย ผู้เขียนได้ นำความหมายจากเอกสาร การสอนและ จากนักวิชาการทั้งไทยและต่างประเทศ ได้แก่ โอภาส เอี่ยมสิริวงศ์. (2560) Laudon & Laudon (2013). มาอธิบายไว้ดังนี้

ระบบสารสนเทศ (Information System) คือระบบที่มีการจัดการและประมวลผลข้อมูล เพื่อให้สามารถใช้ในการบริหารจัดการ การวิเคราะห์ และการสื่อสารข้อมูลระหว่างหน่วยงานต่าง ๆ ภายในองค์กร ซึ่งประกอบด้วยระบบคอมพิวเตอร์ ระบบเครือข่ายข้อมูล ระบบสารสนเทศสำหรับบริหารจัดการ และระบบสารสนเทศสำหรับการสื่อสารข้อมูลระหว่างหน่วยงานภายในองค์กร ระบบสารสนเทศสามารถช่วยให้องค์กรสามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ

นอกจากนี้ระบบสารสนเทศ (Information System) ยังหมายถึง ระบบที่มีการเก็บรวบรวม ประมวลผล และแจกจ่ายข้อมูลต่าง ๆ ภายในองค์กร โดยระบบสารสนเทศจะเป็นส่วนหนึ่งของการ บริหารจัดการองค์กร โดยมีการใช้อุปกรณ์คอมพิวเตอร์และระบบสารสนเทศอื่น ๆ ร่วมกันเพื่อช่วยใน การประมวลผลข้อมูล และการสื่อสารข้อมูลระหว่างส่วนต่าง ๆ ขององค์กร ซึ่งระบบสารสนเทศจะ แบ่งเป็นระบบสารสนเทศภายใน (Internal Information System) และระบบสารสนเทศภายนอก (External Information System) ระบบสารสนเทศภายใน (Internal Information System) คือ ระบบสารสนเทศที่มีการใช้ภายในองค์กร โดยมีการเก็บรวบรวม ประมวลผล และแจกจ่ายข้อมูลระหว่าง ส่วนต่าง ๆ ขององค์กร ซึ่งระบบสารสนเทศภายในจะถูกนำไปใช้ภายในองค์กรเพื่อช่วยในการ ประมวลผลข้อมูล และการสื่อสารข้อมูลระหว่างส่วนต่าง ๆ ขององค์กร ระบบสารสนเทศภายในจะมี ประโยชน์ต่อองค์กรโดยที่จะช่วยให้การบริหารจัดการองค์กรทำงานได้อย่างมีประสิทธิภาพ

ระบบสารสนเทศภายนอก (External Information System) คือ ระบบสารสนเทศที่มีการใช้ ภายนอกองค์กร โดยมีการเก็บรวบรวม ประมวลผล และแจกจ่ายข้อมูลระหว่างองค์กรกับภายนอก ซึ่ง ระบบสารสนเทศภายนอกจะมีประโยชน์ต่อองค์กรโดยที่จะช่วยให้องค์กรสามารถเข้าถึงข้อมูลที่ เกี่ยวข้องกับการดำเนินงานขององค์กรจากภายนอก และช่วยให้องค์กรสามารถประยุกต์ใช้ข้อมูลจาก ภายนอกในการดำเนินงานขององค์กรได้

1.2 ประเภทของระบบสารสนเทศ (Type of Information Systems)

ประกอบไปด้วย 6 ชนิด (ดังรูป 1.1) คือ (Laudon & Laudon, 2013)

1. ระบบสารสนเทศแบบประมวลรายการ (Transaction Processing Systems: TPS)
2. ระบบสารสนเทศเพื่อการจัดการ (Management Information Systems: MIS)
3. ระบบสารสนเทศสำนักงานอัตโนมัติ (Office Information System: OAS/OIS)

4. ระบบสารสนเทศภูมิศาสตร์ (Geographic Information System: GIS)
5. ระบบสนับสนุนผู้บริหารระดับสูง (Decision Support System – DSS หรือ Executive Support Systems: ESS) หรือระบบธุรกิจอัจฉริยะ (Business Intelligence: BI)
6. ระบบปัญญาประดิษฐ์และระบบผู้เชี่ยวชาญ (Artificial Intelligence and Expert system: AI)

1. ระบบประมวลผลรายการประจำวัน (Transaction Processing Systems: TPS)

ระบบประมวลผลรายการ หรือทีพีเอส (Transaction Processing System – TPS) เป็นระบบปฏิบัติการ ที่บันทึกและประมวลผลข้อมูลที่เกิดจาก ธุรกิจหรือการปฏิบัติงานประจำหรืองานขั้นพื้นฐานขององค์กร เพื่อใช้ในการจัดการระดับต้นและการปฏิบัติงานของพนักงาน เช่น การบันทึกรายการบัญชี การบันทึกการขาย รายการฝาก-ถอนเงิน TPS ถือเป็นระบบสารสนเทศพื้นฐานขององค์กร รวบรวมและประมวลผลข้อมูลรายละเอียดที่ใช้ในการปฏิบัติงานต่างๆ และข้อมูลจากระบบนี้เป็นฐานให้กับระบบสารสนเทศอื่น ๆ ด้วย ระบบที่สนับสนุนการทำงาน ยังประกอบด้วย ระบบสำนักงานอัตโนมัติ ซอฟต์แวร์สนับสนุนการทำงานกลุ่ม (Groupware) งานส่วนบุคคล (Personal Productivity Software) การทำงานของระบบ TPS การเก็บ (Storage) การบันทึกเหตุการณ์ที่มีผลต่อการปฏิบัติงาน อาจจำเป็นต้องเก็บรักษาข้อมูลไว้ โดยเฉพาะข้อมูลบางประเภทที่จำเป็นต้องเก็บรักษาไว้ตามกฎหมาย ที่จริงแล้ว TPS เกี่ยวข้องกับงานทุกระดับในองค์กร แต่งานส่วนใหญ่ของ TPS จะเกิดขึ้นในระดับปฏิบัติการมากกว่า แม้ว่า TPS จะจำเป็นในการปฏิบัติงานในองค์กร แต่ระบบ TPS ก็ไม่เพียงพอในการสนับสนุนในการตัดสินใจของผู้บริหาร ดังนั้นองค์กรจึงจำเป็นต้องมีระบบอื่นสำหรับช่วยผู้บริหาร

การประมวลผลข้อมูลของ TPS มี 3 วิธี คือ

1. Batch Processing การประมวลผลเป็นชุดโดยมีการรวบรวมข้อมูลที่เกิดจากการทำธุรกรรมที่เกิดขึ้นและรวมไว้เป็นกลุ่มหรือเป็นชุด (Batch) เพื่อตรวจสอบความถูกต้อง หรือจัดลำดับให้เรียบร้อยก่อนที่จะส่งไปประมวลผล โดยการประมวลผลนี้จะกระทำเป็นระยะๆ (อาจจะทำทุกคืน ทุก 2-3 วัน หรือทุกสัปดาห์)

2. Online Processing คือ ข้อมูลจะได้รับการประมวลผลและทำให้เป็นเอาต์พุตทันทีที่มีการป้อนข้อมูลของธุรกรรมเกิดขึ้น เช่น การเบิกเงินจากตู้ ATM จะประมวลผลและดำเนินการทันที เมื่อมีลูกค้าใส่รหัสและป้อนข้อมูลและคำสั่งเข้าไปในเครื่อง

3. Hybrid Systems เป็นวิธีการผสมผสานแบบที่ 1) และ 2) โดยอาจมีการรวบรวมข้อมูลที่เกิดขึ้นทันทีแต่การประมวลผลจะทำในช่วงระยะเวลาที่กำหนด เช่น แคชเชียร์ที่ป้อนข้อมูล การซื้อขายจากลูกค้าเข้าคอมพิวเตอร์ ณ จุดขายของ แต่การประมวลผลข้อมูลจากแคชเชียร์ทุกคนอาจจะทำหลังจากนั้น (เช่น หลังเลิกงาน)

ตัวอย่าง ระบบประมวลผลรายการประจำวัน (TPS)

- งานเงินเดือน (Payroll)
 - การบันทึก เวลาการทำงานของพนักงาน
 - การบันทึก การหักภาษี ค่าประกัน หรือค่าใช้จ่ายอื่นๆ ของพนักงาน
 - การบันทึกการขาย (Sales) ของพนักงาน
 - การบันทึกการออกใบเสร็จรับเงินหรือบิลส่งสินค้า
 - การบันทึกการจ่ายหนี้
 - การบันทึกข้อมูลการส่งสินค้าหรือบริการไปยังลูกค้า
- การสั่งซื้อสินค้า (Purchasing)
 - การบันทึกข้อมูลสั่งซื้อหรือบริการต่างๆ
 - การบันทึกข้อมูล การส่งสินค้าหรือบริการจากซัพพลายเออร์
 - การบันทึกข้อมูลเกี่ยวกับรายรับ (Finance and Accounting)
 - การบันทึกข้อมูลเกี่ยวกับภาษี บัญชี
 - การบันทึกการเบิก จ่าย และควบคุมค่าใช้จ่ายต่างๆ

2. ระบบสารสนเทศเพื่อการจัดการ (Management Information Systems: MIS)

MIS คือระบบสารสนเทศเพื่อการบริหาร และคือกระบวนการประมวลผลข้อมูลโดยใช้เทคโนโลยีคอมพิวเตอร์เข้ามาใช้จัดการข้อมูลอย่างเป็นระบบเพื่อให้ได้สารสนเทศที่มีประสิทธิภาพ โดยจะประมวลผลจากทั้งสารสนเทศภายในและภายนอก มาจัดทำรายงาน หรือสรุปสารสนเทศเพื่อนำไปใช้ในการตัดสินใจในการจัดการงานประจำด้านต่างๆ เช่น การผลิต การจำหน่าย การเงิน การบริหารสินค้าคงคลัง สนับสนุนการทำงานของผู้บริหารในการวางแผนเชิงกลยุทธ์ ซึ่งในปัจจุบัน ระบบ MIS ได้ประยุกต์นำซอฟต์แวร์ที่นำข้อมูลที่มีอยู่เพื่อจัดทำรายงานในรูปแบบต่างๆ ที่เหมาะสมกับมุมมองในการวิเคราะห์ และตรงตามความต้องการของผู้ใช้งาน และซอฟต์แวร์นั้นก็คือ ระบบ BI (Business Intelligence) ใช้สำหรับวิเคราะห์ข้อมูล ของงานในมุมมองต่างๆ เช่น วิเคราะห์ เพื่อการตัดสินใจ

ด้านการลงทุนสำหรับผู้บริหาร วิเคราะห์เพื่อขยายสาขาหรือยุบสาขา วิเคราะห์และวางแผนการขาย และการตลาด

ลักษณะเด่นของ MIS คือ วางแผน ควบคุม และการตัดสินใจซึ่งมักจะนำข้อมูลมาจากระบบ TPS มาทำการประมวลผล ในลักษณะรายงาน ประจำสัปดาห์ เดือน ซึ่งเป็นข้อมูลสรุปหรือเปรียบเทียบ เช่น ระบบการตัดสต็อกสินค้า ระบบการสั่งซื้อสินค้า ที่มีการนำข้อมูลมาสรุปออกเป็นรายงาน เป็นต้น ฯลฯ

3. ระบบสารสนเทศสำนักงานอัตโนมัติ (Office Information System: OAS/OIS)

เป็นระบบสารสนเทศที่ออกแบบขึ้นเพื่อช่วยให้การทำงานในสำนักงานมีประสิทธิภาพมากขึ้น ส่วนใหญ่เป็นการนำคอมพิวเตอร์มาเชื่อมโยงเป็นเครือข่ายกับอุปกรณ์ต่างๆ ของสำนักงาน เพื่อประโยชน์ในการใช้งาน ระบบสำนักงานอัตโนมัติประกอบด้วยเช่น ระบบจัดการเอกสาร ระบบสนับสนุนสำนักงาน ระบบประเมินออนไลน์ ระบบจัดการด้านข่าวสาร ระบบประชุมทางไกล ระบบการควบคุมการไหลของงาน (Workflow) ระบบการลาออนไลน์ (Leave Online System) หรือ Google Doc เป็นต้น

โดยสามารถจำแนกระบบสำนักงานอัตโนมัติออกตามหน้าที่เป็น

3.1 การสื่อสารภายในสำนักงาน เช่น ไปรษณีย์อิเล็กทรอนิกส์หรืออีเมล (E-mail) ไปรษณีย์เสียงหรือเมลเสียง (Voice Mail) การประชุมทางไกล (Teleconferencing) โทรสาร (Fax) เป็นต้น

3.2 เพิ่มประสิทธิภาพการทำงานสำนักงาน เช่นการประมวลคำ (Word Processing) การประมวลภาพลักษณ์ (Image Processing) การพิมพ์ตั้งโต๊ะ ซึ่งเป็นการผลิตเอกสารที่มีคุณภาพดีการออกแบบกราฟิกและรูปแบบประเภทต่างๆ (Desktop Publishing) การแปลงภาพและเอกสารในรูปดิจิทัล (Digitization)

3.3 เพิ่มประสิทธิภาพในการประสานงาน ปฏิบัติงานร่วมกัน เช่น ทางไกล ระบบการควบคุมการไหลของงาน (Workflow) กรุปแวร์ (Groupware) เป็นซอฟต์แวร์ที่ใช้สนับสนุนการทำงานร่วมกัน มีองค์ประกอบที่สำคัญ 3 ประการ คือ การสื่อสาร การร่วมมือ (Collaboration) และการประสานงาน (Coordination) เช่น ระบบ Google Doc ระบบอินทราเน็ตที่เป็นทางเลือกที่สำคัญของการใช้ กรุปแวร์ และเป็นระบบที่เพิ่มประสิทธิภาพงานด้านการจัดการสำนักงานและการสื่อสาร พนักงานในองค์กรต้องใช้คอมพิวเตอร์ในการจัดการกับเอกสาร

4. ระบบสารสนเทศภูมิศาสตร์ (Geographic Information System: GIS) คือระบบสารสนเทศที่ทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ โดยข้อมูลมีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ ประกอบไปด้วยข้อมูลเช่น ที่อยู่ บ้านเลขที่ สัมพันธ์กับตำแหน่งในแผนที่ ตำแหน่ง เส้นรุ้ง เส้นแวง ข้อมูลและแผนที่ ระบบสารสนเทศภูมิศาสตร์ในปัจจุบัน เป็นรูปแบบแทนสิ่งต่างๆ บนโลกที่เป็นลายเส้นและพื้นที่ด้วยสัญลักษณ์แบบ 1. จุด (Point) 2. เส้น (Line) 3. โพลีกอน (Polygon) และข้อมูลจะเป็นลักษณะของ META DATA ที่มีการบันทึก แก่ไขรายวัน เช่น ข้อมูลในระบบแผนที่ของ Google Map

ประโยชน์ของระบบสารสนเทศภูมิศาสตร์ เช่น การวางแผนการเดินทาง การวางแผนผังเมือง การวางแผนล่วงหน้าในการติดตั้งระบบไฟฟ้า ประปา ท่อน้ำทิ้ง บ่อบำบัดน้ำเสีย เป็นต้น

5. ระบบสนับสนุนผู้บริหารระดับสูง (Decision Support System - DSS หรือ Executive Support Systems: ESS) หรือระบบธุรกิจอัจฉริยะ (Business Intelligence: BI) หรือบางครั้งอาจเรียกว่าระบบสารสนเทศเพื่อผู้บริหาร เพื่อช่วยสนับสนุนกระบวนการของการตัดสินใจทางการจัดการให้มีประสิทธิภาพ โดยใช้ตัวแบบ (Model) ด้านสถิติ คณิตศาสตร์ เศรษฐศาสตร์ ฯลฯ วิเคราะห์ข้อมูล และการจำลองสถานการณ์เปรียบเทียบทางเลือกต่างๆ เพื่อการตัดสินใจเลือกทางเลือกที่ดีที่สุด และเหมาะสมที่สุดในสถานการณ์นั้น เปรียบเทียบ เพื่อใช้ประโยชน์ต่อการตัดสินใจ โดยจะเป็นระบบ

- ที่สามารถให้ทางเลือกโดยอาศัยข้อมูลประกอบการตัดสินใจ
- นำข้อมูลจาก TPS MIS มาใช้งานภายในระบบ
- นำสารสนเทศจากภายนอกเข้าสู่ระบบ เช่น การนำข้อมูลราคาหุ้นของตลาดหุ้นมาประกอบการพิจารณา หรือการนำราคาผลิตภัณฑ์ของคู่แข่งมาประกอบการพิจารณา เป็นต้น

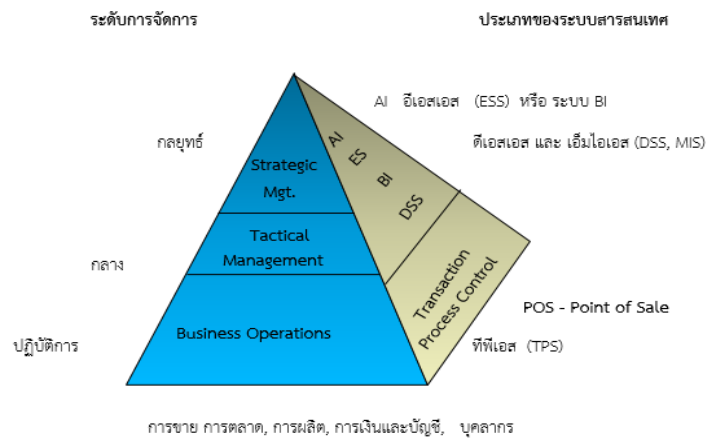
- เป็นการตัดสินใจแบบกึ่งโครงสร้าง หรือไม่มีโครงสร้างหน้าที่หลักในการประมวลผลข้อมูล
 - การสร้างแบบจำลอง เช่น การพยากรณ์ยอดขาย
 - การวิเคราะห์แบบ What-If Analysis
 - วิธีนี้เป็นกึ่งมีโครงสร้าง และไม่มีโครงสร้าง
 - การวิเคราะห์แบบ Goal Seeking เป็นกระบวนการที่ผู้ตัดสินใจกำหนดผลลัพธ์
- การวิเคราะห์แบบ Risk Analysis เป็นการวิเคราะห์ทางเลือกในการตัดสินใจ

- การวิเคราะห์แบบ Graphical Analysis คือการวิเคราะห์และแสดงผลลัพธ์

ตัวอย่างระบบสนับสนุนการตัดสินใจ (DSS) บริษัท ค็อกพิท (Cockpit) มีระบบสนับสนุนการตัดสินใจในการออกแบบยางรถยนต์ยี่ห้อใหม่

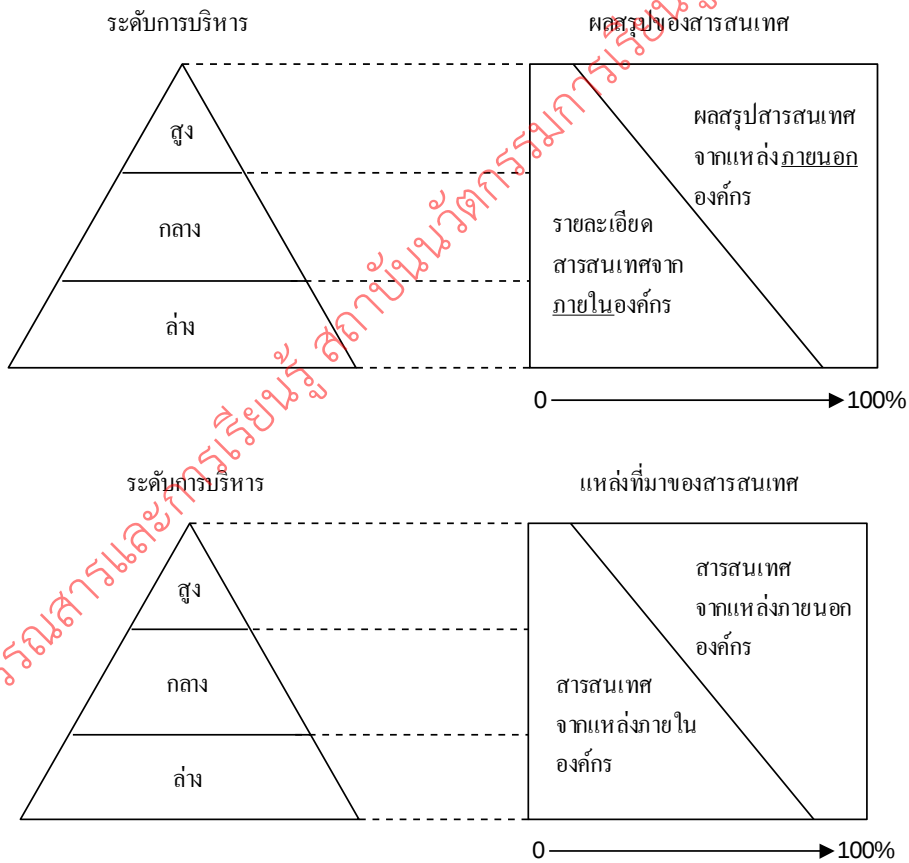
- ระบบช่วยให้นักวิเคราะห์มองเห็นถึงความสัมพันธ์ระหว่างผลทางการเงินในอดีตกับตัวแปรภายนอก เช่น จำนวนรถยนต์ที่ผลิตทั้งหมด ผลิตภัณฑ์มวลรวมภายในประเทศแล้วนำมาสร้างโมเดลในการพยากรณ์การขายด้วยโมเดลต่างๆ ที่มีอยู่ในระบบ
- สามารถสร้างฐานข้อมูลที่บรรจุยางของคู่แข่งทั้งหมด 200 ยี่ห้อ รวมทั้งข้อมูลในการผลิต แรงกด ปริมาณ และประมาณการขายผู้บริหารสามารถใช้ข้อมูลในฐานข้อมูลนี้สำหรับการกำหนดกลยุทธ์ในการแข่งขัน
- ระบบช่วยให้องค์กรสามารถนำเอาประเด็นด้านเทคโนโลยีมาผนวกกับ ด้านการเงินในการตัดสินใจที่เกี่ยวกับผลิตภัณฑ์และช่วยสนับสนุนให้มีการตัดสินใจร่วมกันของหน่วยงานตามหน้าที่ต่างๆ ในองค์กร

ส่วนระบบ EIS- อีไอเอส (Executive Information System – EIS) หรือระบบ BI (Business Intelligence) ออกแบบเพื่อให้สารสนเทศที่มีความสำคัญลำดับสูงต่อองค์กร เพื่อนำไปสู่การตัดสินใจการแก้ไขปัญหา และการวางแผนระยะยาว โดยมีข้อมูลทั้งจากภายในและภายนอกองค์กร และมีการนำเสนอที่ชัดเจน ระบบ ESS, BI คล้ายกับระบบ DSS แต่จะแตกต่างกันที่ระบบ ESS เป็นการตัดสินใจในระดับกลยุทธ์และนโยบายผู้บริหารระดับสูง นำสารสนเทศจากภายนอกมาประกอบการตัดสินใจเชิงกลยุทธ์ และการตัดสินใจแบบไม่มีโครงสร้าง ส่วนระบบ BI เป็นการนำเอาข้อมูล มาวิเคราะห์และประมวลผลข้อมูล ใช้ในการทำธุรกิจ ทั้งการบริหารจัดการ โดยการเสนองานผ่านซอฟต์แวร์ เช่น Power BI, Tableau เพื่อให้เห็นมุมมองทั้งในอดีต ปัจจุบัน และการคาดการณ์ในอนาคตของการดำเนินธุรกิจ



รูปที่ 1. 1 ประเภทของระบบสารสนเทศจำแนกตามระดับชั้นองค์กร

ที่มา : Laudon and Laudon (2013)



รูปที่ 1. 2 ส่วนผสมของข้อมูลในแต่ละระดับชั้นองค์กร

ตัวอย่างระบบสนับสนุนผู้บริหารระดับสูง (ESS)

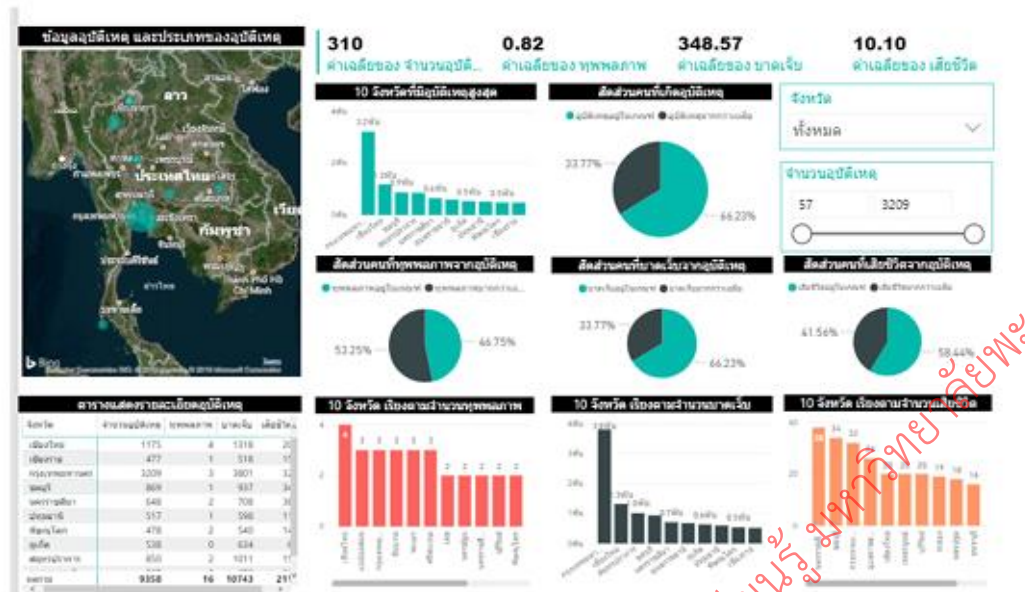
บริษัท Cisco

- กำหนดให้ซัพพลายเออร์ (Supplier) เข้ามามีส่วนร่วมในการกำหนดอุปสงค์ผ่านทางระบบการใช้ข้อมูลร่วมกัน
- ห้ส่วนแต่ละรายของซัสโกก็สร้างโมเดลการทำนายอุปสงค์อุปทานของตัวเอง โดยอาศัยการคัดลอกข้อมูลที่เหลือมาจาก Supply Chain เพื่อช่วยลดภาระสินค้าคงเหลือของซัพพลายเออร์ให้เหลือน้อยที่สุด (บุญเลิศ วงศ์พรหม, 2544)
- ซัสโกมีเว็บไซต์สำหรับการติดต่อระหว่างซัพพลายเออร์ (Supplier) และลูกค้า ทำให้สามารถที่จะเก็บรวบรวมข้อมูลเกี่ยวกับแนวโน้มของผู้บริโภค ว่าลูกค้าต้องการอะไรตลอดจนความเคลื่อนไหวในอุตสาหกรรมเดียวกันได้อย่างมีประสิทธิภาพ

Employee Self-Service (ESS) คือ ระบบบริหารงานบุคคล ที่ช่วยสนับสนุนการทำงานของเช่นระบบ Pro-soft HRMI โดยเป็นระบบงานที่มีส่วนเกี่ยวข้องกับพนักงานโดยตรง และมีความถี่ในการใช้งานบ่อยเข้าสู่ระบบ ESS พนักงานสามารถจัดการข้อมูลต่างๆ ได้ด้วยตนเอง ผ่านเครือข่ายออนไลน์ (Web Browser) เช่น การเข้าถึงข้อมูลข่าวสารขององค์กร การปรับปรุงข้อมูลส่วนตัว บันทึกร้องขออนุมัติต่างๆ ซึ่งผู้บริหารสามารถเข้าไปพิจารณาอนุมัติการร้องขอผ่านระบบ ESS ได้เช่นกัน ทั้งนี้การเข้าถึงข้อมูลจะถูกกำหนดสิทธิ์ของผู้ใช้งานอย่างเป็นระบบเพื่อป้องกันความปลอดภัยของฐานข้อมูล

ตัวอย่างระบบ BI (Business Intelligence)

สถิติอุบัติเหตุ



สถิติ Suppliers, Clients, Employees, Freight Cost



รูปที่ 1. 3 Power BI dashboard and Report Example (ภาพจากการเรียนการสอนของผู้เขียน)

ระบบธุรกิจอัจฉริยะ (Business Intelligence) จะประกอบไปด้วยระบบข้อมูล และใช้โปรแกรม ด้านการวิเคราะห์ เช่น ระบบคลังข้อมูล (Data Warehouse) ดาต้ามาร์ท (Data Mart) การทำเหมืองข้อมูล (Data Mining) การแก้ปัญหาทางคณิตศาสตร์ (Operations Research & Numerical Methods) เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลในหลายมิติ (OLAP) โดยจัดทำรายงานในรูปแบบต่างๆ

ที่เหมาะสมกับมุมมองในการวิเคราะห์ ตัวโปรแกรมมีความสามารถในการให้ผู้ใช้ เปลี่ยนแปลงรายงาน ได้โดยไม่ต้องมีการนำเข้าสู่ข้อมูลใหม่ ซึ่งผู้ใช้สามารถถาม วิเคราะห์มุมมองในเวลาอันรวดเร็ว

6. ระบบปัญญาประดิษฐ์และระบบผู้เชี่ยวชาญ (Artificial Intelligence and Expert system: AI) ระบบปัญญาประดิษฐ์คือ ศาสตร์แขนงหนึ่งของวิทยาศาสตร์ คอมพิวเตอร์ ที่เกี่ยวข้องกับวิธีการทำให้คอมพิวเตอร์มีความสามารถคล้ายมนุษย์หรือความสามารถในการคิดเองได้ หรือมีปัญหา นั่นเอง ตัวอย่างเช่น ระบบการประมวลผลภาษาธรรมชาติ (Natural Language Processing:NLP) ชื่อ “SIRI” ที่ใช้ใน Smart Phone ระบบหุ่นยนต์ ที่ใช้งานในโรงงาน ส่วนระบบผู้เชี่ยวชาญ เป็นระบบที่รวบรวมความรู้ความเชี่ยวชาญเฉพาะด้านในแต่ละสาขา เป็นระบบที่ข้อมูลภายในระบบจะเป็นความรู้ ที่เกิดขึ้นในแต่ละสาขา อาทิเช่น ประสบการณ์ ความรู้เชิงประจักษ์

ตัวอย่างของระบบผู้เชี่ยวชาญ (Expert System :ES) ที่นำไปใช้ในงานด้านต่างๆ

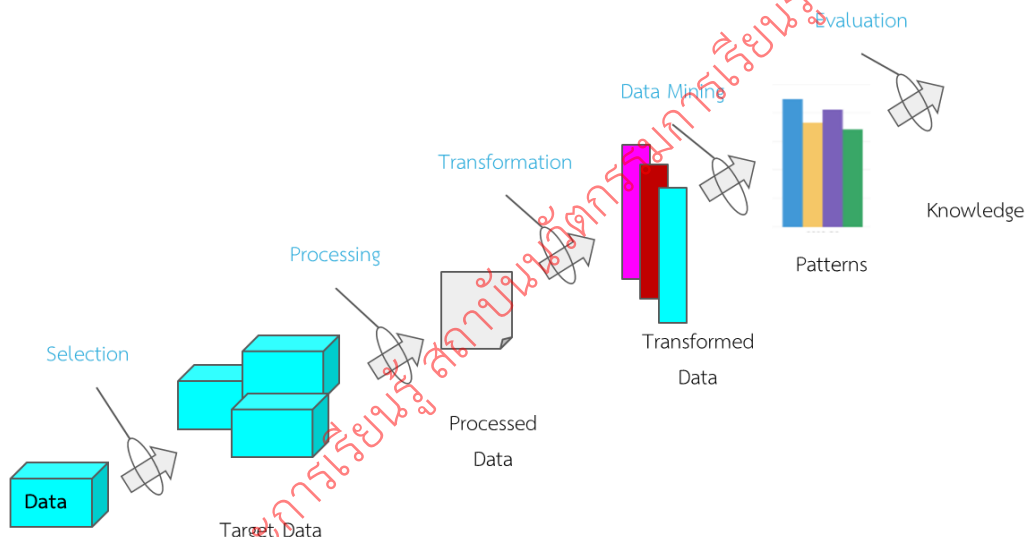
1. ด้านการแพทย์: การให้คำแนะนำแก่หมอในการสั่งยาปฏิชีวนะให้คนไข้ซึ่งต้องคำนึงถึงปัจจัยต่างๆ หลายประการ เช่น ประวัติการเจ็บป่วยของคนไข้ แหล่งติดเชื้อ ราคาของยา
2. ด้านการผลิต: การให้คำแนะนำแก่โรงงานในการผลิตผลิตภัณฑ์ต่างๆ เช่น ชิ้นส่วนเครื่องบิน
3. ด้านธรณีวิทยา: ให้คำแนะนำแก่นักธรณีวิทยาในการวิเคราะห์ดินและน้ำมัน เพื่อพิจารณาในการขุดเจาะน้ำมัน

นอกจากนี้แล้วองค์กรธุรกิจสมัยใหม่ ได้นำระบบวิเคราะห์ข้อมูลเชิงลึกหรือการทำเหมืองข้อมูล มาใช้หาแนวทางการตัดสินใจและคาดการณ์ผลลัพธ์ที่ได้แม่นยำความวิธีการทางสถิติเดิมๆ ซึ่งตัวอย่างการนำเหมืองข้อมูลไปใช้งานในธุรกิจ เช่น การทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่ การทำนายยอดขายเมื่อมีการลดราคาสินค้า การทำนายกลุ่มลูกค้าที่น่าจะใช้สินค้าของเรา หากธุรกิจสามารถรู้ล่วงหน้า และวิ่งนำหน้าผู้บริโภค ด้วยการใช้ข้อมูลการสั่งซื้อ ประวัติลูกค้าทั้งเก่าและใหม่ เพื่อนำมาวิเคราะห์หาความน่าจะเป็น คาดคะเนจากเทคนิคเหมืองข้อมูล และแปลผลรูปแบบข้อมูลที่ได้จากการวิเคราะห์ เพื่อให้จับต้องได้ จะทำให้สามารถนำมาพัฒนากลยุทธ์ขององค์กร ซึ่งผู้ที่เข้าใจการทำ Data Mining (เหมืองข้อมูล)

1.3 ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) เป็นการค้นพบองค์ความรู้ใหม่จากฐานข้อมูล (Knowledge Discovery in Database: KDD) กล่าวคือ เป็นกระบวนการที่กระทำกับข้อมูลจำนวนที่มีขนาดใหญ่ เพื่อค้นหารูปแบบ (Patterns) และความสัมพันธ์ (Associations) ที่ซ่อนอยู่ในชุดข้อมูลนั้นๆ

รวมถึงการบูรณาการศาสตร์ต่างๆ เข้าด้วยกัน เช่น ระบบฐานข้อมูล (Database System) ระบบคลังข้อมูล (Data Warehouse) กระบวนการวิเคราะห์ออนไลน์ (Online Analytical Processing : OLAP) เป็นต้น ซึ่งจัดเป็นเทคโนโลยีการจัดเก็บ รวบรวม และเตรียมข้อมูลที่ใช้ เพื่อทำเหมืองข้อมูล การเรียนรู้ของเครื่อง (Machine Learning) โดยใช้อัลกอริทึมหลักที่ใช้ในการค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูล หลักสถิติ (Statistic and Data Analysis Methods) สำหรับวิเคราะห์ข้อมูลเบื้องต้น ซึ่งอาจชี้ให้เห็นถึงรูปแบบและความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ได้ รวมถึงการประยุกต์ใช้วิทยาศาสตร์สารสนเทศ (Information Science) โปรแกรมทางคณิตศาสตร์ (Mathematical Programming) และการคำนวณประสิทธิภาพสูง (High Performance Computing) ทั้งนี้ จากข้อมูลที่มีขนาดใหญ่ จึงทำให้การทำเหมืองข้อมูลมักใช้เวลาานาน ดังนั้นจึงจำเป็นต้องมีการคำนวณที่รวดเร็วรองรับ รวมถึงการทำจินตทัศน์ (Visualization) เพื่อแสดงผลลัพธ์ รูปแบบ และความสัมพันธ์ของข้อมูลออกมาให้ผู้ใช้งานเข้าใจง่ายที่สุด ซึ่งเป็นประโยชน์ต่อการตีความและนำผลลัพธ์นั้นไปใช้ต่อไปทางธุรกิจ (จิราภา เลาหะวรรณัทและคณะ, 2561)



รูปที่ 1. 4 แผนภาพแสดงกระบวนการทำเหมืองข้อมูล (Data Mining Process) ปรับปรุงจาก (IBM, 2011)

ทำไมจึงต้องเรียนรู้ Data Mining (เหมืองข้อมูล)? ในยุคที่เทคโนโลยีพัฒนาขึ้นอย่างรวดเร็ว ธุรกิจต้องวิ่งไล่ตามความต้องการของผู้บริโภค การรับรู้ข่าวสารที่รวดเร็วยิ่งขึ้น ผู้บริโภครุ่นใหม่เป็นกลุ่มที่เติบโตมากับความ ‘เรียลไทม์’ แคนั่งอยู่หน้าจอ ก็ได้ทุกอย่างแม้กระทั่งการสั่งอาหาร ข้อมูลผ่านโซเชียลมีมามากมายมหาศาล ส่งผลให้พฤติกรรมของผู้บริโภคหันไปใช้จ่ายออนไลน์มากขึ้น ผ่านการตัดสินใจจากข้อมูลจากรีวิว และการนำเสนอผ่านระบบข้อมูลเชิงลึกของธุรกิจออนไลน์ เช่น YouTube Netflix SHOPEE LAZADA GRAP LINEMAN นั่นคือความสะดวกสบายของการช้อปปิ้งออนไลน์ ถัดมาคือการส่งของที่รวดเร็ว นี่เป็นอีกประเด็นที่น่าสนใจ เพราะปัจจุบันพบว่าธุรกิจเกี่ยวกับโลจิสติกส์ต่างๆ

พัฒนาให้มีบริการที่หลากหลายมากขึ้น การทำเหมืองข้อมูล (Data Mining) เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลแบบเดิมไม่ซับซ้อน และไม่มีข้อมูลแบบภาพ เสียง มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้ในการตัดสินใจหาข้อมูลเชิงลึก ที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล เช่น การนำข้อมูลขนาดใหญ่ มาทดลองประมวลผลและการสังเกตการณ์รูปแบบของพฤติกรรมผู้บริโภค โดยการสกัดหรือแยกข้อมูลที่เป็นประโยชน์จากข้อมูล การวางแผนทรัพยากรขององค์กรโดยสามารถวิเคราะห์ทางสถิติและตรรกะของข้อมูลขนาดใหญ่เป็นการมองหารูปแบบที่สามารถช่วยการตัดสินใจได้ เช่น จากงานวิจัยผู้เขียนที่วิเคราะห์ข้อมูลนักศึกษาของสถาบันแห่งหนึ่ง เพื่อประเมินผลการเรียน การสอนย้อนหลัง 10 ปี โดยผลการศึกษา ใช้สำหรับการตัดสินใจวางแผนการศึกษาในอนาคต และสำหรับองค์กรในการตัดสินใจวางแผน รูปแบบหลักสูตร โดยทำให้การตัดสินใจทางธุรกิจที่ดีขึ้นจากการให้ข้อมูลเชิงลึกในข้อมูลขนาดใหญ่จะลูกค้านำมาขึ้นทำให้ทราบว่าตอนนี้ลูกค้าต้องการอะไร แล้วลูกค้าประเภทไหนที่ซื้อสินค้าหรือใช้บริการ รวมถึงสามารถทำนายพฤติกรรมของนักศึกษาหรือผู้บริโภค และออกแบบสินค้าและบริการ ให้ตอบโจทย์กลุ่มนักศึกษาหรือลูกค้าประจำและกลุ่มลูกค้าใหม่ได้ และเห็นมุมมองในการทำตลาดที่กว้างขึ้น จากการมีข้อมูลในมือจำนวนมากและหลายรูปแบบ ทำให้มุมมองในการวางแผนการตลาดก็ยิ่งกว้างมากขึ้น (Chertchom P, 2017)

กรณีศึกษาด้านระบบฐานข้อมูลและการจัดการข้อมูล

กรณีศึกษา: การจัดการข้อมูลลูกค้าในบริษัทออนไลน์

ที่มาของปัญหา: บริษัทออนไลน์แห่งหนึ่งขายสินค้าผ่านเว็บไซต์พบว่าบริษัทมีข้อมูลลูกค้าจำนวนมากที่เก็บรวบรวมมาตลอดระยะเวลา แต่บริษัทไม่สามารถนำข้อมูลนี้มาใช้ในการตัดสินใจและการวางแผนการตลาดได้อย่างมีประสิทธิภาพ

วิธีการ: บริษัทตัดสินใจจัดการข้อมูลลูกค้าโดยใช้กระบวนการทำเหมืองข้อมูลเพื่อวิเคราะห์และค้นหาความรู้จากข้อมูลลูกค้า บริษัทเริ่มต้นด้วยขั้นตอนการเตรียมข้อมูลที่ถูกต้องและสมบูรณ์ จากนั้นพวกเขาใช้เทคนิคการจัดกลุ่มข้อมูล (Clustering) เพื่อแบ่งลูกค้าออกเป็นกลุ่มตามพฤติกรรมการซื้อสินค้า และใช้เทคนิคการจำแนก (Classification) เพื่อคาดการณ์พฤติกรรมลูกค้าในอนาคต

ผลลัพธ์: ผลลัพธ์ที่ได้คือการแบ่งกลุ่มลูกค้าเป็นกลุ่มที่มีความสำคัญและพฤติกรรมการซื้อที่คล้ายกัน นอกจากนี้ก็สามารถทำนายพฤติกรรมการซื้อของลูกค้าในอนาคตได้ ทำให้บริษัทสามารถปรับแผนการตลาดและกลยุทธ์การขายได้อย่างมีประสิทธิภาพ

ประโยชน์: การจัดการข้อมูลลูกค้าในลักษณะนี้ช่วยให้บริษัทสามารถตอบสนองต่อความต้องการของลูกค้าได้ดีขึ้น มีการปรับแผนการตลาดที่เน้นไปที่กลุ่มลูกค้าที่สนใจสินค้าของบริษัทมากที่สุด และมีการพัฒนากลยุทธ์การขายที่มีเสถียรภาพ

ในกรณีศึกษาด้านการจัดการข้อมูลนี้ จะเห็นว่าการนำหลักการและเทคนิคทางการทำเหมืองข้อมูลมาประยุกต์ใช้ในการจัดการข้อมูลลูกค้า ช่วยให้บริษัทมีการตัดสินใจและวางแผนที่มีความ

แม่นยำและประสิทธิภาพสูงขึ้น
ต้องการของลูกค้าในอนาคต

และช่วยในการปรับกลยุทธ์การตลาดและการขายให้เข้ากับความต้องการ

สรุปเนื้อหาบทที่ 1 "ระบบสารสนเทศเพื่อการจัดการธุรกิจการค้าสมัยใหม่" เป็นการเสนอแนะเกี่ยวกับวิธีที่ระบบสารสนเทศมีบทบาทสำคัญในการจัดการธุรกิจและการค้าในยุคปัจจุบัน ที่เรียกว่า "สมัยใหม่" หรือ "ยุคดิจิทัล" โดยเน้นการใช้เทคโนโลยีสารสนเทศในการจัดการและการตัดสินใจทางธุรกิจอย่างมีประสิทธิภาพและมีความสามารถในการปรับตัวกับการเปลี่ยนแปลงของสถานะตลาดและสภาพแวดล้อมทางธุรกิจอย่างรวดเร็ว บทที่ 1 จะให้ความเข้าใจเกี่ยวกับความสำคัญของระบบสารสนเทศในการจัดการธุรกิจในปัจจุบัน การนำเทคโนโลยีสารสนเทศเข้ามาช่วยในการเก็บรวบรวมข้อมูลที่สำคัญและใช้ข้อมูลเหล่านั้นในการวิเคราะห์และการตัดสินใจทางธุรกิจ เพื่อให้สามารถตอบสนองต่อการเปลี่ยนแปลงที่เกิดขึ้นในตลาดได้อย่างต่อเนื่อง นอกจากนี้ยังได้เน้นถึงความสำคัญของการใช้ข้อมูลเพื่อสร้างความเข้าใจเชิงลึกเกี่ยวกับกลุ่มลูกค้า แนวโน้มของตลาด และความต้องการของลูกค้า เพื่อให้ธุรกิจสามารถปรับกลยุทธ์และการตลาดให้เข้ากับความต้องการของตลาดได้อย่างเหมาะสม ในสรุป, บทที่ 1 เน้นความสำคัญของระบบสารสนเทศในการจัดการธุรกิจในยุคสมัยใหม่โดยต้องมีความรู้พื้นฐานด้านประเภทของระบบสารสนเทศ และการใช้เทคโนโลยีสารสนเทศในการเก็บรวบรวมและวิเคราะห์ข้อมูลเช่น ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูล เพื่อการตัดสินใจทางธุรกิจที่มีความสำคัญต่อความสำเร็จของธุรกิจในสถานะที่เปลี่ยนแปลงอย่างรวดเร็วในปัจจุบัน

คำศัพท์ บทที่ 1

Artificial Intelligence and Expert system	ระบบปัญญาประดิษฐ์และระบบผู้เชี่ยวชาญ
Batch processing	การประมวลผลเป็นชุด
Business Intelligence	ระบบธุรกิจอัจฉริยะ
Credit Card	บัตรเครดิต
Data Mining	การทำเหมืองข้อมูล
Decision Support System	ระบบสนับสนุนผู้บริหารระดับสูง
E-commerce	อี-คอมเมิร์ซ
Employee Self-Service	ระบบบริหารงานบุคคล
External Information System	ระบบสารสนเทศภายนอก
Franchise	ธุรกิจแฟรนไชส์
Geographic Information System	ระบบสารสนเทศภูมิศาสตร์
High Performance Computing	การคำนวณประสิทธิภาพสูง

Information Science	วิทยาศาสตร์สารสนเทศ
Information System	ระบบสารสนเทศ
Internal Information System	ระบบสารสนเทศภายใน
Knowledge Discovery in Database	การค้นพบองค์ความรู้ใหม่จากฐานข้อมูล
Machine Learning	การเรียนรู้ของเครื่อง
Management Information Systems	ระบบสารสนเทศเพื่อการจัดการ
Mathematical Programming	โปรแกรมทางคณิตศาสตร์
Modern Business Management	ธุรกิจการค้าสมัยใหม่
Modern Trade	ธุรกิจค้าปลีกสมัยใหม่
Money Wallet	กระเป๋าเงิน
Office Information System	ระบบสารสนเทศสำนักงานอัตโนมัติ
Online processing	การประมวลผลออนไลน์
Payroll	เงินเดือน
Platform	แพลตฟอร์ม
Prompt-Pay	พร้อมเพย์
Purchasing	การสั่งซื้อสินค้า
Sale Forecast	การคาดคะเนยอดขาย
Self Service	บริการตนเอง
Social Media Marketing	การทำตลาดในสังคมออนไลน์
Stock Management	การควบคุมสต็อก
supplier	ซัพพลายเออร์
Transaction Processing Systems	ระบบสารสนเทศแบบประมวลรายการ
Visualization	การทำจินตทัศน์
Modern Business Management	ธุรกิจการค้าสมัยใหม่
Modern Trade	ธุรกิจค้าปลีกสมัยใหม่
E-commerce	อี-คอมเมิร์ซ
Franchise	ธุรกิจแฟรนไชส์
Credit Card	บัตรเครดิต
Money Wallet	กระเป๋าเงิน
Prompt-Pay	พร้อมเพย์
Self Service	บริการตนเอง

Sale Forecast	การคาดคะเนยอดขาย
Stock Management	การควบคุมสต็อก
Social Media Marketing	การทำตลาดในสังคมออนไลน์
Platform	แพลตฟอร์ม
Information System	ระบบสารสนเทศ
Internal Information System	ระบบสารสนเทศภายใน
External Information System	ระบบสารสนเทศภายนอก
Transaction Processing Systems	ระบบสารสนเทศแบบประมวลรายการ
Management Information Systems	ระบบสารสนเทศเพื่อการจัดการ
Office Information System	ระบบสารสนเทศสำนักงานอัตโนมัติ
Geographic Information System	ระบบสารสนเทศภูมิศาสตร์
Decision Support System	ระบบสนับสนุนผู้บริหารระดับสูง
Business Intelligence	ระบบธุรกิจอัจฉริยะ
Artificial Intelligence and Expert system	ระบบปัญญาประดิษฐ์และระบบผู้เชี่ยวชาญ
Batch processing	การประมวลผลเป็นชุด
Online processing	การประมวลผลออนไลน์
Payroll	เงินเดือน
Purchasing	การสั่งซื้อสินค้า
supplier	ซัพพลายเออร์
Employee Self-Service	ระบบบริหารงานบุคคล
Data Mining	การทำเหมืองข้อมูล
Knowledge Discovery in Database	การค้นพบองค์ความรู้ใหม่จากฐานข้อมูล
Machine Learning	การเรียนรู้ของเครื่อง
Information Science	วิทยาศาสตร์สารสนเทศ
Mathematical Programming	โปรแกรมทางคณิตศาสตร์
High Performance Computing	การคำนวณประสิทธิภาพสูง
Visualization	การทำจินตทัศน์

คำถามและแบบฝึกหัดท้ายบทที่ 1

1. จงอธิบายความหมายของระบบหมายถึงอะไร และระบบที่ดีควรมีลักษณะอย่างไร
2. จงอธิบายถึงความแตกต่างของ ข้อมูล (Data) และ สารสนเทศ (Information)

3. จงอธิบายส่วนประกอบด้วยส่วนประกอบระบบสารสนเทศเพื่อการจัดการ
4. จงอธิบายประโยชน์ของระบบสารสนเทศ
5. จงอธิบายส่วนประกอบด้วยส่วนประกอบระบบสารสนเทศเพื่อการจัดการ
6. จงอธิบายประโยชน์ของระบบสารสนเทศ
7. ทำไมจึงต้องเรียนรู้ Data Mining (เหมืองข้อมูล)

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

บทที่ 2 การพัฒนาและวิเคราะห์ระบบสารสนเทศสมัยใหม่

(Modern Information System Development)

ในภาวะสังคมปัจจุบันที่เปลี่ยนไป ธุรกิจการค้าสมัยใหม่ได้เกิดขึ้นพร้อมกับการก้าวสู่ยุคสารสนเทศ ดังนั้นความรู้ด้านคอมพิวเตอร์และการส่งเสริมการนำเทคโนโลยีคอมพิวเตอร์มาประยุกต์งานต่างๆ จึงสำคัญมากขึ้น การบริหารธุรกิจและการพัฒนาระบบสารสนเทศ เพื่อใช้ในองค์กรด้วยการเก็บข้อมูล ประมวลผลและวิเคราะห์ข้อมูล แล้วนำผลลัพธ์มาช่วยในการวางแผนและตัดสินใจ การทำความเข้าใจวิธีการที่ใช้ในการสร้างระบบสารสนเทศขึ้นมาใหม่ในธุรกิจใดธุรกิจหนึ่งหรือระบบย่อยของธุรกิจ นอกจากการสร้างระบบสารสนเทศใหม่แล้ว การวิเคราะห์ระบบช่วยในการแก้ไขระบบสารสนเทศเดิมที่มีอยู่แล้วให้ดีขึ้นด้วยก็ได้ ก็มีความสำคัญอย่างมาก

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจ ขั้นตอนที่ใช้ในการวิเคราะห์และออกแบบระบบ และอธิบายหน้าที่ความรับผิดชอบของนักวิเคราะห์ระบบได้
2. สามารถอธิบายประเภทของโมเดล หรือแบบจำลองการพัฒนาซอฟต์แวร์ได้และนำมาประยุกต์ใช้ในการวิเคราะห์และออกแบบระบบสารสนเทศสมัยใหม่

2.1 ระบบสารสนเทศพื้นฐาน

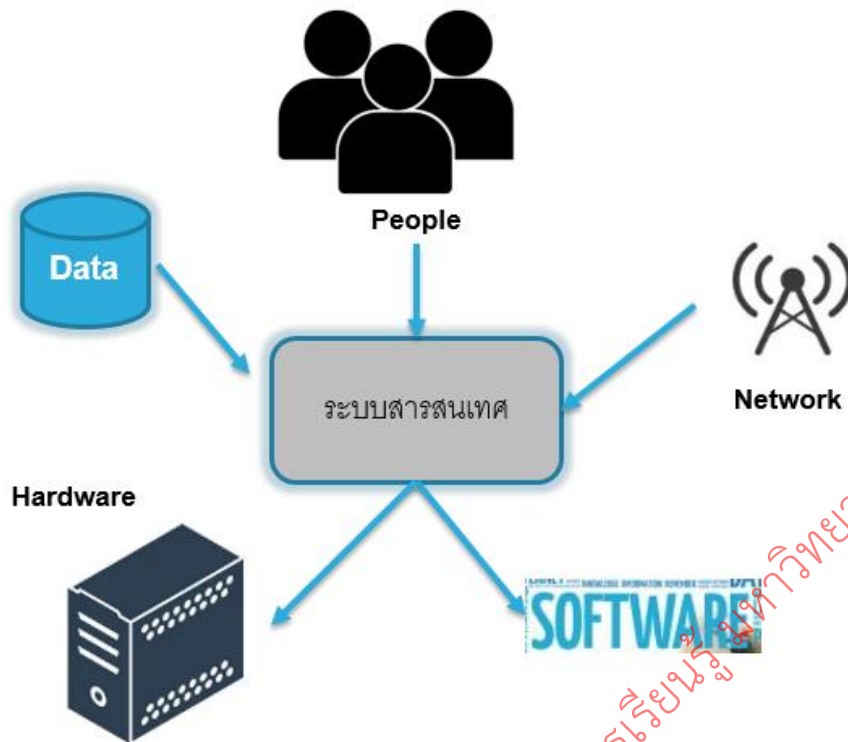
องค์กรและธุรกิจในทุกอุตสาหกรรม กำลังค้นหาวิธีการใหม่ๆและมีประสิทธิภาพมากขึ้นในการใช้ข้อมูลนี้เพื่อสร้างกลยุทธ์ทางธุรกิจ ในขณะที่ประเด็นทางจริยธรรมที่เกี่ยวข้องกับก้า ก็มาพร้อมกับการจัดการข้อมูลส่วนบุคคลจำนวนมาก ในขณะที่โครงสร้างพื้นฐานด้านดิจิทัลทั่วโลกของเราเติบโตขึ้น ด้านการจัดการข้อมูลก็กำลังพัฒนาไปอย่างรวดเร็วควบคู่ไปด้วย ปัจจุบันข้อมูลเป็นสิ่งสำคัญมาก ธุรกิจได้นำข้อมูลไปใช้ประโยชน์มากมาย ในการปฏิบัติงาน โดยข้อมูลทุกวันนี้ มีหลายรูปแบบ เช่น เป็นตัวเลข ข้อความ รูปภาพ เสียงต่าง ๆ ธุรกิจทุกขนาดจะมีระบบสารสนเทศในระดับหนึ่ง ไม่ว่าจะเป็นระบบบันทึกข้อมูลง่ายๆ เช่นการจัดเก็บหมายเลขโทรศัพท์ของลูกค้า หรือจัดเก็บโปรไฟล์ลูกค้า โดยระบบสมาชิก (Membership System)

จากการรวบรวมความหมายเกี่ยวกับระบบสารสนเทศเพื่อการจัดการจากผู้เขียนหนังสือด้านสารสนเทศได้แก่ ศิริลักษณ์ วิจารณ์กิจอำนวย. (2545) โอภาส เอี่ยมสกุล. (2560) Laudon & Laudon (2013). ได้อธิบายไว้ดังนี้ ระบบสารสนเทศเป็นระบบที่รวบรวม ประมวลผล จัดเก็บและแจกจ่ายข้อมูลต่างๆ ที่ใช้ในการตัดสินใจ องค์กรและธุรกิจ ต้องปรับตัวให้เข้ากับการเปลี่ยนแปลงของกระแสโลก หรือ

กระแสวิกฤต ดังนั้นองค์กรและธุรกิจ ต้องนำเทคโนโลยีสารสนเทศและการสื่อสาร ICT เข้ามา สนับสนุน การตัดสินใจการบริหารงาน ต้องพัฒนาระบบสารสนเทศและนำซอฟต์แวร์ มาใช้ในการบริหารงาน วาง แผนการจัดการและควบคุมองค์กรทางธุรกิจ เพื่อช่วยในการทำงานประจำวันและการวางแผนงานใน อนาคต โดย ธุรกิจการค้าสมัยใหม่ (Modern Trade and Business) จะมีรูปแบบการค้าที่ใช้เทคโนโลยี เข้ามาช่วยจัดการและดำเนินการ ในการดำเนินธุรกิจ เพื่อให้สามารถนำเสนอสินค้าและบริการให้แก่ ผู้บริโภคได้ทั้งในรูปแบบออนไลน์และออฟไลน์หรือแบบมีหน้าร้าน

โดยระบบสารสนเทศพื้นฐาน ประกอบไปด้วย (รูป 2.1)

1. บุคลากรผู้ใช้งานระบบ
2. ระบบการประมวลผลข้อมูล (Data Processing) ซึ่งช่วยในการเรียงและแบ่ง
ชั้นของข้อมูล
3. ระบบการสื่อสารข้อมูล (Data Communication)
4. ระบบจัดเก็บข้อมูลและการเข้าถึงข้อมูล เพื่อที่จะเก็บข้อมูลในรูปแบบที่
เหมาะสม และเพื่อให้แน่ใจว่าข้อมูลข่าวสารสามารถนำมาใช้ได้เมื่อต้องการ
5. ฮาร์ดแวร์และระบบรักษาความปลอดภัยและ การกำหนดสิทธิ์การเข้าถึง
ข้อมูล รวมทั้งกำหนดสิทธิ์ ผู้ใช้เข้า



รูปที่ 2.1 องค์ประกอบของระบบสารสนเทศพื้นฐาน

โดยการการพัฒนาเทคโนโลยีสารสนเทศยุคใหม่หรือปรับเปลี่ยนระบบงานเดิมที่มีอยู่แล้วให้สามารถทำงานเพื่อแก้ปัญหาการดำเนินงานทางธุรกิจที่มีการเปลี่ยนแปลงอย่างรวดเร็ว และตรงตามความต้องการของพฤติกรรมของผู้ใช้งานทั้งยุคเก่าและยุคใหม่นั้น มีแนวโน้มที่จะพัฒนาคอมพิวเตอร์ให้มีความสามารถใกล้เคียงกับมนุษย์ เช่น การเข้าใจภาษาสื่อสารของมนุษย์ โครงข่ายประสาทเทียม ระบบจำลอง ระบบเสมือนจริง เป็นต้น โดยใช้เทคโนโลยี เช่น เทคโนโลยีคลาวด์ (Cloud Computing) การวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Analysis) หรือ อินเทอร์เน็ตของสิ่ง (Internet of Things) เป็นต้น อย่างไรก็ตาม ผู้ที่เกี่ยวข้องจำเป็นต้องเข้าใจหลักการเบื้องต้นเกี่ยวกับ วงจรการวิเคราะห์ระบบงาน (System Development Cycle) เพื่อให้สามารถบริหารจัดการโครงการพัฒนาระบบสารสนเทศได้ อย่างสำเร็จและมีประสิทธิภาพ

2.2 วงจรการวิเคราะห์ระบบงาน

ในการพัฒนาระบบสารสนเทศนั้น มีขั้นตอนที่เป็นแนวทางในผู้จัดการโครงการพัฒนาระบบสารสนเทศและนักวิเคราะห์ระบบ ปฏิบัติงานได้โดยมีข้อบกพร่องน้อยที่สุด เพราะงานการ วิเคราะห์ระบบในปัจจุบันมีความซับซ้อน มากกว่าสมัยก่อน นักวิเคราะห์ระบบจึงต้องการมาตรฐานในการพัฒนาระบบดังกล่าว จึงได้มีการคิดค้นวงจรการพัฒนาระบบงานขึ้นเพื่อตอบสนองความต้องการของนักวิเคราะห์ระบบ (System Analysis: SA)

วงจรการพัฒนาระบบงานสำหรับระบบงานทั่วไป (System Development Model) (Boyde, 2014)

วงจรการพัฒนาระบบงาน (System Development Model) สำหรับระบบทั่วไปที่ได้มีการคิดค้นขึ้นมามีดังเดิมนั้นสามารถแบ่งออกเป็นลำดับขั้นได้ 4 ขั้นตอน คือ

1. การวิเคราะห์ระบบงาน เป็นขั้นตอนของการศึกษา ระบบงานเดิมที่ใช้ในปัจจุบัน (Current System) เพื่อคือ การหาความต้องการ (Requirements Gathering) ค้นหาปัญหาที่เกิดจากระบบงานเดิม ตลอดจนการศึกษาถึงความต้องการของธุรกิจ (Business Needs and Requirements) ว่าต้องการเพิ่มเติมอะไรเข้ามาในระบบ พร้อมกับการประเมินเหตุการณ์ต่าง ๆ เพื่อหาทางเลือกที่เหมาะสมมาแก้ปัญหา

2. การออกแบบและวางระบบงาน เป็นขั้นตอนหลังจาก การวิเคราะห์ระบบงาน เป็นการนำเอาความต้องการของระบบมาร่างออกแบบด้วยเครื่องมือการทำโมเดล (Modeling) ซึ่งเป็นขั้นตอนที่นักวิเคราะห์ระบบจะต้องวางโครงสร้าง ของระบบงาน ในรูปลักษณะที่ผู้เกี่ยวข้องในธุรกิจเข้าใจระบบงาน และในรูปลักษณะเฉพาะที่ใช้ในการพัฒนาระบบ ซึ่งจะเข้าใจโดย ทีมพัฒนาระบบเช่น โปรแกรมเมอร์ โดยมีการแจกแจงระบบงานย่อยของระบบที่ได้ออกแบบขึ้นจะถูกส่งต่อไปให้กับโปรแกรมเมอร์ เพื่อจะได้ทำการเขียนโปรแกรมให้เป็นระบบที่ปฏิบัติงานได้จริงในขั้นตอนต่อไป

3. การนำระบบเข้าสู่ธุรกิจหรือผู้ใช้ เป็นขั้น ตอนที่น่าเอาระบบงานมาติดตั้ง (Install) ให้กับผู้ใช้ ซึ่งอาจจะเป็นการติดตั้งแบบคู่ขนานกับระบบเดิมหรือ แทนที่ระบบเดิมโดยสิ้นเชิง หรือการติดตั้งทีละส่วนของระบบ และเพื่อให้แน่ใจว่าระบบงานสามารถปฏิบัติงานได้ตรงตามวัตถุประสงค์ ระบบงานจะ ต้องถูกทำการตรวจสอบมาอย่างดี พร้อมกับการฝึกอบรม (Education and Training) ให้ผู้ใช้ระบบสามารถใช้ระบบงานได้อย่างมีประสิทธิภาพและถูกต้อง

4. การดำเนินการสนับสนุนภายหลังการติดตั้งระบบงาน เป็น ขั้นตอนที่ระบบงานใหม่ได้ถูกนำมาติดตั้งแล้วผู้ใช้ระบบอาจจะยังไม่คุ้นเคย ในส่วนนั้นนอกจากติดตั้งระบบใช้งานแล้ว ยังต้องมีการจัดเตรียมขั้นตอนการสนับสนุนส่งเสริมการใช้งานให้สามารถใช้งานได้อย่างสมบูรณ์ โดย

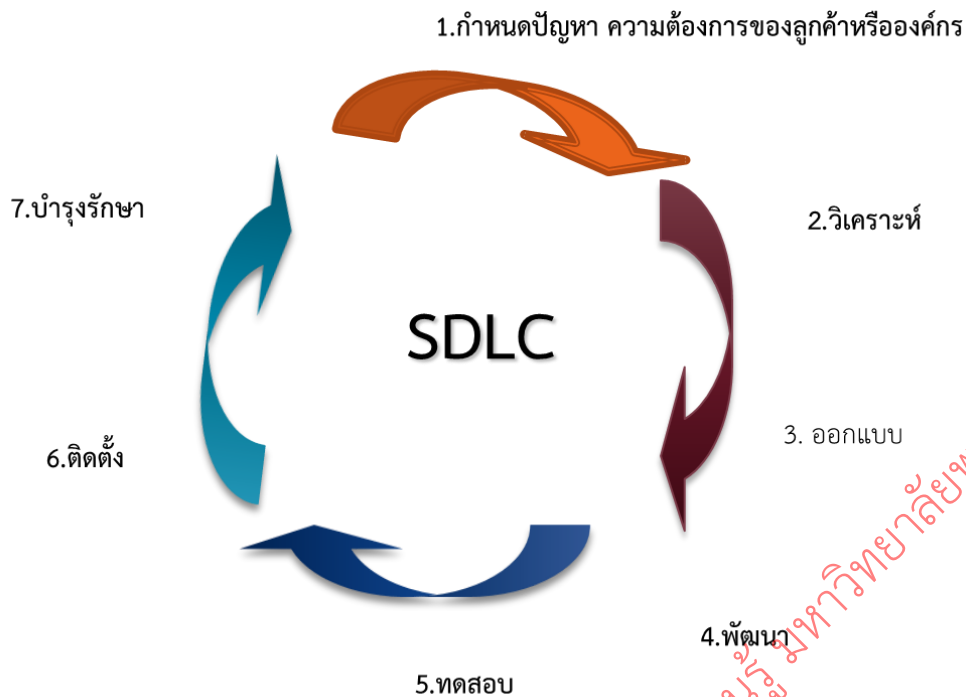
ทีมงานพัฒนาระบบจะต้องจัดทำคู่มือระบบ (User Manual) หลักสูตรฝึกอบรมผู้ใช้งาน (User Training) เอกสารประกอบระบบ (Documentation) เพื่อช่วยเหลือผู้ใช้งานในการปฏิบัติงาน ทั้งนี้รวมถึงเอกสารการบำรุงรักษาระบบงาน (System Maintenance) และเอกสารประกอบการปรับปรุงระบบงาน (System Improvement Documentation)

เมื่อความต้องการเปลี่ยนแปลงไป และระบบงานที่กำลังปฏิบัติอยู่เป็นประจำ จำเป็นต้องปรับปรุงใหม่ นักวิเคราะห์ระบบจะต้องกลับไปเริ่มต้นที่ขั้นที่ 1 ใหม่ และจะเป็นเช่นนี้เรื่อย ๆ ไป ถ้าเกิดความเปลี่ยนแปลงของระบบ

2.3 วงจรการพัฒนากระบวนการสำหรับระบบสารสนเทศ (Laudon & Laudon, 2013)

วงจรการพัฒนากระบวนการ (System Development Life Cycle: SDLC) ของระบบสารสนเทศ ได้มีการคิดค้นขึ้นมาโดยมีขั้นตอนที่แตกต่างไปจากวงจรการพัฒนากระบวนการสำหรับระบบงานทั่วไป ตรงที่มีขั้นตอนในการพัฒนาระบบงานที่ละเอียดกว่าถึง 7 ขั้นตอน ซึ่งนักวิเคราะห์ระบบต้องทำความเข้าใจว่าในแต่ละขั้นตอนว่าทำอะไรและทำอย่างไร สามารถแบ่งออกเป็นลำดับขั้นตอนดังนี้ คือ ภาพที่ 2.2 แสดงวงจรกิจกรรมต่าง ๆ ในแต่ละขั้นตอน ตั้งแต่ต้นจนเสร็จเป็นระบบงานที่ใช้ได้

1. ค้นหาปัญหา โอกาสและเป้าหมาย (Identifying Problems, Opportunity and Objective)
2. ศึกษาความเป็นไปได้ (Feasibility Study)
3. วิเคราะห์ความต้องการของระบบ (Analysing System Needs)
4. การออกแบบระบบ (Designing the Recommended System)
5. พัฒนาซอฟต์แวร์และจัดทำเอกสาร (Developing and Documenting Software)
6. ทดสอบและบำรุงรักษาระบบ (Testing and Maintaining the System)
7. ดำเนินงานและประเมินผล (Implementing and Evaluating the System)



รูปที่ 2.2 วงจรการพัฒนาแบบ SDLC (Laudon & Laudon, 2013)

2.4 รายละเอียดวงจรการพัฒนาระบบงาน

2.4.1 ขั้นตอนการค้นหาคำถาม โอกาสและเป้าหมาย (Identifying Problems, Opportunity and Objective) และ ศึกษาความเป็นไปได้ (Feasibility Study) โดยระบบสารสนเทศ จะเกิดขึ้นได้ก็ต่อเมื่อผู้บริหารหรือผู้ใช้ตระหนักว่าต้องการระบบสารสนเทศหรือและตัดสินใจว่าจะ พัฒนาสร้างระบบสารสนเทศใหม่หรือไม่โดยมีขั้นตอนดังนี้

2.4.1.1 นักวิเคราะห์และออกแบบระบบ ต้องศึกษาระบบโดยละเอียด เพื่อให้ เข้าใจถึงปัญหาที่เกิดขึ้นในองค์กร ตัวอย่างปัญหา เช่น

- บริษัทสะดวกซื้อสะดวกขายแห่งหนึ่งเปิดสาขาเพิ่มมากขึ้นจาก 10 สาขา เป็น 100 สาขา ระบบเดิมไม่ได้ครอบคลุมถึงการขยายตัวของบริษัท ไม่สามารถประมวลผล ระบบ POS (Point of Sale) ได้รวดเร็ว ทำให้ลูกค้าต้องรอคิวการชำระเงินนาน
- บริษัทสะดวกซื้อสะดวกขายแห่งหนึ่งเก็บข้อมูลผู้ขายได้เพียง 1,000,000 ราย แต่ปัจจุบันระบบนี้มีข้อมูลผู้ขาย 9,000,000 ราย และในอนาคตจะมีเกิน 10,000,000 ราย
- ระบบสารสนเทศในองค์กรหลาย ๆ แห่งในปัจจุบัน ที่ใช้มานานแล้ว และใช้เพื่อควบคุมการปฏิบัติงานรายวันเท่านั้น ไม่ได้มีจุดประสงค์เพื่อให้เป็นสารสนเทศเพื่อ

ตัดสินใจ หรือการวางแผนกลยุทธ์ ไม่มีระบบ เช่น Business Intelligence (BI) ทำให้ไม่สามารถใช้ข้อมูลเพื่อการวิเคราะห์ทางกลยุทธ์ได้ทันต่อสถานการณ์ทางธุรกิจ

2.4.1.2 ศึกษาความเป็นไปได้ (Feasibility Study) โดยที่ผู้บริหารจะตัดสินใจว่าจะดำเนินโครงการต่อไปหรือไม่หรือยกเลิกโครงการ จากการวิเคราะห์ ด้านความคุ้มค่า (Investment) ความเป็นไปได้ในเชิงเทคโนโลยี (Technology) และด้านเวลา (Timing)

2.4.1.3 นักวิเคราะห์และออกแบบระบบ ต้องมีความรู้ด้านธุรกิจนั้นพอสมควร ต้องมองเป้าหมายให้ชัดเจน ต้องคาดคะเนความต้องการของระบบและแนวทางแก้ไขปัญหาเพื่อจะได้รู้ทิศทางของการทำระบบให้เป็นไปตามเป้าหมายที่วางไว้ เช่น นักวิเคราะห์และออกแบบระบบต้องเข้าใจวิธีการการลงทุนในการผลิตสินค้า เข้าใจเป้าหมายในการนำระบบสารสนเทศ เข้าไปใช้ในการเก็บข้อมูลต่อกับวัตถุดิบและประมวลผลการส่งวัตถุดิบ เพื่อสร้างความได้เปรียบเชิงกลยุทธ์ขององค์กร

2.4.2 ศึกษาความเป็นไปได้ (Feasibility Study)

2.4.2.1 กำหนดว่าปัญหาคืออะไร และตัดสินใจว่าจะพัฒนาระบบสารสนเทศใหม่หรือการแก้ไขระบบสารสนเทศเดิมมี ความเป็นไปได้หรือไม่ โดยเสียค่าใช้จ่ายและเวลาน้อยที่สุด

2.4.2.2 นักวิเคราะห์และออกแบบระบบ ต้องกำหนดให้ได้ว่า การแก้ปัญหา นั้น

(1) ความเป็นไปได้ในเชิงเทคโนโลยี (Technology Feasibility) มีความเป็นไปได้ทางเทคนิคหรือไม่ เช่น จำนวนเครื่องคอมพิวเตอร์ที่มีอยู่เพียงพอหรือไม่ ซอฟต์แวร์แก้ไขได้หรือไม่

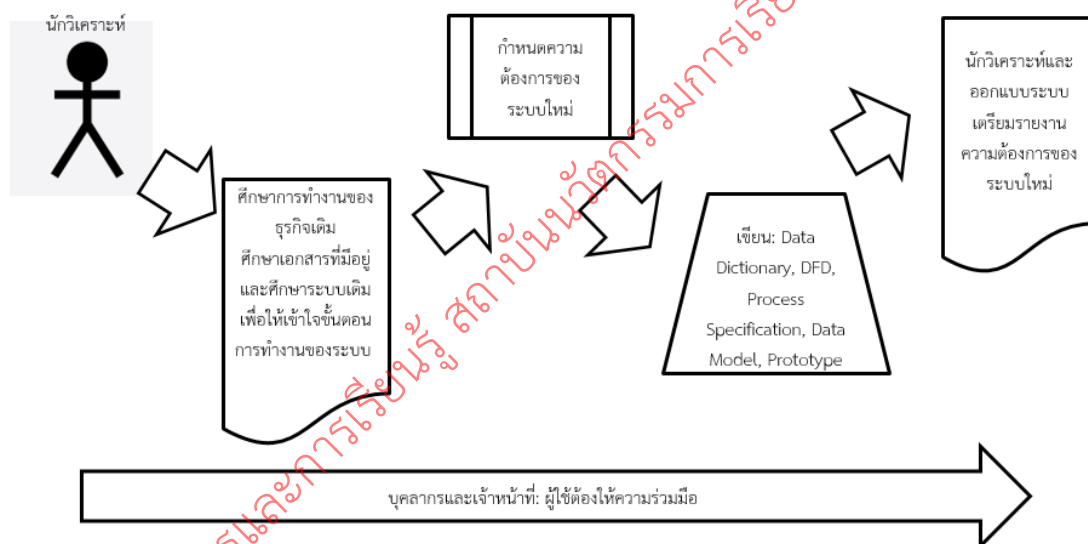
(2) ด้านทรัพยากรบุคคล (Human Resource) มีความเป็นไปได้ทางบุคลากรหรือไม่ เช่น มีบุคคลที่เหมาะสมที่จะพัฒนาและติดตั้งระบบหรือไม่ ผู้ใช้มีความคิดเห็นอย่างไรกับการเปลี่ยนแปลง

(3) ด้านความคุ้มค่า (Investment) มีความเป็นไปได้ทางเศรษฐศาสตร์หรือไม่ เช่น มีเงินลงทุนหรือไม่ค่าใช้จ่ายในการวิเคราะห์และออกแบบ ค่าใช้จ่ายในด้านเวลาที่ต้องใช้ในการพัฒนาระบบ ดังนั้นในการศึกษาความเป็นไปได้นั้นสามารถสรุปภาระงานได้ดังตาราง 2.1 คือ

ตารางที่ 2.1 ประเด็นภาระงานสำหรับนักวิเคราะห์และออกแบบระบบ

หน้าที่	กำหนดปัญหาและศึกษาว่าเป็นไปได้หรือไม่ที่จะเปลี่ยนแปลงระบบ กำหนดความต้องการที่แน่ชัด เพื่อใช้ในการวิเคราะห์ระบบ โดยที่ผู้บริหารจะเป็นผู้ตัดสินใจว่าจะดำเนินโครงการต่อไปหรือไม่หรือยกเลิกโครงการ
ผลลัพธ์	รายงานความเป็นไปได้ของการนำระบบใหม่มาใช้
เครื่องมือ	ต้องเก็บรวบรวมข้อมูลทั้งหมดที่จำเป็น เก็บรวบรวมข้อมูลของระบบและคาดคะเนความต้องการของระบบ เช่น Hardware Software ที่จะต้องจัดซื้อ
บุคลากรและหน้าที่รับผิดชอบ	ประเมินความพร้อมของทีมบุคลากรทุกด้าน ทั้งด้านเทคนิคและด้านปฏิบัติงาน

2.4.3 วิเคราะห์ความต้องการของระบบ (Analysing System Needs)



รูปที่ 2.3 ขั้นตอนการวิเคราะห์ความต้องการของระบบ

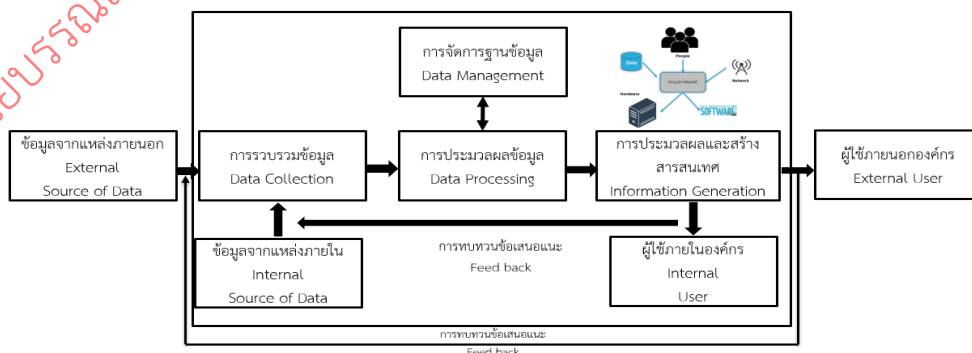
จากรูป 2.3 อธิบายวัตถุประสงค์ของขั้นตอนการวิเคราะห์คือการกำหนดและกำหนดความต้องการของระบบให้เป็นแบบแผน ซึ่งทำได้โดยการกำหนดสิ่งที่ระบบจะต้องถูกพัฒนาตามข้อกำหนดและความคาดหวังของผู้ใช้ โดยนักวิเคราะห์จะเริ่มตั้งแต่ศึกษาการทำงานของธุรกิจเดิม ศึกษาเอกสารที่มีอยู่ และศึกษาระบบเดิมเพื่อให้เข้าใจขั้นตอนการทำงานของระบบ ว่าทำงานอย่างไร รวมทั้งศึกษาเอกสารที่มีอยู่และศึกษาระบบเดิมเพื่อให้เข้าใจขั้นตอนการทำงานของระบบ หลังจากนั้นจะกำหนดความต้องการของระบบใหม่ และนักวิเคราะห์จะต้องสร้างและเขียนเอกสารเพื่อใช้เป็น

เครื่องมือสื่อสารในโครงการพัฒนาระบบสารสนเทศ เช่น เอกสาร Data Dictionary, Data Flow Diagram(DFD) Entity Diagram(ER) Process Specification Data Model Prototype เป็นต้น ซึ่งในขั้นตอนนี้บุคลากรและเจ้าหน้าที่ รวมถึงผู้ใช้จะต้องให้ความร่วมมือ ในการจัดทำเอกสารร่วมกัน

หลังจากนั้นนักวิเคราะห์และออกแบบระบบเตรียมรายงานความต้องการของระบบ ใหม่ เขียนขั้นตอนการวิเคราะห์โครงการ รายงานแสดงผลลัพธ์ที่สำคัญสามประการคือ รายงานความต้องการทางธุรกิจ รายงานแผนการออกแบบระบบสารสนเทศ และเอกสารกลยุทธ์ระดับสูงสำหรับ กระบวนการทั้งหมด

2.4.4. ขั้นตอนการออกแบบระบบ (Designing the Systems Process)

ในขั้นตอนนี้ การออกแบบระบบ (Designing the Systems) นักวิเคราะห์และออกแบบระบบต้องนำผลลัพธ์ที่ได้จากศึกษาความเป็นไปได้ (Feasibility Study) และวิเคราะห์ความต้องการของระบบ (Analysing System Needs) มาพัฒนาเป็นรูปแบบทางกายภาพ (Physical Model) ดังแสดงในรูป 2.4 องค์ประกอบระบบที่ต้องออกแบบ โดยเริ่มจากการออกแบบงานทางด้านฮาร์ดแวร์ และซอฟต์แวร์ ทั้งในส่วนนำข้อมูลเข้า (Input) ทั้งข้อมูลจากแหล่งภายนอกองค์กร (External Source of Data) และข้อมูลจากแหล่งภายใน (Internal Source of Data) ส่วนประมวลผลข้อมูล (Data Processing) ส่วนแสดงผลลัพธ์ (Output) ส่วนจัดเก็บข้อมูล (Storage) การออกแบบจำลองข้อมูล การออกแบบรายงานและช่องทางการสื่อสารกับผู้ใช้ การทบทวนข้อเสนอแนะ (Feed back) ทั้งภายในและภายนอกองค์กร นักวิเคราะห์และออกแบบระบบ ต้องตัดสินใจเลือกฮาร์ดแวร์และซอฟต์แวร์ และออกแบบระบบออกแบบข้อมูลเข้า รายงานการแสดงผลบนหน้าจอออกแบบฐานข้อมูล รวมทั้งออกแบบระบบกำหนดจำนวนบุคลากรในระบบ



รูปที่ 2.4 องค์ประกอบระบบที่ต้องออกแบบ (Systems' Components for Designing)

2.4.5 ขั้นตอนการพัฒนาซอฟต์แวร์และจัดทำเอกสาร (Developing and Documenting Software) ซึ่งรวมถึง การเขียนโปรแกรม การดีไซน์ Functional Software การออกแบบ UX/UI การออกแบบ Workflow และการออกแบบ Backend/Frontend รวมถึงจัดทำคู่มือการใช้โปรแกรม และฝึกอบรมผู้ใช้ที่เกี่ยวข้องในระบบ โดยนักวิเคราะห์และออกแบบระบบต้องจัดเตรียมสถานที่และการติดตั้งเครื่องคอมพิวเตอร์ ดูแลโปรแกรมเมอร์ในการเขียนโปรแกรม และการเขียนคู่มือการใช้โปรแกรมและการฝึกอบรม

2.4.6 ขั้นตอนการทดสอบและบำรุงรักษาระบบ (Testing and Maintaining the System) ในขั้นตอนนี้ นักวิเคราะห์และออกแบบระบบและทีมงานทดสอบโปรแกรม จะมีการทำการทดสอบคุณภาพของระบบ และตัวซอฟต์แวร์ว่าสามารถใช้งานจริงหรือไม่ ซึ่งการทดสอบโปรแกรมหากไม่เป็นไปตามความต้องการ อาจต้องแก้ไขปรับปรุงใหม่ หลักการทดสอบที่นิยมใช้คือ User Acceptance Testing (UAT), System Integration Testing (SIT), และ การทำ Unit Test ที่เป็น Functional Test และการทำ Non-functional Test รวมถึงการทำ Performance Test, Penetration Test และ Smoke Test เป็นต้น โดยผู้ใช้งานจะต้องตรวจสอบว่าโปรแกรมทำงานตามที่ต้องการ ส่วนนบำรุงรักษา ส่วนใหญ่เป็นการแก้ไขโปรแกรมหลังจากใช้งานแล้ว เนื่องจากมีปัญหาในโปรแกรม (Bug) หรือหากธุรกิจเกิดการเปลี่ยนแปลงไป เช่น ธุรกิจขยายตัว ธุรกิจสร้างสินค้าตัวใหม่ความต้องการของระบบก็เพิ่มขึ้น รายงานเพิ่มขึ้น หรือความต้องการผู้ใช้งานเพิ่มขึ้น โดย 40-60 % ของค่าใช้จ่ายในการพัฒนาระบบจะเกิดจากการบำรุงรักษาระบบ

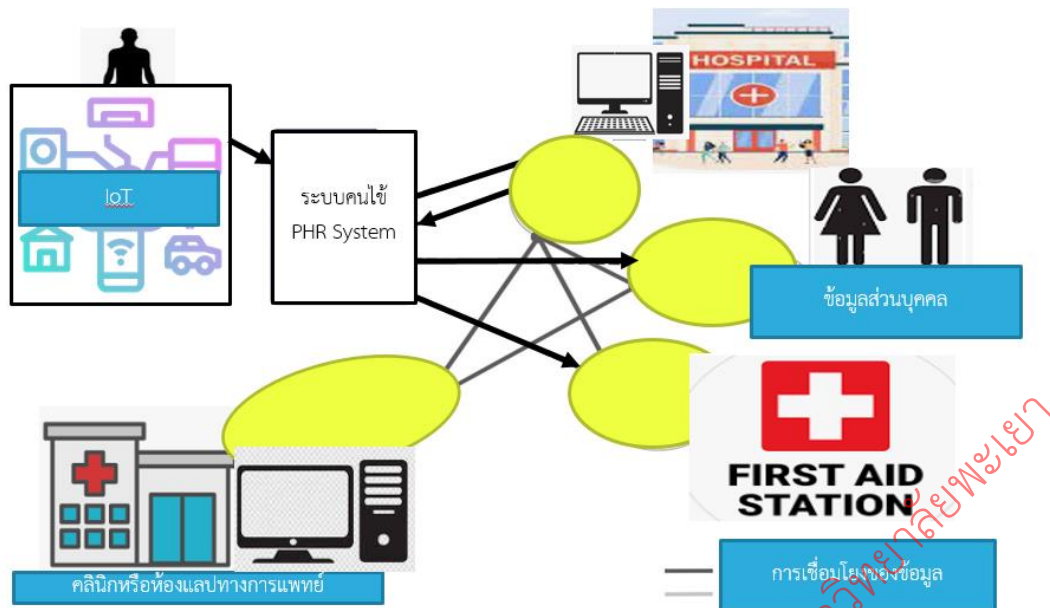
2.4.7 ขั้นตอนการดำเนินงานและประเมิน (Implementing and evaluating the System) หลังจากติดตั้งระบบให้พร้อมใช้ และนำระบบใหม่มาใช้แทนระบบเดิม หรือใช้ระบบใหม่ควบคู่กับระบบเดิมสักระยะหนึ่ง นักวิเคราะห์และออกแบบระบบจะต้องทำการประเมินผล ทำการควบคุมและการประเมินผลผลลัพธ์ว่าตรงกับที่วางไว้หรือไม่ ซึ่งจัดเป็นกระบวนการขั้นสุดท้ายที่ต้องทำในการพัฒนาระบบสารสนเทศ การดำเนินงานโครงการพัฒนาระบบ หากขาดซึ่งการควบคุมและประเมินผลแล้วย่อมไม่สามารถ บ่งบอกถึงความสำเร็จของการทำงานนี้ได้ การประเมินรวมถึงการวิเคราะห์ปัญหาและอุปสรรคที่ประสพรวมถึงการหาวิธีในการแก้ไขปัญหาที่เกิดขึ้น เพื่อให้ทราบถึงความพอใจของผู้ใช้ระบบ หรือสิ่งที่ต้องแก้ไขปรับปรุง หรือปัญหาที่พบ

2.5 แนวทางที่ใช้ในการวิเคราะห์และออกแบบระบบสารสนเทศสมัยใหม่

ในยุคก่อนปี ค.ศ. 2000 การวิเคราะห์และออกแบบระบบสารสนเทศโดยปกติแล้ว จะมีแผนกแยกต่างที่รับผิดชอบในการพัฒนา ทดสอบ และติดตั้ง ชัดเจน ทำให้การส่งมอบโครงการที่ประสบความสำเร็จต้องอาศัยการทำงานร่วมกันระหว่างทีมหรือแผนกอย่างมีประสิทธิภาพ ซึ่งอาจเป็นเรื่องยาก

ในยุคปัจจุบันที่จะบรรลุผลได้ เพราะการพัฒนามีการจ้างงานบางส่วนจากบริษัทภายนอก (Outsource) จึงจำเป็นต้องมีวิธีการเทคนิค เครื่องมือ และมีขั้นตอนการดำเนินงานที่แน่นอนเป็นระบบ เพื่อให้ได้ระบบสารสนเทศที่มีประสิทธิภาพ สามารถตอบสนองต่อวัตถุประสงค์หลักขององค์กรได้ ดังนั้นการวิเคราะห์และออกแบบระบบ (Systems analysis and design :SAD) ระบบสารสนเทศสมัยใหม่ จึงเป็นกระบวนการเชิงโครงสร้าง (Marakas, G. M., & O'Brien, 2013) ที่ใช้สำหรับพัฒนาระบบสารสนเทศ (Information System) เพื่อนำไปใช้ในการแก้ปัญหาขององค์กร

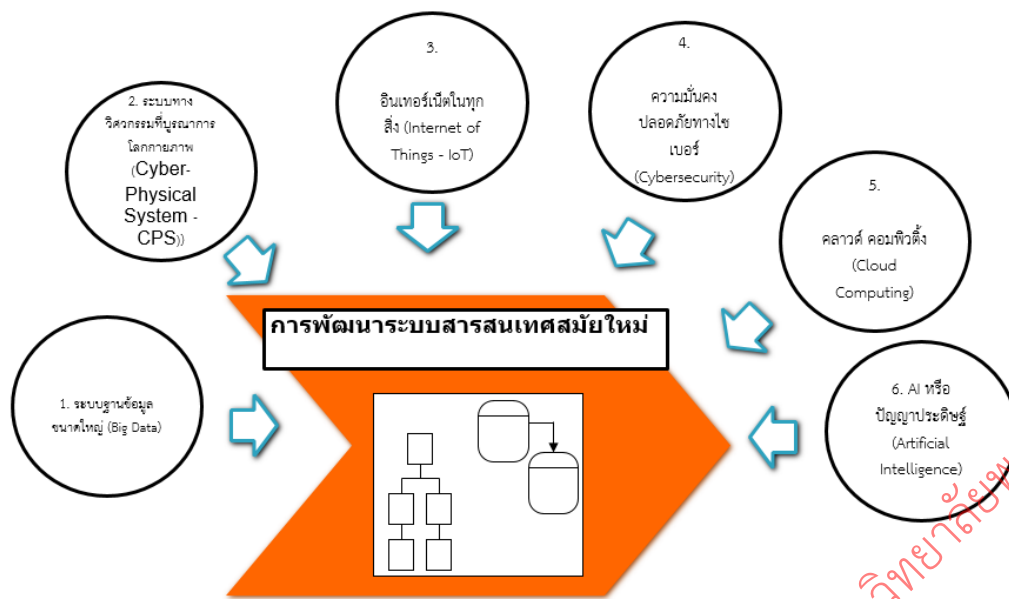
ในช่วงเริ่มต้นของการปฏิวัติอุตสาหกรรมระหว่างปี ค.ศ. 1760–1830 ระบบพื้นฐานของโรงงานผลิตส่วนใหญ่ขึ้นอยู่กับพลังของน้ำและเครื่องยนต์ที่ขับเคลื่อนด้วยไอน้ำ และการปฏิวัติอุตสาหกรรมครั้งที่สองระหว่างปี ค.ศ.1870–1914 การผลิตมุ่งเน้นไปที่การใช้ระบบสายการประกอบ (Assembly Lines) และระบบการผลิตที่ใช้ระบบอัตโนมัติและระบบไฟฟ้า การปฏิวัติอุตสาหกรรมครั้งที่สาม (The Third Industrial Revolution) เริ่มต้นขึ้นในปีค.ศ. 1950 โดยมีเทคโนโลยีการผลิตที่เปลี่ยนผ่านจากเทคโนโลยีอนาล็อก (Analog) เป็นดิจิทัล (Digital) และมีการพัฒนาของการนำเทคโนโลยีสารสนเทศ (Information and Communication Technology) เข้าสู่ระบบการผลิต และส่งผลให้ไปสู่การปฏิวัติอุตสาหกรรมครั้งที่สี่ (The Fourth Industrial Revolution) ซึ่งต่างจากการปฏิวัติครั้งที่ผ่านๆมา การปฏิวัติอุตสาหกรรมครั้งที่ 4 นำไปสู่จุดเริ่มต้นของการผลิตโดยใช้ระบบไซเบอร์ (Cyber-Physical System) ใช้ระบบปัญญาประดิษฐ์ (Artificial System) ระบบการเรียนรู้ของเครื่อง (Machine Learning) มีการนำเอาศาสตร์ด้านการวิเคราะห์ข้อมูลหรือวิทยาการข้อมูล ซึ่งเกิดจากการบูรณาการระหว่างสามสาขาวิชาคือ หนึ่งคณิตศาสตร์และสถิติ สองวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ และสามความรู้ในสาขาวิชาต่างๆ เช่น ถ้าเอาความรู้ทางการแพทย์มารวมก็จะกลายเป็น Medical Business Analytics ตัวอย่างเช่นการวิเคราะห์สาเหตุของโรคจากฐานข้อมูลผู้ป่วยขนาดใหญ่ ถ้าเอาความรู้ทางชีววิทยามารวมก็กลายเป็น Bioinformatics เป็นต้น โดยหลักการออกแบบของระบบสารสนเทศสมัยใหม่จะเป็นลักษณะแบบการกระจายอำนาจ เชื่อมโยงด้วยโครงข่าย มีความโปร่งใสของข้อมูลแบบเรียลไทม์ด้วยมาตรฐานที่แลกเปลี่ยนกันได้แบบทันที มีความถูกต้อง รวดเร็ว ผู้ใช้ (User) สามารถเข้าถึงได้ทุกที่ทุกเวลาโดยการใช้เครื่องคอมพิวเตอร์ Tablet หรือ อุปกรณ์ Smart Devices และระบบวิทยาการข้อมูลช่วยให้วิเคราะห์ข้อมูล โดยเฉพาะข้อมูลขนาดใหญ่ที่มีความหลากหลายและนำไปสู่ความรู้ที่นำไปใช้งานได้จริง (Actionable knowledge) ดังรูป 2.5



รูปที่ 2.5 ระบบฐานข้อมูลขนาดใหญ่ของระบบบันทึกข้อมูลสุขภาพผู้ป่วย (PHR - Personal Health Record) (Roehrs et.al, 2017)

จากรูป 2.5 หากมีการนำแนวทางการออกแบบระบบสารสนเทศสมัยใหม่ เชิงข้อมูลขนาดใหญ่ (Big Data) จะส่งผลดีในระยะยาว คือ

- การนำเทคนิคการวิเคราะห์ ข้อมูลขนาดใหญ่ (Big Data) ด้านสุขภาพในระดับมหภาค จะช่วยแจ้งเตือนสถานการณ์โรคระบาด แจ้งเตือนให้คนไข้เข้ารับการรักษาที่ถูกต้องตามกำหนด และจัดแพทย์ผู้เชี่ยวชาญที่เหมาะสม เพื่อประเมินจุดอ่อนและข้อบกพร่องต่างๆ ที่เกิดขึ้นระหว่างการรักษา
- สามารถประเมินอนาคตได้อย่างแม่นยำ ในการพยากรณ์โรคที่มีแนวโน้มที่จะเกิดขึ้นในอนาคต ตลอดจนกำหนดแนวทางการรักษาที่เหมาะสม
- การเชื่อมโยงแลกเปลี่ยนข้อมูลสุขภาพ จะช่วยในกรณีฉุกเฉิน โดยที่ผู้ป่วยสามารถถูกส่งต่อและทางผู้ให้บริการสามารถแลกเปลี่ยนข้อมูลได้อย่างรวดเร็ว สามารถช่วยให้การให้บริการทางการแพทย์และการดูแลสุขภาพเป็นไปได้อย่างมีประสิทธิภาพและทันที่



รูปที่ 2.6 รูปเทคโนโลยีที่สำคัญ สำหรับการออกแบบสารสนเทศสมัยใหม่ (Agrawal S., Sahu A., & Kumar G., 2022)

จากรูป 2.6 ระบบสารสนเทศสมัยใหม่จะมี เทคโนโลยีที่สำคัญ คือ

1. ระบบฐานข้อมูลขนาดใหญ่ (Big Data System) หมายถึงข้อมูลที่มีขนาดใหญ่ มีปริมาณข้อมูลมหาศาล ทั้งแบบข้อมูลที่มีโครงสร้างและไม่มีโครงสร้าง เกิดขึ้นจากการทำธุรกิจในแต่ละวัน มีการบันทึกเร็ว หรือซับซ้อนจนยากหรือเป็นไปไม่ได้ที่จะประมวลผลโดยใช้วิธีการแบบเดิม เช่น ข้อมูลที่บันทึกจากระบบ Point of Sales (POS) ของร้านสะดวกซื้อ หรือ Modern Trade จำนวน 10,000 สาขา ซึ่งในปัจจุบัน ฐานข้อมูลขนาดใหญ่ (Big Data) ได้กลายเป็นทุนหรือทรัพย์สินไปแล้ว และในกรณีของร้านสะดวกซื้อ หรือ Modern Trade ทางบริษัทที่ใช้ข้อมูลขนาดใหญ่เพื่อการวิเคราะห์และประมวลผลเพื่อเพิ่มมูลค่าของข้อมูลให้สูงขึ้นไปอีก และด้วยความก้าวหน้าทางเทคโนโลยีในปัจจุบัน ส่งเสริมให้ราคาของอุปกรณ์การเก็บข้อมูลและคอมพิวเตอร์ลงแบบก้าวกระโดด ทำให้การเก็บข้อมูลง่ายและมีราคาถูก การเก็บข้อมูลและการเข้าถึงข้อมูลขนาดใหญ่สามารถทำได้ง่ายและมีราคาถูก การออกแบบระบบสารสนเทศสมัยใหม่ซึ่งต้องการตัดสินใจด้านธุรกิจมีความแม่นยำและถูกต้องมากขึ้น จึงจำเป็นต้องใช้ระบบฐานข้อมูลขนาดใหญ่ (Big Data)

2. ระบบทางวิศวกรรมที่บูรณาการโลกกายภาพ (Cyber-Physical System - CPS) คือ ระบบสารสนเทศที่ประกอบด้วยเทคโนโลยีที่เกี่ยวข้องทั้งหลายเช่น Internet of Things (IoT), ระบบฝังตัว (Embedded Systems), ระบบควบคุม (Control Systems), การประมวลผลด้วยเครือข่ายคลาวด์ (Cloud Computing) และ การวิเคราะห์ขั้นสูง (Data Analytics) โดยอุปกรณ์ เครื่องจักร วัสดุ สภาพแวดล้อม หรือสิ่งต่างๆที่จับต้องได้ รวมทั้งมนุษย์ จะถูกเชื่อมโยงเข้ากับระบบดิจิทัล (Digital) ที่

เชื่อมโยงด้วยระบบเครือข่ายคอมพิวเตอร์และข้อมูล สื่อสารแลกเปลี่ยนข้อมูลระหว่างกัน โดยระบบสามารถทำงานร่วมกันได้ สามารถตรวจสอบและสามารถควบคุมเพื่อปรับเปลี่ยนสถานะของระบบตามความเหมาะสมกับสภาพแวดล้อม เช่น โรงเรือนอัจฉริยะ (Smart Farming) โรงงานอัจฉริยะ (Smart Factory) หรือ การผลิตอัจฉริยะ (Smart Manufacturing) เป็นต้น

3. อินเทอร์เน็ตในทุกสิ่ง (Internet of Things - IoT) หมายถึงระบบสารสนเทศที่ อุปกรณ์ต่างๆ (Hardware) สิ่งต่างๆ (Things) ได้ถูกเชื่อมโยงผ่านระบบอินเทอร์เน็ตหรือเครือข่าย ทำให้ผู้ใช้สามารถสั่งการควบคุมการใช้งานอุปกรณ์ แบ่งปันข้อมูลต่างๆ ผ่านแอปพลิเคชัน (Application) ทางเครือข่ายอินเทอร์เน็ตอุปกรณ์และเครื่องจักรที่เชื่อมต่อ IoT ระบบไอโอที (IoT) สามารถปรับปรุงวิธีการทำงานและการใช้ชีวิตของเรา เช่น การสั่งการปิดเปิดไฟฟ้าภายในบ้านด้วยการเชื่อมต่ออุปกรณ์ควบคุม เช่น มือถือ ผ่านทางอินเทอร์เน็ต) การสั่งการปิดเปิดเครื่องมือทางการเกษตร เช่นระบบปั๊มน้ำ โรงเรือนอีแวป (Evaporative cooling system- Evap) ช่วยลดอุณหภูมิภายในโรงเรือนสำหรับการปลูกพืช ผัก ผลไม้ ที่ต้องการดูแลเป็นพิเศษ และระบบให้ปุ๋ย เป็นต้น

4. ความมั่นคงปลอดภัยทางไซเบอร์ (Cyber Security) คือ การออกแบบระบบสารสนเทศสมัยใหม่จะต้องคำนึงถึงเทคโนโลยี กระบวนการและวิธีปฏิบัติด้านความปลอดภัยของระบบ เพื่อปกป้องเครือข่าย อุปกรณ์ โปรแกรมและข้อมูลจากการโจมตี ความเสียหายหรือการเข้าถึงจากบุคคลที่สามโดยไม่ได้รับอนุญาต การปกป้องข้อมูลจากการโจรกรรมของแฮกเกอร์ ตลอดจนการรั่วไหลของข้อมูล ที่อาจจะเกิดขึ้น และสร้างความเสียหายให้กับองค์กรและชื่อเสียงได้ ซึ่งระบบจะต้องออกแบบครอบคลุม

4.1 Network Security: กระบวนการและวิธีปฏิบัติด้านความปลอดภัย ปกป้องเครือข่าย หรือเน็ตเวิร์คจากบุคคลที่สามโดยไม่ได้รับอนุญาต

4.2 Application Security: กระบวนการและวิธีปฏิบัติด้านแอปพลิเคชันที่ต้องการการอัปเดตและการทดสอบอย่างต่อเนื่องเพื่อให้แน่ใจว่า แอปพลิเคชัน (Application) ปลอดภัยจากการโจมตีของบุคคลที่สาม

4.3 Data Security: กระบวนการและวิธีปฏิบัติด้านการรักษาความปลอดภัยข้อมูล การปกป้องข้อมูล ตลอดจนการรั่วไหลของข้อมูล ที่อาจจะเกิดขึ้นกับบริษัท และลูกค้า

5. คลาวด์คอมพิวติ้ง (Cloud Computing) คือการประมวลผลแบบคลาวด์ ผู้ใช้ระบบจะใช้ซอฟต์แวร์ (Software) แอปพลิเคชัน (Application) ระบบ (System) และทรัพยากรของเครื่องคอมพิวเตอร์ของผู้ให้บริการผ่านอินเทอร์เน็ต ด้วยการประมวลผลแบบคลาวด์ผู้ใช้สามารถเข้าถึง ข้อมูลถึงไฟล์ และใช้แอปพลิเคชันจากอุปกรณ์ใดก็ได้ที่สามารถเข้าถึงอินเทอร์เน็ตได้ บริการประมวลผลแบบ

คลาวด์ครอบคลุมถึงการประมวลผล การจัดเก็บข้อมูล และระบบออนไลน์ต่างๆจาก ผู้ให้บริการ เพื่อลดความยุ่งยากในการติดตั้ง ดูแลระบบ ช่วยประหยัดเวลา และลดต้นทุนในการสร้างระบบคอมพิวเตอร์ และ เครือข่ายเอง ซึ่งก็มีทั้งแบบบริการฟรีและแบบเก็บเงิน เช่น ระบบ Gmail Google Sheet Google Doc Google Data Studio ของ Google ผู้ใช้สามารถเข้าถึงเอกสาร ไฟล์ ข้อมูลที่ Google ให้บริการผ่านอินเทอร์เน็ตจากอุปกรณ์ใดก็ได้

6. AI หรือปัญญาประดิษฐ์ (Artificial Intelligence) ระบบสารสนเทศสมัยใหม่ที่เทคโนโลยีที่ถูกออกแบบให้มีระบบทำงานเหมือนกับการทำงานของสมองมนุษย์ โดย AI ที่ผู้รู้จักกันดีคือระบบการขับรถยนต์อัตโนมัติ (Autonomous Vehicles) ของบริษัทเทสลา (Tesla) รถยนต์ไฟฟ้าและพลังงานสะอาดสัญชาติอเมริกัน ระบบจดจำใบหน้า (Face Recognition) สำหรับการเข้าใช้ Smart Phone และ ระบบ AI บนโปรแกรมคอมพิวเตอร์ Siri ของ Apple ระบบ AI จะเรียนรู้มาเพื่อตอบสนองไปตามอัลกอริทึมของระบบ เทคนิคของ AI ได้ถูกนำมาประยุกต์ในหลายศาสตร์ เช่น Recommendation System ของ Youtube ที่ทำการแนะนำเนื้อหาที่ผู้ใช้น่าจะถูกใจ หรือระบบ Google Translation ที่แปลภาษาและมีการเรียนรู้ปรับการแปล อยู่ตลอดเวลา เป็นต้น

