เอกสารประกอบการสอบคัดเลือก

น.ประทวนเลื่อนฐานะเป็น น.สัญญาบัตร

โควตา น.ประทวนทำหน้าที่ในตำแหน่ง น.สัญญาบัตร

วิชาระบบฐานข้อมูล การสื่อสารข้อมูลและเครือข่าย

ระบบฐานข้อมูลเบื้องต้น (Database System)

ความรู้เบื้องต้นเกี่ยวกับข้อมูลและการจัดการข้อมูล

การดำเนินชีวิตในปัจจุบันของแต่ละคนจะเกี่ยวข้องกับข้อมูลและฐานข้อมูลทั้งสิ้น เช่น การซื้อสินค้าจากห้างสรรพสินค้า การลงทะเบียนเรียน และการฝากถอนเงิน เป็นต้น ซึ่งองค์กรหรือหน่วยงานธุรกิจต่างๆ ได้เล็งเห็นถึงความสำคัญของการเก็บรวบรวมข้อมูลอย่างเป็นระบบ เพื่อความสะดวกในการบันทึกข้อมูล ค้นหาข้อมูล ลดความซ้ำซ้อนของการจัดเก็บข้อมูล สามารถเปลี่ยนแปลงและแก้ไขข้อมูลให้ทันสมัยอยู่เสมอ

ข้อมูลและธรรมชาติของข้อมูล

ในการจัดการฐานข้อมูลต้องอาศัยความเข้าใจเกี่ยวกับข้อมูลที่ต้องการจะจัดเก็บ ซึ่งมีอยู่หลายชนิด เช่น ข้อมูลที่เป็นตัวอักขระหรือข้อความ ข้อมูลเชิงจำนวน ข้อมูลรหัส ข้อมูลวันที่ ข้อมูลภาพลักษณ์ ข้อมูลภาพเคลื่อนไหว และข้อมูลเสียง เป็นต้น ส่วนประเภทของข้อมูลก็มีอยู่หลากหลายได้แก่ ข้อมูลเกี่ยวกับการศึกษา การดำเนินธุรกิจ สาธารณสุข สิ่งแวดล้อม และความบันเทิง ซึ่งล้วนแล้วแต่ประกอบด้วยข้อมูลที่แตกต่างกันไป รวมทั้งรายละเอียดของข้อมูลในแต่ละประเภทก็ไม่เหมือนกัน ขึ้นอยู่กับว่าเราสนใจจะจัดเก็บข้อมูลอะไร และที่สำคัญสิ่งที่ต้องการจะจัดเก็บต้องเกี่ยวข้องหรือสัมพันธ์กัน ในการจัดเก็บข้อมูลที่มีปริมาณมากๆ จะใช้คอมพิวเตอร์เข้ามาช่วยในการจัดเก็บและจัดการกับข้อมูล ซึ่งอาจเก็บไว้ในแฟ้มข้อมูล และมีการพัฒนาขึ้นมาเป็นฐานข้อมูล ซึ่งมีการจัดการกับข้อมูลได้ดีกว่า โดยมีการนำโปรแกรมสำเร็จรูปมาใช้ในการจัดการฐานข้อมูลด้วย ซึ่งขึ้นอยู่กับรูปแบบของฐานข้อมูลที่ใช้ และลักษณะของข้อมูลในองค์กร หรือถ้าหากต้องการจะจัดเก็บฐานข้อมูลเกี่ยวกับทรัพยากรสารสนเทศในองค์กร โดยเฉพาะในสถาบันบริการสารสนเทศจะต้องจำแนกได้ว่าองค์กรนั้นมีทรัพยากรสารสนเทศประเภทใดบ้าง ทั้งสื่อสิ่งพิมพ์และไม่ตีพิมพ์ เพื่อนำข้อมูลดังกล่าวมาใช้ในการจัดทำฐานข้อมูลต่อไป นอกจากนั้นยังต้องทำความเข้าใจเกี่ยวกับระบบสารสนเทศประเภทต่างๆ เพื่อให้รู้ว่าระบบนั้นๆ ต้องประกอบด้วยข้อมูลอะไรบ้าง โดยเฉพาะในระบบงานทางธุรกิจ ซึ่งมีฝ่ายต่างๆ อยู่มากมาย ซึ่งต้องอาศัยข้อมูลเพื่อใช้ในการจัดทำสารสนเทศที่นำไปใช้ในการกำหนดนโยบาย หรือวางแผนระบบงานต่อไป เมื่อทราบแล้วว่าข้อมูลที่ใช้ในการจัดการฐานข้อมูลประกอบด้วยอะไร จึงนำข้อมูลเหล่านั้นมาทำการออกแบบฐานข้อมูล ตามหลักการและวิธีที่ถูกต้อง เริ่มตั้งแต่การสร้างแบบจำลองอี-อาร์(E-R Model) ที่ประกอบไปด้วยเอนทิตี(entity) แอตทริบิวต์(attribute) คีย์(key) และความสัมพันธ์ (relationship) รวมทั้งการทำนอร์มัลไลเซชัน (normalization) เพื่อให้ฐานข้อมูลที่ออกแบบไว้เกิดความซ้ำซ้อนน้อยที่สุด และสามารถแก้ไขปรับปรุงข้อมูลได้ง่าย จากนั้นก็นำฐานข้อมูลที่ออกแบบไว้นั้นมาสร้างให้ใช้งานได้จริงด้วยโปรแกรมสำเร็จรูปในการจัดการฐานข้อมูล ซึ่งในเอกสารนี้จะยกตัวอย่างการออกแบบและสร้างฐานข้อมูลด้วยโปรแกรมไมโครซอฟต์แอกเซส (Microsoft Access) เนื่องจากเป็นโปรแกรมที่เข้าใจง่ายและเครื่องคอมพิวเตอร์ส่วนใหญ่จะมีโปรแกรมนี้อยู่แล้ว สำหรับกรณีศึกษาที่ใช้จะเป็นเรื่องที่ใกล้ตัวนักศึกษาที่สุดก็คือเรื่องระบบการลงทะเบียนเรียนของนักศึกษา ซึ่งจะเป็นการออกแบบขั้นพื้นฐาน เพื่อให้ง่ายต่อความเข้าใจ และสามารถนำไปประยุกต์ใช้ได้จริงในการทำงานต่อไป

1. ความหมายของข้อมูล

“ข้อมูล” (data) หมายถึง กลุ่มตัวอักขระที่เมื่อนำมารวมกันแล้วมีความหมายอย่างใดอย่างหนึ่งและมีความสำคัญควรค่าแก่การจัดเก็บเพื่อนำไปใช้ในโอกาสต่อๆ ไป ข้อมูลมักเป็นข้อความที่อธิบายถึงสิ่งใดสิ่งหนึ่ง อาจเป็นตัวอักษร ตัวเลข หรือสัญลักษณ์ใดๆ ที่สามารถนำไปประมวลผลด้วยคอมพิวเตอร์ได้ทั้งนี้ ข้อมูลสามารถทำให้ผู้อ่านทราบความเป็นไปของเหตุการณ์ที่เกิดขึ้น บอกสภาพการณ์หรือปรากฏการณ์ที่เกิดขึ้น และข้อมูลจะยังคงสภาพความเป็นข้อมูลไม่ว่าจะนำไปใช้หรือไม่ใช้ก็ตาม

สรุปได้ว่า “ข้อมูล” คือ ข้อเท็จจริงที่ทำให้ทราบถึงเหตุการณ์ต่างๆ ที่เกิดขึ้นภายในองค์กร บอกสภาพและปรากฏการณ์ที่เกิดขึ้นกับสิ่งแวดล้อม โดยอธิบายถึงสิ่งใดสิ่งหนึ่ง ซึ่งข้อมูลอาจจะเป็นตัวอักษร ตัวเลข ภาพและเสียง ที่สามารถนำมาประมวลผลด้วยคอมพิวเตอร์ได้