สรุปเนื้อหาบทที่ 2 "การพัฒนาและวิเคราะห์ระบบสารสนเทศสมัยใหม่" เป็นการแนะนำเกี่ยวกับกระบวนการที่ใช้ในการพัฒนาและวิเคราะห์ระบบสารสนเทศในยุคสมัยใหม่เช่น วงจรการวิเคราะห์ระบบงาน โดยเน้นไปที่การใช้เทคโนโลยีสารสนเทศในการปรับปรุงระบบสารสนเทศให้เหมาะสมกับความต้องการและการเปลี่ยนแปลงของธุรกิจในปัจจุบัน บทที่ 2 นำเสนอขั้นตอนและข้อคิดเพื่อการพัฒนากระบวนการที่มีประสิทธิภาพ ตั้งแต่การวางแผนและวิเคราะห์ความต้องการ: การเริ่มต้นด้วยการรับรู้และเข้าใจความต้องการของธุรกิจและผู้ใช้งาน แนวโน้มในตลาด และปัญหาที่ต้องการแก้ไข นี่จะช่วยในการกำหนดความต้องการและเป้าหมายของระบบสารสนเทศ การออกแบบระบบ: ขั้นตอนนี้เก็บรวบรวมความต้องการและวางแผนการออกแบบระบบ รวมถึงโครงสร้างของฐานข้อมูล การออกแบบตัวแปรและระบบโมดูลต่างๆ และการกำหนดการทำงานของระบบ การพัฒนาและการสร้างระบบ: ในขั้นตอนนี้ทำการเขียนโปรแกรม สร้างฐานข้อมูล และสร้างส่วนประสานกับผู้ใช้ ระหว่างกระบวนการนี้ทดสอบและปรับปรุงระบบเป็นส่วนสำคัญ การทดสอบและการประเมิน: ระหว่างการพัฒนาจะต้องมีการทดสอบระบบเพื่อตรวจสอบความถูกต้องและประสิทธิภาพ หากพบข้อผิดพลาดจะต้องแก้ไขและทดสอบอีกครั้ง การประเมินคือการตรวจสอบว่าระบบสามารถใช้งานได้ตามความต้องการหรือไม่ การนำเสนอและการใช้งาน: เมื่อระบบสารสนเทศได้รับการอนุมัติและผ่านการทดสอบเพียงพอ ก็จะนำเสนอให้กับผู้ใช้งาน เพื่อให้พวกเขารู้จักกับระบบ และมีการฝึกใช้งานและการสนับสนุน การดูแลและพัฒนาต่อเนื่อง: เมื่อระบบเริ่มทำงาน จะต้องมีการดูแลรักษาและดำเนินการปรับปรุงตามความต้องการและการเปลี่ยนแปลงในองค์กร นอกจากนี้ยังต้องพัฒนาระบบต่อเนื่องเพื่อให้เข้ากับ

การเปลี่ยนแปลงในธุรกิจและเทคโนโลยีใหม่ๆ โดยทั้งหมดนี้เป็นกระบวนการที่ทำซ้ำไปเรื่อยๆ เพื่อให้ระบบสารสนเทศสามารถประสบความสำเร็จในการให้การสนับสนุนแก่ธุรกิจและการตัดสินใจทางธุรกิจอย่างมีประสิทธิภาพ

คำศัพท์ บทที่ 2

Analysing System Needs	วิเคราะห์ความต้องการของระบบ
Application Security	ความปลอดภัยด้านแอปพลิเคชัน
Artificial Intelligence System	ระบบปัญญาประดิษฐ์
Big Data System	ระบบฐานข้อมูลขนาดใหญ่
Business Needs and Requirements	วิเคราะห์ความต้องการของธุรกิจ
Cloud Computing	การประมวลผลด้วยเครือข่ายคลาวด์
Cyber Security	ความมั่นคงปลอดภัยทางไซเบอร์
Cyber-Physical System - CPS	ระบบทางวิศวกรรมที่บูรณาการโลกกายภาพ
Data Communication	การสื่อสารข้อมูล
Data Processing	การประมวลผลข้อมูล
Data Security	ความปลอดภัยด้านการรักษาความปลอดภัยข้อมูล
Developing and Documenting Software	พัฒนาซอฟต์แวร์และจัดทำเอกสาร
Documentation	เอกสารประกอบระบบ
Education and Training	การฝึกอบรม
External Source of Data	ข้อมูลจากแหล่งภายนอกองค์กร
Feasibility Study	ศึกษาความเป็นไปได้
Feed back	การทบทวนข้อเสนอแนะ
Identifying Problems, Opportunity and Objective	ค้นหาปัญหา โอกาสและเป้าหมาย
Implementing and evaluating the System	ดำเนินงานและประเมินผล
Information System Development	การพัฒนาและวิเคราะห์ระบบสารสนเทศสมัยใหม่
Internet of Things	อินเทอร์เน็ตในทุกสิ่ง
Membership System	ระบบสมาชิก
Modelling	การทำโมเดล
Network Security	ความปลอดภัยด้านเครือข่าย
Output	ส่วนแสดงผลลัพธ์

Physical Model	รูปแบบทางกายภาพ
Storage	ส่วนจัดเก็บข้อมูล
System Development Cycle	วงจรการวิเคราะห์ระบบงาน
Testing and Maintaining the System	ทดสอบและบำรุงรักษาระบบ
User manual	คู่มือระบบ

คำถามและแบบฝึกหัดท้ายบทที่ 2

1. จงอธิบายวงจรการพัฒนาระบบ คืออะไร
2. จงอธิบายหน้าที่ความรับผิดชอบของนักวิเคราะห์ระบบมีอะไรบ้าง
3. จงอธิบายนักวิเคราะห์ระบบคือใคร และควรมีคุณสมบัติใดบ้าง
4. จงอธิบาย Requirements คืออะไร และสำคัญต่อการพัฒนาระบบงานอย่างไร
5. จงอธิบายโมเดล หรือแบบจำลองการพัฒนาซอฟต์แวร์ คืออะไร
6. จงอธิบาย SDLC มีขั้นตอนการทำงานใดบ้าง
7. จงอธิบายเทคโนโลยีที่สำคัญ ของระบบสารสนเทศสมัยใหม่
8. จงอธิบายองค์ประกอบของความมั่นคงปลอดภัยทางไซเบอร์ (Cyber Security)

บทที่ 3 ความรู้เบื้องต้นเกี่ยวกับกรอบแนวทางการวางแผนสถาปัตยกรรมระบบ

(Introduction to Modern Information System Architecture)

การออกแบบและวางแผนปฏิรูปกระบวนการสารสนเทศในองค์กรให้รองรับ กระบวนการทำงานทางธุรกิจ (Business processes) ให้บรรลุเป้าหมายที่องค์กรตั้งไว้ แต่ปัญหาอยู่ที่ธุรกิจส่วนใหญ่ไม่มีการปฏิรูปกระบวนการทำงาน มีเพียงการนำเทคโนโลยีใหม่ ๆ ที่ทันสมัยมาสนับสนุนกระบวนการทำงานที่เป็นรูปแบบเดิม ๆ จึงส่งผลให้การลงทุนด้านเทคโนโลยีไม่คุ้มค่าเท่าที่ควร ดังนั้นการออกแบบเทคโนโลยีสารสนเทศโดยองค์รวมขององค์กรในทุกมิติจึงมีความสำคัญมาก

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจความหมายของสถาปัตยกรรมองค์กร (Enterprise Architecture: EA) สำหรับเป็นกรอบแนวทางการวางแผนสถาปัตยกรรมองค์กรในอนาคตด้านสารสนเทศ
2. สามารถเข้าใจกรอบแนวคิดและหลักการในการพัฒนาสำหรับระบบสารสนเทศสมัยใหม่ (TOGAF)

3.1 สถาปัตยกรรมองค์กร (Enterprise Architecture: EA) คืออะไร

สถาปัตยกรรมองค์กร (Enterprise Architecture : EA) ประกอบไปด้วย 2 คำ คือ คำว่า “Enterprise” และ คำว่า “Architecture” เป็นกรอบแนวทางการวางแผนสถาปัตยกรรมองค์กรในอนาคตด้านสารสนเทศ เพื่อให้สอดคล้องกับกลยุทธ์และเป้าหมายขององค์กร โดยความหมายของ 2 คำคือ

- Enterprise หมายถึง องค์กรที่มีการทำธุรกรรม ถูกกำหนดให้เป็นนิติบุคคลที่มีสิทธิดำเนินธุรกิจด้วยตนเอง เช่น เข้าทำสัญญา เป็นเจ้าของทรัพย์สิน ก่อภาระหนี้สิน และสร้างบัญชีธนาคาร อาจเป็นบริษัท กิจบริษัท สถาบันไม่แสวงหาผลกำไร หรือวิสาหกิจที่ไม่ใช่นิติบุคคล
- Architecture คือ การออกแบบเทคโนโลยีสารสนเทศโดยองค์รวมขององค์กรในทุกมิติ มีการวางแผนเพื่อให้สอดคล้องกับกลยุทธ์ในเชิงธุรกิจขององค์กร

ตัวอย่างของกลยุทธ์สารสนเทศเพื่อสร้างความได้เปรียบเชิงธุรกิจ เช่น

- (1) สร้างองค์กรให้เป็นหน่วยงานที่ใช้ข้อมูลเพื่อการตัดสินใจในการทำงาน (Data-driven Organization)

(2) กำหนดแผนการพัฒนาแอปพลิเคชัน (Application) ขององค์กร ที่เน้นการเชื่อมโยงทุกระบบในลักษณะของ Enterprise Resource Planning (ERP) มีระบบผ่าน API เพื่อแลกเปลี่ยนข้อมูลรวมถึงการเปิดข้อมูลให้กับผู้เกี่ยวข้องของระบบธุรกิจ (Stakeholder)

(3) กำหนดแผนพัฒนาสารสนเทศโดยใช้ Cloud Computing เป็นเทคโนโลยีเพื่อการใช้ทรัพยากรร่วมกันและลดต้นทุนการใช้กระดาษให้มากที่สุด

โดยสรุปสถาปัตยกรรมองค์กร (Enterprise Architecture : EA) คือ แผนผังภาพรวมขององค์กร (Scenario) ที่ถูกเขียนขึ้นจากวิสัยทัศน์ขององค์กรหรือผู้บริหาร โดยบูรณาการเทคโนโลยีดิจิทัลให้สอดคล้องกับกระบวนการการทำงาน (Business Process) หรือภารกิจ ขององค์กรอย่างเป็นระบบ ตั้งแต่ในระดับระดับนโยบาย ยุทธศาสตร์ นำไปสู่ทิศทางการขับเคลื่อน (Roadmap) ขององค์กร ทำหน้าที่เป็นพิมพ์เขียวที่มีรายละเอียดมากพอที่จะสร้างสิ่งที่มีความสลับซับซ้อนได้อย่างมีคุณภาพ ช่วยลดการทำงานที่ผิดพลาดที่มักเกิดจากความเข้าใจที่ไม่ตรงกัน เป็นทั้งกรอบแนวทาง (Framework) และ ขั้นตอนวิธี (Method) รวมทั้งกฎกติกาและกระบวนการมาตรฐานต่างๆ ช่วยให้องค์กร บรรลุ ยุทธศาสตร์ทางธุรกิจ ซึ่งตัวอย่างสำหรับกรอบแนวคิดและหลักการที่ผู้เขียนแนะนำต่อไปคือ Open Group Architecture Framework (TOGAF) เป็นเฟรมเวิร์คที่ใช้มากที่สุดสำหรับสถาปัตยกรรมองค์กรในปัจจุบัน โครงการที่ดีต้องเริ่มจากให้ความสำคัญกับกระบวนการทำงาน การไหลของงาน กระบวนการทางธุรกิจ ทำการรวบรวมวิเคราะห์กิจกรรมอะไรที่เกี่ยวข้อง มีการปรับเปลี่ยนอย่างไร มีใครเกี่ยวข้องบ้าง เป็นต้น ซึ่งในปัจจุบันมีการออกแบบ (Framework) ที่ใช้กันแพร่หลาย หลายทฤษฎี ซึ่งหนึ่งในนั้นคือ “TOGAF” (The Open Group Architecture Framework)

3.2 ประโยชน์ของการใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture: EA) ในการออกแบบระบบสารสนเทศสมัยใหม่

ประโยชน์ที่ได้จากการใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture : EA)

3.2.1 การลดต้นทุน (Cost Saving) ในการออกแบบระบบสารสนเทศสมัยใหม่ก็คือ เรื่องของการลดต้นทุน (Cost Saving) ที่จะทำให้เกิดการลงทุนของธุรกิจที่คุ้มค่า ทั้งนี้เมื่อองค์กรจะมีการออกแบบระบบสารสนเทศใหม่ องค์กรสามารถใช้ EA พิจารณา เพื่อให้ได้ทางออก หรือทางเลือกของ สถาปัตยกรรม (Solution Architecture) ที่เหมาะสมตาม สถานการณ์ของอุตสาหกรรม และการเปลี่ยนแปลงของเทคโนโลยีที่รวดเร็ว และเลือกเทคโนโลยีต่างๆตามแผน

3.2.2 เข้าใจสถานะที่เป็นอยู่คืออะไร (As-Is) โดยกรอบ EA จะช่วยสร้างความเข้าใจ และความสัมพันธ์ที่ชัดเจนต่อสถานะปัจจุบันสถาปัตยกรรมองค์กร (Current State of Enterprise Architecture) ทั้ง 2 ด้านหลัก คือ ด้านธุรกิจ (Business) และ ด้านโครงสร้างพื้นฐาน (Infrastructure)

รวมทั้ง 4 ด้านรอง เพราะมุมมองของ EA ที่วางแผนไว้แล้วนั้น จะต้องคำนึงถึงสถานะปัจจุบันของระบบสารสนเทศในองค์กรว่ามีการลงทุนใช้เทคโนโลยีสารสนเทศไปแล้วอย่างไร Infrastructure เป็นอย่างไร ใช้เทคโนโลยีอะไร เช่น ฮาร์ดแวร์ ซอฟต์แวร์ อะไร เครือข่ายการเชื่อมโยงของหน่วยธุรกิจเป็นอย่างไร สถานะที่เป็นอยู่คืออะไร (As-Is) ประสบการณ์การใช้งานเป็นอย่างไร (User Usage/Adoption/Experience) ในแต่ละมิติ ข้อมูล มีเก็บอะไรแล้วบ้าง การบริหารจัดการข้อมูล รวมถึงการนำข้อมูลมาใช้ต่อ

3.2.3 ออกแบบอนาคตของสถาปัตยกรรมองค์กร (Future State of Enterprise Architecture) ได้ครบถ้วน การมีกรอบ EA ทำให้องค์กรสามารถออกแบบโดยการนำเป้าหมายเชิงธุรกิจ (Business Architecture) มา นำเป้าหมายทางสารสนเทศเพราะนำมุมมองของผู้บริหารมาร่วมด้วย จะทำให้เราเห็นว่าองค์กรจะขับเคลื่อนอย่างไร ส่วนไหนคือหัวใจสำคัญของธุรกิจ องค์กรจะเห็นภาพว่าควรนำเทคโนโลยีหรือซื้อระบบที่ทางธุรกิจใดมาใช้ในองค์กร หรือไม่ควรนำเทคโนโลยีที่ทางองค์กรไม่ได้ให้ความสำคัญและไม่ได้นำไปใช้มาอยู่ในกรอบ EA

3.2.4 องค์กรสามารถใช้ EA ประเมินสถานภาพความแข็งแกร่งตามหลักของ Balance Score Card ตามหลักของการบริหารตามแนวทางของ Balanced Score Card คือการแปลงวิสัยทัศน์และกลยุทธ์ขององค์กรให้ออกมาเป็นตัวชี้วัดต่าง ๆ ใน 4 ด้าน ได้แก่ ด้านการเงิน ด้านลูกค้า ด้านกระบวนการภายในธุรกิจหรือระบบสารสนเทศ และด้านการเรียนรู้และการเติบโตขององค์กร โดยองค์กรต้องผลักดันตัวชี้วัดทั้ง 4 ด้านนี้ให้ตอบสนองต่อเป้าหมาย กำไร และความสำเร็จของการบริหารองค์กรเชิงกลยุทธ์ ถ้าองค์กรไม่ได้มีการวางกรอบ EA ก็จะไม่เห็นภาพว่า ขณะนี้กระบวนการภายในธุรกิจหรือระบบสารสนเทศ มีความแข็งแกร่งหรือมีความพร้อมทั้งด้านข้อมูลด้านโครงสร้างพื้นฐาน (Infrastructure) หรือด้านระบบงานแค่ไหน มีปัญหาอะไรอยู่หรือไม่ โดยเฉพาะในปัจจุบันธุรกิจต้องการการตอบสนองแบบเรียลไทม์ (Real Time) การมีกรอบ EA จะทำให้เห็นว่าระบบงานมีความยืดหยุ่นเพียงพอที่จะสนับสนุนความต้องการที่เปลี่ยนแปลงอย่างรวดเร็วแค่ไหนในเชิงกลยุทธ์

โดยสรุปการมีกรอบ EA จะเห็นแผนผังหลัก 4 ด้านคือ

- (1) ด้านสถาปัตยกรรมของธุรกิจองค์กร (Business Architecture)
องค์กรจะสารภาพเห็นกลยุทธ์ขององค์กร กระบวนการทางธุรกิจ (Business Process) ของแต่ละหน่วยได้ชัดเจน
- (2) ด้านสถาปัตยกรรมสารสนเทศ (Information Architecture) EA
แสดงให้เห็นว่าต้องมีข้อมูลอย่างไรในแต่ละหน่วยธุรกิจ

(3) ด้านสถาปัตยกรรมโปรแกรมประยุกต์ (Application Architecture) EA แสดงให้เห็นว่าในองค์กรจะต้องมีระบบ หรือแอปพลิเคชัน (Application) อะไรบ้าง ใน การที่จะตอบโจทย์ของธุรกิจต่างๆ

(4) ด้านสถาปัตยกรรมทางเทคนิค (Technical Architecture) EA แสดงให้เห็นโครงสร้างฮาร์ดแวร์ (Hardware) ซอฟต์แวร์ (Software) และระบบเครือข่าย (Network System) ในองค์กร

3.3 ตัวอย่างองค์ประกอบและเฟรมเวิร์ค (Framework) ของสถาปัตยกรรมองค์กรสำหรับระบบสารสนเทศสมัยใหม่

กรอบสถาปัตยกรรมองค์กร ของ The Zachman The Enterprise Ontology

Classification Names Audience Perspective	อะไร (What)	อย่างไร (How)	ที่ไหน (Where)	ใคร (Who)	เมื่อไหร่ (When)	ทำไม (Why)	Classification Names Model Name
Executive	List of Inventory Types	List of Process Types	List of Distribution Types	List of Business Entity Types	List of Timing Types	List of Allocation Types	Scope Contexts
Business Management	Business Entity	Business Transform	Business Location	Business Role	Business Interval	Business End	Business Concepts
Architect	System Entity	System Transform	System Location	System Role	System Interval	System End	System Logic
Engineer	Technology Entity	Technology Transform	Technology Location	Technology Role	Technology Interval	Technology End	Technology Physics
Technician	Tool Entity	Tool Transform	Tool Location	Tool Role	Tool Interval	Tool End	Tool Components
Enterprise	Operation Entity	Operation Transforms	Operation Locations	Operation Role	Operation Interval	Operation End	Operation Instance
Enterprise Names	Inventory Set	Process Flows	Distribution Network	Enterprise Role Assignment	Timing Cycle	Allocation Instance	

รูปที่ 3.1 กรอบแนวคิดและหลักการในการพัฒนาสถาปัตยกรรมองค์กรสากล เช่น The Zachman Framework (Zachman, 2022)

3.3.1 Zachman Framework คืออะไร

Zachman Framework เป็น เฟรมเวิร์ค (Framework) ที่ได้รับการตั้งชื่อตามผู้สร้าง John Zachman ซึ่งเป็นคนแรกที่พัฒนาแนวคิดนี้ในทศวรรษ 1980 ที่ IBM และมีการปรับปรุงหลายครั้ง Zachman Framework เป็นกรอบการพัฒนาสถาปัตยกรรมองค์กรที่พิจารณาหลักการ 2 มิติ (แสดงด้วยภาพประกอบ 3.2-แกน X) โดยมีมิติแรก คือ 5W1H ประกอบด้วย 6 คำถาม ใคร (Who) ทำอะไร (What) ที่ไหน (Where) เมื่อไร (When) ทำไม (Why) และ อย่างไร (How) คำถามหัวคอลัมน์แนวนอนเหล่านี้จะช่วยรวบรวม และ วิเคราะห์เรื่องราวหรือสิ่งต่างๆ และการหาความสัมพันธ์ในเชิงเหตุผล

มิติที่สอง คือ มุมมองผู้รับฟัง (แกน Y ตาม แนวตั้งของรูป) เป็นมิติที่กล่าวถึงมุมมองของ ผู้ที่เกี่ยวข้อง (Stakeholder) กับสถาปัตยกรรมในมุมมองต่างๆ ประกอบไปด้วย 6 มุมมอง 1. มุมมองผู้บริหาร (Executive) 2. มุมมองผู้จัดการธุรกิจ (Business Management) 3. มุมมองสถาปนิก (Architect) 4. มุมมองวิศวกร (Engineer) 5. มุมมองช่างเทคนิค (Technician) และ 6. มุมมองระดับองค์กร (Enterprise) มุมมองเหล่านี้ จะมองสิ่งที่ได้มาจาก 5W1H จากนามธรรมสู่รูปธรรม มากขึ้นตามลำดับ

Zachman Framework™ เป็นโมเดลหลายมิติ (Metamodel) ที่สามารถประยุกต์ใช้ได้กับหลายกิจกรรมในการวิเคราะห์ นั่นคือ ไม่ว่าผู้ที่จะสร้างโมเดลดั้งเดิม (Primitive Models) การวิเคราะห์ด้านปรัชญา (The Ontological) การเชื่อมวิเคราะห์แบบตัวแปรเดียว (Single-Variable Intersection ระหว่างคำถาม Interrogatives) และ การปรับเปลี่ยน Transformations หรือว่าผู้ที่จะสร้าง Ad hoc การวิเคราะห์แบบหลายตัวแปร (Multi-variable) หรือ โมเดลแบบผสมผสาน (Composite Models) ที่ประกอบขึ้นจากส่วนประกอบ ของ Primitive Models หลายรุ่น เป็นสถาปัตยกรรมที่วิเคราะห์ จากบนลงล่าง จากล่างขึ้นบน ซ้ายไปขวา ขวาไปซ้าย จุดเริ่มต้น ฯลฯ ฯลฯ ได้หลายทิศทาง

ตอนนี้เราอยู่ในยุคข้อมูลข่าวสาร ยุคที่องค์กรที่มีความซับซ้อนเพิ่มมากขึ้น และองค์กรที่กำลังเปลี่ยนแปลง Enterprise Architecture –EA จะช่วยให้การวิเคราะห์ในยุคข้อมูลข่าวสารดีขึ้น โดยเฉพาะนำมาใช้กับการออกแบบระบบสารสนเทศสมัยใหม่ เนื่องจาก EA ให้มุมมอง (View Point) และเฟรมเวิร์ค (Framework) ที่กำหนดขั้นตอน และแนวทางในการจัดทำโครงการสารสนเทศ EA เป็นแบบจำลอง (Model) ที่อธิบายความสัมพันธ์กันระหว่าง ฟังก์ชันทางธุรกิจ กระบวนการทำงาน ข้อมูลที่ใช้ บุคลากรที่เกี่ยวข้อง เทคโนโลยี และตัวองค์กรเอง (โครงสร้าง วัสดุทัศน พลังงาน ยุทธศาสตร์) จึงมีสิ่งที่น่าสนใจวิเคราะห์ต้อง ตอบคำถามแต่ละมุมมองจากมิติของ Framework สำหรับจัดทำ

และนำเสนอระบบเพื่อดำเนินธุรกิจทั้งปัจจุบันและอนาคต เพื่อให้ผู้ที่เกี่ยวข้องสามารถเข้าใจ เข้าถึง และสามารถนำ EA ไปใช้งานได้อย่างเหมาะสม

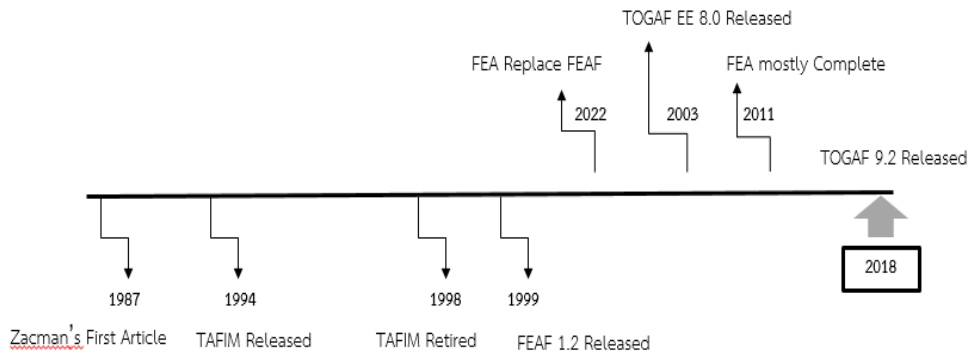
3.3.2 คำถามที่สำคัญของ Zachman Framework™

- คนในองค์กรรวมทั้งผู้บริหารมองเห็นภาพรวมขององค์กรหรือไม่
- คนในองค์กรรู้ปัญหาที่เกี่ยวกับธุรกิจ (Business) ที่ต้องใช้ระบบสารสนเทศ (IT) เข้ามาปรับปรุงส่วนใดบ้าง
- คนในองค์กรรับรู้ถึงความซับซ้อนของงานหรือไม่
- องค์กรขาดความเชื่อมโยงด้านกระบวนการปฏิบัติงานของงานที่เกี่ยวข้องหรือไม่
- องค์กรมีความคุ้มค่าในการลงทุนด้านระบบสารสนเทศหรือไม่
- และแผนธุรกิจและแผนการพัฒนาด้านระบบสารสนเทศเป็นไปในทิศทางเดียวกันหรือไม่

ดังนั้นเพื่อการวิเคราะห์และออกแบบระบบสารสนเทศสมัยใหม่ องค์กรควรนำกรอบแนวทาง Zachman Framework™ ไปประยุกต์ใช้ โดยกรอบ Zachman Framework™ นี้ สามารถ สร้าง ประโยชน์ได้หลายมุม เช่น

- เป็นกรอบการจัดระเบียบและวิเคราะห์ข้อมูลทำให้มองเห็นภาพรวมขององค์กร
- เป็นกรอบสำหรับการออกสถาปัตยกรรมองค์กรทำให้สามารถริเริ่มการกำหนดยุทธศาสตร์ใหม่ได้
- ใช้เป็นกรอบแสดงระบบการจัดประเภทหรือรูปแบบการจัดหมวดหมู่เพื่อการวิเคราะห์ ทำให้การคิดค้นรูปแบบบริการใหม่ๆ ได้รวดเร็วขึ้น
- เป็นแบบจำลองสองมิติสำหรับการวิเคราะห์ ทำให้เห็นผลกระทบและสะท้อนระหว่างมิติต่างๆ ในองค์กร
- เป็นตารางสองมิติ ใช้เพื่อจัดระเบียบในการนำเสนอโดยละเอียดขององค์กร ทำให้ระบุปัญหา สาเหตุและแนวทางการแก้ไขปัญหาได้เร็วขึ้น และทำให้วางแผนเห็นความคุ้มค่าในการลงทุน ลดความซับซ้อนของการลงทุน และองค์กรสามารถ ปรับตัวกับการเปลี่ยนแปลงต่างๆ ได้

3.3.3 ไทม์ไลน์ (Timeline) โดยย่อของสถาปัตยกรรมองค์กร



รูปที่ 3.2 ประวัติโดยย่อของสถาปัตยกรรมองค์กร (A Brief History of Enterprise Architecture , 2022)

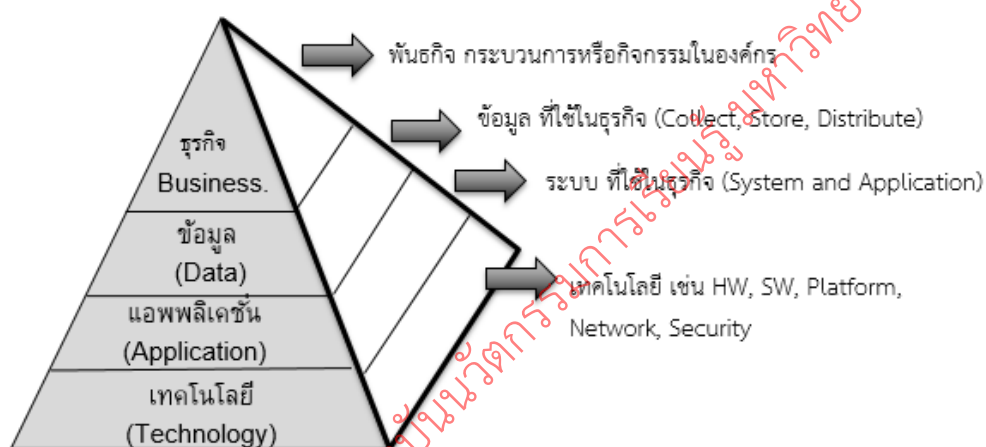
จากรูป 3.2 EA Framework ได้ถูกปรับปรุงเปลี่ยนแปลงมาตลอด ตั้งแต่ ปี ค.ศ. 1987

- ในปี 1994 The Open Group ได้พัฒนา Zachman Framework มาเป็น TAFIM ซึ่งเป็นพื้นฐานสำหรับการพัฒนาเฟรมเวิร์ค (Framework) TOGAF โดยที่สถาปัตยกรรม TOGAF เริ่มใช้สำหรับมุมมองเชิงกลยุทธ์ ทัวทั้งองค์กร มุ่งเน้นไปใช้สำหรับการพัฒนาเทคโนโลยี เพื่อกำหนดรูปแบบแพลตฟอร์ม (Platform) ที่จำเป็นที่องค์กรต้องใช้เพื่อสนับสนุนการทำงานทางธุรกิจ
- ปี 2011 TOGAF ได้ถูกพัฒนาต่อเนื่อง เพื่ออธิบายว่าสถาปัตยกรรมองค์กรที่สมบูรณ์ควรมีโดเมนหลักๆ 4 มุม
 - โดเมนทางธุรกิจซึ่งกำหนดกลยุทธ์ทางธุรกิจการกำกับดูแลองค์กร และกระบวนการทางธุรกิจที่สำคัญขององค์กร
 - โดเมนข้อมูลซึ่งอธิบายโครงสร้างของข้อมูลเชิงตรรกะและทางกายภาพขององค์กรและทรัพยากรที่ใช้จัดการข้อมูล
 - โดเมนแอปพลิเคชันซึ่งแสดงพิมพ์เขียวสำหรับแต่ละระบบที่ใช้ในองค์กร และแสดงความสัมพันธ์กับกระบวนการทางธุรกิจหลักขององค์กรกับกรอบการทำงานโดยรวม
 - โดเมนทางเทคนิคหรือโดเมนเทคโนโลยีซึ่งอธิบายฮาร์ดแวร์ซอฟต์แวร์และโครงสร้างพื้นฐานของเครือข่ายที่จำเป็นในการสนับสนุนระบบและแอปพลิเคชันขององค์กร

3.3.4 The Open Group Architecture Framework (TOGAF)

TOGAF ได้พัฒนาขึ้นในช่วงต้นทศวรรษ 1990 ได้รับการพัฒนาโดย The Open Group ในปี 1995 มีการนำเสนอ TOGAF (TOGAF 1.0) เวอร์ชันแรก ซึ่งพัฒนาจากเทคนิค

สถาปัตยกรรมรอบสำหรับการจัดการข้อมูล (TAFIM) การพัฒนาเริ่มต้นในช่วงปลายทศวรรษ 1980 โดยกระทรวงกลาโหมสหรัฐ ในปี 2001 TOGAF เวอร์ชัน 7 ได้มีการเผยแพร่ และ TOGAF เวอร์ชัน 8 ("Enterprise Edition") ได้รับการตีพิมพ์ครั้งแรกในเดือนธันวาคมปี 2002 ส่วน เวอร์ชัน 9.1 ได้รับการแนะนำในปี 2011 มีการพัฒนาเชิงวิวัฒนาการจาก TOGAF 8, TOGAF 9 มีลดและยกเลิกซึ่งไม่จำเป็นและเสนอตัวอย่างและเทมเพลตอื่น ๆ มากมาย และเวอร์ชันล่าสุดคือ TOGAF 9.2 ประกาศใช้เมื่อปี ค.ศ. 2018



รูปที่ 3.3 องค์ประกอบของสถาปัตยกรรมองค์กร TOGAF

สถาปัตยกรรมองค์กร TOGAF เป็นทั้งกรอบแนวทางและขั้นตอนวิธีที่ใช้ในการวิเคราะห์และออกแบบพิมพ์เขียว ที่เป็นอนาคตขององค์กร รวมทั้งเป็นเครื่องมือช่วยในการจัดทำแผนการจัดทำโครงการระบบสารสนเทศสมัยใหม่โดย เชื่อมโยง “วิสัยทัศน์และยุทธศาสตร์” หรือ “แผนธุรกิจและ แผนสารสนเทศ” ขององค์กรให้เป็นรูปธรรมที่จับต้องได้ โดยมี 2 มุมมองหลักประกอบไปด้วยคั่นดังนี้

3.3.4.1 มุมมองเชิงกลยุทธ์ (Business Strategy) จะประกอบไปด้วย มุมมองของผู้บริหาร (Executive) และมุมมองผู้จัดการธุรกิจ (Business Management) ตามกรอบแนวคิดและหลักการในการพัฒนาสถาปัตยกรรมสารสนเทศ องค์กรสากลจะพิจารณาถึง ผู้มีส่วนได้ส่วนเสีย (Stakeholder) ในการสร้างระบบนิเวศของระบบสารสนเทศ ซึ่งในระบบนิเวศของผู้มีส่วนได้ส่วนเสีย (Stakeholder) สามารถแบ่งได้ เป็น 3 กลุ่ม คือ กลุ่มที่ 1 ผู้มีส่วนได้ส่วนเสียหลัก (Key Stakeholders)

กลุ่มที่ 2 ผู้มีส่วนได้ส่วนเสียขั้นพื้นฐาน (Primary Stakeholders) และกลุ่มที่ 3 ผู้มีส่วนได้ส่วนเสียขั้นรอง (Secondary Stakeholders)

3.3.4.2 มุมมองด้านโครงสร้างพื้นฐาน (Infrastructure) มุมมองส่วนนี้อาจจะแยกย่อยออกมาเป็นหลายส่วน เช่น มุมมองสถาปนิก (Architect) มุมมองวิศวกร (Engineer) มุมมองช่างเทคนิค (Technician) ซึ่งในระบบนิเวศของมุมมองด้านโครงสร้างพื้นฐาน จะสามารถแยกย่อยได้อีก 4 ด้านคือ

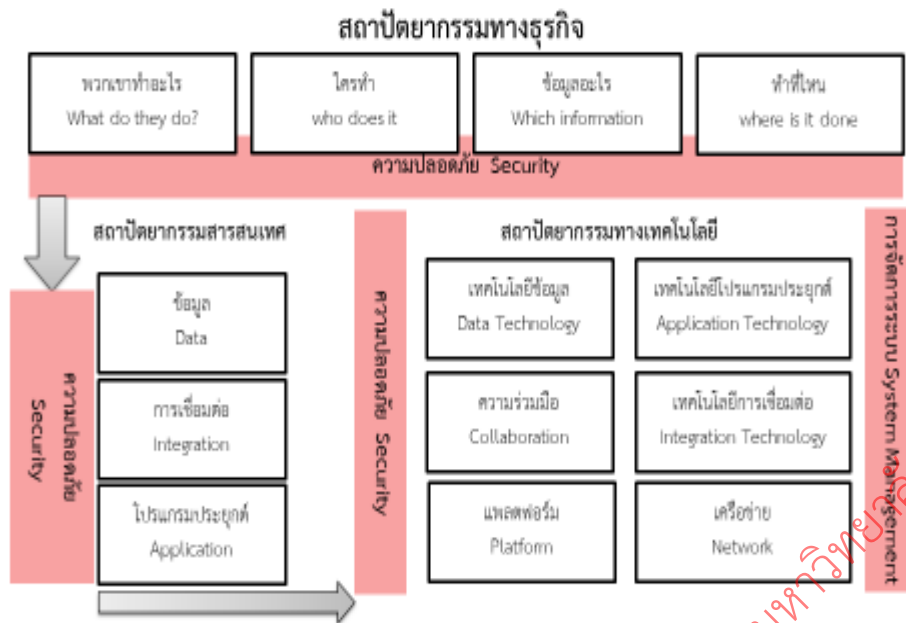
(1) มุมมองด้านแอปพลิเคชัน (Application) คือมุมมองที่มองระบบงานที่เป็นด้าน แอปพลิเคชัน (Application) ทั้งหมดที่เกี่ยวข้องเพื่อรองรับกลยุทธ์ขององค์กรซึ่งรวมทั้งแอปพลิเคชัน (Application) เฉพาะด้านและแอปพลิเคชัน (Application) ที่เป็นพื้นฐานสนับสนุนงานทั้งองค์กร

(2) มุมมองด้านข้อมูล คือมุมมอง ข้อมูล (Data) ทั้งหมดที่เกี่ยวข้องที่สนับสนุนในการทำธุรกิจ

(3) มุมมองด้านความมั่นคงปลอดภัย (Security) เป็นมุมมองด้านการรักษาความปลอดภัยทางไซเบอร์ เพื่อให้ระบบสารสนเทศมีความปลอดภัยสำหรับผู้ใช้งานและเจ้าของระบบนั้นจำเป็นจะต้องมีการคัดกรองผู้ใช้งานและยืนยันตัวตนเพื่อประกอบการพิจารณาของระบบว่าผู้เข้าใช้งานดังกล่าวคือบุคคลที่ระบบได้รองรับไว้ หรือเป็นผู้ไม่ประสงค์ดีที่ต้องการสร้างความเสียหายให้เกิดขึ้นกับระบบ

(4) มุมมองด้านเวลา คือมุมมองเชิงแผนงานของโครงการ และแผนอนาคต (Roadmap) ด้านสารสนเทศให้สอดคล้องกับแผนการด้านธุรกิจ เพื่อสร้างความได้เปรียบในการแข่งขันทั้งปัจจุบันและอนาคต

3.3.5 แนวทางการออกแบบสถาปัตยกรรมระบบสารสนเทศตามเฟรมเวิร์ค TOGAF



รูปที่ 3.4 แนวทางการใช้งานตามสถาปัตยกรรมองค์กร TOGAF (Harrison, 2018)

สำหรับแนวทางการออกแบบสถาปัตยกรรมระบบสารสนเทศตามเฟรมเวิร์ค TOGAF นั้น ให้ความสำคัญกับพันธกิจ (Mission) ขององค์กร โดยเน้นการปรับกระบวนการ จากบนลงล่าง (Top-Down) แล้วจึง ค่อยปรับระบบพื้นฐานโครงสร้างทางเทคโนโลยีสารสนเทศ เช่น ด้านข้อมูล (Data) การเชื่อมโยง (Integration) และระบบหรือแอปพลิเคชัน ตามกระบวนการที่เปลี่ยนไป โดยต้องอาศัยการวิเคราะห์สถาปัตยกรรมธุรกิจ (Business Architecture) โดยเบื้องต้น โดยต้องตอบคำถาม

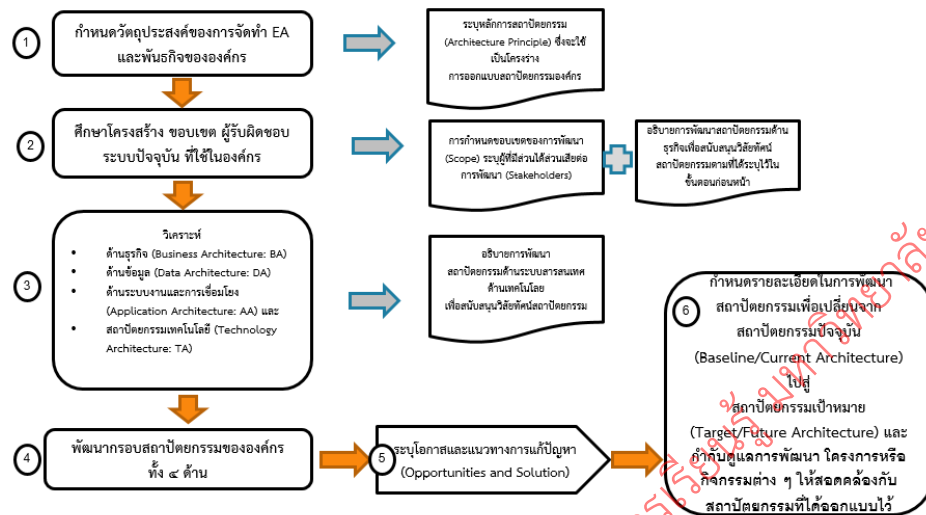
- What do they do (we do)? องค์กรเราทำอะไร หรือคู่แข่งกำลังทำอะไร
- Who does it? ใครคือผู้เกี่ยวข้องหรือรับผิดชอบ
- Which information? มีสารสนเทศใดบ้างที่เกี่ยวข้อง
- Where is it done? เราจะเริ่มอย่างไร ที่ไหน

เพื่อออกแบบการประยุกต์ใช้เทคโนโลยีสารสนเทศ ให้เชื่อมโยงกับการดำเนินงานได้ อย่างสอดคล้องและมีประสิทธิภาพ ทั้งในระดับสถาปัตยกรรม ไปจนถึง แนวทางการพัฒนา (Roadmap) ของ องค์กร เพื่อผลักดันให้องค์กรสามารถ ดำเนินการตามนโยบายและวิสัยทัศน์ ขององค์กรที่กำหนดไว้ ทั้ง 4 องค์ประกอบ คือ

- 1) ด้านธุรกิจ (Business Architecture: BA)
- 2) ด้านข้อมูล (Data Architecture: DA)
- 3) ด้านโปรแกรมประยุกต์ และการเชื่อมโยง (Application Architecture: AA) และ

4) สถาปัตยกรรมเทคโนโลยี (Technology Architecture: TA) ดังในภาพที่ 3.5

โดยดำเนินกิจกรรมดังรูป 3.5 ขั้นตอนการออกแบบสถาปัตยกรรมองค์กรตามแนวทาง TOGAF โดยย่อ (ปรับปรุงจากขั้นตอนการพัฒนาสถาปัตยกรรมองค์กรของ TOGAF)



รูปที่ 3.5 ขั้นตอนการออกแบบสถาปัตยกรรมองค์กรตามแนวทาง TOGAF



จากภาพ 3.5 ผู้เขียนได้ สรุปงานขั้นตอนที่ต้องทำให้เหลือ 6 กระบวนการ กล่าวคือ

1. ระยะเตรียมการ (Preliminary) องค์กรจะกำหนดวัตถุประสงค์ของการจัดทำ EA และพันธกิจขององค์กร กำหนดแนวทางและคุณลักษณะของสถาปัตยกรรมองค์กรที่ควรจะเป็นโดยศึกษาบริบทและขอบเขตขององค์กรและส่วนที่เกี่ยวข้องทั้งหมด ทั้งภายในและภายนอก
2. ระยะที่สอง Architecture Vision และ Business Architecture และ ทบทวนเป้าหมาย พันธกิจ กลยุทธ์ขององค์กร โดยทบทวนทั้งในมิติด้านธุรกิจ ข้อมูล แอปพลิเคชัน และเทคโนโลยี โดยศึกษาโครงสร้าง ขอบเขต ผู้รับผิดชอบ ระบบปัจจุบัน ที่ใช้ในองค์กร ทำความเข้าใจและอธิบายสถาปัตยกรรมทางธุรกิจในปัจจุบัน (As-Is) พร้อมกับออกแบบสถาปัตยกรรมที่ต้องการ (To-Be) ที่สอดคล้องตามแผนภาพแนวคิดของสถาปัตยกรรมที่อยากได้ โดยแสดงแผนภาพความสัมพันธ์ของกระบวนการทางธุรกิจ ผู้เกี่ยวข้อง รวมถึงหน่วยงานที่รับผิดชอบในทุก ๆ มิติ (View Point)

3. ระยะเวลา Information Systems Architectures คือกิจกรรมที่ต้องวิเคราะห์สถาปัตยกรรมย่อย เช่น สถาปัตยกรรมด้านข้อมูล (Data Architecture: DA) ด้านโปรแกรมประยุกต์ และการเชื่อมโยง (Application Architecture: AA) และสถาปัตยกรรมเทคโนโลยี (Technology Architecture: TA) โดยพิจารณาว่าองค์กรมีข้อมูลใดบ้าง มีการสร้าง จัดเก็บ รายงาน และส่งต่อระหว่างกระบวนการอย่างไร มีแอปพลิเคชันใดบ้างที่เกี่ยวข้อง แต่ละแอปพลิเคชันส่งผ่าน ข้อมูลใดต่อกันบ้าง

4. ระยะเวลาพัฒนารอบสถาปัตยกรรมขององค์กร ทั้ง 4 ด้านให้สอดคล้องตามแนวคิดสถาปัตยกรรมที่ได้ออกแบบไว้

5. ระบุโอกาสและแนวทางการแก้ปัญหา (Opportunities and Solution) เป็นขั้นตอนของวิเคราะห์โอกาส วางแนวทางและแผนงานในการนำสถาปัตยกรรมที่ได้ออกแบบไว้ ไปใช้งานจริง โดยใช้ผลการวิเคราะห์ความต่างระหว่างสถาปัตยกรรมปัจจุบัน (Baseline) และสถาปัตยกรรมเป้าหมายที่ได้ ออกแบบไว้ (Target) โดยคำนึงถึง ผู้เกี่ยวข้องทุกหน่วยกิจกรรม ซึ่งจะได้รับผลกระทบจากการเปลี่ยนแปลง

6. กำหนดรายละเอียดในการพัฒนาสถาปัตยกรรมเพื่อเปลี่ยนจากสถาปัตยกรรมปัจจุบัน (Baseline/Current Architecture) ไปสู่สถาปัตยกรรมเป้าหมาย (Target/Future Architecture) และกำกับดูแลการพัฒนา โครงการหรือกิจกรรมต่าง ๆ ให้สอดคล้องกับสถาปัตยกรรมที่ได้ออกแบบไว้ และจัดการลดข้อจำกัด รวมทั้งรับฟังความต้องการจากผู้เกี่ยวข้องหลัก และปรับให้เหมาะสมกับการใช้งานในแต่ละระยะของการออกแบบสถาปัตยกรรมองค์กร

3.4 สรุปประโยชน์ของสถาปัตยกรรมองค์กร TOGAF

โดยสรุป สถาปัตยกรรมองค์กร TOGAF ซึ่งประกอบด้วย สถาปัตยกรรมองค์กรย่อย 4 ด้านนั้นทำหน้าที่

1. แสดงกลยุทธ์ (Strategy) ขององค์กร แผนผังองค์กร (Organization Chart) รวมทั้งกระบวนการธุรกิจ (Business Process) และ การไหลของงาน (Workflows) ของแต่ละฝ่าย ทำให้เป็นเครื่องมือบริหาร จัดการองค์กรธุรกิจยุคใหม่ที่จะช่วยให้ได้เปรียบทางการแข่งขัน หรือเพื่อการบริหารจัดการที่มีประสิทธิภาพ และในรายละเอียดของสถาปัตยกรรมยังแสดงโครงสร้างการจัดความรับผิดชอบและหน้าที่ของพนักงาน ใครทำงานอะไร และ สายบังคับบัญชา ลักษณะของงานที่ต้องทำ และทำอะไรด้วยกระบวนการทำงานอย่างไร (Business Processes) ต้องใช้ข้อมูลแบบใด (Business Objects) เพื่อให้บริการแก่ผู้อื่นทั้งบุคคลภายนอกและภายใน ในเรื่องอะไรบ้าง

2. แสดงโครงสร้างข้อมูล (Information Architecture) แสดงการจัดเก็บข้อมูล การแลกเปลี่ยนข้อมูล โดยทำหน้าที่เป็นพิมพ์เขียวที่มี รายละเอียด เช่น โครงสร้างของข้อมูลที่ใช้ในแต่ละ

ระบบงาน (Data Architecture) และ โครงสร้างด้านโปรแกรมประยุกต์ (Application Architecture) และ แสดงการเชื่อมโยงการสนับสนุนการทำงานของกระบวนการ (Business processes) ใน สถาปัตยกรรมของธุรกิจ Business Architecture

3. เป็นเครื่องมือสื่อสารระหว่างผู้ที่เกี่ยวข้องกับการออกแบบและพัฒนาระบบให้มีความ เข้าใจตรงกัน โดยภาพระบบจะถูกนำเสนอโดยแบบพิมพ์เขียว (Blueprints) หลายรูปแบบ เพื่อ ประโยชน์ของบุคลากรที่มีหน้าที่แตกต่างกัน อันจะทำให้เห็นภาพรวมทั้งหมด ช่วยลดการทำงานที่ ผิดพลาดที่มักเกิดจากความเข้าใจ ที่ไม่ตรงกัน ตลอดจนนำเสนอข้อมูลที่จำเป็นกับการปรับเปลี่ยน ยุทธศาสตร์และใช้ในการตัดสินใจเมื่อมีการเปลี่ยนแปลงที่จะเกิดขึ้นในอนาคต

4. สถาปัตยกรรมองค์กรเช่น TOGAF ช่วยแก้ปัญหาของการลงทุนระบบที่ซ้ำซ้อนอันเกิด จากการบริหารจัดการระบบสารสนเทศอย่างไม่เป็นระบบ แยกฝ่าย แยกหน่วยงานกันทำ ขาดการ เชื่อมโยงระหว่างข้อมูลในองค์กร โดยเฉพาะขาดข้อมูลที่ซ้ำซ้อนหรือซ้ำซ้อนไม่อัปเดต (Redundancy)

สรุปเนื้อหาบทที่ 3 "ความรู้เบื้องต้นเกี่ยวกับกรอบแนวทางการวางแผนสถาปัตยกรรมระบบ" สถาปัตยกรรมองค์กร (Enterprise Architecture: EA) เป็นการอธิบายและสรุปเนื้อหาที่เกี่ยวข้องกับ กระบวนการวางแผนสถาปัตยกรรมระบบสารสนเทศในองค์กร โดยเน้นไปที่การกำหนดโครงสร้างและ แนวทางที่จะใช้ในการพัฒนาระบบสารสนเทศที่เหมาะสมกับความต้องการของธุรกิจ รวมถึงประโยชน์ ของการใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture: EA)

1. **การเริ่มต้นด้วยการวางแผน:** กระบวนการการวางแผนสถาปัตยกรรมระบบเริ่มต้นด้วยการ ระบุนิยามวัตถุประสงค์ของระบบ รวมถึงเป้าหมายและความคาดหวังที่ต้องการให้ระบบสารสนเทศ เข้ามาช่วย
2. **การกำหนดและอธิบายรูปแบบสถาปัตยกรรมระบบ:** การเลือกและกำหนดรูปแบบ สถาปัตยกรรมระบบที่เหมาะสมเพื่อให้ระบบสารสนเทศทำงานได้อย่างมีประสิทธิภาพ รวมถึง การอธิบายหลักการและเหตุผลของการเลือกรูปแบบดังกล่าว
3. **การกำหนดโครงสร้างและส่วนประกอบ:** ส่วนสำคัญของการวางแผนสถาปัตยกรรมระบบคือ การกำหนดโครงสร้างทั้งระบบ ซึ่งรวมถึงส่วนประกอบต่างๆ ที่ประกอบระบบสารสนเทศ โดย การกำหนดฐานข้อมูล การสร้างโมดูล และการกำหนดความสัมพันธ์ระหว่างโมดูลต่างๆ
4. **การกำหนดเทคโนโลยีและแพลตฟอร์ม:** การเลือกและกำหนดเทคโนโลยีที่เหมาะสมสำหรับ การพัฒนาระบบสารสนเทศ รวมถึงการกำหนดแพลตฟอร์มที่เหมาะสมกับความต้องการ ของระบบ

5. **การกำหนดกระบวนการพัฒนาและการจัดการโครงการ:** การกำหนดกระบวนการที่ใช้ในการพัฒนาระบบสารสนเทศ รวมถึงการกำหนดวิธีการจัดการโครงการ เพื่อให้การพัฒนาระบบเป็นไปได้อย่างเป็นระบบและมีประสิทธิภาพ
6. **การวางแผนการกระจายทรัพยากร:** การกำหนดแผนการใช้ทรัพยากรทั้งมนุษย์และเทคโนโลยีในกระบวนการพัฒนา รวมถึงการกำหนดแผนการจัดการความเสี่ยงและปัญหาที่อาจเกิดขึ้นระหว่างการพัฒนา
7. **การสรุปและการพิจารณาต่อไป:** บทที่ 3 จบลงด้วยการสรุปและการบัญชาเกี่ยวกับความสำคัญของการวางแผนสถาปัตยกรรมระบบสารสนเทศให้เหมาะสมกับความต้องการและธุรกิจขององค์กร

โดยบทที่ 3 เป็นการนำเสนอความรู้เบื้องต้นเกี่ยวกับกรอบแนวทางการวางแผนสถาปัตยกรรมระบบสารสนเทศ ซึ่งเน้นไปที่ขั้นตอนและข้อคิดที่เกี่ยวข้องกับการกำหนดโครงสร้าง

คำศัพท์ บทที่ 3

Analytics Layer	เลเยอร์การวิเคราะห์
Application	แอปพลิเคชัน
Application Architecture	สถาปัตยกรรมแอปพลิเคชัน
As-Is	เข้าใจสถานะที่เป็นอยู่คืออะไร
Big Data	ฐานข้อมูลขนาดใหญ่
Blueprints	แบบพิมพ์เขียว
Business Architecture	สถาปัตยกรรมของธุรกิจองค์กร
Business Process	กระบวนการทางธุรกิจ
Business Strategy	กลยุทธ์ธุรกิจ
Cloud Computing	การประมวลผลด้วยเครือข่ายคลาวด์
Composite Models	โมเดลแบบผสมผสาน
Cost Saving	การลดต้นทุน
Data Aggregation Layer	เลเยอร์การรวมข้อมูล
Data Governance Layer	ศึกษาความเป็นไปได้
Data Layer	เลเยอร์ข้อมูล
Data-driven Organization	หน่วยงานที่ใช้ข้อมูลเพื่อการตัดสินใจในการทำงาน
Enterprise Architecture	สถาปัตยกรรมองค์กร
Enterprise Resource Planning (ERP)	การวางแผนทรัพยากรองค์กรโดยรวม
Information Architecture	สถาปัตยกรรมสารสนเทศ

Information Architecture	โครงสร้างข้อมูล
Information Exploration Layer	เลเยอร์การสำรวจข้อมูล
Primitive Models	โมเดลดั้งเดิม
Roadmap	แผนอนาคต
Security	ความมั่นคงปลอดภัย
Single-Variable Intersection	การเชื่อมวิเคราะห์แบบตัวแปรเดียว
Stakeholder	ผู้เกี่ยวข้องของระบบธุรกิจ หรือผู้มีส่วนได้ส่วนเสียกับธุรกิจ
Technical Architecture	สถาปัตยกรรมทางเทคนิค
View Point	มุมมอง
Workflows	การไหลของงาน

คำถามและแบบฝึกหัดท้ายบทที่ 3

1. จงอธิบาย สถาปัตยกรรมองค์กร (Enterprise Architecture: EA) คืออะไร
2. จงอธิบายมุมมองหลัก 2 มุมของสถาปัตยกรรมองค์กรสำหรับระบบสารสนเทศสมัยใหม่
3. จงอธิบายโครงสร้างพื้นฐาน ทั้ง 4 ด้าน
4. จงอธิบายประโยชน์ของการใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture: EA)
5. จงอธิบายสถาปัตยกรรมองค์กรแบบ TOGAF
6. จงอธิบายประโยชน์ของสถาปัตยกรรมองค์กร TOGAF

บทที่ 4 แนวทางการออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศสมัยใหม่

(Guidelines for new Information System Architecture Design)

ปัจจุบันในยุคสมัยของโลกสังคมดิจิทัลและธุรกิจ มีการพึ่งพาเทคโนโลยีสารสนเทศมากยิ่งขึ้น วิทยาการข้อมูลได้เข้ามามีบทบาทสำคัญ แนวคิดของการออกแบบสถาปัตยกรรมสารสนเทศสมัยใหม่ เป็นศาสตร์ที่ว่าด้วยการบริหารจัดการข้อมูลอย่างเป็นระบบ ต้องอาศัยความรู้ทางด้านการจัดการแบบบูรณาการ ซึ่งจำเป็นต้องให้ความสำคัญต่อการออกแบบโครงสร้างและสถาปัตยกรรมข้อมูลที่ดี เพื่อให้การใช้ข้อมูลสารสนเทศเกิดประโยชน์สูงสุดต่อองค์กร

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจปัจจัยเพื่อพิจารณาสำหรับการบริหารโครงการสารสนเทศสมัยใหม่
2. สามารถประยุกต์ใช้แบบจำลองและสัญลักษณ์ต่างๆ เพื่อออกแบบสถาปัตยกรรมองค์กร และสำหรับการพัฒนาระบบสารสนเทศสมัยใหม่

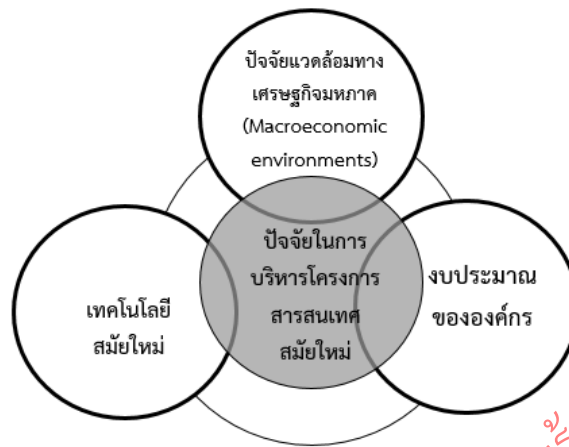
4.1 บทนำแนวทางการออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศ

แนวทางการปรับปรุงสถาปัตยกรรมสำหรับระบบสารสนเทศ ตามมาตรฐานและแนวปฏิบัติที่ดี (Best Practices) จำเป็นต้องมีวิธีการ เทคนิค เครื่องมือ และมีขั้นตอนการดำเนินงานที่แน่นอนเป็นระบบ เพื่อให้ได้ระบบสารสนเทศที่มีประสิทธิภาพ สามารถตอบสนองต่อวัตถุประสงค์หลักขององค์กร

สถาปัตยกรรมองค์กร (Enterprise Architecture) เป็นกระบวนการเชิงโครงสร้างที่ใช้สำหรับพัฒนาระบบสารสนเทศ (Information System) ที่ทันสมัยเป็นหลักการที่สามารถนำมาใช้ร่วมกับวงจรการพัฒนากระบวนการ (SDLC) เพื่อนำไปใช้ในการวางแผนและแก้ปัญหาขององค์กร โดยการประยุกต์ใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture Frameworks) จะต้องครอบคลุมถึงองค์ประกอบหลักทั้ง 5 ด้าน ประกอบด้วยด้านธุรกิจ (Business Architecture) ด้านข้อมูล (Data Architecture) ด้านโปรแกรมประยุกต์ (Application Architecture) ด้านเทคโนโลยี (Technology Architecture) และด้านความปลอดภัยของข้อมูล (Security Architecture) รวมถึงการใช้ร่วมกับกรอบการบริหารโครงการสารสนเทศสมัยใหม่

4.1.1 ปัจจัยเพื่อพิจารณาสำหรับการบริหารโครงการสารสนเทศสมัยใหม่ สำหรับการแนวทางการนำสถาปัตยกรรมสำหรับระบบสารสนเทศ มาใช้ในองค์กรนั้นจะต้องจัดทำเป็นพิมพ์เขียว (Blueprint) เพื่อกำหนดภาพรวมการบริหารจัดการองค์กรและระบบเทคโนโลยีดิจิทัลให้สอดคล้อง

กับวิสัยทัศน์ (Vision) พันธกิจ (Mission) และ สถานการณ์ของธุรกิจปัจจุบันของอุตสาหกรรมนั้นๆ โดยจะต้องพิจารณาถึงปัจจัยในการบริหารโครงการ 3 ปัจจัย (ดังรูป 4.1) คือ

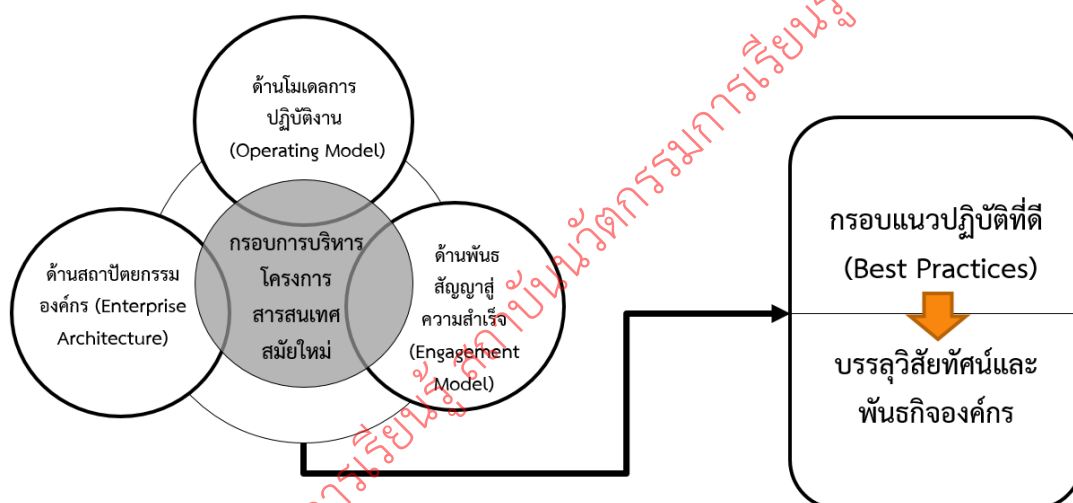


รูปที่ 4.1 ปัจจัยเพื่อพิจารณาสำหรับการบริหารโครงการสารสนเทศสมัยใหม่

- ปัจจัยแวดล้อมทางเศรษฐกิจมหภาค (Macroeconomic Environments) ตัวอย่างเช่น อัตราการขยายตัวทางเศรษฐกิจ การเปลี่ยนแปลงโครงสร้างจากการพึ่งพาภาคอุตสาหกรรมการผลิตไปสู่ภาคบริการ ระบบห่วงโซ่มูลค่าโลกที่หดสั้นลง ทำให้เพิ่มแรงกดดันด้านการแข่งขัน การรวมกลุ่มการค้าโลกมุ่งสู่การสร้างพันธมิตรภายในภูมิภาคเดียวกัน เป็นต้น
- ปัจจัยด้านเทคโนโลยีสมัยใหม่ที่กำลังเข้ามาปฏิรูปโครงสร้างภาคอุตสาหกรรมการผลิตและบริการ ได้แก่ The Internet of Things (IoT) หุ่นยนต์ที่ทำงานได้โดยอัตโนมัติและมีความแม่นยำสูง (Robotics) การใช้ปัญญาประดิษฐ์ในภาคอุตสาหกรรม (Artificial intelligence (AI)) เครือข่าย 5G (5G technology) อากาศยานไร้คนขับที่ควบคุมได้จากระยะไกล (Drone) ระบบการเก็บและบันทึกข้อมูลโดยกระจายฐานข้อมูลให้ผู้ที่อยู่ในเครือข่ายแบ่งปันข้อมูลกันได้ (Blockchain) ระบบประมวลผล Big data ที่เชื่อมโยงข้อมูลด้านความต้องการของผู้บริโภคไปสู่การปรับเปลี่ยนสายการผลิตหรือรูปแบบธุรกิจแบบทันเวลา (Real time) (Edge computing) การพัฒนาระบบประมวลผลของคอมพิวเตอร์ผ่านอนุภาคย่อยของอะตอมแทนการใช้เลขฐานสอง (Quantum Computing) เทคโนโลยีการพิมพ์ 3 มิติ (3D Printing) Synthetic Biology คือ ส่วนใหญ่ประยุกต์ใช้เพื่อผลิตเนื้อสังเคราะห์ประเภท Cultured Meat ซึ่งผลิตเนื้อสัตว์จากสเต็มเซลล์ของสัตว์ต้นแบบ และการวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Analysis)

- ปัจจัยด้านงบประมาณขององค์กร องค์กรจะต้องมีการวางแผนงบประมาณ โดยพิจารณา ทบทวนบทบาท ภารกิจให้ชัดเจน เพื่อกำหนดโครงสร้างแผนงานให้สอดคล้องกับการ จัดสรรงบประมาณรายจ่ายประจำปี รวมทั้งต้องมีการวางแผนถึงงบประมาณที่ใช้ในอนาคตเพราะระบบ มีการปรับปรุงเพิ่มเติมอยู่ตลอดเวลา การวางแผนงบประมาณมีประสิทธิภาพได้มาตรฐาน จะต้อง มี แผนงบประมาณที่ครอบคลุม แผนงบประมาณระยะปานกลาง โดยเน้นการใช้งบประมาณแบบการ บริหารจัดการเชิงรุก มีการจัดทำงบประมาณมุ่งเน้นผลงาน

4.1.2 กรอบการบริหารโครงการสารสนเทศสมัยใหม่ โดยกรอบแนวทางทางการ ออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศ นอกจากการพิจารณาปัจจัย 3 มุม ด้าน เศรษฐกิจ เทคโนโลยีและเงินลงทุนหรืองบประมาณ และต้องใช้กรอบหลักการบริหารที่สำคัญของ องค์กรประกอบในการออกแบบซึ่งแบ่งเป็น 3 กรอบหลัก (ดังรูปที่ 4.2) ดังนี้



รูปที่ 4.2 กรอบการบริหารโครงการสารสนเทศสมัยใหม่

1. กรอบโมเดลการปฏิบัติงาน (Operating Model) : เป็นโมเดลที่พิจารณามาตรฐานด้าน กระบวนการงาน (Process Integration and Standardization) โลกธุรกิจสมัยใหม่นาขนาดของ จะเป็น อย่างไรในอีก 5 ปี 10 ปี ไม่มีใครวางแผนได้ล่วงหน้าชัดเจน เพราะทุกอย่างขึ้นอยู่กับปัจจัยที่ เปลี่ยนแปลงตลอดเวลาและควบคุมไม่ได้ เช่น เรื่องความต้องการของลูกค้า รูปแบบการทำตลาดที่แปร ผันไปมา หรือเหตุการณ์อื่นๆ เช่น การระบาดของโรค ภัยสงคราม ภัยธรรมชาติ ดังนั้นกรอบโมเดลนี้ได้ เน้นถึงแนวทางการวางแผนเพื่อการส่งมอบบริการและผลิตภัณฑ์ที่ดีให้แก่ลูกค้า โดยทุกคนในองค์กร

ตั้งแต่ผู้บริหารรวมถึงผู้ปฏิบัติงานจะต้องมีส่วนในการพัฒนาแผนและปรับปรุงให้เกิดกระบวนการงานที่ต่อเนื่องจากต้นสู่ปลาย (End-to-End Process) มีการตัดสินใจในรูปแบบทันต่อเหตุการณ์และเชื่อมโยงกับโลกภายนอกมากกว่าแค่กรอบทางธุรกิจเดิมๆ รวมถึงการพิจารณาใช้เทคโนโลยีสมัยใหม่ เช่น เซนเซอร์ (Sensor), ข้อมูล (Data), AI และการวิเคราะห์ และใช้งานข้อมูลร่วมกันในทุกหน่วยงานขององค์กร

2. ดำเนินการตามกรอบสถาปัตยกรรมองค์กร (Enterprise Architecture) โดยบูรณาการ กระบวนการและมาตรฐานในการปฏิบัติงานตามรูปแบบการปฏิบัติงาน (Operating Model) โดยประโยชน์ของสถาปัตยกรรมองค์กรจะช่วยให้เห็นถึงมุมมองของการพัฒนากระบวนการ ระบบงานและเทคโนโลยีทั้งหมด

3. ดำเนินงานตามกรอบรูปแบบพันธสัญญาสู่ความสำเร็จ (Engagement Model) เป็นแนวทางการออกแบบระบบสารสนเทศที่คำนึงถึงงานสารสนเทศในแต่ละหน่วยงาน และระดับองค์กร (Enterprise Systems) ที่ใช้งานร่วมกันในหลายหน่วยกิจกรรมทั้งองค์กร และจะต้องพิจารณาความเชื่อมโยงระหว่างเทคโนโลยีดิจิทัล กระบวนการและกระบวนการตัดสินใจของผู้บริหาร เพื่อให้ตอบโจทย์การบริหารระดับการตัดสินใจขั้นสูงสุดขององค์กร

โดยสรุปการเริ่มต้นออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศสมัยใหม่จะต้องคำนึงถึงปัจจัย 3 ด้านและกรอบการบริหารโครงการสารสนเทศสมัยใหม่ดังที่ได้กล่าวมาแล้วข้างต้น หลังจากการพิจารณาสิ่งเหล่านี้แล้ว ผู้บริหารโครงการจำเป็นต้องเลือกแนวทางการใช้โมเดลการพัฒนาสารสนเทศ เช่น โมเดลวงจรพัฒนาระบบ SDLC (System Development Life Cycle) โมเดลโครงสร้างแบบน้ำตก (Waterfall Model), โมเดลโครงสร้างแบบกังหัน (Spiral Model) วิธีการพัฒนาซอฟต์แวร์แบบคล่องแคล่วว่องไว (Agile Software Development) เป็นต้น เพื่อเป็นโครงร่างหรือแนวทางวิธีการ ในการทำความเข้าใจและเพื่อใช้เป็นเครื่องมือสื่อสารขั้นตอนการพัฒนาไปยังทีมงานทั้งหมด

4.2 โมเดลที่ใช้ในการวิเคราะห์และออกแบบระบบ

โมเดลการวิเคราะห์และออกแบบระบบการพัฒนาสารสนเทศ เป็นมีวิธีการ เทคนิค เครื่องมือ และมีขั้นตอนการดำเนินงานที่แน่นอนเป็นระบบ เป็นแนวทางเชิงโครงสร้าง ที่ใช้เพื่อการพัฒนาออกแบบระบบสารสนเทศให้ได้ระบบมีประสิทธิภาพ สามารถตอบสนองต่อวัตถุประสงค์หลักขององค์กรได้ โดยประกอบด้วย วิธีการ เทคนิค และเครื่องมือ

4.2.1 วิธีการ (Methodologies) หมายถึง รูปแบบการดำเนินงานที่ใช้สำหรับการวิเคราะห์ การออกแบบ และการพัฒนาระบบสารสนเทศ วิธีการที่ใช้ สำหรับโมเดลวงจรพัฒนา

ระบบ (Systems Development Model) เป็นวิธีการแบบเดิมที่ใช้มากกว่า 30 ปี มีขั้นตอนที่กำหนดไว้อย่างแน่นอน โมเดลวงจรพัฒนาระบบแต่ละจะมีเค้าโครงเหมือนกันจะแตกต่างกันไปเพียงการแบ่งขั้นตอนและรายละเอียดของแต่ละขั้นตอน มีการแบ่งงานระหว่างผู้เกี่ยวข้องอย่างชัดเจน ขณะที่ผู้ใช้มีความรับผิดชอบในระดับจำกัดมาก โมเดลวงจรพัฒนาระบบยังคงมีใช้อยู่สำหรับการพัฒนาระบบงานขนาดเล็กหรือขนาดใหญ่ที่มีระบบที่สลับซับซ้อน มีข้อกำหนดคุณสมบัติต่าง ๆ อย่างเป็นทางการและค่อนข้างคงที่ ข้อด้อยคือเป็นโมเดลวงจรพัฒนาระบบเสนอวิธีการที่ต้องการการพัฒนาระบบที่ใช้เวลานานและทรัพยากรมาก กว่าจะเสร็จระบบที่ได้ก็อาจไม่ตรงกับความต้องการในปัจจุบันได้ นอกจากนี้ ในแต่ละขั้นตอนของการพัฒนาระบบ จะต้องใช้เทคนิคการรวบรวมข้อมูล เครื่องมือแบบจำลอง และแผนภาพชนิดต่าง ๆ อธิบายลักษณะการทำงานของระบบ และข้อมูลที่ใช้ในระบบ ซึ่งเครื่องมือเหล่านี้จะช่วยลดระยะเวลาการทำงานบางขั้นตอนลงได้ตลอดจนมีข้อผิดพลาดน้อยลง และเครื่องมือบางอย่างเป็นโปรแกรมอัตโนมัติ ช่วยสร้างแผนภาพ สร้างรายงานและแบบฟอร์ม และสร้างรหัสโปรแกรมให้ด้วย

4.2.2 เทคนิค (Techniques) หมายถึง กระบวนการและรูปแบบวิธีปฏิบัติ (Processes and Procedures) ที่ใช้ในการวิเคราะห์และออกแบบระบบ เพื่อให้การพัฒนาระบบมีวิธีการที่เป็นระบบ มี มาตรฐาน และมีประสิทธิภาพ เทคนิคที่ใช้กันได้แก่

- เทคนิคการรวบรวมข้อมูล (Data Gathering) เป็นการรวบรวมข้อเท็จจริงเกี่ยวกับระบบ
- เทคนิคการจัดการโครงการ (Project Management) ช่วยในการวางแผนการพัฒนาระบบ

4.2.3 เครื่องมือแบบจำลอง (Model Tools) หรือเครื่องมือหมายถึงการสร้างแบบจำลองกราฟิกในการนำเสนออธิบายกระบวนการ (Process) ข้อมูล (Data) วัตถุ (Object) และตรรกะ (Logic) ที่เกิดขึ้นในระบบ สำหรับแบบจำลองเชิงวัตถุจะมีเครื่องมือสำหรับการวิเคราะห์การออกแบบและการเขียนโปรแกรมเชิงวัตถุ เครื่องมือแบบจำลองได้แก่ ผังงาน (Flowchart) แผนภาพกระแสข้อมูล (Data Flow Diagram - DFD) และเครื่องมือที่นิยมใช้ควบคู่กับแผนภาพกระแสข้อมูล คือ พจนานุกรมข้อมูล (Data Dictionary) ภาษาโครงสร้าง (Structured Language) ตารางตัดสินใจ (Decision Table) และผังต้นไม้ (Decision Tree) นอกจากนี้ยังมีเครื่องมือพื้นฐานแบบจำลองข้อมูลได้แก่ แผนภาพความสัมพันธ์ระหว่างเอนทิตี (Entity-Relationship Diagram - ERD) Use Case Diagram Class Diagram Sequence Diagram Gantt Chart/Pert แบบจำลองเชิงวัตถุ ได้แก่ ยูเอ็มแอล (Unified Modeling Language) โมเดลวิเคราะห์ทางการเงิน เช่น NPV, ROI เป็นต้น

4.3 ตัวอย่างเครื่องมือแบบจำลอง

4.3.1 ตัวอย่างที่ 1 ผังงาน (Flowchart) ผังงาน (Flowchart) เป็นรูปภาพหรือ สัญลักษณ์ (ดังรูปที่ 4.1) ที่ใช้เขียนแทนขั้นตอนของงานหรือสิ่งที่กำลังอธิบาย ที่อาจประกอบไปด้วยคำอธิบาย ข้อความ หรือคำพูด แบบย่อ ของงานหนึ่งๆ โดยใช้สัญลักษณ์ที่เป็นมาตรฐานเดียวกัน เพราะการนำเสนอขั้นตอนของงานให้เข้าใจตรงกันระหว่างผู้เกี่ยวข้อง ด้วยคำพูด หรือ ข้อความทำได้ยากกว่า ผังงานสามารถนำไปใช้ได้หลากหลายประเภทของงาน เช่น

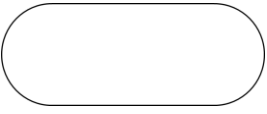
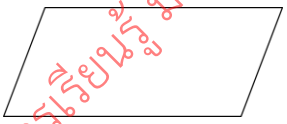
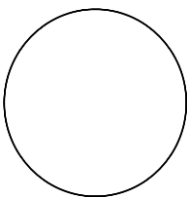
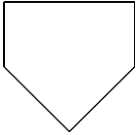
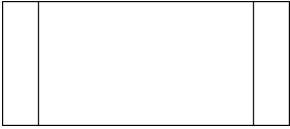
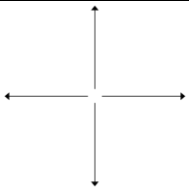
- การทำผังงานขั้นตอนการทำงานในองค์กร
- การทำผังงานของระบบเครือข่ายในการสื่อสารดิจิทัล หรือระบบสารสนเทศ
- การทำผังงานของขั้นตอนการออกกฎหมาย หรือการปฏิบัติงานทางกฎหมาย

ลักษณะของผังงานที่ดี

- ใช้สัญลักษณ์ตามที่กำหนดไว้
- มีการแบ่ง Swimming Lane หรือ BOX เพื่อระบุ คน หรือหน่วยงานที่เกี่ยวข้อง
- ใช้ลูกศรแสดงทิศทางการไหลของข้อมูลจากบนลงล่าง หรือ จากซ้ายไปขวา
- คำอธิบายในภาพควรสั้นกะทัดรัด และ เข้าใจง่าย
- ทุกแผนภาพ ผังงาน ต้องมีลูกศรแสดงทิศทางเข้า – ออก
- ไม่ควรโยงเส้นเชื่อมผังงานที่อยู่ไกลมาก ๆ ควรใช้สัญลักษณ์จุดเชื่อมต่อแทน
- การเขียนชื่อ ผังงาน ควรเขียน วันเวลา และวันแก้ไข ผังงาน รวมทั้งผู้จัดทำ ผังงาน และชื่อเอกสาร (File Name)

- ขนาดของสัญลักษณ์ ก็ต้องมีขนาด เท่ากันหรือใกล้เคียงกัน
- ลูกศรแสดงทิศทางเข้า – ออก ต้องไม่ ทับกัน จน มองเส้นทางของ งาน ไม่ชัดเจน

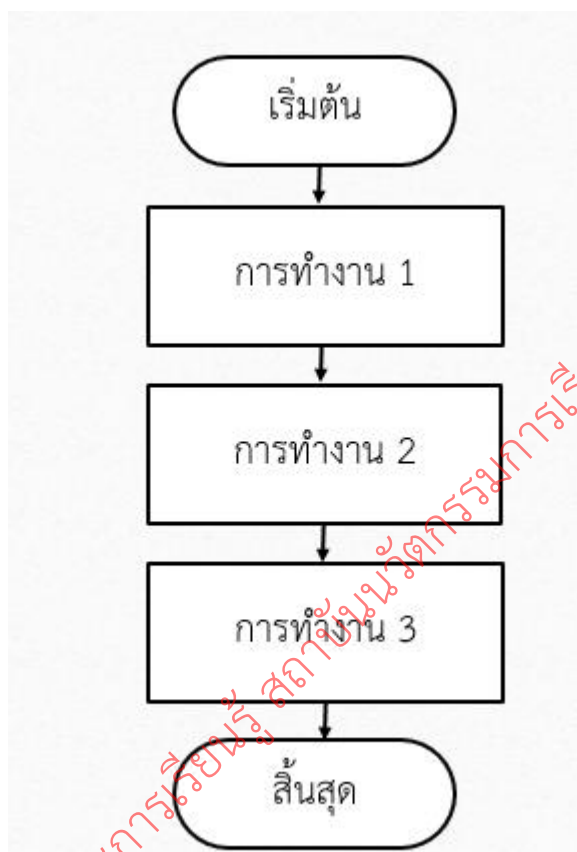
ตารางที่ 4.1 ตัวอย่างสัญลักษณ์ของผังงาน (Flow Chart)

Start/Terminator สัญลักษณ์แทนจุดเริ่มต้นและจุดสิ้นสุดของการทำงาน	
Process สัญลักษณ์แทนการดำเนินการ / ประมวลผลข้อมูล เช่น การดำเนินการค่าคงที่ การกำหนดค่าข้อมูลให้กับตัวแปร การคูณ	
Input/Output สัญลักษณ์แทนการรับและแสดงข้อมูล ในกรณีที่ไม่วิธีการรับข้อมูล อาจรับข้อมูลจากคีย์บอร์ด พิมพ์ข้อมูล หรือแสดงผลข้อมูลทางจอภาพ เครื่องพิมพ์ หรือพิมพ์ข้อมูลก็ได้	
Connector สัญลักษณ์แทนจุดเชื่อมของผังงาน เพื่อให้ผังงานดูเข้าใจง่ายขึ้น ในกรณีที่มีทางเลือกของการดำเนินการแล้วกลับมาดำเนินการต่อไปด้วยขั้นตอนเดียวกัน	
Between-page Connector สัญลักษณ์แทนจุดเชื่อมของผังงาน กรณีขึ้นหน้ากระดาษใหม่	
Predefined Process สัญลักษณ์แทนการเรียกใช้ฟังก์ชันหรือการทำงานย่อยที่มีการกำหนดไว้แล้ว	
Flow line/Direction เส้นแสดงทิศทางของการทำงาน (แต่ละเส้นต้องมีหัวลูกศรเดียวเท่านั้น)	

รูปแบบของผังงาน

รูปแบบของผังงาน (Flowchart) มี 3 รูปแบบ ดังนี้

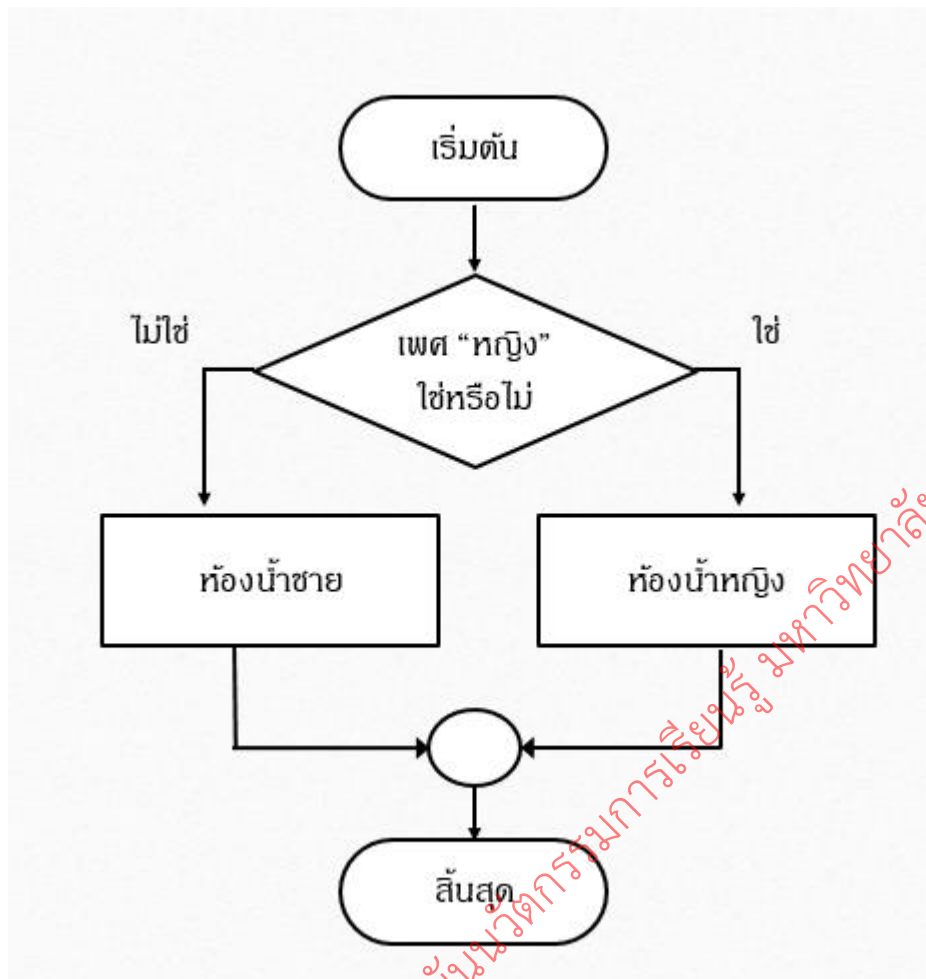
1. ผังงานเรียงลำดับ (Sequence Flowchart) คือการเขียนให้เป็นลำดับ ดังรูปที่ 1. ไม่ใช่เขียนข้ามไปข้ามมาดังรูปที่ 4.3 เป็นรูปแบบผังงานที่ง่ายที่สุด ไม่ซับซ้อน และไม่มีการเปรียบเทียบเงื่อนไขใด ๆ โดยแสดงขั้นตอนการทำงานไปตามลำดับตั้งแต่ต้นจนสิ้นสุดการบวนการ



รูปที่ 4.3 ผังงานเรียงลำดับ (Sequence)

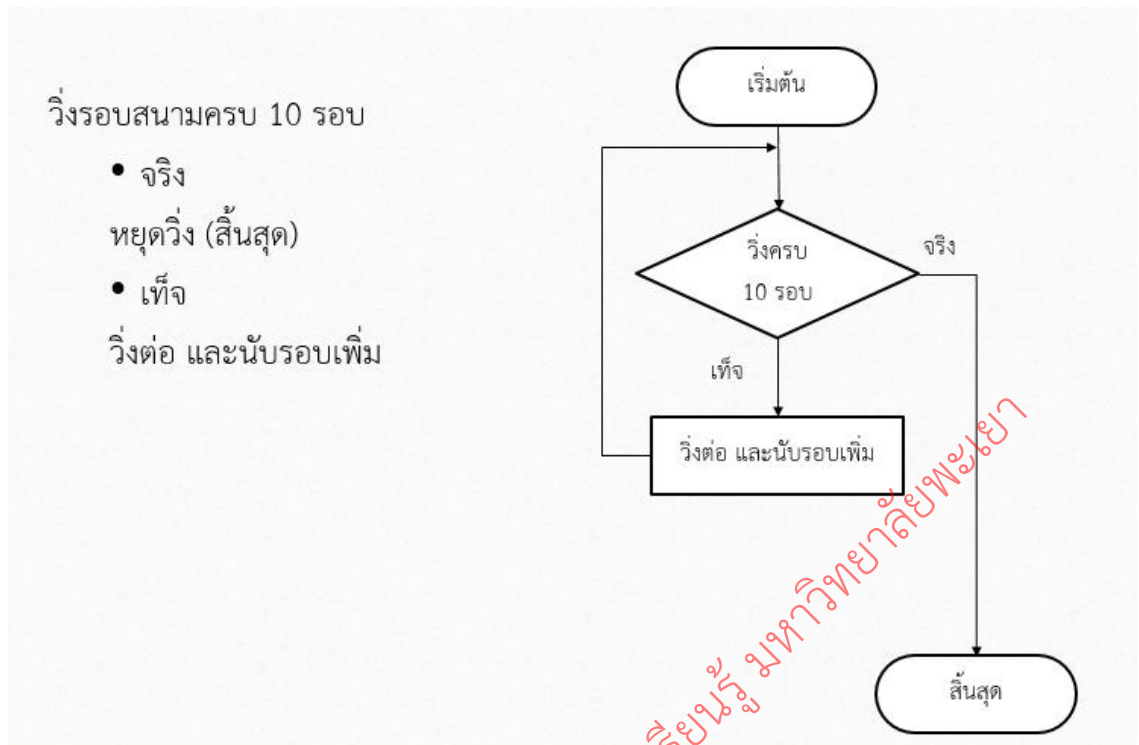
2. ผังงานแบบมีเงื่อนไข (Selection Flowchart) ดังรูปที่ 4.4 เป็นทางเลือกของโปรแกรมซึ่ง จะต้องมีเพียงสองทางเลือกเท่านั้น และ หลังจากนั้นทางเลือกทั้งสองต้องมาพบกัน และทำงานในขั้นตอนต่อไป เป็นรูปแบบที่มีการสร้างเงื่อนไขเพื่อเลือกทำงานในขั้นต่อไป โดยใช้เงื่อนไขการตัดสินใจเพียง 2 ทางเลือก เช่น

- จริง หรือ เท็จ
- ใช่ หรือ ไม่ใช่
- ถูก หรือ ผิด

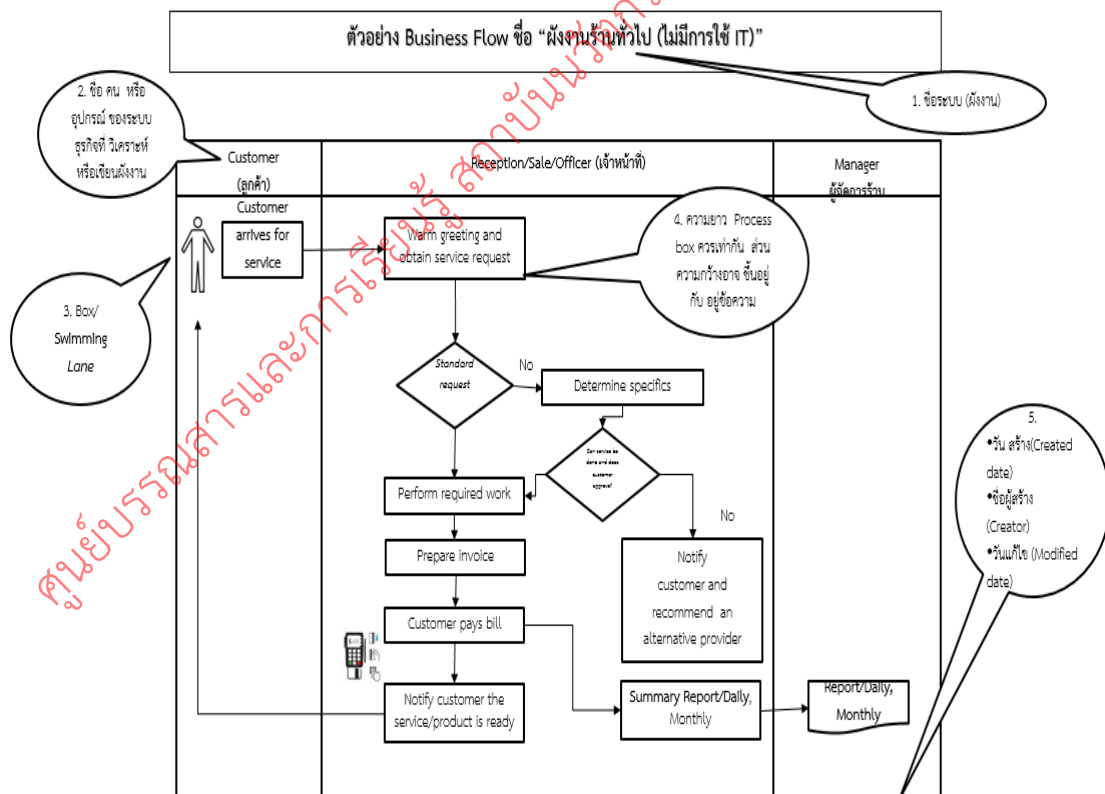


รูปที่ 4.4 ผังงานแบบมีเงื่อนไข SELECTION

3. ผังงานแบบทำซ้ำ (Iteration Flowchart) ดังรูปที่ 4.5 คือการทำซ้ำ เป็นการเขียน Flowchart ให้กลับมาทำงานในขั้นตอนอย่างเก่า จะเห็นว่า Flowchart มีลักษณะวน ซึ่ง เรียกว่า Loop และจะสังเกตว่า การวน Loop ดัง รูปที่ 4 จะไม่มีทางออกไปทำงานในขั้นตอนต่อไปได้เลย เพื่อที่จะทำให้ออกจาก Loop ได้จะต้องมีการ เช็คเพื่อออกจาก Loop ดังจะได้กล่าวต่อไป เป็นรูปแบบที่มีการกระทำกระบวนการหนึ่งหลายครั้ง โดยมีเงื่อนไขในการควบคุม จะแตกต่างกับรูปแบบมีเงื่อนไข ตรงที่เงื่อนไขจะถูกทดสอบจนกว่าจะเป็นเท็จถึงจะหยุดการทำซ้ำ



รูปที่ 4.5 ผังงานแบบทำซ้ำ ITERATION



Dr.Prajak Chertchom Created Date: 6-2-2565 Modified Date: DDMMYY

รูปที่ 4.6 ตัวอย่างผังงานผร้านทั่วไป (ไม่มีการใช้ IT)

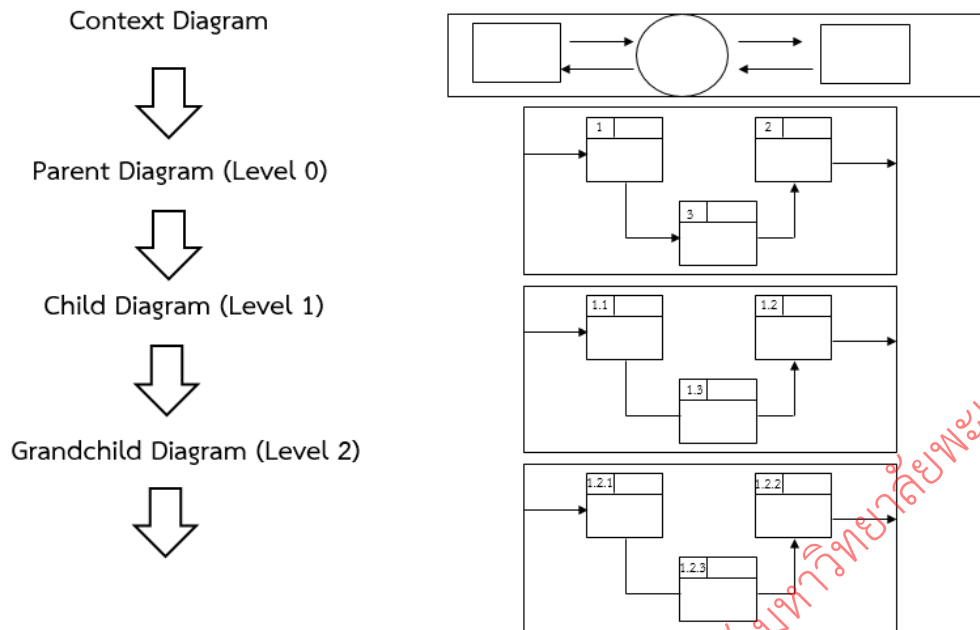
4.3.2 ตัวอย่างที่ 2 แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD) แผนภาพแสดงการไหลของข้อมูลจากแหล่งข้อมูลภายนอกเข้ามาในระบบ เพื่อแสดงวิธีการไหลของข้อมูลจากกระบวนการหนึ่งไปอีกกระบวนการหนึ่ง ซึ่งมีสัญลักษณ์ 4 แบบคือ สี่เหลี่ยม สี่เหลี่ยมมุมมน ลูกศร และสี่เหลี่ยมปลายเปิดข้างหนึ่ง โดยแผนภาพกระแสข้อมูลจะมีหลายระดับ (Level) ตามความลึกของข้อมูล

สัญลักษณ์ของ DFD

ทฤษฎี	External Entity	Process	Data Store	Data Flow
Yourdon				
Gane and Sarson				
SSADM				

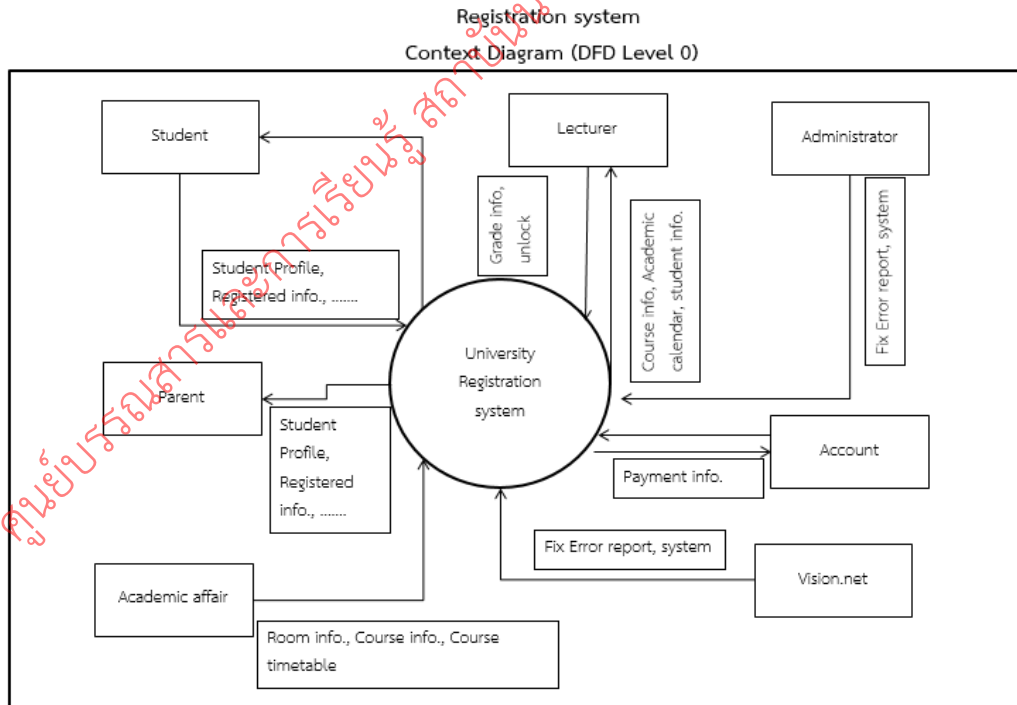
รูปที่ 4.7 สัญลักษณ์ของ (Data Flow Diagram: DFD)

ระดับของ แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level)



รูปที่ 4.8 ระดับของ แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level)

ตัวอย่างแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level 0) ของระบบทะเบียนนักศึกษา Level 0 DFD (event-partitioned system model or diagram 0) แสดง processes หลักๆของ system หรือ subsystem

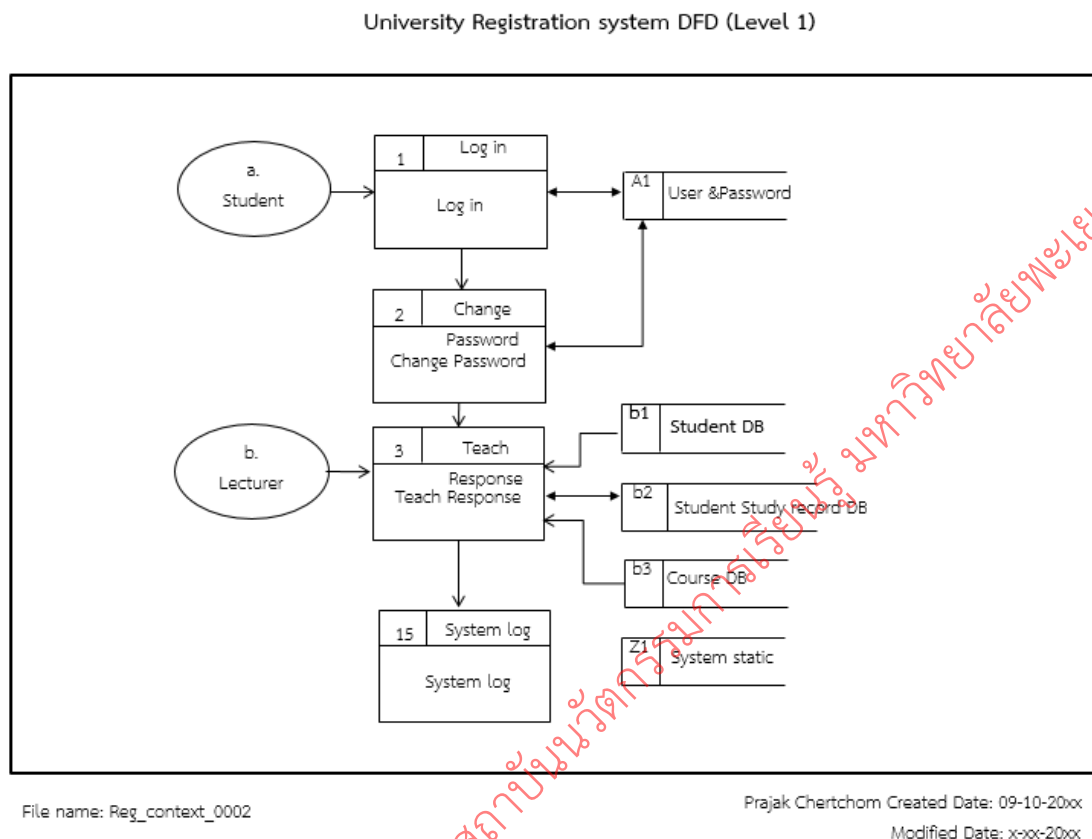


File name: Reg_context_0001

Prajak Chertchom Created Date: 09-10-20xx
Modified Date: x-xx-20xx

รูปที่ 4.9 ตัวอย่างแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level 0) ของระบบทะเบียนนักศึกษา

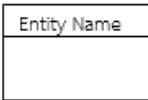
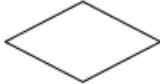
ตัวอย่างแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level 1) Level 1 DFD แสดงการแตกรายละเอียดกระบวนการ (Process) ใน DFD fragment ใดๆ ให้เห็นการประมวลผลในกระบวนการ (Process) ย่อยๆ

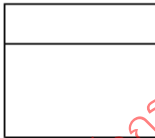
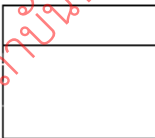
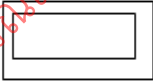


รูปที่ 4.10 ตัวอย่างแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD Level 1) ของระบบทะเบียนนักศึกษา

4.3.3 ตัวอย่างที่ 3 แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Entity Relationship Diagram: E-R Diagram) เป็นแบบจำลองข้อมูล (Data Model) คือแผนภาพที่แสดงการจำลองข้อมูลที่เกิดขึ้นทั้งหมดในระบบ พร้อมทั้งจำลองความสัมพันธ์ระหว่างข้อมูลที่เกิดขึ้นนั้น โดยใช้ แผนภาพ ER-Diagram ประกอบไปด้วย Entity (แทนกลุ่มของข้อมูลที่เป็นเรื่องเดียวกัน/เกี่ยวข้องกัน) และความสัมพันธ์ระหว่างข้อมูล (Data Relationship) ที่เกิดขึ้นทั้งหมดในระบบ

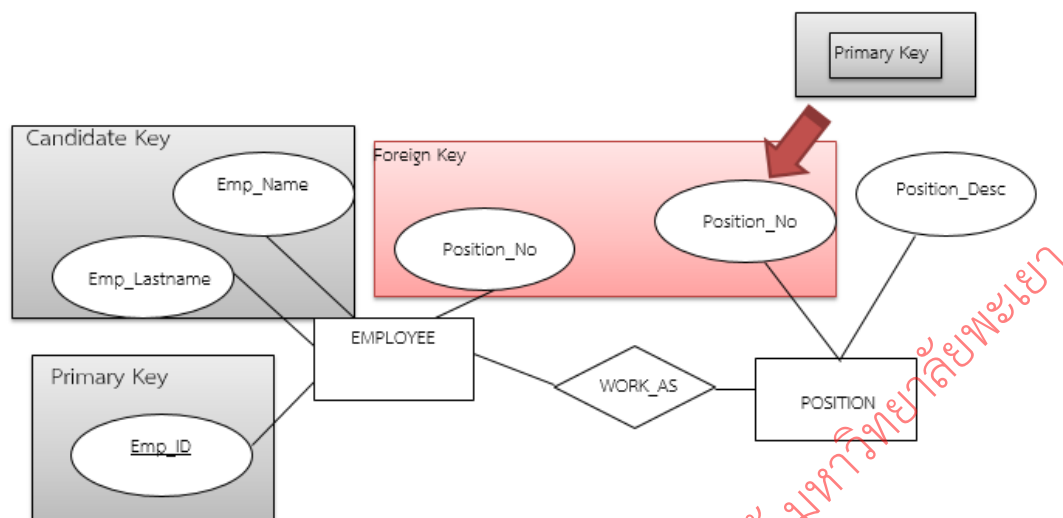
สัญลักษณ์ของ ER-Diagram แยกตามทฤษฎี 2 ประเภท (1. Chen Model 2. Crow's Foot Model)