2. ชนิดและประเภทของข้อมูล

2.1 ชนิดของข้อมูล

ข้อมูลที่คอมพิวเตอร์สามารถเก็บบันทึกได้นั้นสามารถจำแนกได้หลายชนิดที่สำคัญๆได้แก่

2.1.1 ข้อมูลที่เป็นตัวอักขระหรือข้อความ (character data หรือ text) เป็นข้อมูลที่อาจเป็นตัวอักษร ตัวเลข และสัญลักษณ์ โดยข้อมูลที่เป็นตัวเลขจะไม่สามารถนำไปคำนวณได้แต่สามารถนำไปจัดเรียงตามลำดับตัวอักขระ เปลี่ยนแปลงแก้ไขได้ และเปรียบเทียบได้ว่าเหมือนกับข้อมูลอื่นหรือไม่

2.1.2 ข้อมูลเชิงจำนวน (numerical data) มีลักษณะเป็นตัวเลขซึ่งคอมพิวเตอร์สามารถนำไปคำนวณได้ ตัวอย่างเช่น เงินเดือน คะแนน และยอดขาย เป็นต้น โดยข้อมูลที่เป็นตัวเลขอาจจำแนกได้ 2 รูปแบบ คือ ข้อมูลที่เป็นเลขจำนวนเต็ม (integer number) เช่น 3, 5 และ-7 เป็นต้น และข้อมูลที่เป็นเลขทศนิยม เช่น 20.5, 125.75 และ -0.001 เป็นต้น

2.1.3 ข้อมูลรหัส (code data) อาจเป็นตัวอักขระหรือข้อมูลเชิงจำนวนก็ได้ซึ่งมักจะมีการกำหนดขนาดความยาวไว้จำกัด เช่น รหัสที่ใช้ระบุเพศอาจเป็นตัวเลขหนึ่งตัวหรือตัวอักขระหนึ่งตัว ได้แก่ 1 แทนด้วยเพศชาย และ 2 แทนด้วยเพศหญิง หรือรหัสนักศึกษาของมหาวิทยาลัยราชภัฏสวนดุสิตกำหนดให้เป็นตัวเลขจำนวน 11 ตัว เป็นต้น ซึ่งข้อมูลรหัสนั้นไม่ได้นำไปใช้เพื่อการคำนวณ แต่ใช้เพื่อการเปรียบเทียบ นับหรือจัดกลุ่มข้อมูลที่มีรหัสตรงกับที่กำหนด

2.1.4 ข้อมูลวันที่ (date data) เป็นข้อมูลซึ่งกำหนดขึ้นเป็นพิเศษเพื่อให้สามารถแสดง วัน เดือน ปี ที่กำหนดไว้เป็นมาตรฐานได้ ซึ่งข้อมูลวันที่จะมีประโยชน์มากในการคำนวณอายุ หรือหาช่วงเวลาระหว่างวันที่หนึ่งไปยังอีกวันที่หนึ่ง

2.1.5 ข้อมูลภาพลักษณ์(image data) เป็นข้อมูลภาพ เช่น ภาพถ่าย หรือภาพที่จัดเก็บไว้ในระบบคอมพิวเตอร์ด้วยเครื่องสแกนเนอร์เป็นข้อมูลที่แสดงความเข้มและสีของรูปภาพหรือเอกสารที่เป็นผลจากการใช้เครื่องสแกนเนอร์บันทึกข้อมูลไว้ในคอมพิวเตอร์ โดยข้อมูลภาพลักษณ์จะมีลักษณะเป็นจุดภาพ สามารถนำมาแสดงทางจอภาพ ย่อ ขยาย หรือตัดต่อได้ และมักเก็บไว้ใช้เพื่อการอ้างอิงเท่านั้น

2.1.6 ข้อมูลภาพเคลื่อนไหว (moving data) เป็นภาพที่เกิดจากการนำภาพนิ่งหลาย ๆ ภาพ ซึ่งแต่ละภาพมีความแตกต่างจากภาพก่อนหน้าเพียงเล็กน้อยมาแสดงต่อเนื่องกันโดยเร็ว ทำให้ผู้ชมเห็นเป็นภาพเคลื่อนไหว

2.1.7 ข้อมูลเสียง (voice data) เป็นข้อมูลที่นิยมนำมาใช้ในงานประมวลผลภาษาธรรมชาติ (natural language processing) ซึ่งช่วยให้คอมพิวเตอร์สามารถแยกเสียงที่ได้รับออกว่ามีความหมายอะไร เพื่อให้คอมพิวเตอร์สามารถทำงานตามที่มนุษย์บอกหรือออกคำสั่งได้

2.2 ประเภทของข้อมูล

ข้อมูลที่คอมพิวเตอร์สามารถจัดเก็บได้มีหลายรูปแบบที่กล่าวไปแล้ว ซึ่งข้อมูลแต่ละแบบก็มีหลายประเภท ที่สำคัญๆ และเกี่ยวข้องกับชีวิตของเรา ได้แก่ ข้อมูลเกี่ยวกับการศึกษาการบริหารและการจัดการ สาธารณสุข และสิ่งแวดล้อม เป็นต้น

การจัดการข้อมูลด้วยแฟ้มข้อมูล

การจัดการข้อมูลเริ่มจากการบันทึกข้อมูล ซึ่งอาจจะเป็นการบันทึกข้อมูลลงกระดาษหรือสมุด เพื่อช่วยในการจดจำ แต่เมื่อมีข้อมูลเพิ่มขึ้น รูปแบบของการจัดเก็บข้อมูลก็เปลี่ยนไปให้มีระบบระเบียบมากขึ้น โดยมีการบันทึกข้อมูลลงแฟ้มเอกสารต่างๆ ที่จัดไว้เป็นหมวดหมู่ และมีการพัฒนาการจัดเก็บข้อมูลลงในสื่อบันทึกข้อมูลด้วยคอมพิวเตอร์ โดยจัดเก็บเป็นแฟ้มข้อมูล เพื่อให้เครื่องคอมพิวเตอร์สามารถอ่านข้อมูลจากแฟ้มนั้นได้

1. ความหมายของแฟ้มข้อมูล

“แฟ้มข้อมูล” (file) หมายถึง ข้อสนเทศหรือข้อมูลทั้งหมดที่เก็บไว้ในสื่อที่มีคุณสมบัติเป็นแม่เหล็กไม่ว่าจะเป็นจานบันทึกธรรมดาหรือจานแข็ง (hard disk) ก็ตาม ข้อสนเทศที่นำไปเก็บนั้นจะถูกนำไปเก็บไว้เป็นเรื่องๆ ไป อาจจะเป็นโปรแกรม ข้อมูล หรือภาพ (graphics) ก็ได้ แต่ละเรื่องต่างก็ต้องมีชื่อเป็นของตนเอง ที่ต้องไม่ซ้ำกัน

สรุปได้ว่า “แฟ้มข้อมูล” คือ ข้อมูลที่ผู้ใช้ต้องการจะจัดเก็บหรือรวบรวม ซึ่งบันทึกไว้ในหน่วยเก็บข้อมูลในคอมพิวเตอร์ เช่น แฟ้มข้อมูลรายงาน แฟ้มข้อมูลรูปภาพ และแฟ้มข้อมูลโปรแกรมต่างๆ เป็นต้น

2. โครงสร้างแฟ้มข้อมูล

“โครงสร้างแฟ้มข้อมูล” (data structure) หมายถึง รูปแบบของการจัดระเบียบของข้อมูล ซึ่งมีอยู่หลายรูปแบบ ประกอบด้วยโครงสร้างพื้นฐานที่ลำดับจากหน่วยที่เล็กที่สุดไปยังหน่วยที่ใหญ่ขึ้นตามลำดับต่อไปนี้

2.1บิท (bit : binary digit) คือ หน่วยของข้อมูลที่เล็กที่สุดที่เก็บอยู่ในหน่วยความจำภายในคอมพิวเตอร์ ซึ่ง Bit จะแทนด้วยตัวเลขหนึ่งตัว คือ 0 หรือ 1 อย่างใดอย่างหนึ่ง เรียกตัวเลข 0 หรือ 1 ว่าเป็นบิท 1 บิท

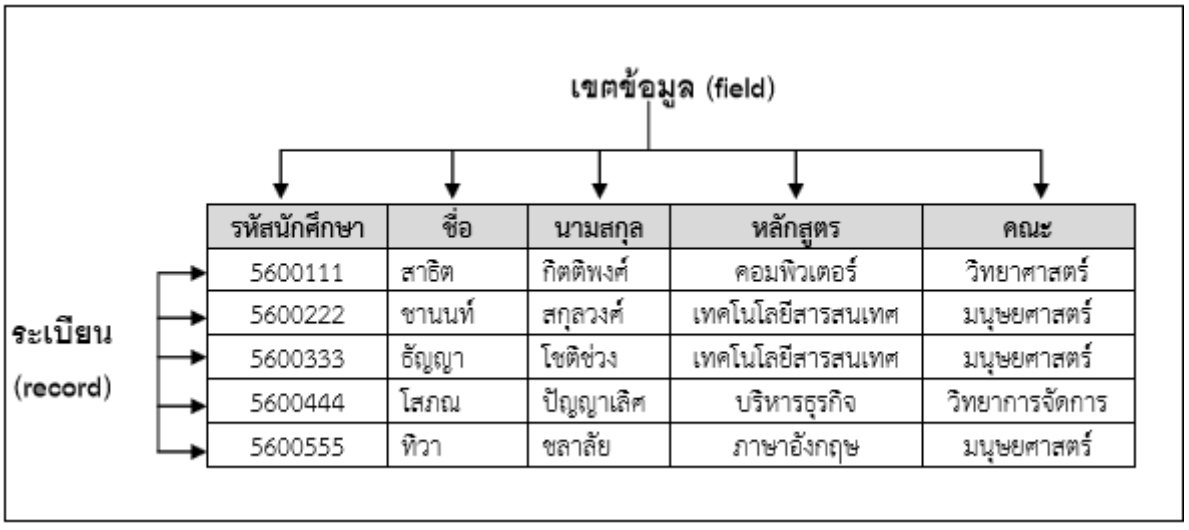
2.2ไบท์ (byte) คือ หน่วยของข้อมูลที่นำบิทหลายๆ บิทมารวมกัน แทนตัวอักษรแต่ละตัว เช่น A, B, ..., Z, 0, 1, 2, ... ,9 และสัญลักษณ์พิเศษอื่นๆ เช่น $, &, +, -, *, / ฯลฯ โดยตัวอักษร 1 ตัวจะแทนด้วยบิท 7 หรือ 8 บิท (1 ไบท์แทนด้วยตัวอักษร 7 หรือ 8 บิท) ซึ่งตัวอักษรแต่ละตัวจะเรียกว่า ไบท์ เช่น ตัว A เมื่อเก็บอยู่ในคอมพิวเตอร์จะเก็บเป็น 1000001 ส่วนตัว B จะเก็บเป็น 1000010 เป็นต้น

2.3เขตข้อมูล (field) คือ หน่วยของข้อมูลที่เกิดจากการนำไบท์หรือตัวอักขระหลายๆตัวมารวมกันเป็นคำที่มีความหมาย เช่น รหัสนักศึกษา ชื่อนักศึกษา นามสกุล ที่อยู่ คณะ และสาขาวิชา เป็นต้น

2.4ระเบียน (record) คือ หน่วยของข้อมูลที่มีการนำเขตข้อมูลหลายๆ เขตข้อมูล ที่มีความสัมพันธ์กันมารวมกัน หรือค่าของข้อมูลในแต่ละเขตข้อมูล เช่น ระเบียนนักศึกษาคนที่ 1 ประกอบด้วยเขตข้อมูล รหัสนักศึกษา :5600111, ชื่อ : สาธิต, นามสกุล : กิตติพงศ์, หลักสูตร :คอมพิวเตอร์, คณะ : วิทยาศาสตร์ เป็นต้น

2.5แฟ้มข้อมูล (file) คือ หน่วยของข้อมูลที่มีการนำระเบียนหลายๆ ระเบียนที่มีความสัมพันธ์กันมารวมกัน เช่น แฟ้มข้อมูลนักศึกษา ซึ่งประกอบไปด้วย ระเบียนจำนวน 5 ระเบียน หรือ 5 แถว ซึ่งก็คือรายละเอียดของนักศึกษาจำนวน 5 คน นั่นเอง

2.6ฐานข้อมูล (database) คือ หน่วยของข้อมูลที่มีการนำแฟ้มข้อมูลหลายๆแฟ้มข้อมูล ที่มีความสัมพันธ์กันมารวมกัน เช่น ฐานข้อมูลในระบบทะเบียนนักศึกษา จะประกอบด้วยแฟ้มข้อมูลรายวิชา นักศึกษา การลงทะเบียน ผลการเรียน และอาจารย์ผู้สอน เป็นต้น



3. ประเภทของแฟ้มข้อมูล

แฟ้มข้อมูลจะถูกแบ่งแยกประเภทตามการใช้งาน ซึ่งแบ่งออกเป็น 5 ประเภทต่างๆดังนี้

3.1แฟ้มข้อมูลรายการหลัก (master file) ทำหน้าที่จัดเก็บข้อมูลที่ไม่มีการเปลี่ยนแปลงหรือมีสภาพค่อนข้างคงที่ เช่น แฟ้มข้อมูลประวัตินักศึกษา จะประกอบด้วยข้อมูลต่างๆ เช่นรหัสนักศึกษา ชื่อ-นามสกุล ที่อยู่ คณะ และหลักสูตร เป็นต้น ซึ่งการปรับปรุงแก้ไขข้อมูลในแฟ้มข้อมูลรายการหลัก ให้ทันสมัยสามารถทำได้ 3 รูปแบบคือ การเพิ่ม (add) การลบออก (delete) และการแก้ไข (modify) เช่น การเพิ่มระเบียนของ

นักศึกษาในกรณีที่เป็นนักศึกษาใหม่ การลบเขตข้อมูล (field) ระเบียน (record) ระเบียนของนักศึกษาในกรณีที่นักศึกษาลาออก และการเปลี่ยนแปลงที่อยู่ของนักศึกษา เป็นต้น อีกตัวอย่างหนึ่งที่เป็นแฟ้มข้อมูลรายการหลัก ก็คือ ข้อมูลของลูกค้าธนาคาร เช่น เลขที่บัญชี ชื่อ-นามสกุล ที่อยู่ ยอดเงินคงเหลือ ซึ่งจะถูกแก้ไขเมื่อมีรายการฝากถอนเงินจากลูกค้า โดยการแก้ไขแฟ้มข้อมูลอาจทำได้โดยตรงหรือแก้ไขโดยใช้ข้อมูลจากแฟ้มข้อมูลรายการเปลี่ยนแปลงซึ่งจะกล่าวในข้อต่อไป

3.2แฟ้มข้อมูลรายการเปลี่ยนแปลง (transaction file) ทำหน้าที่จัดเก็บข้อมูลที่มักมีการเคลื่อนไหวหรือมีการเปลี่ยนแปลงอยู่เสมอ เช่น แฟ้มข้อมูลการลงทะเบียนเรียนของนักศึกษาที่จะต้องมีการลงทะเบียนเรียนในทุกๆ ภาคการศึกษา แฟ้มข้อมูลรายการฝาก-ถอนเงินในบัญชีลูกค้าธนาคาร หรือแฟ้มข้อมูลการขายสินค้าประจำวัน เป็นต้น

3.3 แฟ้มข้อมูลรายงาน (report file) ทำหน้าที่เก็บรายงานที่ได้จากคอมพิวเตอร์ไว้เนื่องจากการเก็บแฟ้มข้อมูลรายงานไว้ในรูปของแฟ้มข้อมูลในหน่วยความจำสำรอง มีข้อดีคือจัดเก็บได้สะดวกและทนทานกว่าการเก็บเป็นกระดาษ อีกทั้งสามารถสั่งพิมพ์เมื่อใดและปริมาณเท่าใดก็ได้