Chen Model	Crow's Foot Model	ความหมาย
		ใช้แสดง Entity
		Relationship Line เส้นเชื่อมความสัมพันธ์ระหว่าง Entity
		Relationship ใช้แสดงความสัมพันธ์ระหว่าง Entity สำหรับ Crow's Foot Model ใช้ตัวอักษรเขียนแสดงความสัมพันธ์

Chen Model	Crow's Foot Model	ความหมาย
		Attribute ใช้แสดง Attribute ของ Entity
		ใช้แสดงคีย์หลัก (Identifier)
		Associative Entity
		Weak Entity

รูปที่ 4.11 ตัวอย่างแผนภาพกระแสดข้อมูลสัญลักษณ์ของ ER-Diagram

ตัวอย่าง E-R Diagram

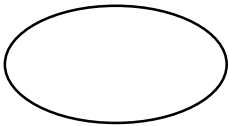
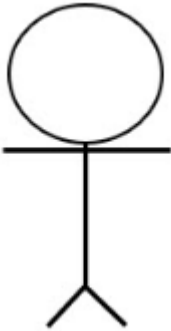
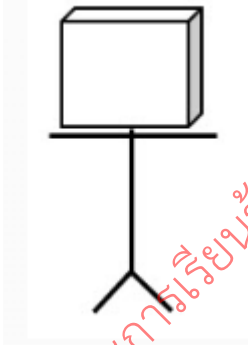


รูปที่ 4.12 ตัวอย่างแผนภาพ E-R Diagram

4.3.4 ตัวอย่างที่ 4 แผนภาพ Use Case Diagram คือแบบจำลองแผนภาพที่ใช้

ดงปฏิสัมพันธ์ระหว่างกิจกรรมที่เกิดขึ้นในระบบระบบงานและ Actor หรือผู้ที่กระทำกับระบบ และแสดงให้เห็นถึงส่วนประกอบทั้งหมด หรือ ภาพรวมของระบบ เป็นรากฐานในการเริ่มต้นการวิเคราะห์ระบบ โดยค้นหาว่าระบบทำอะไร โดยไม่สนใจกลไกการทำงานหรือเทคนิคการทำงาน โดยเป็นแบบจำลองพื้นฐานที่สามารถอธิบายสิ่งต่าง ๆ ได้โดยใช้รูปภาพที่ไม่ซับซ้อนดังรูป 4.12 และใช้สัญลักษณ์ ดังตาราง 4.11

ตาราง 4.11 สัญลักษณ์ Use Case

สัญลักษณ์	คำอธิบาย
	วงรี ใช้แทน Use Case
	คน แทน Actor
	หัวบล็อก ใช้แทน Actor ที่เป็นระบบ
	สัญลักษณ์แทน User
	สัญลักษณ์แทน Extends

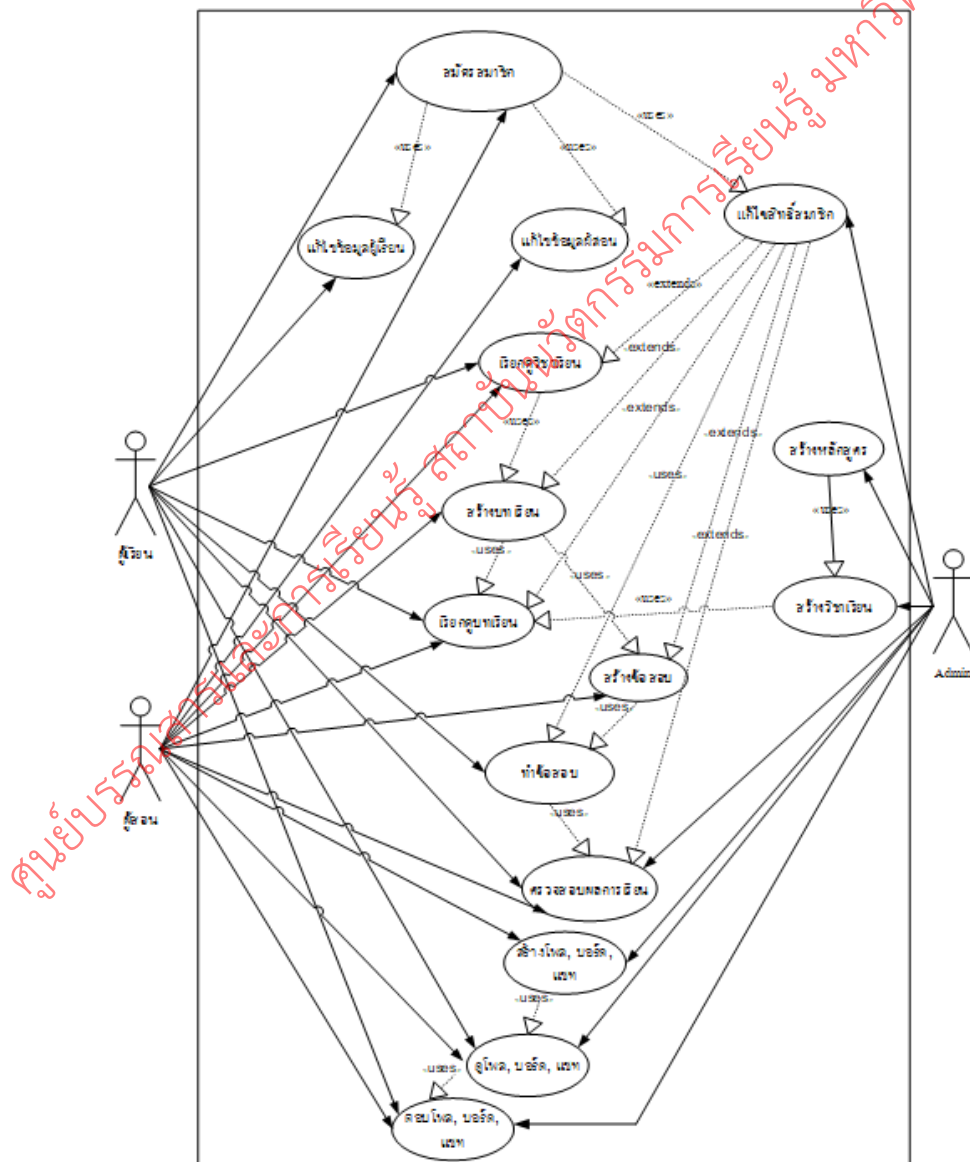
ความสัมพันธ์ของ Use Case มี 2 แบบคือ

1. User
2. Extends

ความหมายของ User คือ การที่ User ทำการเรียกใช้งานโปรแกรม ซึ่งแสดงด้วยสัญลักษณ์ Use Case โดยลูกศรหัวสามเหลี่ยมใส่ ซี่ไปยัง Use Case ที่ถูกเรียกใช้งาน และมีคำว่า << user >> กำกับอยู่บนเส้น ส่วน Use case และ Actor ของระบบ (ผู้จัดการระบบ-System Administrator) คือ การทำงาน 2 ประเภท คือ 1. การตรวจสอบ User (Validate user) และการตรวจสอบรหัสผ่าน (Check password)

สำหรับการใช้สัญลักษณ์ Extends หมายถึงการที่ Use case หนึ่งไปมีผลต่อการทำงานตามปกติของอีก Use case หนึ่ง หรือการทำงานของโปรแกรมใดโปรแกรมหนึ่งถูกรบกวนหรือมีการสะดุด หรือมีการเปลี่ยนแปลงกิจกรรมไป

ตัวอย่างแผนภาพ E-R Diagram ระบบ e-Learning



รูปที่ 4.13 ตัวอย่างแผนภาพ E-R Diagram ระบบ e-Learning

จากรูปที่ 4.13 แสดงกิจกรรมที่เกิดขึ้นกับผู้ใช้งานในระบบซึ่งประกอบด้วย Admin, ผู้สอน และผู้เรียน ดังต่อไปนี้

1. ผู้สอนและผู้เรียนทำการสมัครสมาชิก เพื่อเข้าสู่การทำงานภายในระบบ
2. ผู้เรียนสามารถทำการแก้ไขข้อมูลส่วนตัวได้
3. ผู้สอนสามารถทำการแก้ไขข้อมูลส่วนตัวได้
4. Admin ทำการแก้ไขสิทธิ์และอนุมัติสิทธิ์เกี่ยวกับการเรียนให้กับสมาชิก
5. Admin จะทำการสร้างหลักสูตรเพื่อแบ่งหมวดหมู่ก่อนการสร้างรายวิชาเรียน
6. Admin จะเป็นผู้สร้างวิชาเรียนในแต่ละวิชา
7. ผู้เรียนและผู้สอนสามารถเรียกดูข้อมูลวิชาเรียนได้
8. ผู้สอนทำการสร้างบทเรียนภายใต้วิชาที่ได้รับผิดชอบหรือได้รับสิทธิ์ในการดูแลวิชานั้นๆ
9. ผู้เรียนและผู้สอนสามารถเรียกดูข้อมูลบทเรียนได้ ซึ่งผู้เรียนจะศึกษาบทเรียนในรายวิชาที่ได้รับสิทธิ์ในการเรียนวิชานั้น
10. ผู้สอนสามารถสร้างข้อสอบในรายวิชาที่รับผิดชอบ
11. ผู้เรียนจะสามารถทำข้อสอบในวิชาเรียนที่ผู้เรียนได้สิทธิ์ในการเรียนวิชานั้น
12. Admin ผู้สอน และผู้เรียนสามารถเข้าไปตรวจสอบรายงานข้อมูลผลการเรียนได้ โดย Admin สามารถตรวจสอบข้อมูลสรุปการเรียนการสอนทั้งหมด และผู้สอนสามารถตรวจสอบข้อมูลสรุปการสอนของตนเอง และผู้เรียนสามารถตรวจสอบข้อมูลสรุปผลการเรียนของตนเอง
13. Admin และผู้สอนสามารถสร้างข่าวหรือแก้ไขข่าวสาร
14. ผู้สอนสามารถตรวจสอบข้อมูลข่าวสารต่างๆ ของทางระบบที่ Admin และผู้สอนได้ทำการสร้างขึ้น
15. ผู้สอนสามารถสร้างโพล บอร์ด และแชท ในแต่ละรายวิชาได้
16. ผู้เรียนและผู้สอนสามารถเข้าดูและตอบโพล บอร์ด และแชท ในแต่ละรายวิชาได้

4.3.5 ตัวอย่างที่ 5 Data Dictionary Data Dictionary จะกระทำควบคู่ไปกับการเขียน Data Flow Diagram เพื่อระบุรายละเอียดของข้อมูลต่างๆในแฟ้มข้อมูลซึ่งอยู่ใน Data Flow Diagram

Data Dictionary จะประกอบไปด้วย ส่วนประกอบพื้นฐานที่ไม่สามารถแบ่งแยกให้เล็กลงไปได้ อีก นั่นก็คือ ส่วนย่อยที่สุดของข้อมูล (Data Element) ตั้งแต่ 1 ตัวขึ้นไปโครงสร้างข้อมูล (Data Structure) ที่ประกอบไปด้วยหลายๆส่วนย่อยของข้อมูล (Data Element) ดังตาราง

ตาราง 4.12 สัญลักษณ์ Data Dictionary

Table Name : Member				
Description : ตารางสมาชิก คือตารางรายชื่อและรายละเอียดต่างๆของสมาชิก				
File	Data Type	Length	Constraint	Description
Member_ID	Int	11	PK	รหัสสมาชิก
M_FName	Varchar	100	-	ชื่อจริง
M_LName	Varchar	100	-	นามสกุล
M_Address	Varchar	200	-	ที่อยู่
M_Amphur	Varchar	50	-	เขต/อำเภอ
M_Province	Varchar	50	-	จังหวัด
M_ZipCode	Char	5	-	รหัสไปรษณีย์
M_Phone	Varchar	50	-	หมายเลขโทรศัพท์
UserName	Varchar	25	FK	ชื่อผู้ใช้สมาชิก

4.3.6 ตัวอย่างที่ 6 MINISPECIFICATION และ Pseudocode จาก Data Flow Diagram ข้อมูลนำเข้า (Input) เป็นการเขียนอธิบายการตัดสินใจหรือเงื่อนไขต่างๆ ซึ่งกระบวนการตัดสินใจว่าทำอะไรในแต่ละหน่วยประมวลผลหรือโปรแกรมนี้จะอธิบายโดย Minispecification หรือ Pseudocode เป็นการเขียนอธิบายอัลกอริทึมและลอจิก (Algorithm and Logic) ต่าง ๆ ของโปรแกรมที่เราจะเขียน โดยเขียนคำสั่งที่จำลองความคิดเป็นลำดับขั้นตอนโดยใช้ประโยคภาษาอังกฤษอธิบายอย่างง่าย ดังนั้นการเขียนโปรแกรมที่ดีจึงต้องการเขียนขั้นตอนเหล่านี้ก่อน เพื่อให้มองภาพรวมของงานทั้งหมด ดังนั้นถ้าเป็นไปได้ ควรที่จะเขียน Flowchart และ MINISPECIFICATION และ Pseudocode กันก่อนที่จะเริ่มเขียนโปรแกรมจริงกัน ดังตัวอย่างที่ 1

```

if "1"
    print response
    "I am case 1"
if "2"
    print response
    "I am case 2"

```

ตัวอย่างที่ 2

Algorithm Cal_1

Variables: mLoop, Sum, Score, average

```
Begin
<Label Text="Start developing now" FontSize="Title"
    Input mLoop
    Sum = 0
    For I = 1 to mLoop
        Input Score
        Sum = Sum + Score
    Next
    average = Sum / mLoop
    Print average
End Cal_1
```

4.3.7 ตัวอย่างที่ 7 แบบจำลองเชิงวัตถุ (Object Modeling) หรือภาษาแบบจำลองเชิงวัตถุ เรียกว่ายูเอ็มแอล (Unified Modeling Language-UML) เป็นเครื่องมือเช่นเดียวกับการใช้แผนภาพ ดีเอฟดีและอีอาร์ดี ในการวิเคราะห์เชิงโครงสร้าง โดยที่ยูเอ็มแอลประกอบด้วยเครื่องมือสร้างแผนภาพที่แสดงมุมมองต่าง ๆ ในการอธิบาย แสดงรายละเอียดจำลองการสร้างและการจัดการกับเอกสารต่าง ๆ ในระบบ

แผนภาพยูเอ็มแอล จะมีสัญลักษณ์ต่าง ๆ ใช้อธิบาย ประกอบด้วยแผนภาพใช้ในการวิเคราะห์ และออกแบบ ตัวอย่างเช่น

1. Use case diagram แสดงความต้องการของระบบ
2. Class diagram แสดงองค์ประกอบและความสัมพันธ์ของคลาส
3. Object diagram แสดงองค์ประกอบและความสัมพันธ์ของวัตถุส่วนที่แสดงพฤติกรรมและบทบาท
4. Sequence diagram จำลองพฤติกรรมของคลาสต่าง ๆ หรือขั้นตอนการทำงาน
5. Collaboration diagram จำลองพฤติกรรมหรือความสัมพันธ์เชิงกิจกรรม
6. State chart diagram แสดงเหตุการณ์ของแต่ละสถานะ

7. Activity diagram แสดงขั้นตอนของกิจกรรมในการปฏิบัติงานภาพ
ใช้ในการสร้างระบบ
8. Component diagram แสดงโครงสร้างทางกายภาพสำหรับการเขียน
โปรแกรม
9. Deployment diagram แสดงความสัมพันธ์ของชุดโปรแกรมและ
ฮาร์ดแวร์ที่ใช้

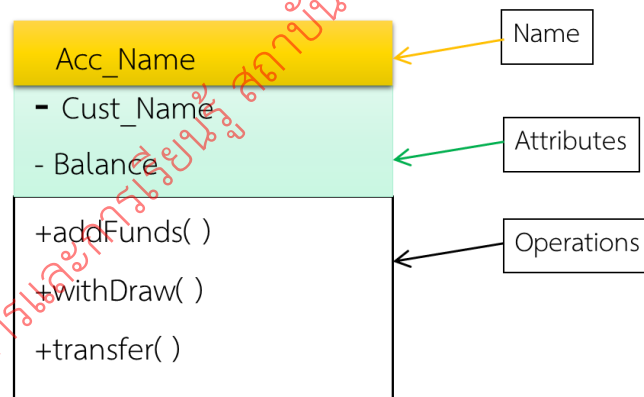
ตัวอย่าง Class diagram แสดงองค์ประกอบและความสัมพันธ์ของคลาส ซึ่งแต่ละ Class จะแสดงด้วยสี่เหลี่ยมผืนผ้าที่แบ่งออกเป็นสามช่อง คือ Name Attributes และ Operations

Modifiers ใช้เพื่อระบุการมองเห็นแอตทริบิวต์และการดำเนินการ ประกอบไปด้วย

‘+’ ใช้เพื่อแสดงถึงการเปิดเผยต่อสาธารณะ (is used to denote Public visibility (everyone)

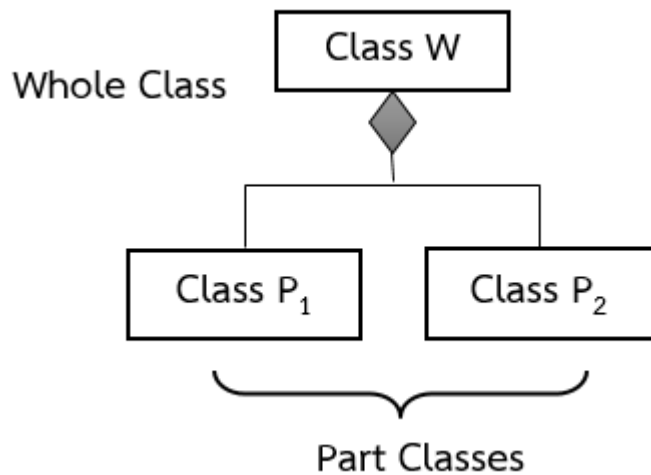
‘#’ ใช้เพื่อแสดงถึงการมองเห็นที่ได้รับการป้องกัน is used to denote Protected visibility (friends and derived)

‘-’ ใช้เพื่อแสดงถึงการมองเห็นส่วนตัว is used to denote Private visibility (no one)

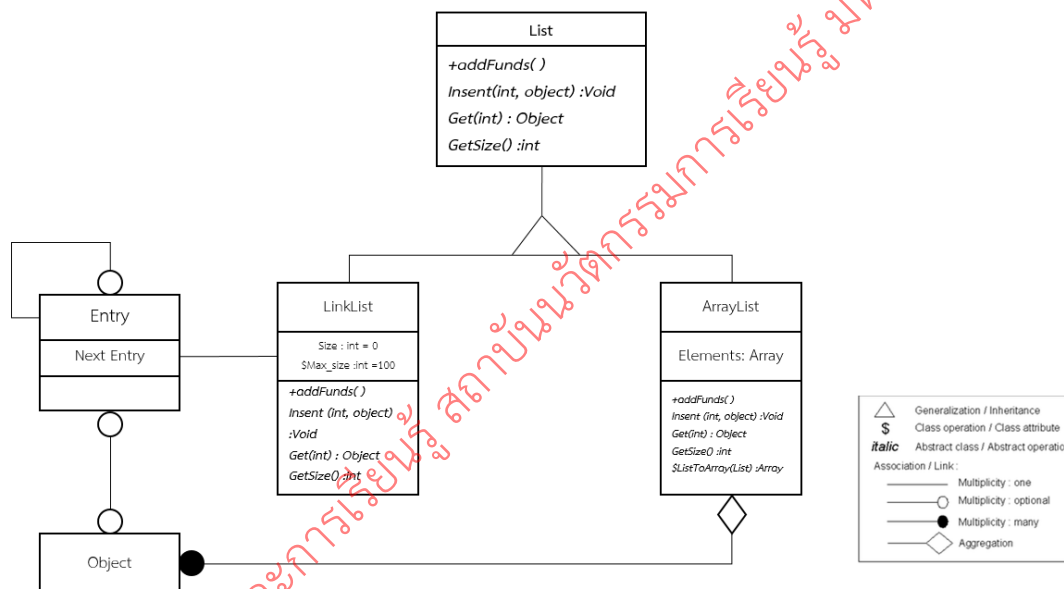


รูปที่ 4.14 ตัวอย่างแผนภาพ Class diagram

โดยการเขียนจะแสดงความสัมพันธ์เชิงกิจกรรมและจำลองพฤติกรรมของคลาสต่าง ๆ หรือขั้นตอนการทำงานดังรูป 4.15



รูปที่ 4.15 ตัวอย่างแผนภาพ Class diagram จำลองความสัมพันธ์



รูปที่ 4.16 ตัวอย่างโมเดลแบบจำลองเชิงวัตถุ (Modelling object)

โดยสรุป การสร้างแบบจำลอง (Model) โดยการทำงานไปพร้อมๆกัน เพื่อให้ได้แบบจำลองการวิเคราะห์ที่เป็นโครงสร้าง (Structure Analysis Model) โดยเริ่มจาก Flowchart Diagram ที่แสดงถึงภาพรวมของระบบ (Over-View of System) จากนั้นจึงสร้าง Data Flow Diagram (DFD) เพื่อแสดงการไหลของข้อมูล จากหน่วยประมวลผลหรือโปรแกรมหนึ่งไปยังหน่วยประมวลผลอื่นๆ นอกจากนั้นยังรวมถึงหน่วยเก็บข้อมูล (Data Store) และตัวแปรภายนอก (External Entities) จากขั้นตอนนี้เอง เราได้เรียนรู้การติดตามการไหลของข้อมูลจากหน่วยประมวลผลต่างๆ และกำหนดชื่อให้กับข้อมูลแต่ละตัว ขั้นตอนต่อไปก็คือการอธิบายรายละเอียดหรือขยายความข้อมูลโดยการสร้าง Data

Dictionary ซึ่งจะได้ออกมาเป็นโครงสร้างข้อมูล (Data Structure) จากนั้นจึงนำมาทำการลดความซ้ำซ้อนลง ในขณะเดียวกันเราก็สามารถสร้าง System Structure Chart และกำหนดความสัมพันธ์ระหว่างโครงสร้างข้อมูล (Data Structure) โดยการสร้าง Data Structure Diagram และ Data Access Diagram ในอีกด้านหนึ่งก็อธิบายการทำงานของแต่ละหน่วยประมวลผลใน Data Flow Diagram ด้วย Use Case Diagram และ Minispecification หรือ Psudocode ตามด้วย Object Modeling เมื่อทำงานทุกขั้นตอนเสร็จ เราก็จะได้แบบจำลองของระบบหรือที่เรียกว่า System Model

สรุปเนื้อหาบทที่ 4 "แนวทางการออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศสมัยใหม่" เป็นการเสนอแนะแนวทางและขั้นตอนในการออกแบบสถาปัตยกรรมระบบสารสนเทศที่เหมาะสมสำหรับธุรกิจในยุคสมัยใหม่ โมเดลที่ใช้ในการวิเคราะห์และออกแบบระบบ โดยการใช้เทคโนโลยีและความรู้ทางเทคนิคเพื่อสร้างระบบสารสนเทศที่มีประสิทธิภาพและสามารถตอบสนองต่อความต้องการขององค์กรได้อย่างเหมาะสม รวมถึงทำความเข้าใจเครื่องมือแบบจำลองสำหรับการออกแบบสถาปัตยกรรมสารสนเทศ

1. **การเข้าใจความต้องการของระบบสารสนเทศ** ขั้นตอนแรกเน้นให้เข้าใจความต้องการและเป้าหมายของระบบสารสนเทศ ซึ่งรวมถึงการระบุความต้องการของผู้ใช้งานและองค์กร เพื่อให้การออกแบบเป็นไปในทิศทางที่ถูกต้อง
2. **การกำหนดและสร้างโครงสร้าง** กระบวนการนี้เน้นการกำหนดและสร้างโครงสร้างสถาปัตยกรรมของระบบสารสนเทศ ซึ่งรวมถึงการออกแบบฐานข้อมูล การกำหนดหน้าที่และฟังก์ชันของแต่ละส่วน และการสร้างเครื่องมือต่างๆ ที่เหมาะสม
3. **การเลือกเทคโนโลยีและแพลตฟอร์ม** ขั้นตอนนี้เน้นการเลือกและนำเข้าเทคโนโลยีที่เหมาะสมกับระบบสารสนเทศ เพื่อให้ระบบสามารถทำงานได้อย่างมีประสิทธิภาพและเชื่อมต่อกับองค์กรได้
4. **การสร้างและออกแบบข้อมูล** การสร้างและออกแบบโครงสร้างข้อมูลที่เหมาะสมสำหรับการเก็บรวบรวมและประมวลผลข้อมูลในระบบสารสนเทศ ซึ่งรวมถึงการสร้างฐานข้อมูลและการกำหนดวิธีการจัดเก็บข้อมูล
5. **การออกแบบส่วนประสานผู้ใช้** การสร้างและออกแบบอินเทอร์เฟซให้ผู้ใช้งานให้ใช้งานได้อย่างง่ายและมีประสิทธิภาพ รวมถึงการจัดการกระบวนการทำงานของผู้ใช้ในระบบ
6. **การประเมินและทดสอบ** ขั้นตอนนี้เน้นการประเมินและทดสอบระบบสารสนเทศที่ได้ออกแบบไป โดยการตรวจสอบความถูกต้องและประสิทธิภาพของระบบ

7. การปรับปรุงและพัฒนาต่อเนื่อง หลังจากการใช้งานจริง เน้นการนำเสนอและการตอบรับจากผู้ใช้ เพื่อทำการปรับปรุงและพัฒนาระบบสารสนเทศต่อไป

โดยบทที่ 4 เป็นการนำเสนอแนวทางการออกแบบสถาปัตยกรรมสารสนเทศสำหรับระบบสารสนเทศในยุคสมัยใหม่ โดยเน้นการเลือกและนำเข้าเทคโนโลยีที่เหมาะสม การสร้างโครงสร้างและอินเทอร์เน็ตที่สอดคล้องกับความต้องการขององค์กร

คำศัพท์ บทที่ 4

Agile Software Development	วิธีการพัฒนาซอฟต์แวร์แบบคล่องแคล่วรวดเร็ว
Artificial intelligence	ปัญญาประดิษฐ์
Blueprint	พิมพ์เขียว
Data Dictionary	พจนานุกรมข้อมูล
Data Flow Diagram - DFD	แผนภาพกระแสข้อมูล
Data gathering	เทคนิคการรวบรวมข้อมูล
Decision table	ตารางตัดสินใจ
Decision tree	ผังต้นไม้
Developing and Documenting Software	การพัฒนาซอฟต์แวร์และการจัดทำเอกสาร
End-to-End Process	กระบวนการที่ต่อเนื่องจากต้นสู่ปลาย
Engagement Model	รูปแบบพันธสัญญาสู่ความสำเร็จ
Entity-relationship diagram - ERD	แผนภาพความสัมพันธ์ระหว่างเอนทิตี
Flowchart	ผังงาน
Macroeconomic environments	ปัจจัยแวดล้อมทางเศรษฐกิจมหภาค
Methodologies	วิธีการ
Operating Model	กรอบโมเดลการปฏิบัติงาน
Project management	การจัดการโครงการ
Pseudocode	ภาษาการเขียนอธิบายอัลกอริทึมและลอจิกอย่างง่าย
Spiral Model	โมเดลโครงสร้างแบบก้นหอย
Structured Language	ภาษาโครงสร้าง
Techniques	เทคนิค
Unified Modelling Language	ยูเอ็มแอล
Waterfall Model	โมเดลโครงสร้างแบบน้ำตก

คำถามและแบบฝึกหัดท้ายบทที่ 4

1. จงอธิบายปัจจัยเพื่อพิจารณาสำหรับการบริหารโครงการสารสนเทศสมัยใหม่ มีอะไรบ้าง
2. จงอธิบายกรอบการบริหารโครงการสารสนเทศสมัยใหม่ มีอะไรบ้าง
3. จงอธิบายผังงาน (Flowchart) คืออะไร
4. จงอธิบายลักษณะของผังงานที่ดี มีอะไรบ้าง จงยกตัวอย่าง
5. จงอธิบายสัญลักษณ์เหล่านี้
6. จงอธิบายรูปแบบของผังงาน มีกี่แบบ
7. จงอธิบายแผนภาพกระแสข้อมูล (Data Flow Diagram: DFD)
8. จงอธิบายแผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Entity Relationship Diagram : E-R Diagram)
9. จงอธิบายแผนภาพ Use Case Diagram

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

บทที่ 5 โมเดลการพัฒนาและมาตรฐานสำหรับการสร้างระบบสารสนเทศสมัยใหม่

(Developing Model and Standard for New Information System Development)

การเข้าใจกระบวนการ โมเดลการพัฒนาและมาตรฐานสำหรับการสร้างระบบสารสนเทศสมัยใหม่ จะช่วยลดช่องว่างระหว่าง As-Is กับ To-Be ของการพัฒนาระบบสารสนเทศ ในการพัฒนาโครงการระบบสารสนเทศสมัยใหม่ จะต้องมีการพิจารณาร่วมกันระหว่างผู้บริหารและทุกฝ่ายในองค์กรแล้ว ยังต้องพิจารณาความต้องการของผู้รับบริการและผู้มีส่วนได้ส่วนเสียด้วย

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

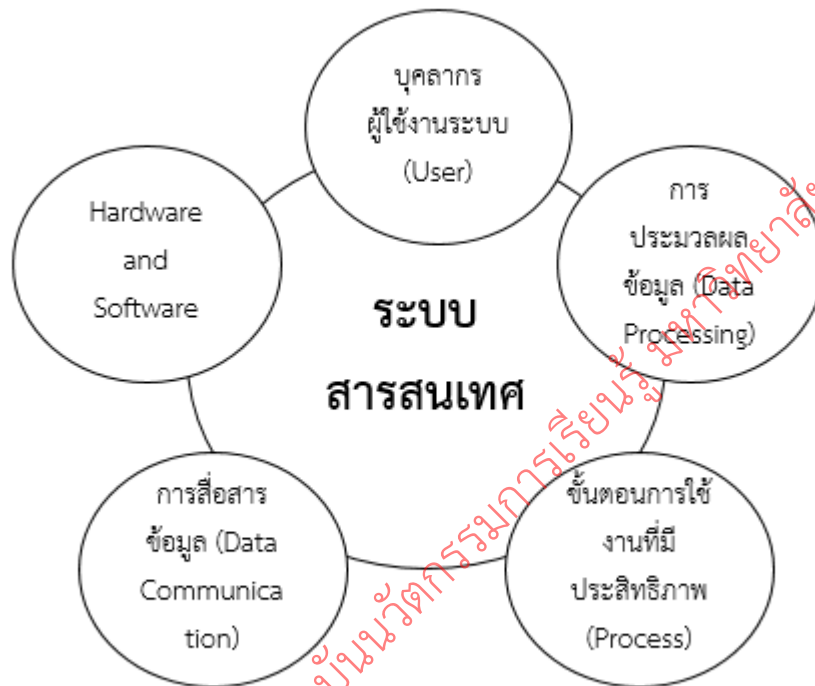
1. สามารถเข้าใจกรรมวิธีการพัฒนาซอฟต์แวร์มาใช้ เพื่อให้กระบวนการพัฒนาซอฟต์แวร์นั้นมีมาตรฐาน
2. สามารถประยุกต์ใช้แนวคิดที่สำคัญของโมเดลแบบจำลองต่างๆ เพื่อออกแบบสถาปัตยกรรมองค์กรและสำหรับการพัฒนาระบบสารสนเทศสมัยใหม่

5.1 บทนำการสร้างระบบสารสนเทศสมัยใหม่ ระบบสารสนเทศเพื่อการจัดการสมัยใหม่เป็นระบบที่บูรณาการด้านธุรกิจโดยนำเอาเทคโนโลยีดิจิทัล (Digital Technology) มาช่วยสนับสนุนการดำเนินงานให้เกิดประสิทธิภาพสูงสุด โดยทำหน้าที่รวบรวม ประมวลผล จัดเก็บและแจกจ่ายข้อมูลต่างๆ ที่ใช้ในการตัดสินใจ สำหรับการวางแผนการจัดการและควบคุมองค์กรทางธุรกิจ ช่วยในการทำงานประจำวันและการวางแผนงานในอนาคต ทั้งด้านการผลิตสินค้าและการให้บริการ นอกจากนี้ฟังก์ชันต่างๆ ของระบบ จะต้องตอบสนองทางธุรกิจ ช่วยสนับสนุนกระบวนการทำงาน บันทึกเก็บรวบรวมข้อมูลที่ใช้ บุคลากรที่เกี่ยวข้อง เทคโนโลยี และ และสอดคล้องกับโครงสร้าง วัสดุทัศน์ พันธกิจ ยุทธศาสตร์องค์กร

ระบบสารสนเทศจะสมบูรณ์แบบต้องประกอบไปด้วย 5 องค์ประกอบ ดังรูป 5.1 คือ

1. บุคลากรผู้ใช้งานระบบ (User)
 2. การประมวลผลข้อมูล (Data Processing)
 3. การสื่อสารข้อมูล (Data Communication)
 4. ที่จัดเก็บข้อมูลและการเข้าถึงข้อมูล เพื่อที่จะเก็บข้อมูลในรูปแบบที่เหมาะสม
- และเพื่อให้แน่ใจว่าข้อมูลข่าวสารสามารถนำมาใช้ได้เมื่อต้องการ (Hardware and Software)

5. ขั้นตอนการใช้งานที่มีประสิทธิภาพ เพื่อรวบรวมข้อมูลทุกส่วน การประมวลผลข้อมูล การสื่อสารข้อมูลที่จัดเก็บข้อมูลและการเข้าถึงข้อมูล รวมทั้งผู้เข้าใช้เป็นการจัดการสารสนเทศที่มีประสิทธิภาพ (Process)



รูปที่ 5.1 องค์ประกอบของระบบสารสนเทศ

โดยการสร้างและพัฒนาระบบสารสนเทศสมัยใหม่นั้น ทีมพัฒนาควรใช้วงจรการพัฒนา ระบบงานหรือขั้นตอนการพัฒนาซอฟต์แวร์ (Software Development) ควบคู่ไปกับการประยุกต์ใช้ มาตรฐาน COBIT (Control Objectives for Information and Related Technology)

โดยขั้นตอนการพัฒนาซอฟต์แวร์ (Software) การพัฒนาคือแนวทางการสร้างผลิตภัณฑ์ให้ ตรงตามข้อกำหนดด้านสารสนเทศ ด้วยการนำกรรมวิธีการพัฒนาซอฟต์แวร์มาใช้ เพื่อให้ กระบวนการพัฒนาซอฟต์แวร์นั้นมีมาตรฐาน และตัวผลิตภัณฑ์ซอฟต์แวร์มีคุณภาพ มีขั้นตอนการ ตรวจสอบความถูกต้อง (Software Validation) ของ software ให้ตรงกับความต้องการ เพื่อให้แน่ใจว่า ซอฟต์แวร์ที่พัฒนานี้เป็นผลิตภัณฑ์ที่ตรงกับความต้องการของลูกค้าหรือผู้ใช้งาน รวมถึงกระบวนการทำ ให้ระบบสารสนเทศ มีวิวัฒนาการ เป็นการปรับเปลี่ยนเพิ่มสิ่งดีๆเข้ามาและเอาสิ่งที่ไม่ดีหรือไม่จำเป็น ออกไป (Software Evolution) ซึ่งจะเปลี่ยนแปลงไปตามเทคโนโลยีหรือตามความต้องการของผู้ใช้ ซึ่ง

ซอฟต์แวร์ควรออกแบบไว้เพื่อรองรับวิวัฒนาการที่สามารถเปลี่ยนแปลงไปตามความต้องการของผู้ใช้งานได้อย่างเหมาะสม ซึ่งคุณสมบัติซอฟต์แวร์ที่มีคุณภาพ ประกอบด้วย มีความถูกต้อง, มีความน่าเชื่อถือ, ใช้งานง่าย, บำรุงรักษาง่าย, สามารถนำกลับมาใช้งานใหม่ได้, มีความคงทน, มีประสิทธิภาพ, มีความสะดวกในการเคลื่อนย้าย และมีความปลอดภัย โดยคุณสมบัติของระบบสารสนเทศที่ดี คือ

- Understandability มีขอบเขตกระบวนการที่ชัดเจนและง่ายต่อความเข้าใจ
- Visibility สนับสนุนให้กิจกรรม กระบวนการธุรกิจทำงานอย่างมีประสิทธิภาพที่สุด
- Supportability เครื่องมือช่วยในการวิเคราะห์ สามารถช่วยสนับสนุนกิจกรรม กระบวนการตัดสินใจในทุกระดับ
- Acceptability กระบวนการที่มีการตรวจสอบ ยอมรับ โดยผู้ใช้
- Reliability กระบวนการถูกออกแบบในแนวทางที่หลีกเลี่ยงที่จะส่งต่อความผิดพลาดของซอฟต์แวร์
- Robustness กระบวนการสามารถทำงานต่อได้ แม้จะมีปัญหาที่ไม่ได้คาดการณ์ไว้เกิดขึ้น
- Maintainability กระบวนการสามารถวิวัฒนาการเพื่อตอบสนองต่อการเปลี่ยนแปลงความต้องการขององค์กร
- Rapidity กระบวนการสามารถทำให้ส่งมอบผลิตภัณฑ์ได้รวดเร็ว ตรงตามคุณลักษณะที่ได้กำหนดไว้

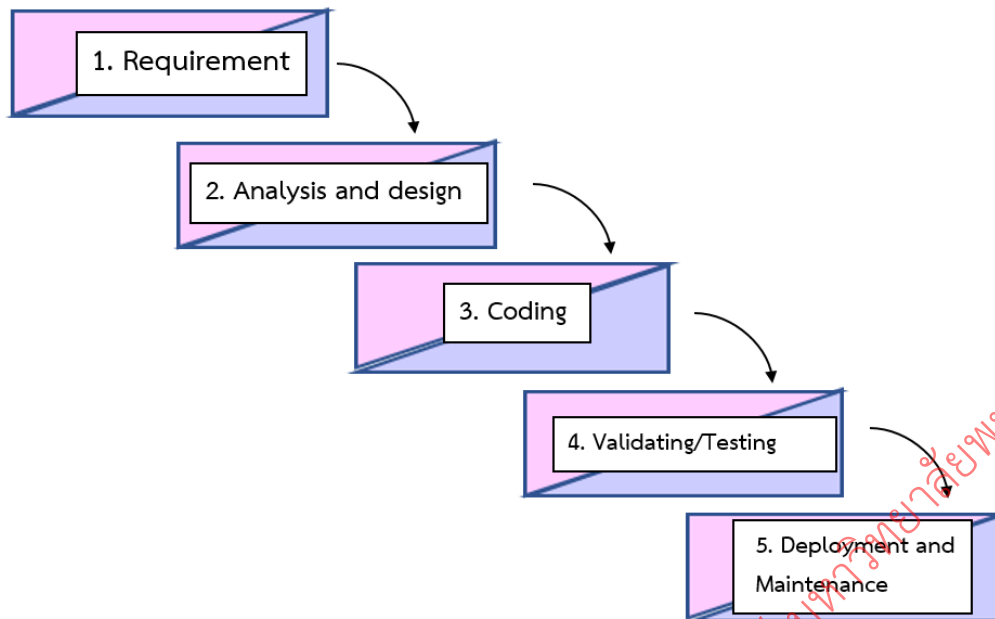
ทั้งนี้โครงการพัฒนาระบบสารสนเทศควรใช้แนวทางตามมาตรฐาน COBIT มากำกับควบคุม ซึ่งเป็นแนวทางในการสร้างความมั่นใจว่าการใช้ทรัพยากรด้านเทคโนโลยีสารสนเทศ เพื่อให้สอดคล้องกับวัตถุประสงค์เชิงธุรกิจขององค์กร (Business Objectives) และเพื่อให้เกิดการใช้ทรัพยากรอย่างมีประสิทธิภาพอันจะส่งประโยชน์สูงสุดแก่องค์กร รวมทั้งลดความเสี่ยงด้านเทคโนโลยีสารสนเทศ (IT Risk) และได้รับผลตอบแทนต่อการลงทุน (Return of Investment : ROI) ซึ่งกรอบมาตรฐานของ COBIT ประกอบด้วย 5 กลุ่มใหญ่

1. การจัดวางกลยุทธ์ (Strategic Alignment)
2. การนำเสนอคุณค่า (Value Delivery)
3. การจัดการทรัพยากร (Resource Management)
4. การจัดการความเสี่ยง (Risk Management)
5. การวัดผลการดำเนินงาน (Performance Measurement)

5.2 โมเดลขั้นตอนการพัฒนาซอฟต์แวร์ (Software Development Process Models) คือ แบบจำลองกระบวนการที่ใช้ในการผลิตระบบสารสนเทศ ในแบบจำลองนี้จะเป็นขั้นตอนหรือกิจกรรมสำหรับการพัฒนาระบบซอฟต์แวร์หรือระบบสารสนเทศ ซึ่งจากอดีตจะมีขั้นตอนการพัฒนาอยู่ 2 แบบใหญ่ๆ คือ การพัฒนาซอฟต์แวร์แบบไม่เป็นระบบและการพัฒนาซอฟต์แวร์แบบเป็นระบบ โดยการพัฒนาซอฟต์แวร์แบบไม่เป็นระบบนั้นไม่เป็นที่ยอมรับและสนับสนุนในการใช้พัฒนาระบบสารสนเทศสมัยใหม่ เพราะเป็นการพัฒนาซอฟต์แวร์แบบไม่มีการวางแผน ไม่มีการจัดการ ไม่มีการทำเอกสาร วิธีการพัฒนาระบบแบบนี้คือ เขียนโปรแกรมโดยคิดอะไรได้ก็ลงมือเขียน แล้วทำทดสอบการทำงาน ถ้าโปรแกรมไม่สามารถทำงานตามที่คิด ก็แก้ไขเขียนใหม่ เรียกการพัฒนาแบบนี้ว่า Coding & Debug

ในโลกธุรกิจสมัยใหม่ ซอฟต์แวร์หรือระบบสารสนเทศมีขนาดใหญ่และซับซ้อนขึ้น มีระบบข้อมูลขนาดใหญ่ (Big Data) มีการเชื่อมโยงแลกเปลี่ยนข้อมูล ซึ่งต้องใช้หลายคน หลายทีมช่วยกันพัฒนา ซึ่งตั้งแต่อดีตจนถึงปัจจุบันมีหลายทฤษฎีและแนวทางขึ้นอยู่กับปัจจัยหลายอย่าง เช่น ขนาดของโครงการ ข้อกำหนดและกติกาของบริษัทผู้ผลิตซอฟต์แวร์หรือระบบสารสนเทศ โดยรูปแบบของโมเดลขั้นตอนการพัฒนาระบบสารสนเทศมี 4 แบบใหญ่ๆ คือ (Laudon & Laudon, 2013).

5.1.1 แบบ Linear Model เป็นแบบจำลองที่มีขั้นตอนเป็นลำดับ เป็นโมเดลพื้นฐานในการพัฒนาซอฟต์แวร์และระบบสารสนเทศ โดยต้องทำขั้นตอนแรกให้สำเร็จจึงจะดำเนินการขั้นตอนต่อไปได้ แต่ละขั้นจะถูกกำหนดงานและผลผลิตได้จากขั้นตอนนั้นๆ อย่างชัดเจนเช่น โมเดลน้ำตก (Waterfall Model) ข้อเสียจากกระบวนการนี้ คือลูกค้าไม่มีโอกาสที่จะได้ตรวจสอบหรือทบทวน (Review) ระบบเลยจะเห็นอีกทีเมื่อระบบเสร็จสิ้นทำให้โอกาสความผิดพลาดสูงมาก หากในขั้นตอนแรกๆ วิเคราะห์ไม่ดีพอ ดังตัวอย่างรูปที่ 5.2



รูปที่ 5.2 โมเดลน้ำตก (Waterfall Model) (Laudon & Laudon, 2013)

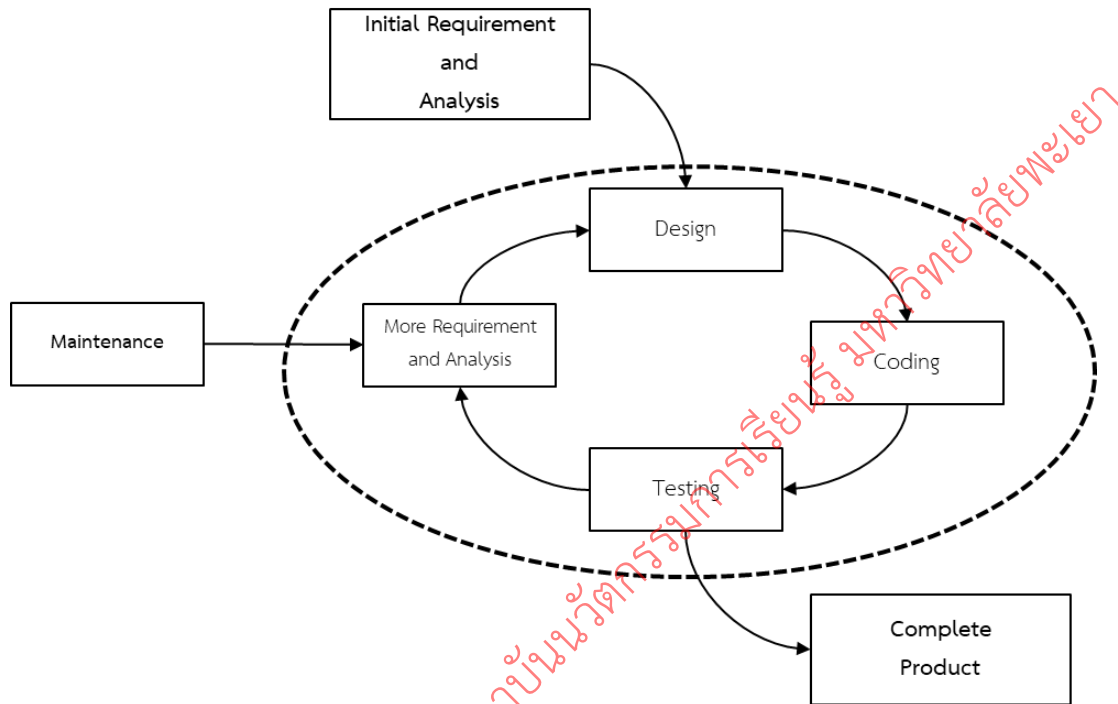
โดยมีรายละเอียดดังนี้

- ขั้นตอนที่ 1 Requirement เป็นขั้นตอนในการเก็บความต้องการของผู้ใช้ระบบสารสนเทศ แล้วว่า ผู้ใช้ระบบสารสนเทศ แล้วต้องการของระบบสารสนเทศ แล้วอะไร ทำงานอย่างไร
- ขั้นตอนที่ 2 Analysis and design เป็นขั้นตอนในการเอาความต้องการของผู้ใช้มาวิเคราะห์ถึงความเป็นไปได้และความเหมาะสม เลือกเครื่องมือที่จะทำการพัฒนาระบบสารสนเทศ แล้วทำการออกแบบ (Design) ระบบสารสนเทศ แล้ว
- ขั้นตอนที่ 3 Coding เป็นขั้นตอนของการเขียนโปรแกรมและระบบสารสนเทศตามที่เรากออกแบบมา
- ขั้นตอนที่ 4 Validating หรือ Testing เป็นขั้นตอนในการเอาโปรแกรมที่เขียนเสร็จแล้วมาทำการทดสอบหาข้อผิดพลาด
- ขั้นตอนที่ 5 Deployment and Maintenance เมื่อทำการทดสอบจนแน่ใจแล้วว่าไม่มีข้อผิดพลาดก็นำโปรแกรมและระบบสารสนเทศ แล้วไปใช้งานจริง รวมถึงเป็นขั้นตอนในการดูแลผู้ใช้ระบบสารสนเทศว่ามีปัญหาในการใช้งานหรือไม่อย่างไร พร้อมให้คำปรึกษาและรับฟังข้อคิดเห็นเพื่อจะนำมาพัฒนาระบบสารสนเทศในรุ่นถัดไป

ข้อเสียของ โมเดลน้ำตก (Waterfall Model) คือ แต่ละขั้นตอนต้องรอให้การออกแบบทั้งหมดเสร็จสมบูรณ์ก่อน จึงจะเริ่มเขียนโปรแกรมหรือพัฒนาระบบได้ ต้องรอให้ระบบงานทั้งหมดเสร็จก่อน

จึงจะใช้งานได้ ไม่สามารถนำระบบบางส่วนมาใช้งานก่อน รวมทั้งใช้เวลามากในการจัดทำเอกสารในแต่ละขั้นตอน

5.1.2 แบบ Iterative Model & Incremental Model เป็นโมเดลที่มีแนวคิดแบ่งงานออกเป็นขั้นเล็กๆ พัฒนางานขึ้นนั้นให้สมบูรณ์ที่สุดใกล้เคียงกับระบบที่สามารถใช้งานได้จริง



รูปที่ 5.3 แบบ Iterative Model & Incremental Model (Laudon & Laudon, 2013)

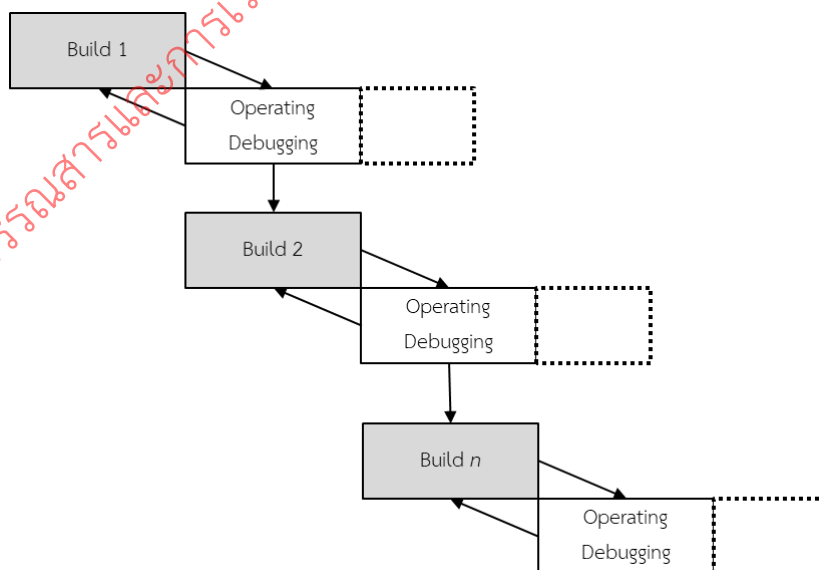
จากรูป 5.3 จะเห็นได้ว่าโมเดลแบบ Iterative Model & Incremental จะแบ่งงานออกเป็น Phase แล้วในแต่ละ Phase งานก็จะถูกแบ่งเป็นงานย่อยๆ เหมือน Waterfall Model เมื่อจบในแต่ละขั้นตอน (Phase) ก็จะทำให้ลูกค้าได้ทบทวน (Review) ระบบงานดังนั้นระบบงานนี้จะทำให้ลดความเสี่ยงในการทำโปรเจกต์ผิดพลาด ไม่ตรงตามความต้องการ และในขั้นตอน (Phase) ต่อไปก็ก็จะมีคามสำคัญ (Priority) ลดลงตามลำดับ โมเดลนี้มีลักษณะเด่น คือ การเก็บข้อมูลความต้องการของผู้ใช้งานในตัวโปรแกรมครั้งแรก แล้วนำ ข้อมูลนั้นมาวิเคราะห์เพื่อพัฒนาต่อ เหมาะสำหรับ โครงการที่มีความเสี่ยงปานกลางถึงต่ำ โครงการที่มีการกำหนด ระยะเวลาการทำงานที่ยาวนาน หรือโครงการที่ใช้เทคโนโลยีใหม่ๆ เข้าช่วยในการพัฒนาระบบสารสนเทศ

- Phase ที่ 1 Initial Requirement and Analysis รวบรวม ความต้องการของผู้ใช้โปรแกรมแล้วนำมาวิเคราะห์ พร้อมทั้ง หาเครื่องมือที่ใช้ในการพัฒนาต่อ

- Phase ที่ 2 Design ทำการออกแบบโปรแกรมหรือระบบตามที่ได้วิเคราะห์ต่อจากขั้นตอนแรก
- Phase ที่ 3 Testing เป็นขั้นตอนทำการหาข้อผิดพลาดของโปรแกรม และตรวจสอบทั้งในด้านผู้พัฒนาและผู้ใช้งาน
- Phase ที่ 4 More Requirement and Analysis หากทำการทดสอบแล้วได้ผลที่ไม่น่าพอใจ จะนำผลการวิเคราะห์นั้น มาทำการทบทวนและออกแบบระบบใหม่
- Phase ที่ 5 Product ถ้าผลการทดสอบเป็นที่น่าพอใจ จะระบบไปติดตั้งใช้งานจริง
- Phase ที่ 6 Maintenance เป็นขั้นตอนการพัฒนาโปรแกรมหรือระบบที่เคยใช้งานมาแล้ว หากพบปัญหาจะนำกลับเข้าสู่ขั้นตอนการพัฒนาปรับปรุงอีกครั้ง โดยส่งต่อเข้าสู่กระบวนการ More Requirement and Analysis

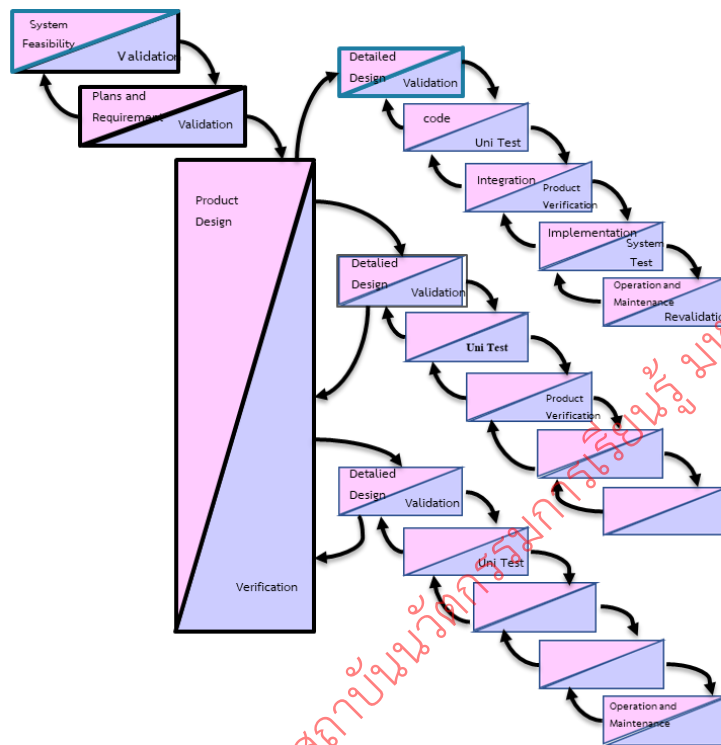
โดยสรุปโมเดลนี้ มีความยืดหยุ่น เพราะเริ่มต้นจากโปรแกรมหรือระบบเล็กๆ ที่สามารถทำได้ก่อน แล้วขยายให้ใหญ่ขึ้น แต่ข้อเสียคือ ยากในการบริหารโครงการโดยรวม ทั้งแง่ค่าใช้จ่ายและระยะเวลาที่ตั้งไว้เนืองด้วย หากทำการทดสอบแล้วได้ผลที่ไม่น่าพอใจ จะนำผลการวิเคราะห์นั้น มาทำการทบทวนและออกแบบระบบใหม่

5.1.3 แบบ Evolutionary Model แนวคิดของ Evolutionary Model เป็นแบบจำลองที่มีการทำกิจกรรมในลักษณะวนซ้ำ โดยสร้างส่วนของระบบงานจากขั้นเล็กๆน้อยๆขยายวงงานไปจนงานเสร็จสมบูรณ์คือ มีการปรับปรุงระบบอย่างต่อเนื่อง ดังรูป 5.4



รูปที่ 5.4 แบบ Evolutionary (Laudon & Laudon, 2013)

จุดสำคัญของโมเดลนี้คือ มีการพัฒนาระบบและตรวจสอบงานไปด้วย (Debugging) เพื่อให้แน่ใจว่าระบบงานที่ได้พัฒนาขึ้นมีความสมบูรณ์และมีคุณภาพเพียงพอที่จะดำเนินงานในขั้นตอนต่อไป ตัวอย่างโมเดลในรูปแบบ 5.5 คือ Incremental Model เป็นโมเดลที่นำ Water Fall Model มาปรับปรุงให้มีประสิทธิภาพยิ่งขึ้น



รูปที่ 5.5 โมเดลแบบ Incremental Model (Laudon & Laudon, 2013)

5.1.4 โมเดลแบบจำลอง Agile แนวคิดของ Agile นั้นหลักๆ มีดังนี้ ไม่เน้นกระบวนการและเอกสาร ยอมรับความเปลี่ยนแปลง ทำงานทีละนิดแต่ทำอย่างสม่ำเสมอ ไม่กลัวที่จะผิดพลาด นำข้อผิดพลาดมาแก้ไข และสุดท้ายคือเน้นการทำงานเป็นทีม โดยแนวคิดแบบ Agile Model ได้ถูกพัฒนาจากกระบวนการยุคแรกๆ ที่ใช้ชื่อโมเดลว่า Crystal FDD และ XP นอกจากนี้ยังมี โมเดลแบบ Hybrid Agile Approach ที่เรียกว่า DSDM ที่ใช้กันแพร่หลายทั่วโลก ยกเว้น ประเทศสหรัฐอเมริกา โดยสรุป แนวคิดที่สำคัญของ โมเดลแบบ Agile (Zhang & Patel, 2010)

- การพัฒนาระบบมีการร่วมมือกับลูกค้าในการเจรจาสัญญา และตอบสนองต่อการเปลี่ยนแปลงตามแผน ยอมรับความเปลี่ยนแปลง เพราะความต้องการของระบบ เช่นฟังก์ชันต่างๆ อาจเปลี่ยนแปลงได้ตลอด

- การพัฒนาระบบไม่เน้นกระบวนการและเอกสาร โดยเน้นไปที่การพัฒนาโปรแกรมให้ดีที่สุด ตรงตามความต้องการลูกค้า โดยพัฒนาโปรแกรมทีละส่วน และทำต่อเนื่องทีละเล็กทีละน้อย สามารถส่งมอบระบบทีละส่วนได้ ทำให้มองเห็นจุดที่ต้องแก้ไขได้เร็ว

- Agile Model มุ่งเน้นการทำให้ขั้นตอนการพัฒนาและการทดสอบคล่องตัว หากพบข้อผิดพลาดจะแก้ไขให้เร็ว เพื่อให้ระบบสามารถนำไปใช้งานได้ (Production)

- โมเดลแบบจำลอง Agile มุ่งเน้นการสื่อสารที่ชัดเจนระหว่างสมาชิกในทีมและลูกค้า ทำให้ระบบมีภาพที่ชัดเจนในทุกระยะของการพัฒนา เน้นการมีปฏิสัมพันธ์ระหว่างบุคคลมากกว่าตามกระบวนการ เช่น การประชุมสามารถนำ ทีมงานไปประชุมกับลูกค้า เพื่อจะได้รับข้อมูลที่ชัดเจน โดยไม่ต้องรอการสื่อสารหรือเอกสารจากนักวิเคราะห์ระบบ (System Analysis)

โดยโมเดลที่นิยมมากในการพัฒนาระบบที่ผ่านมาสองโมเดลคือ โมเดลน้ำตก (Water Fall) ซึ่งเป็นวิธีการแบบดั้งเดิม (Traditional Approach) มีข้อแตกต่างกันคือ

วิธีการแบบ Agile มุ่งเน้นพัฒนาแอปพลิเคชันหรือระบบอย่างรวดเร็ว และมักใช้งานร่วมกับการทำงานกับ Scrum Framework ซึ่งแบบจำลองโมเดลน้ำตก (Water Fall) เป็นกระบวนการพัฒนาระบบแบบต่อเนื่องกัน ซึ่งเห็นความคืบหน้าของงานว่าไหลลงไปเรื่อยๆ แบบ น้ำตก ผ่านขั้นตอนต่างๆ โดยยึดเอกสารเป็นหลัก นอกจากนี้โมเดลแบบจำลอง Agile เป็นแนวทางการพัฒนาซ้ำๆ แบบทีมมีการปรับปรุงจากข้อผิดพลาดตลอดเวลาผ่านการสื่อสาร แทนที่จะพัฒนาระบบจากเอกสาร และกำหนดการพัฒนาระบบจะถูกกำหนดด้วย "กล่องเวลา" "Time-Boxed" และแบ่งออกเป็นขั้นตอนที่เรียกว่า "การวิ่ง" "Sprints" โดยการวิ่งแต่ละครั้งมีระยะเวลาที่กำหนดไว้สั้นๆ (โดยปกติเป็นสัปดาห์) โดยมีการระบุรายการสิ่งที่ส่งมอบและต้องดำเนินการเป็นรายสัปดาห์ หากงานที่วางแผนไว้ไม่สามารถทำได้ เสร็จสมบูรณ์ งานจะถูกจัดลำดับความสำคัญใหม่ และต้องวางแผนการวิ่งต่อไป (แผนการพัฒนาระบบเปลี่ยนได้ตลอดตามการหารือกับลูกค้า) (McCormick, 2012)

โดยแนวคิดแบบ 'Agile' จะถูกกล่าวถึงการทำงานร่วมกับกรอบการทำงาน (Framework) แบบ "Scrum" โดยตามกรอบการทำงานของ Scrum จะประกอบด้วยทีมงาน 3 ตำแหน่ง คือ Product owner, Scrum master และ Development team โดยทำงานกันแบบ "Cross functional" สามารถทำงานข้ามสายงานบนความร่วมมือของทีมงาน โดย

- Product Owner ทำหน้าที่เป็นตัวแทนของทีม โดยบุคคลนี้จะต้อง มีความเข้าใจในธุรกิจ ลูกค้า และทำหน้าที่ทำให้เกิดการสื่อสารภายในทีม ทำให้แต่ละงานที่อยู่ใน Product Backlog มีความชัดเจน ประเมินผลและจัดลำดับความสำคัญ (Priorities and Tasks) ของงานต่างๆ ให้กับทีมด้วย

- Scrum Master ทำหน้าที่ ช่วยแก้ปัญหาต่างๆ ที่เกิดขึ้นในระหว่างการพัฒนา ระบบ โดยช่วยกำจัดอุปสรรคที่ขัดขวางไม่ให้ทีมบรรลุเป้าหมาย

- Development team คือบุคคลที่ทำหน้าที่ Designer, Programmer, UI/UX, Testing โดยทีมหนึ่งจะประกอบไปด้วยทีมงานประมาณ 2 ขึ้นไป ขึ้นอยู่กับขนาดของโครงการ โดยแต่ละคน แต่ละตำแหน่งจะทำงานแบบ Self-Management คือบริหารจัดการด้วยตนเอง รับผิดชอบงานตนเองให้ตรงต่อเวลา และส่งมอบ Tasks and Requirement ที่ได้รับมอบหมายจากการประชุมร่วมกันแต่ละสัปดาห์

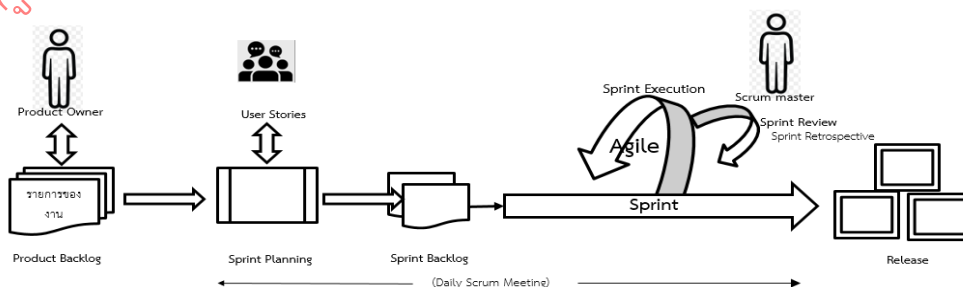
โดยวิธีการทำงานของ “Scrum” จะมีองค์ประกอบที่สำคัญ (Scrum Artifacts) คือ

- Product Backlog หมายถึงรายการของงานที่ต้องทำหรือโปรแกรมที่ต้องพัฒนาไปจนถึงฟังก์ชัน (Functions and Features) รวมถึงข้อกำหนดทางธุรกิจและอื่นๆ โดย Product Owner จะเป็นผู้จัดทำรายการนี้ขึ้นมาและจัดลำดับความสำคัญของงาน (Priorities and Tasks) และเป็นคนตัดสินใจนำรายการ (Tasks) ต่างๆ เหล่านี้เข้าไปใน Sprint โดยคำนึงถึงทรัพยากรที่ต้องใช้ (Effort) รวมถึงประเมินผล (Evaluation) ของงานต่างๆ ให้กับทีม

- Sprint Backlog คือรายการงานที่เลือกมาจากรายการของงานที่ต้องทำ (Product Backlog) โดย Development team จะรวบรวมข้อมูลต่างๆ จากการประชุมทุกๆ เข้า เพื่อใช้ในการทำงาน เช่น User story, Bug fix, สิ่งที่ต้องทำ (To-Do), สิ่งที่กำลังลงมือทำอยู่ (Doing), สิ่งที่ทำเสร็จไปแล้ว (Done) เป็นต้น รายการงาน (Sprint Backlog) จะต้องสามารถเข้าถึงได้ง่ายโดยทุกคนในทีมและสามารถตรวจสอบได้ตลอดเวลา

- Increment (หรือ Sprint Goal) คือเป้าหมายที่ถูกกำหนดขึ้นระหว่างที่ทำการกิจกรรม Sprint Planning โดยปกติจะเป็นรายสัปดาห์ Sprint Goal จะต้องชัดเจนว่าเป้าหมายคืออะไร ใครคือกลุ่มเป้าหมาย และต้องเป็นแผน (กลยุทธ์แบบไหน) เพื่อที่สมาชิกในทีมจะรับทราบและติดตามร่วมกันได้ ให้ความเห็น สื่อสารได้ตลอดเวลา

กระบวนการทำงานของ “Scrum” แสดงโดยรูป 5.6



รูปที่ 5.6 กระบวนการทำงานของ “Scrum” (ผู้เขียน)

จากรูป 5.6 จะมีขั้นตอนการทำงานดังนี้

ขั้นที่ 1 ขั้นตอนการทำ Backlog ขั้นตอนนี้คือการจัดทำรายการหรือลิสต์ของงานที่ต้องทำ รวมถึงข้อกำหนดทางธุรกิจต่างๆ ซึ่ง Product Owner จะเป็นคนจัดทำเพื่อนำเข้าสู่กระบวนการ Sprint Planning

ขั้นที่ 2 Sprint Planning เป็นขั้นตอนการวางแผนงานที่จะต้องดำเนินการใน Sprint การวางแผนนี้ทำขึ้นโดยความร่วมมือของสมาชิกทีมงาน Scrum ทั้งหมด และจะถูกจัดทำเป็น Sprint Backlog มีการกำหนดเป้าหมาย กลยุทธ์ เพื่อให้งานสำเร็จตามแผนงานแต่ละสัปดาห์

ขั้นที่ 3 Daily Scrum Meeting เป็นขั้นตอนการประชุมสั้นๆ ภายในทีมประมาณ 10-15 นาที ตลอดระยะเวลาโครงการที่ทำ Sprint เพื่อปรึกษาทบทวนการทำงาน สิ่งที่ทำไปแล้ว สิ่งที่จะทำต่อไป รวมถึงปัญหาต่างๆ และมีการแก้ไขปัญหาอย่างต่อเนื่อง

ขั้นที่ 4 Sprint Review ขั้นตอนนี้เป็นขั้นตอนเพื่อตรวจสอบ Increment (ผลรวมของงาน) และรับคำแนะนำหรือความต้องการใหม่ๆ (Feedback and Requirement) จากผู้เกี่ยวข้องทั้งหมด (Stakeholder) ในการพัฒนาระบบและลูกค้าสำหรับนำไปปรับปรุงการพัฒนาโปรแกรมหรือแอปพลิเคชันต่อไป

ขั้นที่ 5 Sprint Retrospective ขั้นตอนนี้เป็นขั้นตอนสุดท้าย โดยสรุปสิ่งที่ทำมาทั้งหมด และเป็นการประชุมทีมงานให้ร่วมกันตรวจสอบการทำงานก่อนจะเริ่มต้นทำ Sprint Planning ครั้งถัดไป

โดยข้อมูลเอกสารทั้งหมดทุกกระบวนการจะต้องสามารถเข้าถึงได้ง่ายโดยสมาชิกทุกคนในทีม และตรวจสอบได้ตลอดเวลา ไม่ว่าจะทำงานแบบออนไลน์หรือออฟไลน์ หรือแบบ Hybrid Meeting

สรุปประโยชน์ของการประยุกต์ใช้แบบจำลอง Agile และกรอบการทำงาน Scrum ทำงานร่วมกัน

การประยุกต์ใช้แบบจำลอง แบบ Agile ซึ่งเป็นกระบวนการคิดที่มีความยืดหยุ่น ประกอบกับการทำงานรูปแบบ Scrum ซึ่งจะจัดประชุมได้ทุกวัน มี Daily Meeting หรือวันละ 1-2 รอบก็ตาม ประมาณ 10-15 นาทีตลอดระยะเวลาโครงการ แบบออนไลน์หรือออฟไลน์ก็ได้ (จากเดิมที่อาจจะประชุมกันทั้งวัน) เพื่อประเมินหาข้อผิดพลาดแบบ วันต่อวันรวม การประชุมมีการทบทวนภาพใหญ่คู่ไปด้วย ดังนั้น ถ้าเกิดปัญหาอย่างใดอย่างหนึ่งระหว่างกระบวนการ ทีมจะสามารถแก้ไขได้ทันที โดยสรุปการพัฒนาระบบสารสนเทศสมัยใหม่ที่มีสิ่งเข้ามากระทบต่อการบริหารโครงการอย่างคาดไม่ถึง ทั้งโรคระบาด ภัยพิบัติ ต่างๆ ทำให้การทำงานในรูปแบบปรกติ ไม่สามารถดำเนินงานได้ จำเป็นต้องมี

กระบวนการทำงานและกรอบการทำงานที่ดีอย่าง **แบบจำลอง Agile และกรอบการทำงาน Scrum** ซึ่งมีข้อดีดังนี้

- แบบจำลอง แบบ Agile ให้แนวทางการทำงานด้วยความรวดเร็ว จากการแบ่งโครงการออกเป็นงานหรือขั้นตอนย่อยๆ ช่วยให้จัดการและวางแผนได้ง่ายขึ้น
- กรอบการทำงาน Scrum มีความยืดหยุ่น สามารถปรับเปลี่ยนแผนงานได้ตามความต้องการ (Requirement) เงื่อนไขสถานการณ์ หรือข้อเสนอแนะ (Feedback) จาก Stakeholder และ ลูกค้า
- ระบบหรือโปรแกรมจะถูกพัฒนาและทดสอบในขั้นตอน Sprint Review ทำให้ทีมงานสามารถค้นพบและแก้ไขข้อผิดพลาดได้ทันที ไม่ต้องแก้ไขในขั้นตอนสุดท้ายทีเดียว
- การสื่อสารอย่างสม่ำเสมอในขั้นตอน Daily Scrum Meeting ซึ่งสามารถกระตุ้นให้เกิดการเรียนรู้ได้มากขึ้น ในการทำงานของทีม ทำให้เกิดการเรียนรู้ร่วมกัน (Knowledge Management)

5.3 มาตรฐาน COBIT: Control Objectives for Information and Related Technology (Amorim et al., 2021)

COBIT ย่อมาจาก Control Objectives for Information and Related Technology เป็นกรอบที่สร้างขึ้นโดย ISACA สำหรับการกำกับดูแลและการจัดการด้านสารสนเทศ ใช้เป็นเครื่องมือในการบริหารจัดการเพื่อสร้างความมั่นใจว่าการใช้ทรัพยากรด้านเทคโนโลยีสารสนเทศ ให้สอดคล้องกับวัตถุประสงค์เชิงธุรกิจขององค์กร และช่วยให้การพัฒนาและวางระบบสารสนเทศสามารถเชื่อมระหว่างการจัดการด้านเทคโนโลยี การจัดการความเสี่ยงทางธุรกิจ และเป้าหมายและข้อกำหนดในการดำเนินธุรกิจได้อย่างสมดุล เกิดประสิทธิภาพสูงสุด

โดย COBIT 5 เป็นเฟรมเวิร์ค (Framework) ที่ใช้กันอย่างแพร่หลายสำหรับการควบคุมบริหารโครงการสารสนเทศเป็นหนึ่งในเฟรมเวิร์ค (Framework) ของ GEIT โดย GEIT เป็นกรอบการทำงานที่ทำให้แน่ใจว่าทรัพยากรหน่วยงานด้านสารสนเทศขององค์กรถูกใช้อย่างมีประสิทธิภาพเพื่อตอบสนองความต้องการของผู้มีส่วนได้ส่วนเสียโดย มุ่งเน้นใน 5 ด้าน คือ

- การจัดวางกลยุทธ์ (Strategic Alignment)
- การนำเสนอคุณค่า (Value Delivery)
- การจัดการทรัพยากร (Resource Management)

- การจัดการความเสี่ยง (Risk Management)
- การวัดผลการดำเนินงาน (Performance Measurement)

และมีขอบเขตของกระบวนการ 4 ด้านหลักด้วยกัน คือ

- การวางแผนและการจัดองค์กร Planning and Organization : PO
- การจัดหาและนำไปใช้ Acquisition and Implementation : AI
- การนำส่งและสนับสนุน Delivery and Support : DS
- การตรวจสอบและประเมินผล Monitoring and Evaluation : ME

โดยที่แต่ละด้านนั้นจะมีกระบวนการย่อยอยู่ภายใน ซึ่งจะมีกระบวนการรวมทั้งหมด 34 กระบวนการ และแต่ละกระบวนการจะมีวัตถุประสงค์ย่อย ทั้งนี้ขึ้นอยู่กับ การวางแผนร่วมกันของทีมงานผู้มีส่วนได้ส่วนเสียในการวางแผนพัฒนาระบบสารสนเทศดังตัวอย่างตามตาราง ที่ 5.1

ตารางที่ 5.1 มาตรฐาน COBIT และวัตถุประสงค์ย่อย (DPO Enterprise Architecture 4.0, 2020)

กลุ่ม	วัตถุประสงค์ควบคุม	ตัวอย่างวัตถุประสงค์ย่อย
1. การวางแผนและการจัดการองค์กร (PO - Planning and Organization)	1.1 การจัดทำแผนกลยุทธ์ด้านเทคโนโลยีสารสนเทศ	1) แผนงานระยะยาวด้านเทคโนโลยีสารสนเทศ 2) แผนงานระยะสั้นสำหรับหน่วยงานด้านเทคโนโลยีสารสนเทศ
	1.2 การกำหนดสถาปัตยกรรม ด้านสารสนเทศ	1) ต้นแบบสถาปัตยกรรม 2) แบบแผนการจัดประเภทของข้อมูล 3) ระดับการรักษาความปลอดภัย
	1.3 การกำหนดทิศทางด้าน เทคโนโลยี	1) การวางแผนโครงสร้างพื้นฐานด้านเทคโนโลยี 2) มาตรการด้านเทคโนโลยี

	1.4 การกำหนดโครงสร้างการ จัด องค์กรด้านเทคโนโลยี สารสนเทศ และความสัมพันธ์กับ หน่วยงานอื่น	1) นโยบายและขั้นตอนการความ รับผิดชอบด้านเทคโนโลยี สารสนเทศ 2) ความเป็นเจ้าของข้อมูลและ ระบบ
	1.5 การจัดการด้านการลงทุนใน เทคโนโลยีสารสนเทศ	1) งบประมาณของการดำเนินงาน ด้าน เทคโนโลยีสารสนเทศ 2) การติดตามดูแลความเหมาะสม ของค่าใช้จ่ายและประโยชน์ที่ได้รับ
	1.6 การสื่อสารเป้าหมายและทิศ ทางการจัดการ	1) นโยบายด้านกรอบงานการรักษา ความปลอดภัยและ การควบคุมภายใน 2) การปฏิบัติตามนโยบาย ขั้นตอน การปฏิบัติงาน และมาตรฐาน
	1.7 การจัดการทรัพยากรบุคคล	1) บทบาทและความรับผิดชอบ บุคลากร 2) การฝึกอบรมบุคลากร
	1.8 การปฏิบัติตามข้อกำหนด ขององค์กรภายนอก	1) วิธีการปฏิบัติและระเบียบปฏิบัติ เพื่อให้เป็นไปตามข้อกำหนดของ องค์กร 2) ความเป็นส่วนตัว ทรัพย์สินทาง ปัญญาและกระแสข้อมูล
	1.9 การประเมินความเสี่ยง	1) แผนปฏิบัติการเพื่อจัดการความ เสี่ยง 2) การประเมินความเสี่ยงทางธุรกิจ
	1.10 การจัดการโครงการ	1) แผนงานหลักของโครงการ 2) การบริหารความเสี่ยงของ โครงการอย่างเป็นทางการ

	1.11 การจัดการคุณภาพ	1) การวางแผนการรับรองคุณภาพ 2) กรรมวิธีวงจรการพัฒนา ระบบงาน 3) มาตรฐานการสร้างเอกสาร โปรแกรมและมาตรฐานการทดสอบ โปรแกรม
2. การจัดหาและการนำระบบออก ใช้งานจริง (AI – Acquisition And Implementation)	2.1 การเลือกเทคโนโลยีมาใช้ใน การปฏิบัติงาน	1) การกำหนดความต้องการด้าน สารสนเทศและการศึกษาความเป็น ไปได้ด้านเทคโนโลยี ด้าน เศรษฐกิจ 2) สถาปัตยกรรมสารสนเทศ 3) รายงานการวิเคราะห์ความเสี่ยง
	2.2 การจัดหาและบำรุงรักษา ซอฟต์แวร์ประยุกต์	1) วิธีการออกแบบระบบ 2) การกำหนดความต้องการของ แฟ้มข้อมูล วิธีการเก็บรวบรวม ข้อมูลต้นทาง การนำข้อมูลเข้าสู่ ระบบและการจัดทำเอกสาร 3) การกำหนดเกี่ยวกับการเชื่อมโยง
	2.3 การจัดหาและบำรุงรักษา โครงสร้างพื้นฐานด้านเทคโนโลยี	1) การประเมินฮาร์ดแวร์และ ซอฟต์แวร์ใหม่ และการบำรุงรักษา เชิงป้องกันของฮาร์ดแวร์ 2) การรักษาความปลอดภัยใน ซอฟต์แวร์ระบบ 3) การควบคุมการแก้ไข เปลี่ยนแปลงซอฟต์แวร์ระบบ
	2.4 ระเบียบปฏิบัติในการพัฒนา และบำรุงรักษา	1) คู่มือปฏิบัติงานของผู้ใช้ 2) คู่มือปฏิบัติงานด้านปฏิบัติการ คอมพิวเตอร์

		3) เอกสารประกอบการฝึกอบรม
	2.5 การติดตั้งและรับรองระบบ	1) แผนการนำระบบออกใช้งานจริง 2) กลยุทธ์และแผนการทดสอบระบบ 3) การทดสอบและรับรองความปลอดภัยของระบบ
	2.6 การจัดการการเปลี่ยนแปลง	1) การควบคุมการเปลี่ยนแปลง 2) การจัดทำเอกสารและระเบียบปฏิบัติ 3) นโยบายเกี่ยวกับการนำออกใช้งาน (Release) ของซอฟต์แวร์
3. การส่งมอบและการสนับสนุน (DS – Delivery and Support)	3.1 การกำหนดและการจัดการระดับการให้บริการ	1) หลักเกณฑ์ และระเบียบข้อตกลงเรื่องระดับการให้บริการและการทำงาน 2) การสอบทานข้อตกลงและสัญญาเรื่องระดับการให้บริการ
	3.2 การจัดการการใช้บริการจากบุคคลภายนอก	1) การประสานงานกับผู้ให้บริการ 2) ข้อสัญญาการใช้บริการจากบุคคลภายนอก 3) การตกลงร่วมมือในการรักษาความปลอดภัย
	3.3 การจัดการด้านสมรรถนะและความสามารถของระบบ	1) แผนด้านความพร้อมใช้งานของระบบ

		2) การจัดการเชิงรุกด้านสมรรถนะของระบบ
	3.4 ความต่อเนื่องในการให้บริการ	1) กรอบงานของการดำเนินงานอย่างต่อเนื่องด้านเทคโนโลยีสารสนเทศ 2) การทดสอบแผนการดำรงอยู่ด้านเทคโนโลยีสารสนเทศ
	3.5 การรักษาความปลอดภัยระบบ	1) การจัดการด้านมาตรการการรักษาความปลอดภัย 2) สถาปัตยกรรม firewall และการเชื่อมโยงกับเครือข่ายสาธารณะ
	3.6 การกำหนดและการจัดสรรต้นทุน	1) ระเบียบปฏิบัติในเรื่องต้นทุน 2) การเรียกเก็บค่าบริการและระเบียบปฏิบัติในการคิดค่าบริการกับผู้ใช้งาน
	3.7 การให้ความรู้และการฝึกอบรมแก่ผู้ใช้งาน	1) การบริหารจัดการการฝึกอบรม 2) การฝึกอบรมเรื่องหลักการและความตระหนักถึงการรักษาความปลอดภัย
	3.8 การช่วยเหลือและให้คำปรึกษาแก่ลูกค้า	1) การจัดตั้งศูนย์ช่วยเหลือผู้ใช้งาน 2) ขั้นตอนการแก้ไขปัญหาข้อซักถามของผู้ใช้งาน 3) การวิเคราะห์แนวโน้มของปัญหาและการรายงานผล
	3.9 การจัดการรายละเอียดทรัพย์สิน	1) การบันทึกรายการรายละเอียดทรัพย์สิน 2) การควบคุมรายละเอียดทรัพย์สิน

	3.10 การจัดการปัญหาและ เหตุการณ์ที่เกิดขึ้น	1) ระบบการจัดการปัญหา 2) หลักฐานการตรวจสอบและการ ติดตามปัญหา 3) การอนุมัติให้เข้าถึงระบบในกรณี ฉุกเฉินและชั่วคราว
	3.11 การจัดการข้อมูล	1) ระเบียบปฏิบัติในการเตรียม ข้อมูล 2) ระเบียบปฏิบัติว่าด้วยสิทธิการ บันทึกข้อมูล
	3.12 การจัดการกับสิ่งอำนวยความสะดวก	1) การรักษาความปลอดภัยทาง กายภาพ 2) สถานที่ตั้งของศูนย์คอมพิวเตอร์ที่ ไม่เป็นที่สังเกต 3) การป้องกันภัยจากปัจจัยทาง สภาพแวดล้อม
	3.13 การจัดการการปฏิบัติการ	1) ระเบียบปฏิบัติและคู่มือคำสั่งการ ประมวลผล 2) การรักษาความปลอดภัยของ เอกสาร แบบฟอร์มพิเศษ และ อุปกรณ์ประมวลผลลัพท์
4. การตรวจสอบและประเมินผล Monitoring and Evaluation : ME	4.1 การเฝ้าติดตามกระบวนการ	1) การประเมินผลการดำเนินงาน 2) การประเมินความพึงพอใจของ ลูกค้าและการรายงานสำหรับผู้บริหาร
	4.2 การประเมินความเพียงพอ ของการควบคุมภายใน	1 การดำเนินการของการควบคุม ภายใน 2) การรับรองความปลอดภัยในการ ดำเนินงานและการ

		ควบคุมภายใน
	4.3 การรับรองอย่างเป็นอิสระ	1) การออกใบรับรอง/การรับรอง อย่างเป็นอิสระในด้านการรักษา ความปลอดภัยและการควบคุม ภายในของการให้บริการด้าน เทคโนโลยีสารสนเทศ ด้าน การรักษาความปลอดภัยและการ ควบคุมภายในของการให้บริการ จากบุคคลภายนอก 2) การประเมินประสิทธิภาพอย่าง เป็นอิสระของบริการ ด้านเทคโนโลยีสารสนเทศ 3) การรับรองอย่างเป็นอิสระต่อการ ปฏิบัติตาม กฎหมาย ข้อบังคับ และ ข้อผูกมัดในสัญญา และการรับรอง อย่างเป็นอิสระต่อการปฏิบัติตาม กฎหมาย ข้อบังคับ และข้อผูกมัดใน สัญญากับผู้ให้บริการภายนอก
	4.4 การจัดให้มีผู้ตรวจสอบ อิสระ	1) กฎบัตรการตรวจสอบ 2) การวางแผน 3) การดำเนินงานตรวจสอบ 4) การรายงาน 5) การติดตามผล

แนวทางการนำ COBIT5 ไปใช้งานควบคู่กับโมเดลการพัฒนาซอฟต์แวร์นั้น ผู้นำไปใช้ควร
จะต้องทำความเข้าใจก่อนว่า COBIT5 คือกรอบหรือเฟรมเวิร์คใหญ่ที่ใช้เป็นแนวทางปฏิบัติ ส่วนโมเดล
การพัฒนานั้นเป็นองค์ประกอบย่อยในเฟรมเวิร์ค COBIT5

แนวทางการนำ COBIT5 ไปใช้งานควบคู่กับโมเดลการพัฒนาซอฟต์แวร์นั้น ผู้นำไปใช้ควรจะต้องทำความเข้าใจก่อนว่า COBIT5 คือกรอบหรือเฟรมเวิร์คใหญ่ที่ใช้เป็นแนวทางปฏิบัติ ส่งโมเดลพัฒนานั้นเป็นองค์ประกอบย่อยในเฟรมเวิร์ค COBIT5

โดยจากการวิจัยของผู้เขียนเรื่อง "Strategic Approach to Adopt ICT in SMEs." ได้สรุปปัจจัยหลักสู่ความสำเร็จที่จะช่วยให้การนำ COBIT5 ไปใช้งานประสบผลสำเร็จร่วมกับโมเดลการพัฒนาซอฟต์แวร์นั้น ประกอบด้วย

- 1) ผู้บริหารระดับสูงต้องให้คำมั่นสัญญาและการสนับสนุนอย่างต่อเนื่องอย่างชัดเจน
- 2) ผู้มีส่วนได้ส่วนเสียทุกหน่วยต้องมีความเข้าใจในวัตถุประสงค์ทั้งทางด้านธุรกิจและด้านสารสนเทศ
- 3) สร้างการสื่อสารที่มีประสิทธิภาพให้กับทีมงานทั้งหมด รวมทั้งผู้มีส่วนได้ส่วนเสียทุกหน่วย
- 4) ปรับใช้ COBIT มาตรฐานที่สนับสนุนให้เหมาะสมกับบริบทเฉพาะขององค์กร
- 5) กำหนดวัตถุประสงค์ของการนำไปใช้งาน รวมถึงระยะเวลาของโครงการและผลประโยชน์ที่ต้องการอย่างชัดเจน
- 6) สร้างความตระหนัก และความเห็นชอบร่วมกัน
- 7) จัดสรรทรัพยากรที่เหมาะสม เพื่อสนับสนุนโครงการ และควรกำหนดบทบาทหน้าที่และความรับผิดชอบที่ชัดเจน
- 8) จัดทำโครงสร้างและกระบวนการที่เหมาะสมในการควบคุมดูแลโครงการและกระบวนการ
- 9) ตรวจสอบอยู่เป็นประจำเกี่ยวกับปัญหาประสิทธิภาพด้านสารสนเทศหรือรายงานเกี่ยวกับปัญหาคุณภาพของการให้บริการ

สรุปเนื้อหาบทที่ 5 "โมเดลการพัฒนาและมาตรฐานสำหรับการสร้างระบบสารสนเทศสมัยใหม่" นำเสนอเนื้อหาที่เกี่ยวข้องกับแนวทางการพัฒนาและโมเดลที่ใช้ในการสร้างระบบสารสนเทศในยุคสมัยใหม่ เช่น โมเดลขั้นตอนการพัฒนาซอฟต์แวร์ (Software Development Process Models) รวมถึงการใช้อนุกรมมาตรฐานเพื่อให้การพัฒนาระบบเป็นไปอย่างมีประสิทธิภาพและสอดคล้องกับความต้องการขององค์กร

1. **แนวทางการพัฒนา** การสร้างระบบสารสนเทศสมัยใหม่เริ่มต้นด้วยการระบุแนวทางการพัฒนาที่เหมาะสม เช่น การใช้วิธีการพัฒนาแบบน้ำตก (Agile) หรือวิธีการพัฒนาแบบความร่วมมือ

(Collaborative Development) เพื่อให้การพัฒนาระบบเป็นไปอย่างมีประสิทธิภาพและสามารถปรับตัวกับการเปลี่ยนแปลงได้

2. **โมเดลในการพัฒนา** นำเสนอโมเดลการพัฒนาที่เหมาะสมสำหรับการสร้างระบบสารสนเทศ อาจเป็นโมเดลกระบวนการการพัฒนาซอฟต์แวร์ที่เหมาะสม เช่น โมเดลวงจรชีวิตการพัฒนาซอฟต์แวร์ (Software Development Life Cycle) หรือโมเดล Spiral ที่เน้นการทดลองและปรับปรุงเป็นรอบ
3. **มาตรฐานสำหรับการพัฒนา** เช่นมาตรฐาน COBIT การใช้มาตรฐานในการพัฒนาระบบสารสนเทศสามารถช่วยให้การพัฒนาระบบเป็นไปอย่างเป็นระบบและเข้าใกล้กับความต้องการขององค์กร มาตรฐานอาจเป็นเกี่ยวกับการออกแบบฐานข้อมูล เทคโนโลยีที่ใช้ หรือกระบวนการในการพัฒนา
4. **การทดสอบและการตรวจสอบมาตรฐาน** นำเสนอการใช้การทดสอบและการตรวจสอบเพื่อตรวจสอบว่าระบบสร้างตามมาตรฐานและรายละเอียดที่กำหนดไว้อย่างถูกต้องและเสถียร
5. **การประเมินและการปรับปรุง** เมื่อระบบสารสนเทศสร้างขึ้นแล้ว จะต้องมีการประเมินระบบเพื่อตรวจสอบว่ามีการปฏิบัติตามมาตรฐานและเป้าหมายหรือไม่ และมีการปรับปรุงเพื่อปรับปรุงประสิทธิภาพและความเหมาะสม
6. **การบริหารจัดการและการสนับสนุนระบบ** นำเสนอแนวทางในการบริหารจัดการและการดูแลระบบสารสนเทศหลังจากการพัฒนา เพื่อให้ระบบสารสนเทศสามารถทำงานได้อย่างเป็น ประสิทธิภาพและรองรับการเปลี่ยนแปลงในอนาคต

โดยสรุปบทที่ 5 เป็นการนำเสนอแนวทางในการพัฒนาและโมเดลและมาตรฐานที่เหมาะสมสำหรับการสร้างระบบสารสนเทศในยุคสมัยใหม่ รวมถึงการใช้มาตรฐานและการทำการประเมินและการบริหารจัดการในระยะยาว

คำศัพท์ บทที่ 5

Acquisition and Implementation	การจัดหาและการนำระบบออกใช้งานจริง
Business Objectives	วัตถุประสงค์เชิงธุรกิจ

Control Objectives for Information and Related Technology	มาตรฐาน COBIT
Data Communication	การสื่อสารข้อมูล
Data Processing	การประมวลผลข้อมูล
Debugging	ตรวจสอบงาน
Delivery and Support	การส่งมอบและการสนับสนุน
Digital Technology	เทคโนโลยีดิจิทัล
Evolutionary Model	แบบจำลองที่มีการทำกิจกรรมในลักษณะวนซ้ำ
Iterative Model & Incremental Model	โมเดลที่มีแนวคิดแบ่งงานออกเป็นขั้นเล็กๆ
Linear Model	แบบจำลองที่มีขั้นตอนเป็นลำดับ
Monitoring and Performance Evaluation	การตรวจสอบและประเมินผล
Performance Measurement	การวัดผลการดำเนินงาน
Planning and Organization	การวางแผนและการจัดการองค์กร
Priority	ลำดับความสำคัญ
Process	ขั้นตอน
Product Backlog	รายการของงานที่ต้องทำหรือโปรแกรมที่ต้องพัฒนา
Resource Management	การจัดการทรัพยากร
Risk Management	การจัดการความเสี่ยง
Software Development Process Models	โมเดลขั้นตอนการพัฒนาซอฟต์แวร์
Software Validation	การตรวจสอบความถูกต้องซอฟต์แวร์
Sprint Backlog	รายการงานที่เลือกมาจากรายการของงานที่ต้องทำ (Product Backlog)
Sprint Goal	เป้าหมายที่ถูกกำหนดขึ้นระหว่างที่ทำกิจกรรม Sprint Planning
Strategic Alignment	การจัดวางกลยุทธ์
System Analysis	นักวิเคราะห์ระบบ
Traditional Approach	วิธีการแบบดั้งเดิม
User	บุคลากรผู้ใช้งานระบบ
Value Delivery	การนำเสนอคุณค่า

คำถามและแบบฝึกหัดท้ายบทที่ 5

1. ระบบสารสนเทศจะสมบูรณ์แบบต้องประกอบไปด้วย 5 องค์ประกอบอะไรบ้าง
2. โมเดลขั้นตอนการพัฒนาซอฟต์แวร์ (Software) คือ
3. คุณสมบัติของระบบสารสนเทศที่ดี คือ

4. กรอบมาตรฐานของ COBIT ประกอบด้วย 5 กลุ่มใหญ่คือ
5. โมเดลน้ำตก (Waterfall Model) จัดอยู่ในโมเดลแบบไหน และจงอธิบาย
6. จงอธิบายโมเดลแบบ Iterative Model & Incremental
7. จงอธิบายจุดสำคัญของโมเดลแบบ Evolutionary Model
8. จงอธิบายแนวคิดที่สำคัญของ โมเดลแบบจำลอง Agile
9. จงอธิบายทีมงาน 3 ตำแหน่งคืออะไร และมีหน้าที่อะไรบ้าง
10. จงอธิบาย Daily Scrum Meeting คืออะไร
11. จงอธิบายมาตรฐาน COBIT คืออะไร

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

บทที่ 6 การประยุกต์ใช้สถาปัตยกรรมองค์กรเพื่อพัฒนาระบบข้อมูลขนาดใหญ่

(Applying the Enterprise Architecture for Big Data System Development)

การศึกษาและการทำความเข้าใจการประยุกต์ใช้สถาปัตยกรรมองค์กร รวมถึงทราบกระบวนการและวิธีการวางแผนเพื่อพัฒนาระบบข้อมูลขนาดใหญ่โดยใช้สถาปัตยกรรมองค์กร และเข้าใจแนวทางการวิเคราะห์ปัญหา รู้จักความต้องการขององค์กรที่มีการเปลี่ยนแปลง ทั้งในเชิงการดำเนินธุรกิจที่มีการแข่งขันรุนแรงมากขึ้น รวมถึงการเข้าใจกระบวนการงานและการจัดการที่ทันสมัย และบริบทการใช้งานระบบสารสนเทศที่เปลี่ยนแปลงไป มีความสำคัญต่อการบริหารโครงการสารสนเทศและการจัดทำแผนในยุคนี้

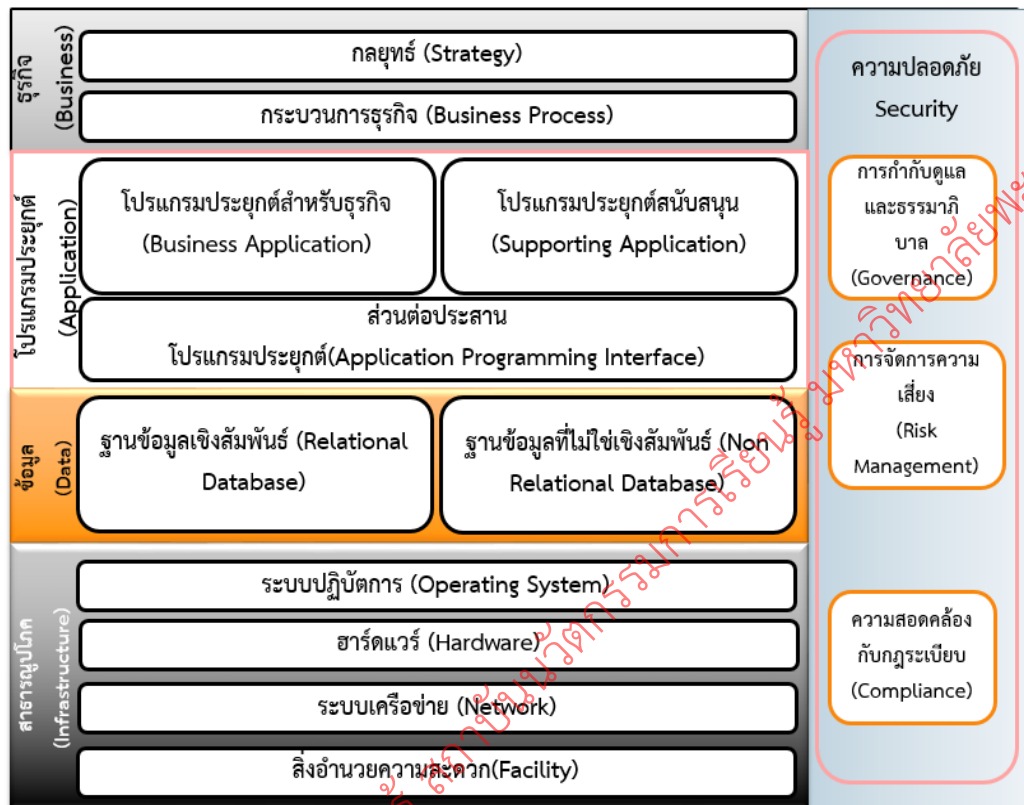
วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจองค์ประกอบที่สำคัญของสถาปัตยกรรมองค์กรต่อการพัฒนาระบบข้อมูลขนาดใหญ่
2. สามารถประยุกต์ใช้แนวคิดสำคัญของสถาปัตยกรรมองค์กรในการพัฒนาระบบข้อมูลขนาดใหญ่

6.1 สถาปัตยกรรมองค์กรเพื่อพัฒนาระบบข้อมูลขนาดใหญ่ (Enterprise Architecture for Big Data System) สถาปัตยกรรมองค์กรที่สมบูรณ์ (Enterprise Architecture Maturity) จะใช้เป็นแนวทางในการบูรณาการทั้งด้านการบริหารจัดการองค์กรสมัยใหม่ การออกแบบกระบวนการงานและเทคโนโลยีดิจิทัลของสถาปัตยกรรมองค์กรจะต้องสอดคล้องกับวิสัยทัศน์ วัตถุประสงค์เชิงธุรกิจและวัตถุประสงค์เชิงปฏิบัติการ และจะต้องสามารถสร้างระดับความสามารถในการก้าวสู่การเป็นองค์กรดิจิทัล (Digital Organization) อย่างยั่งยืน โดยมีองค์ประกอบที่สำคัญของ**สถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่** คือ

- 1) การกำกับดูแลการบริหารองค์กร (Enterprise Governance)
- 2) ส่วนองค์ประกอบของสถาปัตยกรรมและกรอบการบริหารจัดการเทคโนโลยีดิจิทัลประกอบไปด้วย (1) สถาปัตยกรรมด้านธุรกิจ (Business Architecture) (2) สถาปัตยกรรมด้านข้อมูล (Data Architecture) (3) สถาปัตยกรรมด้านโปรแกรมประยุกต์ (Application Architecture) (4) สถาปัตยกรรมด้านเทคโนโลยี (Technology Architecture) (5) สถาปัตยกรรมด้านการรักษาความปลอดภัยข้อมูล (Security Architecture) และ (6) การเฝ้าระวังและบริหารจัดการระบบ (System Monitoring and Management) ดังรูป 6.1 สถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่

(Enterprise Reference Model) โดยสถาปัตยกรรมองค์กร (Enterprise Architecture) จะต้องมีระบบฐานงาน (Platform) ระบบคลังข้อมูล (Data Centre) ระบบข้อมูลขนาดใหญ่ (Big Data) และระบบ Internet of Things (IoT) ดังรูป 6.2 สถาปัตยกรรมของข้อมูลขนาดใหญ่ (Architecture of Big Data)



รูป 6.1 แบบจำลองสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ (DGA, 2022)

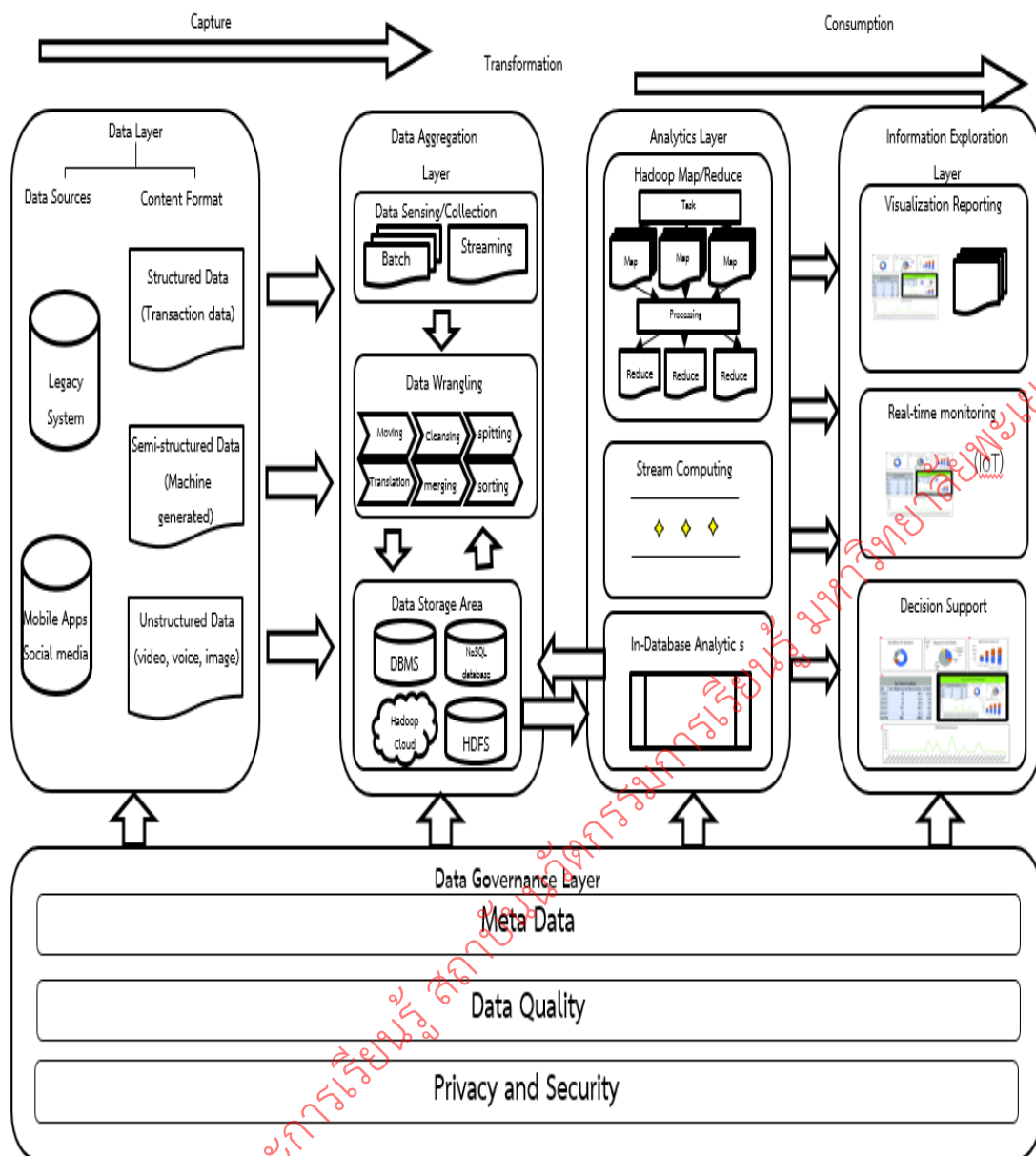
จากรูปที่ 6.1 องค์ประกอบสำหรับสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ควรประกอบประกอบไปด้วย 5 หมวดใหญ่คือกล่าวคือสถาปัตยกรรมที่ถูกออกแบบมาให้มีความสอดคล้องทั้งด้านธุรกิจ (Business) ด้านโปรแกรมประยุกต์ (Application) ด้านข้อมูล (Data) ด้านสาธารณูปโภค (Infrastructure) และด้านความปลอดภัย (Security) โดยด้านธุรกิจแบ่งออกเป็นสองหมวดประกอบไปด้วย ด้านยุทธศาสตร์ (Strategy) และกระบวนการธุรกิจ (Business Process) โดเมนกระบวนการทางธุรกิจที่ดีจะต้องสอดคล้องกับเป้าหมายและทิศทางของธุรกิจที่ตั้งไว้ โดยสามารถทำให้กิจกรรมทางธุรกิจทั้งหมดสามารถดำเนินไปได้อย่างมีประสิทธิภาพและราบรื่น

ด้านโปรแกรมประยุกต์ (Application) แบ่งออกเป็นสามหมวดประกอบไปด้วย โปรแกรมประยุกต์สำหรับธุรกิจ (Business Application) โปรแกรมประยุกต์สำหรับสนับสนุน (Supporting Application) และส่วนต่อประสานโปรแกรมประยุกต์ (Application Programming Interface) แนวคิดสำหรับการออกแบบในส่วนนี้จะต้องคำนึงถึง กระบวนการทำงานขององค์กรซึ่งสามารถแบ่ง

ออกได้เป็นกระบวนการหลัก (Primary Activities) และกระบวนการสนับสนุน (Supporting Activities) ซึ่งการเชื่อมโยงกันของกระบวนการทั้งสองนี้จะก่อให้เกิด "คุณค่า" แก่ผู้รับบริการหรือลูกค้า และแก่ผู้มีส่วนได้ส่วนเสียและต่อองค์กรธุรกิจ ดังนั้นจึงจำเป็นที่จะต้องมีการออกแบบระบบที่สนับสนุนทั้งส่วนหน้าและส่วนสนับสนุนสำหรับการจัดการกระบวนการทางธุรกิจ (Business Process Management) ให้เหมาะสม จึงจะทำให้เกิดเป็นผลสำเร็จในการดำเนินธุรกิจได้อย่างยั่งยืน

ส่วนด้านข้อมูลแบ่งออกเป็นสองหมวดประกอบไปด้วยข้อมูลที่จัดเก็บในรูปแบบฐานข้อมูลเชิงสัมพันธ์ (Relational Database) และข้อมูลที่ไม่ได้จัดเก็บในรูปแบบฐานข้อมูลเชิงสัมพันธ์ (Non-Relational Database) ด้านโครงสร้างพื้นฐานสาธารณูปโภค (Infrastructure) แบ่งออกเป็นสี่หมวดประกอบไปด้วย ระบบปฏิบัติการ (Operating System) ฮาร์ดแวร์ (Hardware) เครือข่าย (Network) และระบบอำนวยความสะดวก (Facility) ด้านความมั่นคงปลอดภัยประกอบไปด้วย การกำกับดูแล และธรรมาภิบาล (Governance) การบริหารจัดการความเสี่ยง (Risk Management) และความสอดคล้องกับกฎระเบียบ (Compliance) ซึ่งการเชื่อมโยงระหว่างส่วนประกอบต่างๆ ในแบบจำลองหรือมุมมองต่างๆ จะช่วยให้การบริหารจัดการมีประสิทธิภาพมากยิ่งขึ้น โดยทำให้สามารถพิจารณาถึงความเชื่อมโยงกิจกรรมทางธุรกิจกับระบบสารสนเทศได้อย่างมีประสิทธิภาพมากขึ้น และเน้นที่การมองเห็น (Visibility) ข้อมูลสารสนเทศตลอดทั้งองค์กร รวมถึงองค์ประกอบทุกด้านไปด้วยกัน

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมและถ่ายทอดเทคโนโลยี มหาวิทยาลัยพะเยา



รูป 6.2 สถาปัตยกรรมของข้อมูลขนาดใหญ่ (Architecture of Big Data) (Roehrs et al., 2017, Moorthy & Nobledhas, 2019)

6.2 สถาปัตยกรรมของข้อมูลขนาดใหญ่ (Architecture of Big Data)

เมื่อพูดถึงสถาปัตยกรรมของข้อมูลขนาดใหญ่ (Architecture of Big Data) องค์การที่ประสบความสำเร็จในยุคปัจจุบัน จะต้องนำข้อมูลมาใช้งานให้เป็นประโยชน์ และต้องออกแบบมีสถาปัตยกรรมด้านสารสนเทศที่ดี มี 3 กระบวนการใหญ่ (รูป 6.2) คือ การเก็บรวบรวมข้อมูล (Data Capture) การแปลงรูปข้อมูล (Data Transformation) และ การนำไปใช้ (Consumption) โดยเริ่มจากการเก็บรวบรวมข้อมูล (Capture) โดยสำรวจแหล่งที่มาของข้อมูลขนาดใหญ่ (Source of Big Data) ว่าแหล่งข้อมูลขององค์กรว่ามาจากไหนบ้าง มีขนาดใหญ่แค่ไหน กระจายอยู่หน่วยงานต่างๆ

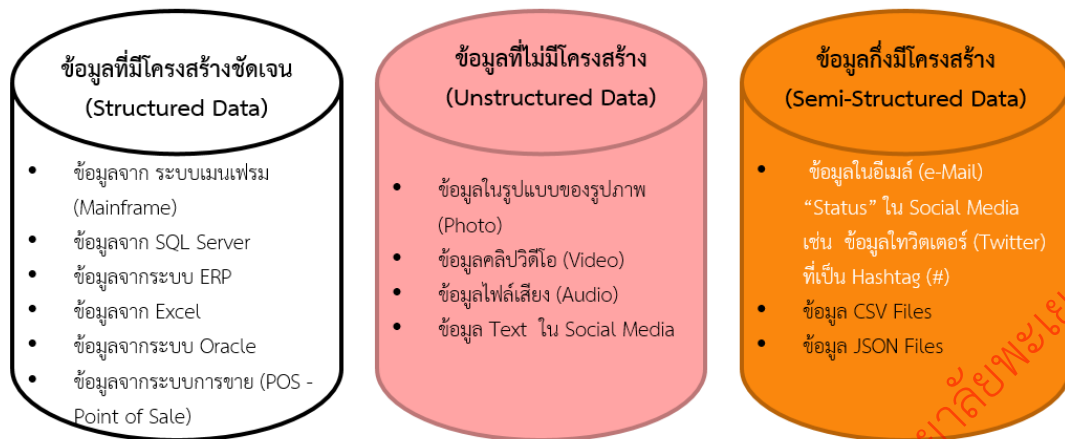
มากแค่ไหน และต้องการนำข้อมูลไปวิเคราะห์อย่างไร แล้วจึงมาวิเคราะห์ถึงเทคโนโลยีใดจะมาตอบ
โจทย์ในการบริหารจัดการ

6.2.1 ประเภทของข้อมูลโลกธุรกิจสมัยใหม่: ข้อมูลขนาดใหญ่ (Big Data)

การทำธุรกิจในยุคที่มีลูกค้าเป็นศูนย์กลาง จำเป็นต้องมีข้อมูลต่างๆ อย่างครบถ้วน เพื่อ
สามารถติดตาม วิเคราะห์ และคาดการณ์ความต้องการของลูกค้าได้อย่างแม่นยำ รวมทั้งสามารถ
ขับเคลื่อนธุรกิจให้เติบโตได้ต่อเนื่อง แม้ปัจจัยภายนอกจะผันผวน หรือต้องเผชิญกับปัจจัยเสี่ยงอย่าง
รอบด้านก็ตาม โดยจากรูป 6.3 จะเห็นได้ว่าข้อมูลของโลกธุรกิจสมัยใหม่จะประกอบไปด้วย

- ข้อมูลที่มีโครงสร้างชัดเจน (Structured Data) หมายถึง ชุดข้อมูลที่มีการ
จัดเรียงโครงสร้างอย่างเป็นระเบียบ มีความชัดเจน หรือเป็นตัวเลข พร้อมใช้งานได้ทันที เช่น ข้อมูลที่มา
จากระบบธุรกิจ ตัวอย่างเช่น จากระบบ POS (Point of Sale) ของร้านสะดวกซื้อ หรือ Modern
Trade ข้อมูลที่อยู่ในรูปของไฟล์ CSV หรือไฟล์ Excel ซึ่งอยู่ในรูปแบบตาราง ข้อมูลของประชากรที่
รัฐบาลที่ทำการจัดเก็บไว้ในระบบทะเบียนราษฎร หรือทะเบียนขนส่ง ตัวอย่างเช่นข้อมูลใบขับขี่ ข้อมูล
บัตรประจำตัวประชาชน ข้อมูลของธนาคารที่ ตัวอย่างเช่น ข้อมูลลูกค้าบัญชีเงินฝาก ข้อมูลเครดิตทาง
การเงิน ข้อมูลการใช้จ่ายผ่าน Mobile App เป็นต้น
- ข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) หมายถึง ชุดข้อมูลที่มี
โครงสร้างไม่ชัดเจน หรือไม่สามารถระบุความแน่นอนของข้อมูลนั้นๆ ได้ การประมวลผลจำเป็นต้องใช้
เครื่องมือเพื่อแปลงให้เป็นข้อมูล Structure data ก่อนนำไปใช้งาน เช่น ข้อมูลในรูปแบบของรูปภาพ
คลิปวิดีโอ ไฟล์เสียง ข้อมูลเหล่านี้จะมีจำนวนมากกว่าข้อมูลในรูปแบบ Structured data และเป็น
ข้อมูลที่ทำให้เกิด Big Data
- ข้อมูลกึ่งมีโครงสร้าง (Semi-Structured Data) หมายถึง ชุดข้อมูลที่มี
โครงสร้างระดับหนึ่งแต่ยังไม่สมบูรณ์ ซึ่งอยู่ตรงกลางระหว่าง Structured data และ Unstructured
Data โดยมีตัวกำกับไว้ว่าข้อมูลนี้คืออะไร ข้อมูลนี้อยู่ในลำดับขั้นต่อจากอะไร ทำให้สามารถเอา
ประมวลผลได้ง่ายกว่าข้อมูลแบบไร้โครงสร้าง เช่น ข้อมูลในอีเมล (e-Mail) “Status” ใน Social
Media เช่น ข้อมูลทวีตเตอร์ (Twitter) โดยเป็นข้อมูลที่ไม่มีโครงสร้าง แต่ในกรณีที่มี Hashtag (#)
เข้ามาช่วยในการจัดหมวดหมู่

ประเภทของข้อมูลสารสนเทศสมัยใหม่ และแหล่งที่มา



รูป 6.3 ประเภทของข้อมูลสารสนเทศสมัยใหม่ (ผู้เขียน)

6.2.2 คุณลักษณะ 7 ประการของข้อมูลสารสนเทศสมัยใหม่

ในการบริหารองค์กรสมัยใหม่ ข้อมูลถือว่าเป็นทรัพยากรที่สำคัญเพิ่มเติมจาก ที่ดิน ไฟฟ้า น้ำ และสาธารณูปโภค ที่จำเป็นต่อธุรกิจ สำหรับภาคธุรกิจแล้ว ข้อมูลสารสนเทศสมัยใหม่ขนาดใหญ่หรือที่เรียกว่า "Big Data" เป็นข้อมูลที่เก็บรวบรวมพฤติกรรมและกิจกรรมของลูกค้าที่เกิดขึ้นจริง เป็นเสมือนขุมทรัพย์ที่องค์กรธุรกิจสามารถนำไปวิเคราะห์เพื่อให้ได้ ข้อมูลเชิงลึก (Insight Data) ที่เป็นประโยชน์สำหรับการกำหนดกลยุทธ์ นำไปคาดการณ์พฤติกรรมผู้บริโภคได้มากกว่าการวิเคราะห์ข้อมูลแบบดั้งเดิมในอดีต โดยตัวอย่างจากการทำวิจัยของผู้เขียนและทีมศาสตราจารย์จากมหาวิทยาลัยประเทศญี่ปุ่น เกี่ยวกับการจัดเก็บข้อมูลจาก นาฬิกาหรืออุปกรณ์พกพา ในลักษณะของข้อมูลชีวิตประจำวัน "Lifelogs" (Chertchom et al., 2017) ซึ่งคือการเก็บข้อมูลส่วนตัวในการใช้ชีวิตประจำวัน บริษัทผู้ให้บริการต่างๆ ก็ได้ออกแบบฐานข้อมูลซึ่งสามารถจัดเก็บ เพื่อบันทึก/บันทึกเหตุการณ์ในชีวิตของตนผ่านทางคอมพิวเตอร์ส่วนบุคคลและอุปกรณ์เคลื่อนที่ แอปพลิเคชันและอุปกรณ์รุ่นใหม่ได้รับการพัฒนาสำหรับบุคคลในการเขียนอัตชีวประวัติดิจิทัลเพื่อระลึกถึงเหตุการณ์ ทบทวนความคืบหน้า และปรับปรุงตนเองเพื่อชีวิตที่ดีขึ้น โดยปกติ เทคโนโลยีเหล่านี้จะรวมอยู่ในอุปกรณ์สวมใส่อัจฉริยะสำหรับการบันทึกข้อมูลชีวิตประจำวัน "Lifelog" และมีข้อมูลครบทั้ง 3 ประเภทเช่น สถิติการดูแลสุขภาพส่วนบุคคล หรือสุขภาพ หรือข้อมูลที่เก็บพิกัด (Location) จากการใช้ชีวิตประจำวัน โดยยี่ห้อที่เป็นที่นิยมได้แก่ Garmin Apple Waych Samsung The 37° Bracelet Fitbit One™ SAGA Google Glass Google Fit และ tagCAM เป็นตัวอย่างของอุปกรณ์/แอปที่สวมใส่ได้อัจฉริยะสำหรับการจัดการสุขภาพ เทคโนโลยีบันทึกชีวิต อุปกรณ์เหล่านี้ได้รับการพัฒนาขึ้นเพื่อช่วยเหลือบุคคลอย่างต่อเนื่องในการบันทึก

กิจกรรมประจำวันโดยอัตโนมัติตลอดจนในการจับภาพช่วงเวลาที่สำคัญในชีวิต และยังสามารถแบ่งปันรายละเอียดของบันทึกชีวิตของผู้ใช้กับเพื่อน ครอบครัว และผู้ติดตามในรูปแบบเปิดและปิด แต่อีกด้านหนึ่งคือ ผู้ให้บริการได้จัดเก็บข้อมูลไว้บนระบบคลาวด์ (Cloud System) ทั้งนี้เพื่อใช้ในการทำเหมืองข้อมูล (Data Mining) และนำผลมาใช้ในการปรับปรุงผลิตภัณฑ์ รวมทั้งเข้าใจพฤติกรรมของผู้บริโภคแบบเรียลไทม์

ซึ่งข้อมูลที่ถูกกล่าวมานี้จะถูกเรียกในชื่อว่า “7Vs” มีลักษณะดังต่อไปนี้ ดังรูป 6.4



รูป 6.4 คุณลักษณะของข้อมูลสารสนเทศสมัยใหม่ (ผู้เขียน)

1. ปริมาณ (Volume) หมายถึง ข้อมูลในสมัยปัจจุบันเป็นข้อมูลปริมาณมหาศาลที่เกิดจากการเก็บของระบบหลายๆระบบ ตัวอย่างเช่น จากอุปกรณ์ Smart Phone Smart Watch แอปพลิเคชัน เช่น Google Youtube Twitter Shopee หรืออุปกรณ์ IoT (Internet of Things) หรือระบบ POS (Point of Sale) ของร้าน Modern Trade ทำให้เมื่อนำปริมาณมหาศาล ที่เก็บไว้ เช่น ข้อมูลอายุ เพศ ของลูกค้าส่วนใหญ่ พฤติกรรมการซื้อ การใช้ชีวิตมาวิเคราะห์แล้วจะได้ความรู้เชิงลึก

ทำให้องค์กรธุรกิจเห็นรูปแบบของลูกค้าหรือเห็นรูปแบบโพรล์ (Demographic Profile) ของลูกค้าที่ถูกต้องได้

2. ความหลากหลาย (Variety) หมายถึง รูปแบบของข้อมูลในแต่ประเภทที่มีความหลากหลาย ทั้งแบบโครงสร้าง (Structured) กึ่งโครงสร้าง (Semi Structured) ไม่มีโครงสร้าง (Unstructured)

3. ความเร็ว (Velocity) หมายถึง คุณลักษณะข้อมูลที่ถูกสร้างขึ้นอย่างรวดเร็วต่อเนื่องและเป็นแบบวินาที (Real Time) ซึ่งในที่นี้จะหมายถึงความเร็วในการเพิ่มขึ้นของข้อมูลจากแหล่งข้อมูล ทำให้องค์กรธุรกิจเห็นข้อมูลและสามารถวิเคราะห์ข้อมูลแบบ Real-Time

4. ความถูกต้อง (Veracity) หมายถึง ข้อมูลต้องมีความน่าเชื่อถือ ทราบแหล่งที่มา และมีความถูกต้อง มีกระบวนการในการตรวจสอบและยืนยันความถูกต้องของข้อมูล ซึ่งมีผลโดยตรงกับการวิเคราะห์ข้อมูล เนื่องจากระบบจะประมวลผลจากข้อมูลที่นำเข้าเท่านั้น การนำขยะเข้าจะได้ผลลัพธ์ขยะออก (Garbage In Garbage Out, GIGO) เป็นคำที่มักกล่าวเตือนสติในงานวิเคราะห์ของระบบสารสนเทศ “ถ้าใส่ข้อมูลที่ไร้คุณค่าลงไปในระบบงานคอมพิวเตอร์ก็จะได้ข้อมูลที่ไม่มีค่ากลับมา”

5. คุณค่า (Value) หมายถึง ข้อมูลที่มีคุณค่าต่อการนำไปใช้งาน มีประโยชน์ต่อการตัดสินใจของธุรกิจ โดยข้อมูลที่มีประโยชน์จะต้องตอบวัตถุประสงค์ทางธุรกิจ แก้ไขปัญหา และใช้วางแผนกลยุทธ์

6. ความแปรผันได้ (Variability) หมายถึง ความแปรปรวนในบริบทของข้อมูลขนาดใหญ่ มีบางสิ่งที่แตกต่างกัน เช่น จำนวนของความไม่สอดคล้องกันในข้อมูล จำเป็นต้องค้นพบสิ่งเหล่านี้โดยวิธีการตรวจจับความผิดปกติและค่าผิดปกติเพื่อให้การวิเคราะห์ที่มีความหมายเกิดขึ้นข้อมูล โดยคุณลักษณะของข้อมูลแบบนี้ คือสามารถในการเปลี่ยนแปลงรูปแบบไปตามการใช้งานของธุรกิจและมุมมองการวิเคราะห์ โดยสามารถคิดวิเคราะห์ข้อมูลที่มีอยู่ได้จากหลายแง่มุม ข้อมูลขนาดใหญ่ยังแปรผันได้เนื่องจากมีข้อมูลจำนวนมากที่เกิดจากชนิดข้อมูลและแหล่งที่มาที่แตกต่างกันหลายประเภท ความแปรปรวนยังหมายถึงความเร็วที่ไม่สอดคล้องกันที่ไหลของข้อมูลขนาดใหญ่ลงในฐานข้อมูล

7. สร้างภาพได้ (Visualization) จากรูป 6.2 งานหลักในกล่องสุดท้ายสำหรับระบบประมวลผล Big Data คือการแปลงขนาดมหึมา ให้เข้าใจได้ง่ายและนำไปใช้ได้ สำหรับการจัดการของธุรกิจ ซึ่งก็คือการแปลงเป็นรูปแบบกราฟิก ดังนั้นข้อมูลที่ได้มาต้องสามารถนำมาเสนอในรูปแบบของรายงานที่หลากหลายได้โดยเฉพาะในรูปแบบกราฟิก (Graphical Formats)

6.2.3 การจัดเก็บข้อมูล (Capture) และ การแปลงข้อมูล (Transformation) จากรูป 6.2 จากกล่องหรือเลเยอร์แหล่งที่มาของข้อมูล (Capture/Data Layer) ข้อมูลอาจจะมาจากระบบจากโปรแกรมหรือจากมนุษย์สร้างให้เกิดข้อมูลขึ้นมา ข้อมูลต่างๆ จะมาจากแหล่งข้อมูลที่หลากหลาย

และมีความยากลำบากในการจัดการโครงสร้าง หรือจัดเตรียมให้ข้อมูลให้พร้อมใช้ต่อไป ดังนั้นขั้นตอนต่อไป คือกระบวนการการตรวจจับข้อมูลหรือการรวบรวมข้อมูลที่เกี่ยวข้อง (Data collection) จากอุปกรณ์ ระบบ หรือจากการติดต่อของมนุษย์ กับมนุษย์ ผ่านระบบคอมพิวเตอร์ สำหรับการรวบรวมข้อมูล ข้อมูลถูกนำไปแปลง (Transformation) ให้อยู่ในรูปแบบที่ประมวลผลได้ เทคนิคเริ่มต้นนี้เรียกว่า การทำ ETL (Extract-Transform-Load) เป็นการนำข้อมูล (Extract) จากหลากหลายแหล่งข้อมูล เช่น จากทุกแหล่งของแหล่งข้อมูล (Data Sources) เช่นจากระบบการซื้อขาย ระบบการตลาด จากระบบ CRM (Customer Relationship) จากโซเชียลมีเดีย (Social Media) เป็นต้น มาแปลง (Transform) รูปแบบให้เหมาะสม ปรับแต่งข้อมูลให้อยู่ในรูปแบบเดียวกัน นำไปประมวลผล แล้วจัดเก็บ (Load) เข้าไปในคลังข้อมูล (Data Warehouse) หรือทะเลสาบข้อมูล (Data Lake) โดยมีกระบวนการที่เรียกว่า ขั้นตอนการนำข้อมูลดิบ (Data ingestion) คือขั้นตอนการนำข้อมูลดิบ (Raw data) จากแหล่งต่างๆ ของแหล่งข้อมูล (Data Sources) เข้ามาในกระบวนการ ซึ่งอาจต้องใช้เทคโนโลยีที่แตกต่างกันสำหรับแหล่งข้อมูลหลากหลายชนิดนำไปสู่กระบวนการ ส่วน Data Storage คือขั้นตอนการเก็บข้อมูลมาไว้ที่เดียวกัน ซึ่งเทคโนโลยีในการเก็บข้อมูลจากหลายแหล่งอาจเป็นคลังข้อมูล (Data Warehouse) หรือทะเลสาบข้อมูล (Data Lake) โดยจะต้องมีการแปลงข้อมูล (Data Wrangling) คือกระบวนการแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกับการทำการวิเคราะห์ข้อมูล (Data Analysis/Business Analytics) และ การแสดงข้อมูลในรูปแบบรูปภาพ (Data Visualization) ก่อนที่จะนำไปใช้ในเชิงธุรกิจต่อไป (Service)

ในเลเยอร์ข้อมูล (Data Layer) เป็นส่วนที่รวมแหล่งข้อมูลทั้งหมดที่จำเป็นเพื่อนำเข้าสู่ระบบข้อมูลขนาดใหญ่ ที่จำเป็นต่อสนับสนุนการดำเนินงานประจำวันและแก้ปัญหาทางธุรกิจ ข้อมูลที่เข้ามามีคุณลักษณะหลากหลายและมีหลายประเภท หลากหลายโครงสร้างเช่น

- ข้อมูลแบบโครงสร้าง (Traditional DB) ตัวอย่างเช่น ข้อมูลจากระบบปฏิบัติการประจำวัน (Transaction System)
- ข้อมูลกึ่งโครงสร้าง เช่น ข้อมูลจากรายการคอมพิวเตอร์ (Log) ข้อมูลจากอุปกรณ์ตรวจสอบ

- และข้อมูลที่ไม่มีโครงสร้าง เช่น ข้อมูลภาพ ข้อมูลเสียง เป็นต้น ข้อมูลทางเหล่านี้จะถูกเก็บรวบรวมจาก แหล่งทั้งภายในและภายนอกองค์กร โดยขึ้นอยู่กับชนิดของข้อมูล และรูปแบบเนื้อหา

เลเยอร์การรวมข้อมูล (Data Aggregation Layer) เลเยอร์นี้มีหน้าที่จัดการข้อมูลจากแหล่งข้อมูลต่างๆ โดยข้อมูลจะถูกแยกย่อย โดยมีการดำเนินการสามขั้นตอน การเก็บข้อมูล (Data Acquisition) การแปลงข้อมูล (Data Transformation) และการจัดเก็บข้อมูล (Data Storage)

โดยเฉพาะการแปลงข้อมูลใน เลเยอร์นี้ เครื่องมือที่ใช้ (The Transformation Engine) จะต้องมี
ความสามารถในการ Moving Cleaning Splitting Translating Merging Sorting และ Validating
Data ได้อย่างมีประสิทธิภาพ

เลเยอร์การวิเคราะห์ (Analytics Layer) เลเยอร์นี้มีหน้าที่ในการประมวลผลข้อมูลทุก
ประเภท และมีเครื่องมือวิเคราะห์ ข้อมูลขนาดใหญ่

เลเยอร์การสำรวจข้อมูล (Information Exploration Layer) เลเยอร์นี้มีเครื่องมือ
สร้างรายงานผลลัพธ์ เช่น รายงานข้อมูลแบบเรียลไทม์และข้อมูลเชิงลึกทางธุรกิจ

เลเยอร์การกำกับดูแลข้อมูล (Data Governance Layer) เลเยอร์นี้ประกอบด้วย
จัดการข้อมูลหลัก (Master Data Management: MDM) การจัดการวงจรข้อมูล และการจัดการความ
ปลอดภัยของข้อมูลและความเป็นส่วนตัว

โดยตาราง 6.1 ได้นำเสนอเครื่องมือเพิ่มเติมสำหรับสถาปัตยกรรมระบบสารสนเทศสมัยใหม่
ดังนี้

ตารางที่ 6.1 เครื่องมือสำหรับการจัดการข้อมูลขนาดใหญ่ (Moorthy & Nobledhas, 2019)

กิจกรรม (Task)	เครื่องมือและแอปพลิเคชัน Tools Merits & its Applications
1. การรวบรวมข้อมูล (Data Integration)	เครื่องมือที่ใช้ในการรวบรวมข้อมูลทั้งหมดที่มีในแหล่งต่างๆ มาเก็บไว้ในพื้นที่ ปลายทาง เช่น ฐานข้อมูล แพลตฟอร์มจัดการข้อมูล Data Lake เพื่อความสะดวก ในการค้นหาและการนำไปใช้งานได้ต่อ โดยเครื่องมือต้องสามารถยืดหยุ่น สำหรับ ใช้งาน เช่น การทำ Service ที่ใช้กับ Extract-Transform-Load (ETL) หรือ การ จัดเตรียมข้อมูล ให้อยู่ในรูปแบบที่เหมาะสม (Data Pipeline)
2. การค้นหาและประมวลผล (Searching and Processing)	เครื่องมือที่ใช้เปลี่ยนข้อมูลให้อยู่ในรูปแบบที่ตรงกับความต้องการ ซึ่งจะครอบคลุม ไปในส่วนของการเก็บรวบรวม การบันทึก การจัดเก็บ การวางโครงสร้าง การนำ ข้อมูลมาใช้ เพื่อให้สามารถนำไปประมวลผล วิเคราะห์ จัดเก็บ และนำเสนอได้ใน รูปแบบที่ต้องการ
3. การเรียนรู้ของเครื่อง (Machine Learning)	เครื่องมือสำหรับการเรียนรู้ของเครื่องและการวิเคราะห์ขั้นสูงของชุดข้อมูลขนาด ใหญ่ด้วยความเร็วสูง
4. การประมวลผลข้อมูลแบบ ต่อไหลต่อเนื่อง (Stream Data Processing)	เครื่องมือสำหรับการสืบค้นหรือประมวลผลข้อมูลภายในหน้าต่างแสดงเวลา Rolling หรือในบันทึกข้อมูลล่าสุด การประมวลผลข้อมูลจะต้องใช้ 2 เลเยอร์ คือ เลเยอร์การจัดเก็บข้อมูล และเลเยอร์การประมวลผล เลเยอร์การจัดเก็บข้อมูลต้อง รองรับการจัดเรียงบันทึกและความสม่ำเสมอเป็นอย่างมากเพื่อให้อ่านและบันทึก ค่าสตรีม (Streaming Data) ข้อมูลขนาดใหญ่ได้อย่างรวดเร็ว

5. การนำเสนอข้อมูลวิเคราะห์ด้วยภาพ (Visual Data Analytics)	เครื่องมือสำหรับการวิเคราะห์ข้อมูล ใช้ Data Visualization เป็นเครื่องมือหรือสื่อ เพื่อให้ผู้ใช้งานสามารถสำรวจ ค้นหา และเข้าใจข้อมูลได้ตามที่ต้องการ
--	---

6.2.4 เลเยอร์การวิเคราะห์ข้อมูล (Analytics Layer) และ เลเยอร์การค้นหาความรู้จากข้อมูล (Information Exploration Layer) เลเยอร์นี้มีกระบวนการของการประมวลผลข้อมูลทุกประเภทและดำเนินการวิเคราะห์ที่เหมาะสม ในเลเยอร์นี้การวิเคราะห์ข้อมูลสามารถแบ่งออกเป็นสามองค์ประกอบหลักคือ หนึ่ง การจัดสรรทรัพยากรในการจัดเก็บและประมวลผล ซึ่งถ้าหากองค์กรใช้เทคโนโลยี Hadoop นั่นก็คือกระบวนการ “Hadoop Map/Reduce” สอง การประมวลผลแบบสตรีม (Stream Computing) และสามการวิเคราะห์ในฐานข้อมูล (In-Database Analytics) ซึ่งขึ้นอยู่กับประเภทของข้อมูลและวัตถุประสงค์ของการวิเคราะห์ “Hadoop Map/Reduce” เป็นรูปแบบการเขียนโปรแกรมที่ใช้บ่อยที่สุดในปัจจุบัน โปรแกรมหนึ่งในการวิเคราะห์ข้อมูลขนาดใหญ่ ซึ่งให้ความสามารถในการประมวลผลข้อมูลปริมาณมากในรูปแบบแบทช์ (Batch) ได้อย่างมีประสิทธิภาพตลอดจนช่วยให้สามารถวิเคราะห์ข้อมูลทั้งที่ไม่มีโครงสร้างและที่มีโครงสร้างในรูปแบบการประมวลผลแบบขนานขนาดใหญ่ (A Massively Parallel Processing : MPP)

ในส่วนเลเยอร์การค้นหาความรู้จากข้อมูล (Information Exploration Layer) เลเยอร์ประกอบไปด้วยกระบวนการสร้างผลลัพธ์จากการวิเคราะห์ข้อมูลออกมาเป็น ตัวอย่างเช่น รายงานแสดงด้วยภาพกราฟฟิค ข้อมูลแบบเรียลไทม์สำหรับการตรวจสอบ และมีกระบวนการนำเสนอข้อมูลเชิงลึกทางธุรกิจที่มีความหมาย เพื่อส่งต่อไปยังผู้ใช้ (User) ในองค์กร กระบวนการนี้คล้ายคลึงกับระบบธุรกิจอัจฉริยะเดิม (Business Intelligence Platform) ซึ่งจะมีเครื่องมือการวิเคราะห์ข้อมูลขนาดใหญ่ที่ทำให้ข้อมูลถูกแสดงเป็นภาพในลักษณะที่เป็นประโยชน์เพื่อสนับสนุนการดำเนินงานประจำวันของผู้ใช้และช่วยให้ผู้บริหารสามารถตัดสินใจได้เร็วและดีขึ้น อย่างไรก็ตามในระบบสารสนเทศสมัยใหม่สิ่งที่สำคัญคือ ข้อมูลเชิงลึกที่เป็นข้อมูลแบบเรียลไทม์ใช้ในการแจ้งเตือนเตือนเชิงรุก การแสดงตัวบ่งชี้ประสิทธิภาพหลักในการปฏิบัติงาน (KPI) แบบเรียลไทม์ เป็นต้น

6.2.5 เลเยอร์การกำกับดูแลข้อมูล (Data governance layer) เลเยอร์นี้ประกอบด้วยกระบวนการจัดการข้อมูลหลัก (Master Data Management: MDM) การจัดการวงจรชีวิตของข้อมูล และการรักษาความปลอดภัยข้อมูล (Data Security) และการจัดการความเป็นส่วนตัว (Data Privacy) เลเยอร์นี้เน้น "วิธีการ" เช่นเดียวกับวิธีควบคุมข้อมูลในองค์กร องค์ประกอบแรกของการกำกับดูแลข้อมูลและการจัดการกำกับดูแลข้อมูล ถือเป็นกระบวนการที่ใช้กำกับดูแลกำหนดนโยบาย สร้างกรอบมาตรฐาน และเครื่องมือสำหรับการจัดการข้อมูล ข้อมูลจะถูกทำให้เป็นมาตรฐาน มีนโยบายในการจัดการลบข้อมูล นำเข้านำออก และรวมเข้าด้วยกันเพื่อสร้างความรวดเร็ว

ความสมบูรณ์ ความถูกต้อง และความพร้อมใช้งานของข้อมูลสำหรับสนับสนุนการวิเคราะห์ข้อมูลและการตัดสินใจ

นอกจากนี้ในलेयरการกำกับดูแลข้อมูลนี้ มีเครื่องมือและกระบวนการในการจัดการวงจรชีวิตของข้อมูล ซึ่งเป็นกระบวนการเริ่มตั้งแต่การจัดเก็บข้อมูลแบบถาวร ไปจนถึงการบำรุงรักษาค้างข้อมูล การทดสอบและการส่งมอบระบบแอปพลิเคชันต่างๆ ไปจนถึงการลบและกำจัดข้อมูลเพื่อให้ระบบสารสนเทศมีประสิทธิภาพสูงสุด นอกจากนี้ยังมีกระบวนการการจัดการความปลอดภัยของข้อมูล และความเป็นส่วนตัว (Security and Privacy) ซึ่งเป็นกระบวนการ การเฝ้าติดตาม การตรวจสอบ และการป้องกัน เพราะในโลกธุรกิจมีความซับซ้อนในการจัดการข้อมูล องค์กรต้องเผชิญกับความท้าทายด้านจริยธรรม กฎหมาย และกฎระเบียบด้วยการกำกับดูแลข้อมูลที่ออกมาตลอดเวลาเพื่อปกป้องผู้บริโภคและระบบธุรกิจ

6.3 ประโยชน์ของสถาปัตยกรรมองค์กรสำหรับระบบข้อมูลขนาดใหญ่

ในการขับเคลื่อนองค์กรยุคใหม่และนำใช้นวัตกรรมด้านดิจิทัลที่มีการนำเทคโนโลยีหลากหลายเช่น ปัญญาประดิษฐ์ (Artificial intelligence: AI) ข้อมูลขนาดใหญ่ (Big Data) ไอโอที (IoT) ซอฟต์แวร์โรบอทที่เลียนแบบพฤติกรรมของมนุษย์ (Robotic Process Automation: RPA) ระบบคลาวด์ (Cloud) รวมถึงกรอบแนวทางในการพัฒนาด้านดิจิทัลเข้ามาใช้ ล้วนแล้วมีความเชื่อมโยงกันในองค์กร โดยที่สถาปัตยกรรมองค์กร (Enterprise Architecture :EA) จะให้ประโยชน์ได้หรือไม่เพียงใดนั้นขึ้นอยู่กับความสามารถของสถาปนิก (Architect) ที่จะออกแบบแผนพิมพ์เขียว (Blueprint) ให้กลุ่มผู้เกี่ยวข้อง (Stakeholders) ได้เข้าใจกระบวนการธุรกิจทั้งหมด เข้าใจระบบงานและรู้จักผู้รับผิดชอบ และต้องสัมพันธ์กับเทคโนโลยีที่ใช้และทำงานร่วมกับส่วนงานอื่นได้อย่างมีประสิทธิภาพ วัตถุประสงค์ของสถาปัตยกรรมองค์กรคือการสร้างแผนที่ของระบบสารสนเทศและกระบวนการทางธุรกิจให้สอดคล้องกับกลยุทธ์ทางธุรกิจจึงควรมองภาพรวมทั่วทั้งองค์กร ตั้งแต่วิสัยทัศน์ ยุทธศาสตร์ กระบวนการทำงาน มีการบูรณาข้อมูลต่างๆ ดังกล่าวกับเทคโนโลยีสารสนเทศขององค์กร เพื่อให้เกิดความรู้ ความเข้าใจที่ตรงกันก่อนที่จะปรับปรุงและขับเคลื่อนไปสู่องค์กรดิจิทัลต่อไป การจัดทำสถาปัตยกรรมองค์กร (Enterprise Architecture:EA) หรือที่เรียกว่า พิมพ์เขียวองค์กรดิจิทัล (Digital Enterprise Blueprint) จึงมีความสำคัญอย่างยิ่งที่จะทำให้ผู้นำเกิดความเข้าใจที่ตรงกัน เพื่อออกแบบการประยุกต์ใช้เทคโนโลยีสารสนเทศให้เชื่อมโยงกับการดำเนินงาน ดำเนินการพัฒนาบริการใหม่ ๆ ยกเลิกกระบวนการที่ไม่จำเป็นและจัดทำแผนที่นำทางองค์กรดิจิทัลในระยะต่าง ๆ (Digital Enterprise Roadmap) ให้ตรงกับเป้าหมายอนาคตองค์กร และยังช่วยลดความซ้ำซ้อนของการทำงานและการลงทุนอีกด้วย ซึ่งจะสามารถผลักดันให้องค์กรสามารถดำเนินการตามนโยบายและวิสัยทัศน์ขององค์กรที่กำหนดไว้ได้อย่างมีประสิทธิภาพดังต่อไปนี้

- **ลดความซับซ้อน (Decreasing Complexity)** การออกแบบสถาปัตยกรรมองค์กรจะช่วยลดความซับซ้อนของไอที ระบบที่ซับซ้อนซึ่งจัดการได้ยาก อาจทำให้เกิดข้อผิดพลาดได้มากมาย รวมทั้งลดประสิทธิภาพของการดำเนินการทางธุรกิจ การออกแบบสถาปัตยกรรมองค์กรจะช่วยลดความซับซ้อนของระบบสารสนเทศ การรวมข้อมูลและแอปพลิเคชัน และเพื่อการทำงานร่วมกัน ของระบบ เป็นวิธีแก้ปัญหาความซับซ้อนของระบบในองค์กร ดังนั้น การลดความซับซ้อนของระบบจึงเป็นหนึ่งในความคาดหวังที่ใหญ่ที่สุดขององค์กร สถาปัตยกรรมองค์กรช่วยลดความยุ่งยากในการจัดการระบบด้วยเครื่องมือที่เหมาะสม และลดความซับซ้อนของระบบโดยการตรวจจับและหยุดกระบวนการที่ไม่ส่งผลต่อเป้าหมาย
- **สร้างมาตรฐานให้กับองค์กร (Standardization)** พิมพ์เขียวของสถาปัตยกรรมองค์กรสร้างกระบวนการขององค์กรและมาตรฐานกระบวนการในการดำเนินธุรกิจ ช่วยกำหนดระเบียบวินัยและการกำหนดมาตรฐานของกระบวนการทางธุรกิจ กระบวนการมาตรฐานจะช่วยเพิ่มประสิทธิภาพ โครงสร้างพื้นฐานด้านระบบสารสนเทศที่ดีขึ้นจะส่งให้ธุรกิจประสบความสำเร็จมากขึ้น นอกจากนี้ยังช่วยเพิ่มความเร็วของการกำหนดความต้องการและความถูกต้องของข้อกำหนด ผ่านการระบบการจัดทำเอกสาร
- **ลดเวลาที่เสียไปและลดต้นทุน (Reducing Wasted Time and Cost Saving)** สถาปัตยกรรมองค์กรช่วยประหยัดเวลาสำหรับพนักงานที่บริหารจัดการระบบสารสนเทศ ทำให้สามารถตรวจพบสาเหตุของปัญหาได้เร็วขึ้น ซึ่งช่วยลดการสูญเสียเวลาและแรงงานที่เกิดจากการหยุดชะงัก การออกแบบที่ดีจะช่วยจัดการกับกระบวนการที่ไม่เป็นประโยชน์ ซึ่งรวมถึงกระบวนการที่ทำซ้ำโดยไม่จำเป็น ส่งผลให้ประหยัดทรัพยากรและต้นทุนได้อย่างมาก นอกจากนี้พิมพ์เขียวองค์กรดิจิทัล (Digital Enterprise Blueprint) ยังช่วยประหยัดงบประมาณด้านสารสนเทศด้วยการปรับโครงสร้างพื้นฐานด้านไอทีให้เหมาะสมตามมุมมองที่สอดคล้องกับวัตถุประสงค์ทางธุรกิจ นอกจากนี้สถาปัตยกรรมองค์กรช่วยสร้างระบบการสื่อสารระหว่างผู้มีส่วนได้ส่วนเสียของโครงการ สนับสนุนการกำหนดขอบเขตของโครงการต่างๆ อย่างมีประสิทธิภาพ
- **การวิเคราะห์การเปลี่ยนแปลง (Change Analysis)** พิมพ์เขียวของสถาปัตยกรรมองค์กรเป็นเครื่องมือช่วยวิเคราะห์และตอบสนองต่อการเปลี่ยนแปลงในอุตสาหกรรมได้เร็วขึ้น สามารถใช้ประกอบการตัดสินใจเชิงกลยุทธ์และปรับกระบวนการด้านสารสนเทศให้เข้ากับโมเดลทางธุรกิจใหม่หรือสร้างความยั่งยืนเพื่อรองรับตามการเปลี่ยนแปลงของโลกธุรกิจ
- **ความปลอดภัยที่แข็งแกร่ง (Strong Security)** สถาปัตยกรรมองค์กรเป็นพิมพ์เขียวที่อธิบายโครงสร้างของระบบที่ช่วยระบุมাত্রการที่ใช้กับการละเมิดความปลอดภัยในโลกไซเบอร์ในปัจจุบัน เป็นหนึ่งในเครื่องมือที่สำคัญที่สุด ในการปกป้องข้อมูลด้วยการเสริมความปลอดภัยทาง

ไซเบอร์ของธุรกิจ องค์กรจำเป็นต้องตื่นตัวเพื่อป้องกันภัยคุกคามที่อาจเข้าสู่ระบบจากจุดอ่อนได้ตลอดเวลา การรั่วไหลและการขโมยข้อมูลที่เป็นไปได้ จะทำให้เกิดความสูญเสียอย่างมหาศาลทางธุรกิจ

กรณีศึกษา: การปรับปรุงกระบวนการธุรกิจโดยการออกแบบสถาปัตยกรรมองค์กร

ที่มาของปัญหา: บริษัทในอุตสาหกรรมบริการทางการเงินพบว่ากระบวนการธุรกิจทั้งหมดไม่เชื่อมต่อกันอย่างเพียงพอ มีการกระจายข้อมูลและขั้นตอนการทำงานที่ซับซ้อน ส่งผลให้มีการล่าช้าในการบริหารงานและการตอบสนองต่อลูกค้าที่ช้า

วิธีการ: บริษัทตัดสินใจทำการปรับปรุงกระบวนการธุรกิจโดยการออกแบบสถาปัตยกรรมองค์กรใหม่ พวกเขาเริ่มต้นด้วยการวิเคราะห์กระบวนการทางธุรกิจที่มีอยู่และที่ที่ต้องการปรับปรุง จากนั้นพวกเขาออกแบบโครงสร้างองค์กรใหม่ที่เน้นการเชื่อมต่อและการกระทำที่มีประสิทธิภาพ พวกเขาใช้หลักการของ **Enterprise Architecture (EA)** เพื่อกำหนดเป้าหมายและระบบใหม่

ผลลัพธ์: ผลลัพธ์ที่ได้คือการออกแบบสถาปัตยกรรมองค์กรใหม่ที่ช่วยให้กระบวนการธุรกิจทั้งหมดเชื่อมต่อกันได้อย่างรวดเร็วและมีประสิทธิภาพ ระบบเทคโนโลยีสารสนเทศ ถูกออกแบบให้รองรับกระบวนการธุรกิจที่เปลี่ยนแปลงได้อย่างยืดหยุ่น การปรับปรุงนี้ช่วยลดเวลาในการดำเนินงาน ลดความซับซ้อนในการจัดการและเพิ่มความพร้อมในการตอบสนองต่อลูกค้า

ประโยชน์: การปรับปรุงกระบวนการธุรกิจโดยการออกแบบสถาปัตยกรรมองค์กรช่วยให้บริษัทสามารถดำเนินงานอย่างมีประสิทธิภาพและสอดคล้องกับความต้องการของลูกค้า มีการปรับปรุงความสามารถในการตอบสนองต่อการเปลี่ยนแปลงในตลาดและแวดวงธุรกิจ นอกจากนี้ยังช่วยประหยัดค่าใช้จ่ายในการดำเนินงานและบำรุงรักษาระบบเทคโนโลยีสารสนเทศ

ในกรณีศึกษาด้านการปรับปรุงกระบวนการโดยการออกแบบสถาปัตยกรรมองค์กรนี้ เราเห็นว่าการใช้หลักการ **Enterprise Architecture (EA)** และการออกแบบโครงสร้างองค์กรใหม่ช่วยเพิ่มประสิทธิภาพและความเชื่อมต่อในกระบวนการธุรกิจ การปรับปรุงนี้ช่วยให้องค์กรเข้ากับสภาวะการเปลี่ยนแปลงและสามารถตอบสนองต่อความต้องการของตลาด

สรุปเนื้อหาบทที่ 6 "การประยุกต์ใช้สถาปัตยกรรมองค์กรเพื่อพัฒนาระบบข้อมูลขนาดใหญ่" นำเสนอเนื้อหาเกี่ยวกับการนำสถาปัตยกรรมองค์กรมาประยุกต์ใช้ในการพัฒนาระบบข้อมูลขนาดใหญ่ ประโยชน์ของสถาปัตยกรรมองค์กรสำหรับระบบข้อมูลขนาดใหญ่ โดยเน้นการใช้สถาปัตยกรรมที่เป็น

เฟรมเวิร์กที่ช่วยในการออกแบบและพัฒนาระบบให้สามารถทำงานได้อย่างมีประสิทธิภาพ ประกอบด้วยหัวข้อ

1. **สถาปัตยกรรมองค์กร** นำเสนอแนวคิดเกี่ยวกับสถาปัตยกรรมองค์กรซึ่งเป็นโครงสร้างและกรอบงานในการออกแบบและพัฒนาระบบในองค์กร สถาปัตยกรรมช่วยกำหนดรูปแบบของข้อมูล กระบวนการทำงาน และแนวทางการพัฒนา
2. **การเลือกสถาปัตยกรรม** อธิบายขั้นตอนในการเลือกสถาปัตยกรรมที่เหมาะสมสำหรับระบบข้อมูลขนาดใหญ่ ซึ่งรวมถึงการประเมินความเหมาะสมของสถาปัตยกรรมเทคโนโลยีต่างๆ ตามความต้องการของโครงการ
3. **การออกแบบและพัฒนาระบบข้อมูล** อธิบายกระบวนการในการออกแบบและพัฒนาระบบข้อมูลขนาดใหญ่ ซึ่งรวมถึงการแบ่งโมดูล การกำหนดความสัมพันธ์ระหว่างโมดูล การติดต่อสื่อสาร และการจัดการกับปัญหาที่เกิดขึ้นระหว่างการพัฒนา
4. **การจัดการเรื่องปัญหาและการป้องกัน** นำเสนอแนวทางในการจัดการกับปัญหาที่อาจเกิดขึ้นระหว่างการพัฒนา ระบบข้อมูล รวมถึงการวางแผนการป้องกันและการจัดการเมื่อเกิดปัญหา
5. **การจัดการความเสี่ยงและความปลอดภัย** อธิบายการจัดการความเสี่ยงและการพิจารณาเรื่องความปลอดภัยในระบบข้อมูลขนาดใหญ่ เพื่อป้องกันปัญหาที่อาจเกิดขึ้น

โดยสรุปบทที่ 6 เป็นการนำเสนอแนวทางในการประยุกต์ใช้สถาปัตยกรรมองค์กรในการพัฒนาระบบข้อมูลขนาดใหญ่ เน้นการเลือกสถาปัตยกรรมที่เหมาะสม และกระบวนการที่ช่วยให้การพัฒนาระบบเป็นไปอย่างมีประสิทธิภาพและป้องกันปัญหาที่อาจเกิดขึ้นได้อย่างเหมาะสม

คำศัพท์ บทที่ 6

Analytics Layer	เลเยอร์การวิเคราะห์
Application	โปรแกรมประยุกต์
Application Architecture	สถาปัตยกรรมด้านโปรแกรมประยุกต์
Application Programming Interface	ส่วนต่อประสานโปรแกรมประยุกต์
Architecture of Big Data	สถาปัตยกรรมของข้อมูลขนาดใหญ่
Business Architecture	สถาปัตยกรรมด้านธุรกิจ

Business Process Management	กระบวนการทางธุรกิจ
Capture	การเก็บรวบรวมข้อมูล
Change Analysis	การวิเคราะห์การเปลี่ยนแปลง
Cloud System	ระบบคลาวด์
Compliance	ความสอดคล้องกับกฎระเบียบ
Consumption	การนำไปใช้
Data Aggregation Layer	เลเยอร์การรวมข้อมูล
Data Architecture	สถาปัตยกรรมด้านข้อมูล
Data Governance Layer	เลเยอร์การกำกับดูแลข้อมูล
Data ingestion	การนำเข้าข้อมูลดิบ
Data Lake	ทะเลสาบข้อมูล
Data Layer	เลเยอร์ข้อมูล
Data Mining	การทำเหมืองข้อมูล
Data Transformation	การแปลงข้อมูล
Data Warehouse	คลังข้อมูล
Data Warehouse	คลังข้อมูล
Data Wrangling	การแปลงข้อมูล
Decreasing Complexity	ลดความซับซ้อน
Digital Organization	องค์กรดิจิทัล
Enterprise Architecture	สถาปัตยกรรมองค์กร
Enterprise Governance	การกำกับดูแลการบริหารองค์กร
Extract-Transform-Load (ETL)	การนำเข้าข้อมูลเข้า
Graphical Formats	รูปแบบกราฟิก
Information Exploration Layer	เลเยอร์การสำรวจข้อมูล
Internet of Things (IoT)	ระบบไอโอที
Machine Learning	การเรียนรู้ของเครื่อง
Master Data Management: MDM	การจัดการข้อมูลหลัก
Non Relational Database	ฐานข้อมูลเชิงไม่สัมพันธ์
Primary Activities	กระบวนการหลัก
Raw data	ข้อมูลดิบ
Reducing Wasted Time and Cost Saving	ลดเวลาที่เสียไปและลดต้นทุน
Relational Database	ฐานข้อมูลเชิงสัมพันธ์
Risk Management	การบริหารจัดการความเสี่ยง
Searching and Processing	การค้นหาและประมวลผล
Security Architecture	สถาปัตยกรรมด้านการรักษาความปลอดภัยข้อมูล
Semi-Structured Data	ข้อมูลกึ่งมีโครงสร้าง

Social Media	โซเชียลมีเดีย
Source of Big Data	แหล่งที่มาของข้อมูลขนาดใหญ่
Stream Data Processing	การประมวลผลข้อมูลแบบต่อเนื่อง
Strong Security	ความปลอดภัยที่แข็งแกร่ง
Structured Data	ข้อมูลที่มีโครงสร้างชัดเจน
Supporting Activities	กระบวนการสนับสนุน
Supporting Application	โปรแกรมประยุกต์สำหรับสนับสนุน
System Monitoring and Management	การเฝ้าระวัง และบริหารจัดการระบบ
Technology Architecture	สถาปัตยกรรมด้านเทคโนโลยี
Transaction System	ระบบปฏิบัติการประจำวัน
Transformation	การแปลงรูปข้อมูล
Unstructured Data	ข้อมูลที่ไม่มีโครงสร้าง
Value	คุณค่า
Variability	ความแปรผันได้
Variety	ความหลากหลาย
Velocity	ความเร็ว
Veracity	ความถูกต้อง
Visual Data Analytics	การนำเสนอข้อมูลวิเคราะห์ด้วยภาพ
Visualization	สร้างภาพ
Volume	ปริมาณ

คำถามและแบบฝึกหัดท้ายบทที่ 6

1. จงอธิบายองค์ประกอบที่สำคัญของสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่
2. จงอธิบาย ประเภทของข้อมูลสารสนเทศสมัยใหม่ประกอบไปด้วยอะไรบ้าง
3. จงอธิบายคุณลักษณะ 7 ประการของข้อมูลสารสนเทศสมัยใหม่
4. จงอธิบายการทำ ETL (Extract-Transform-Load)
5. เลเยอร์การวิเคราะห์ข้อมูล (Analytics Layer) และ เลเยอร์การค้นหาคำถามจากข้อมูล (Information Exploration Layer) ทำหน้าที่อะไรบ้าง

บทที่ 7 กระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่

(The Process of Developing Modern Corporate Architecture and Information Systems)

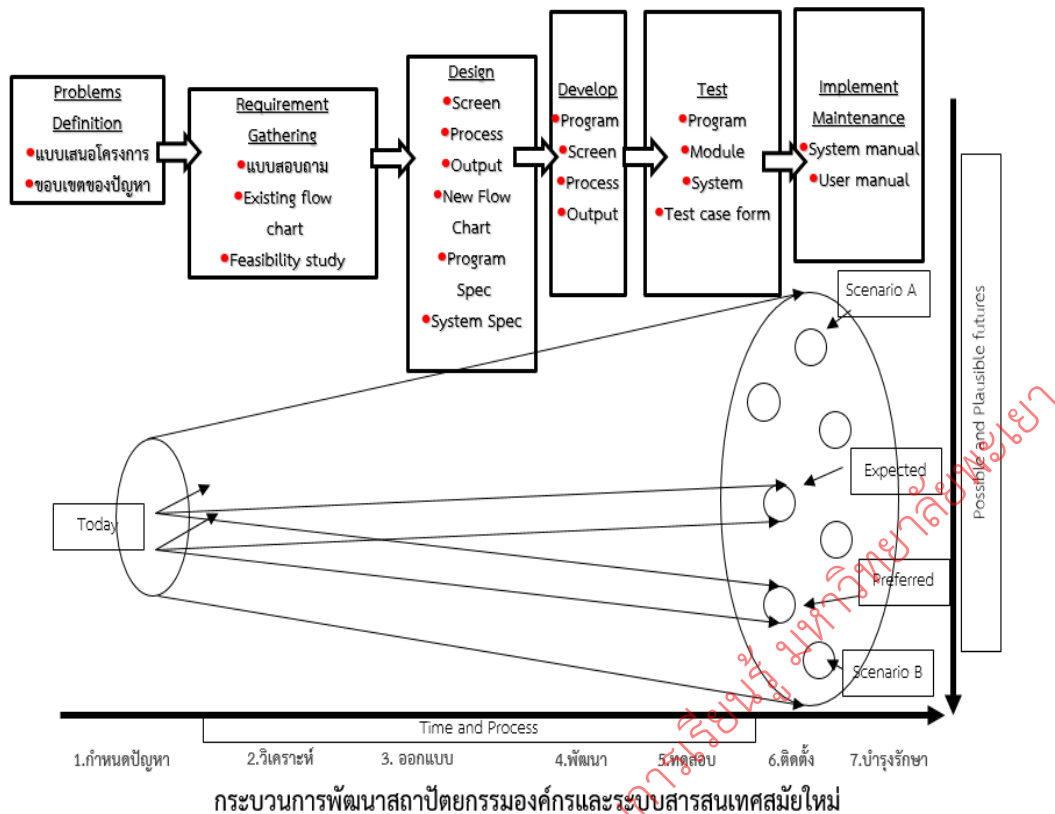
ปัจจุบันเทคโนโลยีสารสนเทศมีความสำคัญในการพัฒนา และการขับเคลื่อนธุรกิจสมัยใหม่ ความสำคัญของการมีสถาปัตยกรรมองค์กร (Enterprise Architecture) จะให้การวางแผนกำหนดกระบวนการ มีการเชื่อมโยงข้อมูลที่เป็นระบบมาตรฐาน และมีประสิทธิภาพ ส่งผลต่อการดำเนินธุรกิจองค์กรในยุคดิจิทัล และการบูรณาการเทคโนโลยีสารสนเทศเข้ากับวิสัยทัศน์และกลยุทธ์ทางธุรกิจอย่างเป็นระบบจึงมีความสำคัญอย่างมาก

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจกรรมวิธี กิจกรรมและกระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่
2. สามารถประยุกต์ใช้กระบวนการเพื่อออกแบบสถาปัตยกรรมองค์กรและสำหรับการพัฒนาระบบสารสนเทศสมัยใหม่ และเข้าใจบทบาทหน้าที่ของผู้ที่เกี่ยวข้อง

7.1 บทนำกระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่

กระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่คือกระบวนการที่จัดทำขึ้นมาเพื่อใช้เป็นแนวทางในการพัฒนาสถาปัตยกรรมองค์กรเพื่อให้ได้ผลลัพธ์ (Work Product) ตามเป้าหมายที่กำหนดขึ้น โดยกระบวนการนั้นจะต้องเลือกใช้แบบจำลองหรือวงจรการพัฒนาระบบงานให้เหมาะสมกับขนาดของโครงการ



รูปที่ 7.1 ความคาดหวังกับกระบวนการพัฒนาระบบสารสนเทศสมัยใหม่

จากรูป 7.1 คือแบบจำลองหรือวงจรการพัฒนาระบบงานตามหลักตามทฤษฎี SDLC (System Development Life Cycle) ซึ่งประกอบด้วย 7 ขั้นตอนคือ

ขั้นตอนที่ 1. กำหนดปัญหา จุดมุ่งหมาย และเป้าหมาย เป็นการวิเคราะห์ สถานการณ์ปัจจุบัน โดยพิจารณาปัจจัยภายในและปัจจัยภายนอกที่มีความสัมพันธ์กับการดำเนินกิจกรรมขององค์กร ทำความเข้าใจวิสัยทัศน์ (Vision) พันธกิจ (Mission) เช่นนโยบายและแผนงานที่เกี่ยวข้อง กฎระเบียบด้านความมั่นคง ปลอดภัย กระบวนการทำงาน โปรแกรมประยุกต์ (Application) ข้อมูลที่ใช้ในการดำเนินงานขององค์กร เป็นต้น

ขั้นตอนที่ 2. วิเคราะห์กำหนดความต้องการสารสนเทศของผู้ใช้และผู้เกี่ยวข้องทั้งหมด วิเคราะห์กระบวนการปฏิบัติงานเพื่อสนับสนุนพันธกิจหลัก (Supporting Process) ประกอบด้วย 2 กระบวนการย่อย ได้แก่ กระบวนการปฏิบัติการภายใน (Internal Process) และ กระบวนการปฏิบัติการภายนอก (External Process) หรือจะใช้ Value Chain หรือ ห่วงโซ่คุณค่า เป็นกรอบการวิเคราะห์เพื่อให้เห็นภาพรวมของกระบวนการในองค์กรที่เริ่มตั้งแต่การจัดหาวัตถุดิบไปจนถึงขั้นตอน

สุดท้ายที่ส่งมอบสินค้าให้แก่ลูกค้า โดยการวิเคราะห์ความต้องการต้องสำรวจให้ครบทั้ง 2 กิจกรรมหลัก คือ กระบวนการปฏิบัติการภายใน (Internal Process) ประกอบไปด้วย

- กิจกรรม Support Activity คือกิจกรรมที่สนับสนุนการดำเนินงานหลังบ้านของบริษัท (Back End Process) เช่น ฝ่ายบัญชี (Accounting) ฝ่ายการพัฒนาเทคโนโลยี (Technology Development) ฝ่ายทรัพยากรบุคคล (Human Resource Management) ฝ่ายการจัดซื้อ (Purchasing) ฝ่ายคลังสินค้า (Inventory) เป็นต้น
- กิจกรรม Primary Activity คือกิจกรรมหลักของบริษัท เช่น การนำวัตถุดิบเข้ามา (Inbound Logistics) ผลิต (Operation) เตรียมจัดจำหน่าย (Outbound Logistics) การตลาด (Marketing) การขาย และบริการหลังการขาย (Sale and Service) เป็นต้น

กระบวนการปฏิบัติการภายนอก (External Process) ตัวอย่างเช่น กระบวนการบริหารลูกค้าสัมพันธ์ (Customer Relationship Management Process) การเปลี่ยนแปลงของกฎหมายและระเบียบข้อบังคับทางธุรกิจ ระบบเศรษฐกิจ ระบบการจ้างงาน วิเคราะห์สภาพอุตสาหกรรม วิเคราะห์สภาพสังคม และปัจจัยทางเทคโนโลยี เป็นต้น

ขั้นตอนที่ 3. ออกแบบระบบจากการรวบรวมข้อมูลจากความต้องการ เป็นกระบวนการออกแบบระบบ ฟังก์ชันการใช้งานของระบบ จะมีการจัดทำเอกสารอธิบายอย่างละเอียด รวมถึงโครงร่างหน้าจอ (Screen Design) กฎเกณฑ์ทางธุรกิจ (Business Rules) ไดอะแกรมกระบวนการ (Logical Diagram) และเอกสารประกอบอื่นๆ ผลผลิตของเอกสารในขั้นตอนนี้จะอธิบายระบบใหม่แสดงถึงชุดของโมดูลหรือระบบย่อย (Modules or Subsystems) ขั้นตอนการออกแบบจะใช้เป็นเอกสารเริ่มต้นของการพัฒนาระบบสารสนเทศ โดยเอกสารที่ได้จะได้อาจมาจากการสัมภาษณ์ (Interview) เวิร์คช็อป (Workshop) และการสร้างต้นแบบ (Prototype) เพื่อใช้ในการเขียนโปรแกรมในขั้นตอนต่อไป โดยเอกสารในขั้นตอนนี้จะต้องทำให้ละเอียดที่สุด เพื่อจะได้ไม่ต้องเปลี่ยนแปลงมากในการตอนการเขียนโปรแกรม

ขั้นตอนที่ 4. พัฒนาซอฟต์แวร์และจัดทำเอกสารโดยจะทำงานร่วมกันระหว่างโปรแกรมเมอร์และนักวิเคราะห์ระบบ โดยขั้นตอนการพัฒนาคือส่วนที่นักพัฒนาเขียนโปรแกรม (Coding) และสร้างแอปพลิเคชันตามเอกสารการออกแบบก่อนและข้อกำหนดจากขั้นตอนที่ 3 โปรแกรมจะถูกสร้างขึ้นตามข้อกำหนดของเอกสารการออกแบบ ตามทฤษฎีแล้ว การวางแผนและโครงร่างก่อนหน้าทั้งหมดควรทำให้ขั้นตอนการพัฒนาจริงไม่มีการปรับแก้ไขมากนัก นักพัฒนาซอฟต์แวร์หรือโปรแกรมเมอร์จะเขียนโค้ดหรือโปรแกรม (Coding) ตามที่องค์กรกำหนดและใช้เครื่องมือต่างๆ เช่น Compilers Debuggers และ Interpreters เป็นต้น ภาษาที่ใช้เขียนโปรแกรมตัวอย่างเช่น เช่น C++, PHP JAVA

HTML และภาษาอื่นๆ นักพัฒนาจะเลือกเขียนโปรแกรมที่เหมาะสมเพื่อใช้ตามข้อกำหนดและข้อกำหนดของโครงการ

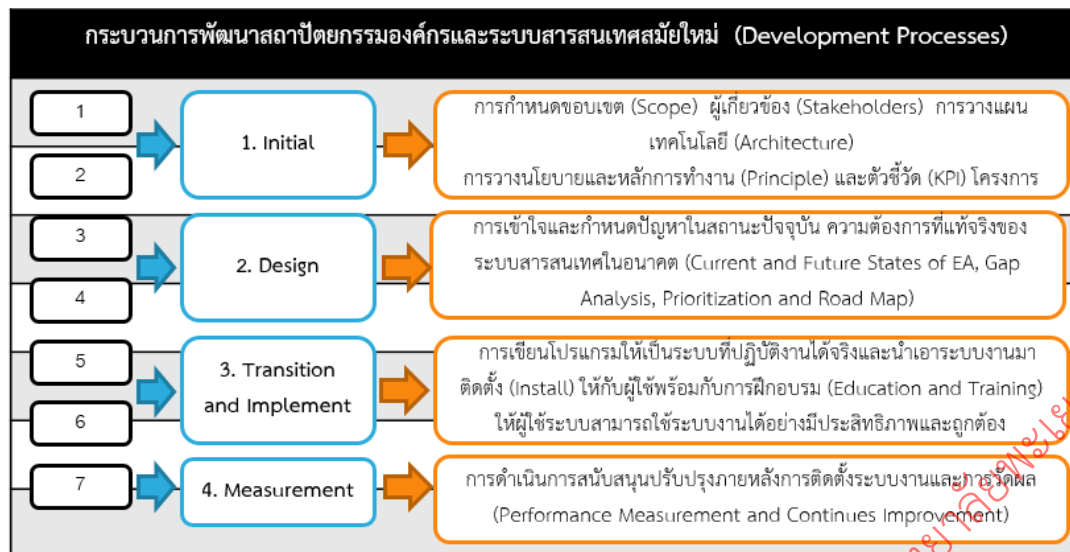
ขั้นตอนที่ 5. การเขียนโปรแกรมไม่ใช่จุดสิ้นสุด จะต้องมีการทดสอบโดยจะต้องใช้ข้อมูลจากการปฏิบัติงานจริงมาทดสอบ หรือจำลองสถานการณ์จริงเพื่อให้แน่ใจว่าไม่มีข้อบกพร่องใดๆ และผู้ใช้จะไม่ได้รับผลกระทบจากการผิดพลาดของโปรแกรม ในระหว่างขั้นตอนการทดสอบ นักพัฒนาจะตรวจสอบซอฟต์แวร์ของตนโดยละเอียด เพื่อสังเกตข้อบกพร่อง แก้ไข และทดสอบซ้ำแล้วซ้ำอีกจนแน่ใจว่าไม่มีช่องโหว่ของระบบ

ขั้นตอนที่ 6. หลังจากทดสอบแล้ว จะดำเนินงานติดตั้งระบบและประเมินผล จะต้องอบรมผู้ใช้งานก่อนใช้งานจริงและควรคำนึงถึงผลกระทบต่อผู้ใช้งานและองค์กรนั้น ระบบกับข้อมูลจะถูกติดตั้งให้เข้ากับสภาพแวดล้อมการทำงาน หลังจากผ่านขั้นตอนนี้แล้ว ระบบก็พร้อมสำหรับกิจกรรมทางธุรกิจขององค์กร

ขั้นตอนที่ 7. บำรุงรักษาและปรับปรุงระบบให้เหมาะสมกับกิจกรรมทางธุรกิจอย่างต่อเนื่อง ทีมนักพัฒนามีหน้าที่รับผิดชอบในการดำเนินการปรับปรุงซอฟต์แวร์หรือระบบที่อาจต้องการการปรับใช้ ซึ่งอาจรวมถึงการจัดการบั๊ก (Bugs) ที่ไม่สามารถแก้ไขก่อนการติดตั้งหรือปรับปรุงแก้ไข ปัญหาใหม่ที่เกิดขึ้นจากการรายงานของผู้ใช้ ระบบที่ใหญ่กว่าอาจต้องใช้เวลา และขั้นตอนการบำรุงรักษาที่นานกว่าเมื่อเทียบกับระบบที่เล็กกว่า

7.2 แนวทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่

โดยจากรูป 7.1 จะเห็นได้ว่า จุดเริ่มต้น (Today) และความคาดหวัง (Expected) ได้มีการตั้งเป้าหมายที่ชัดเจนไว้ แต่อย่างไรก็ตามการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่อาจได้รับผลกระทบจากปัจจัยหลายๆ อย่างทำให้ผลลัพธ์ (Work Product) ไม่ตรงกับความคาดหวัง (Expected) โดยโครงการอาจจะสำเร็จตามแผนภาพ A (Scenario A) หรือ แผนภาพ B (Scenario B) เนื่องจากปัจจัยด้าน เทคโนโลยี ที่เปลี่ยนแปลงไปอย่างรวดเร็ว การเปลี่ยนแปลงทีม เป็นต้น ดังนั้นในการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ จำเป็นต้องมีการวางแผนที่ดีและเข้าใจกระบวนการเพื่อให้ผู้เกี่ยวข้องทั้งหมดเห็นทางเดินและเป้าหมายไปด้วยกันดังที่ผู้เขียนจะได้เสนอแนะตามรูปที่ 7.2 ซึ่งเป็นการย่อ 7 ขั้นตอนให้เหลือเพียง 4 ขั้นตอน



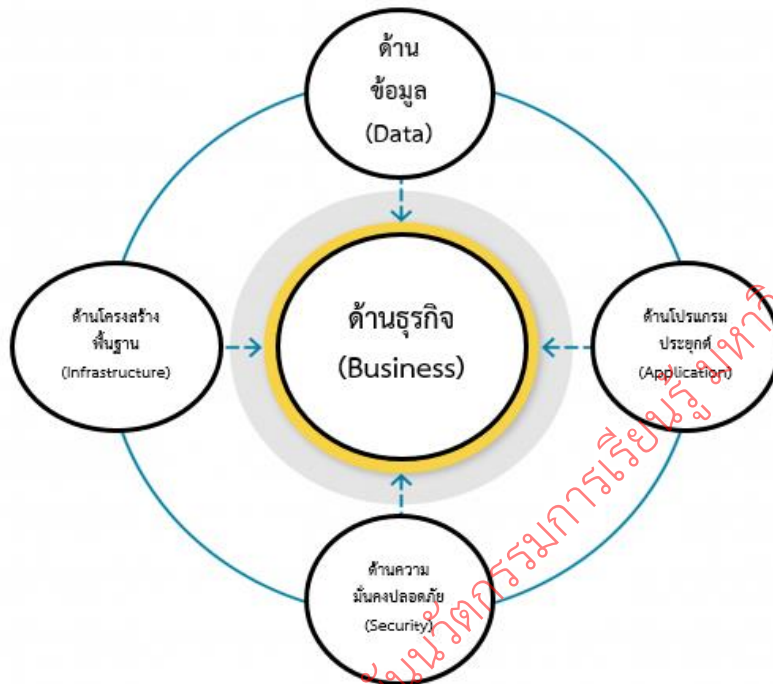
รูปที่ 7.2 แนวทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่

แต่อย่างไรก็ตาม ในปัจจุบันสถานการณ์ของธุรกิจเปลี่ยนแปลงไปอย่างรวดเร็ว และผู้บริโภคก็มีพฤติกรรมที่เปลี่ยนไปตามกาเข้ามาของเทคโนโลยี ทำให้ทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ ต้องแข่งขัน “เร็ว” และต้อง “สำเร็จ” ให้ทันต่อการเปลี่ยนแปลง ดังนั้นกระบวนการขั้นตอนที่มากไปอย่างวงจร SDLC อาจจะไม่เหมาะสมกับยุคปัจจุบัน จึงเป็นที่มาของแนวทางการทำงานแบบ **Agile** ที่มีความหมายว่า “ความคล่องตัว” ซึ่งจากการวิจัยและประสบการณ์ของผู้เขียนในการทำงานด้านการวิเคราะห์และออกแบบระบบมา จึงได้ย่อกระบวนการให้เหลือ 4 ขั้นตอนได้แก่

ขั้นตอนที่ 1 กระบวนการเตรียมการ (Initial) ประกอบไปด้วยกำหนดปัญหา จุดมุ่งหมาย และเป้าหมาย วิเคราะห์ความต้องการสารสนเทศของผู้ใช้ เพื่อกำหนดขอบเขตของสถาปัตยกรรมองค์กร (Scope) ผู้ที่เกี่ยวข้อง (Stakeholders) แนวทางสถาปัตยกรรม (Architecture Principle) ซึ่งขั้นตอนนี้จะทำให้เห็นโครงร่างของ สถาปัตยกรรมองค์กรในอนาคต (Future State Outline) และตัวชี้วัดผลสำเร็จขององค์กร (Key Performance Indicator)

โดยกระบวนการนี้คือ การค้นหาปัญหา โอกาสและเป้าหมาย (Identifying Problems, Opportunity and Objective) ระบบสารสนเทศจะเกิดขึ้นได้ก็ต่อเมื่อผู้บริหารหรือผู้ใช้ตระหนักว่าต้องการ ระบบสารสนเทศหรือต้องแก้ไขระบบเดิมโดย ทีมพัฒนาจะต้องศึกษาระบบโดยละเอียดเพื่อให้เข้าใจถึงปัญหาที่เกิดขึ้นในองค์กร มองเป้าหมายให้ชัดเจน เพื่อจะได้รู้ทิศทางของการทำระบบให้เป็นไปตามเป้าหมายที่วางไว้ เช่น ต้องการสร้างระบบเพื่อแข่งขันกับคู่แข่งในเรื่องการลดต้นทุนในการผลิตสินค้า หรือการสร้างฐานลูกค้าใหม่ เป็นต้น ทีมงานออกแบบสถาปัตยกรรมองค์กรและระบบ

สารสนเทศสมัยใหม่จะต้องทำการสำรวจและศึกษาข้อมูลให้ครบ 5 ด้านอันประกอบไปด้วย ด้านธุรกิจ (Business) ด้านแอปพลิเคชันหรือโปรแกรมประยุกต์ (Application) ด้านข้อมูล (Data) ด้านโครงสร้างพื้นฐาน (Infrastructure) และด้านความมั่นคงปลอดภัย (Security)

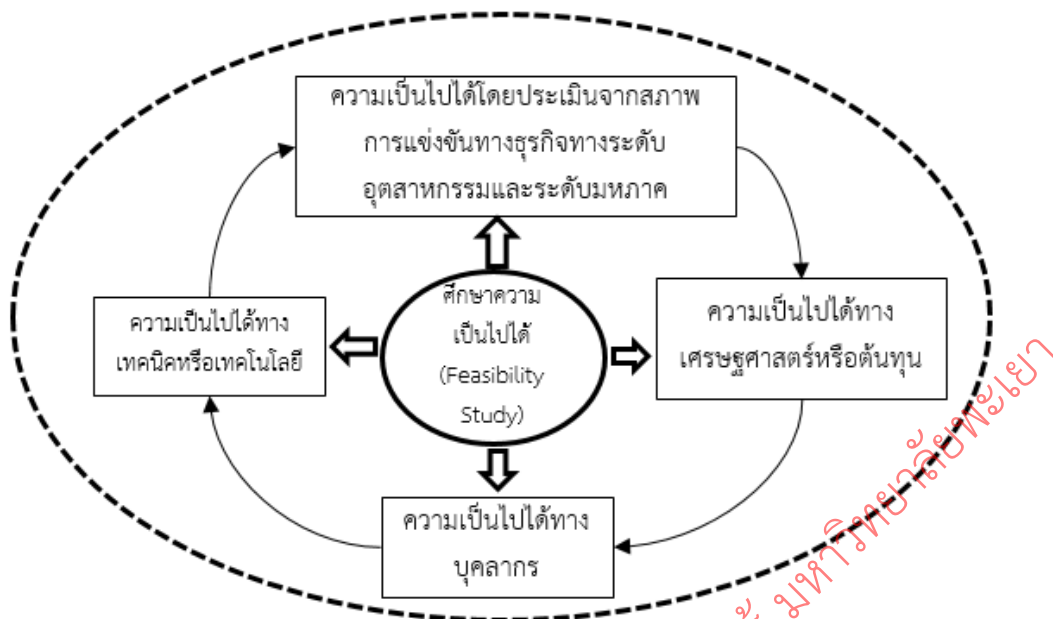


รูปที่ 7.3 มุมการวิเคราะห์ขอบเขตของสถาปัตยกรรมองค์กร (Scope)

ทีมงานจะต้องศึกษาความเป็นไปได้ (Feasibility Study) ดังรูป 7.4 กำหนดว่าปัญหาของธุรกิจคืออะไร และตัดสินใจว่าจะพัฒนาระบบสารสนเทศใหม่หรือการแก้ไขระบบสารสนเทศเดิม โดยเสียค่าใช้จ่ายและเวลาน้อยที่สุด โดยต้องคำนึงถึง

- ความเป็นไปได้ทางเทคนิคหรือเทคโนโลยี
- ความเป็นไปได้ทางบุคลากร
- ความเป็นไปได้ทางเศรษฐศาสตร์หรือต้นทุน
- ความเป็นไปได้โดยประเมินจากสภาพการแข่งขันทางธุรกิจทางระดับ

อุตสาหกรรมและระดับมหภาค



รูปที่ 7.4 มุมการศึกษาความเป็นไปได้ (Feasibility Study)

การศึกษาความเป็นไปได้นี้เป็นการประเมินของแผนการปฏิบัติการหรือของโครงการที่ได้วางแผนไว้ โดยการวิเคราะห์ปัจจัยความเป็นไปได้ทางเทคนิค เศรษฐกิจ กฎหมาย การดำเนินงาน บุคลากร ทีมงานและระยะเวลา เช่นเดียวกับกรที่เราวางแผนการเดินทางไปท่องเที่ยวข้ามทวีปว่า "เป็นไปได้ไหม" หรือการคิดตัวอย่างของคำถามเช่น เราสามารถในการสร้างเทคโนโลยีเพื่อทำสิ่งที่เสนอได้หรือไม่? องค์กรของเรามีบุคลากร เครื่องมือ และทรัพยากรที่จำเป็นหรือไม่? และโครงการนี้จะทำให้องค์กรได้รับผลตอบแทนการลงทุนหรือ ROI (Return of Investment) ที่คาดหวังหรือไม่

ผู้จัดการโครงการควรทำการศึกษาความเป็นไปได้เมื่อใด คำตอบคือควรทำในช่วงเวลาขั้นแรกของการศึกษากรณีทางธุรกิจว่าจะดำเนินโครงการพัฒนาระบบหรือไม่ หรือสถาปัตยกรรมองค์กรที่วางแผนไว้ควรดำเนินการต่อหรือไม่ ปัจจัยที่ควรศึกษาความเป็นไปได้อาจมีดังนี้

- ความเป็นไปได้ทางเทคนิคหรือเทคโนโลยี: ประกอบด้วยการพิจารณาว่าองค์กรมีทรัพยากรทางเทคนิค และความเชี่ยวชาญทางเทคโนโลยีที่ตรงตามข้อกำหนดของโครงการหรือไม่
- ความเป็นไปได้ทางเศรษฐกิจหรือต้นทุน: องค์กรจะต้องทำการประเมินปัจจัยทางเศรษฐกิจของโครงการ เพื่อกำหนดศักยภาพทางการเงินของโครงการ องค์กรสามารถใช้การวิเคราะห์ต้นทุนและผลประโยชน์เพื่อเปรียบเทียบต้นทุนทางการเงินกับผลประโยชน์ที่คาดการณ์ไว้
- ความเป็นไปได้โดยประเมินจากสภาพการแข่งขันทางธุรกิจทางระดับอุตสาหกรรมและระดับมหภาคและความเป็นไปได้ทางกฎหมาย: องค์กรจะต้องทำการประเมินการ

แข่งขันในอุตสาหกรรม เพื่อทราบถึงจุดแข็งและจุดอ่อนขององค์กรและคู่แข่ง จำนวนผู้แข่งขันในอุตสาหกรรม ผู้เข้าตลาดรายใหม่ หรือแม้กระทั่งการออกจากตลาดในอุตสาหกรรมสำหรับใช้ในการวางแผนเชิงกลยุทธ์ของสถาปัตยกรรมองค์กร รวมถึงศึกษาความเป็นไปตามข้อกำหนดทางกฎหมาย ซึ่งรวมถึงนโยบายรัฐบาล ข้อบังคับต่างๆที่ใช้กับกิจกรรมและงานทั้งหมดในขอบเขตโครงการ

- ความเป็นไปได้ทางบุคลากรและความเป็นไปได้ในการดำเนินงาน ความเป็นไปได้ในการดำเนินงานหมายถึงโครงการขององค์กรตรงกับกระบวนการวางแผนทุกระยะขององค์กร ทีมงานทรัพยากร เป้าหมายเชิงกลยุทธ์และวัตถุประสงค์ทางธุรกิจ รวมถึงความเป็นไปได้ด้านเวลา ประเมินการเวลาที่จะใช้ในการดำเนินการโครงการและกำหนดเส้นตาย (Deadline) จากนั้นลองคิดว่าไทม์ไลน์ (Time Line) โครงการขององค์กรเหมาะสมกับการดำเนินงานปัจจุบันอย่างไร

การประเมินความเป็นไปได้เป็นขั้นตอนที่สำคัญมาก องค์กรควรตรวจสอบและวิเคราะห์ข้อมูลทั้งหมด มุมในการวิเคราะห์ที่กล่าวมาทั้งหมดมีความสำคัญ เพื่อให้แน่ใจว่าทุกอย่างเป็นไปตามที่ควรเป็น และควรมีแผนบริหารจัดการความเสี่ยงและแผนฉุกเฉิน

ขั้นตอนที่ 2 กระบวนการออกแบบ (Design) บางทีก็เรียกว่าการออกแบบสเปคของระบบ (System Design Specification) หรือ การออกแบบภายในระบบ (System Internals Design) กระบวนการออกแบบจะประกอบไปด้วย การออกแบบระบบจากการรวบรวมข้อมูลจากความต้องการของผู้เกี่ยวข้องทั้งหมด โดยจะร่างแบบสถานะปัจจุบันและอนาคตของ สถาปัตยกรรมองค์กร (Current and Future States of Enterprise Architecture) เขียนรายงานวิเคราะห์ความแตกต่างของทั้งสองสถานะ (Gap Analysis) การจัดเรียงลำดับและเลือกโครงการที่ต้องการจะปรับเปลี่ยน (Prioritization) และจัดทำแผนการดำเนินงาน (Roadmap) ในขั้นตอนนี้ ทีมพัฒนาสถาปัตยกรรมองค์กรจะต้องดำเนินการทำกิจกรรมดังต่อไปนี้

- จัดทำแผนการดำเนินโครงการ (Project Plan)
- จัดทำแผนการติดตั้งระบบคอมพิวเตอร์ และระบบงานคอมพิวเตอร์
- จัดทำเอกสารการออกแบบระบบ (Prototype)
- ออกแบบระบบ (Designing the Systems) ให้สอดคล้องกับความต้องการของผู้ใช้ทั้งหมดและผู้บริหาร
- ออกแบบระบบโดยตัดสินใจเลือกฮาร์ดแวร์ (Hardware) และซอฟต์แวร์ (Software) ที่เหมาะสมกับสถานการณ์ทางธุรกิจรวมทั้งการขยายตัวในอนาคตทั้งฐานข้อมูลและรูปแบบของธุรกิจ

- ออกแบบฐานข้อมูล ออกแบบรายงานการแสดงผลบนหน้าจอ ออกแบบระบบโดยการกำหนดจำนวนบุคลากรผู้ใช้งานในระบบ รวมถึงสิทธิ์การใช้งานในระบบต่างๆ
- ออกแบบระบบรักษาความมั่นคงปลอดภัยของระบบสารสนเทศ
- ออกแบบการติดตั้งสายสัญญาณ การเชื่อมต่อระบบเครือข่ายเพื่อให้สามารถเชื่อมโยงระบบงานคอมพิวเตอร์กับระบบงานอื่นๆในองค์กร
- ออกแบบการติดตั้งอุปกรณ์และการเชื่อมต่อ (Shop drawing) ระบบคอมพิวเตอร์และระบบเครือข่าย

• ออกแบบการติดตั้งอุปกรณ์และการเชื่อมต่อ (As uilt) เป็นต้น
ซึ่งในกระบวนการนี้ทีมงานจะต้องสร้างแบบจำลองกราฟิกในการนำเสนออธิบายกระบวนการ (Process) ข้อมูล (Data) วัตถุ (Object) และตรรกะ (Logic) ที่เกิดขึ้นในสถาปัตยกรรมระบบ สำหรับแบบจำลองเชิงวัตถุจะมีเครื่องมือสำหรับการวิเคราะห์การออกแบบและการเขียนโปรแกรม เครื่องมือแบบจำลองจำแนก ตัวอย่างเช่น

- ผังงาน (Flowchart)
- แบบจำลองกระบวนการ ได้แก่ ดีเอฟดีหรือแผนภาพกระแสข้อมูล (Data Flow Diagram - DFD) และเครื่องมือที่นิยมใช้ควบคู่กับแผนภาพกระแสข้อมูล คือ พจนานุกรมข้อมูล (Data Dictionary) ภาษาโครงสร้าง (Structured Language) ตารางตัดสินใจ (Decision Table) และผังต้นไม้ (Decision tree) เป็นต้น
- แบบจำลองข้อมูล ได้แก่ อีอาร์ดี หรือแผนภาพความสัมพันธ์ระหว่างเอนทิตี (Entity-Relationship Diagram - ERD)
- Use Case Diagram
- Class Diagram
- Sequence Diagram
- Gantt Chart/Pert
- แบบจำลองเชิงวัตถุ ได้แก่ ยูเอ็มแอล (Unified modeling language)
- โมเดลวิเคราะห์ทางการเงิน เช่น NPV, ROI

สำหรับการใช้แบบจำลองเหล่านี้สามารถนำซอฟต์แวร์สำเร็จที่มีอยู่ค่อนข้างมาก เช่น เคส (Computer-Aided Software Engineering-CASE) มาใช้สนับสนุนการสร้างแบบจำลอง รวมถึงเครื่องมือสมัยใหม่ที่จะช่วยในการวิเคราะห์ทางการเงิน NPV, ROI

ก่อนจะเริ่มทำการพัฒนาระบบนั้น องค์กรอาจต้องกระทำบางขั้นตอนซ้ำแล้วซ้ำเล่าจนกว่าจะได้ข้อสรุปที่ถูกต้องตรงกัน ระหว่างผู้ใช้และนักวิเคราะห์ระบบ เมื่อได้ทำการศึกษาและมีความเข้าใจในขั้นตอนการวิเคราะห์และออกแบบระบบอย่างดีแล้ว องค์กรก็สามารถพัฒนาระบบใหม่ได้อย่างมีประสิทธิภาพและเป็นขั้นตอนรวมทั้งระบบที่ได้ก็มีความสมบูรณ์และง่ายต่อการบำรุงรักษาอีกด้วย

ขั้นตอนที่ 3 กระบวนการการปรับเปลี่ยน (Transition) และกระบวนการติดตั้ง (Implement) คือการพัฒนาซอฟต์แวร์และจัดทำเอกสาร โดยจะทำงานร่วมกันระหว่างโปรแกรมเมอร์และนักวิเคราะห์ระบบ ทีมงานจะจัดทำแผนการปรับเปลี่ยน (Transition Plan) และการกำกับดูแลการพัฒนาโครงการรวมทั้งนโยบายการใช้งาน ธรรมชาติขององค์กรในการใช้ระบบสารสนเทศ (Implementation Governance)

ในขั้นตอนนี้ทีมงานจะดำเนินการเขียนโปรแกรม จัดทำคู่มือการใช้โปรแกรม และฝึกอบรมผู้ใช้ที่เกี่ยวข้องในระบบ โดยนักวิเคราะห์และออกแบบ (System Analysts) จะวางแผนการเตรียมระบบ เตรียมสถานที่และเตรียมการติดตั้งเครื่องคอมพิวเตอร์ นอกจากนี้ ยังมีหน้าที่ในการดูแลการเขียนคู่มือการใช้โปรแกรมและการฝึกอบรม รวมถึงวางแผนการทดสอบและบำรุงรักษาระบบ (Testing and Maintaining the System) โดยจะต้องทำงานร่วมกับผู้ใช้ โดยตรวจสอบว่าโปรแกรมทำงานตรงตามที่ต้องการที่ได้ระบุไว้ตั้งแต่ขั้นตอนแรกหรือไม่ กรณีเกิดข้อผิดพลาดของโปรแกรมให้ปรับปรุงแก้ไขหรือเมื่อทดสอบโปรแกรมแล้ว โปรแกรมไม่เป็นไปตามความต้องการ อาจต้องแก้ไขปรับปรุงใหม่ นอกจากนี้ยังต้องวางแผนการบำรุงรักษา เพราะหากมีการใช้งานจริง จะต้องเจอสิ่งเหล่านี้คือ

- มีปัญหาในโปรแกรม (Bug)
- ธุรกิจเกิดการเปลี่ยนแปลงไป เช่น ธุรกิจขยายตัว ธุรกิจสร้างสินค้าตัวใหม่ ความต้องการของระบบก็เพิ่มขึ้น รายงานเพิ่มขึ้น
- การเปลี่ยนแปลงของกฎหมาย หรือระเบียบมาตรฐานของแต่ละธุรกิจ เช่น ธุรกิจประกันภัยหรือประกันชีวิต ที่มักจะมีแบบประกันใหม่ออกอยู่เสมอ และระเบียบจากกรมประกันภัยที่ออกมาบังคับใช้ เพื่อปกป้องผู้บริโภค ทำให้ระบบต่างๆ ในธุรกิจประกันภัยต้องมีการปรับปรุงระบบอยู่ตลอดเวลา เป็นต้น
- การเปลี่ยนแปลงทางฮาร์ดแวร์ (Hardware) และซอฟต์แวร์ (Software)
- ความต้องการผู้ใช้มีเพิ่มขึ้น หรือลูกค้ามีจำนวนมากขึ้น

ซึ่งองค์กรต่างๆ จำเป็นต้องวางแผนงบประมาณอีกประมาณ 20-50 % จากงบประมาณ
ค่าใช้จ่ายในการพัฒนาระบบ เพื่อสำรองไว้ใช้ในการบำรุงรักษาระบบ และการปรับปรุงอย่างต่อเนื่อง

ขั้นตอนที่ 4 การวัดประสิทธิภาพองค์กร (Measurement and Continuous Improvement) คือขั้นตอนการบำรุงรักษาและปรับปรุงระบบให้เหมาะสมกับกิจกรรมทางธุรกิจอย่างต่อเนื่อง (Continual Improvement) รวมถึงการวัดประสิทธิภาพองค์กร (Performance Measurement) ซึ่งจะต้องสอดคล้องกับตัวชี้วัดผลสำเร็จขององค์กรตามที่ระบุไว้ในกระบวนการแรก

องค์กรต้องวางแผนและดำเนินการในการเฝ้าติดตาม การตรวจวัด การวิเคราะห์ กระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่และการปรับปรุงที่จำเป็นเพื่อ

- แสดงให้เห็นถึงความสอดคล้องของสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่กับเป้าหมายขององค์กร
- สร้างความมั่นใจถึงความสอดคล้องกับวิสัยทัศน์ (Vision) พันธกิจ (Mission) และกิจกรรมขององค์กร การเชื่อมโยงระหว่างหน่วยงานต่างๆ ในการดำเนินธุรกิจ
- เพื่อให้มีการปรับปรุงประสิทธิภาพของระบบบริหารคุณภาพอย่างต่อเนื่อง

สิ่งที่ต้องติดตามและการตรวจวัด (Monitoring and Measurement) ในกระบวนการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ประกอบด้วย

- การตรวจวัดความพึงพอใจของผู้เกี่ยวข้องทั้งหมดรวมถึงลูกค้า (Stakeholders and Customer Satisfaction) โดยองค์กรต้องสร้างระบบเฝ้าติดตามข้อมูลที่เกี่ยวข้องกับการรับรู้ผู้เกี่ยวข้องทั้งหมดรวมถึงลูกค้า ว่าองค์กรได้ดำเนินการตามความต้องการที่เขียนไว้ในแผน
- การตรวจประเมินภายใน (Internal audit) องค์กรต้องตรวจประเมิน การดำเนินกิจกรรมว่าเป็นไปตามสถาปัตยกรรมองค์กรที่วางแผนไว้หรือไม่ ขั้นตอนการทำงาน การเชื่อมโยงของแต่ละระบบ รวมถึงข้อมูล ตรวจระบบรักษาความปลอดภัย ว่าได้ถูกนำไปดำเนินการและเก็บรักษาไว้อย่างมีประสิทธิภาพ

หลังจากการตรวจประเมินองค์กรต้องดำเนินการ ปรับปรุงระบบโดยลด ความไม่สอดคล้องตามแผนสถาปัตยกรรมองค์กร เพื่อ

- ทบทวนสิ่งที่ไม่เป็นไปตามแผนสถาปัตยกรรมองค์กร (รวมทั้งการร้องเรียนจากลูกค้า)
- การหาสาเหตุของสิ่งที่ไม่เป็นไปตามแผนสถาปัตยกรรมองค์กร

- แก้ไขสิ่งที่ไม่เป็นไปตามแผนสถาปัตยกรรมองค์กร
- ทบทวนสิ่งที่ได้แก้ไขที่ได้ดำเนินไปแล้ว

7.3 องค์ประกอบของทีมงานพัฒนาระบบในองค์กรปัจจุบัน



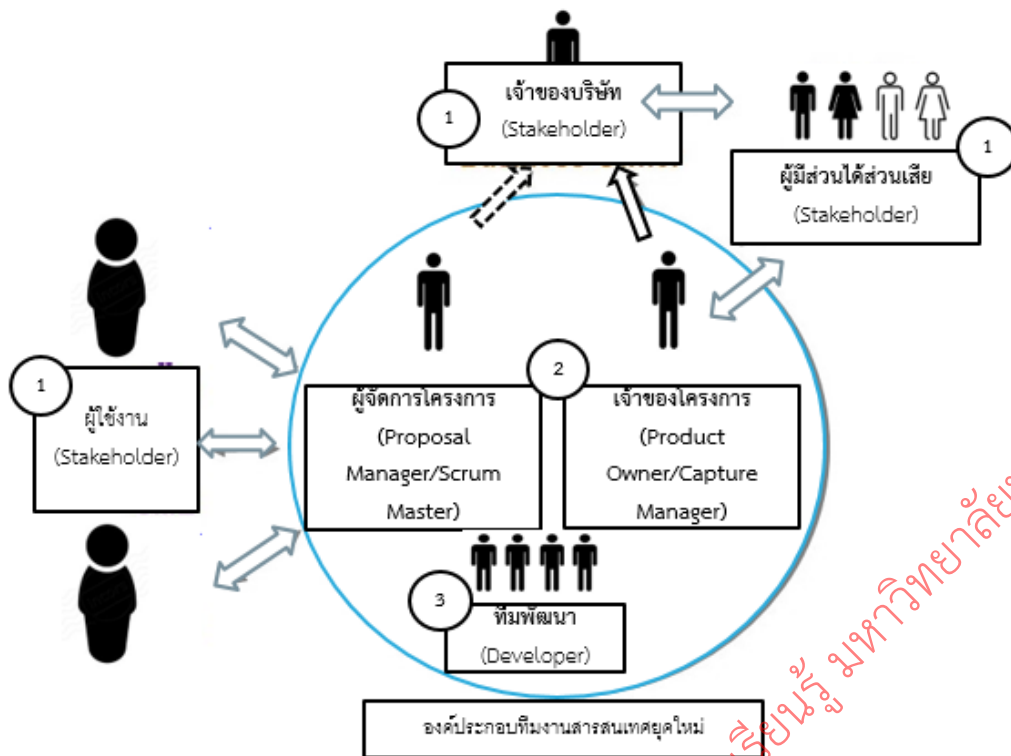
รูปที่ 7.5 ผังโครงสร้างองค์กรส่วนงานสารสนเทศยุคใหม่

องค์กรส่วนงานสารสนเทศยุคใหม่ (Modern Organization for Modern Information System) การจัดรูปแบบโครงสร้างเพื่อให้ทำงานร่วมกันอย่างเป็นระบบเพื่อให้ได้บรรลุเป้าหมายอย่างใดอย่างหนึ่ง โดยมีทิศทางการทำงานร่วมกัน คือ ประการที่หนึ่งมีวัตถุประสงค์หรือเป้าหมายร่วมจากพันธกิจขององค์กร ประการที่สองมีวัตถุประสงค์หรือเป้าหมายของตนเอง ประการที่สามมีคนร่วมกันทำงาน ประการที่สามมีการจัดโครงสร้างงานแบ่งงานหน้าที่รับผิดชอบของคนในองค์กรที่ทำงานส่งผลลัพท์ค่อนข้างเร็ว วัตถุประสงค์ง่าย และที่สำคัญจัดลำดับความสำคัญปรับทิศทางการได้รวดเร็ว

โครงสร้างตามรูป 7.5 จึงเป็นเพียงกรอบการแบ่งขอบเขตภาระงานและบทบาทที่ต้องรับผิดชอบ แต่ต้องปฏิบัติงานบนพื้นฐาน Agile ตามรูปที่ 7.6 คือ

- รูปแบบผังองค์กรที่เน้นการสื่อสารและปฏิสัมพันธ์กันระหว่างทีมงาน
- เน้นการพัฒนาระบบหรือแอปพลิเคชัน มากกว่าการทำเอกสาร
- เน้นตอบสนองผู้ใช้งาน สามารถปรับเปลี่ยนไปตาม และเน้นการปรับปรุงพัฒนา ตามความต้องการของธุรกิจได้ทันต่อการแข่งขันมากกว่าการยึดตามสัญญา หรือเอกสาร

ซึ่งองค์ประกอบของผู้ที่เกี่ยวข้องตามหลักการทำงานของ Agile คือ



รูปที่ 7.6 องค์ประกอบทีมงานสารสนเทศยุคใหม่

จากรูปที่ 7.6 บทบาทที่ตำแหน่งที่สำคัญในทีมงานสารสนเทศยุคใหม่

1. ผู้มีส่วนได้ส่วนเสีย (Stakeholder) ผู้ใช้งาน และเจ้าของบริษัท คือผู้ให้การสนับสนุนโครงการ มีเป้าหมายความสนใจในความสำเร็จของโครงการเป็นหลักมากกว่า จะตรวจสอบเรื่องกระบวนการพัฒนาระบบสารสนเทศ ผู้มีส่วนได้ส่วนเสีย (Stakeholder) เป็นผู้สนับสนุนทรัพยากร รวมทั้ง จัดสรรการเงินที่จำเป็นให้กับโครงการตามขอบเขตและผลลัพธ์ที่คาดหวัง
2. เจ้าของโครงการ (Product Owner: PO) ผู้จัดการโครงการ (Proposal Manager/Scrum Master) โดยทั่วไปแล้ว เจ้าของโครงการ (Product Owner/Capture Manager) จะมีบทบาทหลายอย่าง รวมถึงนักวางกลยุทธ์ทางธุรกิจ ผู้ออกแบบผลิตภัณฑ์ นักวิเคราะห์ตลาด ผู้ประสานงานกับลูกค้า และผู้จัดการโครงการ รวมถึงการกำหนดวิสัยทัศน์ วางแผนงาน จัดลำดับความสำคัญของความต้องการ ดูแลขั้นตอนการพัฒนา คาดการณ์ความต้องการของลูกค้า ทำหน้าที่เป็นผู้ประสานงานระหว่างทีมและผู้มีส่วนได้ส่วนเสีย และประเมินความคืบหน้าในการดำเนินโครงการ
3. ทีมพัฒนา (Developer: Dev) คือผู้พัฒนาระบบสารสนเทศตามที่ออกแบบไว้ให้เกิดขึ้นจริง หรือกลุ่มบุคคลที่ทำงานร่วมกันเพื่อส่งมอบระบบ ตามกำหนดเวลาและด้วยคุณภาพที่กำหนดไว้ บุคคลภายในทีมพัฒนา มักจะมีทักษะและความสนใจเฉพาะทาง อย่างไรก็ตาม ในการเพิ่มประสิทธิภาพการทำงาน คือการมีทีมที่มีทักษะที่สมดุลเพื่อรับมือกับความท้าทายที่เปลี่ยนแปลงตลอดเวลา

ซึ่งด้วยกระบวนการทำงานแบบ Agile การสร้างองค์ประกอบของทีม ความสำเร็จจะเกิดจากการพัฒนาแบบย้อนกลับ (Iterative Development) และการทำงานร่วมกัน บวกกับรับฟังข้อเสนอแนะของผู้มีส่วนได้ส่วนเสีย จะสนับสนุนกระบวนการพัฒนาข้อเสนอที่ประสบความสำเร็จ นอกจากนี้การส่งเสริมสภาพแวดล้อมการทำงานร่วมกันมากขึ้นในระหว่างการตรวจสอบ สามารถช่วยส่งเสริมขวัญกำลังใจในเชิงบวกมากขึ้นและเปลี่ยนการทบทวนหนึ่งในการบวนการทำงาน แทนที่นำโดยหัวหน้าโครงการคนเดียวและการตำหนิซึ่งกันและกัน การจัดเวลาสำหรับการประชุม การทบทวนอย่างสม่ำเสมอสามารถช่วยทีมในการเพิ่มประสิทธิภาพและประสิทธิผลด้วยการเรียนรู้และปรับปรุงอย่างต่อเนื่อง

7.4 บทบาทของนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile

นักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile เป็นบุคคลที่ผู้ควบคุมดูแลการพัฒนาระบบสารสนเทศโดยรวม โดยทำการศึกษาปัญหาซับซ้อนที่เกิดขึ้นในระบบและแยกแยะปัญหาเหล่านั้นอย่างมีหลักเกณฑ์ นักวิเคราะห์ระบบหรือที่เราเรียกกันว่า System Analyst หรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile จะต้องหาวิธีการมาแก้ไขปัญหานั้น พร้อมทั้งให้เหตุผลประกอบด้วย นักวิเคราะห์ระบบหรือผู้จัดการโครงการจะต้องกำหนดขอบเขตของการวิเคราะห์ และจะต้องกำหนดจุดมุ่งหมายหรือเป้าหมายในการวิเคราะห์หรือออกแบบระบบสารสนเทศนั้นด้วย นอกจากนี้และยังต้องทำความเข้าใจในโครงสร้างลักษณะขององค์กรนั้นๆ ในด้านต่างๆ เช่น การไหลของงาน อำนาจการบังคับบัญชา เป็นต้น

สิ่งที่ต้องศึกษาอีกสิ่งหนึ่ง คือ ลักษณะโครงสร้างของข้อมูล ที่มีอยู่ในการทำงานขององค์กร และที่สำคัญที่นักวิเคราะห์ระบบจะมองข้ามไปไม่ได้ นั่นก็คือ คนหรือบุคลากรที่ทำงานกับระบบสารสนเทศที่กำลังทำการวิเคราะห์ ต้องทำการศึกษาว่าคนเกี่ยวข้องกับระบบอย่างไร เกี่ยวตรงไหน ทำอะไร เพราะคนเป็นปัจจัยที่สำคัญที่สุด ถ้าขาดความร่วมมือจากบุคลากร ที่ทำงานอยู่ในระบบที่จะศึกษาก็ถือว่าล้มเหลวไปแล้วครั้งหนึ่งดังนั้นจะมองข้ามคนไปไม่ได้

นักวิเคราะห์ระบบหรือผู้จัดการโครงการควรมีคุณลักษณะดังต่อไปนี้

7.4.1 ผู้เชี่ยวชาญในทักษะทางเทคนิคและเทคโนโลยีที่จำเป็นสำหรับโครงการ ควรจะมีความรู้ทางด้านภาษาการเขียนโปรแกรมมาก่อน

7.4.2 เป็นนักสื่อสารที่ดีที่จะช่วยบริหารงานทีมให้ประสบความสำเร็จ ควรจะมีมนุษยสัมพันธ์ที่ดี เนื่องจากนักวิเคราะห์ระบบจะต้องเกี่ยวข้องกับคนในทุกระดับในองค์กร รวมถึงระบบปฏิบัติการ ช่างเทคนิค พนักงานบัญชี เลขานุการ พนักงานธุรการ

7.4.3 เป็นนักวางแผนที่ดีเพื่อให้งานพัฒนาสามารถดำเนินการได้ตรงเวลาในแต่ละช่วงของวงจรการพัฒนา

7.4.4 นักวิเคราะห์ผู้จัดการโครงการ ระบบเปรียบเหมือนผู้จัดการทั่วไป จะเป็นผู้ที่ตัดสินใจในการกำหนดออกแบบระบบทั้งหมด เป็นผู้ติดตามประเมินระบบที่ออกแบบ และติดตั้ง ว่าจะได้รับผลตามวัตถุประสงค์ที่วางไว้ตั้งแต่ต้นหรือเปล่า รวมทั้งการประเมินออกมาเป็นตัวเลข เพื่อชี้แจงให้ผู้จ้างออกแบบระบบเข้าใจ

ดังนั้น นักวิเคราะห์ระบบผู้จัดการโครงการ จึงควรมีทักษะด้านมนุษยสัมพันธ์ ความเชี่ยวชาญด้านเทคนิค ด้านการจัดการ และการวิเคราะห์ ความรับผิดชอบของพวกเขาค่อนข้างหลากหลายและมีความสำคัญต่อความสำเร็จในท้ายที่สุดของโครงการพัฒนาระบบสารสนเทศตามที่กำหนด นักวิเคราะห์ระบบมักจะถูกคาดหวังให้:

- ☐ รวบรวมข้อเท็จจริงและข้อมูลเพื่อใช้ในโครงการพัฒนาระบบสารสนเทศ
- ทำการตัดสินใจ จะจัดลำดับความสำคัญของโครงการ
- เสนอทางเลือกในการบริหารจัดการโครงการ
- วาดข้อกำหนดส่วนผู้ใช้งานเพื่อให้โปรแกรมเมอร์เข้าใจได้ง่าย
- สามารถประเมินและแก้ไขระบบผลลัพธ์ได้ตามเป้าหมายของโครงการ
- ช่วยวางแผนและกำหนดเป้าหมายของโครงการโดยละเอียด

สรุปเนื้อหาบทที่ 7 "กระบวนการพัฒนาระบบสารสนเทศสมัยใหม่" เป็นการอธิบายกระบวนการที่เกี่ยวข้องกับแนวทางการพัฒนาระบบสารสนเทศสมัยใหม่ ซึ่งเน้นการนำเสนอกระบวนการที่มีความเป็นระบบและเป็นขั้นตอนเพื่อให้การพัฒนาระบบเป็นไปอย่างมีประสิทธิภาพ รวมถึงองค์ประกอบของทีมงานพัฒนาระบบในองค์กรปัจจุบันได้แก่

1. **การวางแผนและการกำหนดเป้าหมาย** ขึ้นแรกเน้นในการวางแผนและการกำหนดเป้าหมายของกระบวนการพัฒนา ซึ่งรวมถึงการระบุความต้องการของระบบสารสนเทศและแนวทางการพัฒนาที่เหมาะสม
2. **การวิเคราะห์และออกแบบ** ขั้นตอนนี้เน้นการวิเคราะห์ความต้องการของระบบและการออกแบบโครงสร้างสถาปัตยกรรม รวมถึงการกำหนดขั้นตอนการทำงานและการจัดการข้อมูล

3. การปรับปรุงและพัฒนาต่อเนื่อง: ในยุคสมัยใหม่ เทคโนโลยีและความต้องการเปลี่ยนแปลงอย่างรวดเร็ว ขั้นตอนนี้เน้นการปรับปรุงและพัฒนาต่อเนื่องของระบบเพื่อให้ระบบสอดคล้องกับสภาพแวดล้อมที่เปลี่ยนแปลง
4. การวิจัยและนวัตกรรม: เน้นการวิจัยและนวัตกรรมเพื่อพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศให้เข้ากับเทคโนโลยีและแนวโน้มใหม่

คำศัพท์ บทที่ 7

Business Rules	กฎเกณฑ์ทางธุรกิจ
Current and Future States of Enterprise Architecture	สถานะปัจจุบันและอนาคตของ สถาปัตยกรรมองค์กร
Customer Relationship Management Process	กระบวนการบริหารลูกค้าสัมพันธ์
Deadline	กำหนดเส้นตาย
Developer	ทีมพัฒนา
External Process	กระบวนการปฏิบัติการภายนอก
Feasibility Study	ศึกษาความเป็นไปได้
Human Resource Management	ฝ่ายทรัพยากรบุคคล
Iterative Development	การพัฒนาแบบย้อนกลับ
Implementation Governance	ธรรมาภิบาลขององค์กรในการใช้ระบบสารสนเทศ
Internal Process	กระบวนการปฏิบัติการภายใน
Key Performance Indicator	ตัวชี้วัดผลสำเร็จขององค์กร
Modules or Subsystems	โมดูลหรือระบบย่อย
Performance Measurement	การวัดประสิทธิภาพ
Primary Activity	กิจกรรมหลักของบริษัท
Product Owner	เจ้าของโครงการ
Prototype	การสร้างต้นแบบ
Return of Investment : ROI	ผลตอบแทนการลงทุน
Screen Design	ออกแบบหน้าจอ
Stakeholder	ผู้มีส่วนได้ส่วนเสีย
Support Activity	กิจกรรมที่สนับสนุนการดำเนินงานหลังบ้านของบริษัท
Supporting Process	กระบวนการปฏิบัติงานเพื่อสนับสนุนพันธกิจหลัก
System Design Specification	การออกแบบสเปคของระบบ
System Internals Design	การออกแบบภายในระบบ

Technology Development	ฝ่ายการพัฒนาเทคโนโลยี
Transition Plan	แผนการปรับเปลี่ยน
Value Chain	ห่วงโซ่คุณค่า
Work Product	ได้ผลลัพธ์

คำถามและแบบฝึกหัดท้ายบทที่ 7

1. จงอธิบายกิจกรรม Support Activity
 2. จงอธิบายกิจกรรม Primary Activity
 3. จงอธิบายแนวทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ 4
- ขั้นตอน
4. จงอธิบาย ผู้มีส่วนได้ส่วนเสีย (Stakeholder) ได้แก่
 5. จงอธิบายบทบาทของเจ้าของโครงการ (Product Owner: PO)
 6. จงอธิบาย ทีมพัฒนา (Developer: Dev)
 7. จงอธิบายนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile
 8. จงอธิบาย คุณลักษณะนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master)
 9. จงอธิบายความคาดหวังต่อนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master)

บทที่ 8 ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูลในธุรกิจการค้าสมัยใหม่

(Introduction to Data Mining for Modern Trading Business)

การวิเคราะห์ข้อมูลสำหรับธุรกิจการค้าสมัยใหม่ ต้องอาศัยเทคนิคการวิเคราะห์ที่โดดเด่น ซับซ้อนกว่าการวิเคราะห์ทางสถิติแบบเดิม และการทำเหมืองข้อมูลมีบทบาทหลายด้าน เทคนิคการทำเหมืองข้อมูลแอบแฝงอยู่ในโปรแกรมประยุกต์ทางธุรกิจและบนแพลตฟอร์มของผู้ให้บริการบนอินเทอร์เน็ตมากมาย

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจนิยามและความหมายการทำเหมืองข้อมูล
2. สามารถเข้าใจกระบวนการทำเหมืองข้อมูลและเทคนิคต่างๆ เพื่อประยุกต์ใช้ในการออกแบบระบบสารสนเทศสมัยใหม่

8.1 ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูล

ในยุคที่เทคโนโลยีพัฒนาขึ้นอย่างรวดเร็ว ธุรกิจต้องวิ่งไล่ตามความต้องการของผู้บริโภค การรับรู้ข่าวสารที่รวดเร็วยิ่งขึ้น ผู้บริโภครุ่นใหม่เป็นกลุ่มที่เติบโตมากับความ ‘เรียลไทม์’ แคนั่งอยู่หน้าจอ ก็ได้ทุกอย่างแม้กระทั่งการสั่งอาหาร ข้อมูลผ่านโซเชียลออนไลน์ (Online Social Media) มีมากมาย มหาศาล ส่งผลให้พฤติกรรมของผู้บริโภคหันไปใช้จ่ายออนไลน์มากขึ้น ผ่านการตัดสินใจจากข้อมูลจากรีวิว และการนำเสนอผ่านระบบข้อมูลเชิงลึกของธุรกิจออนไลน์ เช่น YouTube Netflix SHOPEE LAZADA GRAP LINEMAN นั่นคือความสะดวกสบายของการช้อปปิ้งออนไลน์ ถัดมาคือการส่งของที่รวดเร็ว นี่เป็นอีกประเด็นที่น่าสนใจ เพราะปัจจุบันพบว่าธุรกิจเกี่ยวกับโลจิสติกส์ต่างๆ พัฒนาให้มีบริการที่หลากหลายมากขึ้น การทำเหมืองข้อมูล (Data Mining) เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บ และตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลแบบเดิมไม่ซับซ้อน และไม่มีข้อมูลแบบภาพ เสียง มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้ในการตัดสินใจหาข้อมูลเชิงลึก ที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล เช่น การนำข้อมูลขนาดใหญ่ มาทดลองประมวลผลและการสังเกตการณ์รูปแบบของพฤติกรรมผู้บริโภค โดยการสกัดหรือแยกข้อมูลที่เป็นประโยชน์จากข้อมูล การวางแผนทรัพยากรขององค์กรโดยสามารถวิเคราะห์ทางสถิติและตรรกะของข้อมูลขนาดใหญ่เป็นการมองหารูปแบบที่สามารถช่วยการตัดสินใจได้

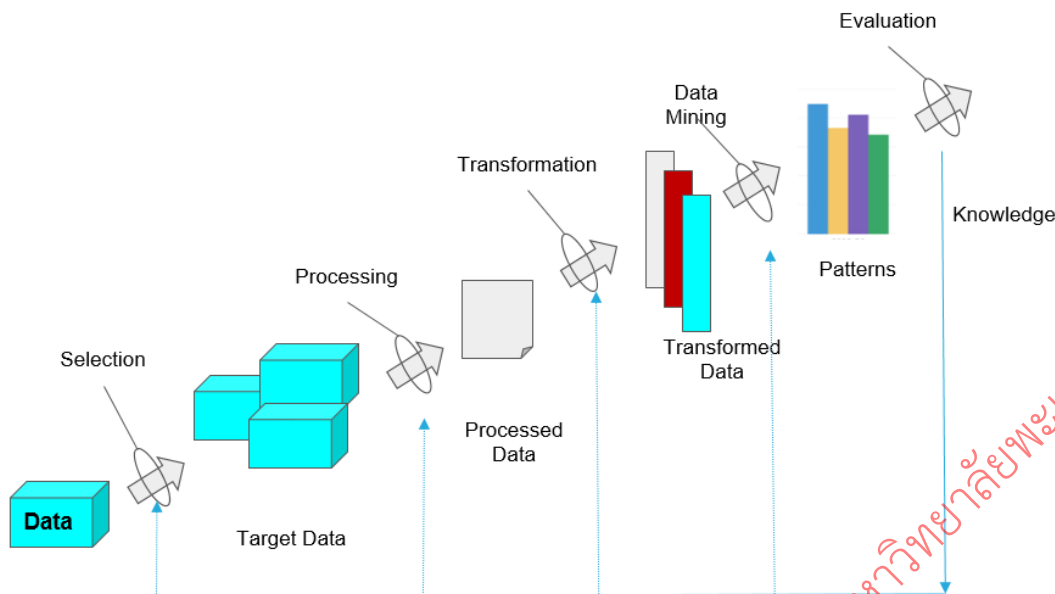
ดังนั้นประโยชน์ของการเรียนรู้ด้านการทำเหมืองข้อมูล (Data Mining) จะทำให้ ผู้ศึกษา สามารถหาแนวทางการตัดสินใจและคาดการณ์ผลลัพธ์ที่ได้แม่นยำความวิธีการทางสถิติเดิมๆ ซึ่ง

ตัวอย่างการนำเหมืองข้อมูลไปใช้งานในธุรกิจ เช่น การทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่ การทำนายยอดขายเมื่อมีการลดราคาสินค้า การทำนายกลุ่มลูกค้าที่นำจะใช้สินค้าของเรา

กล่าวโดยสรุปหากธุรกิจ สามารถรู้ล่วงหน้า และวินิจฉัยหน้าผู้บริโภค ด้วยการใช้ข้อมูลการสั่งซื้อ ประวัติลูกค้าทั้งเก่าและใหม่ เพื่อนำมาวิเคราะห์หาความน่าจะเป็น คาดคะเนจากเทคนิคเหมืองข้อมูล และแปลผลรูปแบบข้อมูลที่ได้จากการวิเคราะห์ เพื่อให้จับต้องได้ จะทำให้สามารถนำมาพัฒนากลยุทธ์ขององค์กร ซึ่งผู้ที่เข้าใจการทำเหมืองข้อมูล (Data Mining)

เช่น จะลูกค้ามากขึ้นทำให้คุณทราบว่าตอนนี้ลูกค้าต้องการอะไร แล้วลูกค้าประเภทไหนที่ซื้อสินค้าหรือใช้บริการ รวมถึงสามารถทำนายพฤติกรรมของผู้บริโภค และออกแบบสินค้าและบริการให้ตอบโจทย์กลุ่มลูกค้าประจำและกลุ่มลูกค้าใหม่ได้ และเห็นมุมมองการทำการตลาดที่กว้างขึ้น จากการมีข้อมูลในมือจำนวนมากและหลายรูปแบบ ทำให้มุมมองในการวางแผนการตลาดก็ยิ่งกว้างมากขึ้น

8.1.1 นิยามและความหมายการทำเหมืองข้อมูล การทำเหมืองข้อมูล (Data Mining) เป็นการค้นพบองค์ความรู้ใหม่จากฐานข้อมูล (Knowledge Discovery in Database: KDD) กล่าวคือ เป็นกระบวนการที่กระทำกับข้อมูลจำนวนที่มีขนาดใหญ่ เพื่อค้นหารูปแบบ (Patterns) และความสัมพันธ์ (Associations) ที่ซ่อนอยู่ในชุดข้อมูลขนาดใหญ่ (ศึกษาความหมายของข้อมูลขนาดใหญ่ได้ในบทที่ 6) รวมถึงการบูรณาการศาสตร์ต่างๆ เข้าด้วยกัน เช่น ฐานข้อมูล Database Systems Data Warehouses Online Analytical Processing (OLAP) เป็นต้น ซึ่งจัดเป็นเทคโนโลยีการจัดเก็บ รวบรวม และเตรียมข้อมูลที่ใช้ เพื่อทำเหมืองข้อมูล การเรียนรู้ของเครื่อง (Machine Learning) โดยใช้อัลกอริทึมหลักที่ใช้ในการค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูล หลักสถิติ (Statistic and Data Analysis Methods) สำหรับวิเคราะห์ข้อมูลเบื้องต้น ซึ่งอาจชี้ให้เห็นถึงรูปแบบและความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ได้ รวมถึงการประยุกต์ใช้วิทยาศาสตร์สารสนเทศ (Information Science) โปรแกรมทางคณิตศาสตร์ (Mathematical Programming) และการคำนวณประสิทธิภาพสูง (High Performance Computing) ทั้งนี้ จากข้อมูลที่มีขนาดใหญ่ จึงทำให้การทำเหมืองข้อมูลมักใช้เวลาานาน ดังนั้นจึงจำเป็นต้องมีการคำนวณที่รวดเร็วรองรับ รวมถึงการทำจินตทัศน์ (Visualization) เพื่อแสดงผลลัพธ์ รูปแบบ และความสัมพันธ์ของข้อมูลออกมาให้ผู้ใช้งานเข้าใจง่ายที่สุด ซึ่งเป็นประโยชน์ต่อการตีความและนำผลลัพธ์นั้นไปใช้ต่อไปทางธุรกิจ (Olson & Delen, 2008)



รูปที่ 8.1 ภาพแสดงกระบวนการทำเหมืองข้อมูล (Data Mining) (Olson & Delen, 2008)

8.1.2 กลยุทธ์หลักในการทำเหมืองข้อมูลและประเภทของแบบจำลอง

แนวคิดพื้นฐานของการทำเหมืองข้อมูลคือการค้นหาหรือเลือกสิ่งสำคัญจากตัวแปรหรือแอตทริบิวต์ และค้นหาความหมายของตัวแปร แต่สารสนเทศโดยทั่วไปจะมีบางอย่างซ่อนอยู่ ดังนั้นถ้าสามารถค้นหาบางอย่างที่ซ่อนอยู่จะถูกกระทำโดย การสร้างแบบจำลองการทำนายซึ่งประกอบด้วย 2 ประเภทใหญ่ๆ คือเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Machine Learning) และ แบบเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Machine Learning)

8.1.2.1 แบบจำลองเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Machine Learning) เป็นผลลัพธ์ที่สร้างจากการอนุมาน (Inference) ชุดข้อมูลปัจจุบัน เพื่อใช้ในการทำนายประเภทตัวอย่างในอนาคต การสร้างโมเดลแบบการเรียนรู้แบบมีผู้สอน (Supervised Machine Learning) จะต้องมีข้อมูลเป้าหมาย (Target) เพื่อทำนายข้อมูลที่เข้ามาใหม่ เช่น ถ้าต้องการทำนายลูกค้าว่าลูกค้านี้เป็นลูกค้าเครดิตที่ดี หรือไม่ดี จะต้องมีการระบุข้อมูลตัวอย่างไว้ว่าลูกค้าแบบไหนเป็นลูกค้าเครดิตที่ดี และแบบไหนเป็นลูกค้าเครดิตไม่ดี และต้องมีการแบ่งข้อมูลที่ใช้สำหรับสร้างโมเดลและทดสอบโมเดล จากนั้นจะต้องทำการประเมินโมเดลว่าโมเดลที่สร้างขึ้นมีประสิทธิภาพหรือไม่ก่อนที่จะนำโมเดลไปพัฒนาเป็นโปรแกรมเพื่อการประเมินลูกค้า

8.1.2.2 แบบจำลองแบบเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Machine Learning) เป็นการหาความสัมพันธ์ต่างๆ (Association) รวมถึงการจัดกลุ่มข้อมูล (Clustering) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย เช่น ต้องการจัดกลุ่มลูกค้าตามกลุ่มพฤติกรรมที่เหมือนกันอัตโนมัติ ซึ่งเทคนิคที่ใช้ในแบบจำลองนี้คือ “K-means Clustering” ซึ่งคือการเรียนรู้แบบไม่

ต้องสอน เป็นเทคนิคการแบ่งกลุ่ม แบบการจัดกลุ่ม (Clustering) โดยใช้พื้นฐานทางสถิติ มีตัวเลขประกอบ อย่างน้อย 2 ตัวแปรขึ้นไป เป็นต้น

ทั้งนี้ขั้นตอนการทำเหมืองข้อมูล (Data Mining) ที่นิยมใช้ คือกระบวนการทำเหมืองข้อมูล “CRISP-DM” ย่อมาจาก Cross-Industry Standard Process for Data Mining ซึ่งใช้ในแปลงปัญหาทางธุรกิจให้เป็นขั้นตอนต่างๆของการทำเหมืองข้อมูล (Data Mining) และแยกจากส่วนของโปรแกรมประยุกต์ (Application) และการใช้เทคโนโลยี โดยมี 6 ขั้นตอน ดังตารางที่ 8.1

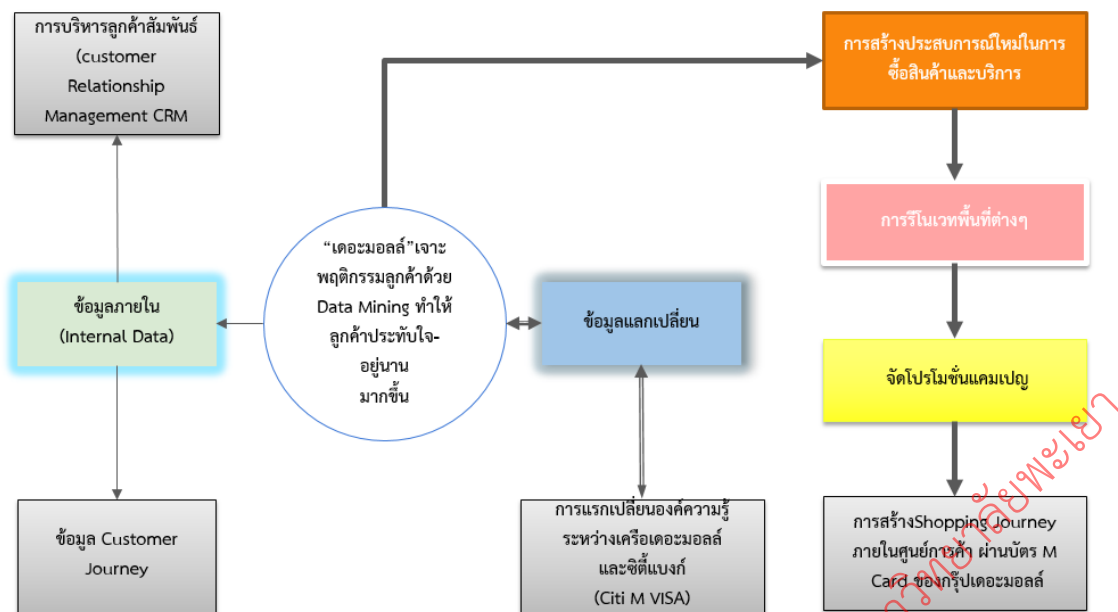
ตารางที่ 8.1 กระบวนการทำเหมืองข้อมูล “CRISP-DM” Cross-Industry Standard Process for Data Mining

ขั้นตอน	รายละเอียด
1. การทำความเข้าใจประเด็นทางธุรกิจ Business Understanding)	● การรวบรวมข้อมูลพื้นฐานของสภาพธุรกิจในปัจจุบัน ● การจัดหาจุดประสงค์ทางธุรกิจที่เฉพาะเจาะจง ● การหาเกณฑ์ที่ใช้ในการวัดความสำเร็จของการทำเหมืองข้อมูล จากมุมมองเชิงธุรกิจ
2. การทำความเข้าใจข้อมูล (Data Understanding)	● การจัดเก็บข้อมูลเบื้องต้นทั้งที่มาจากข้อมูลเดิม ชื่อข้อมูล สำนวน หรือเริ่มต้นตามข้อมูลเพิ่มเติม ● การอธิบายข้อมูลทั้งเชิงปริมาณและคุณภาพ ● การรายงานและสำรวจข้อมูลที่ประกอบด้วยตาราง กราฟ และเครื่องมือ Visualization ต่างๆ ● การตรวจสอบความถูกต้องของข้อมูลว่ามีข้อมูลที่ผิดพลาดรวมอยู่ด้วยหรือไม่
3. การเตรียมข้อมูล (Data Preparation)	● การรวบรวมข้อมูล และเลือกข้อมูลตัวอย่างที่จะนำไปใช้ ● การสร้างตัวแปรใหม่ที่ถูกคำนวณค่ามาจากตัวแปรเดิมผ่านการใช้อัลกอริทึม ● จัดเตรียมข้อมูลสำหรับการทำแบบจำลอง ● การแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของแบบจำลอง
4. การทำโมเดล (Modeling)	● การเรียนรู้แบบมีผู้สอน (Supervised Learning) (1) Classification เป็นการจัดกลุ่มข้อมูลด้วยตัวแปรที่กำหนดไว้แล้ว เพื่อสร้างแบบจำลองคาดการณ์ข้อมูลในอนาคต ได้แก่ เทคนิค Decision Tree Naïve Bayes K-Nearest Neighbors (K-NN) Neural Network เป็นต้น (2) Regression เป็นการประมาณค่าข้อมูล คล้ายกับ Classification แต่คำตอบจะเป็นตัวเลขเท่านั้น ● การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

	(1) เทคนิคการจัดกลุ่ม (Clustering) เป็นการจัดข้อมูลที่มีลักษณะคล้ายกันในกลุ่มเดียวกัน (2) เทคนิคการหาความสัมพันธ์ (Association) เป็นการค้นหาความสัมพันธ์ของข้อมูลในเหตุการณ์เดียวกัน
5. การประเมิน (Evaluation)	● การประเมินผลการวิเคราะห์กับเกณฑ์ที่ตั้งไว้ ● ผลที่ได้ ได้แก่ แบบจำลองที่คัดเลือกและข้อสรุปหรือสิ่งที่ได้จากแบบจำลอง
6. การนำไปใช้ (Deployment)	● การใช้องค์ความรู้ใหม่ที่ค้นพบมาพัฒนาองค์กร หรือการแสดงถึงการเปลี่ยนแปลงที่เกิดขึ้นในองค์กรได้ ● การวางแผนและควบคุมการนำผลไปใช้งาน รวมถึงเขียนสรุปและทบทวนโครงการ

8.1.3 กรณีศึกษาเกี่ยวกับการทำเหมืองข้อมูล (Data Mining)

การทำธุรกิจในยุคที่มีลูกค้าเป็นศูนย์กลาง จำเป็นต้องมีข้อมูลต่างๆ อย่างครบถ้วน เพื่อสามารถติดตาม วิเคราะห์ และคาดการณ์ความต้องการของลูกค้าได้อย่างแม่นยำ รวมทั้งสามารถขับเคลื่อนธุรกิจให้เติบโตได้ต่อเนื่อง แม้ปัจจัยภายนอกจะผันผวน หรือต้องเผชิญกับปัจจัยเสี่ยงอย่างรอบด้านก็ตาม จากการติดตามข้อมูลและการทำเหมืองข้อมูล (Data Mining) ภายในเครือข่ายมอลล์พบว่า ลูกค้าที่เข้ามาใช้บริการภายในห้าง ส่วนใหญ่จะมีพฤติกรรมในการเลือกซื้อสินค้าและบริการเดิม ไม่เปลี่ยนแปลงมากนัก เช่น การไปซื้อสินค้าร้านเดิม เลือกซื้อสินค้าแบรนด์เดิม โดยหากไม่มีสิ่งกระตุ้นเพื่อดึงดูดให้ลูกค้าเข้ามาใช้บริการทุกวันหรือเพิ่มความถี่ในการซื้อสินค้าและบริการเพิ่มขึ้น อาจทำให้ลูกค้าเข้ามาใช้บริการลดลง หรือใช้เวลาแต่ละครั้งน้อยลงได้ เนื่องจากปัจจุบันลูกค้ามีแนวโน้มเปลี่ยนแปลงตลอดเวลา และเบื่อง่ายมากยิ่งขึ้นกว่าเดิม จากปัจจัยที่สำคัญนี้ทำให้เครือข่ายมอลล์จำเป็นต้องมีความเข้าใจและมีข้อมูลต่างๆ ของลูกค้าอย่างรอบด้าน ผ่านการบริหารลูกค้าสัมพันธ์ (Customer Relationship Management: CRM) อย่างต่อเนื่อง โดยใช้กลยุทธ์การนำลูกค้าเป็นศูนย์กลาง (Customer Centric) เพื่อให้รู้จักลูกค้าแบบ 360 องศา และสามารถนำข้อมูลที่ได้ไปปรับใช้กับกลยุทธ์ของทางห้างให้เกิดประโยชน์สูงสุดต่อองค์กรเองดังรูป 8.2 (Brand Buffet, 2563)



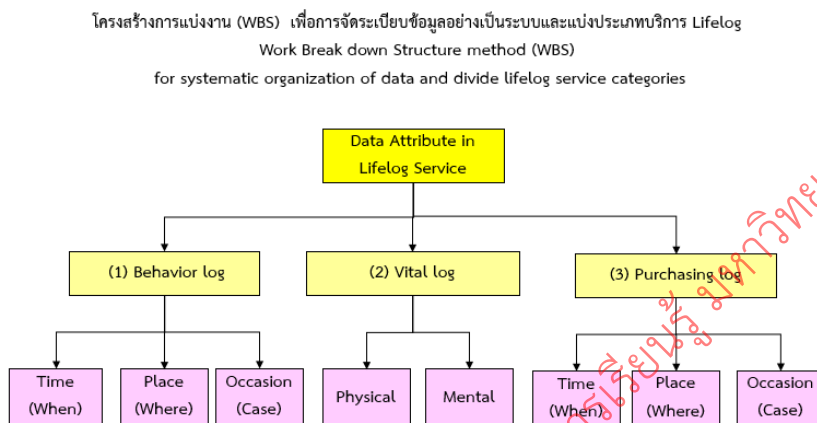
รูปที่ 8.2 กรณีศึกษา “เดอะมอลล์” จะใช้พฤติกรรมลูกค้าด้วย Data Mining ทำให้ลูกค้าประทับใจ อยู่ นานมากขึ้น

จากรูป 8.2 “เดอะมอลล์” นำข้อมูลภายในจากระบบการบริหารลูกค้าสัมพันธ์และข้อมูลภายใน (Internal Data) ซึ่งเป็นลูกค้าผู้ถือบัตร “M Card” ของเดอะมอลล์ มีจำนวน 3.5 ล้านราย มีอัตรา Active 60% และนำข้อมูลแลกเปลี่ยนจากพันธมิตรอย่างธนาคารซีทีแบงก์มาทำเหมืองข้อมูล โดยการสร้างแบบจำลองเพื่อหากลยุทธ์ในการการสร้างประสบการณ์ใหม่ในการซื้อสินค้าและบริการ จัดโปรโมชั่นแคมเปญ และ สร้าง Shopping Journey ภายในศูนย์การค้า ผ่านบัตร “M Card” ของกลุ่มเดอะมอลล์

นอกจากนี้จากการทำวิจัยของผู้เขียน เกี่ยวกับการจัดเก็บข้อมูลจาก นาฬิกาหรืออุปกรณ์พกพา ในลักษณะของข้อมูลชีวิตประจำวัน “Lifelogs” (Chertchom et al., 2017) ซึ่งคือการเก็บข้อมูลส่วนตัวในการใช้ชีวิตประจำวัน บริษัทผู้ให้บริการต่างๆ ก็ได้ออกแบบฐานข้อมูลซึ่งสามารถจัดเก็บตามรูป 8.3 เพื่อบันทึก/บันทึกเหตุการณ์ในชีวิตของตนผ่านทางคอมพิวเตอร์ส่วนบุคคลและอุปกรณ์เคลื่อนที่ แอปพลิเคชันและอุปกรณ์รุ่นใหม่ได้รับการพัฒนาสำหรับบุคคลในการเขียนอัตชีวประวัติดิจิทัลเพื่อระลึกถึงเหตุการณ์ ทบทวนความคืบหน้า และปรับปรุงตนเองเพื่อชีวิตที่ดีขึ้น โดยปกติ เทคโนโลยีเหล่านี้จะรวมอยู่ในอุปกรณ์สวมใส่อัจฉริยะสำหรับการบันทึกข้อมูลชีวิตประจำวัน “Lifelog” เช่น สถิติการดูแลสุขภาพส่วนบุคคลหรือสุขภาพ The 37° Bracelet,, Fitbit One™, SAGA, Google Glass, Google Fit และ tagCAM เป็นตัวอย่างของอุปกรณ์/แอปที่สวมใส่ได้อัจฉริยะสำหรับการจัดการสุขภาพ เทคโนโลยีบันทึกชีวิต อุปกรณ์เหล่านี้ได้รับการพัฒนาขึ้น เพื่อช่วยเหลือบุคคลอย่างต่อเนื่องในการบันทึกกิจกรรมประจำวันโดยอัตโนมัติตลอดจนในการจับภาพช่วงเวลาที่สำคัญหรือน่าทึ่งในชีวิต และยัง

สามารถแบ่งปันรายละเอียดของบันทึกชีวิตของพวกเขาให้กับเพื่อน ครอบครัว และผู้ติดตามในรูปแบบเปิดและปิด แต่อีกด้านหนึ่งคือ ผู้ให้บริการได้จัดเก็บข้อมูลไว้บนระบบคลาวด์ (Cloud System) ทั้งนี้เพื่อใช้ในการทำเหมืองข้อมูล (Data Mining) และนำผลมาใช้ในการปรับปรุงผลิตภัณฑ์ รวมทั้งเข้าใจพฤติกรรมของผู้บริโภคแบบเรียลไทม์

ลำดับชั้นข้อมูล (Data hierarchy of LIFELOG DATA PORTFOLIO :LLADP)



รูปที่ 8.3 กรณีสึกษาการออกแบบการจัดเก็บข้อมูลเพื่อการวิเคราะห์พฤติกรรมผู้บริโภคบนนาฬิกาอัจฉริยะ (Wearable Smart Watch) โดยผู้เขียน

การเก็บข้อมูลชีวิตประจำวัน (Lifelog) ดังรูป 8.3 จำนวนมากในแบบเรียลไทม์ (Real time) สามารถนำไปใช้ทำโมเดลการจำแนกประวัติสุขภาพ โดยวิธีการทางเหมืองข้อมูล ใช้เป็นข้อมูลเพื่อสำรวจและนำเสนอแนวทางปฏิบัติด้านการดูแลสุขภาพที่ดีขึ้น หรือเพื่อใช้จำแนกโรคเรื้อรัง จากข้อมูลบันทึกชีวิตจึงเป็นส่วนสำคัญของการสร้างแบบจำลองในการป้องกัน และช่วยในการควบคุมด้านสาธารณสุขได้

8.2 นิยามและความหมายนักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)

นักวิเคราะห์ข้อมูล (Data Analyst) หมายถึง บุคคลที่ใช้ข้อมูลในการวิเคราะห์โอกาสทางธุรกิจที่สอดคล้องกับแนวโน้มในการสร้างผลประกอบการตามค่านิยมของตลาดในปัจจุบัน หรือแก้ปัญหาจากสิ่งที่ผิดแปลกไปจากแนวโน้มเดิม โดยใช้หลัก “สถิติ” ทางคณิตศาสตร์หรือหลักเศรษฐศาสตร์มาประกอบแนวคิดในการวิเคราะห์ข้อมูลเชิงการสร้างมูลค่าทางธุรกิจในด้านต่างๆ ตามวัตถุประสงค์ที่ตั้งไว้ ทั้งนี้การวิเคราะห์ต่างๆ ต้องอาศัยประสบการณ์ องค์กรความรู้และมุมมองที่เฉียบขาด ส่วนนักวิทยาศาสตร์ข้อมูล (Data Scientist) หมายถึง บุคคลที่นำข้อมูลจากหลายๆ แหล่งองค์กรความรู้ นำมาผ่านวิธีการต่างๆ เช่นการทำเหมืองข้อมูล (Data Mining) การเรียนรู้ของเครื่อง (Machine Learning)

การเพิ่มประสิทธิภาพ (Optimization) เพื่อวิเคราะห์หาโอกาสความน่าจะเป็นภายใต้แนวคิดหรือมุมมองที่ทันสมัยใหม่ ซึ่งมุ่งเน้นผลลัพธ์ของคำตอบใหม่ๆ ที่ใช้ในการนำไปวิเคราะห์และสร้างสูตรทางการตลาดและการเติบโตในธุรกิจที่มุ่งหวัง (Apipoj, 2563)

8.2.1 เปรียบเทียบระหว่างนักวิเคราะห์ข้อมูล นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)

งานของนักวิทยาศาสตร์ข้อมูล (Data Scientist) ส่วนใหญ่คือ การสร้างแบบจำลอง คิดสมการใหม่ๆ เพื่อส่งต่อให้โครงการด้านข้อมูล (Data Project) ซึ่งเป็นงานประเภทโครงการ (Project Base) จากนั้นนักวิเคราะห์ข้อมูล (Data Analyst) จะเข้ามารับช่วงต่อ โดยเป็นผู้ที่ใช้ข้อมูล และวิธีการต่างๆ เพื่อวิเคราะห์สิ่งที่เกิดขึ้นด้วยตัวเลข เพื่อดูผลกระทบที่มีต่อธุรกิจต่อไป ดังนั้น แม้ว่า นักวิเคราะห์ข้อมูล จะไม่สามารถสร้างโมเดลได้อย่างที่นักวิทยาศาสตร์ข้อมูล ทำ แต่นักวิเคราะห์ข้อมูล สามารถเป็นได้ทั้งผู้ให้ข้อกำหนดความต้องการ (Requirement) และผู้ใช้งาน (User) รวมถึงอาจเป็นผู้จัดการโครงการ (Project Manager) อีกด้วย

ทั้งนี้นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist) มีความแตกต่างของในแง่ของลักษณะในการวิเคราะห์ ความลึกในการวิเคราะห์ข้อมูลอย่างเช่น นักวิเคราะห์ข้อมูล (Data Analyst) เน้นการวิเคราะห์เชิงธุรกิจมากกว่า มีเครื่องมือไว้ใช้จัดทำกระดานนำเสนอข้อมูลหรือ Dashboard เพื่อหาข้อมูล หรือ ประเมินโครงการ หรือ การพยากรณ์ข้อมูลในอนาคต ในส่วนของ นักวิทยาศาสตร์ข้อมูล (Data Scientist)จะเน้นงานที่มีความซับซ้อน เช่น การทำระบบแนะนำ (Recommendation System) หรือ การทำการเรียนรู้ของเครื่อง (Machine learning) เพื่อหาข้อมูลใหม่และความรู้ใหม่ เพื่อการใช้งานของข้อมูลต่างๆ ในเชิงการวิเคราะห์ โดยทั้ง นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist) มีการใช้ทักษะต่างๆในการวิเคราะห์ที่คล้ายคลึงกัน แต่แตกต่างกันเพียงแค่เทคนิคการวิเคราะห์และการนำข้อมูลมาใช้งานทั้งด้านการใช้เทคโนโลยีคอมพิวเตอร์ การใช้ทักษะทางคณิตศาสตร์และวิทยาศาสตร์ รวมไปถึงการนำไปใช้เชิงธุรกิจอีกด้วย ทักษะที่ใช้ในการวิเคราะห์ของนักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist) ซึ่งสามารถเปรียบเทียบได้ ดังตารางต่อไปนี้ (Wardah, 2016)

ตารางที่ 8.2 การเปรียบเทียบความแตกต่างระหว่าง นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)

ความแตกต่าง	นักวิเคราะห์ข้อมูล (Data Analyst)	นักวิทยาศาสตร์ข้อมูล (Data Scientist)
ความหมาย	ผู้ที่ทำหน้าที่รวบรวมข้อมูลจากหลายแหล่งข้อมูล เพื่อนำมาวิเคราะห์ เป็น report และ visualization	ผู้ที่ไม่ได้เพียงแค่ทำหน้าที่รวบรวม และแปลผลข้อมูลเท่านั้น แต่ยังทำหน้าที่ดึงข้อมูลเชิงลึกที่สามารถนำมาประยุกต์ใช้ได้ หรือการใช้ข้อมูลเพื่อ business transformation
ข้อมูลที่ใช้	เน้นที่การเปลี่ยนแปลง และ การแปลผลของข้อมูล ซึ่งมักจะมุ่งไปที่ข้อมูลในปัจจุบันและอดีต	เน้นที่การสรุปข้อมูล และ คาดการณ์ข้อมูลในอนาคตบนพื้นฐานของข้อมูลในอดีตและปัจจุบัน
วิธีการเพื่อตอบคำถามทางธุรกิจ	เขียน Custom Query เช่น SQL เพื่อตอบคำถามที่ซับซ้อนในเชิงธุรกิจ ยังไม่ถึงขั้นใช้ในการตัดสินใจทางธุรกิจได้	สร้าง Statistical Model ที่ใช้ในการตัดสินใจบนพื้นฐานของข้อมูล ซึ่งเป็นไปได้ทั้งการตัดสินใจขั้นสูง เช่น การ Block Page และ ขั้นพื้นฐาน เช่น การให้คะแนนความเสี่ยงกับ Webpage
รูปแบบการวิจัย	เข้าใจและต่อยอดจากส่วนที่เคยเข้าใจไม่เพียงพอเกี่ยวกับธุรกิจและผลิตภัณฑ์	สร้างการทดสอบสมมติฐานที่เป็นเหตุเป็นผลกัน (Causality Experiments) ที่จะพยายามหาต้นเหตุของกรณีศึกษา
ผลลัพธ์	ระบุได้เพียงประเด็นของคุณภาพข้อมูล เช่น Data Gap หรือ Bias ในการรวบรวมข้อมูล จะทำงานร่วมกับ Engineering ส่วนที่เหลือ เพื่อมองหาการรวบรวมข้อมูลแบบใหม่ๆ	ระบุถึงผลิตภัณฑ์ หรือ ลักษณะของผลิตภัณฑ์ใหม่ๆ ที่มาจากการถอดมูลค่าของข้อมูล เป็นผู้นำทางความคิดเกี่ยวกับมูลค่าของข้อมูล เช่น ระบบ Recommender System ของ Amazon ที่เคยทำออกมาครั้งแรกให้กับผู้ใช้งานทั่วไป

ทักษะด้านเทคนิค	Excel Tableau SQL SAS SPSS	Python R Java Julia Matlab SQL Hive Hadoop Tableau Spark
ทักษะด้านอื่น	ความรู้ทางสถิติ ความรู้ทางธุรกิจ และ ทักษะการสื่อสาร	ความอยากรู้ทางปัญญา ความรู้ทางธุรกิจ และ ทักษะการสื่อสาร

8.3 กรณีศึกษาเกี่ยวกับนักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)

ตัวอย่างการทำงานของ นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist) เพื่อจัดการและวิเคราะห์ข้อมูลด้านการบริหารจัดการน้ำในเขื่อน ซึ่งประกอบด้วย ข้อมูลอุทกนิยามวิทยา ข้อมูลน้ำท่า ข้อมูลประเพณีน้ำ ข้อมูลการบริหารจัดการอ่างเก็บน้ำ แบบจำลองการปฏิบัติการอ่างเก็บน้ำ และเงื่อนไขในการจัดการอ่างเก็บน้ำ

ด้านนักวิเคราะห์ข้อมูล (Data Analyst) เป็นผู้ใช้ข้อมูลในการวิเคราะห์แนวโน้ม หรือ แก้ปัญหาจากสิ่งที่ผิดแปลกไปจากแนวโน้มเดิม โดยใช้ หลัก สถิติ ทั้งนี้การวิเคราะห์ต่างๆ ต้องอาศัย ประสบการณ์ และ มุมมองที่ชัดเจน สะท้อนข้อมูลเป็นการนำเสนอข้อมูลเสมือน (Data Visualization) เพื่อสามารถประยุกต์ใช้ในเชิงธุรกิจ นำไปสู่การแก้ปัญหาที่ตรงจุด รวมถึงการติดตามสถานการณ์น้ำ เพื่อควบคุมและเฝ้าระวังในการบริหารการจัดการน้ำในเขื่อน ดังนั้น นักวิเคราะห์ข้อมูล นักวิเคราะห์ข้อมูล (Data Analyst) จึงจำเป็นต้องมีความเชี่ยวชาญเฉพาะด้าน เช่น การวิเคราะห์ด้านธุรกิจ การตลาด การขนส่ง การผลิต เป็นต้น เพื่อให้คำแนะนำจากประสบการณ์กับนักวิทยาศาสตร์ข้อมูล (Data Scientist) และ การวิเคราะห์ธุรกิจอย่างต่อเนื่อง

ด้านนักวิทยาศาสตร์ข้อมูล (Data Scientist) เป็นผู้นำข้อมูลจากหลายๆ แหล่ง มาผ่านวิธีการต่างๆ เช่นการทำเหมืองข้อมูล (Data Mining) การเรียนรู้ของเครื่อง (Machine Learning) การเพิ่มประสิทธิภาพ (Optimization) เพื่อหามุมมอง และคำตอบใหม่ๆ โดยการออกแบบแบบจำลองข้อมูล ผ่านกระบวนการเรียนรู้ของเครื่อง (Machine Learning) หรือเขียนโปรแกรมขึ้นมาเพื่อทำความเข้าใจกับข้อมูล ทั้งนี้ นักวิทยาศาสตร์ข้อมูล (Data Scientist) ต้องสามารถนำข้อมูลจากหลายๆ แหล่ง มาประกอบในแบบจำลอง เพื่อหาคำตอบผ่านแบบจำลองที่สร้างได้ หลังจากการออกแบบแบบจำลองสำเร็จ และแบบจำลองมีความเสถียรแล้ว นักวิทยาศาสตร์ข้อมูล (Data Scientist) อาจหมอบบทบาทตรงส่วนนั้นไป ในการทำงานต้องอาศัยคำแนะนำจากประสบการณ์ของนักวิเคราะห์ข้อมูล

นอกจากนี้แล้วจากงานวิจัยของผู้เขียน ใน เรื่อง “Output Based Upon Input, the Result from Data Analytics and Visualization of TNI Registration System’s Data” (Chertchom P., 2017) ได้พูดถึงนักวิเคราะห์ข้อมูล (Data Analyst) ไว้ว่า การวิเคราะห์ข้อมูลเป็น

เครื่องมือที่ช่วยให้องค์กรเข้าใจลูกค้า กิจกรรมทางธุรกิจ เพื่อใช้ในการวางแผนการตลาด การวิเคราะห์ตนเอง และการคาดการณ์สิ่งที่อาจเกิดขึ้นในอนาคต การใช้การวิเคราะห์ข้อมูลยังใช้เพื่อจัดหมวดหมู่ โปรไฟล์และอธิบายข้อมูลประชากรของลูกค้าเพื่อทำความเข้าใจว่าเกิดอะไรขึ้นกับสถานการณ์ทางธุรกิจ ในปัจจุบัน การวิเคราะห์ข้อมูลและการแสดงภาพ (Visualization) ได้ถูกนำมาใช้กันอย่างแพร่หลายในการค้นหาความรู้ ซึ่งกระบวนการทำวิเคราะห์ข้อมูล

8.4 ความรู้เบื้องต้นเกี่ยวกับ Data Preprocessing

การเตรียมข้อมูล (Data Preparation หรือ Data Preprocessing) เป็นขั้นตอนในการเตรียมข้อมูลดิบ (Raw Data) ให้อยู่ในรูปแบบที่เหมาะสมกับการนำมาวิเคราะห์ข้อมูลต่อไป ซึ่งมีความสำคัญมากในการทำเหมืองข้อมูล (Data Mining) เนื่องจากบางครั้งข้อมูลดิบที่รวบรวมมา อาจจะเป็นข้อมูลที่ไม่สมบูรณ์, มีค่าผิดพลาดหรือไม่อยู่ในรูปแบบที่สอดคล้อง และเป็นไปในทิศทางเดียวกัน ทำให้การทำเหมืองข้อมูล เป็นไปได้ยาก หากไม่มีการเตรียมข้อมูลที่ดี ผลลัพธ์จากการทำเหมืองข้อมูลออกมาจะไม่ถูกต้อง และไม่น่าเชื่อถือ (DataPreparator, 2022)

การเตรียมข้อมูลที่มีประสิทธิภาพ ควรมีลักษณะสำคัญดังนี้

- ให้ผลลัพธ์ที่ครบถ้วนสมบูรณ์
- ให้ความสำคัญกับนิยามข้อมูล
- มีการจัดบันทึกขั้นตอนการเตรียมข้อมูลโดยละเอียด
- ปรับกระบวนการให้เป็นอัตโนมัติให้มากที่สุด

การเตรียมข้อมูลประกอบด้วยวิธีการหลักๆ 4 ขั้นตอน (Techopedia, 2022)

8.4.1 การทำความสะอาดข้อมูล (Data cleaning) เป็นขั้นตอนในการแยกข้อมูลที่รบกวนออกไป เช่น การเติมข้อมูลที่ขาดหายไป (Missing Values) การปรับเปลี่ยนข้อมูลที่มีผิดพลาด (Noisy Data) การแยกข้อมูลที่มีค่าผิดปกติออกมา (Outliers) และการแก้ไขข้อมูลที่ไม่สอดคล้องกัน (Inconsistent Data)

8.4.2 การรวมข้อมูล (Data Integration) เป็นขั้นตอนที่รวมแหล่งข้อมูลหลายๆ แห่งเข้าด้วยกัน เพื่อไว้ในที่เดียวกัน เพื่อลดความซ้ำซ้อนของข้อมูล และเพิ่มความเร็วและคุณภาพในการทำเหมืองข้อมูล

8.4.3 การแปลงข้อมูล (Data Transformation) เป็นขั้นตอนการแปลงรูปข้อมูลเพื่อให้เหมาะสมกับการทำเหมืองข้อมูล โดย Normalization (การแปลงค่าข้อมูลให้อยู่ในช่วงสั้นๆ ที่นำมาใช้ประมวลผลได้) เป็นวิธีที่พบบ่อย เช่น Min-Max Decimal Scaling หรือ Z-Score เป็นต้น

8.4.3 การลดรูปข้อมูล (Data Reduction) คือการลดรูปข้อมูลให้มีรูปแบบกระชับรัดและกินเนื้อที่น้อยลง เพื่อที่จะนำไปวิเคราะห์แล้วเกิดผลลัพธ์ที่เท่ากันหรือเทียบเท่ากันเมื่อเทียบกับการวิเคราะห์จากข้อมูลทั้งหมด

8.4.3.1 การรวมข้อมูล (Data Aggregation หรือ Dimensionality Reduction) เป็นการลดรูปข้อมูลที่พบในดาต้าคิวบ์ (Data Cube) ซึ่งเป็นแบบจำลองข้อมูลหลายมิติ (Multidimensional Data Model)

8.4.3.2 Numerosity Reduction มี 2 ประเภท คือ

(1) วิธีพารามิเทริก (Parametric Methods) คือการสร้างแบบจำลองข้อมูลด้วยการประมาณค่าพารามิเตอร์ของแบบจำลอง จัดเก็บค่าพารามิเตอร์ เช่น การสร้างแบบจำลองข้อมูลด้วยเทคนิคทางสถิติความถดถอยในการประมาณค่าตัวแปรตาม

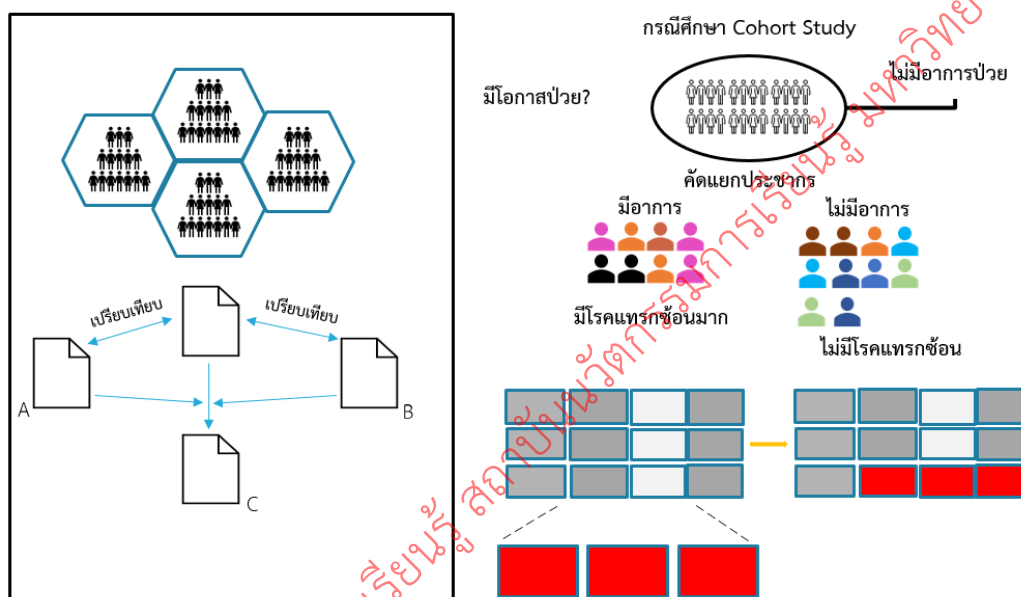
(2) วิธีการที่ไม่ใช่พารามิเตอร์ (Non-parametric Methods) ซึ่งเป็นวิธีที่ไม่ได้สร้างแบบจำลองข้อมูล แต่ใช้เครื่องมือ เช่น แผนภาพฮิสโตแกรม (Histogram Diagram) วิธีการจัดกลุ่ม (Clustering) แสดงการกระจายข้อมูล และเก็บค่าตัวแทนกลุ่มแทนค่าข้อมูลจริง, และการสุ่มตัวอย่าง (Sampling) เป็นต้น

8.4.3.3 การทำให้ข้อมูลเป็นแบบข้อมูลแบบไม่ต่อเนื่อง (Discretization Data) คือการลดรูปข้อมูลดิบโดยแบ่งข้อมูลดิบออกเป็นช่วงๆ และใช้ค่าช่วง (Interval Label) แทนค่าของแต่ละช่วง ในกรณีที่อัลกอริทึมรับเฉพาะข้อมูลแบบ การจัดลำดับชั้นแนวคิด (Categorical Concept Hierarchy Generation) แบ่งข้อมูลดิบเป็นแต่ละช่วงจากแนวคิดระดับล่างไประดับที่สูงขึ้น เช่น อายุ แบ่งเป็นกลุ่ม วัยเยาว์ วัยกลางคน วัยชรา

8.5 กรณีศึกษาเกี่ยวกับ Data Preprocessing

Big Data University (Cognitive Class, 2022) ผู้ดำเนินการสอนคลาสเรียนออนไลน์เกี่ยวกับข้อมูลขนาดใหญ่ (Big Data) จากไอบีเอ็ม (IBM) ได้อธิบายเกี่ยวกับกระบวนการทำการจัดเตรียมข้อมูล (Data Preparation) หรือการทำความสะอาดข้อมูล (Data Cleaning) ว่าเปรียบเสมือนการล้างผัก เพราะเป็นการกำจัดข้อมูลที่ใช้งานไม่ได้ออกจากข้อมูลที่เราจะนำมาใช้งานจริงๆ เพื่อป้องกันการผิดพลาดของผลลัพธ์ที่ต้องการ โดยยกตัวอย่างจากกรณีการศึกษาหาเกณฑ์การรับผู้ป่วยเข้าใหม่ (Readmission criteria) ของโรคหัวใจล้มเหลว (Congestive Heart Failure : CHF) โดยกำหนดให้

ติดตามผู้ป่วยที่กลับเข้ามารับการรักษ หรือโรคกำเริบอีกครั้งในระยะเวลา 30 วัน ซึ่งจะนำข้อมูลมาจากคลินิกแห่งหนึ่ง โดยใช้ข้อมูลประเภทต่างๆ มาจำแนกและแบ่งลงในคอลัมน์ (Transform the Data to 1 Table) เช่น อายุ เพศ น้ำหนัก ส่วนสูง และโรคแทรกซ้อนอื่นๆ เป็นต้น และรวบรวมเป็นข้อมูล 1 Record ต่อ 1 ผู้ป่วย (Aggregating Data) ทั้งนี้หลังจากที่ได้ข้อมูลแล้วจะทำการทบทวนวรรณกรรม (Literature Review) เพื่อเทียบเคียงว่าข้อมูลที่ได้มาเป็นอย่างไร เพื่อคัดเลือกคุณลักษณะ (Attributes) หรือปัจจัย (Factor) ที่มีผลต่อการทำนายความเสี่ยงของการเข้ารับรักษาใหม่ (Readmission) ในผู้ป่วยโรคหัวใจล้มเหลว “CHF” เมื่อได้ทำการเตรียมข้อมูลแล้วก็นำมาเทรนนิ่ง (Training) ทดสอบ (Test) และสรุปผลต่อไปดังรูปที่ 8.4



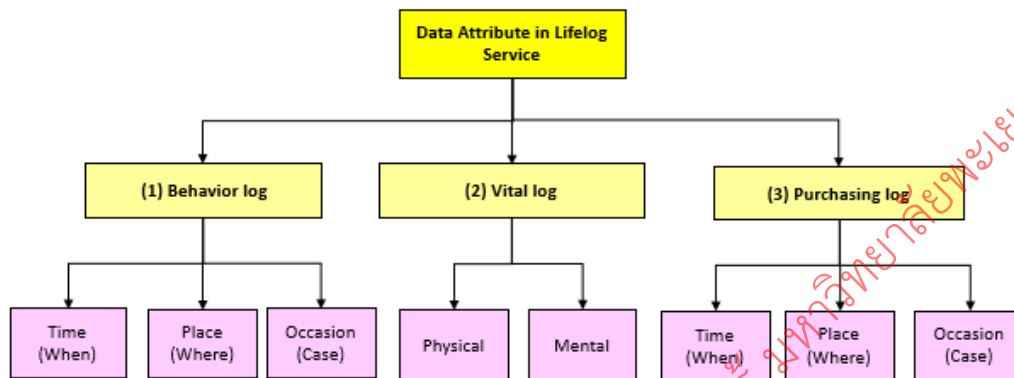
รูปที่ 8.4 กรณีศึกษากระบวนการจัดเตรียมข้อมูลผู้ป่วยโรคหัวใจล้มเหลว (Congestive Heart Failure : CHF)

จากกรณีศึกษา 2 จากการวิจัยของผู้เขียน (Chertchom, 2017) ในการ ศึกษาข้อมูลชีวิตประจำวัน “LIFELOG” ซึ่งเป็นการบันทึก ข้อมูลการใช้ชีวิตประจำวัน โดยใช้เทคโนโลยี ซึ่งส่วนใหญ่ ใช้เซ็นเซอร์ จากอุปกรณ์ที่สวมใส่ เช่น นาฬิกาอัจฉริยะ (Smart Watch) โดยกระบวนการการรวบรวมข้อมูลจาก เหตุการณ์ในชีวิตของผู้ใช้และแชร์กับเพื่อนที่กำหนด ซึ่งการบันทึกข้อมูล จะต้องเริ่มต้นจาก การรวบรวม วัตถุข้อมูลซึ่งแบ่งเป็นคุณลักษณะ (Attributes) ต่างๆ โดยชนิดของข้อมูลอาจแบ่งได้ดังนี้

1. ข้อมูลที่เป็นตัวเลข (Quantitative Attributes) คือข้อมูลที่ แบ่งเป็น Discrete Attributes เป็นข้อมูลที่ไม่ต่อเนื่อง ส่วน Continuous Attributes เป็นข้อมูลที่ต่อเนื่อง
2. ข้อมูลที่เป็นหมวดหรือกลุ่ม (Categorical Attributes) เช่น ข้อมูลกลุ่มพฤติกรรม “Behavior Log” ข้อมูลกลุ่มสุขภาพ “Vital Log” ข้อมูลกลุ่มการใช้จ่าย ดังรูป 8.5

3. ข้อมูลที่เป็นนามบัญญัติ (Nominal Categorical) คือข้อมูลที่มีมาตรวัดเป็นนามบัญญัติและจัดเป็นกลุ่ม เช่น ข้อมูลคุณลักษณะต่าง ๆ ข้อมูลเหตุการณ์ต่างๆ หรือข้อมูลสิ่งต่าง ๆ ยกตัวอย่างเช่น เพศ เชื้อชาติ อาชีพ ศาสนา เป็นต้น โดยข้อมูลในแต่ละกลุ่มเป็นอิสระต่อกัน

4. ข้อมูลที่เป็นลักษณะลำดับขั้นหรือชั้น (Ordinal categorical) คือข้อมูลที่ถูกกำหนดเป็นตัวเลขหรือสัญลักษณ์เพื่อชี้ถึงอันดับ โดยที่แต่ละกลุ่มไม่เป็นอิสระต่อกัน



รูปที่ 8.5 กรณศึกษากระบวนการจัดเตรียมข้อมูลชีวิตประจำวัน “LIFELOG”

โดยทั่วไปข้อมูลชีวิตประจำวัน “LIFELOG” คือข้อมูลของตนเองที่เกี่ยวกับสุขภาพ ข้อมูลการออกกำลังกาย และข้อมูล เนื่องจากคนทั่วไป ต้องการติดตามและปรับปรุงพฤติกรรมและความเป็นอยู่ที่ดีของตนเอง นอกจากนี้ ข้อมูลที่บันทึกชีวิตอาจได้มาจากข้อมูลการทำงาน มีทั้งข้อมูลกิจกรรมทางกาย และข้อมูลทางอารมณ์ เช่น อัตราการเต้นของหัวใจ ความดันโลหิตขณะเดินทาง อัตราการหายใจ ความเหนื่อยล้า คุณภาพการนอนหลับ การใช้ระบบขนส่งสาธารณะ ความสม่ำเสมอของการใช้ชีวิต เป็นต้น

สรุปเนื้อหาบทที่ 8 “ความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูลในธุรกิจการค้าสมัยใหม่” เป็นการนำเสนอความรู้พื้นฐานเกี่ยวกับกระบวนการทำเหมืองข้อมูลในภาคธุรกิจและการค้าในยุคสมัยใหม่ โดยเน้นการนำข้อมูลที่มีอยู่ในองค์กรมาวิเคราะห์และแปรผลเพื่อช่วยในการตัดสินใจและการวางแผน ประกอบไปด้วย

1. **การทำเหมืองข้อมูล (Data Mining)** เบื้องต้นอธิบายความหมายและหลักการของการทำเหมืองข้อมูล ซึ่งเป็นกระบวนการวิเคราะห์ข้อมูลเพื่อค้นหาความรู้ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในข้อมูล
2. **การเตรียมข้อมูล (Data Preparation)** อธิบายการกระบวนการเตรียมข้อมูลก่อนนำเข้าสู่กระบวนการทำเหมืองข้อมูล รวมถึงการทำความสะอาดข้อมูล เช่น การจัดระเบียบ การลบข้อมูลซ้ำ และการแปลงรูปแบบข้อมูล

3. **เทคนิคการทำเหมืองข้อมูล** นำเสนอเทคนิคที่ใช้ในการทำเหมืองข้อมูล เช่น การค้นหาความสัมพันธ์ (Association Rule Mining) การจัดกลุ่มข้อมูล (Clustering) และการจำแนกข้อมูล (Classification) เพื่อค้นหาแนวโน้มและข้อมูลที่มีความหมาย
4. **การประยุกต์ใช้การทำเหมืองข้อมูลในธุรกิจ** อธิบายการนำเสนอและการประยุกต์ใช้การทำเหมืองข้อมูลในธุรกิจและการค้า ซึ่งสามารถช่วยในการตัดสินใจในการวางแผนธุรกิจ การเข้าใจลูกค้า และการตลาดสินค้า
5. **การวิเคราะห์และการสร้างความรู้** อธิบายการวิเคราะห์ผลลัพธ์จากการทำเหมืองข้อมูล เพื่อสร้างความรู้และเข้าใจแนวโน้ม รูปแบบ และความสัมพันธ์ที่อาจเกิดขึ้น
6. **ความเป็นมาตรฐานและความเป็นส่วนตัว** อธิบายเรื่องความเป็นมาตรฐานในการใช้ข้อมูล และการทำเหมืองข้อมูล รวมถึงการคำนึงถึงความเป็นส่วนตัวของข้อมูลที่ใช้ในกระบวนการทำเหมือง
7. **การใช้เทคโนโลยีในการทำเหมืองข้อมูล** นำเสนอวิธีการใช้เทคโนโลยีต่างๆ เช่น การใช้เครื่องมือซอฟต์แวร์สำหรับการทำเหมืองข้อมูล และการนำเอาเทคโนโลยีใหม่ในการทำเหมือง

โดยสรุปบทที่ 8 เป็นการนำเสนอความรู้เบื้องต้นเกี่ยวกับการทำเหมืองข้อมูลในธุรกิจการค้าสมัยใหม่ ซึ่งเน้นการนำข้อมูลมาวิเคราะห์และประยุกต์เพื่อช่วยในการตัดสินใจและการวางแผนในการดำเนินธุรกิจ รวมถึงการศึกษาเกี่ยวกับนักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)

คำศัพท์ บทที่ 8

Association	การหาความสัมพันธ์
Business Understanding	การทำความเข้าใจประเด็นทางธุรกิจ
Categorical Attributes	ข้อมูลที่เป็นหมวดหรือกลุ่ม
Categorical Concept Hierarchy Generation	การจัดลำดับชั้นแนวคิด
Cloud System	ระบบคลาวด์
Clustering	การจัดกลุ่มข้อมูล
Clustering	การจัดกลุ่ม
Congestive Heart Failure : CHF	โรคหัวใจล้มเหลว
“CRISP-DM” Cross-Industry Standard Process for Data Mining	ขั้นตอนการทำเหมืองข้อมูล “CRISP-DM”
Customer Centric	ลูกค้าเป็นศูนย์กลาง
Data Aggregation	การรวมข้อมูล
Data Analyst	นักวิเคราะห์ข้อมูล
Data cleaning	การทำความสะอาดข้อมูล
Data Integration	การรวมข้อมูล
Data Mining	การทำเหมืองข้อมูล
Data Preparation	การเตรียมข้อมูล
Data Project	โครงการด้านข้อมูล
Data Reduction	การลดรูปข้อมูล
Data Scientist	นักวิทยาศาสตร์ข้อมูล
Data Transformation	การแปลงข้อมูล
Data Understanding	การทำความเข้าใจข้อมูล
Data Warehouse	ระบบคลังข้อมูล
Database System	ระบบฐานข้อมูล
Deployment	การนำไปใช้
Discretization	ข้อมูลแบบไม่ต่อเนื่อง
Evaluation	การประเมิน
High Performance Computing	การคำนวณประสิทธิภาพสูง
Histogram Diagram	แผนภาพฮิสโตแกรม
Inconsistent Data	ข้อมูลที่ไม่สอดคล้องกัน
Information Science	วิทยาศาสตร์สารสนเทศ
Interval Label	ค่าช่วง
Knowledge Discovery in Database: KDD	การค้นพบองค์ความรู้ใหม่จากฐานข้อมูล
Lifelogs	ข้อมูลชีวิตประจำวัน

Machine Learning	การเรียนรู้ของเครื่อง
Mathematical Programming	โปรแกรมทางคณิตศาสตร์
Missing Values	ข้อมูลที่ขาดหายไป
Modeling	การทำโมเดล
Multidimensional Data Model	แบบจำลองข้อมูลหลายมิติ
Noisy Data	ข้อมูลที่ผิดพลาด
Nominal Categorical	ข้อมูลที่เป็นนามบัญญัติ
Non-parametric Methods	วิธีการที่ไม่ใช้พารามิเตอร์
Online Analytical Processing : OLAP	กระบวนการวิเคราะห์ออนไลน์
Online Social Media	โซเชียลมีเดียออนไลน์
Optimization	การเพิ่มประสิทธิภาพ
Ordinal categorical	ข้อมูลที่เป็นลักษณะลำดับขั้นหรือชั้น
Outliers	ค่าผิดปกติออกมา
Parametric Methods	วิธีพารามิเตอร์
Quantitative Attributes	ข้อมูลที่เป็นตัวเลข
Raw Data	ข้อมูลดิบ
Readmission criteria	การรับผู้ป่วยเข้าใหม่
Real time	แบบเรียลไทม์
Recommendation System	ระบบแนะนำ
Sampling	การสุ่มตัวอย่าง
Smart Watch	นาฬิกาอัจฉริยะ
Supervised Machine Learning	เทคนิคการเรียนรู้แบบมีผู้สอน
Unsupervised Machine Learning	เทคนิคการเรียนรู้แบบไม่มีผู้สอน
Visualization	การทำจินตทัศน์

คำถามและแบบฝึกหัดท้ายบทที่ 8

1. จงอธิบายประโยชน์ของการเรียนรู้ด้านการทำเหมืองข้อมูล (Data Mining)
2. จงอธิบายเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Machine Learning)
3. จงอธิบายแบบจำลองแบบเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Machine Learning)
4. จงอธิบาย กระบวนการทำเหมืองข้อมูล “CRISP-DM” Cross-Industry Standard Process for Data Mining
5. จงอธิบายนักวิเคราะห์ข้อมูล (Data Analyst)

6. จงอธิบายนักวิทยาศาสตร์ข้อมูล (Data Scientist)
7. จงอธิบายงานของนักวิทยาศาสตร์ข้อมูล (Data Scientist) ส่วนใหญ่
8. จงอธิบายความแตกต่างระหว่าง นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)
9. จงอธิบายการเตรียมข้อมูลที่มีประสิทธิภาพ ควรมีลักษณะสำคัญดังนี้
10. จงอธิบายการทำความสะอาดข้อมูล (Data cleaning)
11. จงอธิบายการรวมข้อมูล (Data Integration)
12. จงอธิบายการแปลงข้อมูล (Data Transformation)
13. จงอธิบายการลดรูปข้อมูล (Data Reduction)
14. จงอธิบายวิธีพารามетริก (Parametric Methods)
15. จงอธิบายวิธีการที่ไม่ใช่พารามетริก (Non-parametric Methods)

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

บทที่ 9 เทคนิคและ อัลกอริทึม สำหรับการทำเหมืองข้อมูล

(Techniques and Algorithms for Data Mining)

การทำเหมืองข้อมูลคือกระบวนการทางเทคนิคที่ใช้คอมพิวเตอร์และอัลกอริทึมในการค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ (Big Data) ช่วยในการวิเคราะห์เพื่อประมวลผลและสำรวจชุดข้อมูลขนาดใหญ่ เมื่อใช้เทคนิคและอัลกอริทึมที่เหมาะสมแล้ว องค์กรจะสามารถค้นพบรูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในข้อมูลขององค์กร การทำเหมืองข้อมูลแปลงข้อมูลดิบเป็นความรู้เชิงปฏิบัติ องค์กรธุรกิจจะใช้ความรู้นี้ในการแก้ไขปัญหา รวมถึงวิเคราะห์ผลกระทบในอนาคตของการตัดสินใจเพื่อกำหนดกลยุทธ์

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจเทคนิคและอัลกอริทึมในการค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ (Big Data)
2. สามารถเข้าใจประยุกต์เทคนิคและอัลกอริทึมต่างๆ เพื่อประยุกต์ใช้ในการออกแบบระบบสารสนเทศสมัยใหม่

9.1 ความรู้เบื้องต้นเกี่ยวกับเทคนิคและ อัลกอริทึม สำหรับการทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data Mining) เป็นงานที่เกี่ยวข้องกับการรวบรวมข้อมูลจากหลายๆ แห่ง มีการเตรียมการข้อมูล การสร้างโมเดลเพื่อการวิเคราะห์ การแสดงภาพ (Visualization) การประเมินผลโมเดล (Model Evaluation) การจัดการนำไปใช้ (Deployment) และการเก็บรักษาข้อมูลจำนวนมาก การทำเหมืองข้อมูล (Data Mining) จะต้องมีความรู้หลากหลายสาขาเช่น ด้านระบบฐานข้อมูล วิทยาการคอมพิวเตอร์ การนำเสนอทางธุรกิจ รวมถึงทักษะที่ไม่ใช่ทางคณิตศาสตร์ก็เป็นสิ่งจำเป็นในการทำเหมืองข้อมูล (Data Mining) นอกจากนี้ นักวิทยาศาสตร์ข้อมูลและนักวิเคราะห์ข้อมูล ต้องมีส่วนร่วมในการเก็บข้อมูล ออกแบบการเก็บรักษาข้อมูลที่รวบรวมไว้ในรูปแบบที่ทำให้สามารถนำกลับมาใช้ใหม่ได้สะดวก รวมทั้งต้องมีทักษะการเรียนรู้อย่างรวดเร็วในการประยุกต์ใช้ข้อมูล ในบริบทเฉพาะของธุรกิจ สามารถการสื่อสารกับผู้ใช้ข้อมูล มีทักษะที่ดีในการเรียนรู้ความต้องการและความชอบของผู้ใช้ การแปลผลกลับไปกลับมาระหว่างการใช้เทคนิคทางวิทยาการคอมพิวเตอร์กับความต้องการของธุรกิจ

สำหรับเทคนิคการทำเหมืองข้อมูล (Data Mining) เป็นกระบวนการหรือเทคนิคที่ใช้ค้นหา รูปแบบในข้อมูลหรือเพื่อหาข้อมูลเชิงลึกที่ตอบโจทย์ทางธุรกิจ ปัจจุบันมีเทคนิคการทำเหมืองข้อมูล

(Data Mining) มากมายที่สามารถนำมาใช้เพื่อเปลี่ยนข้อมูลดิบให้เป็นข้อมูลเชิงลึกที่นำไปปฏิบัติได้ ในการดำเนินกลยุทธ์ทางธุรกิจ ผู้เขียนได้คัดเลือกเทคนิคที่เป็นที่นิยมมาแนะนำเสนอดังตาราง 9.1

ตาราง 9.1 เทคนิคการทำเหมืองข้อมูล (Data Mining)

เทคนิคการทำเหมืองข้อมูล (Data Mining Techniques)	คำอธิบาย	ตัวอย่างเทคนิค
1. รูปแบบการติดตาม (Tracking patterns)	เป็นหนึ่งในเทคนิคพื้นฐานที่สุดในการทำเหมืองข้อมูลคือการเรียนรู้จดจำรูปแบบในชุดข้อมูล เช่น ความคลาดเคลื่อนบางอย่างในข้อมูลที่เกิดขึ้นในช่วงเวลาหนึ่ง หรือการลดลงและการไหลของตัวแปร หรือข้อมูลบางตัวในช่วงเวลาหนึ่ง ตัวอย่างเช่น การลดลงการขายผลิตภัณฑ์บางอย่าง แต่ดูเหมือนจะมียอดขายพุ่งขึ้นก่อนวันหยุด หรือสังเกตเห็นว่าสภาพอากาศที่ร้อนขึ้นทำให้ผู้คนมาที่เว็บไซต์ของธุรกิจมากขึ้น	อนุกรมเวลา (Time Series Data Mining: TSDM)
2. การจำแนกประเภท (Classification)	การจำแนกข้อมูล (Classification) หรือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) คือประเภทของ Data Mining ที่ทราบค่าประเภทของตัวอย่างในชุดข้อมูล สอนล่วงหน้า เป็นการสร้างโมเดลหรือตัวจำแนกข้อมูล (Classifier) เพื่อทำนายหมวดหมู่ของข้อมูล (Categories/ Class) โดยมีตัวอย่างในชุดข้อมูลสอน (Training Set) ที่ใช้ จะมีคุณลักษณะหนึ่ง ซึ่งบอกค่า	● ต้นไม้แห่งการตัดสินใจ (Decision Tree) ● Random Forests. ● นาอิว เบย์ (Naïve Bayesian) ● ซัพพอร์ตเวกเตอร์ แมชชีน (Support Vector Machine) ● การถดถอยโลจิสติกส์ (Logistic Regression)

	ประเภทของตัวอย่างนั้น (Class Label) ซึ่งเป็นค่าข้อมูลแบบ Category	
3. การหาความสัมพันธ์ (Association)	กฎความสัมพันธ์ (Association Rules) เป็นกระบวนการหนึ่งในการทำ Data Mining ที่ได้รับความนิยมมาก โดยจะใช้ Association Rules ในการหาความสัมพันธ์ของข้อมูลสองชุดหรือมากกว่าสองชุดขึ้นไปภายในกลุ่มข้อมูลที่มีขนาดใหญ่	Apriori
4. การตรวจจับค่าผิดปกติ (Outlier detection)	การตรวจหาค่าผิดปกติจากข้อมูล คือ: การใช้ระยะทางและความหนาแน่นของจุดข้อมูลสำหรับการตรวจจับค่าผิดปกติ หรือการสร้างแบบจำลองเพื่อคาดการณ์การกระจายจุดข้อมูลและการเน้นค่าผิดปกติที่ไม่ตรงตามเกณฑ์ที่ผู้ใช้งานกำหนด	• Numeric • Z-Score • DBSCAN • Isolation Forest
5. การจัดกลุ่ม (Clustering)	การหากลุ่มของวัตถุที่มีความคล้ายคลึงกันในกลุ่มแต่แตกต่างจากวัตถุในกลุ่มอื่น ในกระบวนการจัดกลุ่มในการวิเคราะห์ข้อมูล ชุดข้อมูลจะถูกแบ่งออกเป็นกลุ่มหรือคลาสตามความคล้ายคลึงของข้อมูล	• K-Means clustering • Mean-Shift Clustering • DBSCAN – Density-Based Spatial Clustering of Applications with Noise. ... • EM using GMM – Expectation-Maximization (EM) Clustering using Gaussian Mixture Models (GMM) • Agglomerative Hierarchical Clustering.

6. ระบบการแนะนำ (Recommendation System)	เป็นเทคนิคที่สร้างโมเดลการแนะนำที่เกี่ยวข้องกับการชอบและไม่ชอบของผู้ใช้หรือลูกค้า วัตถุประสงค์หลักคือเพื่อแนะนำรายการให้กับผู้ใช้ที่หรือลูกค้า ซึ่งมีโอกาสสูงที่จะชอบหรือต้องการผู้ใช้ที่หรือลูกค้า รายใดรายหนึ่งจากการซื้อครั้งก่อนของเขา	• Collaborative Filtering. • Content-Based Filtering. • Hybrid Recommendation Systems. • Summing Up.
7. การทำนาย (Prediction) หรือโครงข่ายประสาทเทียม (Neural Networks)	เป็นตัวแทนในการทำนายแบบไม่ใช่เชิงเส้นตรง ซึ่งหาได้จากการสร้างตัวแบบและลอกเลียนเครือข่ายประสาทเชิงชีวภาพ	• Artificial Neural Networks (ANN) • Convolution Neural Networks (CNN) • Recurrent Neural Networks (RNN)

9.2 ความรู้เบื้องต้นเกี่ยวกับเทคนิคการจำแนกประเภท (Classification)

การจำแนกประเภท (Classification) เป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ ประเภทของ Data Mining ที่ทราบค่าประเภทของตัวอย่างในชุดข้อมูลสอนล่วงหน้า เป็นการสร้างโมเดลหรือตัวจำแนกข้อมูล (Classifier) เพื่อทำนายหมวดหมู่ของข้อมูล (Categories/ Class) โดยมีตัวอย่างในชุดข้อมูลสอน (Training Set) ที่ใช้ จะมีคุณลักษณะหนึ่ง ซึ่งบอกค่าประเภทของตัวอย่างนั้น (Class Label) ซึ่งเป็นค่าข้อมูลแบบ Category ทั้งนี้ การประยุกต์ใช้งานแบบ Classification มักพบในการขออนุมัติค้ำขอมีบัตรเครดิต (Credit Approval), การตลาดลูกค้ากลุ่มเป้าหมาย (Target Marketing), การตรวจวินิจฉัย (Medical Diagnosis) และการวิเคราะห์ประสิทธิผลการรักษา (Treatment Effectiveness Analysis) (จิราภา เลหาะวรรณน์ และคณะ, 2561)

การเรียนรู้แบบต้นไม้ตัดสินใจ (Decision Tree Learning) เป็นเทคนิคที่สามารถสืบย้อนหลังถึงการตัดสินใจและใช้งานตั้งแต่ปี 2506 กับผลงานของ Morgane และ Sonquist (Morgan, J. N., & Sonquist, J. A., 1963) ซึ่งเป็นกลุ่มแรกที่ใช้แผนผังการตัดสินใจในกระบวนการทำนายและจำแนกประเภทการตรวจจับแรงดึงดูด (AID Automatic Interaction Detection) ซึ่งวิธีนี้ได้รับการพัฒนาด้วยเทคนิค CART (Classification และ Regression Tree) โดยแบบจำลองสำหรับการพยากรณ์เทคนิคการเรียนรู้แบบต้นไม้ตัดสินใจ (Decision Tree) เป็นประเภทของการเรียนรู้ของเครื่องแบบมีผู้สอน (Supervised Learning) ประเภทหนึ่ง (นักวิเคราะห์ต้องอธิบายว่าอินพุต (Input) คืออะไรและผลลัพธ์

(Output) ที่สอดคล้องกันคืออะไรในข้อมูลเรียนรู้ (Training Data)) โดยที่ข้อมูลจะถูกแบ่งอย่างต่อเนื่องตามพารามิเตอร์บางอย่าง ต้นไม้สามารถอธิบายได้ด้วยสองเอนทิตี (Entity) ได้แก่ โหนดการตัดสินใจและใบไม้ ใบไม้คือการตัดสินใจหรือผลลัพธ์สุดท้าย และโหนดการตัดสินใจเป็นที่ที่ข้อมูลถูกแยกออกดังตัวอย่างตารางที่ 9.1

ตารางที่ 9.1 การเปรียบเทียบเทคนิคการจำแนกประเภท (Classification) ขั้นพื้นฐาน (Kumari and Godara, 2011)

เทคนิค	Decision Tree	Naïve Bayesian	Logistic Regression
ความหมาย	ต้นไม้ตัดสินใจ - ใช้ตรวจสอบข้อมูลและสร้างต้นไม้เพื่อการพยากรณ์	นาอิว เบย์ - ใช้ในการคัดแยกประเภทข้อมูลผ่านหลักความน่าจะเป็นที่อยู่บนพื้นฐานของทฤษฎี Bayes โดยเหตุการณ์เป็นอิสระต่อกัน	การถดถอยโลจิสติกส์ - การพยากรณ์ความน่าจะเป็นของผลลัพธ์ที่ขึ้นกับตัวแปรที่ส่งผลต่อเหตุการณ์นั้นๆ
ลักษณะ	คล้ายโครงสร้างต้นไม้ทั่วไป แตกแขนงไปตามเงื่อนไขหรือเส้นทางของกิ่งไม้ และข้อมูลที่คาดคะเนไว้ว่าจะเกิดขึ้น ซึ่งจะใช้กฎในรูปแบบ “ถ้า (เงื่อนไข) แล้ว (ผลลัพธ์)”	ไม่มีการหมุนวนที่ซับซ้อน ส่งผลให้สามารถทำงานได้ดีและมีประโยชน์กับชุดข้อมูลขนาดใหญ่อย่างมาก ทำให้เทคนิคนี้ได้รับความนิยมไว้อย่างรวดเร็วและถูกใช้งานอย่างกว้างขวาง	ผลลัพธ์ที่จะเกิดขึ้น สามารถเป็นได้เพียง 2 ค่า เช่น ใช่หรือไม่ใช่ โดยการพยากรณ์จะขึ้นกับตัวแปรที่ส่งผลต่อเหตุการณ์นั้นๆ ซึ่งอาจจะมีเพียงหนึ่งตัวหรือมากกว่า
การทำงาน	คัดเลือก Attribute ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็น Root Node จากนั้นทำการแตกกิ่งไปเรื่อยๆ จนสามารถแบ่งข้อมูลออกเป็นคลาสได้ชัดเจน	หลักการคำนวณของเทคนิคนี้จะอ้างอิงจากทฤษฎีโดยสันนิษฐานว่า ผลลัพธ์หรือค่าที่เกิดจากตัวที่ใช้ทำนาย (Predictor) เป็นอิสระต่อกัน	สร้าง Curve จากการใช้ลอการิทึม “Odds” ในตัวแปรเป้าหมาย ซึ่งจำกัดค่า 0 ถึง 1
ข้อดี	เร็ว ใช้กับการค้นหาข้อมูล SQL ใช้กับข้อมูลที่มี Error หรือ Missing Value ใช้ดีกับ Output ที่เป็น Discrete	ง่ายต่อการนำไปใช้ เพราะเป็นการคำนวณที่ง่าย ได้ผลลัพธ์ที่นำไปประยุกต์ใช้ได้ดี ใช้กับข้อมูลที่มีขนาดใหญ่	เร็ว สามารถแปลผลได้ง่าย ค่าตัวแปรอิสระ ไม่จำเป็นต้องมี Normal Distribution

ข้อเสีย	อาจเกิด Overfitting หากมีหลายกึ่ง ความถูกต้องน้อยสำหรับตัวอย่างที่ไม่เคยเห็นมาก่อน	ใช้ได้กับ Attribute ที่เป็นอิสระต่อกันเท่านั้น	มี Overfitting ใช้ไม่ดีกับ Output แบบ Continuous หรือข้อมูลที่มี Outliers และ Missing Value
---------	--	--	---

กรณีศึกษาเกี่ยวกับเทคนิคการจำแนกประเภท (Classification)

หลักการสร้างต้นไม้ตัดสินใจจะประกอบไปด้วยสองกระบวนการหลักคือ ขั้นตอนหนึ่งคือการกระบวนการสร้างตัวจำแนกข้อมูล ขั้นตอนที่สองของการจำแนกข้อมูล ลักษณะของข้อมูลแต่ละแถวซึ่งข้อมูลที่น่ามาวิเคราะห์จะประกอบไปด้วยเซตของแอตทริบิว (Attribute) ที่บ่งบอกถึงคุณลักษณะของบุคคลหรือเหตุการณ์ดังตารางที่ 1 ซึ่งแสดงหมวดหมู่ของบุคคลนั้นๆว่า จะตัดสินใจเล่นเบสบอลหรือไม่ โดยกระบวนการสร้างตัวจำแนกข้อมูลมักถูกเรียกว่า ‘Learning’ หรือ ‘Training’ ที่เกิดจากการนำเอาข้อมูลมาดำเนินการจำแนก

ในงานวิจัยของผู้เขียนได้เขียนอธิบายเทคนิคการจำแนกประเภท (Classification) (ประจักษ์ เฉิดโถม, 2566) ไว้ว่า ในชุดข้อมูลประกอบไปด้วยเซตของแอตทริบิว $X = (x_1, x_2, \dots, x_n)$, และข้อมูลเรคคอร์ด X หนึ่งๆ (แถวหนึ่งๆ) จะบ่งบอกถึงคุณลักษณะต่างๆของข้อมูล เช่น แถวที่ 1 ถ้าสภาพอากาศ (Sunny, Hot, High, windy(False)) จะไม่เล่นเบสบอล โดยชุดข้อมูลจะมีแอตทริบิว (Attribute) ที่บ่งบอกถึงหมวดหมู่ของข้อมูล (Class Label Attribute) โดย (Attribute) หมวดหมู่ข้อมูลจะเป็นข้อมูลแบบไม่ต่อเนื่อง (Discrete-Valued) ซึ่งจากตารางก็คือ (Attribute) “Play” โดยข้อมูลที่เข้ามานั้นอาจมีความผิดพลาดได้บ้าง และต้นไม้การตัดสินใจจะมีกระบวนการที่จะไม่นำความผิดพลาดนั้นมาพิจารณาเรียกว่าการตัดแต่งกิ่ง (Post-pruning) ดังตารางที่ 9.3

ตารางที่ 9.2 ชุดข้อมูลคุณลักษณะสำหรับการตัดสินใจเล่นเบสบอล

ข้อมูลชุด No.	สภาพอากาศ (Outlook)	อุณหภูมิ (Temperature)	มีความชื้น (Humidity)	ลมแรง (Windy)	เล่น (Play)
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes

5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	mild	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes

โดยสรุปลักษณะของข้อมูล ชุดข้อมูลที่เป็นอินพุต (Input) สำหรับทำการสร้างตัวจำแนกข้อมูล จะถูกเรียกว่า “ชุดข้อมูลสำหรับการสอน (Training data) การเตรียมข้อมูล นักวิเคราะห์ข้อมูลต้องอธิบายและจำแนกค่า (คลาส, กิริยา) ที่ใช้เป็นหมวดหมู่ ตัวอย่างของอัลกอริทึม (Algorithm) การเรียนรู้แบบต้นไม้ตัดสินใจ (Decision Tree Learning) ได้แก่ the C4.5 algorithm the C5.0 algorithm the AID tree the C5 tree the CHAID tree CART and QUEST เป็นต้น

สำหรับสูตรที่ใช้ในการเรียนรู้แบบต้นไม้ตัดสินใจ (Decision Tree Learning) โดยเบื้องต้นจะประกอบด้วยสูตรการคำนวณ 3 สูตร ดังตารางที่ 9.3 ที่สุด (Rokach & Maimon, 2005, Muhajir, 2022, Sharma et.al., 2021)

ตารางที่ 9.3 สูตรพื้นฐานสำหรับการเรียนรู้ต้นไม้การตัดสินใจ (Sharma et.al., 2021)

Formulas		
1. Information Gain	$I(S_1, S_2, \dots, S_n) = - \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S}$	I หมายถึงจำนวนของแต่ละคลาสในฐานข้อมูล S หมายถึงจำนวนทั้งหมดในฐานข้อมูล
2. Entropy	$E(A) = - \sum_{j=1}^n \frac{S_{ij} + \dots + S_{nj}}{S} I(S_1, S_2, \dots, S_n)$ Or $\text{Entropy}(S) = - \sum_{i=1}^c p_i \log_2 p_i$	S แสดงจำนวนข้อมูลทั้งหมดสำหรับแต่ละคลาสในข้อมูลของคอลัมน์นั้น S หมายถึงจำนวนของแต่ละคลาสของคอลัมน์นั้นในฐานข้อมูล)

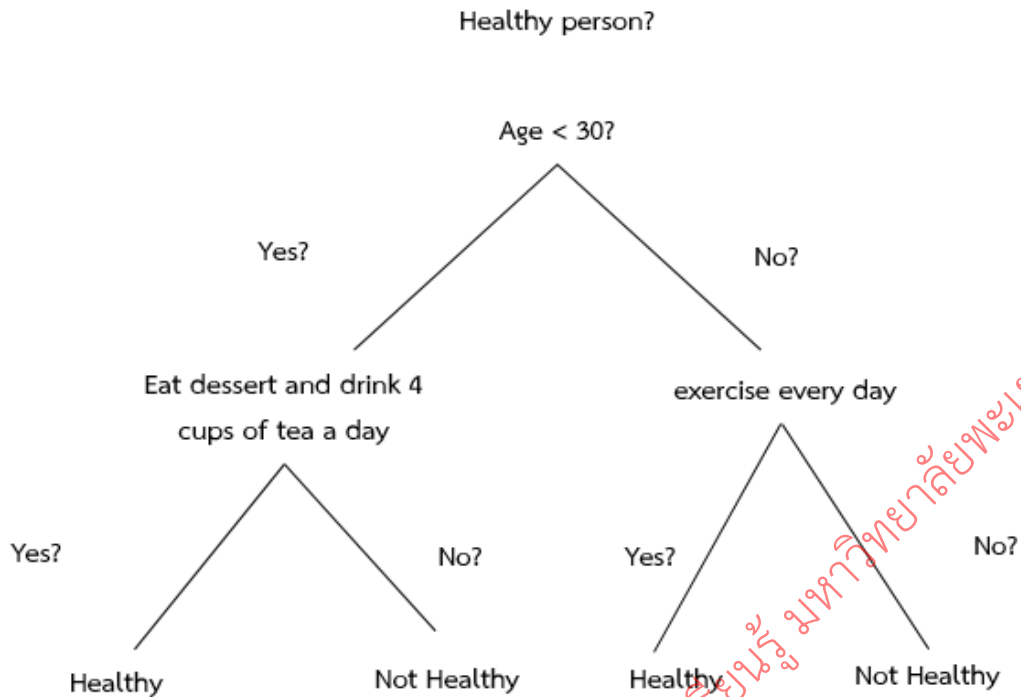
		S หมายถึงจำนวนทั้งหมดในฐานข้อมูล
3. Gain	Gain = Information Gain – Entropy	

โดยเทคนิคต้นไม้การตัดสินใจ (Decision Tree) เป็นโมเดล (Model) แบบ Rule-Based คือ สร้างกฎ if-else จากค่าของแต่ละ feature โดยไม่มีสมการมากำกับความสัมพันธ์ระหว่าง Feature & Target สิ่งที่สำคัญในการสร้าง Decision Tree คือ การเลือก Split ค่า Feature แต่ละครั้ง จะต้อง Minimize ค่าของ Cost Functionให้น้อยที่สุด (Rokach & Maimon, 2005, Muhajir, 2022) ข้อดีของเทคนิคต้นไม้การตัดสินใจ (Decision Tree) คือ เป็นอัลกอริทึมที่ถูกนำมาใช้ในหลายๆด้าน ในด้านของธุรกิจก็นำมาช่วยในการตัดสินใจเพื่อช่วยในการสร้างแผนงาน และเป็นพื้นฐานหนึ่งในการทำเหมืองข้อมูล และยังสามารถนำมาใช้ในการทำนายเหตุการณ์การตัดสินใจได้ดี ง่ายต่อการนำไปใช้ เพราะเป็นการคำนวณที่ง่าย ได้ผลลัพธ์ที่นำไปประยุกต์ใช้ได้ดี ใช้กับข้อมูลที่มีขนาดใหญ่ และที่สำคัญ เหมาะกับข้อมูลวิจัยที่ แอททริบิวต์ (Attribute) เป็นอิสระต่อกันเท่านั้น (Kumari & Godara, 2011)

อีกหนึ่งตัวอย่างการจำแนกประเภท (Classification) โดยใช้เทคนิคต้นไม้ตัดสินใจ (Decision Tree) ของข้อมูลว่าลูกค้ามีสุขภาพดีหรือไม่จากพฤติกรรมและอายุ ดังรูปที่ 9.1

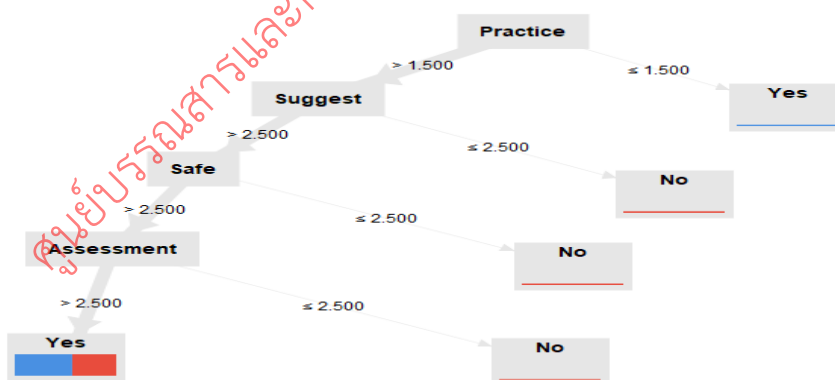
- (1) ลูกค้าที่อายุน้อยกว่า 30 ปี
- (2) ลูกค้าที่อายุน้อยกว่า 30 ปี และทานขนมและเครื่องดื่มประเภทชา 4 แก้วต่อวัน
- (3) เพื่อทำนายว่าโอกาสที่สุขภาพไม่ดี จะมาจากเส้นไหม เป็นจำนวนเท่าไร จากชุด

ข้อมูลที่เรียนรู้



รูปที่ 9.1 กรณีศึกษาเทคนิคการจำแนกประเภท (Classification) แบบ Decision Tree เพื่อจำแนกผู้มีสุขภาพดี

กรณีศึกษาที่ 2 ในด้านเทคนิคการจำแนกประเภท (Classification) จากการทดลองของผู้เขียนในงานวิจัยค้นหาเหตุผลการเลือกเรียนออนไลน์มาจากเหตุผลใด ด้วยเทคนิคการเรียนรู้แบบต้นไม้ตัดสินใจ (Decision Tree) พบว่าผู้ที่เลือกเรียนออนไลน์จากข้อมูล 200 คนนั้นสามารถสร้างแบบจำลองได้ดังรูปที่ 9.2



รูปที่ 9.2 กรณีศึกษาเทคนิคการจำแนกประเภท (Classification) แบบ Decision Tree เพื่อจำแนก

หาเหตุผลการเลือกเรียนออนไลน์ (รูปจากงานวิจัยผู้เขียน) (Prajak, 2565)

โดยเหตุผลการเลือกเรียนออนไลน์มาจากเหตุผลดังต่อไปนี้

1. รูปแบบการสอนออนไลน์ที่มีกิจกรรมทบทวนความรู้เชิงปฏิบัติ
2. ทางสถาบันการศึกษามีประกาศและให้คำแนะนำในการเรียนออนไลน์ครบถ้วนชัดเจน
3. ผู้เรียนตระหนักว่าสามารถประหยัดค่าใช้จ่ายจากการเดินทางไปเรียนสถาบัน (On-site)
4. และมีความพึงพอใจที่มีต่อการสอนของครูต่อการวัดผลและประเมินผลที่เหมาะสมกับเนื้อหาวิชา

ดังนั้นประโยชน์ของการเรียนรู้ด้าน Data Mining (เหมืองข้อมูล) จะทำให้ ผู้ศึกษาสามารถหาแนวทางการตัดสินใจและคาดการณ์ผลลัพธ์ที่ได้แม่นยำความวิธีการทางสถิติเดิมๆ ซึ่งตัวอย่างการนำเหมืองข้อมูลไปใช้งานในธุรกิจ เช่น การทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่ การทำนายยอดขายเมื่อมีการลดราคาสินค้า การทำนายกลุ่มลูกค้าที่น่าจะใช้สินค้าของเรา กล่าวโดยสรุปหากธุรกิจสามารถรู้ล่วงหน้า และวิ่งนำหน้าผู้บริโภค ด้วยการใช้ข้อมูลการสั่งซื้อ ประวัติลูกค้าทั้งเก่าและใหม่ เพื่อนำมาวิเคราะห์หาความน่าจะเป็น คาดคะเนจากเทคนิคเหมืองข้อมูล และแปลผลรูปแบบข้อมูลที่ได้จากการวิเคราะห์ เพื่อให้จับต้องได้ จะทำให้สามารถนำมาพัฒนากลยุทธ์ขององค์กร ซึ่งผู้ที่เข้าใจการทำ Data Mining (เหมืองข้อมูล) เช่น จะลูกค้ามากขึ้นทำให้คุณทราบว่าตอนนี้ลูกค้าต้องการอะไร แล้วลูกค้าประเภทไหนที่ซื้อสินค้าหรือใช้บริการ รวมถึงสามารถทำนายพฤติกรรมของผู้บริโภค และออกแบบสินค้าและบริการ ให้ตอบโจทย์กลุ่มลูกค้าประจำและกลุ่มลูกค้าใหม่ได้ และเห็นมุมมองในการทำตลาดที่กว้างขึ้น จากการมีข้อมูลในมือจำนวนมากและหลายรูปแบบ ทำให้มุมมองในการวางแผนการตลาดก็ยิ่งกว้างมากขึ้น

9.3 ความรู้เบื้องต้นเกี่ยวกับเทคนิคการแบ่งกลุ่ม (Clustering)

การแบ่งกลุ่ม (Clustering) เป็นเทคนิคที่ใช้จำแนกหรือแบ่งประเภท (คน สัตว์ สิ่งของ หรือ องค์กร ฯลฯ) หรือ แบ่งตัวแปรออกเป็นกลุ่มย่อย ๆ ตั้งแต่ 2 กลุ่มขึ้นไป โดยประเภทที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่เหมือนกันหรือคล้ายกัน ส่วนประเภทที่อยู่ต่างกลุ่มกัน จะมีลักษณะที่แตกต่างกัน ดังนั้น การพิจารณาเลือกลักษณะหรือตัวแปรที่นำมาใช้ในการแบ่งกลุ่มประเภท จึงมีความสำคัญ นอกจากนั้น ประเภทใดประเภทหนึ่งจะต้องอยู่ในกลุ่มหนึ่งเพียงกลุ่มเดียว ถ้านำเทคนิคการแบ่งกลุ่ม (Clustering Analysis) มาใช้ในการแบ่งกลุ่มตัวแปร จะให้ตัวแปรอยู่ในกลุ่มเดียวกันมีความสัมพันธ์กันมากกว่าตัวแปรที่อยู่ต่างกลุ่มกัน ตัวแปรที่อยู่ต่างกลุ่มกันมีความสัมพันธ์กันน้อยหรือไม่มีความสัมพันธ์กันเลย ส่วนใหญ่การแบ่งกลุ่มตัวแปรจะใช้เทคนิค Factor ส่วนการแบ่งกลุ่มประเภท (คน สัตว์ สิ่งของ) จะใช้เทคนิค เทคนิคการแบ่งกลุ่ม (Clustering)

การแบ่งกลุ่มข้อมูลจะแตกต่างจากเทคนิคการจำแนกประเภท (Classification) โดยจะแบ่งกลุ่มข้อมูลจากความคล้าย โดยไม่มีการกำหนดประเภทของข้อมูลไว้ก่อน จึงกล่าวได้ว่าการแบ่งกลุ่มข้อมูล เป็นการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) วัตถุประสงค์ของการแบ่งกลุ่ม (Clustering) เพื่อการลดขนาดข้อมูล (Data Reduction) ด้วยเทคนิคที่แม่นยำและมีประสิทธิภาพเป็นอีกแนวทางหนึ่งเพื่อจัดการกับปัญหาดังกล่าว หลักการที่สำคัญของการลดขนาดข้อมูลคือ การทำให้ข้อมูลตั้งต้นมีขนาดลดลงโดยสูญเสียลักษณะสำคัญของข้อมูลน้อยที่สุด เนื่องจากข้อมูลแต่ละตัวจะมีความสำคัญต่อการจัดกลุ่มข้อมูลไม่เท่ากัน ด้วยเทคนิคการเลือกข้อมูลที่ดีจะทำให้สามารถเลือกข้อมูลที่มีความสำคัญและสามารถใช้เป็นตัวแทนของข้อมูลส่วนใหญ่ได้ ข้อมูลที่มีการรวมกลุ่มกันอย่างหนาแน่นจะเป็นข้อมูลที่มีความสำคัญต่อการจัดกลุ่มข้อมูลในอนาคต

คุณสมบัติของเทคนิควิธีการแบ่งกลุ่ม (Clustering)

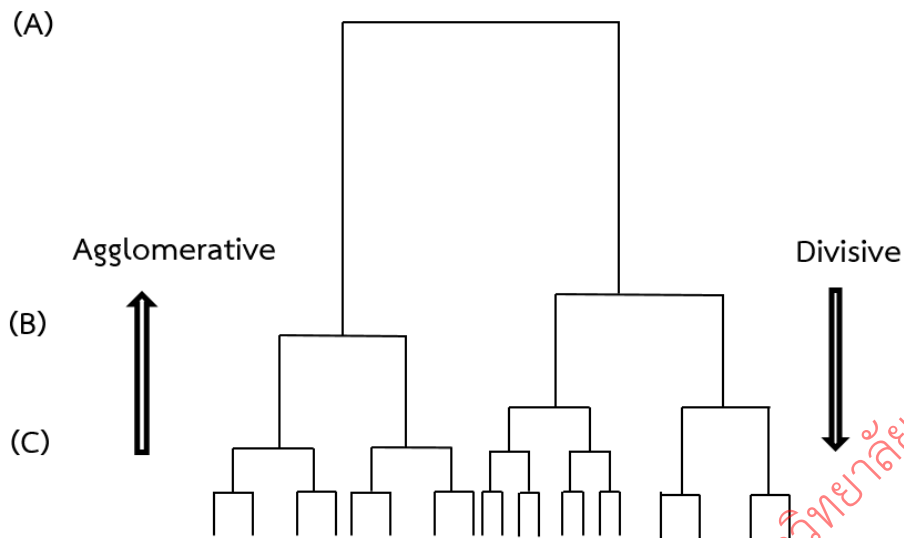
(1) ความต้องการทางด้านข้อมูล สำหรับการวิเคราะห์จัดกลุ่มหน่วยวิเคราะห์ ผู้วิจัยอาจใช้ข้อมูลที่ระบุหน่วยวิเคราะห์และตัวแปรตามที่ได้เก็บมาได้เลย เช่น การวิเคราะห์ที่ได้ กล่าวมาแล้วข้างต้น ส่วนการวิเคราะห์จัดกลุ่มตัวแปร ผู้วิจัยไม่อาจจะใช้เพิ่มข้อมูลดังกล่าวได้โดยใช้เมตริกแสดงความสัมพันธ์ระหว่างตัวแปรแทนได้

(2) แนวคิดพื้นฐาน สิ่งสำคัญที่สุดของการวิเคราะห์การจัดกลุ่มคือ ตัวแปรที่ใช้หากผู้วิจัยไม่ได้เก็บข้อมูลเกี่ยวกับตัวแปรที่สำคัญ ๆ ผลที่ได้ก็จะไม่ดีหรือทำให้ไขว้เขวได้ ทั้งนี้เพราะตัวแปรที่เลือกไว้ตั้งแต่แรกจะเป็นสิ่งที่กำหนดคุณสมบัติของสิ่งที่ระบุความเป็นกลุ่มย่อย เช่นในการจัดกลุ่มโรงเรียนในเมือง หากผู้วิจัยไม่เก็บข้อมูลเกี่ยวกับจำนวนนักศึกษาและครู ขนาดของโรงเรียนก็ไม่อาจเป็นเกณฑ์ในการแบ่งกลุ่มได้

(3) ความคล้ายกันของหน่วย ความคิดเกี่ยวกับความคล้ายของหน่วยศึกษาเป็นเทคนิคของการวิเคราะห์ทางสถิติหลายวิธี โดยทั่วไปการวัดความคล้ายจะพิจารณาจากความห่างระหว่างวัตถุ หรือพิจารณาจากความคล้ายกัน ซึ่งจะกล่าวโดยละเอียดในหัวข้อต่อไป

(4) การวัดความห่าง วิธีการวัดความห่างสามารถวัดได้หลายวิธี วิธีการหนึ่งที่นิยมวัดกันมากก็คือ วิธีที่เรียกว่า ระยะห่างเชิงยูคลิดยกกำลังสอง (Squared Euclidean Distance) คือ ผลรวมของผลต่างยกกำลังสองของทุกตัวแปร เช่น ต้องการดูความห่างกันของเบียร์ 2 ยี่ห้อ ซึ่งทราบราคา

การวิเคราะห์หรือจำแนกกลุ่มข้อมูล สามารถแบ่งออกเป็น 2 ประเภทตาม Algorithm ที่ใช้ได้แก่ วิธีการแบบลำดับขั้น (Hierarchical Algorithms) ดังรูป 9.3 และ วิธีการแบบไม่เป็นลำดับขั้น (Non-Hierarchical Algorithms) หรือที่เรียกว่า “Partitional clustering” ดังรูป 9.4 (Pratya, 2022)