3.4แฟ้มข้อมูลเก็บผลลัพธ์ (output file) โปรแกรมส่วนมากจะมีการรับข้อมูลเข้ามาประมวลผล และได้ผลลัพธ์เป็นข้อมูลใหม่ออกมา ข้อมูลใหม่อาจแสดงออกทางหน่วยแสดงผลหรือจัดเก็บไว้ในแฟ้มข้อมูลก็ได้ เรียกแฟ้มข้อมูลที่เก็บข้อมูลใหม่นี้ว่า แฟ้มข้อมูลเก็บผลลัพธ์ และสามารถนำแฟ้มข้อมูลนี้ไปเป็นข้อมูลนำเข้าของโปรแกรมอื่นได้ต่อไป

3.5แฟ้มข้อมูลสำรอง (backup file) ใช้เก็บสำรองข้อมูลในแฟ้มข้อมูลที่มีความสำคัญสูง การสำรองข้อมูลเป็นสิ่งสำคัญมากในการใช้งานคอมพิวเตอร์ เนื่องจากสื่อที่เก็บข้อมูลต่างๆ อาจเกิดปัญหาได้โดยที่ผู้ใช้คาดไม่ถึง ดังนั้นเราควรจัดเก็บข้อมูลลงบนสื่อบันทึกข้อมูลอื่นด้วย เพื่อเป็นการสำรองข้อมูลในกรณีที่มีปัญหาซึ่งจะช่วยให้ได้ข้อมูลที่สำคัญกลับมาใช้ใหม่

4. ข้อดีของการจัดการข้อมูลด้วยแฟ้มข้อมูล

4.1 การประมวลผลข้อมูลได้รวดเร็ว เนื่องจากมีการแยกข้อมูลไว้เป็นแฟ้มต่างๆ

4.2 ลงทุนต่ำในเบื้องต้น อาจไม่จำเป็นต้องใช้คอมพิวเตอร์ที่มีความสามารถมาก ก็สามารถทำการประมวลผลข้อมูลได้

4.3สามารถออกแบบแฟ้มข้อมูลและทำการพัฒนาได้ง่าย เนื่องจากมีขั้นตอนไม่สลับซับซ้อนมากนัก

5.ข้อเสียของการจัดการข้อมูลด้วยแฟ้มข้อมูล

ในการจัดการข้อมูลด้วยแฟ้มข้อมูลที่เป็นอิสระและกระจัดกระจายนั้น พบว่าแต่ละแผนกแต่ละงานมีแฟ้มข้อมูลของตัวเองแยกเก็บในแต่ละแผนก เมื่อพิจารณาแล้วพบว่าอาจมีความคล่องตัวสูงแต่ในขณะเดียวกันอาจเกิดปัญหาต่างๆ ขึ้น ดังนี้

5.1เกิดความซ้ำซ้อนของข้อมูล(data redundancy) เนื่องจากแต่ละฝ่ายมีแฟ้มข้อมูลของตนเอง คือ ข้อมูลชุดเดียวกันมีการจัดเก็บในแฟ้มข้อมูลที่ต่างกัน หรือข้อมูลชุดเดียวกันถูกจัดเก็บอยู่ในสองแฟ้มข้อมูลหรือมากกว่า ซึ่งจะทำให้เป็นการสิ้นเปลืองเนื้อที่และแรงงานในการจัดเก็บข้อมูลที่ซ้ำซ้อนนั้น

5.2ลำบากต่อการแก้ไข (updating difficulties) ความซ้ำซ้อนของข้อมูลจะทำให้ยากต่อการแก้ไขข้อมูลเหล่านั้น เนื่องจากถ้ามีข้อมูลใดเปลี่ยนแปลงจะต้องทำการเปลี่ยนแปลงข้อมูลทุกแฟ้มข้อมูลที่มีข้อมูลซ้ำกันทั้งหมด ทำให้อาจเกิดข้อผิดพลาดได้และเกิดความสับสนหากข้อมูลในแต่ละแฟ้มข้อมูลไม่ตรงกัน รวมทั้งสิ้นเปลืองแรงงานในการเปลี่ยนแปลงแก้ไขข้อมูลที่ซ้ำซ้อนนั้นด้วย

5.3 เกิดความขัดแย้งของข้อมูล (data inconsistency) เป็นปัญหาที่เกิดจากการจัดเก็บข้อมูลที่ซ้ำซ้อน เนื่องจากการจัดเก็บข้อมูลชุดเดียวกันในหลายแฟ้มข้อมูล อาจทำให้ข้อมูลชุดเดียวกันมีค่าที่แตกต่างกันได้ในแต่ละแฟ้มข้อมูล ถ้ามีการแก้ไขปรับปรุงข้อมูลไม่ครบถ้วน ซึ่งทำให้ไม่ทราบว่าข้อมูลชุดใดคือข้อมูลที่ถูกต้องที่สุด

5.4เกิดการผูกติดกับข้อมูล (data dependence) เมื่อโปรแกรมได้ถูกพัฒนาสำหรับใช้กับแฟ้มข้อมูลใดโดยเฉพาะ จะทำให้เกิดการผูกติดกันกับรูปแบบของข้อมูล กล่าวคือ ถ้าโปรแกรมหรือรูปแบบข้อมูลฝ่ายหนึ่งฝ่ายใดเกิดการเปลี่ยนแปลง จะทำให้อีกฝ่ายต้องมีการแก้ไขด้วย ซึ่งเป็นการเพิ่มต้นทุนในการพัฒนาโปรแกรม

5.5การกระจัดกระจายของข้อมูล (data dispersion) ถ้าข้อมูลถูกจัดเก็บอยู่ในแหล่งต่าง ๆ อย่างกระจัดกระจายและไม่เป็นระบบ โดยมีโครงสร้างและรูปแบบของข้อมูลผูกติดอยู่กับโปรแกรมที่ใช้งานข้อมูลเหล่านั้น จะทำให้เกิดความยากในการใช้ข้อมูลร่วมกันของโปรแกรมอื่นเนื่องจากต้องมีการพัฒนาโปรแกรมใหม่เพื่อให้ใช้รูปแบบข้อมูลที่แตกต่างกันได้

5.6 การใช้ประโยชน์จากข้อมูลลดลง (underutilization of data) เนื่องจากต้องมีการพัฒนาโปรแกรมโดยเฉพาะสำหรับการใช้งานข้อมูลเพื่องานนั้นๆ โดยไม่สามารถใช้โปรแกรมเดิมที่ใช้งานอยู่ได้ ทำให้ผู้ใช้เข้าถึงข้อมูลเหล่านั้นได้ยาก และปฏิเสธที่จะใช้งานในที่สุด

จากข้อเสียดังกล่าวของการจัดการข้อมูลด้วยแฟ้มข้อมูล จึงเป็นที่มาของการพัฒนาระบบการจัดการข้อมูลอีกรูปแบบหนึ่ง เพื่อแก้ปัญหาที่เกิดขึ้นของการจัดการข้อมูลในระบบแฟ้มข้อมูล ซึ่งเรียกว่าระบบการจัดการฐานข้อมูล

การจัดการข้อมูลด้วยระบบการจัดการฐานข้อมูล

การจัดเก็บข้อมูลที่มีปริมาณมากๆ ในระบบคอมพิวเตอร์โดยการเก็บไว้ในแฟ้มข้อมูลแต่ละแฟ้ม อาจเกิดปัญหาข้อมูลซ้ำซ้อนกัน เมื่อมีการแก้ไขข้อมูลที่มีอยู่หลายแฟ้มข้อมูล จนทำให้ข้อมูลมีความขัดแย้งกันเอง จึงได้มีการเปลี่ยนการจัดเก็บข้อมูลให้อยู่ในรูปของฐานข้อมูลแทนเพื่อความสะดวกในการบันทึกข้อมูล แก้ไขข้อมูล และค้นหาข้อมูล

ความหมายของฐานข้อมูลและระบบการจัดการฐานข้อมูล

"ฐานข้อมูล" (database) หมายถึง การจัดรวบรวมข้อนิเทศหรือข้อมูลของเรื่องต่างๆไว้ในรูปแบบที่จะเรียกมาใช้ได้ทันทีเมื่อต้องการ ในการเรียกนั้น อาจเรียกเพียงส่วนใดส่วนหนึ่งมาใช้ประโยชน์เป็นครั้งเป็นคราวก็ได้ ฐานข้อมูลที่ดีควรจะได้รับการปรับให้ทันสมัยอยู่เสมอ