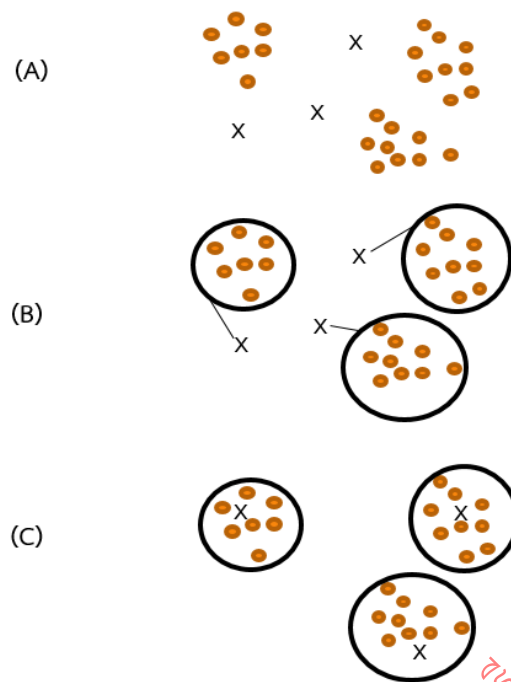


รูปที่ 9.3 วิธีการแบบลำดับชั้น (Hierarchical Algorithms)

1. Agglomerative ("bottom-up"): อัลกอริทึมที่อธิบายการรวมกลุ่มเริ่มต้นด้วยแต่ละองค์ประกอบในฐานะคลัสเตอร์ที่แยกจากกัน และรวมเข้ากับคลัสเตอร์ที่ใหญ่ขึ้นตามลำดับ
2. Divisive ("จากบนลงล่าง"): อัลกอริทึมที่อธิบายการแบ่งกลุ่มเริ่มต้นด้วยทั้งชุดและดำเนินการแบ่งออกเป็นกลุ่มเล็กๆ ตามลำดับ

ส่วน Partitional clustering เป็นวิธีที่จะต้องกำหนดเองว่าจะต้องแบ่งเป็นกี่กลุ่ม เช่น k กลุ่ม

ดังรูป 9.7



รูปที่ 9.4 วิธีการแบบไม่เป็นลำดับชั้น (Non-Hierarchical Algorithms) หรือ Partitional clustering

พื้นฐานการคำนวณ K-Means Clustering

การคำนวณ Partitional clustering (Nonhierarchical Cluster Analysis หรือบางครั้งเรียกว่า K – Means Cluster Analysis) แบ่งเป็น 3 เทคนิคย่อยคือ (Pulkit, 2022)

1. Euclidean distance โดยมีสูตรดังนี้

$$d = ((p_1 - q_1)^2 + (p_2 - q_2)^2)^{1/2}$$

2. Manhattan distance โดยมีสูตรดังนี้

$$d = |p_1 - q_1| + |p_2 - q_2|$$

3. Minkowski Distance โดยมีสูตรดังนี้

$$D = \left(\sum_{i=1}^n |p_i - q_i|^p \right)^{1/p}$$

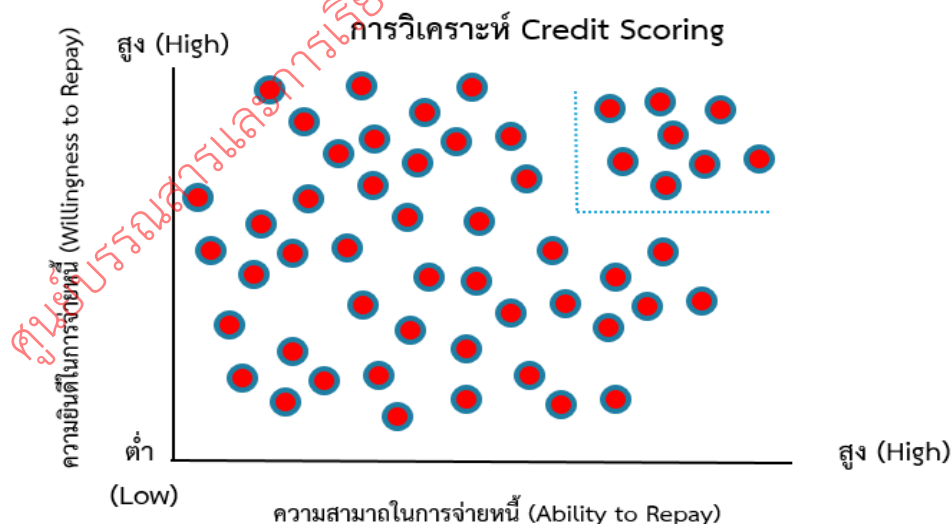
โดยการคำนวณที่ง่ายสำหรับการคำนวณด้วยมือคือสูตร Manhattan distance ซึ่งขั้นตอนการวิเคราะห์ของวิธี K-Means สามารถสรุปได้ดังนี้

ขั้นตอนการจัดกลุ่มดังนี้

1. สุ่มค่าเริ่มต้น จำนวน k ค่า Cluster Center (Centroid) การสุ่มที่ดี จะทำให้ การคำนวณ (Iteration) น้อยลง (ส่งผลให้การประมวลผล เร็วขึ้น) ในกรณีการทำ Big data Analysis
2. ทำการหาค่าระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางมากที่สุด หากข้อมูลไหนใกล้ค่าจุดศูนย์กลาง k มากที่สุดจะอยู่ในกลุ่มนั้น
3. หาค่าเฉลี่ย (Mean) แต่ละกลุ่ม กำหนดให้เป็นค่าจุดกลาง (Centroid) ใหม่
4. กลับไปทำข้อ 2 จนกว่าค่าเฉลี่ยในแต่ละกลุ่มจะไม่เปลี่ยนแปลง

กรณีศึกษาเกี่ยวกับเทคนิคการจัดกลุ่ม (Clustering)

Ricult เป็น Startup ที่มีเป้าหมายมุ่งเน้นเรื่องการสนับสนุนกลุ่มเกษตรกรโดยเฉพาะชื่อของ Ricult มาจากคำว่า Ag(ricult)ture โดยปัญหาในปัจจุบันคือ ผู้ให้สินเชื่อจะไม่ค่อยให้สินเชื่อกับเกษตรกรรายย่อย เนื่องจากไม่สามารถวัดความเสี่ยงได้ อาจเกิดหนี้สูญ สิ่งที่ Ricult ทำคือการให้คะแนนเครดิต (Credit Scoring) กับเกษตรกรรายย่อยในรูปของเกษตรภัณฑ์ต่างๆ เช่น เมล็ดพันธุ์ ปุ๋ย เป็นต้น ด้วยการใช้เทคโนโลยี Machine Learning ในการวิเคราะห์ข้อมูลต่างๆของเกษตรกรเกือบทุกด้าน และนำมาสร้างเป็น Credit Score ผ่านการใช้เทคนิคการแบ่งกลุ่ม (Clustering) โดยยึด 2 หัวข้อนี้เป็นหลักประกันความเสี่ยง ได้แก่ เรามีความตั้งใจจริงที่จะชำระเงินคืนไหม? (Willingness to Pay Back) หรือเรามีความสามารถที่จะชำระเงินคืนไหม? (Ability to Pay Back) (Oravee, 2022) ดังรูป



รูปที่ 9.5 กรณีศึกษาเทคนิคการจัดกลุ่ม (Clustering) เพื่อแบ่งกลุ่มลูกค้าและประเมิน Credit Score ของบริษัท Ricult

กรณีศึกษาที่ 2 เทคนิคการจัดกลุ่ม (Clustering) จากการทดลองของผู้เขียนในงานวิจัยชื่อ “K-means Cluster Analysis of Tourist Package and Social Network Analysis for Promotion Recommendation” เป็นงานวิจัยค้นหาประเภทของแพ็คเกจและราคากรุ๊ปทัวร์ของประเทศไทยของนักท่องเที่ยวที่ชอบเที่ยวญี่ปุ่นโดยเฉพาะตามแลนด์มาร์กต่างๆ ในภูมิภาคคันโต โดยแยกประเภทกรุ๊ปทัวร์ออกเป็นต่างๆ เช่น ชอบช้อปปิ้ง ชอบวัฒนธรรมและประวัติศาสตร์ เป็นต้น สำหรับการรวบรวมข้อมูล ทีมวิจัยรวบรวมข้อมูลจากประเทศไทยเว็บไซต์ทัวร์และหน่วยงานเช่น บริษัท H.I.S. บริษัท Quality Express บริษัท Gusto บริษัท World Tour เป็นต้น ซึ่งเป็นบริษัทการท่องเที่ยวที่เชื่อถือได้ รวมถึงโซเชียลเน็ตเวิร์กชื่อดังของเมืองไทย อาทิ พันธุ์ทิพย์ เฟสบุ๊ค และอื่นๆ เพื่อให้สามารถจัดกลุ่มและทำนายพฤติกรรมของนักท่องเที่ยวชาวไทย และใช้เทคนิคการทำเหมืองข้อมูลที่เรียกว่าการวิเคราะห์คลัสเตอร์สำหรับการจัดกลุ่มวิเคราะห์ การทำนายราคาและจำนวนวัน และความนิยมของกรุ๊ปทัวร์ไปญี่ปุ่น ซึ่งสามารถแบ่งเป็น 3 กลุ่มใหญ่ๆ ดังตาราง 9.4 (Prajak et al, 2018)

ตารางที่ 9.4 การแบ่งกลุ่มวิเคราะห์ การทำนายราคาและจำนวนวัน และความนิยมของกรุ๊ปทัวร์ไปญี่ปุ่นของนักท่องเที่ยวไทย

K=3 จำนวนกลุ่ม = 3	กลุ่มที่ 1 Cluster no.0	กลุ่มที่ 2 Cluster no.1	กลุ่มที่ 3 Cluster no.2
การใช้จ่ายเฉลี่ย	28,999	32,888	29,888
จำนวนวันท่องเที่ยว	5	6	6

โดยสรุปเทคนิคการจัดกลุ่ม (Clustering - K-means) สามารถใช้ได้เกือบทุกโดเมนธุรกิจ ตั้งแต่ธนาคารไปจนถึงธุรกิจค้าปลีก การจัดกลุ่มเอกสาร ไปจนถึงการแบ่งส่วนรูปภาพ โดยทั่วไปจะใช้กับข้อมูลที่มีขนาดน้อยกว่า เป็นตัวเลขและมีความต่อเนื่องกัน

9.4 ความรู้เบื้องต้นเกี่ยวกับเทคนิคกฎความสัมพันธ์ (Association Rules)

กฎความสัมพันธ์ (Association Rules) เป็นกระบวนการหนึ่งในการทำเหมืองข้อมูล (Data Mining) ที่ได้รับความนิยมมาก โดยจะใช้กฎความสัมพันธ์ (Association Rules) ในการหาความสัมพันธ์ของข้อมูลสองชุดหรือมากกว่าสองชุดขึ้นไปภายในกลุ่มข้อมูลที่มีขนาดใหญ่ ในการหาความสัมพันธ์

นั่นจะมีขั้นตอนวิธีการหาหลายวิธีด้วยกัน แต่ขั้นตอนวิธีที่เป็นที่รู้จักและใช้อย่างแพร่หลายคือ ขั้นตอนวิธี “Apriori” ซึ่งเป็นเทคนิคกฎความสัมพันธ์ (Association Rule) เช่น การเชื่อมโยงสำหรับการค้นหาความสัมพันธ์ระหว่างผลิตภัณฑ์ในข้อมูลธุรกรรมขนาดใหญ่ที่บันทึกโดยระบบ ณ จุดขาย (POS) ในซูเปอร์มาร์เก็ตสามารถเขียนได้ในรูปเซตของสิ่งของ (Item) ที่เป็นเหตุ ไปสู่เซตของสิ่งของ (Item) ที่เป็นผล ตัวอย่างเช่น เมื่อลูกค้าซื้อสินค้าใดสินค้าหนึ่ง พวกเขาก็มักจะซื้อสินค้าชิ้นที่สองที่เกี่ยวข้องกัน และนำไปสู่การจัดโปรโมชั่นสินค้าที่ลูกค้าจะใช้ซื้อร่วมกัน หรือจัดทำเลที่ตั้งของสินค้าในบริเวณใกล้เคียงกันซึ่ง

สูตรการคำนวณจะกำหนดให้แต่ละ Transaction ใน D จะมีหมายเลข Transaction ID ที่ไม่ซ้ำกันและ t เป็น subset ของ “ I ”

$I = \{i_1, i_2, \dots, i_m\}$ เป็น set ของ Items

$D = \{t_1, t_2, \dots, t_m\}$ เป็น set ของ Transaction

การหากฎที่มีความน่าสนใจ

กฎ (rules) If-then เกี่ยวกับการหยิบของใส่ตะกร้า

$\{i_1, i_2, \dots, i_k\} \rightarrow j$ หมายถึง “ถ้าในตะกร้ามีสิ่งของ (if a basket contains all of) $i_1, \dots, i_k$ ดังนั้นมีโอกาสที่จะมี j (then it is likely to contain j)”

โดยค่าความเชื่อมั่นในกฎคำนวณ เกิดจากการหารความน่าจะเป็นของรายการที่เกิดขึ้นด้วยกัน หารด้วยความน่าจะเป็นที่จะเกิดขึ้นก่อน

Confidence of this association rule is the probability of j given $I =$

$\{i_1, \dots, i_k\}$

$$\text{conf}(I \rightarrow j) = \frac{\text{support}(I \cup j)}{\text{support}(I)}$$

ค่าความเชื่อมั่นคือค่าเปอร์เซ็นต์ที่แสดงความถี่ที่เกิดขึ้น เพื่อระบุว่าโมเดลมีความน่าเชื่อถือเพียงใด

พื้นฐานการคำนวณด้วยวิธี Apriori

1. ขั้นตอนแรกหารูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อยๆ ในฐานข้อมูล (Frequent itemset) ประกอบไปด้วย การสร้างรูปแบบของ itemset (join) และ การนับค่า support (count) ดังตารางที่ 9.6

ตารางที่ 9.6 การสร้างรูปแบบของ itemset (join)

Transaction no.	Items
1	Bread, Coke, Milk
2	Beer , Bread
3	Beer , Coke, Diaper , Milk
4	Beer , Bread, Diaper , Milk

การนับค่า Support (count)

Support of {Beer, Bread} = 2 เรียกว่า “Frequent itemsets”

Support of {Beer, Diaper} = 2 เรียกว่า “Frequent itemsets”

Support of {Coke, Milk} = 3 เรียกว่า “Frequent itemsets”

ตารางที่ 9.7 การหาค่า Support ครั้งที่ 1

Items	Transaction no.				Support
	1	2	3	4	
Bread	1	1	0	0	$2/4 = 50\%$
Coke	0	0	1	0	$1/4 = 25\%$
Milk	1	0	1	1	$2/4 = 50\%$
Diaper	0	0	1	1	$2/4 = 50\%$
Beer	0	1	1	1	$3/4 = 75\%$

จากค่า support ที่คำนวณ จะเห็นว่า Coke มีค่า support ต่ำกว่าค่า minimum support ($25\% < 50\%$) ดังนั้น Coke จะถูกตัดออกและไม่นำไปพิจารณาสร้างเป็น itemset ที่มีความยาว 2 ต่อไป ดังรูปที่ 3 และเรียก Bread, Milk, Diaper และ Beer ว่า Frequent itemset

2. ขั้นตอนที่สองสร้าง Association Rule หลังจากที่เราหา Frequent itemset ได้แล้วจะนำรูปแบบที่เราได้มาสร้างเป็นกฎความสัมพันธ์โดย เช่น Beer => Diaper หมายความว่าเมื่อลูกค้าซื้อ Beer แล้วลูกค้าจะซื้อ Diaper ร่วมไปด้วย

นำ Frequent itemset ที่ได้ตารางด้านบน มาสร้างเป็น itemset โดยการ Join 2 items เช่น {Bread, Milk}, {Bread, Diaper}, {Bread, Beer}, {Milk, Diaper}, {Milk, Beer}, {Diaper, Beer}

ตารางที่ 9.8 การหาค่า Support ครั้งที่ 2

Items	Transaction no.				Support
	1	2	3	4	
{Bread, Milk}	1	0	0	1	$2/4 = 50\%$
{Bread, Diaper}	0	0	0	1	$1/4 = 25\%$
{Bread, Beer}	0	1	0	1	$2/4 = 50\%$
{Milk, Diaper}	0	0	1	1	$2/4 = 50\%$
{Milk, Beer}	0	0	1	1	$2/4 = 50\%$
{Diaper, Beer}	0	0	1	1	$2/4 = 50\%$

ตารางที่ 9.9 การนำ Frequent itemset ที่ได้ตารางด้านบน มาสร้างเป็น itemset โดยการ Join 3 items

Items	Transaction no.				Support
	1	2	3	4	
{Bread, Milk, Beer}	0	0	0	1	$1/4 = 25\%$
{Milk, Diaper, Beer}	0	0	1	1	$2/4 = 50\%$

ตารางที่ 9.10 การหาค่าสรุป frequent itemset

จากตารางด้านบน เราไม่สามารถสร้าง itemset ที่มีความยาวมากขึ้นกว่านี้ได้อีก ขั้นตอน
การหา frequent itemset จึงหยุด และได้ frequent itemsets ทั้งหมดดังแสดงใน ตารางถัดไป

Items	Support	Size
		จำนวน Items
Bread	$2/4 = 50\%$	1
Milk	$2/4 = 50\%$	1
Diaper	$2/4 = 50\%$	1
Beer	$3/4 = 75\%$	1
{Bread, Milk}	$2/4 = 50\%$	2
{Bread, Beer}	$2/4 = 50\%$	2
{Milk, Diaper}	$2/4 = 50\%$	2
{Milk, Beer}	$2/4 = 50\%$	2
{Diaper, Beer}	$2/4 = 50\%$	2
{Bread, Milk, Beer}	$1/4 = 25\%$	3
{Milk, Diaper, Beer}	$2/4 = 50\%$	3

ขั้นตอนถัดมาคือการสร้างกฎความสัมพันธ์จาก frequent itemset โดยจะพิจารณา
frequent itemset ที่มีความยาวมากกว่า 2 item ขึ้นไปมาสร้างเป็นกฎความสัมพันธ์ เช่น {Bread,
Milk} จะสร้างกฎความสัมพันธ์ได้เป็น Bread $\Rightarrow$ Milk เป็นต้น โดยกฎความสัมพันธ์ทั่วไปจะแสดงเป็น

LHS $\rightarrow$ RHS

ขั้นตอนต่อมากำหนดค่า confidence แสดงความเชื่อมั่นของกฎความสัมพันธ์ที่เมื่อ
รูปแบบ LHS เกิดขึ้นแล้วรูปแบบ RHS จะเกิดขึ้นด้วยเป็นจำนวนกี่เปอร์เซ็นต์

$$\text{conf}(I \rightarrow j) = \frac{\text{support}(I \cup j)}{\text{support}(I)}$$

ตัวอย่างเช่น

$$\begin{aligned}\text{Conf}(\text{Bread} \rightarrow \text{Milk}) &= \text{Support}(\text{Bread, Milk}) / \text{Support}(\text{Bread}) \\ &= (2/4) / (2/4) = 100\%\end{aligned}$$

ขั้นตอนต่อมาคำนวณหาค่า lift คือค่าที่บ่งบอกว่าการเกิดรูปแบบ LHS และ RHS มีความสัมพันธ์กันแค่ไหน โดยถ้าค่า lift เป็น 1 แสดงว่ารูปแบบ LHS และ RHS ไม่ขึ้นต่อกัน (independent) ค่า lift คำนวณได้จาก

$$\text{Lift}(\text{LHS} \rightarrow \text{RHS}) = \text{support}(\text{LHS, RHS}) / \text{Support}(\text{LHS}) \times \text{support}(\text{RHS})$$

ตัวอย่างเช่น

$$\begin{aligned}\text{Lift}(\text{Bread} \rightarrow \text{Milk}) &= \text{Support}(\text{Bread, Milk}) / \text{Support}(\text{Bread}) \times \text{Support}(\text{Milk}) \\ &= (2/4) / (2/4) \times (2/4) \\ &= 2\end{aligned}$$

ตารางที่ 9.11 การ หาค่า Confidence และค่า Lift ของกฎความสัมพันธ์ทั้งหมดที่สร้างได้ โดยทำการเรียงตามค่า Confidence และ Lift จากมากไปหาน้อย

Rule No.	Items	Confidence	Lift
1	{Bread $\rightarrow$ Milk}	100%	0.50
2	{Milk $\rightarrow$ Diaper}	100%	0.50
3	{Bread $\rightarrow$ Beer}	66%	0.75
4	{Milk $\rightarrow$ Beer}	66%	0.75
5	{Diaper $\rightarrow$ Beer}	66%	0.75
6	{Milk, Diaper $\rightarrow$ Beer}	66%	0.56
7	{Bread, Milk $\rightarrow$ Beer}	33%	0.56

กรณีศึกษาเกี่ยวกับเทคนิค Apriori

ผู้เขียนได้ทำวิจัยเรื่องการหาความสัมพันธ์ของการถูกรีไทร์จากสถาบันการศึกษากับโครงสร้าง การลงทะเบียนเรียนและผลการเรียนด้านวิชาด้านภาษาและเทคโนโลยี จากสถาบันแห่งหนึ่งโดยใช้ ข้อมูลย้อนหลัง 10 ปี โดยพบว่า กฎความสัมพันธ์ที่ได้จากการประมวลผล แสดงถึงการไม่สัมพันธ์กัน ระหว่างผลการเรียนด้านภาษาที่ดีจะส่งผลต่อการเรียนด้านเทคโนโลยีที่ดี (Chertchom et al., 2016)

ในทางธุรกิจการรับรู้พฤติกรรมของลูกค้า หรือทำนายได้ว่า ลูกค้าซื้ออะไร แล้วจะซื้ออะไรต่อ เป็นการนำมาซึ่งโอกาสความได้เปรียบทางธุรกิจอย่างมากมาย ซึ่งตัวอย่างหนึ่งของการใช้ Association Rules ในทางธุรกิจคือ “Market Basket Analysis” (การวิเคราะห์ตะกร้าการซื้อ) เพื่อหาความสัมพันธ์ ของสินค้าที่ลูกค้ามักจะซื้อพร้อมๆ กัน โดยการรวบรวมข้อมูลการซื้อของลูกค้าที่เก็บไว้ในอดีต ซึ่งจะ เก็บอยู่ในรูปของ Transaction แล้วนำข้อมูลเหล่านั้นมาทำการวิเคราะห์ข้อมูล (Data Analysis) ซึ่งผล การวิเคราะห์ข้อมูลสามารถเปิดเผยถึงพฤติกรรมในการซื้อสินค้าของลูกค้าได้ ประโยชน์ของการทำ Association Rule คือสามารถนำไปวางแผนเชิงกลยุทธ์เพื่อเพิ่มยอดขาย การจัดรายการส่งเสริมการ ขาย หรือนำไปพัฒนากลยุทธ์การขายอื่นๆ ต่อไป เช่น การกำหนดตำแหน่ง การวางสินค้า การจัด โปรโมชันกลุ่มสินค้า และการบริหารสินค้าคงคลัง เป็นต้น (Ashish, 2022)

อย่างไรก็ตามเทคนิค Apriori มีข้อเสียคือเมื่อทำงานกับชุดข้อมูลขนาดใหญ่ จะไม่มี ประสิทธิภาพ เนื่องจากใช้ทรัพยากรจำนวนมาก จากการต้องสแกนฐานข้อมูลหลายครั้ง และสร้างชุด ตัวเลือกจำนวนมาก และตรวจสอบกฎความสัมพันธ์ในการคำนวณค่า Confidence และ ค่า Lift ของ แต่ละกฎ

9.5 ความรู้เบื้องต้นเกี่ยวกับเทคนิคระบบแนะนำ (Recommendation System)

ในโลกธุรกิจเจ้าของธุรกิจต้องการรายได้เพิ่มขึ้นจากการขายสินค้าที่มากขึ้น ซึ่งในธุรกิจ ดังเดิมจะต้องใช้พนักงานแนะนำสินค้า ซึ่งอาจจะไม่ตรงตามความต้องการของผู้ซื้อ และสินค้าก็มีมากมาย เกินกว่าลูกค้าจะเห็นได้ทั้งหมดในเวลาที่มีจำกัด ซึ่งปัญหานี้ก็พบในธุรกิจออนไลน์ด้วยเช่นเดียวกัน ดังนั้นเทคนิค ของระบบแนะนำ (Recommendation System) คือระบบที่ช่วยแนะนำสินค้าและ บริการให้กับผู้บริโภคออนไลน์ บริษัทชั้นนำด้าน e-Commerce ใช้ระบบแนะนำในการสร้าง ประสบการณ์ในการใช้แพลตฟอร์ม (Platform) ที่ดีขึ้น เช่น Shopee LAZADA Amazon Netflix YouTube Google เป็นต้น โดยแพลตฟอร์มเหล่านี้พัฒนาโมเดลจากกฎ (Rules) ในการแนะนำสินค้า ซึ่งแบ่งเป็น 2 แบบคือ

แบบที่ 1: User-Based Filtering เป็นเทคนิคที่ใช้ในการทำนายรายการที่ประมวลจากการหาลูกค้าที่พฤติกรรมเหมือนกัน โดยอาจจะพิจารณาจากคะแนนที่เรต (Rating) สินค้ารายการนั้น และประมวลผลจากลูกค้ารายอื่นที่มีรสนิยมคล้ายกันกับของลูกค้าเป้าหมาย

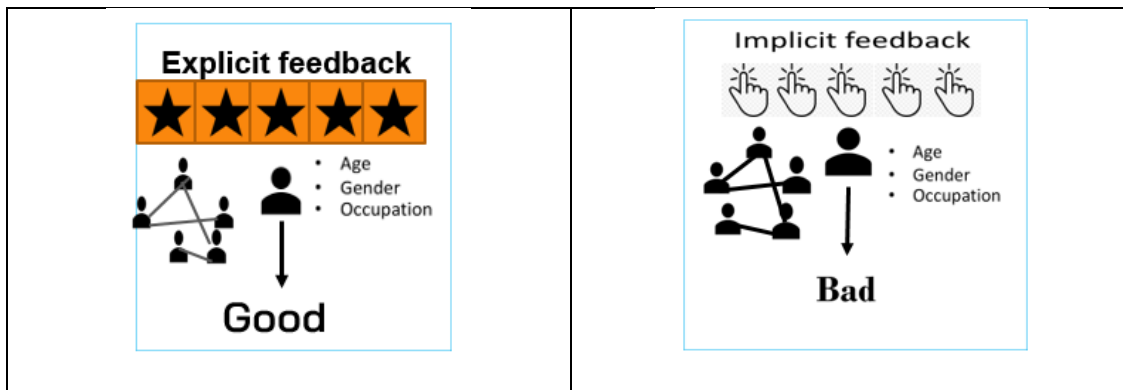
แบบที่ 2: Item-Based Filtering เป็นเทคนิคที่คำนวณจากการให้คะแนนที่ผู้ใช้กลุ่มเดียวกันให้ไว้กับรายการหรือสินค้านั้น

โดยการเรต (Rating) สินค้า หรือให้คะแนนนั้น จะมีอยู่ 3 ลักษณะคือ

1. Explicit Rating คือ การให้คะแนนรีวิวต่างๆ หรือการกดถูกใจ (Like) เช่น ผู้ใช้หรือลูกค้าซื้อสมชายให้คะแนนรีวิวสินค้าผงซึกฟอกยี่ห้อ AA ด้วยคะแนน 9/10 คะแนน เราก็จะสามารถรู้ได้ว่าคุณสมชายนั้นค่อนข้างชอบผงซึกฟอกยี่ห้อ AA หรือ อาจจะให้คะแนนสินค้าผงซึกฟอกยี่ห้อ BB แค่ 3/10 คะแนน เราก็จะพอทราบได้ว่าคุณสมชายไม่ชอบสินค้าผงซึกฟอกยี่ห้อ BB นั้นเท่าไรเป็นต้น ดังนั้นหากแพลตฟอร์มที่ให้บริการมีระบบคะแนนรีวิวหรือการกดถูกใจ (Like) จะสามารถคัดกรองข้อมูล จัดลำดับความสำคัญผู้ใช้ได้โดยตรงว่าชอบหรือไม่ชอบอะไร แต่อย่างไรก็ตามผู้ใช้ส่วนใหญ่ไม่ได้ซื้อหรือใช้สินค้าทั้งหมดทุกยี่ห้อ และผู้ใช้ไม่ชอบให้คะแนน อาจจะทำให้ข้อมูลการเรต (Rating) สินค้ามีจำนวนน้อยมากๆ ทำให้เกิดปัญหาในเรื่องของการ Sparse และ Scalability

ปัญหา “Sparse” และ “Scalability” ในแมชชีนเลิร์นนิง (Machine Learning) คือการเกิดข้อมูลที่กระจัดกระจาย ซึ่งจะทำให้ประสิทธิภาพของอัลกอริทึม (Algorithm) การเรียนรู้ของเครื่องและความสามารถในการคำนวณการคาดการณ์ที่แม่นยำลดลง จากข้อมูลที่เบาบาง หรือจากค่าที่คาดหวังในชุดข้อมูลหายไป และกระจัดกระจาย ส่งผลให้การประมวลผลไม่แม่นยำ

2. Implicit Rating คือการเก็บข้อมูลโดยการสังเกตการกระทำของผู้ใช้ในรายการบนแพลตฟอร์ม (Platform) เช่น เก็บข้อมูลการคลิก (Click) เข้าไปดูสินค้านั้น หรือการตัดสินใจที่จะซื้อสินค้านั้น หรือการที่ผู้ใช้กด Favorite หรือการเฝ้าสังเกต พฤติกรรมจากประวัติการซื้อ การค้นหาสถานที่ และเวลาที่ใช้นบนเว็บในแต่ละแพลตฟอร์ม (Platform) เป็นต้น ซึ่งข้อมูล Implicit Rating จะถูกเก็บโดยแพลตฟอร์ม (Platform) ใหญ่เป็นปกติโดยที่ผู้ใช้หลายคนอาจไม่ทราบ ซึ่งข้อมูล Implicit Rating จะมีปริมาณข้อมูลมากเมื่อเทียบกับ Explicit rating เพราะเก็บทุกครั้งที่ผู้ใช้ Log in เข้ามาใช้บริการ (การเก็บข้อมูลเป็นไปตาม พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562) ซึ่ง Implicit Rating สามารถใช้ได้กับผู้ใช้ใหม่ที่เข้ามา ถึงแม้ว่าจะยังไม่เคยมีประวัติการซื้อมาก่อน แต่อย่างไรก็ตาม Implicit Rating เป็นข้อมูลเชิงคาดเดา อาจทำให้มีการทำนายผิดได้



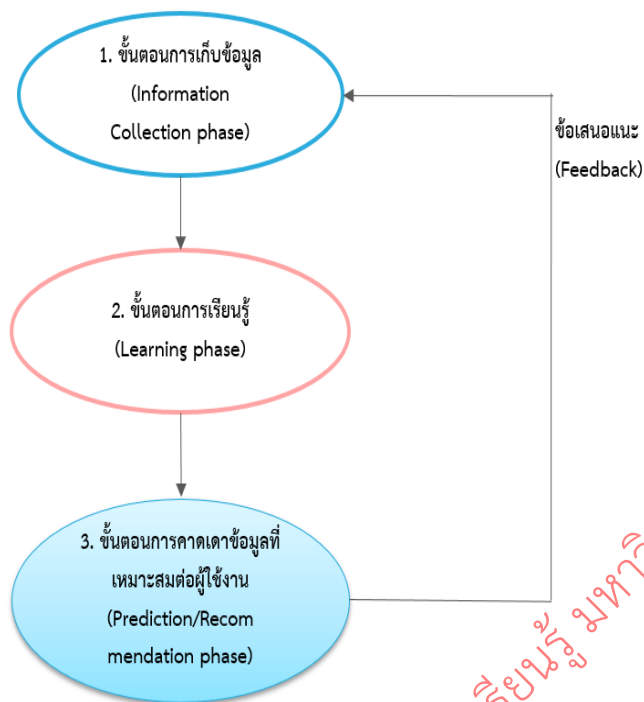
รูปที่ 9.6 Explicit Rating และ Implicit Rating

3. Hybrid Feedback เป็นการนำจุดแข็งของทั้งสองด้านมาใช้ โดยกำหนด Implicit Data เพื่อใช้เป็น Check List ว่าตรงกับ Explicit Rating ที่ลูกค้าให้คะแนนหรือไม่ ดังนั้นจึงจำเป็นต้องมีการคัดกรองข้อมูล จัดลำดับความสำคัญและส่งข้อมูลที่น่าสนใจให้ผู้ซื้อโดยตรง

จากการเติบโตทางด้านของธุรกิจอีคอมเมิร์ซ (E-Commerce) ระบบแนะนำ (Recommendation System) ถือเป็นเทคนิคที่สำคัญ คือเป็นอัลกอริทึม (Algorithm) ที่ทำหน้าที่ค้นหาและคัดกรองรายละเอียดข้อมูล จากข้อมูลที่มีอยู่จำนวนมากหาศาลเพื่อนำมาคาดเดาบริการหรือผลิตภัณฑ์ที่เกี่ยวข้อง หรือตรงกับความต้องการของลูกค้าโดยตรง

9.5.1 กระบวนการของระบบแนะนำ (Recommendation System) (Isinkaye, 2015)

การทำงานของระบบแนะนำ (Recommendation System) ประกอบไปด้วย 3 ขั้นตอนใหญ่ๆ ดังรูปด้านล่าง



รูปที่ 9.7 การทำงานของระบบแนะนำ (Recommendation System)

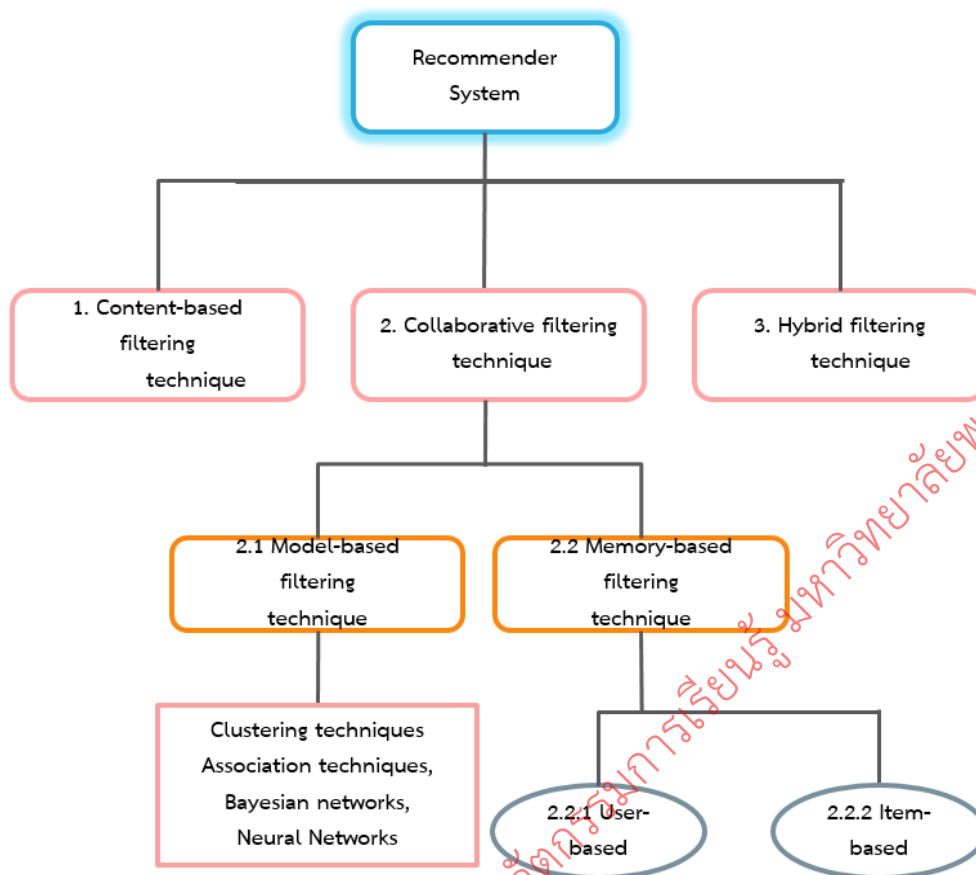
(1) Information Collection Phase: เป็นขั้นตอนการเก็บข้อมูล ที่เกี่ยวข้องจากผู้ใช้งานเพื่อสร้างโปรไฟล์ผู้ใช้งาน เช่น รายละเอียดลูกค้า พฤติกรรมการบริโภค และทรัพยากรที่ผู้ใช้งานมักใช้งานบ่อยครั้ง ระบบต้องการข้อมูลจากผู้ใช้งานมากที่สุดเท่าที่จะเป็นไปได้เพื่อสร้างระบบให้คำแนะนำได้ดีและใกล้เคียงมากที่สุดตั้งแต่ครั้งแรก โดย Recommendation Systems สามารถรับเข้าข้อมูลจากระบบการให้คะแนนและระบบการสังเกตพฤติกรรม 3 ประเภทดังที่กล่าวมาแล้ว

(2) Learning Phase: เป็นขั้นตอนการเรียนรู้ โดยสร้าง Learning Algorithm เพื่อคัดกรองข้อมูล และนำรายละเอียดที่ได้จากลูกค้า เข้ามาสู่กระบวนการบันทึกข้อมูลต่อไป

(3) Prediction / Recommendation Phase: เป็นขั้นตอนการคาดเดาข้อมูลที่เหมาะสมต่อผู้ใช้งานโดยขั้นตอนนี้สามารถดึงข้อมูลมาใช้งานโดยตรงได้จาก Memory Based หรือจาก Model Based ก็ได้ซึ่งจะกล่าวถึงในส่วนของอัลกอริทึม (Algorithm) ต่อไป

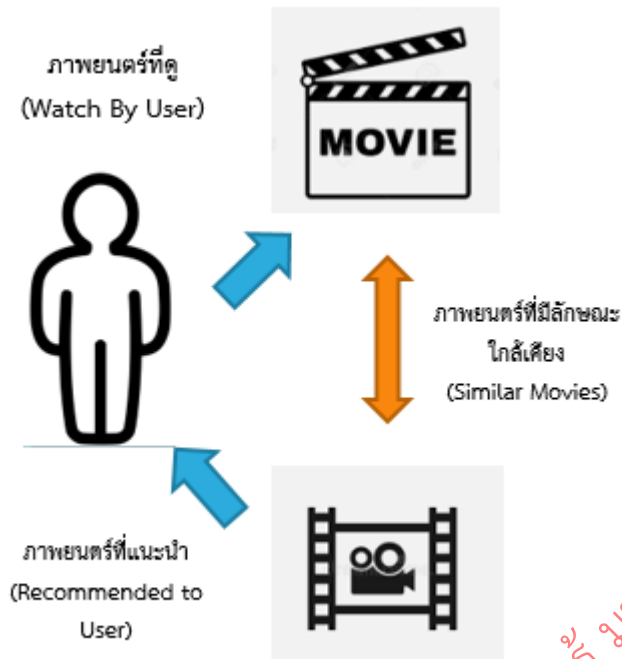
9.5.2 อัลกอริทึม (Algorithm) ในการคัดกรองข้อมูลของ Recommendation System

อัลกอริทึม (Algorithm) ในการคัดกรองข้อมูลของ Recommendation System โดยทั่วไปจะนิยมใช้กันอยู่ 3 วิธีหลักดังรูปด้านล่าง



รูปที่ 9.8 อัลกอริทึม (Algorithm) ในการคัดกรองข้อมูลของ Recommendation System

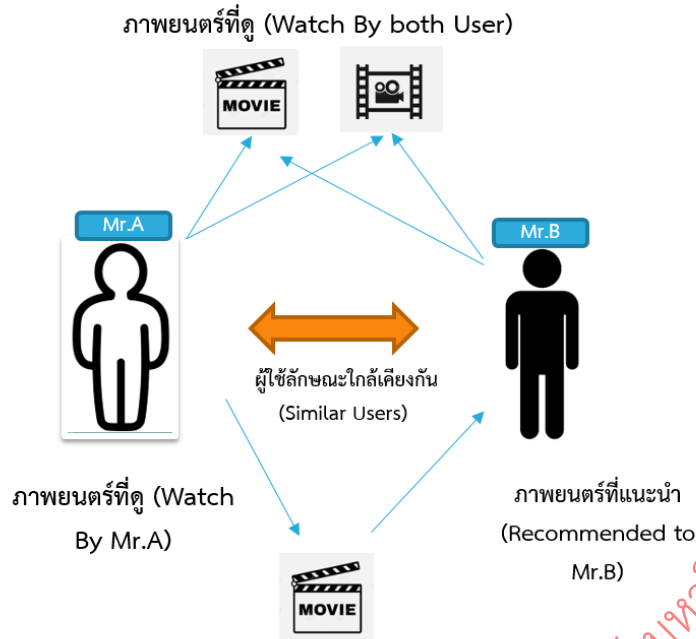
(1) Contented-Based Filtering (CBF): เป็นเทคนิคการสร้างโปรไฟล์ของผู้ใช้แต่ละคนแยกกันไป แล้วแนะนำสินค้าตามโปรไฟล์ของผู้ใช้นั้นเป็นคนๆไป โดยการแนะนำแบบพิจารณาเนื้อหา ที่อ้างอิงจากโดเมนเป็นหลัก (เน้นไปที่การวิเคราะห์รายละเอียดของสินค้าเพื่อการคาดเดา) ส่วนมากนิยมใช้ในด้านของการนำเสนอเว็บไซต์ สื่อสิ่งพิมพ์ และข่าว ดังรูป 9.13 โดยมีค่าความถูกต้องและแม่นยำของการแนะนำสินค้า ขึ้นอยู่กับการประเมินความพอใจของลูกค้า ต่อผลิตภัณฑ์นั้นๆ จากข้อมูลในอดีต (สินค้าใดที่เกี่ยวข้องกับสินค้าที่เคยถูกประเมินในเชิงบวก จะถูกนำเสนอให้ลูกค้าทราบ) โดย CBF จะใช้โมเดลประเภท Vector Space Model เช่น Term Frequency, Inverse Document Frequency หรือโมเดลเชิงสถิติเช่น Naïve Bayes Classifier, Decision Trees, Neural Networks เพื่อหาความสัมพันธ์ระหว่างสินค้า 2 ตัว (กล่าวคือ เป็นการใช้ศาสตร์ทางด้าน Statistical Analysis และ Machine Learning มาช่วยในการสร้างโมเดล)



รูปที่ 9.9 Contented-Based Filtering (CBF)

ข้อดีของวิธีการนี้คือ ลูกค้าสามารถได้รับการ Recommend โดยไม่จำเป็นต้องแชร์ประวัติส่วนตัว และเมื่อโปรไฟล์ของลูกค้าเปลี่ยนไป การแนะนำสินค้าให้ลูกค้าหรือ Recommendations สามารถเปลี่ยนตามได้อย่างรวดเร็ว แต่ข้อเสียคือจำเป็นต้องมีข้อมูลเชิงลึกหรือ Metadata ของตัวสินค้าในฐานข้อมูล ตัวอย่าง CBF ได้แก่ เว็บไซต์ LIBRA สำหรับแนะนำหนังสือ โดยใช้ Naïve Bayes Classifier คัดกรองรายชื่อหนังสือตามโปรไฟล์ของลูกค้าเพื่อจัดอันดับรายชื่อหนังสือที่น่าสนใจสำหรับลูกค้าแต่ละราย ระบบยังอธิบายว่าอะไรเป็นสาเหตุให้หนังสือที่แนะนำนั้นมีคะแนนสูงสุด เพื่อให้ลูกค้าเชื่อมั่นต่อระบบของเว็บไซต์

(2) Collaborative Filtering (CF): การแนะนำแบบพึ่งพาผู้ใช้ร่วม โดยเก็บข้อมูลจากผู้ใช้คนอื่นๆมาช่วย วิธีนี้เป็นการคัดกรองโดยไม่พึ่งโดเมน ใช้สำหรับคอนเทนต์ (Content) ที่ไม่สามารถอธิบายได้ด้วย Metadata เพียงอย่างเดียว เช่น ภาพยนต์ และเพลง โดย CF จะสร้างตารางข้อมูล (User-Item Matrix) ที่เกี่ยวพัน กับประเภทสินค้า และผู้ใช้งาน ซึ่งหลักการสำคัญของวิธีการนี้ คือการจับกลุ่มโปรไฟล์ของลูกค้าที่มีความสนใจคล้ายๆ กัน (เรียกว่า Neighbourhood) ทำให้ลูกค้าสามารถได้รับการแนะนำ (Recommend) สินค้าตัวที่เขาไม่เคยให้คะแนนมาก่อน แต่รับข้อมูลความสนใจมาจากลูกค้าคนอื่นที่อยู่ใน Neighbourhood เดียวกันแทนดังรูป รูปที่ 9.10



รูปที่ 9.10 Collaborative Filtering (CF)

โดยเทคนิคนี้สามารถแบ่งย่อยเทคนิคได้อีก 2 แบบ คือ

(2.1) Memory Based Techniques (หรือ Heuristic-Based Techniques)

ปัจจัยสำคัญของวิธีการนี้คือการนำข้อมูลที่เคยผ่านการให้คะแนนจากผู้ใช้งานมาแล้ว เพื่อหาความเป็น Neighbourhood กัน ซึ่งจะเรียกเฉพาะเจาะจงตามสิ่งที่นำมาเป็นข้อมูล ดังนี้

- User-Based Filtering (ระบบแนะนำแบบคัดกรองผู้ใช้ร่วม) คือ การหาค่าความเหมือนกันของผู้ใช้งาน 2 คน ที่ให้คะแนนไอเทม (Item) ขึ้นเดียวกัน จากนั้นจึง คำนวณหาเฉลี่ยของผู้ใช้งานทุกคนที่มีลักษณะคล้ายกันว่าผลลัพธ์ออกมาเป็นเชิงบวกหรือเชิงลบ ก่อน จะนำเสนอให้กับผู้ใช้งานหน้าใหม่ (Active User)

- Item-Based Filtering (ระบบแนะนำแบบคัดกรองสิ่งของร่วม) คือ การสร้างโมเดลขึ้นมาโมเดลหนึ่งที่รวบรวมข้อมูลในการให้คะแนนสินค้าของผู้ใช้งาน คนปัจจุบันหนึ่ง ท่าน และหาความเหมือนกันของสินค้าแต่ละชิ้น โดยใช้การคำนวณทางคณิตศาสตร์ จากสูตร Correlation หรือสูตร Cosine และใช้ Pearson Correlation Coefficient เพื่อหาน้ำหนักว่ามีค่า ความเหมือนกัน (Similarity) มากน้อยแค่ไหน

(2.2) Model Based Techniques โมเดลนี้จะใช้ Machine learning และ Data Mining ในการพัฒนาระบบคัดกรองข้อมูล โดยจุดเด่นของวิธีการนี้คือ การได้ผลลัพธ์ที่รวดเร็ว เนื่องจากมีโมเดลการคำนวณให้เสร็จสรรพเรียบร้อยแล้ว ตัวอย่างเทคนิค เช่น การใช้ Dimensionality

Reduction Technique จำพวก Singular Value Decomposition (SVD), Matrix Completion Techniques, Latent Semantic Methods, Regression และ Clustering เป็นต้น

(3) Hybrid filtering: วิธีการกรองข้อมูลแบบผสม คือการใช้ Algorithm หลายอันรวมกันเพื่อปิดจุดอ่อนของ Algorithm อันใดอันหนึ่ง แต่อาจทำให้มีความซับซ้อนและใช้ทรัพยากรในการแนะนำสูง โดยในหลายผลงานวิจัยแสดงให้เห็นว่าวิธีนี้มีความแม่นยำมากกว่า Content-Based และ Collaborative Filtering แบบปกติโดยหลักการทำให้ Hybrid สามารถทำได้หลายวิธี เช่น

(3.1) การหาผลลัพธ์ของ Algorithm แต่ละอันแล้ว (แยกส่วน) แล้วจึงนำผลลัพธ์ที่ได้จากทั้ง 2 วิธี มาวิเคราะห์เพื่อหาคำตอบสุดท้ายอีกครั้ง หรือ

(3.2) การนำคุณสมบัติเฉพาะบางอย่างของวิธี Content-Based Filtering มาใช้ประโยชน์ในวิธี Collaborative Approach ด้วย

(3.3) ในทางกลับกัน สามารถใช้ Collaborative Filtering ในแบบ Content-Based Approach ก็ได้เช่นกัน (Ratawan & Nivet, 2019)

9.6 ความรู้เบื้องต้นเกี่ยวกับเทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

โครงข่ายประสาทเทียม (Artificial Neural Network :ANN) คือ โมเดลแบบจำลองทางคณิตศาสตร์ที่มีลักษณะเป็นโครงข่ายประสาทเทียมซึ่งมีแนวคิดมาจากการทำงานสมองของมนุษย์ที่ใช้ความสามารถในการเรียนรู้จากประสบการณ์ ซึ่งประกอบด้วยเซลล์ประสาท (Neurons) และจุดประสานประสาท (Synapses) จำนวนมากทำงานเชื่อมโยงกัน โดยองค์ประกอบของ ANN ดังรูป 9.15 ประกอบไปด้วย

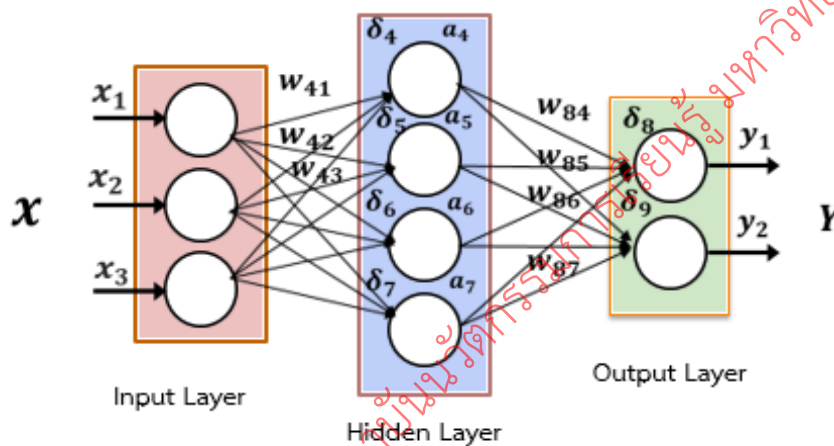
(1) Input Layer เป็นส่วนของข้อมูล Input จำนวนของ Node ขึ้นอยู่กับจำนวนของ Attribute ที่นำเข้ามาคิดในโมเดล และจะมีการกำหนดค่าน้ำหนัก (Weight) ของแต่ละ Attribute โดยเมื่อเริ่มต้นจะเป็นการ Random ขึ้นมา จากนั้นเมื่อโมเดลมีการเรียนรู้ก็จะมีการปรับค่า Weight เพื่อให้ได้คำตอบที่ใกล้เคียงที่สุด

(2) Hidden Layer เป็นชั้นที่อยู่ระหว่างกลาง ซึ่งจะทำหน้าที่คำนวณข้อมูลที่ได้รับมาจาก Layer ก่อนหน้า โดยวิธีการทางคณิตศาสตร์และมี Nonlinearity Function เช่น Sigmoid, Hyperbolic Tangens, Hard Limit, Rectified Linear Unit เป็นต้น เพื่อสร้างเป็น Information ส่งไปให้ Layer ถัดไป โดยผู้สร้างโมเดลสามารถกำหนดจำนวนชั้นของ Hidden Layer ได้ ซึ่งการเพิ่มชั้นและจำนวน Neuron ก็จะส่งผลต่อการทำงานและประสิทธิภาพในการเรียนรู้ของโมเดล ในกรณีที่มีจำนวน Hidden Layer มากกว่า 1 ชั้น จะเรียกโมเดลนั้นเป็น Deep Learning

(3) Output Layer เป็นส่วนของข้อมูลที่ได้จากการคำนวณ จำนวนของ Node ขึ้นอยู่กับรูปแบบของ Output ที่ต้องการ (Siravit, 2021)

การทำงานของเทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network :ANN) คือเริ่มจากมี Input เข้ามายัง Network ก็เอา Input มาคูณกับ Weight ของแต่ละขา ผลที่ได้จาก Input ทุก ๆ ขาของ Neuron จะเอามารวมกันแล้วก็เอามาเทียบกับ Threshold ที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่า Threshold แล้ว Neuron ก็จะส่ง Output ออกไป Output นี้ก็จะถูกส่งไปยัง Input ของ Neuron อื่น ๆ ที่เชื่อมกันใน Network ถ้าค่าน้อยกว่า Threshold ก็จะไม่เกิด Output เขียนออกมาได้ดังนี้

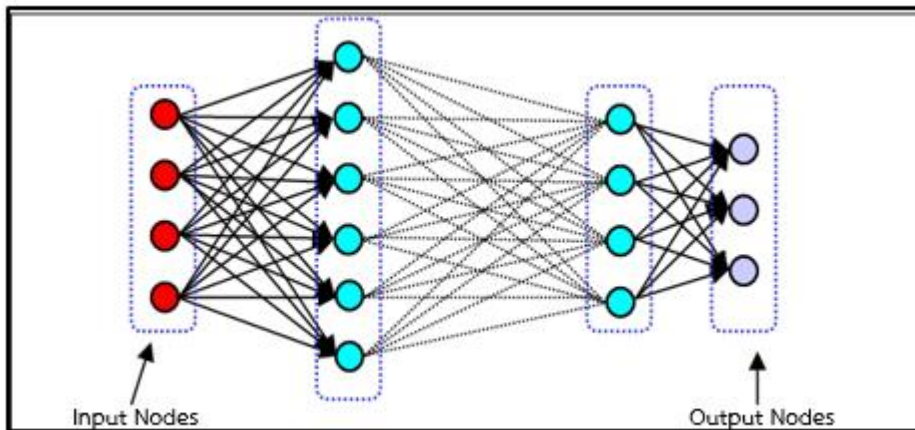
if (sum(input * weight) > threshold) then output



รูปที่ 9.11 องค์ประกอบของโครงข่ายประสาทเทียม (Artificial Neural Network)

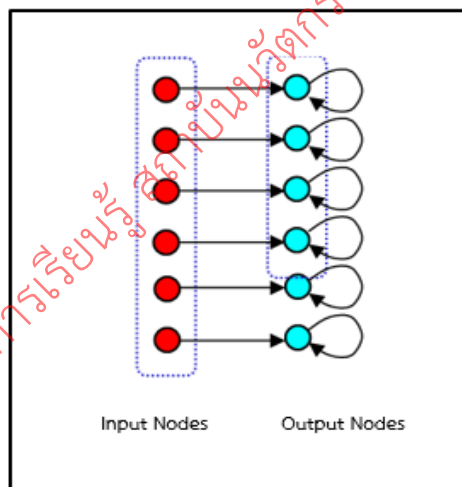
9.6.1 ประเภทของโครงข่ายประสาทเทียม (Artificial Neural Network)

(1) แบบป้อนไปข้างหน้า (Feed Forward) อาจจัดได้เป็นสองแบบย่อย คือ แบบมีชั้นของเซลล์ประสาทชั้นเดียวและแบบมีชั้นของเซลล์ประสาทหลายชั้น โดยปกติแล้วการเชื่อมโยงจะถูกกำหนดขึ้นระหว่างชั้นที่ติดกันโดยจะมีการเชื่อมโยงระหว่างเซลล์ประสาทเทียมทุกตัวจากชั้นหนึ่ง ๆ ไปยังเซลล์ประสาทเทียมทุกตัวในชั้นต่อไป ในบางสถาปัตยกรรมอาจมีการเชื่อมโยงข้ามชั้นก็ได้ดังรูป 9.12



รูปที่ 9.12 โครงข่ายประสาทเทียม (Artificial Neural Network) แบบป้อนไปข้างหน้า (Feed Forward)

(2) แบบมีการป้อนไปเวียนกลับ (Recurrent) ในสถาปัตยกรรมบางแบบโครงข่ายประสาทเทียมอาจมีการเชื่อมโยงที่ถูกระบุขึ้นระหว่างเซลล์ประสาทเทียมในชั้นหนึ่ง ๆ ย้อนกลับไปยังชั้นอื่น ๆ ก่อนหน้านั้นหรือแม้แต่ภายในชั้นเดียวกันเองดังรูป 9.13



รูปที่ 9.13 โครงข่ายประสาทเทียม (Artificial Neural Network) แบบป้อนไปข้างหน้า (Feed Forward)

การประยุกต์ใช้งานโครงข่ายประสาทเทียม (Artificial Neural Network) (Vanshika, 2022)

เนื่องจากโครงข่ายประสาทเทียม (Artificial Neural Network) มีความซับซ้อนจากข้อมูลที่ป้อนให้เรียนรู้ การประยุกต์ใช้โครงข่ายประสาทเทียม (Artificial Neural Network) จึงเป็นทางเลือกใหม่ในการควบคุม ซึ่งมีผู้นำมาประยุกต์ใช้งานหลายประเภท ได้แก่

1. ใช้ในงานการจัดจำรูปแบบที่มีความไม่แน่นอน เช่นระบบจดจำใบหน้า จับคู่ใบหน้ามนุษย์ และเปรียบเทียบกับภาพดิจิทัล ใช้ในระบบตรวจสอบใบหน้ามนุษย์ รวมถึงระบบตรวจสอบลายมือลายเซ็น และตัวอักษร

2. ใช้ในการทำนายเส้นทางการขนส่ง การวิเคราะห์การโจมตีด้วยอาวุธ และการทำนายตำแหน่งของวัตถุ และการควบคุมโดรน (Drone) อัตโนมัติ

3. ใช้เพื่อศึกษาพฤติกรรมของผู้ใช้โซเชียลมีเดีย (Social Media) โดยใช้ข้อมูลที่แชร์ทุกวันผ่านการสนทนา เพื่อการวิเคราะห์พฤติกรรมของคนในสังคมออนไลน์

4. ใช้ในการทำนาย เช่นพยากรณ์อากาศ พยากรณ์หุ้น โดยใช้เทคนิค “Multilayer Perceptron: MLP” โดยทำนายผลการเปลี่ยนแปลงราคาที่ผ่านมาของหุ้น ทำนายผลตอบแทนประจำปี เป็นต้น

5. การประยุกต์ใช้ขายงานระบบประสาทควบคุมกระบวนการทางเคมีโดยวิธีพยากรณ์แบบจำลอง (Model Predictive Control) และใช้ในการทำนายสิ่งแวดล้อมที่เปลี่ยนแปลงอยู่เสมอ (วงจรข่ายโครงข่ายประสาทเทียม (Artificial Neural Network) สามารถปรับตัวเองได้)

สรุปเนื้อหาบทที่ 9 "เทคนิคและอัลกอริทึมสำหรับการทำเหมืองข้อมูล" เป็นการนำเสนอเนื้อหาเกี่ยวกับเทคนิคและอัลกอริทึมที่ใช้ในกระบวนการทำเหมืองข้อมูล เพื่อวิเคราะห์และค้นหาความรู้จากข้อมูล

- 1. เทคนิคในการทำเหมืองข้อมูล:** นำเสนอเทคนิคต่างๆ ที่ใช้ในการทำเหมืองข้อมูล เช่น การตัดสินใจ (Decision Trees) การแบ่งกลุ่ม (Clustering) การจำแนก (Classification) และการทำนาย (Prediction) เพื่อให้เหมาะสมกับวัตถุประสงค์ของการวิเคราะห์
- 2. อัลกอริทึมที่ใช้ในการทำเหมืองข้อมูล:** อธิบายและสาธิตตัวอย่างของอัลกอริทึมที่ใช้ในการทำเหมืองข้อมูล เช่น อัลกอริทึม Apriori สำหรับการค้นหาความสัมพันธ์ในข้อมูลที่เกิดขึ้นบ่อยๆ และอัลกอริทึม K-Means สำหรับการแบ่งกลุ่มข้อมูล
- 3. การปรับแต่งและการค้นหาพารามิเตอร์:** อธิบายกระบวนการปรับแต่งและการค้นหาพารามิเตอร์ในอัลกอริทึมเพื่อให้การทำเหมืองข้อมูลมีประสิทธิภาพและความแม่นยำสูงสุด
- 4. การจัดการกับข้อมูลที่ขาดหาย:** อธิบายเทคนิคที่ใช้ในการจัดการกับข้อมูลที่ขาดหาย ซึ่งอาจเป็นการแทนค่าที่ขาดหายด้วยค่าเฉลี่ย ค่ามากที่สุด หรือค่าน้อยสุด
- 5. การประเมินและการทดสอบ:** นำเสนอเนื้อหาเกี่ยวกับวิธีการประเมินความสามารถของอัลกอริทึมในการทำเหมืองข้อมูล รวมถึงการทดสอบเพื่อวัดประสิทธิภาพและความเหมาะสมของอัลกอริทึม

6. การค้นหาความสัมพันธ์และข้อบ่งชี้: อธิบายเทคนิคในการค้นหาความสัมพันธ์และข้อบ่งชี้ที่ได้จากการทำเหมืองข้อมูล เพื่อใช้ในการสร้างความรู้และการตัดสินใจ

โดยสรุปบทที่ 9 เป็นการนำเสนอเนื้อหาเกี่ยวกับเทคนิคและอัลกอริทึมที่ใช้ในกระบวนการทำเหมืองข้อมูล เพื่อวิเคราะห์และค้นหาความรู้จากข้อมูล ซึ่งเป็นขั้นตอนสำคัญในการนำข้อมูลมาใช้ในการตัดสินใจและวางแผนในองค์กร

คำศัพท์ บทที่ 9

Artificial Neural Network (ANN)	โมเดลที่มีลักษณะเป็นโครงข่ายประสาทเทียมซึ่งมีแนวคิดมาจากการทำงานสมองของมนุษย์ ซึ่งประกอบด้วยเซลล์ประสาท (Neurons) และจุดประสานประสาท (Synapses) จำนวนมากทำงานเชื่อมโยงกัน
Association Rules	กฎความสัมพันธ์
Class Label	คุณลักษณะหนึ่ง ซึ่งบอกค่าประเภทของตัวอย่างนั้น
Classification	การจำแนกข้อมูล
Classification	การจำแนกข้อมูล
Collaborative Filtering (CF)	การแนะนำแบบพึ่งพาผู้ใช้ร่วม
Contented-Based Filtering (CBF)	การแนะนำแบบพิจารณาเนื้อหา
Convolution Neural Networks (CNN)	โครงข่ายประสาทเทียม (CNN)
Credit Approval	การขออนุมัติคำขอมีบัตรเครดิต
Decision Tree	เทคนิคการทำเหมืองข้อมูลแบบต้นไม้ตัดสินใจ - ใช้ตรวจสอบข้อมูลและสร้างต้นไม้เพื่อการพยากรณ์
Deployment	การจัดการนำไปใช้
Explicit Rating	ผลตอบรับแบบชัดเจน / โดยตรง
Feed Forward	แบบป้อนไปข้างหน้า
Frequent Item set	รูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อยๆ ในฐานข้อมูล
Hierarchical Algorithms	อัลกอริทึมแบบวิธีการแบบลำดับชั้น
Hybrid Feedback	ผลตอบรับแบบทั้งสองด้านมาใช้ โดยกำหนด Implicit Data เพื่อใช้เป็น Check List ว่าตรงกับ Explicit Rating ที่ลูกค้าให้คะแนนหรือไม่
Hybrid filtering	วิธีการกรองข้อมูลแบบผสม
Hybrid Recommendation Systems	ระบบแนะนำแบบผสมผสาน
Implicit Rating	ผลตอบรับแบบไม่ชัดเจน
Information Collection Phase	ขั้นตอนการเก็บข้อมูล
Item-Based Filtering	ระบบแนะนำแบบคัดกรองสิ่งของร่วม

Learning Phase	ขั้นตอนการเรียนรู้ เป็นการสร้าง
Logistic Regression	เทคนิคการทำเหมืองข้อมูลแบบการพยากรณ์ความ
Market Basket Analysis	การวิเคราะห์ตะกร้าการซื้อ
Medical Diagnosis	การตรวจวินิจฉัย
Memory Based Techniques/ Heuristic-Based Techniques	การนำข้อมูลที่เคยผ่านการให้คะแนนจากผู้ใช้งานมาแล้ว เพื่อหาความเป็น Neighbourhood กัน ซึ่งจะเรียกเฉพาะเจาะจงตามสิ่งที่นำมาเป็นข้อมูล
Model Evaluation	การประเมินผลโมเดล
Naïve Bayesian	เทคนิคการทำเหมืองข้อมูลแบบการใช้ในการคัดแยกประเภทข้อมูลผ่านหลักความน่าจะเป็นที่อยู่บนพื้นฐานของทฤษฎี Bayes โดยเหตุการณ์เป็นอิสระต่อกัน
Neural Network	เทคนิคการทำเหมืองข้อมูลแบบโครงข่ายประสาท – ใช้โมเดลทางคณิตศาสตร์มา
Non-Hierarchical Algorithms	อัลกอริทึมแบบวิธีการแบบไม่เป็นลำดับชั้น
Prediction / Recommendation Phase	ขั้นตอนการคาดเดาข้อมูลที่เหมาะสมต่อผู้ใช้งานโดยขั้นตอนนี้สามารถดึงข้อมูลมาใช้งานโดยตรงได้จาก Memory Based หรือจาก Model Based
Recurrent	แบบมีการป้อนไปเวียนกลับ
Recurrent Neural Networks (RNN)	โครงข่ายประสาทที่เกิดซ้ำ (RNN)
Squared Euclidean Distance	ระยะทางเชิงยูคลิดยกกำลัง
Supervised Learning	การเรียนรู้แบบมีผู้สอน
Support Vector Machine	เทคนิคการทำเหมืองข้อมูลแบบการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้น Hyperplane แบ่งแยกกลุ่ม
Target Marketing	การทำตลาดลูกค้ากลุ่มเป้าหมาย
Time Series	อนุกรมเวลา
Tracking patterns	รูปแบบการติดตาม
Training Set	ชุดข้อมูลสอน
Treatment Effectiveness Analysis	การวิเคราะห์ประสิทธิผลการรักษา
Unsupervised Learning	การเรียนรู้แบบไม่มีผู้สอน
User-Based Filtering	ระบบแนะนำแบบคัดกรองผู้ใช้ร่วม

คำถามและแบบฝึกหัดท้ายบทที่ 9

1. จงอธิบายเทคนิคการจำแนกประเภท (Classification)

2. จงอธิบายประโยชน์ของเทคนิคการจำแนกประเภท (Classification)
3. จงอธิบายเทคนิคการแบ่งกลุ่ม (Clustering)
3. จงอธิบายเทคนิคกฎความสัมพันธ์ (Association Rules)
4. จงอธิบายเทคนิคระบบแนะนำ (Recommendation System)
5. จงอธิบายปัญหา “Sparse” และ “Scalability” ในแมชชีนเลิร์นนิง (Machine Learning)
6. จงอธิบาย Implicit Rating
7. จงอธิบาย 3 กระบวนการของระบบแนะนำ (Recommendation System)
7. จงอธิบายเทคนิค Contented-Based Filtering (CBF)
 8. จงอธิบาย 3 กระบวนการของระบบแนะนำ (Recommendation System)
 9. จงอธิบายเทคนิค Contented-Based Filtering (CBF)
 10. จงอธิบายเทคนิค Collaborative Filtering (CF)
 11. จงอธิบายเทคนิค โครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

บทที่ 10 เครื่องมือและเทคนิคที่สำคัญสำหรับการทำเหมืองข้อมูล

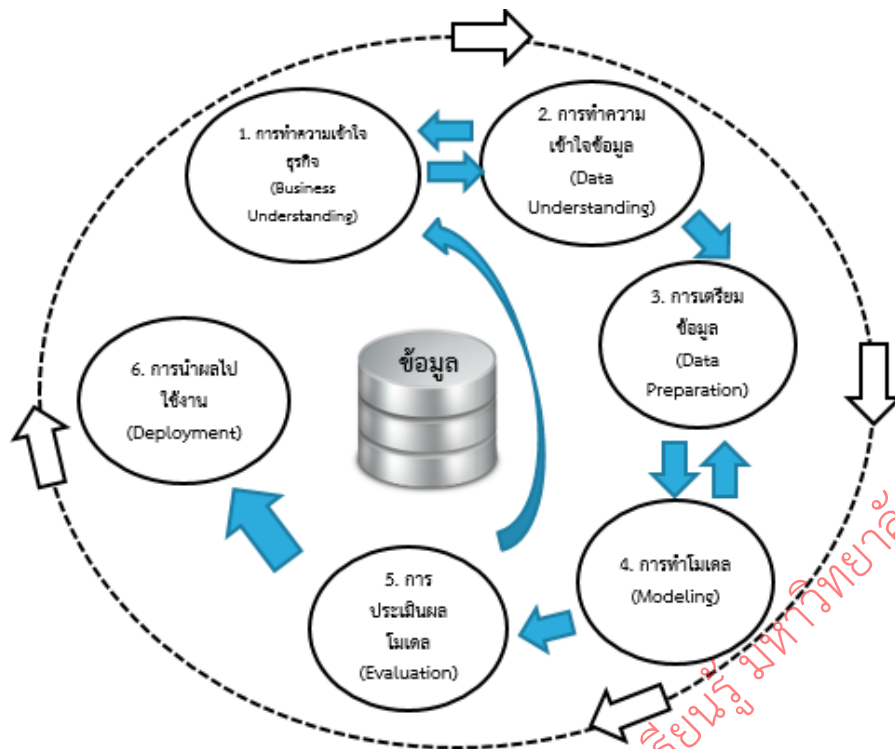
(Important Tools and Techniques for Data Mining)

ในโลกดิจิทัล (Digital World) ทุกวันนี้ ธุรกิจถูกรายล้อมไปด้วยข้อมูลขนาดใหญ่ (Big Data) ที่คาดการณ์ว่าจะเติบโต 40% ต่อปีในทศวรรษหน้า ข้อเท็จจริงคือ ธุรกิจสมัยใหม่ถูกบังคับให้เริ่มข้อมูล แต่ ข้อมูลทั้งหมดนี้สร้าง ยากต่อการขุด ซึ่งคือธุรกิจสมัยใหม่ไม่เข้าใจการค้นหาคำรู้จากข้อมูลขนาดใหญ่ (Big Data) โดยพื้นฐานแล้ว ธุรกิจสมัยใหม่ได้สร้างข้อมูลจำนวนมาก แต่ประสบปัญหาในการริเริ่มทำการขุดเหมือง ข้อมูลขนาดใหญ่ (Big Data) หาคำรู้ที่ฝังลึกอยู่ในชุดข้อมูล หากธุรกิจสมัยใหม่ไม่มีเครื่องมือหรือเทคนิคที่มีประสิทธิภาพในการขุดข้อมูลดังกล่าว ก็เป็นไปได้ที่จะได้รับประโยชน์จากข้อมูลขนาดใหญ่ (Big Data) ดังกล่าว (Hansen & Birkinshaw, 2007)

วัตถุประสงค์ของบทนี้ เพื่อให้ผู้อ่านมีความสามารถดังต่อไปนี้คือ

1. สามารถเข้าใจมุมมองที่ต้องทำความเข้าใจ 5 มิติสำหรับการทำเหมืองข้อมูล (Data Mining)
2. สามารถอธิบายขั้นตอนการประเมินประสิทธิภาพของโมเดล ในการทำเหมืองข้อมูล (Data Mining)
3. รู้จักเครื่องมือและซอฟต์แวร์ที่ใช้สำหรับการทำเหมืองข้อมูล (Data Mining)

ความสำคัญของระบบธุรกิจอัจฉริยะ (Business Intelligence) และการวิเคราะห์ข้อมูล (Data Analytics) นั้นคงกลายเป็นสิ่งที่องค์กรต่างๆ ไม่สามารถปฏิเสธได้อีกต่อไป เพราะทุกวันนี้การแข่งขันทางธุรกิจไม่ว่าจะเป็นเชิงกลยุทธ์ การตลาด เทคโนโลยีหรือแม้แต่การคัดเลือกพนักงาน ต่างก็ต้องอาศัยการตัดสินใจที่แม่นยำจากการวิเคราะห์ข้อมูลและสถิติเพื่อประกอบการตัดสินใจ ดังนั้นเพื่อให้การพัฒนาองค์กรธุรกิจสมัยใหม่เป็นไปอย่างมีแบบแผนและสามารถบูรณาการทั้งด้านการบริหารจัดการ กระบวนการและเทคโนโลยีดิจิทัลให้สอดคล้องกับวิสัยทัศน์ วัตถุประสงค์เชิงธุรกิจ จำเป็นต้องอาศัยเครื่องมือและเทคนิคที่สำคัญสำหรับการวิเคราะห์ธุรกิจ โดยเริ่มต้นจากกระบวนการวิเคราะห์ข้อมูล CRISP-DM ย่อมาจาก Cross-industry Standard Process for Data Mining มาตรฐานซึ่งเป็นเหมือนพิมพ์เขียว (Blueprint) ที่ใช้กันอย่างกว้างขวาง เช่นเดียวกันกับกระบวนการ ISO : International Organization for Standardization ในโรงงานอุตสาหกรรม หรือกระบวนการ CMMI : Capability Maturity Model Integration ซึ่งเป็นมาตรฐานในการพัฒนาซอฟต์แวร์ โดยในกระบวนการ CRISP-DM นี้จะประกอบด้วย 6 ขั้นตอนดังในรูป 10.1



รูปที่ 10.1 โมเดลกระบวนการวิเคราะห์ข้อมูล (The CRISP-DM model) (Chapman et al., 1999)

ทั้งนี้ขั้นตอนการทำเหมืองข้อมูล (Data Mining) ที่นิยมใช้คือกระบวนการวิเคราะห์ข้อมูล (Cross-Industry Standard Process for Data Mining : CRISP-DM) ซึ่งใช้ในการแปลงปัญหาทางธุรกิจให้เป็นขั้นตอนต่างๆ และแยกออกจากส่วนของการพัฒนาโปรแกรมประยุกต์ (Application) และการพัฒนาแผนเทคโนโลยีในองค์กร โดยมี 6 ขั้นตอนตามที่ได้กล่าวไว้เบื้องต้นในบทที่ 8 (ตาราง 8.1) โดยกระบวนการวิเคราะห์ข้อมูล (Cross-Industry Standard Process for Data Mining : CRISP-DM) คือ กระบวนการแปลงปัญหาทางธุรกิจให้เป็นขั้นตอนต่างๆของการทำเหมืองข้อมูล (Data Mining) และแยกออกจากส่วนของการเขียนโปรแกรมประยุกต์ (Application) และการใช้เทคโนโลยี ประกอบด้วย 6 กระบวนการคือ

1. การทำความเข้าใจธุรกิจ (Business Understanding)
2. การทำความเข้าใจข้อมูล (Data Understanding)
3. การเตรียมข้อมูล (Data Preparation)
4. การทำโมเดล (Modeling)
5. การประเมินผลโมเดล (Evaluation)
6. การนำผลไปใช้งาน (Deployment)

โดยแต่ละขั้นตอนจะมีความต่อเนื่องกันคือขั้นตอนถัดไปจะรอผลลัพธ์จากขั้นตอนก่อนหน้าซึ่งแสดงด้วยลูกศรที่เชื่อมระหว่างวงกลม ตัวอย่างเช่นเมื่อได้ผลลัพธ์จาก ขั้นตอนที่หนึ่งการทำความเข้าใจธุรกิจ (Business Understanding) ซึ่งเป็นการรวบรวมข้อมูลพื้นฐานของสภาพธุรกิจในปัจจุบัน การจัดหาจุดประสงค์ทางธุรกิจที่เฉพาะเจาะจง และการหาเกณฑ์ที่ใช้ในการวัดความสำเร็จของการทำเหมืองข้อมูลจากมุมมองเชิงธุรกิจ ตัวอย่างเช่น วัดความสำเร็จจากการลดปัญหาที่เกิดขึ้นนับเป็นจำนวนเคส หรือการลดอุปสรรคหรือข้อจำกัดของโครงการ หรือวัดความสำเร็จจากผลกำไรและความยั่งยืนของธุรกิจเป็นตัวเลขที่วัดได้ จากนั้นถึงดำเนินการขั้นตอนการทำความเข้าใจข้อมูล (Data Understanding) และการเตรียมข้อมูล (Data Preparation) ซึ่งเป็นการจัดเก็บข้อมูลเบื้องต้นทั้งที่มาจากข้อมูลเดิม ข้อมูลสำรวจ หรือเริ่มติดตามข้อมูลเพิ่มเติมแล้วจึงนำไปสร้างโมเดล (Modeling) หรือการวิเคราะห์ด้วยอัลกอริทึม (Algorithm) โดยกระบวนการในขั้นตอน Modeling นั้นอาจจะย้อนกลับมาเปลี่ยนแปลงข้อมูลให้ถูกต้องมากขึ้นเพื่อหวังว่าจะได้โมเดลที่ให้ความถูกต้องมากขึ้นก็ได้ เป็นต้น จากนั้นทำการประเมินผลโมเดล (Evaluation) เป็นการประเมินผลการวิเคราะห์กับเกณฑ์ที่ตั้งไว้ และวัดผลที่ได้ โดยจะได้แบบจำลองที่คัดเลือก และข้อสรุปหรือสิ่งที่ได้จากแบบจำลองเพื่อนำไปเขียนโปรแกรมหรือใช้ทำนายผลทางธุรกิจ และขั้นตอนสุดท้ายคือการนำผลไปใช้งาน (Deployment) คือการใช้องค์ความรู้ใหม่ที่ค้นพบมาพัฒนาองค์กร หรือการแสดงถึงการเปลี่ยนแปลงที่เกิดขึ้นในองค์กรธุรกิจสมัยใหม่ได้และวางแผนควบคุมการนำผลไปใช้งาน รวมถึงเขียนสรุปและทบทวนโครงการ

10.1 การประยุกต์ใช้โมเดลกระบวนการวิเคราะห์ข้อมูล (The CRISP-DM model)

ในส่วนรายละเอียดของโมเดลกระบวนการวิเคราะห์ข้อมูล (The CRISP-DM model) ในขั้นตอนที่หนึ่งการทำความเข้าใจธุรกิจ (Business Understanding) จะประกอบไปด้วยการมุมมองที่ต้องทำความเข้าใจใน 5 มิติดังนี้

10.1.1 ทำความเข้าใจวัตถุประสงค์เชิงธุรกิจ (Business Objectives) โดยธุรกิจอาจจะกำหนดเป้าหมายผลตอบแทนจาก ด้านการลดต้นทุนขององค์กรและต้นทุนด้านเทคโนโลยีดิจิทัล จากเป้าหมายด้านการปรับปรุงกระบวนการระหว่างฝ่ายและหน่วยธุรกิจ (Cross Functional Process Improvement) ซึ่งจะต้องประเมินด้วยผลลัพธ์ในด้านผลกำไร (Bottom Line) ความพึงพอใจของลูกค้า (Customer Satisfaction) หรือการเพิ่มประสิทธิภาพในการใช้ทรัพยากรขององค์กร (Efficiency in Resource Utilization) รวมถึงการเพิ่มประสิทธิภาพในการปฏิบัติงาน (Operational Efficiency) หรือการเพิ่มความสามารถในการแข่งขัน อาทิ ความรวดเร็วในการผลิตสู่การตลาด (Speed to Market) การสร้างเครือข่าย การทำงานร่วมกันระหว่างองค์กรพันธมิตร (Collaborative Commerce) และการเพิ่มความคล่องแคล่วว่องไวในการดำเนินธุรกิจ (Strategic Agility) เป็นต้น แต่

โดยทั่วไปแล้วธุรกิจจะกำหนดเป้าหมายผลตอบแทนจากรายได้ กำไร (Bottom Line) และลูกค้าเป็นหลัก

10.1.2 ทำความเข้าใจลำดับความสำคัญของการลงทุน (Funding Priorities) เป็นการรวบรวมข้อมูลการลงทุนในระบบระดับองค์กร (Enterprise System) ที่มีการบูรณาการด้านเทคโนโลยีและการร่วมใช้งานทั้งด้านข้อมูล (Shared Data) และด้านโครงสร้างพื้นฐาน (Shared Infrastructure) มากยิ่งขึ้นเพื่อตอบโจทย์ในเชิงการจัดการองค์กรทั้งด้านการปรับปรุงประสิทธิภาพการเพิ่มขีดความสามารถในการแข่งขัน และการตอบสนองต่อกลยุทธ์ในการดำเนินธุรกิจสมัยใหม่

10.1.3 ทำความเข้าใจขีดความสามารถในการบริหารจัดการ (Management Capabilities) คือการรวบรวมข้อมูลขีดความสามารถในการบริหารจัดการกระบวนการระหว่างฝ่ายและหน่วยธุรกิจ (Cross Functional Process Improvement) ในด้านการบริหารจัดการมาตรฐานของงานเทคโนโลยีดิจิทัล (Technology Standardization) การบริหารจัดการมาตรฐานข้อมูล (Data Standardization) และการบริหารจัดการมาตรฐานกระบวนการ (Process Standardization) รวมถึงความสามารถในการพัฒนาระบบเทคโนโลยีดิจิทัล การบำรุงรักษาระบบ ขององค์กรธุรกิจ โดยสำรวจรวบรวมการวางแนวทางที่สอดคล้อง (Alignment) และความพร้อมสมบูรณ์ในด้านกระบวนการที่เป็นส่วนจำเพาะทางของกระบวนการธุรกิจในแต่ละหน่วย (Business Process Modules)

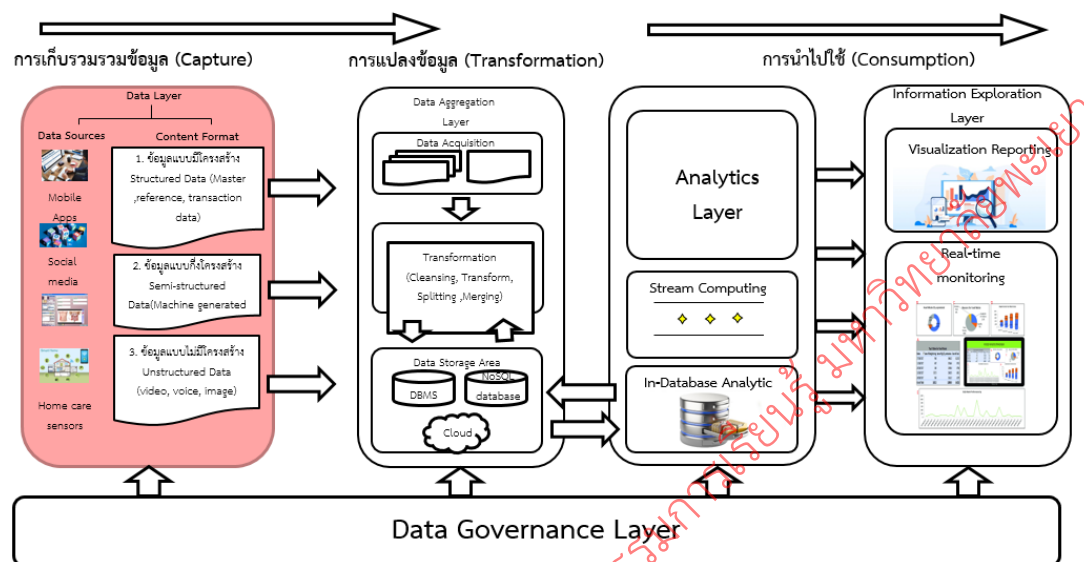
10.1.4 ทำความเข้าใจผู้รับผิดชอบในการกำหนดการใช้งานของระบบ (Who Defines Applications) ในขั้นตอนนี้ยังต้องรวบรวมผู้รับผิดชอบในการกำหนดการใช้งานของระบบ (Who Defines Applications) รวบรวมผู้ทำร่วมพัฒนางานและผู้ใช้งานที่จะต้องอยู่ในกระบวนการวิเคราะห์ข้อมูลทั้ง 6 ขั้นตอน

10.1.5 ทำความเข้าใจการกำกับดูแลด้านเทคโนโลยีดิจิทัล (Digital Governance) ประกอบไปด้วยการทำความเข้าใจความรับผิดชอบ (Responsibility) กลยุทธ์ (Strategy) การจัดซื้อจัดหา (Acquisition) ผลการดำเนินงาน (Performance) ความสอดคล้องกัน (Conformance) และพฤติกรรมบุคคล (Human Behavior) โดยมุ่งเน้นการกำหนดกรอบทิศทางการกำกับดูแลด้านการบริหารจัดการเทคโนโลยีดิจิทัลขององค์กร ให้สามารถบริหารจัดการทรัพยากรเทคโนโลยีดิจิทัลได้อย่างเหมาะสม

10.2 การทำความเข้าใจข้อมูล (Data Understanding) การเตรียมข้อมูล (Data Preparation) ในบางมาตรฐานกระบวนการการวิเคราะห์ข้อมูลอาจจะรวมไว้เป็นกระบวนการเดียวกัน ซึ่งจากการเก็บรวบรวมข้อมูลโครงการด้านการวิเคราะห์ข้อมูลขนาดใหญ่ พบว่านักวิเคราะห์ข้อมูลใช้เวลา 70-80% ในการเตรียมข้อมูลในขั้นตอนที่สองและสาม ที่เหลืออีก 20-30% คือเวลาสำหรับวิเคราะห์

ข้อมูลและกระบวนการอื่นๆ (Press, 2016) ระยะเวลาที่ใช้ในการเตรียมข้อมูลสำหรับการวิเคราะห์นั้นขึ้นอยู่กับความสมบูรณ์ของข้อมูลโดยตรงเช่น ความสมบูรณ์ของข้อมูล จำนวนค่าที่ขาดหายไป (Missing Values) ความสะอาดและความคลาดเคลื่อนเป็นอย่างไร

สำหรับที่มาของข้อมูลจะมีมาจากหลายแหล่งดังรูป 10.2



รูปที่ 10.2 มุมมองสถาปัตยกรรมฐานข้อมูลขนาดใหญ่ (Roehrs et al., 2017)

จากรูป 10.2 ในกล่องของการเก็บรวบรวมข้อมูล (Capture) ประกอบไปด้วย เลเยอร์ข้อมูล (Data Layer) ในเลเยอร์นี้ เป็นส่วนที่รวมแหล่งข้อมูลทั้งหมดที่จำเป็นเพื่อนำเข้าสู่ระบบข้อมูลขนาดใหญ่ที่จำเป็นต่อสนับสนุนการดำเนินงานประจำวันและแก้ปัญหาทางธุรกิจ ข้อมูลมีการแบ่ง หลากหลายโครงสร้างเช่น

(1) ข้อมูลแบบโครงสร้าง (Traditional DB) เช่น ข้อมูลจากระบบอีอาร์พี (Enterprise Resource Planning: ERP) ระบบลูกค้าสัมพันธ์ (Customer Relationship Management: CRM) ระบบบัญชี (Accounting) และระบบอื่น ๆ เช่น ระบบที่ใช้ในโรงพยาบาล เช่น เวชระเบียนอิเล็กทรอนิกส์ (EHRs)

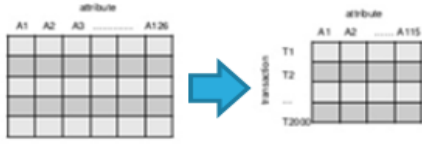
(2) ข้อมูลกึ่งโครงสร้าง เช่น ข้อมูลจากรายการคอมพิวเตอร์ (Log) จากอุปกรณ์ตรวจสอบข้อมูลที่นามสกุลเป็น XML, JSON เป็นต้น

(3) และข้อมูลที่ไม่มีโครงสร้าง เช่น ข้อมูลภาพ

ข้อมูลทางเหล่านี้จะถูกเก็บรวบรวมจาก แหล่งทั้งภายในและภายนอกองค์กร โดยขึ้นอยู่กับชนิดของข้อมูล และรูปแบบเนื้อหา ซึ่งปัญหาที่พบส่วนใหญ่จะเกิดจากการนำข้อมูลภายนอกมา และจำเป็นต้องมีการทำการเตรียมข้อมูล (Data Preparation) เช่น

- (1) มีข้อมูลบางฟิลด์ที่ล้าสมัยหรือซ้ำซ้อน (Fields that are obsolete or redundant)
- (2) ข้อมูลไม่สมบูรณ์และขาดหาย (Missing values)
- (3) ข้อมูลมีค่าผิดปกติ (Outliers) หรือ ข้อมูลมีค่าผิดพลาด (Error and Noisy data)
- (4) ข้อมูลในรูปแบบที่ไม่เหมาะกับแบบจำลอง (Model) การทำเหมืองข้อมูล (Data in a form not suitable for the data mining models)
- (5) ข้อมูลไม่สอดคล้องกัน (Values not consistent with policy or common sense, Inconsistent data) เช่น ข้อมูลเดียวกัน แต่ตั้ง ชื่อต่างกัน หรือใช้ค่าแทนข้อมูลที่แตกต่างกัน เช่น ข้อมูลผู้ใช้เกิดปี 2000 แต่ถูกบันทึกอายุเป็น 46 ปี เป็นต้น

สำหรับกล่องการแปลงข้อมูล (Transformation) จะประกอบไป 4 ขั้นตอนหลักนี้ดังรูป 10.3

(1) Data Cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่เป็นส่วนรบกวนหรือข้อมูลที่ไม่เกี่ยวข้องออกไป	
(2) Data Integration เป็นขั้นตอนการรวมแหล่งข้อมูล ซึ่งมีข้อมูลหลายแห่งมารวมไว้ที่เดียวกัน	
(3) Data Transformation เป็นขั้นตอนการแปลงข้อมูลในขั้นตอนการคัดเลือก ให้เหมาะสำหรับขั้นตอนการทำเหมืองข้อมูล	$-2, 32, 100, 59, 48 \rightarrow -0.02, 0.32, 1.00, 0.59, 0.48$
(4) Data Reduction เป็นขั้นตอนการลดมิติข้อมูล เพื่อเป็นตัวแทนจำนวนข้อมูลทั้งหมด เป็นต้น	

รูปที่ 10.3 4 เทคนิคการแปลงข้อมูล (Data Transformation)

(1) Data Cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่เป็นส่วนรบกวนหรือข้อมูลที่ไม่เกี่ยวข้องออกไป เป็นกระบวนการตรวจสอบและการแก้ไข (หรือลบ) รายการข้อมูลที่ไม่ถูกต้องออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูล ซึ่งมีเทคนิคเบื้องต้นได้แก่ การแจกแจงข้อมูล (Data Parsing) การแก้ไขข้อมูลที่ผิดพลาด (Data Correcting) การทำข้อมูลให้เป็นรูปแบบเดียวกัน (Data Standardizing) และ การลบชุดข้อมูลซ้ำซ้อนทิ้ง (Data Duplicate Elimination)

(2) Data Integration กรณีที่ระบบต้องการใช้ข้อมูลจากหลายแหล่ง หรือมีหลายแฟ้มข้อมูล ต้องทำการเชื่อมโยงข้อมูลจากหลายแหล่งเข้าด้วยกัน โดยใช้คุณลักษณะหรือแอตทริบิวต์ ที่มีอยู่ร่วมกันของหลายแหล่งข้อมูลเป็นตัวเชื่อมโยง โดยเป็นขั้นตอนการรวมแหล่งข้อมูล ซึ่งมีข้อมูลหลายแห่งมารวมไว้ที่เดียวกัน

(3) Data Transformation เป็นขั้นตอนการแปลงข้อมูลในขั้นตอนการคัดเลือก เป็นการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมสำหรับการประมวลผล และให้เหมาะสำหรับขั้นตอนการทำเหมืองข้อมูล

(4) Data Reduction เป็นขั้นตอนการลดมิติข้อมูล เพื่อเป็นตัวแทนขจำนวนข้อมูลทั้งหมด เป็นต้น

นอกเหนือจากนั้นในขั้นตอนการเตรียมข้อมูล (Data Preparation) นักวิเคราะห์อาจจะเจอปัญหาอื่นๆ อีกมากมายเช่น

(1) ปัญหาเรื่องช่องทางการเข้าถึง (Access) และสิทธิในการเข้าถึง (Authorization) ในกรณีที่ทำงานระหว่างบริษัท ระหว่างฝ่ายงาน ระหว่างหน่วยงาน ระหว่างกรม ต่างกระทรวง เป็นต้น

(2) ปัญหาเรื่องความต้องการด้านการวิเคราะห์ที่เปลี่ยนแปลงอยู่ตลอดเวลา ซึ่งเป็นปัญหาปกติของการทำงานด้านระบบสารสนเทศ ทำให้โครงการล่าช้าเพราะต้องกลับไปทำการวิเคราะห์และจัดเตรียมข้อมูลใหม่ ซึ่งในการแก้ปัญหาดังกล่าวทางเทคนิคสำหรับโครงการการทำเหมืองข้อมูลนั้น อาจจะต้องจัดเตรียมบุคลากรที่ทำหน้าที่จัดเตรียมข้อมูลโดยเฉพาะซึ่งก็คือตำแหน่ง วิศวกรข้อมูล (Data Engineer) หรือการฝึกอบรมให้กับผู้ใช้งาน (Business User) ให้มีทักษะด้านไอทีที่จำเป็นสำหรับการทำการเตรียมข้อมูลเบื้องต้นด้วยตนเองระดับหนึ่งก่อน (Self Service Data Preparation) ซึ่งในกรณีนี้ผู้เขียนเคยอบรมการเตรียมข้อมูลให้กับผู้ใช้งาน (Business User) โดยใช้โปรแกรม Excel, Tableau และ Power BI เพราะมีเครื่องมือ (Tool) เบื้องต้นสำหรับการทำ Data Transformation

สำหรับกล่องการนำไปใช้ (Consumption) จะประกอบไปด้วยเลเยอร์การวิเคราะห์ (Analytics Layer) เลเยอร์นี้มีหน้าที่ในการประมวลผลข้อมูลทุกประเภท และมีเครื่องมือวิเคราะห์

ข้อมูลขนาดใหญ่ สำหรับเลเยอร์การสำรวจข้อมูล (Information Exploration Layer) เลเยอร์นี้มีเครื่องมือสร้างรายงานผลลัพธ์ เช่น รายงานข้อมูลแบบเรียลไทม์และข้อมูลเชิงลึกทางธุรกิจ

และสำหรับกลไกการกำกับดูแลข้อมูล (Data Governance Layer) เลเยอร์นี้ประกอบด้วยการจัดการข้อมูลหลัก (Master Data Management: MDM) การจัดการวงจรข้อมูล และการจัดการความปลอดภัยของข้อมูลและความเป็นส่วนตัว

10.3 ความท้าทายในการเตรียมข้อมูล (Data Preparation)

การทำเหมืองข้อมูลจำนวนมากจะใช้เวลาพอสมควรในการเตรียมข้อมูล (Data Preparation) ผู้เชี่ยวชาญด้านข้อมูลใช้เวลามากถึง 60% ของโครงการในการทำความสะอาดข้อมูลสำหรับการวิเคราะห์ ซึ่งเป็นการสิ้นเปลืองทรัพยากรอย่างมาก เนื่องจากทักษะและความรู้ของนักวิเคราะห์ข้อมูลสามารถสร้างมูลค่าเพิ่มได้มาก หากกิจกรรมที่มุ่งเน้นไปที่งานระดับไฮเอนด์ (Hi-end) มากขึ้น เช่นงานที่นำเสนอโมเดลเพื่อสร้างกลยุทธ์ให้กับธุรกิจ ดังนั้นการลดทีมงานให้สอดคล้องกับระยะเวลาของโครงการเป็นหนึ่งในความท้าทายที่ใหญ่ที่สุดที่องค์กรต้องเผชิญในระหว่างขั้นตอนการเตรียมข้อมูล ดังนั้นสิ่งที่ต้องตระหนักในกระบวนการของการวิเคราะห์ข้อมูลสำหรับการจัดการระบบสารสนเทศสมัยใหม่ได้แก่ (Selerity, 2022)

10.3.1 การละเลยบริบท (Neglecting context) องค์กรหลายแห่งอาจไม่มีแนวทางการปฏิบัติในการวิเคราะห์ข้อมูลที่สมบูรณ์ มักจะละเลยการกำหนด วัตถุประสงค์ของงานให้สอดคล้องกับวิสัยทัศน์ พันธกิจ หรือเป้าประสงค์ขององค์กรในขณะนั้น ทีมงานด้านการวิเคราะห์หลายครั้งใช้เวลาไปกับการจัดการชุดข้อมูลให้สมบูรณ์ โดยบางครั้งพบว่าข้อมูลที่กำลังจัดเตรียมไม่มีบริบทที่เกี่ยวข้องกับวิสัยทัศน์ พันธกิจ หรือเป้าประสงค์ขององค์กร ซึ่งการทำงานสำหรับโครงการสารสนเทศสมัยใหม่จำเป็นต้องคำนึงถึงมิติหลายๆ ด้าน ดังนั้น เขียนแนะนำให้ทุกโครงการสารสนเทศศึกษาหลักการในการพัฒนาสำหรับระบบสารสนเทศสมัยใหม่ (TOGAF) ซึ่งกล่าวไว้ในบทที่ 3

10.3.2 การระบุคุณภาพของข้อมูล (Spotting Data Quality Issues)

การสร้างแผนมาตรฐานด้านคุณภาพข้อมูล เป็นสิ่งที่ละเลยกันอย่างมากในการพัฒนาระบบสารสนเทศ ตั้งแต่การจัดเก็บ การออกแบบฐานข้อมูล การออกแบบการรับเข้าข้อมูลประเภทต่างๆ เนื่องจากองค์กรส่วนใหญ่ต้องจัดการกับข้อมูลขนาดใหญ่ทุกวัน ดังนั้นคุณภาพของข้อมูลจะต้องถูกกำหนดให้เป็นมาตรฐานในเชิงของการออกแบบสถาปัตยกรรมฐานข้อมูล ซึ่งการจัดเก็บนี้ต้องสัมพันธ์กับการออกแบบกับสถาปัตยกรรมเทคโนโลยีที่ได้วางแผนไว้ ตัวอย่างเช่น ต้องมีฮาร์ดแวร์ซอฟต์แวร์ ระบบฐานข้อมูลที่วางแผนควบคู่ไปกับประเภทของข้อมูลที่จะจัดเก็บ ขนาดและการเติบโตของข้อมูลและ นโยบายการดูแลรักษาข้อมูล (Data Governace)

โดยสรุปการแก้ปัญหาในการจัดเตรียมข้อมูล สามารถจัดการได้ หากองค์กรได้มีการรอบ และแผนการทำงานที่ดี มีการจัดสรรทรัพยากรด้านเทคโนโลยีและบุคลากร ที่เหมาะสม และเนื่องจาก การวิเคราะห์ข้อมูลมีบทบาทมากขึ้นในการจัดการธุรกิจสมัยใหม่ องค์กรต่างๆจะต้องใช้เวลาและความ พยายามในการจัดการกับความท้าทายที่ขัดขวางการเตรียมข้อมูล (Data Preparation) และการ วิเคราะห์ข้อมูล (Data Analytics) วิธีที่ดีที่สุดในการแก้ไขปัญหาความท้าทายเหล่านี้ อาจจะต้องใช้ หลายวิธีผสมผสานร่วมกัน โดยหลักสำคัญในการเตรียมข้อมูล (Data Preparation) จะต้องคำนึงถึงการ ลงทุนในเทคโนโลยีที่เหมาะสม การระดมทรัพยากรขององค์กร และการปรับโครงสร้างการดำเนินงาน ทั้งหมดเพื่อจัดความไว้ประสิทธิภาพในการเตรียมข้อมูล ผ่านการวางแผนอย่างรอบคอบ

10.4 การนำไปใช้ (Consumption)

ในขั้นตอนต่อจากการเตรียมข้อมูล (Data Preparation) จะเป็นการประมวลผลข้อมูลทุก ประเภท และมีเครื่องมือวิเคราะห์ข้อมูลขนาดใหญ่ (Analytics Layer) โดยจะมีกระบวนการ การสำรวจข้อมูล (Information Exploration Layer) มีเครื่องมือสร้างรายงานผลลัพธ์ เช่น รายงาน ข้อมูลแบบเรียลไทม์และข้อมูลเชิงลึกทางธุรกิจ ซึ่งในกระบวนการทำเหมืองข้อมูล CRISP-DM ก็คือการ ทำโมเดล (Modelling) นั่นเอง ซึ่งผู้อ่านสามารถกลับไปทบทวนเนื้อหาได้ที่บทที่ 9 ความรู้เบื้องต้น เกี่ยวกับเทคนิคและ อัลกอริทึม สำหรับการทำ เหมืองข้อมูล โดยในส่วนนี้จะกล่าวถึงการประเมิน ประสิทธิภาพของเทคนิคการทำเหมืองข้อมูล (Data Mining) หรือที่ภาษาอังกฤษเรียกว่า Model Evaluation เป็นกระบวนการที่ดำเนินการต่อจากขั้นตอนการทำ Modelling

10.4.1 ความรู้เบื้องต้นเกี่ยวกับ Model Evaluation

Model Evaluation คือขั้นตอนการประเมินประสิทธิภาพของเทคนิคการทำเหมือง ข้อมูล (Data Mining) ซึ่งเป็นขั้นตอนที่สำคัญในการพิจารณาคุณภาพของโมเดลหรือเทคนิคที่ใช้ เพื่อ ปรับแต่งคุณลักษณะ (Attribute) ในกระบวนการทำซ้ำ (Iterative Process) และเพื่อคัดเลือกโมเดล (Model) ที่ดีที่สุดจากโมเดลทั้งหมดที่มี ทั้งนี้การประเมินการทำงานของแต่ละโมเดลจะมีวิธีการที่ แตกต่างกันไป (R Sub-sampling, 2011)

ตัวอย่างเช่น วิธีการประเมินสำหรับโมเดลการจำแนกประเภท (Classification) ต้อง มีการวัดอัตราผิดพลาด (Error Rate) ซึ่งเป็นสัดส่วนของจำนวนที่ทำนายหรือแนะนำที่สำเร็จ (Success) ต่อจำนวนข้อมูลทั้งหมด โดยมีวิธีการวัดดังต่อไปนี้

(1) Cross-Validation คือ เทคนิคที่ใช้ในการประเมินการทำงานของโมเดลการ ทำนาย (Predictive Model) ซึ่งจะต้องมีการแบ่งข้อมูล (Data) ออกเป็นชุดข้อมูลที่ใช้โมเดลสอน

(Training Dataset) กับชุดข้อมูลที่ใช้ทดสอบ (Testing Dataset) โดยจะเลือกโมเดลที่ทำให้เกิดความเสียหายที่ค่าน้อยที่สุด

(2) Holdout Method คือ เทคนิคที่ขึ้นกับการแบ่ง (Split) ของข้อมูล เป็นวิธีที่ง่ายที่สุดของ Cross Validation โดยจะแบ่งข้อมูลเป็นชุดข้อมูลที่ใช้โมเดลสอน (Training Dataset) กับชุดข้อมูลที่ใช้ทดสอบ (Testing Dataset) เช่นกัน โดยจะใช้คุณลักษณะ (Attribute) หรือฟังก์ชัน (Function) Approximator กับชุดข้อมูลสอน (Training Data) เท่านั้นในการทำนายผลลัพธ์ (Output) ที่เกิดขึ้นในชุดข้อมูลที่ใช้ทดสอบ (Testing Dataset) โดยที่ไม่เห็นค่าผลลัพธ์ (Output) มาก่อน ทั้งนี้ค่าผิดพลาดจะถูกสะสมมาก่อน ซึ่งวัดได้จาก Mean Absolute Test Set Error ซึ่งวิธีนี้จะคำนวณผลเร็วและมักใช้กับการคำนวณหาค่าผิดพลาด (Error) ของส่วนที่เป็นเศษ (Residual) แต่จะมีความแปรปรวน (Variance) สูงในชุดที่ใช้ทดสอบ (Test Set) ดังนั้นสิ่งที่สำคัญที่สุดของวิธีนี้คือการแบ่งสัดส่วนของ ชุดข้อมูลที่ใช้สอน (Training Dataset) กับชุดข้อมูลที่ใช้ทดสอบ (Testing Dataset)

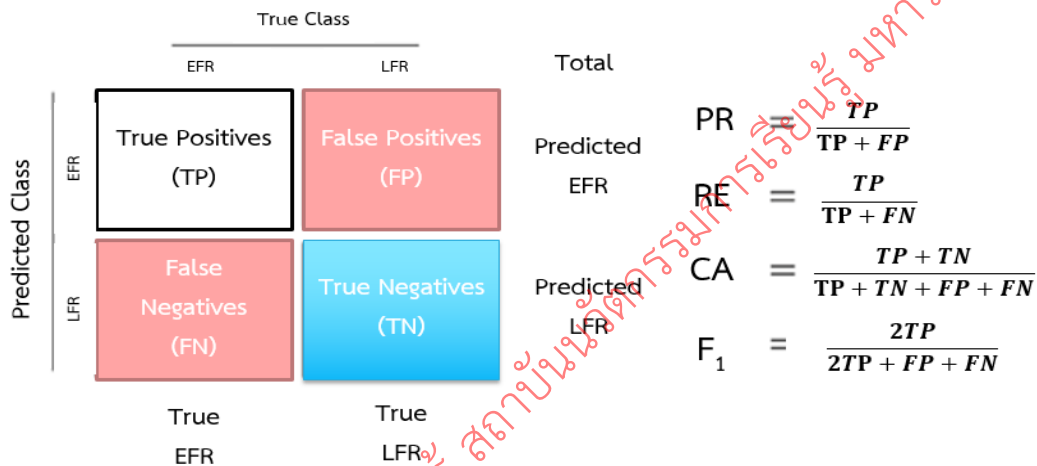
(3) Random Sub-Sampling คือ การทำซ้ำหลายครั้ง เพื่อให้การประเมินการทำงานของตัวแปรดีขึ้น แต่วิธีนี้มีปัญหาเดียวกับ Holdout Method เพราะไม่ได้ใช้ข้อมูลเพียงพอสำหรับ ข้อมูลที่ใช้สอน (Training Dataset) และไม่มีการควบคุมจำนวนครั้งที่แต่ละเรคคอร์ด (Record) ถูกใช้สำหรับทดสอบ (Testing) และเรียนรู้ (Training) ทำให้ผิดพลาดได้หากบางเรคคอร์ด (Record) ถูกใช้ในการเรียนรู้ (Train) มากกว่าเรคคอร์ด (Record) อื่นๆ

(4) K-Fold Cross Validation เป็นอีกวิธีในการแก้ปัญหาของ Holdout Method โดยข้อมูลจะถูกแบ่งเป็น k Subset และจะนำวิธี Holdout Method มาทำซ้ำ k ครั้ง โดยที่ส่วนหนึ่งของ k Subset จะใช้ ทดสอบ (Testing) และที่เหลือใน $k-1$ subset จะถูกรวมกันเอามาการเรียนรู้ (Train) ซึ่งจะวัดผลจากการคำนวณค่าผิดพลาด (Error) จากการทดลองทั้งหมด k ครั้ง ทำให้วิธีนี้ลดผลกระทบจากการขึ้นกับสัดส่วนการแบ่งข้อมูล เพราะทุกๆส่วนของข้อมูลจะถูกทดสอบโดยตรงในครั้งแรก และนำไปเรียนรู้อีก $k-1$ ครั้ง ทำให้ค่าความแปรปรวน (Variance) ลดลง ตามจำนวนครั้ง k ที่เพิ่มขึ้น แต่ข้อเสียคือต้องการคำนวณซ้ำๆ k ครั้ง

(5) Leave-One-Out Method จะใช้จำนวน N ครั้งที่แยกจากกันในการทดสอบข้อมูล และ Function Approximator จะถูกใช้การเรียนรู้ (Train) ในทุกๆ ข้อมูล ยกเว้นไว้หนึ่งจุดที่จะถูกใช้ในการทำนายแทน ซึ่งจะมีการวัด Leave-One-Out Cross Validation Error (LOO-XVE) แต่ข้อเสียคือค่าใช้จ่ายในการคำนวณสูง

(6) Bootstrap วิธีนี้จะต่างวิธีอื่นๆ เพราะการสุ่มตัวอย่าง (Sampling) ของข้อมูลที่ใช้เรียนรู้ จะถูกทำโดยมีการแทนที่ กล่าวคือ ข้อมูลที่เคยใช้การเรียนรู้ (Train) ไปแล้วจะถูกดึงกลับมาใน Original Pool ทำให้ทุกข้อมูลมีโอกาสที่จะนำมาใช้ใหม่เท่าๆกัน

(7) Confusion Matrix คือการแบ่งชั้นของผลลัพธ์ (Output Class) ที่ทำนายได้ออกเป็น 2 กลุ่มย่อยของแต่ละกลุ่มหลัก คือถูก (True) กับผิด (False) จึงทำให้เกิดผลในการ จำแนกประเภท (Classification) ได้ทั้งหมด 4 กลุ่ม คือค่าจริงเป็นบวกและทำนายเป็นบวก (True Positive: TP) ค่าจริงเป็นลบและทำนายเป็นลบ (True Negative: TN) ค่าจริงเป็นลบ แต่ทำนายเป็นบวก (False Positive: FP) ค่าจริงเป็นบวก แต่ทำนายเป็นลบ (False Negative: FN) โดยจะวัดจากค่า ดังรูปที่ 10.4

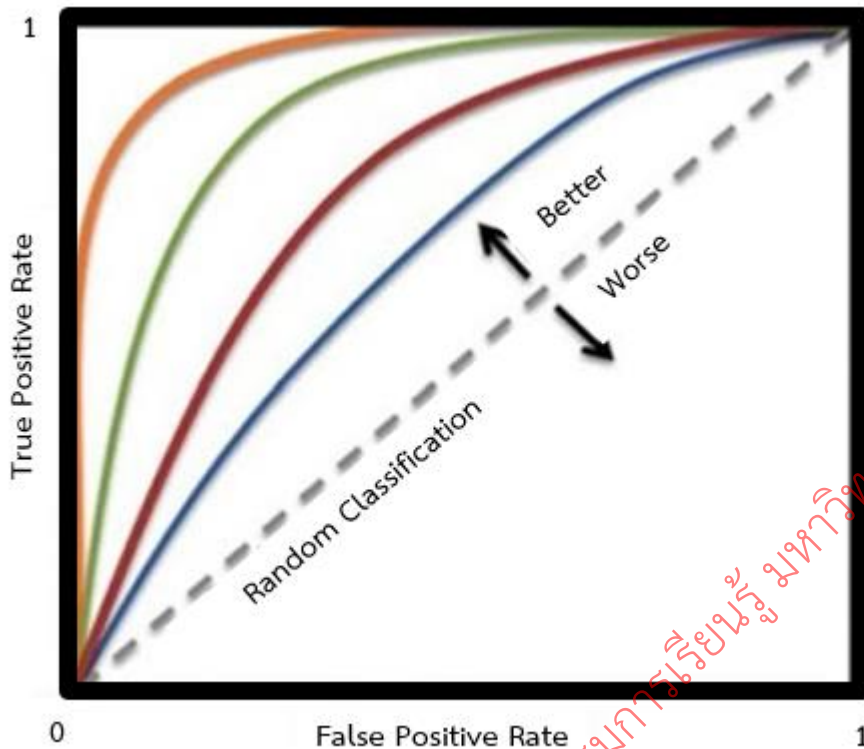


รูปที่ 10.4 Confusion Matrix

(8) Receiver Operating Curves (ROC) วิธีการนี้สามารถบอกจุด Cut-Off ที่เหมาะสมที่สุดได้ว่า ณ จุดไหนที่มีความแม่นยำของ Test เกิดขึ้นสูงที่สุด และน่าเชื่อถือที่สุด ทำให้สามารถใช้เลือกจุดนั้นมาเป็นจุด Cut-Off ได้เลยในการแบ่งกลุ่มข้อมูลเพื่อให้มีความถูกต้องมากที่สุด และผิดพลาดน้อยที่สุด นอกจากนี้ยังสามารถสร้าง ROC Curve มาได้หลายเส้นเพื่อใช้เปรียบเทียบกันระหว่างการทดสอบ (Test) ที่แตกต่างกัน ว่าในสถานการณ์เดียวกันการทดสอบ (Test) ไหนจะเหมาะสมที่สุด ทั้งนี้ ROC Curve สร้างมาจากการเขียน (Plot) กราฟระหว่าง

$$\text{True Positive Rate} = \text{Recall} = \frac{TP}{TP + FN}$$

$$\text{False Positive Rate} = \frac{FP}{FP + TN}$$



รูปที่ 10.5 ROC Curve

จากรูปที่ 10.5 การที่กราฟเข้าใกล้จุดซ้ายบน (Perfect Performance) ซึ่งจะแสดงออกมาเป็นค่าเข้าใกล้ 1 จะยิ่งอธิบายได้ว่าโมเดลนั้นมีประสิทธิภาพ (Mikut & Reischl, 2011)

นอกจากนี้ ยังสามารถประเมินค่าประสิทธิภาพการทำงานของ Model จาก Parameter ดังต่อไปนี้

(1) Main Criterion มักใช้ในการเปรียบเทียบ และใช้เฉพาะกระบวนการที่ Performance Vector ถูกเปรียบเทียบ เช่น Attribute Selection หรือ การทำ Meta Optimization หากไม่มีการเลือก Main Criterion ตัว Criterion ตัวแรกใน Performance Vector จะถูกนำไปใช้เป็น Main Criterion

(2) Accuracy คือจำนวนที่ทำนายได้ถูกต้อง เมื่อเทียบกับข้อมูลทั้งหมด

(3) Classification Error จำนวนที่ทำนายผิด เมื่อเทียบกับข้อมูลทั้งหมด

(4) Kappa จะวัดได้ดีกว่าค่าอื่น เพราะคำนึงถึงจำนวนการทำนายถูกต้อง ที่เกิดขึ้นโดย

บังเอิญด้วย

(5) Recall คือ จำนวนบวกที่ทำนายได้ถูกต้อง เมื่อเทียบกับจำนวนบวกที่แท้จริงทั้งหมด (TP+FN)

(6) Precision คือ จำนวนบวกที่ทำนายได้ถูกต้อง เมื่อเทียบกับจำนวนบวกที่แสดงทั้งหมด (TP+FP)

(7) Squared Correlation คือ ค่าความสัมพันธ์กำลังสองระหว่าง Label (Output) กับ Prediction Attribute

(8) การวัดค่า Error ต่างๆ ดังนี้

Mean Squared error	MSE	$= \frac{1}{n} \sum_{t=1}^n e_t^2$
Root mean Squared error	RMSE	$= \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}$
Mean absolute error	MAE	$= \frac{1}{n} \sum_{t=1}^n e_t $
Mean absolute percentage error	MAPE	$= \frac{100\%}{n} \sum_{t=1}^n \left \frac{e_t}{y_t} \right $

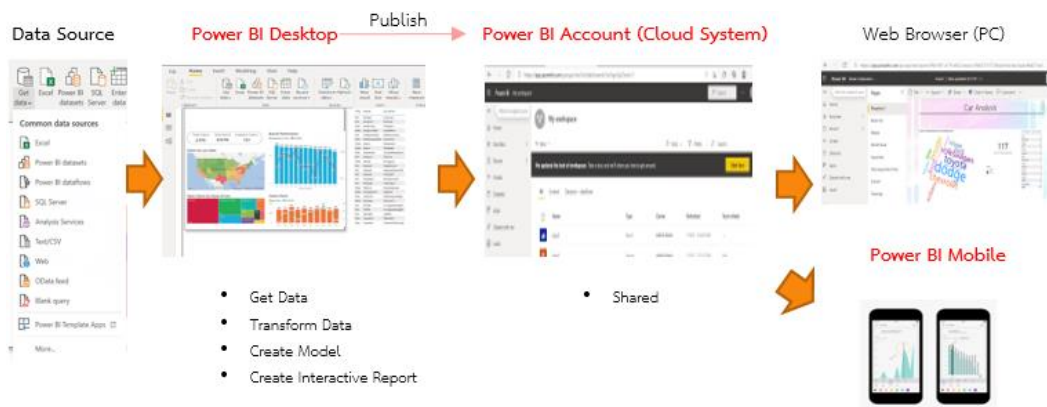
รูปที่ 10.6 ค่าความผิดพลาด (Error) ที่ใช้วัดประเมินผลโมเดล (Model Evaluation)

10.4.2 การสำรวจข้อมูล (Information Exploration Layer)

ในเลเยอร์นี้ ถ้ากล่าวโดยภาษาไทยแบบไม่ได้แปลตามภาษาอังกฤษก็คือ ขั้นตอนการนำเสนอข้อมูลเพื่อให้หน่วยงานทางด้านธุรกิจเข้าใจ และนำไปวางแผนเชิงกลยุทธ์เพื่อให้ชนะทางธุรกิจ ซึ่งอีกความหมายหนึ่งที่ภาษาด้านการทำเหมืองข้อมูลและวิทยาศาสตร์ข้อมูลใช้คือ การเล่าเรื่องด้วยข้อมูล (Data Storytelling) ซึ่งเป็นหนึ่งในองค์ประกอบสำคัญและปัจจัยแห่งความสำเร็จของการนำผลการวิเคราะห์และโมเดลที่ได้ไปใช้ เป็นขั้นตอนของการผสมผสานระหว่างศาสตร์ของวิทยาศาสตร์ข้อมูล (Data Science) และความรู้ในเรื่องโดเมนธุรกิจที่กำลังทำการวิเคราะห์อยู่ (Domain Expertise) ซึ่งในปัจจุบันมีซอฟต์แวร์หลากหลายที่ให้บริการด้านการวิเคราะห์ข้อมูลตั้งแต่ระดับธุรกิจอัจฉริยะ (Business Intelligence: BI) ไปจนถึงเครื่องมือที่ทำด้านการทำเหมืองข้อมูลโดยตรงและมีเครื่องมือช่วยนำเสนอด้วยยกตัวอย่างดังต่อไปนี้

(1) เครื่องมือ Power BI คือ ซอฟต์แวร์หรือ แอปพลิเคชันที่เป็นชุดของเครื่องมือในการวิเคราะห์ข้อมูลธุรกิจ (Business Analytics Tool) สามารถนำเข้าข้อมูลจากแหล่งข้อมูล (Data Source) สามารถทำงานร่วมกันกับไฟล์ข้อมูลประเภทต่างๆ เพื่อเชื่อมโยงข้อมูลที่ไม่เกี่ยวข้องให้เป็นข้อมูลเชิงลึกที่สอดคล้องกัน โดยใช้เครื่องมือที่ชื่อว่า Power Query เพื่อเตรียมข้อมูลให้พร้อมใช้งานสามารถใช้สร้างรายงาน สามารถสร้างกระดานนำเสนอข้อมูล (Dashboard) โดยใช้กราฟหลากหลายรูปแบบผสมผสานในหน้าจอเดียวได้อย่างน่าสนใจ สามารถทำโมเดลข้อมูล (Data Model) สามารถสร้างความสัมพันธ์ของข้อมูล (Relationships) สร้างสูตรการคำนวณด้วยฟังก์ชัน DAX และสร้างกราฟแบบโต้ตอบได้ (Interactive) Power BI เป็นเครื่องมือระดับธุรกิจอัจฉริยะ (Business Intelligence: BI) แบบรวมศูนย์ สามารถอัปเดต (Update) ได้อย่างทันที อีกทั้งยังสามารถดูได้จากทุกๆ อุปกรณ์ ทั้งคอมพิวเตอร์ส่วนบุคคล (Personal Computer) โทรศัพท์มือถือ (Mobile) แท็บเล็ต (Tablet) ผู้ใช้สามารถทำการเลือกดูข้อมูลในมุมมองที่ต้องการ (จากการพัฒนาของผู้สร้างการนำเสนอ) เพื่อที่จะหาคำตอบ เพื่อตัดสินใจในการทำธุรกิจดังรูป 10.7 โดยส่วนประกอบของ Power BI ประกอบไปด้วย (Microsoft, 2561)

- Power BI Desktop คือส่วนการทำงานบนคอมพิวเตอร์ทั่วไป ซึ่งสามารถดาวน์โหลดมาใช้งานได้ฟรีระดับหนึ่งจาก <https://www.powerbi.com> หลังจากดาวน์โหลดมาเราจะเริ่มต้นใช้งาน Power BI Desktop ซึ่งขั้นตอนแรกคือการเชื่อมต่อกับแหล่งข้อมูลต่างๆ นอกจากนี้ Power BI มีแผนภาพชนิดต่าง ๆ จำนวนมากให้เลือกจากใน Power BI Desktop เพื่อสร้างหรือเปลี่ยนแผนภาพ สามารถสร้างรายงาน ซึ่งมีหนึ่งหน้าหรือหลายหน้าก็ได้
- Power Query คือ ส่วนของเตรียมข้อมูล ก่อนนำไปสร้างกราฟและ Data Model โดย Power Query สามารถทำการเปลี่ยนแปลงข้อมูล เช่น เปลี่ยนชนิดของข้อมูล เอาคอลัมน์ออก หรือรวมข้อมูลจากหลายแหล่ง
- Data Model คือส่วนของการคำนวณ DAX การจัดการโมเดล ทั้งการจัดรูปแบบ การกำหนดลำดับชั้น ความสัมพันธ์
- Power BI Service คือส่วนที่ใช้เก็บข้อมูลบนระบบ Cloud
- Power BI Mobile คือแอปพลิเคชันที่แสดงผลรายงาน (Dashboard) ในมือถือ สามารถ Download ได้ฟรีทั้ง iOS และ Android
- Data Gateway หรือ On Premise Data Gateway คือส่วนของการนำข้อมูลจาก On-Premise ขึ้นระบบ Cloud ให้แชร์รายงาน (Dashboard) และแชร์แหล่งข้อมูล (Data Source) ให้กับคนที่เกี่ยวข้องแบบ Real Time



รูปที่ 10.7 ส่วนประกอบของ Power BI

(2) Tableau (อ่านว่า ทาโบลว หรือ แท็บโบลว) เป็นคำศัพท์ที่มีรากศัพท์มาจากภาษาฝรั่งเศสมีความหมายเกี่ยวกับ รูปภาพ ภาพวาด หรือการระบายสี เป็นต้น (AiTeam, 2565) เป็นซอฟต์แวร์ที่สามารถนำข้อมูลจำนวนมากที่มีหลากหลายประเภทในองค์กรมาทำการวิเคราะห์เพื่อเป็นข้อมูลเชิงธุรกิจสำหรับการตัดสินใจ หากต้องการใช้งานต้องซื้อสิทธิ์การใช้ซอฟต์แวร์จากบริษัท ตัวแทนปีต่อปี สามารถใช้สร้างการนำเสนอด้านข้อมูล (Data Visualization) ด้วยกราฟฟิคหลายประเภทในหน้าจอเดียว โดยมุ่งเน้นไปที่อุตสาหกรรมที่ทำงานในด้านของ Business Intelligent สิ่งสำคัญที่สุดของ Tableau คือความสามารถในการเชื่อมต่อกับฐานข้อมูล (Database) สเปรดชีต (Spreadsheets) OLAP (Online Analytical Processing) Cube เป็นต้น (รูป 10.9) สำหรับฟังก์ชันหลักๆ ของ Tableau ประกอบไปด้วย

- Tableau Prep ทำหน้าที่เป็นเครื่องมือสำหรับการเชื่อมต่อข้อมูล (Connect Data) เชื่อมโยงข้อมูล (Data Blending) และทำความสะอาดข้อมูล (Data Cleansing) ติดตั้งในบนคอมพิวเตอร์ส่วนบุคคล (Personal Computer/Desktop) หรือแล็ปท็อป (Laptop)
- Tableau Desktop ทำหน้าที่เป็นเครื่องมือสำหรับสร้างการนำเสนอ ด้านข้อมูล (Data Visualization) บน Dashboards และทำการเชื่อมต่อข้อมูล (Connect Data) ติดตั้งในบนคอมพิวเตอร์ส่วนบุคคล (Personal Computer/Desktop) หรือแล็ปท็อป (Laptop)
- Tableau Server/Online ทำหน้าที่เป็นเครื่องมือสำหรับการแชร์ (Share) หรือการนำเสนอด้านข้อมูล (Data Visualization) บน Dashboard แบบออนไลน์ หรือทำหน้าที่เป็น ศูนย์ข้อมูล (Data Center) ในการเชื่อมโยงข้อมูลบนระบบ Cloud โดยติดตั้งบนเครื่อง

เซิร์ฟเวอร์ (Server) หรือกรณีผู้ใช้ออฟไลน์งานบน Tableau Online ก็จะใช้เครื่องเซิร์ฟเวอร์ (Server) ของ Tableau ดังนั้นองค์กรจึงไม่ต้องเตรียมเครื่องเซิร์ฟเวอร์ (Server)

สามารถดาวน์โหลด

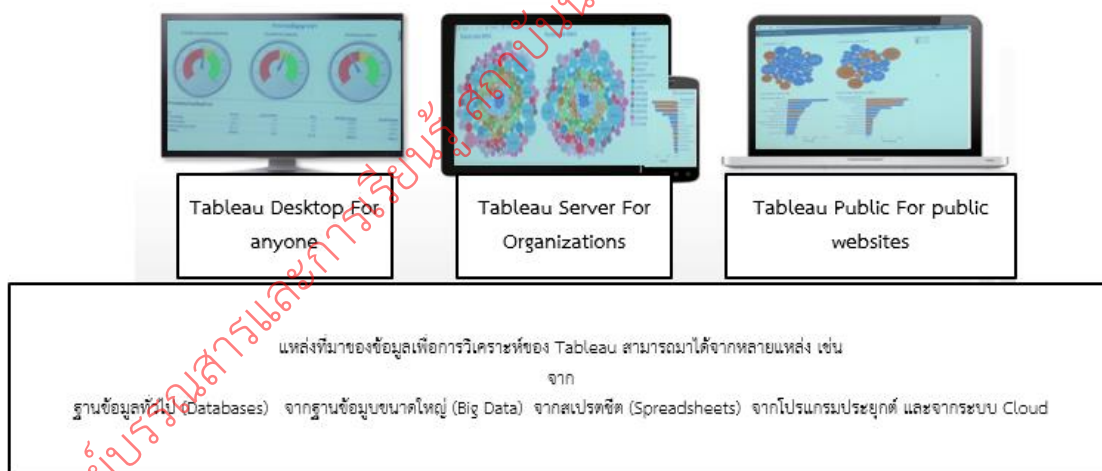
Tableau

มาทดลองใช้งานได้ที่

<https://www.tableau.com/>



รูปที่ 10.8 ห้องการเรียนการสอนด้วย Tableau ด้านการวิเคราะห์และนำเสนอข้อมูลของผู้เขียน



รูปที่ 10.9 ประเภทของสิทธิ์การใช้งาน Tableau 3 ประเภท

(3) NLTK (Natural Language Toolkit) พัฒนาขึ้นโดย Steven Bird และ Edward Lope ในคณะวิทยาคอมพิวเตอร์และสารสนเทศที่ มหาวิทยาลัยของเพนซิลวาเนีย (NLTK, 2022) เป็นที่ได้รับความนิยมมากที่สุดในนักวิทยาศาสตร์ข้อมูล (Data Scientist) เพราะเกี่ยวข้องกับการพัฒนาแบบจำลองทางสถิติที่ช่วยให้คอมพิวเตอร์เข้าใจภาษามนุษย์ แบบจำลองทางสถิติเหล่านี้เป็น

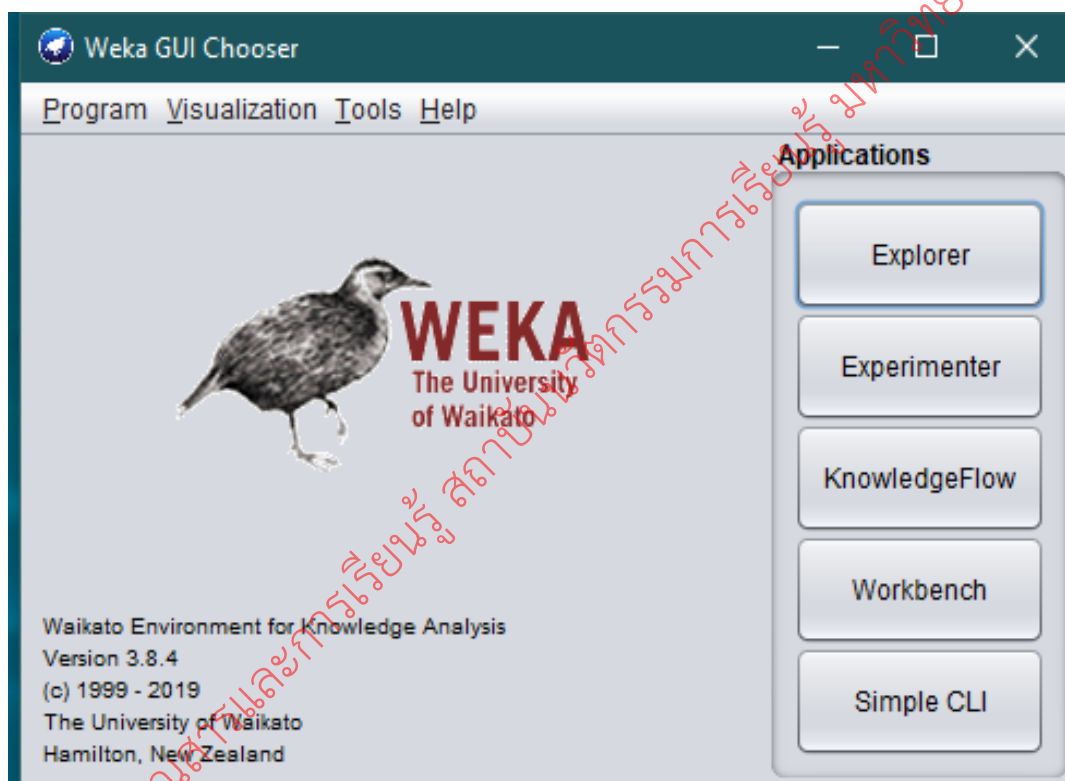
ส่วนหนึ่งของการเรียนรู้ของเครื่อง (Machine Learning และมีอัลกอริทึม (Algorithm) ที่สามารถช่วยให้คอมพิวเตอร์ทำความเข้าใจภาษาธรรมชาติ โดยภาษา Python มาพร้อมกับ Library ที่เรียกว่า Natural Language Toolkit (NLTK) ที่พัฒนาขึ้นเพื่อวัตถุประสงค์นี้เท่านั้น

(4) SAS เป็นคำย่อมาจาก Statistical Analysis System เป็นหนึ่งในเครื่องมือที่ได้รับการออกแบบมาสำหรับการทำงานทางสถิติ SAS เป็นซอฟต์แวร์แบบ Open Source ที่องค์กรขนาดใหญ่ใช้ในการวิเคราะห์ข้อมูล ซึ่งใช้สำหรับการสร้างแบบจำลองทางสถิติ ถูกใช้อย่างกว้างขวางโดยผู้เชี่ยวชาญและบริษัทที่ทำงานเกี่ยวกับซอฟต์แวร์เชิงพาณิชย์ที่เชื่อถือได้ SAS มี Library และเครื่องมือ (Tools) ทางสถิติที่สามารถใช้สำหรับการสร้างแบบจำลองและการจัดระเบียบข้อมูล SAS เป็นโปรแกรมที่จะใช้งานด้านการวิเคราะห์ระดับสูง (Advance Analytic) เช่นการวิเคราะห์เหมืองข้อมูล (Data Mining) หรือข้อมูลที่มีขนาดใหญ่มากๆ หรือข้อมูลที่เป็น Transaction ซึ่งผู้เชี่ยวชาญแนะนำให้ใช้แต่ก็มีราคาแพงและใช้เฉพาะในอุตสาหกรรมขนาดใหญ่เท่านั้น

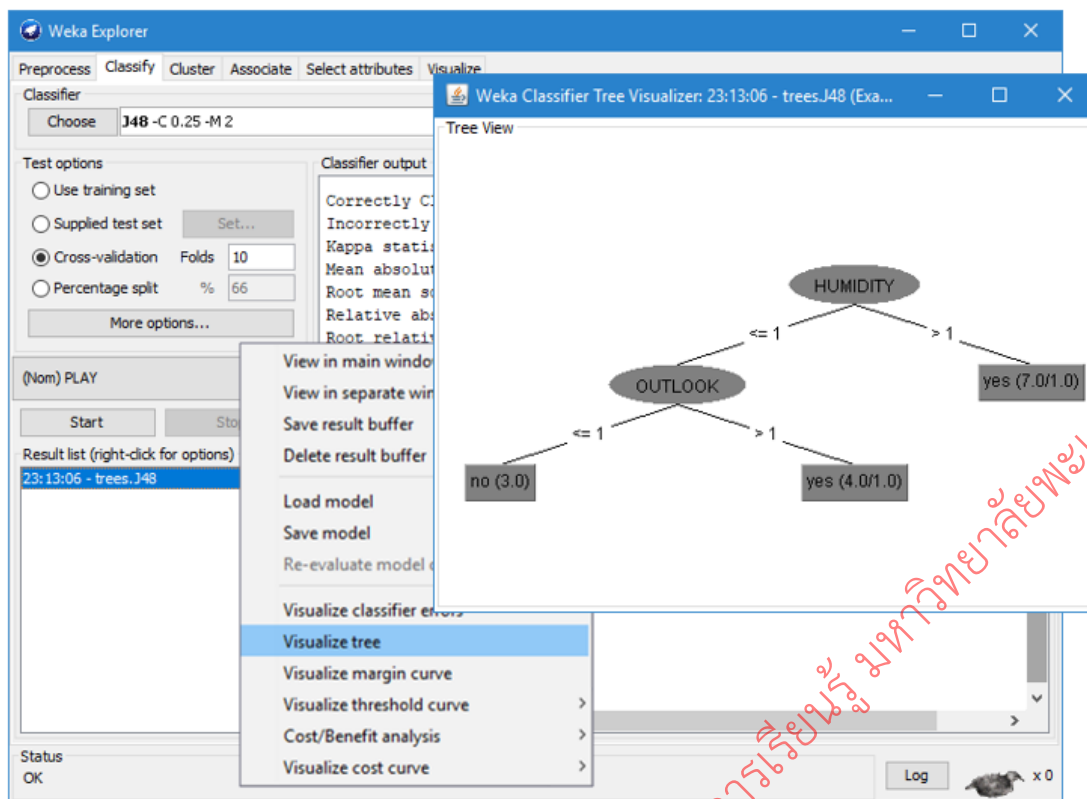
(5) MATLAB เป็นโปรแกรมของบริษัท The Mathworks, Inc. (<http://www.mathworks.com>) ใช้สำหรับการคำนวณทางคณิตศาสตร์ทั้งการคำนวณเชิงตัวเลข และการคำนวณเชิงสัญลักษณ์ (ซอฟต์แวร์เสริมชื่อ Maple) นอกจากนี้ชุดคำสั่งของ MATLAB ยังเป็นภาษาคอมพิวเตอร์ ที่มีความสามารถในการคำนวณ Matrix ซึ่งช่วยในการทำงานด้านวิทยาศาสตร์หลายด้านๆ ใช้สำหรับการคำนวณเชิงตัวเลขหลายกระบวนการสำหรับการประมวลผลข้อมูลทางคณิตศาสตร์ นอกจากนี้ MATLAB เป็นซอฟต์แวร์ Closed Source ที่อำนวยความสะดวกในการทำงานของ Matrix การใช้ Algorithm และการสร้างแบบจำลองทางสถิติของข้อมูล ใช้สำหรับจำลองเครือข่ายประสาทและตรรกศาสตร์คลุมเครือ การใช้ MATLAB Graphic Library สามารถนำเสนอข้อมูล (Data Visualization) MATLAB ยังสามารถใช้ในการประมวลผลภาพและสัญญาณ ซึ่งทำให้ MATLAB เป็นเครื่องมือเนกประสงค์สำหรับ นักวิทยาศาสตร์ข้อมูล (Data Scientist) เนื่องจากสามารถแก้ไขปัญหาทั้งหมดได้ตั้งแต่การทำความสะอาดข้อมูล (Data Cleaning) และวิเคราะห์ไปจนถึงการทำการวิเคราะห์และเรียนรู้ข้อมูลเชิงลึก (Deep Learning) สามารถดาวน์โหลด MATLAB มาทดลองใช้งานได้ที่ <https://matlab.mathworks.com/>

(6) Weka หรือโปรแกรมเวก้า (รูป 10.10) (Waikato Environment for Knowledge Analysis: Weka) พัฒนาโดยมหาวิทยาลัย Waikato ประเทศ นิวซีแลนด์ Weka เป็นซอฟต์แวร์ Machine Learning ที่เขียนด้วยภาษา Java เป็นซอฟต์แวร์ที่ผู้เชี่ยวชาญใช้มากที่สุดในงานวิจัยเพราะสามารถดาวน์โหลดได้ฟรีและสามารถใช้วิเคราะห์ข้อมูลได้ระดับหนึ่ง Weka มีชุด Machine Learning Algorithm สำหรับการทำเหมืองข้อมูล (Data Mining) โดย Weka ประกอบด้วย

เครื่องมือการเรียนรู้ (Machine Learning Tools) ต่าง ๆ เช่น การจำแนกประเภท (Classification) การจัดกลุ่ม (Clustering) การหาความสัมพันธ์ด้วยสมการถดถอย (Regression) มีเครื่องมือแนะนำด้านข้อมูล (Data Visualization) และการเตรียมข้อมูล (Data preparation) ระดับหนึ่งแต่ไม่ได้ดีมากเท่ากับซอฟต์แวร์ที่ต้องจ่ายเพื่อซื้อสิทธิ์ on โดย Weka เป็นซอฟต์แวร์ประเภทเปิดให้ใช้งาน (Open Source GUIX ที่อนุญาตให้ใช้งานเครื่องมือการเรียนรู้ (Machine Learning Tools) อัลกอริทึม (Algorithm) ได้ง่ายขึ้นผ่านแพลตฟอร์ม โดยที่สามารถเข้าใจการทำงานของเครื่อง (Machine Learning) กับข้อมูลโดยไม่ต้องเขียนโค้ด (Coding) เหมาะสำหรับนักวิทยาศาสตร์ข้อมูล (Data Scientist) ที่เพิ่งเริ่มต้นในการทำเหมืองข้อมูล (Data Mining) สามารถดาวน์โหลดได้ที่ <https://waikato.github.io/weka-site/index.html>



10.10 โปรแกรม WEKA



10.11 ตัวอย่างการวิเคราะห์และนำเสนอด้วยโปรแกรม WEKA ด้วยเทคนิค Classify J48

(7) TensorFlow เป็นเครื่องมือมาตรฐานสำหรับงานการเรียนรู้ (Machine Learning Tools) (Thai Programmer, 2022) มีเครื่องมือสำหรับการเรียนรู้ของเครื่อง (Machine Learning deep learning library) ที่พัฒนาโดย Google และถูกใช้อย่างกว้างขวางสำหรับพัฒนาในการทำเหมืองข้อมูลเชิงลึกขั้นสูง (Deep Learning Developer) Tensors ซึ่งเป็น Array หลายมิติ เป็นเครื่องมือประเภทฟรี (Open Source) มี ฟังก์ชันหลัก 3 ฟังก์ชันคือด้านการเตรียมประมวลผลข้อมูล ด้านการสร้างแบบจำลอง และด้านการฝึกและประเมินแบบจำลอง นอกจากนี้ TensorFlow เป็นที่รู้จักในด้านประสิทธิภาพและความสามารถในการคำนวณขั้นสูง TensorFlow สามารถทำงานได้ทั้งบน CPU และ GPU และเพิ่งขึ้นบนแพลตฟอร์ม TPU ที่ดีกว่า ทำให้ได้เปรียบเป็นประวัติการณ์ในแง่ของพลังการประมวลผลของการเรียนรู้ของอัลกอริทึมต่างๆ (Machine Learning Algorithm) ซึ่ง Google เองก็ใช้การเรียนรู้ของเครื่อง (Machine Learning) เพื่อเพิ่มประสิทธิภาพกับผลิตภัณฑ์มากมาย ไม่ว่าจะเป็นเครื่องมือค้นหา (Search engine) การแปลภาษา (Google Translation) คำบรรยายภาพ (Image Captioning) และเครื่องมือช่วยการเสนอแนะ (Recommendations) สามารถดาวน์โหลดใช้งานฟรีได้ที่ <https://www.tensorflow.org/>

10.5 แนวทางในการเล่าเรื่องด้วยข้อมูล (Data Storytelling)

สำหรับขั้นตอนสุดท้ายของการทำเหมืองข้อมูลตามกระบวนการทำเหมืองข้อมูล “CRISP-DM” คือการนำเอาข้อมูลที่เป็นผลลัพธ์จากทั้งหมดจากการทำโมเดล มาลองปฏิบัติจริงกับธุรกิจในองค์กร โดยแปลงแนวคิดที่มีให้เกิดเป็นสารสนเทศที่ทุกคนเข้าใจ เพื่อให้ผู้บริหารหรือนักการตลาดเข้าใจและสามารถนำไปใช้ประโยชน์ในการวางแผนทางธุรกิจได้จริง และใช้ในการติดตามประเมินผลที่ได้เพื่อนำกลับไปปรับปรุงการทำเหมืองข้อมูล (Data Mining) อย่างต่อเนื่องต่อไป โดยอาจจัดทำเป็นรูปแบบของรายงาน (Report) หรือแผนภาพ (Dashboard) ที่พร้อมให้ฝ่ายต่างๆ นำไปใช้ประโยชน์ในการวางแผน กำหนดกลยุทธ์ ซึ่งการประเมินผลสามารถทำได้หลายทางเช่น วัดจากส่วนแบ่งของตลาด วัดจากปริมาณลูกค้า หรือ วัดจากกำไรสุทธิ เป็นต้น โดยเทคนิคสุดท้ายที่นักวิเคราะห์จำเป็นต้องมีในการใช้เครื่องมือดังที่กล่าวมาแล้วในหัวข้อ 10.4 ก็คือทักษะของการเล่าเรื่องด้วยข้อมูล (Data Storytelling) ซึ่งมีแนวทางการนำเสนอเบื้องต้นดังนี้

10.5.1 การเขียนแนวทางแบบมีเรื่องราว (Story) เนื้อหาการวิเคราะห์ก็จะมีความแตกต่างไม่ซ้ำใคร ดูน่าเชื่อถือ และเข้าถึงได้ดี ได้เห็นภาพความเปลี่ยนแปลงต่างๆ เป็นฉากๆ เป็นเรื่องราว ข้อดีคือ ถ้านำเสนอดี จะทำให้เข้าใจภาพธุรกิจที่ดีและวางกลยุทธ์ได้ง่าย

10.5.2 การเขียนแนวทางแบบอธิบายคำจำกัดความ (Definition) หาอ้างอิง (Reference) ได้ง่าย เขียนให้ได้ประโยชน์แบบตรงไปตรงมา แต่ข้อเสียคือ เนื้อหาจะไม่ค่อยโดดเด่นแตกต่าง ยากต่อการจดจำ และไม่รู้ว่าจะช่วยให้เข้าใจภาพธุรกิจได้มากขนาดไหน

10.5.3 การนำเสนอแบบเล่าเรื่อง (Narrative) แนวทางนี้ยังไม่ได้ถึงกับมาเป็นเรื่องเป็นราว ข้อดีคือ สามารถออกแบบการนำเสนอให้น่าสนใจในความยาวกำลังดีได้ ถ้าดีก็ช่วยช่วยให้เข้าใจภาพธุรกิจได้มาก แต่ถ้ายังไม่ดี ก็อาจจะไม่เข้าใจได้เช่นกัน

10.5.4 การนำเสนอแนวทางแบบเชิงเล่าเปรียบเทียบ (Analogy) เน้นที่การเปรียบเทียบกับอีกสิ่งหนึ่ง เพื่อใช้สิ่งนั้นช่วยทำให้คนเกิดความเข้าใจการเปรียบเทียบ (Analogy) ที่ดี จะทำให้เข้าใจภาพธุรกิจเข้าใจคอนเซ็ปต์ของเรื่องยากๆ ได้ในเวลาอันรวดเร็ว ในขณะที่ถ้าเลือกการเปรียบเทียบที่ไม่ดี ก็อาจจะทำให้คนเกิดความเข้าใจผิด หรือไม่เห็นด้วยได้

ทั้งนี้ ในด้านการทำงานจริงนักวิทยาศาสตร์ด้านข้อมูล (Data Scientist) นอกเหนือที่จะต้องพัฒนาในส่วนทักษะด้านเทคนิค (Technical Skill) จากการพัฒนาทักษะด้านการเรียนรู้ของเครื่อง (Machine learning) การเรียนรู้เชิงลึก (Deep learning) ปัญญาประดิษฐ์ (Artificial Intelligence: AI) หรืออัลกอริทึม (Algorithm) แล้วยังจำเป็นอย่างยิ่งที่จะต้องพัฒนาความรู้ด้านธุรกิจ (Business) เพื่อนำข้อมูลไปวิเคราะห์หาข้อมูลเชิงลึก (Value Insight) ในทางด้านธุรกิจ ยกตัวอย่างเช่น การอธิบายผ่านโมเดลโครงข่ายประสาทเทียม (Artificial neural networks: ANN) อาจเกิดปัญหา

Black-Box ทำให้ไม่สามารถอธิบายในเชิงธุรกิจได้อย่างชัดเจนเท่ากับการอธิบายผ่านโมเดลพื้นฐานอย่าง Logistic Regression หรือ Decision Tree เนื่องจากผู้ตัดสินใจทางธุรกิจจะไม่สนใจเรื่องความซับซ้อนของโมเดล แม้ว่าโมเดลนั้นจะมีความถูกต้อง (Accuracy) สูงถึง 99.99% นอกจากนี้ นักวิทยาศาสตร์ด้านข้อมูล (Data Scientist) ยังควรพัฒนาทักษะในการสื่อสาร (Communication) เพื่อนำไปอธิบายข้อมูลที่เราได้ในเชิงธุรกิจให้ผู้ตัดสินใจได้เข้าใจและสามารถแก้ไข ปัญหาที่เกิดขึ้นได้ สิ่งที่ได้ช่วยได้สำหรับขั้นตอนการสื่อสาร คือการนำเสนอ (Visualization) โดยการหาภาพเสมือน (Visualize) ที่เหมาะสมกับข้อมูล และสิ่งที่เราต้องการอธิบายในทางธุรกิจ อีกทั้ง นักวิทยาศาสตร์ด้านข้อมูล (Data Scientist) ควรรู้จักการพัฒนาตนเอง (Self-Development) หรือการพัฒนาทักษะด้วยตนเอง ด้วยการหาข้อมูลจากโซเชียลเน็ตเวิร์ค (Social Network) กรณีศึกษา (Case Study) และค้นหาเทคโนโลยีและอัลกอริทึมใหม่ๆ

สรุปเนื้อหาบทที่ 10 "เครื่องมือและเทคนิคที่สำคัญสำหรับการทำเหมืองข้อมูล" แนะนำเนื้อหาเกี่ยวกับเครื่องมือและเทคนิคที่ใช้ในกระบวนการทำเหมืองข้อมูล การประยุกต์ใช้โมเดลกระบวนการวิเคราะห์ข้อมูล (The CRISP-DM model) เพื่อวิเคราะห์และค้นหาความรู้จากข้อมูล

1. **เครื่องมือซอฟต์แวร์สำหรับการทำเหมืองข้อมูล** อธิบายและแนะนำเครื่องมือซอฟต์แวร์ที่ใช้ในการทำเหมืองข้อมูล เช่น ซอฟต์แวร์ Weka, RapidMiner และเครื่องมือซอฟต์แวร์อื่นๆ ที่รองรับการทำเหมืองข้อมูล
2. **การสร้างและจัดการกับข้อมูล** อธิบายเทคนิคในการสร้างและจัดการกับข้อมูลที่ใช้ในกระบวนการทำเหมืองข้อมูล เช่น การสร้างฐานข้อมูล เชื่อมต่อกับแหล่งข้อมูล และการจัดรูปแบบข้อมูล
3. **เทคนิคการทำนายและการจัดการความไม่แน่นอน** อธิบายเทคนิคในการทำนายผลลัพธ์และการจัดการความไม่แน่นอนในกระบวนการทำเหมืองข้อมูล เช่น การใช้เทคนิค Naive Bayes และการจัดการกับค่าความไม่แน่นอน
4. **การใช้เทคนิคสถิติในการทำเหมืองข้อมูล** อธิบายเทคนิคสถิติที่ใช้ในการทำเหมืองข้อมูล เช่น การคำนวณค่าเฉลี่ย ค่าสูงสุด ค่าต่ำสุด และการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร
5. **การใช้เทคนิคการจัดกลุ่มและการจำแนก** อธิบายเทคนิคการจัดกลุ่มและการจำแนกข้อมูลที่ใช้ในกระบวนการทำเหมืองข้อมูล เช่น การใช้อัลกอริทึม K-Means สำหรับการจัดกลุ่มข้อมูล

6. การทำนายและการสร้างโมเดล อธิบายการใช้เทคนิคในการทำนายและการสร้างโมเดลทำนายจากข้อมูล เช่น การใช้อัลกอริทึม Decision Trees และการสร้างโมเดลสำหรับการทำนาย

7. การใช้เทคนิคการตัดสินใจ: อธิบายการใช้เทคนิคการตัดสินใจในกระบวนการทำเหมืองข้อมูล เพื่อวิเคราะห์และประเมินผลลัพธ์ที่ได้จากการทำเหมืองข้อมูล

โดยสรุปบทที่ 10 เป็นการนำเสนอเครื่องมือและเทคนิคที่สำคัญในการทำเหมืองข้อมูล เพื่อวิเคราะห์และค้นหาความรู้จากข้อมูล การใช้เครื่องมือและเทคนิคเหล่านี้ช่วยในการเข้าใจแนวโน้มและข้อมูลที่มีความหนา

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

คำศัพท์ บทที่ 10

Alignment	สอดคล้อง
Business Intelligence	ระบบธุรกิจอัจฉริยะ
Business Objectives	วัตถุประสงค์เชิงธุรกิจ
Collaborative Commerce	การทำงานร่วมกันระหว่างองค์กรพันธมิตร
Cross Functional Process Improvement	การปรับปรุงกระบวนการระหว่างฝ่ายและหน่วยธุรกิจ
Cross-Industry Standard Process for Data Mining (CRISP-DM)	กระบวนการวิเคราะห์ข้อมูล
Customer Satisfaction	ความพึงพอใจของลูกค้า
Dashboard	กระดานนำเสนอข้อมูล
Data Aggregation Layer	เลเยอร์การรวมข้อมูล
Data Cleaning	การทำความสะอาดข้อมูล
Data Correcting	การแก้ไขข้อมูลที่ผิดพลาด
Data Duplicate Elimination	การลบชุดข้อมูลซ้ำซ้อนทิ้ง
Data Engineer	วิศวกรข้อมูล
Data Governance Layer	กลไกการกำกับดูแลข้อมูล
Data Parsing	การแจกแจงข้อมูล
Data Standardization	มาตรฐานข้อมูล
Data Standardizing	การทำข้อมูลให้เป็นรูปแบบเดียวกัน
Data Storytelling	การเล่าเรื่องด้วยข้อมูล
Data Visualization	การนำเสนอด้วยข้อมูล
Digital World	โลกดิจิทัล
Error Rate	อัตราผิดพลาด
Funding Priorities	ลำดับความสำคัญของการลงทุน
Information Exploration Layer	เลเยอร์การสำรวจข้อมูล
Iterative Process	กระบวนการทำซ้ำ
Management Capabilities	ขีดความสามารถในการบริหารจัดการ
Model Evaluation	การประเมินประสิทธิภาพของโมเดล
Neglecting context	การละเลยบริบท
Operational Efficiency	ประสิทธิภาพในการปฏิบัติงาน
Process Standardization	มาตรฐานกระบวนการ
Random Sub-Sampling	การทำซ้ำหลายๆครั้ง
Residual	ส่วนที่เป็นเศษ
Self Service Data Preparation	การเตรียมข้อมูลเบื้องต้นด้วยตนเอง

Speed to Market	ความเร็วในการผลิตสู่การตลาด
Spotting Data Quality Issues	การระบุคุณภาพของข้อมูล
Technology Standardization	มาตรฐานของงานเทคโนโลยีดิจิทัล
Variance	ความแปรปรวน

คำถามและแบบฝึกหัดท้ายบทที่ 10

1. จงอธิบายมุมมองที่ต้องทำความเข้าใจ 5 มิติสำหรับการทำเหมืองข้อมูล (Data Mining)
2. จงอธิบายโครงสร้างข้อมูล 3 ประเภท
3. ปัญหาที่พบส่วนใหญ่จะเกิดจากการนำเข้าข้อมูลภายนอกระบบคือ
4. จงอธิบายการแปลงข้อมูล (Transformation) จะประกอบไป 4 ขั้นตอนหลัก
5. จงยกตัวอย่างความท้าทายในการเตรียมข้อมูล (Data Preparation)
6. จงยกตัวอย่างขั้นตอนการประเมินประสิทธิภาพของโมเดล ในการทำเหมืองข้อมูล (Data Mining)
7. จงอธิบาย Power BI คือเครื่องมืออะไร ใช้ทำงานด้านใด และมีจุดเด่นอย่างไร
8. จงอธิบาย MATLAB คือเครื่องมืออะไร ใช้ทำงานด้านใด และมีจุดเด่นอย่างไร
9. จงอธิบาย Tableau คือเครื่องมืออะไร ใช้ทำงานด้านใด และมีจุดเด่นอย่างไร
10. จงอธิบายแนวทางในการเล่าเรื่องด้วยข้อมูล (Data Storytelling)

บรรณานุกรม

- เกรียงไกร ราชกิจ, “โครงข่ายประสาทเทียม หรือ ข่ายงานประสาทเทียม (artificial neural network).” [Online]. Available: <https://erp.mju.ac.th/articleDetail.aspx?qid=764>. [Accessed: 19 พฤษภาคม 2565].
- คณินิจต์ หนูเช็ก, ทศนีย์ ประธาน, พรทิพย์ เสียมหาญ, & สุวัจน์ เพชรรัตน์. (2561). การจัดการสินค้าและบริการใน ธุรกิจค้าปลีกสมัยใหม่ ประเภทซูเปอร์ มาเก็ต. Journal of Management Sciences Surattthani Rajabhat University, 5(2), 131-154.
- จิราภา เลหาหวรรณันท์, รชต ลิ้มสุทธีวันภูมิ และ บัณฑิต ฐานะโสภณ. “การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ.” กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, 2561
- ฐายิกา กลวิทย์อำนาจ. (2561). การบริหารธุรกิจค้าปลีกสมัยใหม่. กรุงเทพฯ: โรงพิมพ์ ซีเอ็ดดูเคชั่น.
- ประจักษ์ เฉิดโฉม, การประเมินการรับรู้ของนักศึกษาเกี่ยวกับการเรียนออนไลน์ (e-Learning): การศึกษาโดยใช้แบบจำลองการวิเคราะห์การตัดสินใจ, In การประชุมวิชาการ The 61st Kasetsart University Annual Conference กรุงเทพฯ 1 มีนาคม 2566 ถึงวันที่ 3 มีนาคม 2566
- ศิริลักษณ์ โรจนกิจอำนาจ (2545). การออกแบบและบริหารฐานข้อมูล : Database Design and Administration. กรุงเทพฯ : โรงพิมพ์ ดวงกมล.
- โอภาส-เอี่ยมสิริวงศ์. (2560). ระบบสารสนเทศเพื่อการจัดการ (ฉบับปรับปรุงเพิ่มเติม). กรุงเทพฯ: โรงพิมพ์ ซีเอ็ดดูเคชั่น.
- A Brief History of Enterprise Architecture, <http://blog.51sec.org/2018/11/ea-enterprise-architecture-framework.html>
- Agrawal, S., Sahu, A., & Kumar, G. (2022). A conceptual framework for the implementation of Industry 4.0 in legal informatics. Sustainable Computing: Informatics and Systems,

- AiTeam, “Tableau คืออะไร? What is Tableau?” [Online]. Available:
<https://www.aiteam.co.th/what-is-tableau/>. [Accessed: 14 พฤศจิกายน 2565].
- Amorim, A. C., da Silva, M. M., Pereira, R., & Gonçalves, M. (2021). Using agile methodologies for adopting COBIT. *Information Systems*, 101, 101496.
- Apipoj P., “Data Science คืออะไร?” [Online]. Available: <https://blog.derlivery.com/data-science-คืออะไร-57d7530554bf>. [Accessed: 19 พฤษภาคม 2565]
- Ashish, “ Apriori algorithm simplified with an example. ” [Online] . Available:
<https://edumine.wordpress.com/2013/09/11/apriori-algorithm-simplified-with-an-example/>. [Accessed: 19 May 2022].
- Boyde, J. (2014). A Down-To-Earth Guide To SDLC Project Management: Getting your system/software development life cycle project successfully across the line using PMBOK adaptively. Joshua Boyde.
- Brand Buffet, “กรณีศึกษา ‘เดอะมอลล์’ เจาะพฤติกรรมลูกค้าด้วย Data Mining ทำให้ลูกค้าประทับใจ - อยู่ น าน ม าก ขึ้น ,” [Online] . Available:
<https://www.brandbuffet.in.th/2017/08/themall-shopping-journey-behavior>. [Accessed: 19 พฤษภาคม 2563].
- Brunet-Thornton, R., & Martinez, F. (Eds.). (2018). *Analyzing the impacts of industry 4.0 in modern business environments*. IGI global.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (1999, March). The CRISP-DM user guide. In 4th CRISP-DM SIG Workshop in Brussels in March (Vol. 1999).
- Chertchom, Prajak, Patsama Charoenpong, and Tanasin Yatsungnoen Connecting STEM education and Japanese skills: The result from data mining of registration system’s data, approaches to increase Japanese competency for IT students” , The International Symposium on Advances in Technology Education (ISATE 2016), 13-16 September 2016, ISATE Sendai, Japan
- Chertchom, P. (2017). Output Based upon Input, the Result from Data Analytics and Visualization of TNI Registration System’s Data. *TNI Journal of Business Administration and Languages*, 5(1), 49-54.
- Chertchom, P., Tanimoto, S., Ohba, H., Kohnosu, T., Kobayashi, T., Sato, H., & Kanai, A. (2017, June). A Lifelog Data Portfolio for Privacy Protection Based on Dynamic Data

- Attributes in a Lifelog Service. In International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (pp. 107-120). Springer, Cham.
- Chertchom, Prajak, Senshu Yoshii, Patsama Charoenpong, and Tanasin Yatsungnoen. "Strategic Approach to Adopt ICT in SMEs." *TNI Journal of Engineering and Technology* 7, no. 1 (2019): 73-80.
- Cognitive Class, "Data Science Methodology 101 - Data Preparation Case Study." [Online]. Available: <https://www.youtube.com/watch?v=5oq1Rxo5QOo>. [Accessed: 19 May 2022].
- DataPreparator Software, "*What is Data Preparation?*" [Online]. Available: http://www.datapreparator.com/what_is_data_preparation.html [Accessed: 19 May 2022].
- (PO Enterprise Architecture 4. 0, 2020) <https://digital.dpo.go.th/wp-content/uploads/2020/05/pdffile003.pdf> [Accessed: 19 May 2022].
- Gabay, J., & Gabay, D. (2008). *UML 2 Analyse et conception: Mise en œuvre guidée avec études de cas*. Dunod.
- Harrison, R. (2018). *Togaf (r) 9 Foundation study guide*. Van Haren.
- IBM, "IBM SPSS modeler CRISP-DM guide," INSEAD Data Analytics Github, pp. 4, 13-17,19, 32, 35, 2011.
- Isinkaye F., Folajimi Y. and Ojokoh B., (2015). Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, vol. 16 no. 3, pp.261-273.
- Kumari M .and Godara S. (2011) "*Comparative study of data mining classification methods in cardiovascular disease prediction 1*"
- Laudon, K. C., & Laudon, J. P. (2013). *Management information systems: Managing the digital firm*. Pearson Educación.
- Hansen, M. T., & Birkinshaw, J. (2007). The innovation value chain. *Harvard business review*, 85(6), 121. [Online]. Available: <https://hbr.org/2007/06/the-innovation-value-chain> [Accessed: 8 November 2022].
- Marakas, G. M., & O'Brien, J. A. (2013). *Introduction to information systems* (p. 624). New York: McGraw-Hill/Irwin.
- McCormick, M. (2012). Waterfall vs. Agile methodology. *MPCS*, N/A, 3. Waterfall vs. Agile methodology
- Microsoft, "Power BI Desktop คืออะไร." [Online] Available:

- <https://docs.microsoft.com/th-th/power-bi/desktop-what-is-desktop>. [Accessed: 24 ตุลาคม 2565].
- Mikut, R., & Reischl, M. (2011). Data mining tools. Wiley interdisciplinary reviews: data mining and knowledge discovery, 1(5), 431-443.
- Moorthy, M., & Nobledhas, M. S. The Application of Big Data Analytics in Healthcare: A Proactive Approach. Vol. 6, Issue. 6, 2019, ISSN 2349-0780
- Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data, and a proposal. Journal of the American statistical association, 58(302), 415-434.
- Moudoubah, L., Mansouri, K., & Qbadou, M. (2021, November). COBIT 5 Concepts: Towards the Development of an Ontology Model. In The International Conference on Information, Communication & Cybersecurity (pp. 247-256). Springer, Cham.
- NLTK, "Natural Language Toolkit" [Online]. Available: <https://www.nltk.org/>. [Accessed: 14 November 2022].
- Olson, D. L., & Delen, D. (2008). Advanced data mining techniques. Springer Science & Business Media.
- Oravee, "รู้จักกับ "Ricult" Agriculture Tech ผู้ใช้ Machine Learning ช่วยเหลือเกษตรกรไทยเรื่องปัญหาสินเชื่อก่อน)." [Online]. Available: <https://techsauce.co/tech-and-biz/ricult-offers-machine-learning-credit-scoring-for-farmers/>. [Accessed: 19 May 2022].
- Power BI dashboard and report example – Antdata แหล่งที่มา : ภาพจาก (<https://www.youtube.com/watch?v=wAAZqJVzXA4>)
- Prajak Chertchom, Adisak Suasaming, Vivat Thongchotcha and Nithikorn Lertchanvuth, K-means Cluster Analysis of Tourist Package and Social Network Analysis for Promotion Recommendation, 12th International Conference on Project Management. Bangkok, Thailand, 27 November - 1 December, 2018
- Pratya N., "เทคนิคการจัดกลุ่ม (Clustering)." [Online]. Available: <http://datamining-techniques.blogspot.com/2012/09/clustering.html>. [Accessed: 19 September 2022]
- Press, G. (2016). Cleaning big data: Most time-consuming, least enjoyable data science task, survey says. Forbes, March, 23, 15. [Online]. Available: <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/?sh=383cc5946f63>. [Accessed: 9 November 2022]

- Pulkit Sharma, “ 4 Types of Distance Metrics in Machine Learning” , [Online]. Available: <https://www.analyticsvidhya.com/blog/2020/02/4-types-of-distance-metrics-in-machine-learning/>. [Accessed: 19 October 2022]
- Roehrs, A., Da Costa, C. A., & da Rosa Righi, R. (2017). OmniPHR: A distributed architecture model to integrate personal health records. *Journal of biomedical informatics*, 71, 70-81.
- Saltz, J. S., & Stanton, J. M. (2017). *An introduction to data science*. Sage Publications.
- Rokach, L., and O. Maimon. (2005) "Decission Trees." *Data Mining and Knowledge Discovery Handbook*: 165-191.
- R Sub-sampling, “Evaluation Measures for Data Mining Tasks.” (2011). *IASRI data mining techniques and tools for knowledge discovery in agricultural dataset*, pp. 145-152.
- Selerity, “Data preparation challenges facing every enterprise”, [Online]. Available: <https://seleritysas.com/blog/2020/08/13/data-preparation-challenges-facing-every-enterprise/> [Accessed: 10 November 2022].
- Sharma, K. K., Verdon-Kidd, D. C., & Magee, A. D. (2021). A decision tree approach to identify predictors of extreme rainfall events–A case study for the Fiji Islands. *Weather and Climate Extremes*, 34, 100405.
- Siravit D., “Artificial Neural Network”, [Online]. Available: <https://coladev.com/machine-learning/neural-network/2017/02/22/neural-network-basic>. [Accessed: 19 May 2021].
- Stephane Eric Collignon, Salman Nazir, Nanda Chingleput Surendra, *Agile systems development: Privacy theoretical lens to challenge the full information disclosure paradigm. , Information & Management, Volume 59, Issue 6, 2022, 103679, ISSN 0378-7206, https://doi.org/10.1016/j.im.2022.103679.*

Techopedia, “ D a t a P r e p r o c e s s i n g . ” [Online] . Available: <https://www.techopedia.com/definition/14650/data-preprocessing>. [Accessed: 19 May 2022].

Thai Programmer, “ ม า ท ำ ค ว า ม ร ู้ จ ัก Tensorflow “ [Online] . Available: <https://www.thaiprogrammer.org/2018/12/%E0%B8%A1%E0%B8%B2%E0%B8%97%E0%B8%B3%E0%B8%84%E0%B8%A7%E0%B8%B2%E0%B8%A1%E0%B8%A3%E0%B8%B9%E0%B9%89%E0%B8%88%E0%B8%B1%E0%B8%81-tensorflow/> [Accessed: 14 November 2022].

Wardah Z., "Data Scientist Skills," pp. 53-54, 56, April 2016.

Vanshika K. “ 8 Applications of Neural Networks. ” [Online] . Available: <https://www.analyticssteps.com/blogs/8-applications-neural-networks>. [Accessed: 1 November 2022].

Zachman, A. J., “ The Zachman Framework for Enterprise Architecture Version 3.0” . Retrieved from <http://www.zachman.com/about-the-zachman-framework>

Zhang, Y., & Patel, S. (2010). Agile model-driven development in practice. IEEE software, 28(2), 84-91.

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

ภาคผนวก

คำถามและแบบฝึกหัดท้ายบทที่ 1

1. จงอธิบายความหมายของระบบหมายถึงอะไร และระบบที่ดีควรมีลักษณะอย่างไร

ตอบ การทำงานของหน่วยย่อยๆ หลายๆ หน่วย ที่ทำงานปฏิสัมพันธ์กันหรือทำงานร่วมกันในวัตถุประสงค์เดียวกัน เพื่อให้การทำงานบรรลุวัตถุประสงค์ที่ได้กำหนดไว้ ระบบที่ดีควรมีลักษณะพิเศษคือ มีมาตรฐานที่สามารถยอมรับได้ ใช้ระบบเดียวกันนี้ใช้ได้ทั้งเป็นมาตรฐานชี้้นำการทำงาน ใช้ระบบเพื่อการปฏิบัติงาน และใช้ระบบเดียวกันนี้เป็นเครื่องตรวจสอบทบทวนผลการทำงาน

2. จงอธิบายถึงความแตกต่างของ ข้อมูล (Data) และ สารสนเทศ (Information)

ข้อมูล (Data) หมายถึง ข้อมูลดิบ ที่ถูกเก็บรวบรวมจากแหล่งต่างๆ ทั้งภายในและภายนอกองค์กร โดยข้อมูลยังไม่มี ความหมายในการนำไปใช้งาน

สารสนเทศ (Information) หมายถึง ผลลัพธ์ที่เกิดจากการประมวลผลข้อมูลดิบที่จัดเก็บไว้ อย่างเป็นระบบ โดยผลลัพธ์สามารถนำไปประกอบการทำงานหรือสนับสนุนการตัดสินใจของผู้บริหาร

3. จงอธิบายส่วนประกอบด้วยส่วนประกอบระบบสารสนเทศเพื่อการจัดการ

ประกอบด้วยส่วนสำคัญ 3 ประการดังต่อไปนี้ (Input - Process - Output)

1. เครื่องมือในการสร้างระบบสารสนเทศเพื่อการจัดการ หมายถึง โครงสร้างพื้นฐานที่รวมเข้าเป็น MIS และช่วยให้สารสนเทศดำเนินงานอย่างมีประสิทธิภาพ ได้แก่ ฐานข้อมูลและเครื่องมือ

2. วิธีการหรือขั้นตอนการประมวลผล โดยจัดลำดับวิธีการประมวลผลข้อมูล เพื่อให้ได้สารสนเทศที่ต้องการ

3. การแสดงผลที่ได้จากการประมวลผลของสารสนเทศอยู่ในรูปแบบรายงานต่างๆ

4. จงอธิบายประโยชน์ของระบบสารสนเทศ

1. ช่วยให้ผู้เข้าถึงสารสนเทศได้อย่างรวดเร็วและทันต่อเหตุการณ์

2. ช่วยให้ผู้ใช้ในการกำหนดเป้าหมายกลยุทธ์และการวางแผนปฏิบัติการ

3. ช่วยผู้ใช้ในการตรวจสอบผลการดำเนินงาน

4. ช่วยผู้ใช้ในการศึกษาและวิเคราะห์สาเหตุของปัญหา

5. ช่วยให้ผู้ใช้สามารถวิเคราะห์ปัญหาหรืออุปสรรคที่เกิดขึ้น

6. ช่วยลดค่าใช้จ่าย

5. จงอธิบายส่วนประกอบด้วยส่วนประกอบระบบสารสนเทศเพื่อการจัดการ

ประกอบด้วยส่วนสำคัญ 3 ประการดังต่อไปนี้ (Input □ Process - □ Output)

1. เครื่องมือในการสร้างระบบสารสนเทศเพื่อการจัดการ หมายถึง โครงสร้างพื้นฐานที่รวมเข้าเป็น MIS และช่วยให้สารสนเทศดำเนินงานอย่างมีประสิทธิภาพ ได้แก่ ฐานข้อมูลและเครื่องมือ

2. วิธีการหรือขั้นตอนการประมวลผล โดยจัดลำดับวิธีการประมวลผลข้อมูล เพื่อให้ได้สารสนเทศที่ต้องการ

3. การแสดงผลลัพธ์ ได้จากการประมวลผลของสารสนเทศอยู่ในรูปแบบรายงานต่างๆ

6. จงอธิบายประโยชน์ของระบบสารสนเทศ

1. ช่วยให้ผู้ใช้เข้าถึงสารสนเทศได้อย่างรวดเร็วและทันต่อเหตุการณ์

2. ช่วยให้ผู้ใช้ในการกำหนดเป้าหมายกลยุทธ์และการวางแผนปฏิบัติการ

3. ช่วยผู้ใช้ในการตรวจสอบผลการดำเนินงาน

4. ช่วยผู้ใช้ในการศึกษาและวิเคราะห์สาเหตุของปัญหา

5. ช่วยให้ผู้ใช้สามารถวิเคราะห์ปัญหาหรืออุปสรรคที่เกิดขึ้น

6. ช่วยลดค่าใช้จ่าย

7. ทำไมจึงต้องเรียนรู้ Data Mining (เหมืองข้อมูล)

การทำเหมืองข้อมูล (Data Mining) เปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลแบบเดิมไม่ซับซ้อน และไม่มีข้อมูลแบบภาพ เสียง มาสู่การจัดเก็บในรูปแบบฐานข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้ในการตัดสินใจเชิงลึก ที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล เช่น การนำข้อมูลขนาดใหญ่ มาทดลองประมวลผลและการสังเกตการณ์รูปแบบของพฤติกรรมผู้บริโภค โดยการสกัดหรือแยกข้อมูลที่เป็นประโยชน์จากข้อมูล การวางแผนทรัพยากรขององค์กรโดยสามารถวิเคราะห์ทางสถิติและตรรกะของข้อมูลขนาดใหญ่เป็นการมองหารูปแบบที่สามารถช่วยการตัดสินใจได้

คำถามและแบบฝึกหัดท้ายบทที่ 2

1. จงอธิบายวงจรการพัฒนาระบบ คืออะไร

ตอบ ขั้นตอน ที่ใช้ศึกษาขอบข่ายการวิเคราะห์และออกแบบระบบ เป็นวงจรที่แสดงถึงกิจกรรมต่างๆ ในแต่ละขั้นตอนตั้งแต่เริ่มจนเสร็จเป็นระบบงานที่สามารถใช้ได้ ประกอบด้วย

1. การทำความเข้าใจกับปัญหา (Problem definition and Feasibility Study)
2. การรวบรวมข้อมูล (Requirement Gathering)
3. การวิเคราะห์ระบบ (Analysis)
4. การออกแบบระบบ (Design)
5. การพัฒนาระบบ และจัดทำเอกสาร (Development)
6. การทดสอบและบำรุงรักษาระบบ (Testing)
7. การส่งเสริมและการประเมินผลระบบ (Implementation and Maintenance)

2. จงอธิบายหน้าที่ความรับผิดชอบของนักวิเคราะห์ระบบมีอะไรบ้าง

ตอบ นักวิเคราะห์ระบบจะมีหน้าที่เหมือนเป็นส่วนช่วยสนับสนุนผู้เชี่ยวชาญต่างๆ ทำหน้าที่ให้คำแนะนำ เป็นตัวแทนการเปลี่ยนแปลง กระตุ้นให้เกิดการเปลี่ยนแปลงขึ้นในองค์กร

3. จงอธิบายนักวิเคราะห์ระบบคือใคร และควรมีคุณสมบัติใดบ้าง

ตอบ ผู้ที่มีหน้าที่ศึกษาปัญหาและความต้องการขององค์กรในการกำหนดบุคคล ข้อมูล การประมวลผล การสื่อสาร และเทคโนโลยีสารสนเทศว่าจัดการหรือปรับปรุงอย่างไร เพื่อสามารถพัฒนาระบบไปสู่ความสำเร็จได้ คุณลักษณะของนักวิเคราะห์ระบบคือ

1. มีความรู้ทางระบบธุรกิจเพื่อจะสามารถนำมาประยุกต์ใช้ในการวิเคราะห์ระบบ
2. มีความเป็นผู้ฟัง
3. มีมนุษยสัมพันธ์ที่ดี
4. มีความสามารถในการแก้ปัญหา
5. เป็นคนที่มองปัญหาว่าเป็นเรื่องท้าทาย
6. มีความสามารถในการวิเคราะห์ด้านต้นทุนและผลตอบแทน

7. มีความรู้ด้านเขียนโปรแกรม
8. ต้องติดตามเทคโนโลยีอย่างสม่ำเสมอ
9. มีประสบการณ์ทำงานด้านการวิเคราะห์และออกแบบระบบ

4. จงอธิบาย Requirements คืออะไร และสำคัญต่อการพัฒนาระบบงานอย่างไร

ตอบ คือความต้องการสารสนเทศของผู้ใช้ ซึ่งไม่ใช่การผูกติดกับทรัพยากรด้านฮาร์ดแวร์ ซอฟต์แวร์ เครือข่าย ข้อมูล และบุคลากร เป็นสิ่งที่จำเป็นต่อการพัฒนาระบบอย่างยิ่ง เพราะการที่นักวิเคราะห์ระบบจะสามารถออกแบบระบบใหม่ได้นั้น จะต้องศึกษาความต้องการใช้งานของบุคลากรหรือองค์กร เพื่อให้เกิดการพัฒนาระบบที่ตอบสนองความต้องการของบุคลากรหรือองค์กรได้อย่างมีประสิทธิภาพมากที่สุด

5. จงอธิบายโมเดล หรือแบบจำลองการพัฒนาซอฟต์แวร์ คืออะไร

ตอบ การถอดแบบหรือจำลองระบบที่แสดงถึงลักษณะความสัมพันธ์เชิงสาเหตุจากระบบจริงที่สามารถนำไปใช้พัฒนาซอฟต์แวร์ได้

6. จงอธิบาย SDLC มีขั้นตอนการทำงานใดบ้าง

ตอบ 7 ขั้นตอน

1. กำหนดปัญหา จุดมุ่งหมาย และเป้าหมาย
2. กำหนดความต้องการสารสนเทศของผู้ใช้
3. วิเคราะห์ระบบโดยรวมรวบรวมข้อมูลจากความต้องการสารสนเทศของผู้ใช้
4. ออกแบบระบบโดยใช้การตัดสินใจจากฝ่ายบริหาร
5. พัฒนาซอฟต์แวร์และจัดทำเอกสารโดยจะทำงานร่วมกันระหว่างโปรแกรมเมอร์และนักวิเคราะห์ระบบ

6. ทดสอบและบำรุงรักษา จะต้องใช้ข้อมูลจากการปฏิบัติงานจริงมาทดสอบ

7. ดำเนินงานและประเมินผล จะต้องอบรมผู้ใช้ระบบก่อนใช้งานจริงและควรคำนึงถึงผลกระทบต่อผู้ใช้งานระบบและองค์กรนั้น

7. ระบบสารสนเทศสมัยใหม่จะมี เทคโนโลยีที่สำคัญ คือ

1. ระบบฐานข้อมูลขนาดใหญ่ (Big Data System)
 2. ระบบทางวิศวกรรมที่บูรณาการโลกกายภาพ (Cyber-Physical System - CPS)
 3. อินเทอร์เน็ตในทุกสิ่ง (Internet of Things - IoT)
 4. ความมั่นคงปลอดภัยทางไซเบอร์ (Cybersecurity)
 5. คลาวด์คอมพิวติ้ง (Cloud Computing)
 6. AI หรือปัญญาประดิษฐ์ (Artificial Intelligence)
8. จงอธิบายองค์ประกอบของความมั่นคงปลอดภัยทางไซเบอร์ (Cyber Security)
- 1 Network Security: กระบวนการและวิธีปฏิบัติด้านความปลอดภัย ปกป้องเครือข่ายหรือเน็ตเวิร์คจากบุคคลที่สามโดยไม่ได้รับอนุญาต
 - 2 Application Security: กระบวนการและวิธีปฏิบัติด้านแอปพลิเคชันที่ต้องการการอัปเดตและการทดสอบอย่างต่อเนื่องเพื่อให้แน่ใจว่า แอปพลิเคชัน(Application) ปลอดภัยจากการโจมตีของบุคคลที่สาม
 - 3 Data Security: กระบวนการและวิธีปฏิบัติด้านการรักษาความปลอดภัยข้อมูล การปกป้องข้อมูล ตลอดจนการรั่วไหลของข้อมูล ที่อาจเกิดขึ้นกับบริษัท และลูกค้า

คำถามและแบบฝึกหัดท้ายบทที่ 3

1. จงอธิบาย สถาปัตยกรรมองค์กร (Enterprise Architecture: EA) คืออะไร
 - Enterprise หมายถึง องค์กรที่มีการทำธุรกรรม ถูกกำหนดให้เป็นนิติบุคคลที่มีสิทธิดำเนินธุรกิจด้วยตนเอง เช่น เจ้าของสัญญา เป็นเจ้าของทรัพย์สิน ก่อภาระหนี้สิน และสร้างบัญชีธนาคาร อาจเป็นบริษัท กิจบริษัท สถาบันไม่แสวงหาผลกำไร หรือวิสาหกิจที่ไม่ใช่นิติบุคคล
 - Architecture คือ การออกแบบเทคโนโลยีสารสนเทศโดยองค์รวมขององค์กรในทุกมิติ มีการวางแผนเพื่อให้สอดคล้องกับกลยุทธ์ในเชิงธุรกิจขององค์กร
2. จงอธิบายมุมมองหลัก 2 มุมของสถาปัตยกรรมองค์กรสำหรับระบบสารสนเทศสมัยใหม่
 - 2.1 มุมมองเชิงกลยุทธ์ (Business Strategy) จะประกอบไปด้วย มุมมองของผู้บริหาร (Executive) และมุมมองผู้จัดการธุรกิจ (Business Management) ตามกรอบแนวคิดและหลักการในการพัฒนาสถาปัตยกรรมสารสนเทศ องค์กรสากลจะพิจารณาถึง ผู้มีส่วนได้ส่วนเสีย (Stakeholder) ในการสร้างระบบนิเวศของระบบสารสนเทศ ซึ่งในระบบนิเวศของผู้มีส่วนได้ส่วนเสีย (Stakeholder) สามารถแบ่งได้ เป็น 3 กลุ่ม คือ กลุ่มที่ 1 ผู้มีส่วนได้ส่วนเสียหลัก (Key Stakeholders) กลุ่มที่ 2 ผู้มี

ส่วนได้ส่วนเสียขั้นพื้นฐาน (Primary Stakeholders) และกลุ่มที่ 3 ผู้มีส่วนได้ส่วนเสียชั้นรอง (Secondary Stakeholders)

2.2 มุมมองด้านโครงสร้างพื้นฐาน (Infrastructure) มุมมองส่วนนี้อาจจะแยกย่อยออกมาเป็นหลายส่วน เช่น มุมมองสถาปนิก (Architect) มุมมองวิศวกร (Engineer) มุมมองช่างเทคนิค (Technician) ซึ่งในระบบนิเวศของมุมมองด้านโครงสร้างพื้นฐาน

3. จงอธิบายโครงสร้างพื้นฐาน ทั้ง 4 ด้าน

3.1 มุมมองด้านแอปพลิเคชัน (Application) คือมุมมองที่มองระบบงานที่เป็นด้านแอปพลิเคชัน (Application) ทั้งหมดที่เกี่ยวข้องเพื่อรองรับกลยุทธ์ขององค์กรซึ่งรวมทั้งแอปพลิเคชัน (Application) เฉพาะด้านและแอปพลิเคชัน (Application) ที่เป็นพื้นฐานสนับสนุนงานทั้งองค์กร

3.2 มุมมองด้านข้อมูล คือมุมมอง ข้อมูล (Data) ทั้งหมดที่เกี่ยวข้องที่สนับสนุนในการทำธุรกิจ

3.3 มุมมองด้านความมั่นคงปลอดภัย (Security) เป็นมุมมองด้านการรักษาความปลอดภัยทางไซเบอร์ เพื่อให้ระบบสารสนเทศมีความปลอดภัยสำหรับผู้ใช้งานและเจ้าของระบบนั้นจำเป็นต้องมีการคัดกรองผู้ใช้งานและยืนยันตัวตนเพื่อประกอบการพิจารณาของระบบว่าผู้ใช้งานดังกล่าวคือบุคคลที่ระบบได้รองรับไว้ หรือเป็นผู้ไม่ประสงค์ดีที่ต้องการสร้างความเสียหายให้เกิดขึ้นกับระบบ

3.4 มุมมองด้านเวลา คือมุมมองเชิงแผนงานของโครงการ และแผนอนาคต (Roadmap) ด้านสารสนเทศให้สอดคล้องกับแผนการด้านธุรกิจ เพื่อสร้างความได้เปรียบในการแข่งขันทั้งปัจจุบันและอนาคต

4. จงอธิบายประโยชน์ของการใช้กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture: EA)

4.1 การลดต้นทุน (Cost Saving) ทำให้เกิดการลงทุนของธุรกิจที่คุ้มค่า ทั้งนี้เมื่อองค์กรจะมีการออกแบบระบบสารสนเทศใหม่ องค์กรสามารถใช้ EA พิจารณา เพื่อให้ได้ทางออก หรือทางเลือกของสถาปัตยกรรม (Solution Architecture) ที่เหมาะสมตาม สถานการณ์ของอุตสาหกรรม และการเปลี่ยนแปลงของเทคโนโลยีที่รวดเร็ว

4.2 เข้าใจสถานะที่เป็นอยู่คืออะไร (As-Is) โดยกรอบ EA จะช่วยสร้างความเข้าใจและความสัมพันธ์ที่ชัดเจนต่อสถานะปัจจุบันสถาปัตยกรรมองค์กร (Current State of Enterprise Architecture) ทั้ง 2 ด้านหลัก คือ ด้านธุรกิจ (Business) และ ด้านโครงสร้างพื้นฐาน (Infrastructure)

4.3 ใช้ออกแบบอนาคตของสถาปัตยกรรมองค์กร (Future State of Enterprise Architecture) ได้ครบถ้วน การมีกรอบ EA ทำให้องค์กรสามารถออกแบบโดยการนำเป้าหมายเชิงธุรกิจ (Business Architecture) มา นำเป้าหมายทางสารสนเทศเพราะนำมุมมองของผู้บริหารมาร่วมด้วย

4.4 องค์กรสามารถใช้ EA ประเมินสถานภาพความแข็งแกร่งตามหลักของ Balance Score Card ตามหลักของการบริหารตามแนวทางของ Balanced Scorecard คือการแปลงวิสัยทัศน์และกลยุทธ์ขององค์กรให้ออกมาเป็นตัวชี้วัดต่าง ๆ ใน 4 ด้าน ได้แก่ ด้านการเงิน ด้านลูกค้า ด้านกระบวนการภายในธุรกิจหรือระบบสารสนเทศ

5. จงอธิบาย 4 องค์ประกอบ ของ TOGAF

5.1 สถาปัตยกรรมธุรกิจ (Business Architecture: BA)

5.2 สถาปัตยกรรมข้อมูล (Data Architecture: DA)

5.3 สถาปัตยกรรมด้านโปรแกรมประยุกต์ (Application Architecture: AA) และ

5.3 สถาปัตยกรรมเทคโนโลยี (Technology Architecture: TA)

6. จงอธิบายประโยชน์ของสถาปัตยกรรมองค์กร TOGAF คือ

6.1 แสดงกลยุทธ์ (Strategy) ขององค์กร แผนผังองค์กร (Organization Chart)

6.2 แสดงโครงสร้างข้อมูล (Information Architecture) แสดงการจัดเก็บข้อมูล การแลกเปลี่ยนข้อมูล โดยทำหน้าที่เป็นพิมพ์เขียวที่มีรายละเอียดครบทุกมุม

6.3 เป็นเครื่องมือสื่อสารระหว่างผู้ที่เกี่ยวข้องกับการออกแบบและพัฒนาระบบให้มีความเข้าใจตรงกัน

6.4 สถาปัตยกรรมองค์กร TOGAF ช่วยแก้ปัญหาของการลงทุนระบบที่ซ้ำซ้อนอันเกิดจากการบริหารจัดการระบบสารสนเทศอย่างไม่เป็นระบบ

คำถามและแบบฝึกหัดท้ายบทที่ 4

1. จงอธิบายปัจจัยเพื่อพิจารณาสำหรับการบริหารโครงการสารสนเทศสมัยใหม่ มีอะไรบ้าง

คำตอบ 1. ปัจจัยแวดล้อมทางเศรษฐกิจมหภาค (Macroeconomic environments) 2. ปัจจัยด้านเทคโนโลยีสมัยใหม่ 3. ปัจจัยด้านงบประมาณขององค์กร+

2. จงอธิบายกรอบการบริหารโครงการสารสนเทศสมัยใหม่ มีอะไรบ้าง

คำตอบ 1. กรอบโมเดลการปฏิบัติงาน (Operating Model) 2. กรอบสถาปัตยกรรมองค์กร (Enterprise Architecture) 3. กรอบรูปแบบพันธสัญญาสู่ความสำเร็จ (Engagement Model)

3. จงอธิบายผังงาน (Flowchart) คืออะไร

คำตอบ รูปภาพ หรือ สัญลักษณ์ ที่ใช้เขียนแทนขั้นตอนของงานหรือสิ่งที่กำลังอธิบาย ที่อาจประกอบไปด้วยคำอธิบาย ข้อความ หรือคำพูด แบบย่อ ของงานหนึ่งๆ โดยใช้สัญลักษณ์ที่เป็นมาตรฐานเดียวกัน

4. จงอธิบายลักษณะของผังงานที่ดี มีอะไรบ้าง จงยกตัวอย่าง

คำตอบ

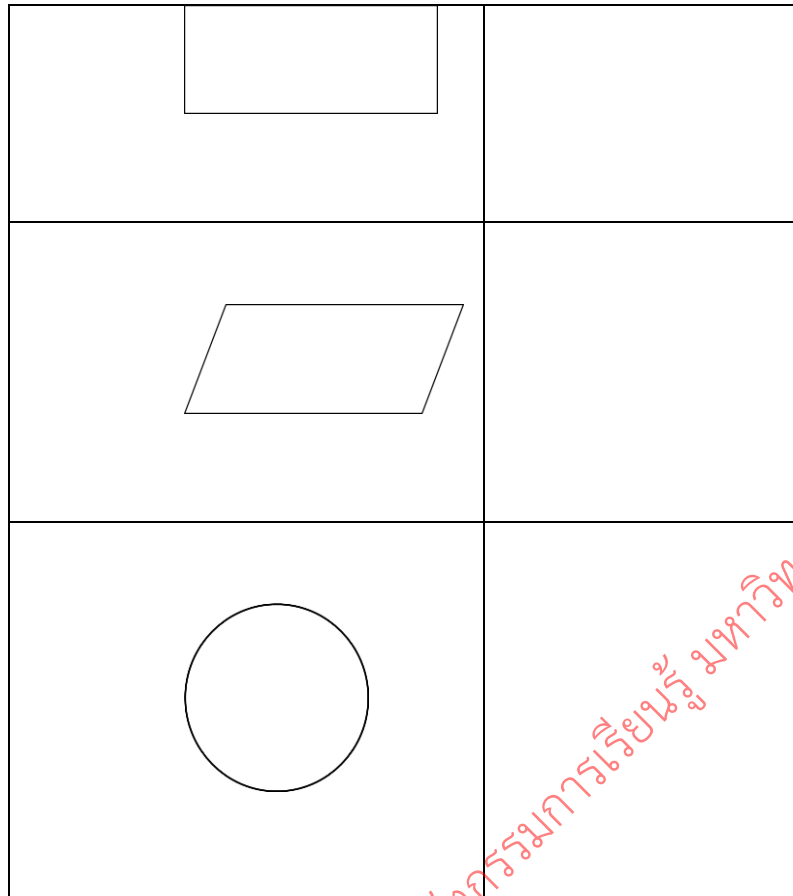
- ใช้สัญลักษณ์ตามที่กำหนดไว้
- มีการแบ่ง Swimming Lane หรือ BOX เพื่อระบุ คน หรือหน่วยงานที่เกี่ยวข้อง
- ใช้ลูกศรแสดงทิศทางการไหลของข้อมูลจากบนลงล่าง หรือ จากซ้ายไปขวา
- คำอธิบายในภาพควรสั้นกะทัดรัด และ เข้าใจง่าย
- ทุกแผนภาพ ผังงาน ต้องมีลูกศรแสดงทิศทางเข้า – ออก
- ไม่ควรโยงเส้นเชื่อมผังงานที่อยู่ไกลมาก ๆ ควรใช้สัญลักษณ์จุดเชื่อมต่อแทน
- การเขียนชื่อ ผังงาน ควรเขียน วันเวลา และวันแก้ไข ผังงาน รวมทั้งผู้จัดทำ

ผังงาน และชื่อเอกสาร (File Name)

- ขนาดของสัญลักษณ์ ก็ต้องมีขนาด เท่ากันหรือใกล้เคียงกัน
- ลูกศรแสดงทิศทางเข้า – ออก ต้องไม่ ทับกัน จน มองเส้นทางของ งาน ไม่ชัดเจน

5. จงอธิบายสัญลักษณ์เหล่านี้คือ



6. จงอธิบายรูปแบบของผังงาน มีกี่แบบ

คำตอบ 3 แบบ คือ 1. ผังงานเรียงลำดับ (Sequence) 2. ผังงานแบบมีเงื่อนไข (Selection) 3. ผังงานแบบทำซ้ำ (Iteration)

7. แผนภาพกระแสข้อมูล (Data Flow Diagram: DFD) คือ

คำตอบ แผนภาพแสดงการไหลของข้อมูลจากแหล่งข้อมูลภายนอกเข้ามาในระบบ เพื่อแสดงวิธีการไหลของข้อมูลจากกระบวนการหนึ่งไปอีกกระบวนการหนึ่ง ซึ่งมีสัญลักษณ์ 4 แบบคือ สี่เหลี่ยม สี่เหลี่ยมมุมมน ลูกศร และสี่เหลี่ยมปลายเปิดข้างหนึ่ง โดยแผนภาพกระแสข้อมูลจะมีหลายระดับ (Level) ตามความลึกของข้อมูล

8. แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล (Entity Relationship Diagram : E-R Diagram) คือ

คำตอบ แบบจำลองข้อมูล (Data Model) คือแผนภาพที่แสดงการจำลองข้อมูลที่เกิดขึ้นทั้งหมดในระบบ พร้อมทั้งจำลองความสัมพันธ์ระหว่างข้อมูลที่เกิดขึ้นนั้น โดยใช้ แผนภาพ ER-Diagram ประกอบไปด้วย Entity (แทนกลุ่มของข้อมูลที่เป็นเรื่องเดียวกัน/เกี่ยวข้องกัน) และ ความสัมพันธ์ระหว่างข้อมูล (Relationship) ที่เกิดขึ้นทั้งหมดในระบบ

9. จงอธิบายแผนภาพ Use Case Diagram คือ

คำตอบ แบบจำลองแผนภาพที่ใช้แสดงปฏิสัมพันธ์ระหว่างกิจกรรมที่เกิดขึ้นในระบบระบบงานและ Actor หรือผู้ที่กระทำกับระบบ และแสดงให้เห็นถึงส่วนประกอบทั้งหมด หรือ ภาพรวมของระบบ เป็นรากฐานในการเริ่มต้นการวิเคราะห์ระบบ โดยค้นหาคำว่าระบบทำอะไร โดยไม่สนใจกลไกการทำงานหรือเทคนิคการทำงาน

คำถามและแบบฝึกหัดท้ายบทที่ 5

1. ระบบสารสนเทศจะสมบูรณ์แบบต้องประกอบไปด้วย 5 องค์ประกอบอะไรบ้าง

คำตอบ

1. บุคลากรผู้ใช้งานระบบ (User)
2. การประมวลผลข้อมูล (Data Processing)
3. การสื่อสารข้อมูล (Data Communication)
4. ที่จัดเก็บข้อมูลและการเข้าถึงข้อมูล เพื่อที่จะเก็บข้อมูลในรูปแบบที่เหมาะสม และเพื่อให้แน่ใจว่าข้อมูลข่าวสารสามารถนำมาใช้ได้เมื่อต้องการ (Hardware and Software)
5. ขั้นตอนการใช้งานที่มีประสิทธิภาพ เพื่อรวบรวมข้อมูลทุกส่วน การประมวลผลข้อมูล การสื่อสารข้อมูลที่จัดเก็บข้อมูลและการเข้าถึงข้อมูล รวมทั้งผู้ใช้เข้าเป็นการจัดการสารสนเทศที่มีประสิทธิภาพ (Process)

2. โมเดลขั้นตอนการพัฒนาซอฟต์แวร์ (Software) คือ

คำตอบ แนวทางการสร้างผลิตภัณฑ์ให้ตรงตามข้อกำหนดด้านสารสนเทศ ด้วยการนำกรรมวิธีการพัฒนาซอฟต์แวร์มาใช้ เพื่อให้กระบวนการพัฒนาซอฟต์แวร์นั้นมีมาตรฐาน และตัวผลิตภัณฑ์ซอฟต์แวร์มีคุณภาพ มีขั้นตอนการตรวจสอบความถูกต้อง (Software Validation) ของ software ให้ตรงกับความต้องการ เพื่อให้แน่ใจว่าซอฟต์แวร์ที่พัฒนานี้เป็นผลิตภัณฑ์ที่ตรงกับความต้องการของลูกค้าหรือผู้ใช้งาน รวมถึงกระบวนการทำให้ระบบสารสนเทศ มีวิวัฒนาการ เป็นการปรับเปลี่ยนเพิ่มสิ่งดีๆเข้ามาและเอาสิ่งที่ไม่ดีหรือไม่จำเป็นออกไป (Software Evolution) ซึ่งจะเปลี่ยนแปลงไปตามเทคโนโลยีหรือตามความต้องการของผู้ใช้

3. คุณสมบัติของระบบสารสนเทศที่ดี คือ

คำตอบ

- Understandability มีขอบเขตกระบวนการที่ชัดเจนและง่ายต่อความเข้าใจ
- Visibility สนับสนุนให้กิจกรรม กระบวนการธุรกิจทำงานอย่างมีประสิทธิภาพที่สุด
- Supportability เครื่องมือช่วยในการวิเคราะห์ สามารถช่วยสนับสนุนกิจกรรม กระบวนการตัดสินใจในทุกระดับ
- Acceptability กระบวนการที่มีการตรวจสอบ ยอมรับ โดยผู้ใช้
- Reliability กระบวนการถูกออกแบบในแนวทางที่หลีกเลี่ยงที่จะส่งต่อความผิดพลาดของซอฟต์แวร์
- Robustness กระบวนการสามารถทำงานต่อได้ แม้ว่ามีปัญหาที่ไม่ได้คาดการณ์ไว้เกิดขึ้น
- Maintainability กระบวนการสามารถพัฒนาการเพื่อตอบสนองต่อการเปลี่ยนแปลงความต้องการขององค์กร
- Rapidity กระบวนการสามารถทำให้ส่งมอบผลิตภัณฑ์ได้รวดเร็ว ตรงตามคุณลักษณะที่ได้กำหนดไว้

4. กรอบมาตรฐานของ COBIT ประกอบด้วย 5 กลุ่มใหญ่คือ

คำตอบ

1. การวางแผนและการจัดการองค์กร (PO - Planning and Organization)
2. การจัดหาและการนำระบบออกใช้งานจริง (AI - Acquisition and Implementation)
3. การส่งมอบและการสนับสนุน (DS - Delivery and Support)
4. การตรวจสอบและประเมินผล (ME - Monitoring and Performance Evaluation)
5. โมเดลน้ำตก (Waterfall Model) จัดอยู่ในโมเดลแบบไหน และจงอธิบาย

คำตอบ แบบ Linear Model เป็นแบบจำลองที่มีขั้นตอนเป็นลำดับ เป็นโมเดลพื้นฐานในการพัฒนาซอฟต์แวร์และระบบสารสนเทศ โดยต้องทำขั้นตอนแรกให้สำเร็จจึงจะดำเนินการขั้นตอนต่อไปได้ แต่ละขั้นจะถูกกำหนดงานและผลิตผลที่ได้จากขั้นตอนนั้นๆอย่างชัดเจน

6. จงอธิบายโมเดลแบบ Iterative Model & Incremental

คำตอบ โมเดลแบบ Iterative Model & Incremental จะแบ่งงานออกเป็น Phase แล้วในแต่ละ Phase งานก็จะถูกแบ่งเป็นงานย่อยๆเหมือน Waterfall Model เมื่อจบในแต่ละ Phase ก็จะทำให้ลูกค้าได้ Review ระบบงานดังนั้นระบบงานนี้จะช่วยลดความเสี่ยงในการทำโปรเจกต์ที่ผิดพลาดไม่ตรงตามความต้องการ และใน Phase ต่อๆไปก็จะมีมีความสำคัญ (Priority) ลดลงตามลำดับ โมเดลนี้มีลักษณะเด่น คือ การเก็บข้อมูลความต้องการของผู้ใช้งานในตัวโปรแกรมครั้งแรก แล้วนำ ข้อมูลนั้นมาวิเคราะห์เพื่อพัฒนาต่อ เหมาะสำหรับ โครงการที่มีความเสี่ยงปานกลางถึงต่ำ โครงการที่มีการกำหนดระยะเวลาการทำงานที่ยาวนาน หรือโครงการที่ใช้เทคโนโลยีใหม่ๆ เข้าช่วยในการพัฒนาระบบสารสนเทศ

7. จุดสำคัญของโมเดลแบบ Evolutionary Model คือ

คำตอบ มีการพัฒนาระบบและตรวจสอบงานไปด้วย (Debugging) เพื่อให้แน่ใจว่าระบบงานที่ได้พัฒนาขึ้นมีความสมบูรณ์และมีคุณภาพเพียงพอที่จะดำเนินงานในขั้นตอนต่อไป เป็นโมเดลที่นำ Water Fall Model มาปรับปรุงให้มีประสิทธิภาพยิ่งขึ้น

8. แนวคิดที่สำคัญของ โมเดลแบบจำลอง Agile คือ

คำตอบ

- การพัฒนาระบบมีการร่วมมือกับลูกค้าในการเจรจาสัญญา และตอบสนองต่อการเปลี่ยนแปลงตามแผน **ยอมรับความเปลี่ยนแปลง** เพราะความต้องการของระบบ เช่นฟังก์ชันต่างๆ อาจเปลี่ยนแปลงได้ตลอด
- การพัฒนาระบบไม่เน้นกระบวนการและเอกสาร โดยเน้นไปที่การพัฒนาโปรแกรมให้ดีที่สุด ตรงตามความต้องการลูกค้า โดยพัฒนาโปรแกรมทีละส่วน และทำต่อเนื่องทีละเล็กทีละน้อย สามารถส่งมอบระบบทีละส่วนได้ ทำให้มองเห็นจุดที่ต้องแก้ไขได้เร็ว
- Agile Model มุ่งเน้นที่การทำให้ขั้นตอนการพัฒนาและการทดสอบคล่องตัว หากพบข้อผิดพลาดจะแก้ไขให้เร็ว เพื่อให้ระบบสามารถนำไปใช้งานได้ (Production)

- โมเดลแบบจำลอง Agile มุ่งเน้นการสื่อสารที่ชัดเจนระหว่างสมาชิกในทีมและลูกค้า ทำให้ระบบมีภาพที่ชัดเจนในทุกระยะของการพัฒนา เน้นการมีปฏิสัมพันธ์ระหว่างบุคคลมากกว่าตามกระบวนการ เช่น การประชุมสามารถนำ ทีมงานไปประชุมกับลูกค้า เพื่อจะได้รับข้อมูลที่ชัดเจน โดยไม่ต้องรอการสื่อสารหรือเอกสารจากนักวิเคราะห์ระบบ (System Analysis)

9. ทีมงาน 3 ตำแหน่งคืออะไร และมีหน้าที่อะไรบ้าง

คำตอบ คือ **Product owner, Scrum master และ Development team** โดยทำงานกันแบบ “Cross functional” สามารถทำงานข้ามสายงานบนความร่วมมือของทีมงาน โดย

- Product Owner ทำหน้าที่เป็นตัวแทนของทีม โดยบุคคลนี้จะต้อง มีความเข้าใจในธุรกิจ ลูกค้า และทำหน้าที่ทำให้เกิดการสื่อสารภายในทีม ทำให้แต่ละงานที่อยู่ใน Product Backlog มีความชัดเจน ประเมินผลและจัดลำดับความสำคัญ (Priorities and Tasks) ของงานต่าง ๆ ให้กับทีมด้วย

- Scrum Master ทำหน้าที่ ช่วยแก้ปัญหาต่างๆ ที่เกิดขึ้นในระหว่างการพัฒนา ระบบ โดยช่วยกำจัดอุปสรรคที่ขัดขวางไม่ให้ทีมบรรลุเป้าหมาย

- Development team คือบุคคลที่ทำหน้าที่ Designer, Programmer, UI/UX, Testing โดยทีมหนึ่งจะประกอบไปด้วยทีมงานประมาณ 2 ขึ้นไป ขึ้นอยู่กับขนาดของโครงการ โดยแต่ละคน แต่ละตำแหน่งจะทำงานแบบ Self-Management คือบริหารจัดการด้วยตนเอง รับผิดชอบงานตนเองให้ตรงต่อเวลา และส่งมอบ Tasks and Requirement ที่ได้รับมอบหมายจากการประชุมร่วมกันแต่ละสัปดาห์

10. Daily Scrum Meeting คืออะไร

คำตอบ เป็นขั้นตอนการประชุมสั้นๆ ภายในทีมประมาณ 10-15 นาทีตลอดระยะเวลาโครงการที่ทำการ Sprint เพื่อปรึกษาทบทวนการทำงาน สิ่งที่ทำไปแล้ว สิ่งที่จะทำต่อไป รวมถึงปัญหาต่างๆ และมีการแก้ไขปัญหาอย่างต่อเนื่อง

11. มาตรฐาน COBIT คืออะไร

คำตอบ COBIT ย่อมาจาก Control Objectives for Information and Related Technology เป็นกรอบที่สร้างขึ้นโดย ISACA สำหรับการกำกับดูแลและการจัดการด้านสารสนเทศ ใช้เป็นเครื่องมือในการบริหารจัดการเพื่อสร้างความมั่นใจว่าการใช้ทรัพยากรด้านเทคโนโลยีสารสนเทศ ให้สอดคล้องกับวัตถุประสงค์เชิงธุรกิจขององค์กร มุ่งเน้นใน 5 ด้าน คือ

- การจัดวางกลยุทธ์ (Strategic Alignment)
- การนำเสนอคุณค่า (Value Delivery)
- การจัดการทรัพยากร (Resource Management)
- การจัดการความเสี่ยง (Risk Management)
- การวัดผลการดำเนินงาน (Performance Measurement)

คำถามและแบบฝึกหัดท้ายบทที่ 6

1. จงอธิบายองค์ประกอบที่สำคัญของสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่

คำตอบ

1) การกำกับดูแลการบริหารองค์กร (Enterprise Governance)

2) ส่วนองค์ประกอบของสถาปัตยกรรมและกรอบการบริหารจัดการเทคโนโลยีดิจิทัล ประกอบไปด้วย สถาปัตยกรรมด้านธุรกิจ (Business Architecture) สถาปัตยกรรมด้านข้อมูล (Data Architecture) สถาปัตยกรรมด้านโปรแกรมประยุกต์ (Application Architecture) สถาปัตยกรรมด้านเทคโนโลยี (Technology Architecture) สถาปัตยกรรมด้านการรักษาความปลอดภัยข้อมูล (Security Architecture) และการเฝ้าระวัง และบริหารจัดการระบบ (System Monitoring and Management)

2. จงอธิบาย ประเภทของข้อมูลสารสนเทศสมัยใหม่ประกอบไปด้วยอะไรบ้าง

คำตอบ

2.1 ข้อมูลที่มีโครงสร้างชัดเจน (Structured Data) หมายถึง ชุดข้อมูลที่มีการจัดเรียงโครงสร้างอย่างเป็นระเบียบ มีความชัดเจน หรือเป็นตัวเลข พร้อมใช้งานได้ทันที

2.2 ข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) หมายถึง ชุดข้อมูลที่มีโครงสร้างไม่ชัดเจน หรือไม่สามารถระบุความแน่นอนของข้อมูลนั้นๆ ได้ การประมวลผลจำเป็นต้องใช้เครื่องมือเพื่อแปลงให้เป็นข้อมูล Structure data ก่อนนำไปใช้งาน

2.3 ข้อมูลกึ่งมีโครงสร้าง (Semi-Structured Data) หมายถึง ชุดข้อมูลที่มีโครงสร้างระดับหนึ่งแต่ยังไม่สมบูรณ์ ซึ่งอยู่ตรงกลางระหว่าง Structured data และ Unstructured Data โดยมี

ตัวกำกับไว้ว่าข้อมูลนี้คืออะไร ข้อมูลนี้อยู่ในลำดับขั้นต่อจากอะไร ทำให้สามารถเอาประมวลผลได้ง่ายกว่าข้อมูลแบบไร้โครงสร้าง

3. จงอธิบายคุณลักษณะ 7 ประการของข้อมูลสารสนเทศสมัยใหม่

คำตอบ

1. ปริมาณ (Volume) หมายถึง ข้อมูลในสมัยปัจจุบันเป็นข้อมูลปริมาณมหาศาลที่เกิดจากการเก็บของระบบหลายๆระบบ

2. ความหลากหลาย (Variety) หมายถึง รูปแบบของข้อมูลในแต่ละประเภทที่มีความหลากหลาย ทั้งแบบโครงสร้าง (Structured) กึ่งโครงสร้าง (Semi Structured) ไม่มีโครงสร้าง (Unstructured)

3. ความเร็ว (Velocity) หมายถึง คุณลักษณะข้อมูลที่ถูกสร้างขึ้นอย่างรวดเร็วต่อเนื่อง และเป็นแบบวินาที (Real Time) ซึ่งในที่นี้จะหมายถึงความเร็วในการเพิ่มขึ้นของข้อมูลจากแหล่งข้อมูล ทำให้องค์กรธุรกิจเห็นข้อมูลและสามารถวิเคราะห์ข้อมูลแบบ Real-Time

4. ความถูกต้อง (Veracity) หมายถึง ข้อมูลต้องมีความน่าเชื่อถือ ทราบแหล่งที่มา และมีความถูกต้อง มีกระบวนการในการตรวจสอบและยืนยันความถูกต้องของข้อมูล ซึ่งมีผลโดยตรงกับการวิเคราะห์ข้อมูล เนื่องจากระบบจะประมวลผลจากข้อมูลที่นำเข้าเท่านั้น การนำขยะเข้าจะได้ผลลัพธ์ขยะออก (Garbage In Garbage Out, GIGO) เป็นคำที่มักกล่าวเตือนสติในงานวิเคราะห์ของระบบสารสนเทศ “ถ้าใส่ข้อมูลที่เราคุณค่าลงไปในระบบงานคอมพิวเตอร์ก็จะได้ข้อมูลที่ไม่มีความกลับมา”

5. คุณค่า (Value) หมายถึง ข้อมูลที่มีคุณค่าต่อการนำไปใช้งาน มีประโยชน์ต่อการตัดสินใจของธุรกิจ โดยข้อมูลที่มีประโยชน์จะต้องตอบวัตถุประสงค์ทางธุรกิจ แก้ไขปัญหา และใช้วางแผนกลยุทธ์