สรุปได้ว่า "ฐานข้อมูล" คือ การรวบรวมข้อมูลที่เราต้องการจะจัดเก็บ ซึ่งต้องมีความสัมพันธ์กันหรือเป็นเรื่องเดียวกันไว้ด้วยกัน เพื่อสะดวกในการใช้งาน

"ระบบการจัดการฐานข้อมูล" (Data Base Management System: DBMS) หมายถึง ซอฟต์แวร์ที่สร้างขึ้นเพื่อรวบรวมข้อมูลให้เป็นระบบ เพื่อจะได้นำไปเก็บรักษา เรียกใช้หรือนำมาปรับปรุงให้ทันสมัยได้ง่าย ทั้งนี้จำเป็นต้องคำนึงถึงการรักษาความปลอดภัยของข้อมูลเป็นเรื่องสำคัญด้วย

สรุปได้ว่า "ระบบการจัดการฐานข้อมูล" คือ โปรแกรมที่ทำหน้าที่ในการกำหนดลักษณะข้อมูลที่จะเก็บไว้ในฐานข้อมูล อำนวยความสะดวกในการบันทึกข้อมูลลงในฐานข้อมูล กำหนดผู้ที่ได้รับอนุญาตให้ใช้ฐานข้อมูลได้ พร้อมกับกำหนดด้วยว่าให้ใช้ได้แบบใด เช่น ให้อ่านข้อมูลได้อย่างเดียวหรือให้แก้ไขข้อมูลได้ด้วย นอกจากนั้นยังอำนวยความสะดวกในการค้นหาข้อมูล และการแก้ไขปรับปรุงข้อมูล ทำให้ผู้ใช้สามารถเข้าถึงข้อมูลได้ง่าย สะดวกและมีประสิทธิภาพ เสมือนเป็นตัวกลางระหว่างผู้ใช้กับฐานข้อมูลให้สามารถติดต่อกันได้

ประวัติความเป็นมาของระบบการจัดการฐานข้อมูล

การจัดการฐานข้อมูลเริ่มต้นจากการที่องค์การบริหารการบินและอวกาศสหรัฐอเมริกาหรือนาซาได้ว่าจ้างบริษัทไอบีเอ็ม (IBM) ประเทศสหรัฐอเมริกา ให้ออกแบบระบบเก็บรวบรวมข้อมูลที่ได้จากการสำรวจดวงจันทร์ในโครงการอะพอลโล (โครงการอะพอลโลเป็นโครงการสำรวจอวกาศอย่างจริงจัง และมีการส่งมนุษย์ขึ้นบนดวงจันทร์ได้สำเร็จด้วยยานอะพอลโล 11) ได้พัฒนาระบบการดูแลข้อมูลเรียกว่า ระบบ GUAM (Generalized Upgrade Access Method) ซึ่งถือเป็นต้นกำเนิดของระบบการจัดการฐานข้อมูลต่อมาบริษัทไอบีเอ็ม ได้พัฒนาระบบการจัดการฐานข้อมูลขึ้นมาใหม่เพื่อให้ใช้งานกับธุรกิจทั่วๆ ไปได้ เรียกว่า DL/I (Data Language/I) จนในที่สุดก็ได้กลายมาเป็นระบบ IMS (Information Management System) ในช่วงปี พ.ศ. 2525 มีการนำระบบฐานข้อมูลเข้ามาใช้กับคอมพิวเตอร์อย่างเต็มที่ ได้มีการคิดค้นและผลิตซอฟต์แวร์เกี่ยวกับฐานข้อมูลออกมามากมาย การเจริญเติบโตของการจัดการฐานข้อมูลรุดหน้าไปอย่างรวดเร็วพร้อมกับระบบคอมพิวเตอร์และมีการพัฒนามาจนถึงทุกวันนี้ปัจจุบันได้มีการนำคอมพิวเตอร์มาใช้ในการเก็บข้อมูล โดยใช้โปรแกรมสำเร็จรูปทั่วไปโดยที่ผู้ใช้ไม่ต้องเขียนโปรแกรมเอง เพียงแต่เรียนรู้คำสั่งการเรียกใช้ข้อมูลหรือการจัดการข้อมูล เช่นการป้อนข้อมูล การบันทึกข้อมูล การแก้ไขและเปลี่ยนแปลงข้อมูล เป็นต้นในอดีตยุคที่มีไมโครคอมพิวเตอร์เกิดขึ้นแรกๆ โปรแกรมสำเร็จรูปทางด้านการจัดการฐานข้อมูลที่นิยมใช้กันอย่างแพร่หลาย คือ Personal Filling System : PFS ต่อมาได้มีโปรแกรมฐานข้อมูลเพิ่มขึ้นหลายโปรแกรม เช่น Datastar DB Master และ dBASE II เป็นต้นโดยเฉพาะโปรแกรม dBASE II ได้รับความนิยมมาก จนกระทั่งในปีพ.ศ. 2528 ผู้ผลิตได้สร้าง dBASE III Plus ออกมาซึ่งสามารถจัดการฐานข้อมูลแบบสัมพันธ์ (relational) เชื่อมโยงแฟ้มข้อมูลต่างๆ เข้าด้วยกันค้นหา และนำมาสร้างเป็นรายงานตามความต้องการได้สะดวก รวดเร็ว ต่อมาได้มีการสร้างโปรแกรมสำเร็จรูปเกี่ยวกับฐานข้อมูลออกมา เช่น ไมโครซอฟต์แอกเซส (Microsoft Access) ออราเคิล (Oracle) ดีเบส (dBase) ฟอกซ์โปร (FoxPro) อินฟอร์มิกซ์ (Informix) มายเอสคิวแอล (MySQL) และไซเบส (Sybase) เป็นต้น

หน้าที่ของระบบการจัดการฐานข้อมูล

ระบบการจัดการฐานข้อมูลมีหน้าที่สำคัญๆ หลายอย่าง เพื่อให้เกิดความถูกต้องและสอดคล้องกันของข้อมูลภายในฐานข้อมูล ได้แก่

1. การจัดการพจนานุกรมข้อมูล ระบบการจัดการฐานข้อมูลจะทำการจัดเก็บนิยามของข้อมูล และความสัมพันธ์ระหว่างข้อมูลไว้ในพจนานุกรมข้อมูล เป็นสารสนเทศที่บอกเกี่ยวกับโครงสร้างของฐานข้อมูล โปรแกรมประยุกต์ทั้งหมดที่ต้องการเข้าถึงข้อมูลในฐานข้อมูลจะต้องทำงานผ่านระบบการจัดการฐานข้อมูล โดยที่ระบบจัดการฐานข้อมูลจะใช้พจนานุกรมข้อมูล เพื่อค้นหาโครงสร้างตลอดจนส่วนประกอบของข้อมูลและความสัมพันธ์ที่ต้องการ นอกจากนั้นแล้วการเปลี่ยนแปลงใดๆ ที่มีต่อโครงสร้างฐานข้อมูลจะถูกบันทึกไว้โดยอัตโนมัติในพจนานุกรมข้อมูล ทำให้เราไม่ต้องเปลี่ยนแปลงแก้ไขโปรแกรมเมื่อโครงสร้างข้อมูลมีการเปลี่ยนแปลง

2. การจัดเก็บข้อมูล ระบบการจัดการฐานข้อมูลจะสร้างโครงสร้างที่จำเป็นต่อการจัดเก็บข้อมูล ช่วยลดความยุ่งยากในการนิยามและการเขียนโปรแกรมที่เกี่ยวข้องกับคุณสมบัติทางกายภาพของข้อมูล ระบบการจัดการฐานข้อมูลในปัจจุบันไม่เพียงแต่จะช่วยในการจัดเก็บข้อมูลเท่านั้น แต่ยังรวมถึงการจัดเก็บกฎเกณฑ์ต่างๆ ที่ใช้ในการตรวจสอบบูรณภาพของข้อมูลอีกด้วย

3. การแปลงและนำเสนอข้อมูล ระบบการจัดการฐานข้อมูลจะทำหน้าที่ในการแปลงข้อมูลที่ได้รับเข้ามา เพื่อให้สอดคล้องกับโครงสร้างในการจัดเก็บข้อมูล ทำให้เราไม่ต้องไปยุ่งเกี่ยวกับความแตกต่างระหว่างรูปแบบของข้อมูลทางตรรกะและทางกายภาพ กล่าวคือทำให้มีความเป็นอิสระของข้อมูล ระบบการจัดการฐานข้อมูลจะแปลงความต้องการเชิงตรรกะของผู้ใช้ ให้เป็นคำสั่งที่สามารถดึงข้อมูลทางกายภาพที่ต้องการ

4.การจัดการระบบความปลอดภัยของข้อมูล ระบบการจัดการฐานข้อมูลจะสร้างระบบรักษาความปลอดภัยของข้อมูล โดยการกำหนดรายชื่อผู้มีสิทธิ์เข้าใช้ระบบ และความสามารถในการใช้ระบบ เช่น การอ่าน เพิ่ม ลบ หรือแก้ไขเปลี่ยนแปลงข้อมูล การจัดการระบบความปลอดภัยของข้อมูลมีความสำคัญมากในระบบฐานข้อมูลแบบที่มีผู้ใช้หลายคน

5. การควบคุมการเข้าถึงข้อมูลของผู้ใช้หลายคน ระบบการจัดการฐานข้อมูลจะใช้หลักการออกแบบโปรแกรมที่เหมาะสม เพื่อให้แน่ใจว่าผู้ใช้หลายคนสามารถเข้าใช้ฐานข้อมูลพร้อมกันได้และข้อมูลมีความถูกต้อง

6. การเก็บสำรองและกู้คืนข้อมูล ระบบการจัดการฐานข้อมูลจะมีโปรแกรมเพื่อสนับสนุนการสำรองและกู้คืนข้อมูล เพื่อให้แน่ใจด้านความปลอดภัยและความมั่นคงของข้อมูลในระบบ ระบบการจัดการฐานข้อมูลจะกู้ข้อมูลในฐานข้อมูลคืนมาหลังจากระบบเกิดความล้มเหลว เช่นเมื่อเกิดกระแสไฟฟ้าขัดข้อง เป็นต้น

7. การควบคุมความถูกต้องของข้อมูล ระบบการจัดการฐานข้อมูลจะสนับสนุนและควบคุมความถูกต้องของข้อมูล ตั้งแต่ลดความซ้ำซ้อนของข้อมูล ไปจนถึงความไม่สอดคล้องกันของข้อมูล ความสัมพันธ์ของข้อมูลที่เก็บไว้ในพจนานุกรมข้อมูลจะถูกนำมาใช้ในการควบคุมความถูกต้องของข้อมูลด้วย

8. ภาษาที่ใช้ในการเข้าถึงฐานข้อมูลและการเชื่อมต่อกับโปรแกรมประยุกต์ ระบบการจัดการฐานข้อมูลสนับสนุนการเข้าถึงข้อมูลโดยผ่านภาษาคิวรี (query language) ซึ่งเป็นคำสั่งที่ใช้ในการค้นคืนข้อมูล

จากฐานข้อมูล โดยผู้ใช้เพียงบอกว่าต้องการอะไร และไม่จำเป็นต้องรู้ว่ามีขั้นตอนอย่างไรในการนำข้อมูลออกมา เพราะระบบการจัดการฐานข้อมูลจะเป็นผู้กำหนดวิธีการในการเข้าถึงข้อมูลอย่างมีประสิทธิภาพเอง

9. การติดต่อสื่อสารกับฐานข้อมูล ระบบการจัดการฐานข้อมูลที่ทันสมัยจะต้องสนับสนุนการใช้งานฐานข้อมูลผ่านทางเครือข่ายอินเทอร์เน็ตได้

<u>ข้อดีของการใช้ฐานข้อมูล</u>

เมื่อมีการนำระบบการจัดการฐานข้อมูลมาใช้ เพื่ออำนวยความสะดวกในการบันทึกข้อมูล แก้ไขปรับปรุงข้อมูล ค้นหาข้อมูล รวมทั้งกำหนดผู้ที่ได้รับอนุญาตให้ใช้ฐานข้อมูล เป็นต้น ทำให้ฐานข้อมูลมีข้อดีมากมาย ได้แก่

1. ลดความจำเจของงานดูแลเอกสาร ซึ่งเป็นงานประจำที่ทำให้ผู้ดูแลรู้สึกเบื่อหน่ายและขาดแรงจูงใจ แต่เราสามารถใช้คอมพิวเตอร์ในการปฏิบัติงานนี้แทนมนุษย์ได้โดยผ่านโปรแกรมสำหรับการจัดการฐานข้อมูล

2. ข้อมูลที่จัดเก็บมีความทันสมัย เมื่อข้อมูลในระบบฐานข้อมูลได้รับการดูแลปรับปรุงอย่างต่อเนื่อง ทำให้ข้อมูลที่จัดเก็บเป็นข้อมูลที่มีความทันสมัย ตรงกับเหตุการณ์ในปัจจุบัน และตรงกับความต้องการอยู่เสมอ

3. ลดความซ้ำซ้อนในการจัดเก็บข้อมูล เนื่องจากการจัดทำฐานข้อมูลจะมีการรวบรวมข้อมูลประเภทต่างๆ เข้ามาจัดเก็บไว้ในระบบและเก็บข้อมูลเพียงชุดเดียว ซึ่งทุกฝ่ายที่เกี่ยวข้องจะสามารถเรียกใช้ข้อมูลที่ต้องการได้ เป็นการประหยัดเนื้อที่ในการจัดเก็บ และทำให้เกิดความรวดเร็วในการค้นหาและจัดเก็บข้อมูลด้วย

4. หลีกเลี่ยงความขัดแย้งของข้อมูลได้เมื่อข้อมูลถูกจัดเก็บในระบบฐานข้อมูล จะทำให้ข้อมูลลดความซ้ำซ้อนลง คือ มีข้อมูลแต่ละประเภทเพียงหนึ่งชุดในระบบ ทำให้ข้อมูลที่เก็บได้ไม่ขัดแย้งกันเอง ในกรณีที่จำเป็นต้องเก็บข้อมูลที่ซ้ำซ้อนกัน เพื่อสาเหตุบางประการ เช่น เพื่อความรวดเร็วในการประมวลผลข้อมูล ระบบจัดการฐานข้อมูลจะเป็นผู้ดูแลข้อมูลที่ซ้ำกันให้มีความถูกต้องตรงกัน

5.ใช้ข้อมูลร่วมกันได้เนื่องจากระบบการจัดการฐานข้อมูลสามารถจัดให้ผู้ใช้แต่ละคนเข้าใช้ข้อมูลในแฟ้มที่มีข้อมูลเดียวกันได้ในเวลาเดียวกัน เช่น ฝ่ายบุคคลและฝ่ายการเงิน สามารถที่จะใช้ข้อมูลจากแฟ้มประวัติพนักงานในระบบฐานข้อมูลได้พร้อมกัน

6.ควบคุมมาตรฐานของข้อมูลได้ เมื่อข้อมูลต่างๆ ในหน่วยงานถูกรวบรวมเข้ามาผู้บริหารระบบฐานข้อมูลสามารถที่จะวางมาตรฐานในการรับข้อมูล แสดงผลข้อมูล ตลอดจนการจัดเก็บข้อมูลได้ เช่น การกำหนดรูปแบบของตัวเลขให้มีทศนิยม 2 ตำแหน่งสำหรับค่าที่เป็นตัวเงินการกำหนดรูปแบบของการรับ และแสดงผลสำหรับข้อมูลที่เป็นวันที่ นอกจากนี้การที่ข้อมูลมีมาตรฐานเดียวกัน ทำให้สามารถแลกเปลี่ยนข้อมูลระหว่างระบบได้อย่างสะดวก