6. ความแปรผันได้ (Variability) หมายถึง ความแปรปรวนในบริบทของข้อมูลขนาดใหญ่ มีบางสิ่งที่แตกต่างกัน เช่น จำนวนของความไม่สอดคล้องกันในข้อมูล จำเป็นต้องค้นพบสิ่งเหล่านี้โดยวิธีการตรวจจับความผิดปกติและค่าผิดปกติเพื่อให้การวิเคราะห์ที่มีความหมายเกิดขึ้นข้อมูล โดยคุณลักษณะของข้อมูลแบบนี้ คือสามารถในการเปลี่ยนแปลงรูปแบบไปตามการใช้งานของธุรกิจและมุมมองการวิเคราะห์ โดยสามารถคิดวิเคราะห์ข้อมูลที่มีอยู่ได้จากหลายแง่มุม ข้อมูลขนาดใหญ่ยังแปรผันได้เนื่องจากมีข้อมูลจำนวนมากที่เกิดจากชนิดข้อมูลและแหล่งที่มาที่แตกต่างกันหลายประเภท ความแปรปรวนยังหมายถึงความเร็วที่ไม่สอดคล้องกันที่ไหลตข้อมูลขนาดใหญ่ลงในฐานข้อมูล

7. สร้างภาพได้ (Visualization) จากรูป 6.2 งานหลักในกล่องสุดท้ายสำหรับระบบประมวลผล Big Data คือการแปลงขนาดมหึมา ให้เข้าใจได้ง่ายและนำไปใช้ได้ สำหรับการจัดการของธุรกิจ ซึ่งก็คือการแปลงเป็นรูปแบบกราฟิก ดังนั้นข้อมูลที่ได้มาต้องสามารถนำมาเสนอในรูปแบบของรายงานที่หลากหลายได้โดยเฉพาะในรูปแบบกราฟิก (Graphical Formats)

4. จงอธิบายการทำ ETL (Extract-Transform-Load)

คำตอบ

เทคนิคในการรวบรวมข้อมูล แปลง (Transformation) ข้อมูลให้อยู่ในรูปแบบที่ประมวลผลได้ในระบบคลังข้อมูล (Data Warehouse) โดยระบบที่ออกแบบเอาไว้จะดึงข้อมูลออกมาจากหลายๆ ที่ มีกระบวนการตรวจสอบคุณภาพของข้อมูลก่อนมาประยุกต์ใช้ มีการเชื่อมโยงและปรับข้อมูลให้เป็นไปในรูปแบบเดียวกันเพื่อให้ ข้อมูลจากหลายๆ แหล่งสามารถใช้งานร่วมกันได้

5. จงอธิบายเลเยอร์การวิเคราะห์ข้อมูล (Analytics Layer) และ เลเยอร์การค้นหาคำรู้จากข้อมูล (Information Exploration Layer) ทำหน้าที่อะไรบ้าง

คำตอบ

เลเยอร์นี้มีกระบวนการของการประมวลผลข้อมูลทุกประเภทและดำเนินการวิเคราะห์ที่เหมาะสม ในเลเยอร์นี้การวิเคราะห์ข้อมูลสามารถแบ่งออกเป็นสามองค์ประกอบหลักคือ หนึ่ง การจัดสรรทรัพยากรในการจัดเก็บและประมวลผล สองการประมวลผลแบบสตรีม (Stream Computing) และสามารถวิเคราะห์ในฐานข้อมูล (In-Database Analytics)

6. จงอธิบายเลเยอร์การกำกับดูแลข้อมูล (Data governance layer) ทำหน้าที่อะไรบ้าง

คำตอบ

เลเยอร์นี้ประกอบด้วยกระบวนการจัดการข้อมูลหลัก (Master Data Management: MDM) การจัดการวงจรชีวิตของข้อมูล และการรักษาความปลอดภัยข้อมูล (Data Security) และการจัดการความเป็นส่วนตัว (Data Privacy) เลเยอร์นี้เน้น "วิธีการ" เช่นเดียวกับวิธีควบคุมข้อมูลในองค์กร องค์ประกอบแรกของการกำกับดูแลข้อมูลและการจัดการกำกับดูแลข้อมูล ถือเป็นกระบวนการที่ใช้กำกับดูแล กำหนดนโยบาย สร้างกรอบมาตรฐาน และเครื่องมือสำหรับการจัดการข้อมูล ข้อมูลจะถูกทำให้เป็นมาตรฐาน มีนโยบายในการจัดการลบข้อมูล นำเข้านำออก และรวมเข้าด้วยกันเพื่อสร้างความ

รวดเร็ว ความสมบูรณ์ ความถูกต้อง และความพร้อมใช้งานของข้อมูลสำหรับสนับสนุนการวิเคราะห์ข้อมูลและการตัดสินใจ

7. จงอธิบายประโยชน์ของสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่มีอะไรบ้าง

คำตอบ

- ลดความซับซ้อน Decreasing Complexity การออกแบบสถาปัตยกรรมองค์กรจะช่วยลดความซับซ้อนของไอที ระบบที่ซับซ้อนซึ่งจัดการได้ยาก
- สร้างมาตรฐานให้กับองค์กร (Standardization) พิมพ์เขียวของสถาปัตยกรรมองค์กรสร้างกระบวนการขององค์กรและมาตรฐานกระบวนการในการดำเนินธุรกิจ ช่วยกำหนดระเบียบวินัยและการกำหนดมาตรฐานของกระบวนการทางธุรกิจ
- ลดเวลาที่เสียไป (Reducing Wasted Time and Cost Saving) สถาปัตยกรรมองค์กรช่วยประหยัดเวลาสำหรับพนักงานที่บริหารจัดการระบบสารสนเทศ ทำให้สามารถตรวจพบสาเหตุของปัญหาได้เร็วขึ้น ซึ่งช่วยลดการสูญเสียเวลาและแรงงานที่เกิดจากการหยุดชะงัก
- การวิเคราะห์การเปลี่ยนแปลง (Change Analysis) พิมพ์เขียวของสถาปัตยกรรมองค์กรเป็นเครื่องมือช่วยวิเคราะห์และตอบสนองต่อการเปลี่ยนแปลงในอุตสาหกรรมได้เร็วขึ้น สามารถใช้ประกอบการตัดสินใจเชิงกลยุทธ์และปรับกระบวนการด้านสารสนเทศให้เข้ากับโมเดลทางธุรกิจใหม่หรือสร้างความยั่งยืนเพื่อรองรับตามการเปลี่ยนแปลงของโลกธุรกิจ
- ความปลอดภัยที่แข็งแกร่ง (Strong Security) สถาปัตยกรรมองค์กรเป็นพิมพ์เขียวที่อธิบายโครงสร้างของระบบที่ช่วยระบุนมาตรการที่ใช้กับการละเมิดความปลอดภัยในโลกไซเบอร์ในปัจจุบัน เป็นหนึ่งในเครื่องมือที่สำคัญที่สุด ในการปกป้องข้อมูลด้วยการเสริมความปลอดภัยทางไซเบอร์ของธุรกิจ

คำถามและแบบฝึกหัดท้ายบทที่ 7

1. จงอธิบายกิจกรรม Support Activity

คำตอบ กิจกรรมที่สนับสนุนการดำเนินงานหลังบ้านของบริษัท (Back End Process) เช่น ฝ่ายบัญชี (Accounting) ฝ่ายการพัฒนาเทคโนโลยี (Technology Development) ฝ่ายทรัพยากร

บุคคล (Human Resource Management) ฝ่ายการจัดซื้อ (Purchasing) ฝ่ายคลังสินค้า (Inventory) เป็นต้น

2. จงอธิบายกิจกรรม Primary Activity

คำตอบ กิจกรรมหลักของบริษัท เช่น การนำวัตถุดิบเข้ามา (Inbound Logistics) ผลิต (Operation) เตรียมจัดจำหน่าย (Outbound Logistics) การตลาด (Marketing) การขาย และบริการหลังการขาย (Sale and Service) เป็นต้น

3. จงอธิบายแนวทางการพัฒนาสถาปัตยกรรมองค์กรและระบบสารสนเทศสมัยใหม่ 4 ขั้นตอน

คำตอบ

ขั้นตอนที่ 1 กระบวนการเตรียมการ (Initial)

ขั้นตอนที่ 2 กระบวนการออกแบบ (Design)

ขั้นตอนที่ 3 กระบวนการการปรับเปลี่ยน (Transition) และกระบวนการติดตั้ง (Implement)

ขั้นตอนที่ 4 การวัดประสิทธิภาพองค์กร (Measurement and Continuous Improvement)

4. จงอธิบาย ผู้มีส่วนได้ส่วนเสีย (Stakeholder) ได้แก่

คำตอบ ผู้ใช้งาน และเจ้าของบริษัท

5. จงอธิบายบทบาทของเจ้าของโครงการ (Product Owner: PO)

คำตอบ เป็นนักวางกลยุทธ์ทางธุรกิจ ผู้ออกแบบผลิตภัณฑ์ นักวิเคราะห์ตลาด ผู้ประสานงานกับลูกค้า และผู้จัดการโครงการ รวมถึงการกำหนดวิสัยทัศน์ วางแผนงาน จัดลำดับความสำคัญของความต้องการ ดูแลขั้นตอนการพัฒนา คาดการณ์ความต้องการของลูกค้า ทำหน้าที่เป็นผู้ประสานงานระหว่างทีมและผู้มีส่วนได้ส่วนเสีย และประเมินความคืบหน้าในการดำเนินโครงการ

6. จงอธิบาย ทีมพัฒนา (Developer: Dev)

คำตอบ ผู้พัฒนาระบบสารสนเทศตามที่ออกแบบไว้ให้เกิดขึ้นจริง หรือกลุ่มบุคคลที่ทำงานร่วมกันเพื่อส่งมอบระบบ ตามกำหนดเวลาและด้วยคุณภาพที่กำหนดไว้

7. จงอธิบายนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile

คำตอบ เป็นบุคคลที่ผู้ควบคุมดูแลการพัฒนาระบบสารสนเทศโดยรวม โดยทำการศึกษา ปัญหาซับซ้อนที่เกิดขึ้นในระบบและแยกแยะปัญหาเหล่านั้นอย่างมีหลักเกณฑ์ นักวิเคราะห์ระบบหรือที่เราเรียกกันว่า System Analyst หรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) ในระบบ Agile จะต้องหาวิธีการมาแก้ไขปัญหาที่แยกแยะเหล่านั้น พร้อมทั้งให้เหตุผลประกอบด้วย นักวิเคราะห์ระบบหรือผู้จัดการโครงการจะต้องกำหนดขอบเขตของการวิเคราะห์ และจะต้องกำหนด จุดมุ่งหมายหรือเป้าหมายในการวิเคราะห์ออกแบบระบบสารสนเทศนั้นด้วย นอกจากนี้และยังต้องทำความเข้าใจในโครงสร้างลักษณะขององค์กรนั้นๆ ในด้านต่างๆ เช่น การไหลของงาน อำนาจการบังคับบัญชา เป็นต้น

8. คุณลักษณะนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) คือ

คำตอบ

- ผู้เชี่ยวชาญในทักษะทางเทคนิคและเทคโนโลยีที่จำเป็นสำหรับโครงการ ควรจะมี ความรู้ทางด้านภาษาการเขียนโปรแกรมมาก่อน
- เป็นนักสื่อสารที่ดีที่จะช่วยบริหารงานทีมให้ประสบความสำเร็จ ควรจะมีมนุษยสัมพันธ์ที่ดี เนื่องจากนักวิเคราะห์ระบบจะต้องเกี่ยวข้องกับคนในทุกระดับในองค์กร รวมถึงระบบปฏิบัติการ ช่วงเทคนิค พนักงานบัญชี เลขานุการ พนักงานธุรการ
- เป็นนักวางแผนที่ดีเพื่อให้งานพัฒนาสามารถดำเนินการได้ตรงเวลาในแต่ละช่วงของ วงจรการพัฒนา
- นักวิเคราะห์ผู้จัดการโครงการ ระบบเปรียบเหมือนผู้จัดการทั่วไป จะเป็นผู้ที่ตัดสินใจ ในการกำหนดออกแบบระบบทั้งหมด เป็นผู้ติดตามประเมินระบบที่ออกแบบ และติดตั้ง ว่าได้รับผลตามวัตถุประสงค์ที่วางไว้ตั้งแต่ต้นหรือเปล่า รวมทั้งการ ประเมินออกมาเป็นตัวเลข เพื่อชี้แจงให้ผู้ที่เกี่ยวข้องออกแบบระบบเข้าใจ

9. ความคาดหวังต่อนักวิเคราะห์ระบบหรือผู้จัดการโครงการ (Proposal Manager/Scrum Master) คือ

คำตอบ

- รวบรวมข้อเท็จจริงและข้อมูลเพื่อใช้ในโครงการพัฒนาระบบสารสนเทศ
- ทำการตัดสินใจ จะจัดลำดับความสำคัญของโครงการ
- เสนอทางเลือกในการบริหารจัดการโครงการ

- วาดข้อกำหนดส่วนผู้ใช้งานและเพื่อให้โปรแกรมเมอร์เข้าใจได้ง่าย
- สามารถประเมินและแก้ไขระบบผลลัพธ์ได้ตามเป้าหมายของโครงการ
- ช่วยวางแผนและกำหนดเป้าหมายของโครงการโดยละเอียด

คำถามและแบบฝึกหัดท้ายบทที่ 8

1. จงอธิบายประโยชน์ของการเรียนรู้ด้านการทำเหมืองข้อมูล (Data Mining)

คำตอบ ผู้ศึกษาสามารถหาแนวทางการตัดสินใจและคาดการณ์ผลลัพธ์ที่ได้แม่นยำตามวิธีการทางสถิติเดิมๆ การนำเหมืองข้อมูลไปใช้งานในธุรกิจ เช่น การทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่ การทำนายยอดขายเมื่อมีการลดราคาสินค้า การทำนายกลุ่มลูกค้าที่น่าจะใช้สินค้าของเรา

2. จงอธิบายเทคนิคการเรียนรู้แบบมีผู้สอน (Supervised Machine Learning)

คำตอบ เป็นผลลัพธ์ที่สร้างจากการอนุมาน (Inference) ชุดข้อมูลปัจจุบัน เพื่อใช้ในการทำนายประเภทตัวอย่างในอนาคต การสร้างโมเดลแบบการเรียนรู้แบบมีผู้สอน (Supervised Machine Learning) จะต้องมีข้อมูลเป้าหมาย (Target) เพื่อทำนายข้อมูลที่เข้ามาใหม่

3. จงอธิบายแบบจำลองแบบเทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Machine Learning)

คำตอบ เป็นการหาความสัมพันธ์ต่างๆ (Association) รวมถึงการจัดกลุ่มข้อมูล (Clustering) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย เช่น ต้องการจัดกลุ่มลูกค้าตามกลุ่มพฤติกรรมที่เหมือนกันอัตโนมัติ ซึ่งเทคนิคที่ใช้ในแบบจำลองนี้คือ “K-means Clustering” ซึ่งคือการเรียนรู้แบบไม่ต้องสอน เป็นเทคนิคการแบ่งกลุ่ม แบบการจัดกลุ่ม (Clustering) โดยใช้พื้นฐานทางสถิติ มีตัวเลขประกอบ อย่างน้อย 2 ตัวแปรขึ้นไป เป็นต้น

4. จงอธิบาย กระบวนการทำเหมืองข้อมูล “CRISP-DM” Cross-Industry Standard Process for Data Mining

คำตอบ

ขั้นตอน	รายละเอียด
1. การทำความเข้าใจประเด็นทางธุรกิจ (Business Understanding)	● การรวบรวมข้อมูลพื้นฐานของสภาพธุรกิจในปัจจุบัน ● การจัดหาจุดประสงค์ทางธุรกิจที่เฉพาะเจาะจง ● การหาเกณฑ์ที่ใช้ในการวัดความสำเร็จของการทำเหมืองข้อมูล จากมุมมองเชิงธุรกิจ
2. การทำความเข้าใจข้อมูล (Data Understanding)	● การจัดเก็บข้อมูลเบื้องต้นทั้งที่มาจากข้อมูลเดิม ชื่อข้อมูล สำนวน หรือเริ่มติดตามข้อมูลเพิ่มเติม ● การอธิบายข้อมูลทั้งเชิงปริมาณและคุณภาพ ● การรายงานและสำรวจข้อมูลที่ประกอบด้วยตาราง กราฟ และเครื่องมือ Visualization ต่างๆ ● การตรวจสอบความถูกต้องของข้อมูลว่ามีข้อมูลที่ผิดพลาดรวมอยู่ด้วยหรือไม่
3. การเตรียมข้อมูล (Data Preparation)	● การรวบรวมข้อมูล และเลือกข้อมูลตัวอย่างที่จะนำไปใช้ ● การสร้างตัวแปรใหม่ที่ถูกลำดับค่ามาจากตัวแปรเดิมผ่านการใช้อัลกอริทึม ● จัดเตรียมข้อมูลสำหรับการทำแบบจำลอง ● การแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของแบบจำลอง
4. การทำโมเดล (Modeling)	● การเรียนรู้แบบมีผู้สอน (Supervised Learning) (1) Classification เป็นการจัดกลุ่มข้อมูลด้วยตัวแปรที่กำหนดไว้แล้ว เพื่อสร้างแบบจำลองคาดการณ์ข้อมูลในอนาคต ได้แก่ เทคนิค Decision Tree Naïve Bayes K-Nearest Neighbors (K-NN) Neural Network เป็นต้น (2) Regression เป็นการประมาณค่าข้อมูล คล้ายกับ Classification แต่คำตอบจะเป็นตัวเลขเท่านั้น ● การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) (1) เทคนิคการจัดกลุ่ม (Clustering) เป็นการจัดข้อมูลที่มีลักษณะคล้ายกันในกลุ่มเดียวกัน (2) การหาความสัมพันธ์ (Associatio) เป็นการค้นหาความสัมพันธ์ของข้อมูลในเหตุการณ์เดียวกัน
5. การประเมิน (Evaluation)	● การประเมินผลการวิเคราะห์กับเกณฑ์ที่ตั้งไว้ ● ผลที่ได้ ได้แก่ แบบจำลองที่คัดเลือกและข้อสรุปหรือสิ่งที่ได้จากแบบจำลอง

6. การนำไปใช้ (Deployment)	● การใช้องค์ความรู้ใหม่ที่ค้นพบมาพัฒนาองค์กร หรือการแสดงถึงการเปลี่ยนแปลงที่เกิดขึ้นในองค์กรได้ ● การวางแผนและควบคุมการนำผลไปใช้งาน รวมถึงเขียนสรุปและทบทวนโครงการ
----------------------------	---

5. จงอธิบายนักวิเคราะห์ข้อมูล (Data Analyst)

คำตอบ บุคคลที่ใช้ข้อมูลในการวิเคราะห์โอกาสทางธุรกิจที่สอดคล้องกับแนวโน้มในการสร้างผลประกอบการตามค่านิยมของตลาดในปัจจุบัน หรือแก้ปัญหาจากสิ่งผิดแปลกไปจากแนวโน้มเดิม โดยใช้หลัก “สถิติ” ทางคณิตศาสตร์หรือหลักเศรษฐศาสตร์มาประกอบแนวคิดในการวิเคราะห์ข้อมูลเชิงการสร้างความมูลค่าทางธุรกิจในด้านต่างๆ ตามวัตถุประสงค์ที่ตั้งไว้ ทั้งนี้การวิเคราะห์ต่างๆ ต้องอาศัยประสบการณ์ องค์ความรู้และมุมมองที่เฉียบขาด

6. จงอธิบายนักวิทยาศาสตร์ข้อมูล (Data Scientist)

คำตอบ บุคคลที่นำข้อมูลจากหลายๆ แหล่งองค์ความรู้ นำมาผ่านวิธีการต่างๆ เช่น การทำเหมืองข้อมูล (Data Mining) การเรียนรู้ของเครื่อง (Machine Learning) การเพิ่มประสิทธิภาพ (Optimization) เพื่อวิเคราะห์หาโอกาสที่น่าจะเป็นภายใต้แนวคิดหรือมุมมองที่ทันสมัยใหม่ ซึ่งมุ่งเน้นผลลัพธ์ของคำตอบใหม่ๆ ที่ใช้ในการนำไปวิเคราะห์และสร้างสูตรทางการตลาดและการเติบโตในธุรกิจที่มุ่งหวัง

7. จงอธิบายงานของนักวิทยาศาสตร์ข้อมูล (Data Scientist) ส่วนใหญ่

คำตอบ การสร้างแบบจำลอง คิดสมการใหม่ๆ เพื่อส่งต่อให้โครงการด้านข้อมูล (Data Project) ซึ่งเป็นงานประเภทโครงการ (Project Base) จากนั้นนักวิเคราะห์ข้อมูล (Data Analyst) จะเข้ามารับช่วงต่อ โดยเป็นผู้ที่ใช้ข้อมูล และวิธีการต่างๆ เพื่อวิเคราะห์สิ่งที่เกิดขึ้นด้วยตัวเลข เพื่อดูผลกระทบที่มีต่อธุรกิจต่อไป

8. จงอธิบายความแตกต่างระหว่าง นักวิเคราะห์ข้อมูล (Data Analyst) และ นักวิทยาศาสตร์ข้อมูล (Data Scientist)

คำตอบ

ความแตกต่าง	นักวิเคราะห์ข้อมูล (Data Analyst)	นักวิทยาศาสตร์ข้อมูล (Data Scientist)
-------------	-----------------------------------	---------------------------------------

ความหมาย	ผู้ที่ทำหน้าที่รวบรวมข้อมูลจากหลายแหล่งข้อมูล เพื่อนำมาวิเคราะห์ เป็น report และ visualization	ผู้ที่ไม่ได้เพียงแค่ทำหน้าที่รวบรวม และแปลผลข้อมูลเท่านั้น แต่ยังทำหน้าที่ดึงข้อมูลเชิงลึกที่สามารถนำมาประยุกต์ใช้ได้ หรือการใช้ข้อมูลเพื่อ business transformation
ข้อมูลที่ใช้	เน้นที่การเปลี่ยนแปลง และ การแปลผลของข้อมูล ซึ่งมักจะมุ่งไปที่ข้อมูลในปัจจุบันและอดีต	เน้นที่การสรุปข้อมูล และ คาดการณ์ข้อมูลในอนาคตบนพื้นฐานของข้อมูลในอดีตและปัจจุบัน
วิธีการเพื่อตอบคำถามทางธุรกิจ	เขียน Custom Query เช่น SQL เพื่อตอบคำถามที่ซับซ้อนในเชิงธุรกิจ ยังไม่ถึงขั้นใช้ในการตัดสินใจทางธุรกิจได้	สร้าง Statistical Model ที่ใช้ในการตัดสินใจบนพื้นฐานของข้อมูล ซึ่งเป็นไปได้ทั้งการตัดสินใจขั้นสูง เช่น การ Block Page และ ขั้นพื้นฐาน เช่น การให้คะแนนความเสี่ยงกับ Webpage
รูปแบบการวิจัย	เข้าใจและต่อยอดจากส่วนที่เคยเข้าใจไม่เพียงพอเกี่ยวกับธุรกิจและผลิตภัณฑ์	สร้างการทดสอบสมมติฐานที่เป็นเหตุเป็นผลกัน (Causality Experiments) ที่จะพยายามหาต้นเหตุของกรณีศึกษา
ผลลัพธ์	ระบุได้เพียงประเด็นของคุณภาพข้อมูล เช่น Data Gap หรือ Bias ในการรวบรวมข้อมูล จะทำงานร่วมกับ Engineering ส่วนที่เหลือ เพื่อมองหากการรวบรวมข้อมูลแบบใหม่ๆ	ระบุถึงผลิตภัณฑ์ หรือ ลักษณะของผลิตภัณฑ์ใหม่ๆ ที่มาจากการถอดมูลค่าของข้อมูล เป็นผู้นำทางความคิดเกี่ยวกับมูลค่าของข้อมูล เช่น ระบบ Recommender System ของ Amazon ที่เคยทำออกมาครั้งแรกให้กับผู้ใช้งานทั่วไป
ทักษะด้านเทคนิค	Excel Tableau SQL SAS SPSS	Python R Java Julia Matlab SQL Hive Hadoop Tableau Spark

ทักษะด้านอื่น	ความรู้ทางสถิติ ความรู้ทางธุรกิจ และ ทักษะการสื่อสาร	ความอยากรู้ทางปัญญา ความรู้ทางธุรกิจ และ ทักษะการสื่อสาร
---------------	--	--

9. จงอธิบายการเตรียมข้อมูลที่มีประสิทธิภาพ ควรมีลักษณะสำคัญดังนี้

คำตอบ

- ให้ผลลัพธ์ที่ครบถ้วนสมบูรณ์
- ให้ความสำคัญกับนิยามข้อมูล
- มีการจัดบันทึกขั้นตอนการเตรียมข้อมูลโดยละเอียด
- ปรับกระบวนการให้เป็นอัตโนมัติให้มากที่สุด

10. จงอธิบายการทำความสะอาดข้อมูล (Data cleaning)

คำตอบ เป็นขั้นตอนในการแยกข้อมูลที่รบกวนออกไป เช่น การเติมข้อมูลที่ขาดหายไป (Missing Values) การปรับเปลี่ยนข้อมูลที่ผิดพลาด (Noisy Data) การแยกข้อมูลที่มีค่าผิดปกติออกมา (Outliers) และการแก้ไขข้อมูลที่ไม่สอดคล้องกัน (Inconsistent Data)

11. จงอธิบายการรวมข้อมูล (Data Integration) คือ

คำตอบ เป็นขั้นตอนที่รวมแหล่งข้อมูลหลายๆแห่งเข้าด้วยกัน เพื่อไว้ในที่เดียวกัน เพื่อลดความซ้ำซ้อนของข้อมูล และเพิ่มความเร็วและคุณภาพในการทำเหมืองข้อมูล

12. จงอธิบายการแปลงข้อมูล (Data Transformation) คือ

คำตอบ เป็นขั้นตอนการแปลงรูปข้อมูลเพื่อให้เหมาะสมกับการทำเหมืองข้อมูล โดย Normalization (การแปลงค่าข้อมูลให้อยู่ในช่วงสั้นๆ ที่นำมาใช้ประมวลผลได้) เป็นวิธีที่พบบ่อย เช่น Min-Max Decimal Scaling หรือ Z-Score เป็นต้น

13. จงอธิบายการลดรูปข้อมูล (Data Reduction) คือ

คำตอบ การลดรูปข้อมูลให้มีรูปแบบกระชับและกินเนื้อที่น้อยลง เพื่อที่จะนำไปวิเคราะห์แล้วเกิดผลลัพธ์ที่เท่ากันหรือเทียบเท่ากันเมื่อเทียบกับการวิเคราะห์จากข้อมูลทั้งหมด

14. จงอธิบายวิธีพารามเทริก (Parametric Methods) คือ

คำตอบ การสร้างแบบจำลองข้อมูลด้วยการประมาณค่าพารามิเตอร์ของแบบจำลอง จัดเก็บค่าพารามิเตอร์ เช่น การสร้างแบบจำลองข้อมูลด้วยเทคนิคทางสถิติความถดถอยในการประมาณค่าตัวแปรตาม

15. จงอธิบายวิธีการที่ไม่ใช่พารามิเตอร์ (Non-parametric Methods) คือ

คำตอบ วิธีที่ไม่ได้สร้างแบบจำลองข้อมูล แต่ใช้เครื่องมือ เช่น แผนภาพฮิสโตแกรม (Histogram Diagram) วิธีการจัดกลุ่ม (Clustering) แสดงการกระจายข้อมูล และเก็บค่าตัวแทนกลุ่มแทนค่าข้อมูลจริง, และ การสุ่มตัวอย่าง (Sampling) เป็นต้น

16. จงอธิบายการทำให้ข้อมูลเป็นแบบข้อมูลแบบไม่ต่อเนื่อง (Discretization Data) คือ

คำตอบ การลดรูปข้อมูลดิบโดยแบ่งข้อมูลดิบออกเป็นช่วงๆ และใช้คาชวง (Interval Label) แทนค่าของแต่ละช่วง ในกรณีที่อัลกอริทึมรับเฉพาะข้อมูลแบบ การจัดลำดับชั้นแนวคิด (Categorical Concept Hierarchy Generation) แบ่งข้อมูลดิบเป็นแต่ละช่วงจากแนวคิดระดับล่างไประดับที่สูงขึ้น เช่น อายุ แบ่งเป็นกลุ่ม วัยเยาว์ วัยกลางคน วัยชรา

คำถามและแบบฝึกหัดท้ายบทที่ 9

1. จงอธิบายเทคนิคการจำแนกประเภท (Classification)

คำตอบ การจำแนกประเภท (Classification) เป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ ประเภทของ Data Mining ที่ทราบค่าประเภทของตัวอย่างในชุดข้อมูลสอนล่วงหน้า เป็นการสร้างโมเดลหรือตัวจำแนกข้อมูล (Classifier) เพื่อทำนายหมวดหมู่ของข้อมูล (Categories/ Class) โดยมีตัวอย่างในชุดข้อมูลสอน (Training Set) ที่ใช้ จะมีคุณลักษณะหนึ่ง ซึ่งบอกค่าประเภทของตัวอย่างนั้น (Class Label) ซึ่งเป็นค่าข้อมูลแบบ Category ทั้งนี้ การประยุกต์ใช้งานแบบ Classification มักพบใน การขออนุมัติคำขอมิบัติเครดิต (Credit Approval), การทำตลาดลูกค้ากลุ่มเป้าหมาย (Target Marketing), การตรวจวินิจฉัย (Medical Diagnosis) และ การวิเคราะห์ประสิทธิภาพการรักษา (Treatment Effectiveness Analysis)

2. จงอธิบายประโยชน์ของเทคนิคการจำแนกประเภท (Classification)

คำตอบ การนำเทคนิคการจำแนกประเภท (Classification) ไปใช้งานในธุรกิจ เช่น การทำนายผลการตอบสนองกับการเปิดตัวสินค้าใหม่ การทำนายยอดขายเมื่อมีการลดราคาสินค้า การทำนายกลุ่มลูกค้าที่น่าจะใช้สินค้าของเรา กล่าวโดยสรุปหากธุรกิจ สามารถรู้ล่วงหน้า และวิ่งนำหน้า

ผู้บริโภค ด้วยการใช้ข้อมูลการสั่งซื้อ ประวัติลูกค้าทั้งเก่าและใหม่ เพื่อนำมาวิเคราะห์หาความน่าจะเป็น คาดคะเนจากเทคนิคเหมืองข้อมูล และแปลผลรูปแบบข้อมูลที่ได้จากการวิเคราะห์ เพื่อให้จับต้องได้ จะทำให้สามารถนำมาพัฒนากลยุทธ์ขององค์กร ซึ่งผู้ที่เข้าใจการทำ Data Mining (เหมืองข้อมูล) เช่น จะลูกค้ามากขึ้นทำให้คุณทราบว่าตอนนี้ลูกค้าต้องการอะไร แล้วลูกค้าประเภทไหนที่ซื้อสินค้าหรือใช้บริการ รวมถึงสามารถทำนายพฤติกรรมของผู้บริโภค และออกแบบสินค้าและบริการ ให้ตอบโจทย์ กลุ่มลูกค้าประจำและกลุ่มลูกค้าใหม่ได้ และเห็นมุมมองการทำการตลาดที่กว้างขึ้น

1. จงอธิบายเทคนิคการแบ่งกลุ่ม (Clustering)

คำตอบ การแบ่งกลุ่ม (Clustering) เป็นเทคนิคที่ใช้จำแนกหรือแบ่งประเภท (คน สัตว์ สิ่งของ หรือ องค์กร ฯลฯ) หรือ แบ่งตัวแปรออกเป็นกลุ่มย่อย ๆ ตั้งแต่ 2 กลุ่มขึ้นไป โดยประเภทที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่เหมือนกันหรือคล้ายกัน ส่วนประเภทที่อยู่ต่างกลุ่มกัน จะมีลักษณะที่แตกต่างกัน ดังนั้น การพิจารณาเลือกลักษณะหรือตัวแปรที่นำมาใช้ในการแบ่งกลุ่มประเภท จึงมีความสำคัญ นอกจากนั้น ประเภทใดประเภทหนึ่งจะต้องอยู่ในกลุ่มหนึ่งเพียงกลุ่มเดียว ถ้านำเทคนิคการแบ่งกลุ่ม (Clustering Analysis) มาใช้ในการแบ่งกลุ่มตัวแปร จะให้ตัวแปรอยู่ในกลุ่มเดียวกันมีความสัมพันธ์กันมากกว่าตัวแปรที่อยู่ต่างกลุ่มกัน ตัวแปรที่อยู่ต่างกลุ่มกันมีความสัมพันธ์กันน้อยหรือไม่มีความสัมพันธ์กันเลย ส่วนใหญ่การแบ่งกลุ่มตัวแปรจะใช้เทคนิค Factor ส่วนการแบ่งกลุ่มประเภท (คน สัตว์ สิ่งของ) จะใช้เทคนิค เทคนิคการแบ่งกลุ่ม (Clustering)

2. จงอธิบายเทคนิคกฎความสัมพันธ์ (Association Rules)

คำตอบ กฎความสัมพันธ์ (Association Rules) เป็นกระบวนการหนึ่งในการทำเหมืองข้อมูล (Data Mining) ที่ได้รับความนิยมมาก โดยจะใช้กฎความสัมพันธ์ (Association Rules) ในการหาความสัมพันธ์ของข้อมูลสองชุดหรือมากกว่าสองชุดขึ้นไปภายในกลุ่มข้อมูลที่มีขนาดใหญ่ ในการหาความสัมพันธ์นั้นจะมีขั้นตอนวิธีการหาหลายวิธีด้วยกัน แต่ขั้นตอนวิธีที่เป็นที่รู้จักและใช้อย่างแพร่หลายคือ ขั้นตอนวิธี “Apriori” ซึ่งเป็นเทคนิคกฎความสัมพันธ์ (Association Rule) เช่น การเชื่อมโยงสำหรับการค้นหาความสัมพันธ์ระหว่างผลิตภัณฑ์ในข้อมูลธุรกรรมขนาดใหญ่ที่บันทึกโดยระบบ ณ จุดขาย (POS) ในซูเปอร์มาร์เก็ตสามารถเขียนได้ในรูปเซตของสิ่งของ (Item) ที่เป็นเหตุ ไปสู่เซตของสิ่งของ (Item) ที่เป็นผล

3. จงอธิบายเทคนิคระบบแนะนำ (Recommendation System)

คำตอบ คือระบบที่ช่วยแนะนำสินค้าและบริการให้กับผู้บริโภคออนไลน์ บริษัทที่ขึ้นชื่อด้าน e-Commerce ใช้ระบบแนะนำในการสร้างประสบการณ์ในการใช้แพลตฟอร์ม (Platform) ที่ดีขึ้น

เช่น Shopee LAZADA Amazon Netflix YouTube Google เป็นต้น โดยแพลตฟอร์มเหล่านี้พัฒนาโมเดลจากกฎ (Rules) ในการแนะนำสินค้า ซึ่งแบ่งเป็น 2 แบบคือ

แบบที่ 1: User-Based Filtering เป็นเทคนิคที่ใช้ในการทำนายรายการที่ประมวลจากการหาสินค้าที่พฤติกรรมเหมือนกัน โดยอาจจะพิจารณาจากคะแนนที่เรต (Rating) สินค้ารายการนั้น และประมวลผลจากลูกค้ารายอื่นที่มีรสนิยมคล้ายกันกับของลูกค้าเป้าหมาย

แบบที่ 2: Item-Based Filtering เป็นเทคนิคที่คำนวณจากการให้คะแนนที่ผู้ใช้กลุ่มเดียวกันให้ไว้กับรายการหรือสินค้านั้น

4. จงอธิบายปัญหา “Sparse” และ “Scalability” ในแมชชีนเลิร์นนิง (Machine Learning)

คำตอบ คือการเกิดข้อมูลที่กระจัดกระจาย ซึ่งจะทำให้ประสิทธิภาพของอัลกอริทึม (Algorithm) การเรียนรู้ของเครื่องและความสามารถในการคำนวณการคาดการณ์ที่แม่นยำลดลง จากข้อมูลที่เบาบาง หรือจากค่าที่คาดหวังในชุดข้อมูลหายไป และกระจัดกระจาย ส่งผลให้การประมวลผลไม่แม่นยำ

5. จงอธิบาย Implicit Rating

คำตอบ คือการเก็บข้อมูลโดยการสังเกตการกระทำของผู้ใช้ในรายการบนแพลตฟอร์ม (Platform) เช่น เก็บข้อมูลการคลิก (Click) เข้าไปดูสินค้านั้น หรือการตัดสินใจที่จะซื้อสินค้านั้น หรือการที่ผู้ใช้กด Favorite หรือการเฝ้าสังเกต พฤติกรรมจากประวัติการซื้อ การค้นหาสถานที่ และเวลาที่ใช้งานเว็บในแต่ละแพลตฟอร์ม (Platform) เป็นต้น

6. จงอธิบาย 3 กระบวนการของระบบแนะนำ (Recommendation System)

คำตอบ การทำงานของระบบแนะนำ (Recommendation System) ประกอบไปด้วย 3 ขั้นตอนใหญ่ๆ คือ

(1) Information Collection Phase: เป็นขั้นตอนการเก็บข้อมูล ที่เกี่ยวข้องจากผู้ใช้งานเพื่อสร้างโปรไฟล์ผู้ใช้งาน เช่น รายละเอียดลูกค้า พฤติกรรมการบริโภค และทรัพยากรที่ผู้ใช้งานใช้บ่อยครั้ง ระบบต้องการข้อมูลจากผู้ใช้งานมากที่สุดเท่าที่จะเป็นไปได้เพื่อสร้างระบบให้คำแนะนำได้ดีและใกล้เคียงมากที่สุดตั้งแต่ครั้งแรก โดย Recommendation Systems สามารถรับเข้าข้อมูลจากระบบการให้คะแนนและระบบการสังเกตพฤติกรรม 3 ประเภทดังที่กล่าวมาแล้ว

(2) Learning Phase: เป็นขั้นตอนการเรียนรู้ โดยสร้าง Learning Algorithm เพื่อคัดกรองข้อมูล และนำรายละเอียดที่ได้จากลูกค้า เข้ามาสู่กระบวนการบันทึกข้อมูลต่อไป

(3) Prediction / Recommendation Phase: เป็นขั้นตอนการคาดเดาข้อมูลที่เหมาะสมต่อผู้ใช้งานโดยขั้นตอนนี้สามารถดึงข้อมูลมาใช้งานโดยตรงได้จาก Memory Based หรือจาก Model Based

7. จงอธิบายเทคนิค Contented-Based Filtering (CBF)

คำตอบ เป็นเทคนิคการสร้างโปรไฟล์ของผู้ใช้แต่ละคนแยกกันไป แล้วแนะนำสินค้าตามโปรไฟล์ของผู้ใช้นั้นเป็นกันไป โดยการแนะนำแบบพิจารณาเนื้อหา ที่อ้างอิงจากโดเมนเป็นหลัก (เน้นไปที่การวิเคราะห์รายละเอียดของสินค้าเพื่อการคาดเดา) ส่วนมากนิยมใช้ในด้านของการนำเสนอเว็บไซต์ สื่อสิ่งพิมพ์ และข่าว ดังรูป 9.13 โดยมีค่าความถูกต้องและแม่นยำของการแนะนำสินค้าขึ้นอยู่กับ การประเมินความพอใจของลูกค้า ต่อผลิตภัณฑ์นั้นๆ จากข้อมูลในอดีต (สินค้าใดที่เกี่ยวข้องกับสินค้าที่เคยถูกประเมินในเชิงบวก จะถูกนำเสนอให้ลูกค้าทราบ) โดย CBF จะใช้โมเดลประเภท Vector Space Model เช่น Term Frequency, Inverse Document Frequency หรือโมเดลเชิงสถิติเช่น Naïve Bayes Classifier, Decision Trees, Neural Networks เพื่อหาความสัมพันธ์ระหว่างสินค้า 2 ตัว (กล่าวคือ เป็นการใช้ศาสตร์ทางด้าน Statistical Analysis และ Machine Learning มาช่วยในการสร้างโมเดล)

8. จงอธิบาย 3 กระบวนการของระบบแนะนำ (Recommendation System)

คำตอบ การทำงานของระบบแนะนำ (Recommendation System) ประกอบไปด้วย 3 ขั้นตอนใหญ่ๆ คือ

(1) Information Collection Phase: เป็นขั้นตอนการเก็บข้อมูล ที่เกี่ยวข้องจากผู้ใช้งานเพื่อสร้างโปรไฟล์ผู้ใช้งาน เช่น รายละเอียดลูกค้า พฤติกรรมการบริโภค และทรัพยากรที่ผู้ใช้งานมักใช้งานบ่อยครั้ง ระบบต้องการข้อมูลจากผู้ใช้งานมากที่สุดเท่าที่จะเป็นไปได้เพื่อสร้างระบบให้คำแนะนำได้ดี และใกล้เคียงมากที่สุดตั้งแต่ครั้งแรก โดย Recommendation Systems สามารถรับเข้าข้อมูลจากระบบการให้คะแนนและระบบการสังเกตพฤติกรรม 3 ประเภทดังที่กล่าวมาแล้ว

(2) Learning Phase: เป็นขั้นตอนการเรียนรู้ โดยสร้าง Learning Algorithm เพื่อคัดกรองข้อมูล และนำรายละเอียดที่ได้จากลูกค้า เข้ามาสู่กระบวนการบันทึกข้อมูลต่อไป

(3) Prediction / Recommendation Phase: เป็นขั้นตอนการคาดเดาข้อมูลที่เหมาะสมต่อผู้ใช้งานโดยขั้นตอนนี้สามารถดึงข้อมูลมาใช้งานโดยตรงได้จาก Memory Based หรือจาก Model Based

9. จงอธิบายเทคนิค Contented-Based Filtering (CBF)

คำตอบ เป็นเทคนิคการสร้างโปรไฟล์ของผู้ใช้แต่ละคนแยกกันไป แล้วแนะนำสินค้าตามโปรไฟล์ของผู้ใช้นั้นเป็นกันไป โดยการแนะนำแบบพิจารณาเนื้อหา ที่อ้างอิงจากโดเมนเป็นหลัก

(เน้นไปที่การวิเคราะห์รายละเอียดของสินค้าเพื่อการคาดเดา) ส่วนมากนิยมใช้ในด้านของการนำเสนอเว็บไซต์ สื่อสิ่งพิมพ์ และข่าว ดังรูป 9.13 โดยมีค่าความถูกต้องและแม่นยำของการแนะนำสินค้า ขึ้นอยู่กับการประเมินความพอใจของลูกค้า ต่อผลิตภัณฑ์นั้นๆ จากข้อมูลในอดีต (สินค้าใดที่เกี่ยวข้องกับสินค้าที่เคยถูกประเมินในเชิงบวก จะถูกนำเสนอให้ลูกค้าทราบ) โดย CBF จะใช้โมเดลประเภท Vector Space Model เช่น Term Frequency, Inverse Document Frequency หรือโมเดลเชิงสถิติเช่น Naïve Bayes Classifier, Decision Trees, Neural Networks เพื่อหาความสัมพันธ์ระหว่างสินค้า 2 ตัว (กล่าวคือ เป็นการใช้ศาสตร์ทางด้าน Statistical Analysis และ Machine Learning มาช่วยในการสร้างโมเดล)

10. จงอธิบายเทคนิค Collaborative Filtering (CF)

คำตอบ คือการแนะนำแบบพึ่งพาผู้ใช้ร่วม โดยเก็บข้อมูลจากผู้ใช้คนอื่นๆมาช่วยด้วย วิธีการนี้เป็นการคัดกรองโดยไม่พึ่งโดเมน ใช้สำหรับคอนเทนต์ (Content) ที่ไม่สามารถอธิบายได้ด้วย Metadata เพียงอย่างเดียว เช่น ภาพยนต์ และเพลง โดย CF จะสร้างตารางข้อมูล (User-Item Matrix) ที่เกี่ยวพัน กับประเภทสินค้า และผู้ใช้งาน ซึ่งหลักการสำคัญของวิธีการนี้ คือการจับกลุ่มโปรไฟล์ของลูกค้าที่มีความสนใจคล้ายๆ กัน (เรียกว่า Neighborhood) ทำให้ลูกค้าสามารถได้รับการแนะนำ (Recommend) สินค้าตัวที่เขาไม่เคยให้คะแนนมาก่อน แต่รับข้อมูลความสนใจมาจากลูกค้าคนอื่นที่อยู่ใน Neighborhood

11. จงอธิบายเทคนิค โครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

คำตอบ คือโมเดลแบบจำลองทางคณิตศาสตร์ที่มีลักษณะเป็นโครงข่ายประสาทเทียมซึ่งมีแนวคิดมาจากการทำงานสมองของมนุษย์ที่ใช้ความสามารถในการเรียนรู้จากประสบการณ์ ซึ่งประกอบด้วยเซลล์ประสาท (Neurons) และจุดประสานประสาท (Synapses) จำนวนมากทำงานเชื่อมโยงกัน ประกอบไปด้วย

(1) Input Layer เป็นส่วนของข้อมูล Input จำนวนของ Node ขึ้นอยู่กับจำนวนของ Attribute ที่นำเข้ามาคิดในโมเดล และจะมีการกำหนดค่าน้ำหนัก (Weight) ของแต่ละ Attribute โดยเมื่อเริ่มต้นจะเป็นการ Random ขึ้นมา จากนั้นเมื่อโมเดลมีการเรียนรู้ก็จะมีการปรับค่า Weight เพื่อให้ได้คำตอบที่ใกล้เคียงที่สุด

(2) Hidden Layer เป็นชั้นที่อยู่ระหว่างกลาง ซึ่งจะทำการคำนวณข้อมูลที่ได้รับมาจาก Layer ก่อนหน้า โดยวิธีการทางคณิตศาสตร์และมี Nonlinearity Function เช่น Sigmoid, Hyperbolic Tangens, Hard Limit, Rectified Linear Unit เป็นต้น เพื่อสร้างเป็น Information ส่งไปให้ Layer

ถัดไป โดยผู้สร้างโมเดลสามารถกำหนดจำนวนชั้นของ Hidden Layer ได้ ซึ่งการเพิ่มชั้นและจำนวน Neuron ก็ส่งผลต่อการทำงานและประสิทธิภาพในการเรียนรู้ของโมเดล ในกรณีที่มีจำนวน Hidden Layer มากกว่า 1 ชั้น จะเรียกโมเดลนั้นเป็น Deep Learning

(3) Output Layer เป็นส่วนของข้อมูลที่ได้จากการคำนวณ จำนวนของ Node ขึ้นอยู่กับรูปแบบของ Output ที่ต้องการ

คำถามและแบบฝึกหัดท้ายบทที่ 10

1. จงอธิบายมุมมองที่ต้องทำความเข้าใจ 5 มิติสำหรับการทำเหมืองข้อมูล (Data Mining)

คำตอบ

มิติที่ 1 ทำความเข้าใจวัตถุประสงค์เชิงธุรกิจ (Business Objectives) ธุรกิจต้องกำหนดเป้าหมายผลตอบแทนจาก ด้านการลดต้นทุนขององค์กรและต้นทุนด้านเทคโนโลยีดิจิทัล จากเป้าหมายด้านการปรับปรุงกระบวนการระหว่างฝ่ายและหน่วยธุรกิจ (Cross Functional Process Improvement) ซึ่งจะต้องประเมินด้วยผลลัพธ์ในด้านผลกำไร (Bottom Line) ความพึงพอใจของลูกค้า (Customer Satisfaction) หรือการเพิ่มประสิทธิภาพในการใช้ทรัพยากรขององค์กร (Efficiency in Resource Utilization) รวมถึงการเพิ่มประสิทธิภาพในการปฏิบัติงาน (Operational Efficiency) หรือการเพิ่มความสามารถในการแข่งขัน

มิติที่ 2 ทำความเข้าใจลำดับความสำคัญของการลงทุน (Funding Priorities) เป็นการรวบรวมข้อมูลการลงทุนในระบบระดับองค์กร (Enterprise System) ที่มีการบูรณาการด้านเทคโนโลยี และการร่วมใช้งานทั้งด้านข้อมูล (Shared Data) และด้านโครงสร้างพื้นฐาน (Shared Infrastructure)

มิติที่ 3 ทำความเข้าใจขีดความสามารถในการบริหารจัดการ (Management Capabilities) คือการรวบรวมข้อมูลขีดความสามารถในการบริหารจัดการกระบวนการระหว่างฝ่ายและหน่วยธุรกิจ (Cross Functional Process Improvement) ในด้านการบริหารจัดการมาตรฐานของงาน เทคโนโลยีดิจิทัล (Technology Standardization) การบริหารจัดการมาตรฐานข้อมูล (Data Standardization) และการบริหารจัดการมาตรฐานกระบวนการ (Process Standardization)

มิติที่ 4 ทำความเข้าใจผู้รับผิดชอบในการกำหนดการใช้งานของระบบ (Who Defines Applications)

มิติที่ 5 ทำความเข้าใจการกำกับดูแลด้านเทคโนโลยีดิจิทัล (Digital Governance) ทำความเข้าใจการกำกับดูแลด้านเทคโนโลยีดิจิทัล (Digital Governance) ประกอบไปด้วยการทำความเข้าใจความรับผิดชอบ (Responsibility) กลยุทธ์ (Strategy) การจัดซื้อจัดหา (Acquisition) ผลการดำเนินงาน (Performance) ความสอดคล้องกัน (Conformance) และพฤติกรรมบุคคล (Human Behavior) โดยมุ่งเน้นการกำหนดกรอบทิศทางในการกำกับดูแลด้านการบริหารจัดการเทคโนโลยีดิจิทัลขององค์กร ให้สามารถบริหารจัดการทรัพยากรเทคโนโลยีดิจิทัลได้อย่างเหมาะสม

2. จงอธิบายโครงสร้างข้อมูล 3 ประเภท

คำตอบ (1) ข้อมูลแบบโครงสร้าง (Traditional DB) เช่น ข้อมูลจากระบบอีอาร์พี (Enterprise Resource Planning: ERP) ระบบลูกค้าสัมพันธ์ (Customer Relationship Management: CRM) ระบบบัญชี (Accounting) และระบบอื่น ๆ เช่น ระบบที่ใช้ในโรงพยาบาล เช่น เวชระเบียนอิเล็กทรอนิกส์ (EHRs)

(2) ข้อมูลกึ่งโครงสร้าง เช่น ข้อมูลจาวารคอมพิวเตอร์ (Log) จากอุปกรณ์ตรวจสอบ ข้อมูลที่นามสกุลเป็น XML, JSON เป็นต้น

(3) และข้อมูลที่ไม่มีโครงสร้าง เช่น ข้อมูลภาพทาง

3. ปัญหาที่พบส่วนใหญ่จะเกิดจากการนำเข้าข้อมูลภายนอกระบบคือ

คำตอบ (1) มีข้อมูลบางฟิลด์ที่ล้าสมัยหรือซ้ำซ้อน (Fields that are obsolete or redundant)

(2) ข้อมูลไม่สมบูรณ์และขาดหาย (Missing values)

(3) ข้อมูลมีค่าผิดปกติ (Outliers) หรือ ข้อมูลมีค่าผิดพลาด (Error and Noisy data)

(4) ข้อมูลในรูปแบบที่ไม่เหมาะกับแบบจำลอง (Model) การทำเหมืองข้อมูล (Data in a form not suitable for the data mining models)

(5) ข้อมูลไม่สอดคล้องกัน (Values not consistent with policy or common sense, Inconsistent data)

4. จงอธิบายการแปลงข้อมูล (Transformation) จะประกอบไป 4 ขั้นตอนหลัก

คำตอบ (1) Data Cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่เป็นส่วนรบกวนหรือข้อมูลที่ไม่เกี่ยวข้องออกไป เป็นกระบวนการตรวจสอบและการแก้ไข (หรือลบ) รายการข้อมูลที่ไม่

ถูกต้องออกไปจากชุดข้อมูล ตารางหรือฐานข้อมูล ซึ่งมีเทคนิคเบื้องต้นได้แก่ การแจกแจงข้อมูล (Parsing) การแก้ไขข้อมูลที่ผิดพลาด (Correcting) การทำข้อมูลให้เป็นรูปแบบเดียวกัน (Standardizing) และ การลบชุดข้อมูลซ้ำซ้อนทิ้ง (Duplicate Elimination)

(2) Data Integration กรณีที่ระบบต้องการใช้ข้อมูลจากหลายแหล่ง หรือมีหลายแฟ้มข้อมูล ต้องทำการเชื่อมโยงข้อมูลจากหลายแหล่งเข้าด้วยกัน โดยใช้คุณลักษณะหรือแอตทริบิวต์ ที่มีอยู่ร่วมกันของหลายแหล่งข้อมูลเป็นตัวเชื่อมโยง โดยเป็นขั้นตอนการรวมแหล่งข้อมูล ซึ่งมีข้อมูลหลายแห่งมารวมไว้ที่เดียวกัน

(3) Data Transformation เป็นขั้นตอนการแปลงข้อมูลในขั้นตอนการคัดเลือก เป็นการเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมสำหรับการประมวลผล และให้เหมาะสมสำหรับขั้นตอนการทำเหมืองข้อมูล

(4) Data Reduction เป็นขั้นตอนการลดมิติข้อมูล เพื่อเป็นตัวแทนจำนวนข้อมูลทั้งหมด เป็นต้น

นอกเหนือจากนั้นในขั้นตอนการเตรียมข้อมูล (Data Preparation) นักวิเคราะห์อาจจะเจอปัญหาอื่นๆ อีกมากมายเช่น

5. จงยกตัวอย่างความท้าทายในการเตรียมข้อมูล (Data Preparation)

คำตอบ

(1) การละเลยบริบท (Neglecting context) องค์กรหลายแห่งอาจไม่มีแนวทางการปฏิบัติในการวิเคราะห์ข้อมูลทึ่มบูรณ์ มักจะละเลยการกำหนด วัตถุประสงค์ของงานให้สอดคล้องกับวิสัยทัศน์ พันธกิจ หรือเป้าประสงค์ขององค์กรในขณะนั้น

(2) ขาดการระบุมาตรฐานคุณภาพของข้อมูล (Spotting data quality issues)

6. จงยกตัวอย่างขั้นตอนการประเมินประสิทธิภาพของโมเดล ในการทำเหมืองข้อมูล (Data Mining)

คำตอบ

(1) Cross-Validation คือ เทคนิคที่ใช้ในการประเมินการทำงานของโมเดลการทำนาย (Predictive Model)

(2) Holdout Method คือ เทคนิคที่ขึ้นกับการแบ่ง (Split) ของข้อมูล เป็นวิธีที่ง่ายที่สุดของ Cross Validation

(3) Random Sub-Sampling คือ การทำซ้ำหลายๆครั้ง เพื่อให้การประเมินการทำงานของตัวแปรดีขึ้น

(4) K-Fold Cross Validation เป็นอีกวิธีในการแก้ปัญหาของ Holdout Method โดยข้อมูลจะถูกแบ่งเป็น k Subset

(5) Leave-One-Out Method จะใช้จำนวน N ครั้งที่แยกจากกันในการทดสอบข้อมูล และ Function Approximator จะถูกใช้การเรียนรู้ (Train) ในทุกๆ ข้อมูล ยกเว้นไว้หนึ่งจุดที่จะถูกใช้ในการทำนายแทน ซึ่งจะมีการวัด Leave-One-Out Cross Validation Error (LOO-XVE) แต่ข้อเสียคือค่าใช้จ่ายในการคำนวณสูง

(6) Bootstrap วิธีนี้จะต่างวิธีอื่นๆ เพราะการสุ่มตัวอย่าง (Sampling) ของข้อมูลที่ใช้เรียนรู้ จะถูกทำโดยมีการแทนที่ กล่าวคือ ข้อมูลที่เคยใช้การเรียนรู้ (Train) ไปแล้วจะถูกดึงกลับมาใน Original Pool ทำให้ทุกข้อมูลมีโอกาสที่จะนำมาใช้ใหม่เท่าๆกัน

(7) Confusion Matrix คือการแบ่งชั้นของผลลัพธ์ (Output Class) ที่ทำนายได้ออกเป็น 2 กลุ่มย่อยของแต่ละกลุ่มหลัก คือถูก (True) กับผิด (False) จึงทำให้เกิดผลในการ การจำแนกประเภท (Classification) ได้ทั้งหมด 4 กลุ่ม คือค่าจริงเป็นบวกและทำนายเป็นบวก (True Positive: TP) ค่าจริงเป็นลบและทำนายเป็นลบ (True Negative: TN) ค่าจริงเป็นลบ แต่ทำนายเป็นบวก (False Positive: FP) ค่าจริงเป็นบวก แต่ทำนายเป็นลบ (False Negative: FN)

(8) Receiver Operating Curves (ROC) วิธีนี้สามารถบอกจุด Cut-Off ที่เหมาะสมที่สุดได้ ณ จุดไหนที่มีความแม่นยำของ Test เกิดขึ้นสูงที่สุด และน่าเชื่อถือที่สุด ทำให้สามารถใช้เลือกจุดนั้นมาเป็นจุด Cut-Off ได้เลยในการแบ่งกลุ่มข้อมูลเพื่อให้มีความถูกต้องมากที่สุด และผิดพลาดน้อยที่สุด

7. จงอธิบาย Power BI

คำตอบ ชุดของเครื่องมือในการวิเคราะห์ข้อมูลธุรกิจ (Business Analytics Tool) สามารถนำเข้าข้อมูลจากแหล่งข้อมูล (Data Source) สามารถทำงานร่วมกันกับไฟล์ข้อมูลประเภทต่างๆ เพื่อเชื่อมโยงข้อมูลที่ไม่เกี่ยวข้องให้เป็นข้อมูลเชิงลึกที่สอดคล้องกัน โดยใช้เครื่องมือที่ชื่อว่า Power

Query เพื่อเตรียมข้อมูลให้พร้อมใช้งานสามารถใช้สร้างรายงาน สามารถสร้างกระดานนำเสนอข้อมูล (Dashboard) โดยใช้กราฟหลากหลายรูปแบบผสมผสานในหน้าจอเดียวได้อย่างน่าสนใจ

8. จงอธิบาย MATLAB

คำตอบ เป็นโปรแกรมใช้สำหรับการคำนวณทางคณิตศาสตร์ทั้งการคำนวณเชิงตัวเลข และการคำนวณเชิงสัญลักษณ์ (ใช้โปรแกรมเสริมชื่อ Maple) นอกจากนี้ชุดคำสั่งของ MATLAB ยังเป็นภาษาคอมพิวเตอร์ ที่มีความสามารถในการคำนวณ Matrix ซึ่งช่วยในการทำงานด้านวิทยาศาสตร์ หลายด้านๆ ใช้สำหรับการคำนวณเชิงตัวเลขหลายกระบวนการสำหรับการประมวลผลข้อมูลทาง คณิตศาสตร์

9. Tableau คือ

คำตอบ เป็นซอฟต์แวร์ที่สามารถนำข้อมูลจำนวนมากที่มีหลากหลายประเภทในองค์กรมาทำการวิเคราะห์เพื่อเป็นข้อมูลเชิงธุรกิจสำหรับการตัดสินใจ หากต้องการใช้งานต้องซื้อสิทธิ์ การใช้ซอฟต์แวร์จากบริษัทตัวแทนปีต่อปี สามารถใช้สร้างการนำเสนอข้อมูล (Data Visualization) ด้วยกราฟพิศหลายประเภทในหน้าจอเดียว โดยมุ่งเน้นไปที่อุตสาหกรรมที่ทำงานในด้านของ Business Intelligent

10. แนวทางในการเล่าเรื่องด้วยข้อมูล (Data Storytelling) คือ

คำตอบ

(1) การเขียนแนวทางแบบมีเรื่องราว (Story) เนื้อหาการวิเคราะห์ก็จะต้องมีความแตกต่าง ไม่ซ้ำใคร ดูน่าเชื่อถือ และเข้าถึงได้ดี ได้เห็นภาพความเปลี่ยนแปลงต่างๆ เป็นฉากๆ เป็นเรื่องราว ข้อดี คือ ถ้านำเสนอดี จะทำให้เข้าใจภาพธุรกิจที่ดีและวางกลยุทธ์ได้ง่าย

(2) การเขียนแนวทางแบบอธิบายคำจำกัดความ (Definition) หาอ้างอิง (Reference) ได้ง่าย เขียนให้ได้ประโยชน์แบบตรงไปตรงมา แต่ข้อเสียคือ เนื้อหาจะไม่ค่อยโดดเด่นแตกต่าง ยากต่อการจดจำ และไม่รู้ว่าจะช่วยให้เข้าใจภาพธุรกิจได้มากขนาดไหน

(3) การนำเสนอแบบเล่าเรื่อง (Narrative) แนวทางนี้ยังไม่ได้ถึงกับมาเป็นเรื่องเป็นราว ข้อดีคือ สามารถออกแบบการนำเสนอให้น่าสนใจในความยาวกำลังดีได้ ถ้าดีก็ช่วยช่วยให้เข้าใจภาพธุรกิจได้มาก แต่ถ้ายังไม่ดี ก็อาจจะไม่เข้าใจได้เช่นกัน

(4) การนำเสนอแนวทางแบบเชิงเล่าเปรียบเทียบ (Analogy) เน้นที่การเปรียบเทียบกับอีกสิ่งหนึ่ง เพื่อใช้สิ่งนั้นช่วยทำให้คนเกิดความเข้าใจการเปรียบเทียบ (Analogy) ที่ดี จะทำให้เข้าใจภาพ

ธุรกิจเข้าใจคอนเซ็ปต์ของเรื่องยากๆ ได้ในเวลาอันรวดเร็ว ในขณะที่ถ้าเลือกการเปรียบเทียบที่ไม่ดี ก็อาจจะทำให้คนเกิดความเข้าใจผิด หรือไม่เห็นด้วยได้

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา

ดัชนี

A

A Massively Parallel Processing : MPP · 109

Accuracy · 196

Acquisition · 188

Actionable knowledge · 29

Activity diagram · 71

Actor · 65

Agglomerative · 162

Aggregating Data · 145

Agile · 120, 127

Agile Model · 84

Agile Software Development · 54

AID Automatic Intraction Detection · 154

Algorithm · 157, 174, 201

Alignment · 188

Analog · 29

Analogy · 204

Analysing System Needs · 26, 27

Analytics Layer · 108, 191, 193

Application · 32, 37, 39, 44, 100, 121, 136, 186

Application Architecture · 39, 45, 47, 51, 99

Application Security · 32

Apriori · 153, 166

Architect · 40, 44

Architecture · 36

Architecture of Big Data · 100, 102

Architecture Principle · 120

Architecture visio · 46

Artificial Intelligence · 33

Artificial intelligence (AI) · 52

Artificial Intelligence and Expert system: AI · 4, 12

Artificial intelligence: AI · 110

Artificial Intelligence: AI · 204

Artificial Neural Network · 180

Artificial Neural Network: ANN · 178

Artificial Neural Networks (ANN) · 154

As built · 124

As-Is · 46, 76

Assembly Lines · 29

Association · 136, 153

Association Rule · 168

Association Rule · 166

Association Rules · 166

Associations · 12

Attributes · 145

Authorization · 191

Autonomous Vehicles · 33

B

Back End Process · 118

Balanced Score Card · 38

Baseline · 47

Batch · 109

Batch processing · 4

Behavior Log · 146

Best Practices · 51

Big Data · 30, 100, 110, 151, 185

Big Data Analysis · 21, 52

Big Data System · 31

Bioinformatics · 29

Blockchain · 52
 Blueprint · 51, 110, 185
 Blueprints · 48
 Bootstrap · 195
 Bottom Line · 187
 Bug · 28
 Bug fix · 85
 Bugs · 119
 Business Architecture · 38, 46, 51, 99
 Business Architecture: BA · 45
 Business Intelligence · 5, 8, 11, 185
 Business Intelligence Platform · 109
 Business Intelligence: BI · 4, 7, 197
 Business Management · 40, 43
 Business Needs and Requirements · 22
 Business Objectives · 78, 187
 Business Objects · 47
 Business Process · 37, 38, 47, 100
 Business Process Management · 101
 Business Process Modules · 188
 Business processes · 36
 Business Processes · 47
 Business Rules · 118
 Business Strategy · 43
 Business Understanding · 136, 186

C
 Capture · 189
 Categorical Attributes · 146
 Categorical Concept Hierarchy Generation · 144
 Categories/ Class · 154
 Centroid · 164
Change Analysis · 111
 Chen Model · 64
 Class diagram · 70
 Class Diagram · 55, 124
 Classification · 8, 152, 154, 155
 Classifier · 154
 Cloud Computing · 21, 31, 32
 Cloud System · 105
 Cluster Center · 164
 Clustering · 136, 144, 160, 178
 CMMI : Capability Maturity Model Integration · 185
 COBIT · 77, 87
 Coding · 118
 Coding & Debug · 79
 Collaboration · 6
 Collaboration diagram · 70
 Collaborative Commerce · 187
 Collaborative Filtering · 154
 Collaborative Filtering (CF) · 176
 Compliance · 101
 Component diagram · 71
 Composite Models · 40
 Computer-Aided Software Engineering-CASE · 124
 Conformance · 188
 Confusion Matrix · 195
 Congestive Heart Failure : CHF · 145
 Consumption · 102, 191
 Content-Based Filtering · 154
 Contented-Based Filtering (CBF) · 175
 Control Objectives for Information and Related Technology · 87
 Control Systems · 31
 Convolution Neural Networks (CNN) · 154
 Coordination · 6

Correlation · 177

Cost Saving · 37

Credit Approval · 154

CRISP-DM · 136, 185

Cross Functional Process Improvement · 187

Cross-industry standard process for data
minin · 185

Cross-Validation · 193

Crow's Foot Model · 64

Current State of Enterprise Architecture · 37

Current System · 22

Customer Centric · 138

Customer Relationship · 107

Customer Relationship Management Process ·
118

Customer Relationship Management: CRM ·
137, 189

Customer Satisfaction · 187

Cyber Security · 32

Cyber-Physical System · 29

Cyber-Physical System - CPS · 31

D

Dashboard · 198

Data · 55

Data Access Diagram · 73

Data Acquisition · 107

Data Aggregation · 144

Data Aggregation Layer · 107

Data Analyst · 139, 142

Data Analytics · 31, 185, 193

Data Architecture · 45, 51, 99

Data Capture · 102

Data Centre · 100

Data cleaning · 143

Data Cleaning · 144, 191, 201

Data Communication · 20, 76

Data Correcting · 191

Data Cube · 144

Data Dictionary · 27, 55, 68, 69

Data Engineer · 191

Data Flow Diagram · 69

Data Flow Diagram - DFD · 55

Data Flow Diagram (DFD) · 72

Data Flow Diagram(DFD) · 27

Data Flow Diagram: DFD · 61, 62

Data Gateway · 198

Data gathering · 55

Data Governance · 192

Data governance layer · 109

Data Governance Layer · 108, 192

Data Integration · 143, 191

Data Lake · 107

Data Layer · 107, 189

Data Mart · 11

Data Mining · 11, 12, 13, 14, 105, 134, 137,
140, 151

Data Mining Process · 13

Data Model · 27, 198

Data Parsing · 191

Data Preparation · 136, 143, 186, 187, 190,
191, 192, 193

Data Preparation · 144

Data Preprocessing · 143, 144

Data Privacy · 109

Data Processing · 20, 27, 76

Data Project · 140

Data Reduction · 144, 161, 191

Data Relationship · 63

Data Scientist · 140, 142, 200

Data Security · 32, 109

Data Source · 198

Data Sources · 107

Data Standardization · 188

Data Store · 72

Data Storytelling · 197, 204

Data Structure · 73

Data Structure Diagram · 73

Data Transformation · 7, 102, 144, 190, 191

Data Understanding · 136, 186, 187, 188

Data Visualization · 107, 142, 199

Data Warehouse · 11, 13, 107

Database System · 13

Data-driven Organization · 36

DBSCAN · 153

Deadline · 123

Debugging · 83

Decision Support System – DSS · 4

Decision table · 55

Decision tree · 55

Decision Tree · 152, 155

Decision Tree Learning · 154

Decision Trees · 175

Decreasing Complexity · 111

Deep Learning Developer · 203

Deployment · 137, 186, 187

Deployment diagram · 71

Designing the Recommended System · 23

Designing the Systems Process · 27

Developer: Dev · 128

Developing and Documenting Software · 23, 28

Development team · 84

Digital · 29, 31

Digital Enterprise Blueprint · 110

Digital Enterprise Roadmap · 110

Digital Governance · 188

Digital Organization · 99

Digital Technology · 76

Digital World · 185

Digitization · 6

Dimensionality Reduction · 144

Dimensionality Reduction Technique · 178

Discretization Data · 144

Divisive · 162

Documentation · 23

Domain Expertise · 197

Drone · 52, 181

E

E-Commerce · 173

Education and Training · 22

Efficiency in Resource Utilization · 187

E-mail · 6

Embedded Systems · 31

Employee Self-Service (ESS) · 10

End-to-End Process · 54

Engagement Model · 54

Engineer · 40, 44

ensorFlow · 203

Enterprise · 36

Enterprise Architecture · 51, 54, 100, 116

Enterprise Architecture : EA · 37

Enterprise Architecture :EA · 110

Enterprise Architecture –EA · 40

Enterprise Architecture Frameworks · 51

Enterprise Architecture Maturity · 99

Enterprise Architecture: EA · 36

Enterprise Governance · 99

Enterprise Resource Planning (ERP) · 37

Enterprise Resource Planning: ERP · 189

Enterprise System · 188

Enterprise Systems · 54

Entity · 63, 155

Entity Diagram(ER) · 27

Entity Relationship Diagram: E-R Diagram ·
63

Entity-relationship diagram - ERD · 55

Entity-Relationship Diagram - ERD · 124

Entropy · 157

ER-Diagram · 63, 64

Error Rate · 193

Euclidean distance · 163

Evaluation · 137, 186

Evaporative cooling system- Evap · 32

Executive · 40, 43

Executive Information System – EIS · 8

Executive Support Systems: ESS · 4

Expert System :ES · 12

Explicit Rating · 172

Extends · 66

External Entities · 72

External Information System · 3

External Process · 117

External Source of Data · 27

Extract-Transform-Load · 107

F

Face Recognition · 33

Factor · 145

False Negative: FN · 195

Fax · 6

Feasibility Study · 23, 25, 27, 121

Feed back · 27

Feed Forward · 179

File Name · 56

Flow Chart · 8, 57

Flowchart · 55, 56, 124

Flowchart Diagram · 72

Framework · 37, 39, 40

Frequent itemset · 167

Functional Test · 28

Functions and Features · 85

Funding Priorities · 188

Future State of Enterprise Architecture · 38

Future State Outline · 120

G

Gain · 158

Gantt Chart · 55, 124

Gap Analysis · 123

Garbage In Garbage Out, GIGO · 106

Geographic Information System: GIS · 4

Goal Seeking · 7

Google Translation · 33, 203

Governance · 101

Graphical Analysis · 8

Groupware · 4

H

Hadoop Map/Reduce · 109

Hardware · 39

Heuristic-Based Techniques · 177

Hidden Layer · 178

Hierarchical Algorithms · 161

High Performance Computing · 13
Histogram Diagram · 144
Holdout Method · 194
Human Behavior · 188
Human Resource · 25
Hybrid Feedback · 173
Hybrid filtering · 178
Hybrid Recommendation Systems · 154
Hybrid systems · 5

I

Identifying Problems, Opportunity and Objective · 23
Image Captioning · 203
Image Processing · 6
Implementation Governance · 125
Implementing and evaluating the System · 23, 28
Increment · 85
Incremental Model · 81, 83
In-Database Analytics · 109
Inference · 135
Information and Communication Technology · 29
Information Architecture · 38, 47
Information Collection Phase · 174
Information Exploration Layer · 108, 192, 193, 197
Information Gain · 157
Information Science · 13, 135
information system · 51
Information System · 3, 29
Information Systems · 3
Information systems architectures · 47

Infrastructure · 38, 44, 100
Input · 27, 154
Input Layer · 178
Insight Data · 104
Internal audit · 126
Internal Data · 138
Internal Information System · 3
Internal Process · 117
Internal Source of Data · 27
Internet of Things · 21
Internet of Things - IoT · 32
Internet of Things (IoT) · 31
Interval Label · 144
Investment · 25
ISO : International Organization for Standardization · 185
Isolation Forest · 153
Item-Based Filtering · 172, 177
Iteration Flowchart · 59
Iterative Development · 129
Iterative Model · 81
Iterative Process · 193

K

Key Performance Indicator · 120
Key Stakeholders · 43
K-Fold Cross Validation · 194
K-means Clustering · 136
Knowledge Discovery in Database: KDD · 12, 134

L

Latent Semantic Methods · 178

Learning Phase · 174
Leave Online System · 6
Leave-One-Out Method · 194
Lifelog · 139
Lifelogs · 104, 138
Line · 7
Linear Model · 79
Logic · 55
Logical Diagram · 118
Logistic Regression · 152, 155

M

Machine learning · 177
Machine Learning · 13, 29, 134, 140, 142, 175, 201
Macroeconomic environments · 52
Management Capabilities · 188
Management Information Systems: MIS · 3
Manhattan distance · 163
Master Data Management: MDM · 109, 192
Mathematical Programming · 13, 135
MATLAB · 201
Matrix Completion Techniques · 178
Mean Absolute Test Set Error · 194
Mean-Shift Clustering · 153
Measurement and Continuous Improvement · 126
Medical Business Analytics · 29
Medical Diagnosis · 154
Membership System · 19
Memory Based Techniques · 177
Metadata · 176
Metamodel · 40
Method · 37

Methodologies · 54
MINISPECIFICATION · 69
Minkowski Distance · 163
Min-Max Decimal Scaling · 144
Missing values · 190
Missing Values · 143
Mission · 52
Model Based Techniques · 177

Model Evaluation · 193

Model Predictive Control · 181
Model tools · 55
Modeling · 22, 137, 186, 187
Modern Business Management · 1
Modern Trade · 1, 31
Modern Trade and Business · 20
Multidimensional Data Model · 144
Multilayer Perceptron: MLP · 181

N

Naïve Bayes Classifier · 175, 176
Naïve Bayesian · 152
Natural Language Processing: NLP · 12
Natural Language Toolkit (NLTK) · 201
Neglecting context · 192
Neighbourhood · 176
Network Security · 32
Network System · 39
Neural Networks · 175
Nominal Categorical · 146
Non- Relational Database · 101
Non-functional Test · 28
Non-Hierarchical Algorithms · 161
Non-parametric Methods · 144
Normalization · 144

O

Object · 55
 Object diagram · 70
 Object Modeling · 70
 Office Information System: OAS/OIS · 3
 OLAP · 11
 OLAP (Online Analytical Processing) · 199
 Online Analytical Processing (OLAP) · 134
 Online Analytical Processing : OLAP · 13
 Online processing · 4
 Online Social Media · 133
 On-Premise · 198
 Open Group Architecture Framework
 (TOGAF) · 37
 Operating Model · 53, 54
 Operating System · 101
 Operational Efficiency · 187
 Operations Research & Numerical Methods ·
 11
 Opportunities and Solution · 47
 Optimization · 140, 142
 Ordinal categorical · 146
 Organization Chart · 47
 Original Pool · 195
 Outlier detection · 153
 Outliers · 190
 Output · 27, 155
 Output Class · 195
 Output Layer · 179
 Outsource · 29
 Over-View of System · 72

P

Parametric Methods · 144
 Partitional clustering · 161
 Patterns · 12
 Payroll · 5
 Penetration Test · 28
 Performance Measurement · 88
 Performance Test · 28
 Personal Computer · 198
 PHR - Personal Health Record · 30
 Physical Model · 27
 Planning and Organization · 88
 Platform · 42, 172
 Point · 7
 Point of Sale · 24
 Point of Sales (POS) · 31
 Polygon · 7
 POS (Point of Sale) · 103
 Power BI · 198
 Power Query · 198
 Precision · 197
 Prediction · 174
 Preliminary · 46
 Primary Stakeholders · 44
 Primitive Models · 40
 Priorities and Tasks · 85
 Process · 55
 Process Integration and Standardization · 53
 Process Specification · 27
 Process Standardization · 188
 Processes and Procedures · 55
 Product Backlog · 85
Product owner · 84
 Product Owner · 84

Product Owner: PO · 128

Project management · 55

Proposal Manager · 129

Prototype · 27, 123

Pseudocode · 69

Purchasing · 5

Python · 142

Q

Quantitative Attributes · 146

Quantum computing · 52

R

Random Forests · 152

Random Sub-Sampling · 194

Raw data · 107

Readmission criteria · 145

Real Time · 38

Recall · 197

Receiver Operating Curves (ROC) · 195

Recommendation System · 33, 140, 154, 173, 174

Recurrent Neural Networks (RNN) · 154

Reducing Wasted Time and Cost Saving · 111

Redundancy · 48

Regression · 178

Regression Tree) · 154

Relational Database · 101

Relationships · 198

Requirements Gathering · 22

Residual · 194

Resource Management · 87

Responsibility · 188

Return of Investment · 122

Return of Investment : ROI · 78

Risk Analysis · 7

Risk Management · 88, 101

Roadmap · 37, 45

Robotic Process Automation: RPA · 110

Robotics · 52

Rule-Based · 158

S

Sampling · 144

Scalability · 172

Scenario · 37

Screen Design · 118

Scrum Framework · 84

Scrum master · 84

Scrum Master · 128

SDLC · 51

Secondary Stakeholders · 44

Security · 121

Security and Privacy · 110

Security Architecture · 51, 99

Selection Flowchart · 58

Self-Development · 205

Sensor · 54

Sequence diagram · 70

Sequence Diagram · 55, 124

Sequence Flowchart · 58

Shared Data · 188

Shared Infrastructure · 188

Similarity · 177

Smart Devices · 29

Smart Factory · 32

Smart Farming · 32
 Smart Manufacturing · 32
 Smart Phone · 33
 Smart Watch · 145
 Smoke Test · 28
 Social Media · 107, 181
 Software · 32, 39
 Software Development · 77
 Software Evolution · 77
 Solution Architecture · 37
 Source of Big Data · 102
 Sparse · 172
 Speed to Market · 187
 Spiral Model · 54
Spotting Data Quality Issues · 192
 Spreadsheets · 199
 Sprint Backlog · 85
 Sprint Goal · 85
 Sprint Planning · 85
 Sprints · 84
 Squared Correlation · 197
 Squared Euclidean Distance · 161
 Stakeholder · 37, 40, 43, 128
 Stakeholders · 120
 State chart diagram · 70
 Statistic and Data Analysis Methods · 13
 Statistical Analysis System · 201
 Storage · 27
 Strategic Agility · 187
 Strategic Alignment · 87
 Strategy · 188
 Stream Computin · 109
 Structure Analysis Model · 72
 Structured Data · 103
 Structured Language · 55
 Supervised Learning · 154
 Supervised Machine Learning · 135
 Supplier · 10
 Supply Chain · 10
 Support Vector Machine · 152
 Supporting Application · 100
 Swimming Lane · 56
 Synapses · 178
Synthetic biology · 52
 System · 32
 System Analysis · 84
 System Analysis: SA · 22
 System Analysts · 125
 System Design Specification · 123
System Development Cycle · 21
 System Development Life Cycle · 54, 117
 System Development Life Cycle: SDLC · 23
 System Development Model · 22
 System Improvement Documentation · 23
 System Integration Testing (SIT) · 28
 System Maintenance · 23
 System Monitoring and Management · 99
 Systems analysis and design :SAD · 29
 Systems development Model · 55
 Systems' Components for Designing · 5, 27

T
 Tableau · 142, 199
 Tablet · 29
 Target Marketing · 154
 Technical Architecture · 39
 Technical Skill · 204
 Technician · 40, 44
 Techniques · 55

Technology · 25

Technology Architecture · 46, 47, 99

Technology Feasibility · 25

Technology Standardization · 188

Teleconferencing · 6

Testing and Maintaining the System · 23, 28

Testing Dataset · 194

The Internet of Things (IoT) · 52

The Ontological · 40

The Open Group · 42

The Third Industrial Revolution · 29

Threshold · 179

Time Line · 123

Time Series Data Mining: TSDM · 152

Time-Boxed · 84

Timeline · 41

Timing · 25

To-Be · 46, 76

TOGAF · 37, 42, 45

Tracking patterns · 152

Traditional Approach · 84

Traditional DB · 189

Training Data · 194

Transaction Processing Systems: TPS · 3

Transform · 107

True Positive: TP · 195

U

Unified modeling language · 124

Unified Modeling Language · 55

Unified Modeling Language-UML · 70

Unit Test · 28

Unstructured Data · 103

Unsupervised Learning · 161

Unsupervised Machine Learning · 135

Use Case · 66

Use case diagram · 70

Use Case Diagram · 55, 65, 124

User · 29, 76

User Acceptance Testing (UAT) · 28

User manual · 23

User story · 85

User Training · 23

User-Based Filtering · 172, 177

User-Item Matrix · 176

UX/UI · 28

V

Value · 106

Value Chain · 117

Value Delivery · 87

Value Insight · 204

Variability · 106

Variety · 106

Vector Space · 175

Velocity · 106

Veracity · 106

view point · 46

View Point · 40

Visibility · 101

Vision · 52

Visualization · 13, 135, 151

Visualize · 205

Visulization · 106

Vital Log · 146

voice mail · 6

Volume · 105

W

Waterfall Model · 54, 79

Web Browser · 10

Weka · 201

What-If Analysis · 7

Word processing · 6

Workflow · 6

Z

Zachman Framework · 40, 41

Z-Score · 144

ก

กฎเกณฑ์ทางธุรกิจ · 118

กฎความสัมพันธ์ · 166

กรอบแนวทาง · 37

กรอบรูปแบบพันธสัญญาสู่ความสำเร็จ · 54

กรอบสถาปัตยกรรมองค์กร · 51

กระบวนการทางธุรกิจ · 38

กระบวนการ · 55

กระบวนการการทำงาน · 37

กระบวนการทำงานทางธุรกิจ · 36

กระบวนการทำซ้ำ · 193

กระบวนการทำเหมืองข้อมูล · 13

กระบวนการธุรกิจ · 100

กระบวนการธุรกิจในแต่ละหน่วย · 188

กระบวนการบริหารลูกค้าสัมพันธ์ · 118

กระบวนการปฏิบัติการภายนอก · 117

กระบวนการปฏิบัติการภายใน · 117

กระบวนการและรูปแบบวิธีปฏิบัติ · 55

กระบวนการและวิธีปฏิบัติด้านการรักษาความ

ปลอดภัยข้อมูล · 32

กระบวนการและวิธีปฏิบัติด้านความปลอดภัย

ปกป้องเครือข่าย · 32

กระบวนการและวิธีปฏิบัติด้านแอปพลิเคชัน · 32

กระบวนการวิเคราะห์ออนไลน์ · 13

กลยุทธ์ · 188

กล่องเวลา · 84

การกำกับดูแลด้านเทคโนโลยีดิจิทัล · 188

การกำกับดูแลและธรรมาภิบาล · 101

การเก็บข้อมูล · 107

การเก็บข้อมูลชีวิตประจำวัน · 139

การเก็บรวบรวมข้อมูล · 102

การเก็บรวบรวมข้อมูล · 189

การแก้ไขข้อมูลที่ผิดพลาด · 191

การแก้ปัญหาทางคณิตศาสตร์ · 11

การค้นพบองค์ความรู้ใหม่จากฐานข้อมูล · 12

การคำนวณประสิทธิภาพสูง · 13

การจัดกลุ่ม · 136

การจัดการกระบวนการทางธุรกิจ · 101

การจัดการความเป็นส่วนตัว · 109

การจัดการโครงการ · 55

การจัดซื้อจัดหา · 188

การจัดเตรียมข้อมูล · 144

การจัดลำดับขั้นแนวคิด · 144

การจำแนกประเภท · 152, 154

การแจกแจงข้อมูล · 191

การตรวจจับค่าผิดปกติ · 153

การตรวจประเมินภายใน · 126

การตรวจวินิจฉัย · 154

การเตรียมข้อมูล · 136, 143, 186, 193

การถดถอยโลจิสติกส์ · 152

การทบทวนข้อเสนอแนะ · 27

การทำความเข้าใจข้อมูล · 136, 186, 188

การทำความเข้าใจธุรกิจ · 186

การทำความเข้าใจประเด็นทางธุรกิจ · 136

การทำความเข้าใจข้อมูล · 143, 144

การทำจินตทัศน์ · 13, 135
การทำตลาดลูกค้ากลุ่มเป้าหมาย · 154
การทำโมเดล · 22, 137, 186
การทำเหมืองข้อมูล · 11, 12, 14, 105, 134, 137, 140, 151, 186
การนำเข้าจะจะได้ผลลัพธ์ขยออก · 106
การนำไปใช้ · 102, 137
การนำผลไปใช้งาน · 186
การบริหารจัดการความเสี่ยง · 101
การบริหารจัดการมาตรฐานกระบวนการ · 188
การบริหารจัดการมาตรฐานข้อมูล · 188
การบริหารลูกค้าสัมพันธ์ · 137
การบำรุงรักษาระบบงาน · 23
การปฏิบัติอุตสาหกรรมครั้งที่สาม · 29
การประชุมทางไกล · 6
การประมวลคำ · 6
การประมวลผลข้อมูล · 76
การประมวลผลด้วยเครือข่ายคลาวด์ · 31
การประมวลผลแบบสตรีม · 109
การประมวลผลเป็นชุด · 4
การประมวลผลภาพลักษณ์ · 6
การประเมิน · 137
การประเมินผลโมเดล · 186
การประสานงาน · 6
การปรับปรุงกระบวนการระหว่างฝ่ายและหน่วยธุรกิจ · 187
การแปลงข้อมูล · 144
การแปลงรูปข้อมูล · 102
การผลิตอัจฉริยะ · 32
การฝึกอบรม · 22
การเฝ้าระวังและบริหารจัดการระบบ · 99
การพัฒนาซอฟต์แวร์ · 77
การเพิ่มประสิทธิภาพ · 140
การมองเห็น · 101
การรวบรวมข้อมูล · 55

การรวมข้อมูล · 143, 144
การร่วมมือ · 6
การระบุคุณภาพของข้อมูล · 192
การรักษาความปลอดภัยข้อมูล · 109
การเรียนรู้ของเครื่อง · 13, 29, 134, 140
การเรียนรู้ของเครื่องแบบมีผู้สอน · 154
การเรียนรู้แบบต้นไม้ตัดสินใจ · 154
การเรียนรู้แบบมีผู้สอน · 154
การเรียนรู้แบบไม่มีผู้สอน · 161
การลดขนาดข้อมูล · 161
การลดต้นทุน · 37
การลดรูปข้อมูล · 144
การละเลยบริบท · 192
การเล่าเรื่องด้วยข้อมูล · 204
การวัดประสิทธิภาพองค์กร · 126
การวิเคราะห์การเปลี่ยนแปลง · 111
การวิเคราะห์ข้อมูล · 185
การวิเคราะห์ข้อมูลขนาดใหญ่ · 21, 52
การวิเคราะห์พื้นฐานข้อมูล · 109
การวิเคราะห์และออกแบบระบบ · 29
การศึกษาถึงความต้องการของธุรกิจ · 22
การสั่งซื้อสินค้า · 5
การสื่อสารข้อมูล · 76
การสุ่มตัวอย่าง · 144
การแสดงข้อมูลในรูปแบบรูปภาพ · 107
การหาความต้องการ · 22
การหาความสัมพันธ์ · 153
การไหลของงาน · 6
การอนุมาน · 135
การออกแบบระบบ · 23
กำหนดเส้นตาย · 123
เกณฑ์การรับผู้ป่วยเข้าใหม่ · 145

ข

ข้อมูล . 55
ข้อมูลกลุ่มพฤติกรรม . 146
ข้อมูลกลุ่มสุขภาพ . 146
ข้อมูลขนาดใหญ่ . 30, 110, 185
ข้อมูลเข้า . 27
ข้อมูลจากแหล่งภายนอกองค์กร . 27
ข้อมูลจากแหล่งภายใน . 27
ข้อมูลชีวิตประจำวัน . 104
ข้อมูลเชิงลึก . 104
ข้อมูลดิบ . 107
ข้อมูลที่เป็นตัวเลข . 146
ข้อมูลที่เป็นนามบัญญัติ . 146
ข้อมูลที่เป็นลักษณะลำดับขั้นหรือชั้น . 146
ข้อมูลที่เป็นหมวดหรือกลุ่ม . 146
ข้อมูลที่มีโครงสร้างชัดเจน . 103
ข้อมูลที่ไม่ได้จัดเก็บในรูปแบบฐานข้อมูลเชิงสัมพันธ์ . 101
ข้อมูลที่ไม่มีโครงสร้าง . 103
ข้อมูลแบบไม่ต่อเนื่อง . 144
ข้อมูลผ่านโซเชียลออนไลน์ . 133
ข้อมูลภายใน . 138
ข้อมูลมีค่าผิดปกติ . 190
ขั้นตอนการดำเนินงานและประเมิน . 28
ขั้นตอนการทดสอบและบำรุงรักษาระบบ . 28
ขั้นตอนการพัฒนาซอฟต์แวร์และจัดทำเอกสาร . 28
ขั้นตอนการออกแบบระบบ . 27
ขั้นตอนวิธี . 37

ค

ค้นหาปัญหา โอกาสและเป้าหมาย . 23
คลังข้อมูล . 107
คลาวด์คอมพิวติ้ง . 32

ความถูกต้อง . 106
ความปลอดภัยของข้อมูลและความเป็นส่วนตัว . 110
ความเป็นไปได้ในเชิงเทคโนโลยี . 25
ความแปรผันได้ . 106
ความพึงพอใจของลูกค้า . 187
ความมั่นคงปลอดภัย . 121
ความมั่นคงปลอดภัยทางไซเบอร์ . 32
ความรวดเร็วในการผลิตสู่การตลาด . 187
ความเร็ว . 106
ความสอดคล้องกัน . 188
ความสอดคล้องกับกฎระเบียบ . 104
ความสัมพันธ์ . 12
ความสัมพันธ์ระหว่างข้อมูล . 63
ความหลากหลาย . 106
ค่าความเหมือน . 177
ค่าช่วง . 144
คุณค่า . 106
คุณลักษณะ . 145
คู่มือระบบ . 23
เครือข่าย 5G . 52
เครื่องมือแบบจำลอง . 55
โครงการด้านข้อมูล . 140
โครงข่ายประสาทเทียม . 180
โครงสร้างข้อมูล . 47, 73
โครงสร้างพื้นฐาน . 38, 121

ง

งานเงินเดือน . 5

จ

จำแนกข้อมูล . 154
จุด . 7
เจ้าของโครงการ . 128

ซ

ชุดข้อมูลที่ใช้ทดสอบ · 194

ชุดข้อมูล · 194

ช

ซอฟต์แวร์ · 39

ซอฟต์แวร์โรบอทที่เลียนแบบพฤติกรรมของมนุษย์ · 110

ชีฟฟลายเออร์ · 10

เซนเซอร์ · 54

โซเซียลมีเดีย · 107, 181

ฐ

ฐานข้อมูลเชิงสัมพันธ์ · 101

ด

ดาต้าคิวบ์ · 144

ดาต้ามาร์ · 11

ด้านความคุ้มทุน · 25

ด้านทรัพยากรบุคคล · 25

ด้านเวลา · 25

ดำเนินงานและประเมินผล · 23

ดิจิทัล · 29

โดรน · 181

ไดอะแกรมกระบวนการ · 118

ต

ต้นไม้แห่งการตัดสินใจ · 152

ตรรกะ · 55

ตัวชี้วัดผลสำเร็จ · 120

ตัวแปรภายนอก · 72

ตารางตัดสินใจ · 55

ท

ทดสอบและบำรุงรักษาระบบ · 23

ทะเลสาบข้อมูล · 107

ทักษะด้านเทคนิค · 204

ทีมพัฒนา · 128

เทคนิค · 55

เทคนิคการแบ่งกลุ่ม · 160

เทคนิคการเรียนรู้แบบมีผู้สอน · 135

เทคนิคการเรียนรู้แบบไม่มีผู้สอน · 135

เทคนิคโครงข่ายประสาทเทียม · 178

เทคโนโลยี · 25

เทคโนโลยีการพิมพ์ 3 มิติ · 52

เทคโนโลยีคลาวด์ · 21

เทคโนโลยีดิจิทัล · 76

เทคโนโลยีฮาร์ดแวร์ · 29

โทรสาร · 6

ธ

ธรรมาภิบาลขององค์กรในการใช้ระบบสารสนเทศ · 125

ธุรกิจการค้าสมัยใหม่ · 1, 20

ธุรกิจค้าปลีกสมัยใหม่ · 1

ธุรกิจอัจฉริยะ · 197

ธุรกิจอีคอมเมิร์ซ · 173

น

นักวิเคราะห์ข้อมูล · 139, 142

นักวิเคราะห์ระบบ · 22, 84

นักวิเคราะห์และออกแบบ · 125

นักวิทยาศาสตร์ข้อมูล · 200

นักวิทยาศาสตร์ข้อมูล · 140, 142

นาฬิกาอัจฉริยะ · 145

แนวทางสถาปัตยกรรม · 120

ในรูปดิจิทัล · 6

บ

แบบจำลองการวิเคราะห์ที่เป็นโครงสร้าง · 72

แบบจำลองข้อมูลหลายมิติ · 144

แบบจำลองเชิงวัตถุ · 55, 70

แบบป้อนไปข้างหน้า · 179

แบบพิมพ์เขียว · 48

แบบมีการป้อนไปเวียนกลับ · 180

ป

ประมวลผลข้อมูล · 27

ประสิทธิภาพในการใช้ทรัพยากรขององค์กร · 187

ปริมาณ · 105

ปัจจัย · 145

ปัจจัยแวดล้อมทางเศรษฐกิจมหภาค · 52

ปัญญาประดิษฐ์ · 33, 52, 110, 204

โปรแกรมทางคณิตศาสตร์ · 13, 135

โปรแกรมประยุกต์ · 100, 121, 136, 186

โปรแกรมประยุกต์สำหรับสนับสนุน · 100

ไปรษณีย์เสียงหรืออีเมลเสียง · 6

ไปรษณีย์อิเล็กทรอนิกส์หรืออีเมล · 6

ผ

ผลลัพธ์ · 27

ผังงาน · 55, 56, 124

ผังงานแบบทำซ้ำ · 59

ผังงานแบบมีเงื่อนไข · 58

ผังงานเรียงลำดับ · 58

ผังต้นไม้ · 55

ผู้จัดการโครงการ · 129

ผู้ใช้ · 29

ผู้ใช้งานระบบ · 76

ผู้ที่เกี่ยวข้อง · 120

ผู้มีส่วนได้ส่วนเสีย · 43, 128

ผู้มีส่วนได้ส่วนเสียขั้นพื้นฐาน · 44

ผู้มีส่วนได้ส่วนเสียชั้นรอง · 44

ผู้มีส่วนได้ส่วนเสียหลัก · 43

แผนที่นำทางองค์กรดิจิทัล · 110

แผนภาพกระแสข้อมูล · 55, 61, 62

แผนภาพความสัมพันธ์ระหว่างเอนทิตี · 55

แผนภาพแสดงความสัมพันธ์ระหว่างข้อมูล · 63

แผนภาพฮิสโตแกรม · 144

ผ

ฝึกอบรมผู้ใช้งาน · 23

พ

พจนานุกรมข้อมูล · 55

พฤติกรรมบุคคล · 188

พัฒนาซอฟต์แวร์และจัดทำเอกสาร · 23

พันธกิจ · 52

พิมพ์เขียวองค์กรดิจิทัล · 110

แพลตฟอร์ม · 42

โพลีกอน · 7

พ

เฟรมเวิร์ค · 39

แพลตฟอร์ม · 172

ภ

ภาพรวมของระบบ · 72

ภาษาโครงสร้าง · 55

ม

มาตรฐานด้านกระบวนการงาน · 53
มุมมองของผู้บริหาร · 43
มุมมองช่างเทคนิค · 40, 44
มุมมองเชิงกลยุทธ์ · 43
มุมมองด้านข้อมูล · 44
มุมมองด้านโครงสร้างพื้นฐาน · 44
มุมมองด้านความมั่นคงปลอดภัย · 44
มุมมองด้านเวลา · 44
มุมมองด้านแอปพลิเคชัน · 44
มุมมองผู้จัดการธุรกิจ · 40, 43
มุมมองผู้บริหาร · 40
มุมมองระดับองค์กร · 40
มุมมองวิศวกร · 40, 44
มุมมองสถาปนิก · 40, 44
โมเดลการปฏิบัติงาน · 53
โมเดลโครงสร้างแบบกันหอย · 54
โมเดลโครงสร้างแบบน้ำตก · 54
โมเดลน้ำตก · 79
โมเดลวงจรพัฒนาระบบ · 55
โมเดลวงจรพัฒนาระบบ SDLC · 54

ย

ยูเอ็มแอล · 70, 124

ร

ระบบ · 32
ระบบการขับเคลื่อนอัตโนมัติ · 33
ระบบการแนะนำ · 154
ระบบการประมวลผลข้อมูล · 20
ระบบการประมวลผลภาษาธรรมชาติ · 12
ระบบการลาออนไลน์ · 6
ระบบการสื่อสารข้อมูล · 20

ระบบข้อมูลขนาดใหญ่ · 100
ระบบคลังข้อมูล · 11, 13, 100
ระบบคลาวด์ · 105
ระบบควบคุม · 31
ระบบเครือข่าย · 39
ระบบงานเดิม · 22
ระบบจดจำใบหน้า · 33
ระบบไซเบอร์ · 29
ระบบฐานข้อมูล · 13
ระบบฐานข้อมูลขนาดใหญ่ · 31
ระบบดิจิทัล · 31
ระบบทางวิศวกรรมที่บูรณาการโลกกายภาพ · 31
ระบบธุรกิจอัจฉริยะ · 4, 7, 11, 109
ระบบธุรกิจอัจฉริยะ · 185
ระบบแนะนำ · 140, 173
ระบบปฏิบัติการ · 101
ระบบประมวลผลรายการประจำวัน · 4
ระบบปัญญาประดิษฐ์และระบบผู้เชี่ยวชาญ · 4, 12
ระบบผู้เชี่ยวชาญ · 12
ระบบฝังตัว · 31
ระบบสนับสนุนผู้บริหารระดับสูง · 4
ระบบสมาชิก · 19
ระบบสารสนเทศ · 3, 29, 51
ระบบสารสนเทศแบบประมวลผลรายการ · 3
ระบบสารสนเทศเพื่อการจัดการ · 3
ระบบสารสนเทศภายนอก · 3
ระบบสารสนเทศภายใน · 3
ระบบสารสนเทศภูมิศาสตร์ · 4
ระบบสารสนเทศสำนักงานอัตโนมัติ · 3
ระบุโอกาสและแนวทางการแก้ปัญหา · 47
ระยะเตรียมการ · 46
ระยะที่สอง · 46
ระยะห่างเชิงยุคคิดยกกำลังสอง · 161
รูปแบบ · 12
รูปแบบการติดตาม · 152

รูปแบบทางกายภาพ · 27

เรียลไทม์ · 38

โรคหัวใจล้มเหลว · 145

โรงงานอัจฉริยะ · 32

โรงเรือนอัจฉริยะ · 32

โรงเรือนอีแวป · 32

วิธีการผสมผสาน · 5

วิธีการพัฒนาซอฟต์แวร์แบบคล่องแคล่วว่องไว · 54

วิธีพยากรณ์แบบจำลอง · 181

วิธีพาราเมตริก · 144

วิศวกรข้อมูล · 191

วิสัยทัศน์ · 52

ล

ลดความซับซ้อน · 111

ลดเวลาที่เสียไปและลดต้นทุน · 111

ลูกค้าเป็นศูนย์กลาง · 138

เลเยอร์การกำกับดูแลข้อมูล · 108, 109

เลเยอร์การรวมข้อมูล · 107

เลเยอร์การวิเคราะห์ · 108

เลเยอร์การสำรวจข้อมูล · 108

เลเยอร์ข้อมูล · 189

โลกดิจิทัล · 185

ว

วงจรการพัฒนาระบบ · 51

วงจรการพัฒนาระบบงาน · 22

วงจรการพัฒนาระบบงาน · 23

วงจรการวิเคราะห์ระบบงาน · 21

วัตถุ · 55

วัตถุประสงค์เชิงธุรกิจ · 78, 187

วิเคราะห์ความต้องการของระบบ · 23, 26, 27

วิทยาศาสตร์สารสนเทศ · 13, 135

วิธีการ · 54

วิธีการจัดกลุ่ม · 144

วิธีการที่ไม่ใช่พารามิเตอร์ · 144

วิธีการแบบดั้งเดิม · 84

วิธีการแบบไม่เป็นลำดับขั้น · 161

วิธีการแบบลำดับขั้น · 161

ศ

ศึกษาความเป็นไป · 25

ศึกษาความเป็นไปได้ · 23, 25, 27, 121

ส

สถานะปัจจุบันสถาปัตยกรรมองค์กร · 37

สถาปัตยกรรมของข้อมูลขนาดใหญ่ · 100, 102

สถาปัตยกรรมของธุรกิจองค์กร · 38

สถาปัตยกรรมด้านการรักษาความปลอดภัยข้อมูล · 99

สถาปัตยกรรมด้านข้อมูล · 99

สถาปัตยกรรมด้านเทคโนโลยี · 99

สถาปัตยกรรมด้านธุรกิจ · 99

สถาปัตยกรรมด้านโปรแกรมประยุกต์ · 99

สถาปัตยกรรมทางเทคนิค · 39

สถาปัตยกรรมโปรแกรมประยุกต์ · 39

สถาปัตยกรรมสารสนเทศ · 38

สถาปัตยกรรมองค์กร · 36, 37, 51, 54, 100, 110, 116

ส่วนจัดเก็บข้อมูล · 27

สาธารณูปโภค · 100

สายการประกอบ · 29

สำหรับเลเยอร์การสำรวจข้อมูล · 192

สิทธิ์ในการเข้าถึง · 191

เส้น · 7

ห

หน่วยเก็บข้อมูล · 72

หมวดหมู่ของข้อมูล · 154

หลักสถิติ · 13

ห่วงโซ่คุณค่า · 117

หุ่นยนต์ · 52

เหมืองข้อมูล · 13

แหล่งข้อมูล · 107, 198

แหล่งที่มาของข้อมูลขนาดใหญ่ · 102

อ

องค์ความรู้ใหม่จากฐานข้อมูล · 134

องค์ประกอบระบบที่ต้องออกแบบ · 5, 27

อนาคตของสถาปัตยกรรมองค์กร · 38

อนุกรมเวลา · 152

อัตราผิดพลาด · 193

อัลกอริทึม · 174

อากาศยานไร้คนขับ · 52

อินเทอร์เน็ตในทุกสิ่ง · 32

อินเทอร์เน็ตออฟติง · 21

อีไอเอส · 8

เอกสารประกอบการปรับปรุงระบบงาน · 23

เอกสารประกอบระบบ · 23

แอปพลิเคชัน · 32

แอปพลิเคชัน · 32, 37, 39

ฮ

ฮาร์ดแวร์ · 39

ศูนย์บรรณสารและการเรียนรู้ สถาบันนวัตกรรมการเรียนรู้ มหาวิทยาลัยพะเยา



ทรัพยากรเล่มนี้ ได้มอบให้ศูนย์บรรณสารและการเรียนรู้

สถาบันวัฒนธรรมการเรียนรู้ มหาวิทยาลัยพะเยา

เพื่อเผยแพร่ผลงานทางวิชาการ

ห้ามผู้ใดทำซ้ำ คัดลอก ลอกเลียน ดัดแปลง ปลอมแปลง จัดเผยแพร่

จำหน่าย ให้เช่า เข้าครอบครอง เรียกถึงข้อมูล บันทึก ส่งผ่าน

หรือกระทำการใด ๆ เกี่ยวกับลิขสิทธิ์ของเล่ม โดยไม่ชอบด้วยกฎหมาย

หรือโดยไม่ได้รับอนุญาตอย่างเป็นทางการจากเจ้าของผลงาน

ศูนย์บรรณสารและการเรียนรู้ สถาบันวัฒนธรรมการเรียนรู้ มหาวิทยาลัยพะเยา