7. จัดทำระบบการรักษาความปลอดภัยของข้อมูลได้ ผู้บริหารระบบฐานข้อมูลสามารถกำหนดรหัสผ่านเข้าใช้งานข้อมูลของผู้ใช้แต่ละราย และให้ผู้ใช้แต่ละรายมีสิทธิ์ในการทำงานกับข้อมูลไม่เท่าเทียมกันได้ โดยระบบการจัดการฐานข้อมูลจะทำการตรวจสอบสิทธิ์ในการทำงานกับข้อมูลทุกครั้ง เช่น การตรวจสอบสิทธิ์ในการเรียกดูข้อมูล การลบข้อมูล การปรับปรุงข้อมูล และการเพิ่มข้อมูลในแต่ละแฟ้มข้อมูล

8.ควบคุมความถูกต้องของข้อมูลได้ ปัญหาเรื่องความขัดแย้งกันของข้อมูลที่มีความซับซ้อน เป็นปัญหาหนึ่งในเรื่องความถูกต้องของข้อมูล ซึ่งเมื่อได้มีการกำจัดความซับซ้อนของข้อมูลออก ปัญหาเรื่องความถูกต้องของข้อมูลที่อาจเกิดขึ้นได้ เช่น อายุโดยปกติของคนงาน ควรอยู่ระหว่าง 18 – 60 ปี ถ้าหากในระบบฐานข้อมูลปรากฏมีพนักงานที่มีอายุ 150 ปีซึ่งเป็นไปไม่ได้ในทางปฏิบัติที่หน่วยงานจะมีการว่าจ้างคนงานที่มีอายุเกิน 60 ปี และอายุของคนในปัจจุบันไม่ควรเกิน 100 ปีผู้บริหารระบบฐานข้อมูลสามารถกำหนดกฎเกณฑ์ในการนำเข้าข้อมูล และระบบจัดการฐานข้อมูลจะคอยควบคุมให้มีการนำเข้าข้อมูล เป็นไปตามกฎเกณฑ์ให้มีความถูกต้อง

ข้อเสียของการใช้ฐานข้อมูล

แม้ว่าการประมวลผลข้อมูลด้วยระบบการจัดการจัดการฐานข้อมูล จะมีข้อดีหลายประการ แต่ก็จะมีข้อเสียอยู่บ้างดังต่อไปนี้

6.1 เสียค่าใช้จ่ายสูง เนื่องจากราคาของโปรแกรมที่ใช้ในระบบการจัดการฐานข้อมูลจะมีราคาค่อนข้างแพง รวมทั้งเครื่องคอมพิวเตอร์ที่มีประสิทธิภาพสูง คือ ต้องมีความเร็วสูง มีขนาดหน่วยความจำและหน่วยเก็บข้อมูลสำรองที่มีความจุมาก ทำให้ต้องเสียค่าใช้จ่ายสูงในการจัดทำระบบการจัดการฐานข้อมูล

6.2 เกิดการสูญเสียข้อมูลได้เนื่องจากข้อมูลต่างๆ ภายในฐานข้อมูลจะถูกจัดเก็บอยู่ในที่เดียวกัน ดังนั้นถ้าที่เก็บข้อมูลเกิดมีปัญหา อาจทำให้ต้องสูญเสียข้อมูลทั้งหมดในฐานข้อมูลได้ ดังนั้นการจัดทำฐานข้อมูลที่ดีจึงต้องมีการสำรองข้อมูลไว้เสมอ

กรอบแนวคิดและทฤษฎีเกี่ยวกับฐานข้อมูล

กรอบแนวคิดและทฤษฎีเกี่ยวกับฐานข้อมูลนำไปสู่การพัฒนาวิธีการออกแบบฐานข้อมูลให้เหมาะสมกับหน่วยงานและองค์กรต่างๆ บุคลากรที่เกี่ยวข้องกับการจัดการฐานข้อมูลแต่ละคนต้องเข้าใจบทบาทและหน้าที่ความรับผิดชอบของตนเอง โดยเฉพาะอย่างยิ่งผู้ออกแบบฐานข้อมูลที่จะต้องทำให้การติดต่อระหว่างฐานข้อมูลกับผู้ใช้เกิดความสะดวกในการใช้งาน ข้อมูลที่จัดเก็บมีความถูกต้องและปลอดภัย ดังนั้นจึงต้องเข้าใจถึงสถาปัตยกรรมของฐานข้อมูลในแต่ละระดับและรูปแบบของฐานข้อมูลแบบต่างๆ

สถาปัตยกรรมฐานข้อมูล

ระบบฐานข้อมูลจะมีประโยชน์ก็ต่อเมื่อผู้ใช้สามารถเข้าถึงข้อมูลหรือจัดการกับข้อมูลได้อย่างมีประสิทธิภาพ นอกจากนั้นการที่ผู้ใช้ฐานข้อมูลส่วนใหญ่ไม่ใช่ผู้เชี่ยวชาญทางด้านคอมพิวเตอร์ในการออกแบบฐานข้อมูลจึงได้มีการซ่อนรายละเอียดที่ซับซ้อนต่างๆ ไว้ภายใน โดยผู้ใช้จะได้เห็นข้อมูลในเชิงนามธรรมเท่านั้น และสามารถมองเห็นได้ในมุมมองที่แตกต่างกัน ในปี ค.ศ. 1975 American National Standard Institute ได้นำเสนอสถาปัตยกรรมฐานข้อมูลที่แบ่งเป็น 3 ระดับดังต่อไปนี้

1. ระดับภายใน (Internal Level) เรียกอีกอย่างหนึ่งว่า ระดับกายภาพ (physical level) เป็นข้อมูลเชิงนามธรรมในระดับล่างสุด ซึ่งอยู่ในความรับผิดชอบของผู้ออกแบบฐานข้อมูล โดยเป็นผู้พิจารณาว่าจะเก็บข้อมูลไว้ในอุปกรณ์ใด การจัดเก็บข้อมูลจริงๆ ควรทำอย่างไร วิธีการเข้าถึงข้อมูลเพื่อค้นหาหรือปรับปรุงข้อมูลจะใช้วิธีการใด รวมถึงวิธีการบำรุงรักษา และการเพิ่มประสิทธิภาพของฐานข้อมูล ผู้ใช้งานฐานข้อมูลทั่วไปจะไม่เกี่ยวข้องกับการจัดการข้อมูลในระดับนี้

2. ระดับแนวคิด (Conceptual Level) เรียกอีกอย่างหนึ่งว่า ระดับตรรกะ (logical level) เป็นข้อมูลเชิงนามธรรมในระดับที่สูงขึ้นมาอีกระดับหนึ่ง ใช้ในการอธิบายว่ามีข้อมูลอะไรบ้างที่ถูกจัดเก็บไว้ในฐานข้อมูล และข้อมูลเหล่านี้มีความสัมพันธ์กันอย่างไร ผู้ใช้ในระดับตรรกะไม่มีความจำเป็นต้องไปคำนึงถึงความยุ่งยากต่างๆ ในระดับกายภาพเลย ข้อมูลในระดับตรรกะนี้จะถูกกำหนดโดยผู้ดูแลฐานข้อมูล ซึ่งต้องตัดสินใจว่าข้อมูลใดบ้างที่จะถูกจัดเก็บลงในฐานข้อมูล ดังนั้นจึงต้องมีการวิเคราะห์ความต้องการของผู้ใช้เพื่อจะนำข้อมูลที่ได้มาออกแบบฐานข้อมูล

3. ระดับภายนอก (External Level) เรียกอีกอย่างหนึ่งว่า ระดับวิว (view level) เป็นข้อมูลเชิงนามธรรมระดับสูงสุด ใช้อธิบายเกี่ยวกับบางส่วนของฐานข้อมูล เนื่องจากผู้ใช้ฐานข้อมูลส่วนใหญ่ไม่มีความเกี่ยวข้องกับข้อมูลทั้งหมดที่มีอยู่ในฐานข้อมูล แต่มีความจำเป็นต้องเข้าถึงฐานข้อมูลเท่านั้น ดังนั้นเพื่อทำให้การติดต่อกับฐานข้อมูลสามารถกระทำได้ง่ายขึ้น จึงมีการกำหนดข้อมูลเชิงนามธรรมในระดับวิวขึ้น โดยที่ในฐานข้อมูลหนึ่งอาจมีการกำหนดวิวได้หลายๆ วิวที่แตกต่างกันได้ นอกจากนั้นการกำหนดข้อมูลเชิงนามธรรมในระดับวิว ยังเป็นการช่วยรักษาความปลอดภัยและความมั่นคงให้กับฐานข้อมูลอีกด้วยเนื่องจากผู้ใช้แต่ละคนจะมองเห็นเฉพาะข้อมูลที่ตนเองได้รับอนุญาตเท่านั้น ทั้งๆ ที่ในฐานข้อมูลมีข้อมูลอื่นๆ ที่เป็นความลับที่อนุญาตให้เฉพาะบุคคลที่มีหน้าที่รับผิดชอบในเรื่องนั้นๆ สามารถมองเห็นและใช้ข้อมูลได้รวมอยู่ด้วย

เมื่อเข้าใจแล้วว่าฐานข้อมูลระดับต่างๆ มีความสำคัญอย่างไร ก็ต้องมีการพัฒนาระบบการจัดการฐานข้อมูลให้สอดคล้องกับข้อมูลที่มีอยู่ในปัจจุบัน และคำนึงถึงสถาปัตยกรรมของฐานข้อมูลด้วย ซึ่งมีแนวคิดในการพัฒนาหลายรูปแบบโดยมีการถ่ายทอดออกมาในลักษณะของแบบจำลองต่างๆ ดังจะกล่าวในรายละเอียดต่อไป

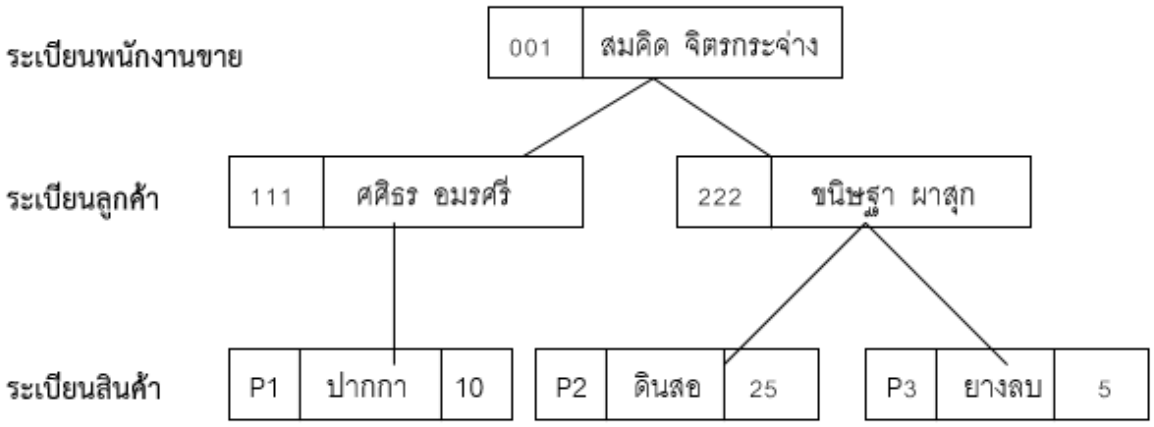
9.1 การเตรียมข้อมูล

<u>รูปแบบของฐานข้อมูล</u>

ฐานข้อมูลเป็นรากฐานสำคัญในการพัฒนาการจัดการข้อมูล จึงมีการสร้างแบบจำลองฐานข้อมูล (database model) ขึ้นมา เพื่อเก็บข้อมูลจริงๆ ตามแนวคิดต่างๆ เพื่อนำเสนอข้อมูลและความสัมพันธ์ระหว่างข้อมูลในรูปแบบที่เข้าใจได้ง่าย โดยแบบจำลองข้อมูลที่ได้รับการยอมรับก็จะถูกนำมาสร้างเป็นระบบฐานข้อมูลยี่ห้อต่างๆ ให้เลือกใช้กัน ในการพัฒนาเป็นฐานข้อมูลมีหลายรูปแบบดังรายละเอียดต่อไปนี้

1. ฐานข้อมูลแบบลำดับชั้น (The Hierarchical Database Model)

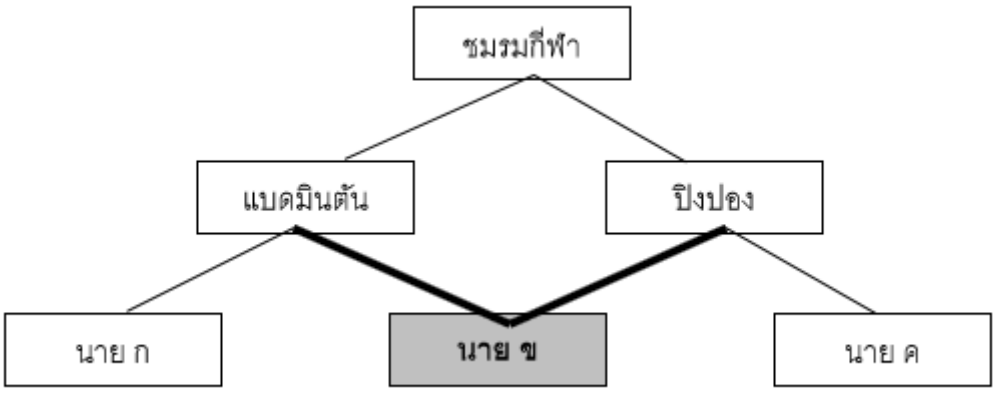
มีลักษณะเป็นแผนภูมิต้นไม้ (tree) ความสัมพันธ์เป็นแบบพ่อกับลูก (parent/child relation) คือระเบียนที่อยู่ด้านบนของโครงสร้างหรือพ่อนั้น สามารถมีลูกได้มากกว่าหนึ่งคน แต่ลูกจะไม่สามารถมีพ่อมากกว่าหนึ่งคนได้ ดังนั้นความสัมพันธ์แบบนี้จึงช่วยลดปัญหาจากการเก็บข้อมูลที่ซ้ำซ้อนกันได้ แต่การที่จะใช้งานข้อมูลได้นั้น ผู้ใช้จะต้องรู้โครงสร้างการเก็บข้อมูลเช่นเดียวกับการเก็บข้อมูลในไฟล์ โดยจะต้องรู้ว่าแผนภูมิต้นไม้นั้นๆ มีข้อมูลอะไรอยู่บ้าง และต้องรู้ด้วยว่าแต่ละลำดับชั้นนั้นเก็บข้อมูลอะไรอยู่



จากตัวอย่างข้างต้นจะเห็นว่าลูกค้าแต่ละคนจะไม่สามารถได้รับบริการจากพนักขายมากกว่าหนึ่งคนได้ เนื่องจากลูกค้าแต่ละคนถือว่าเป็นระเบียนลูก และพนักงานถือว่าเป็นระเบียนพ่อของลูกค้า สินค้าแต่ละชนิดก็จะถูกซื้อโดยลูกค้าเพียงคนเดียวเท่านั้น เนื่องจากสินค้าแต่ละชนิดเป็นระเบียนลูกของระเบียนลูกค้า เป็นต้น ซึ่งในความเป็นจริง ลูกค้าคนหนึ่งๆ สามารถได้รับบริการจากพนักงานมากกว่าหนึ่งคนได้ และสินค้าแต่ละชนิดก็จะถูกซื้อโดยลูกค้าหลายคนได้ แสดงว่าลักษณะของฐานข้อมูลแบบลำดับชั้นมีความสัมพันธ์แบบหนึ่งต่อหนึ่ง (one-to-one) และหนึ่งต่อกลุ่ม (oneto-many) แต่ไม่มีความสัมพันธ์แบบกลุ่มต่อกลุ่ม (many-to-many)

2. ฐานข้อมูลแบบเครือข่าย (The Network Database Model)

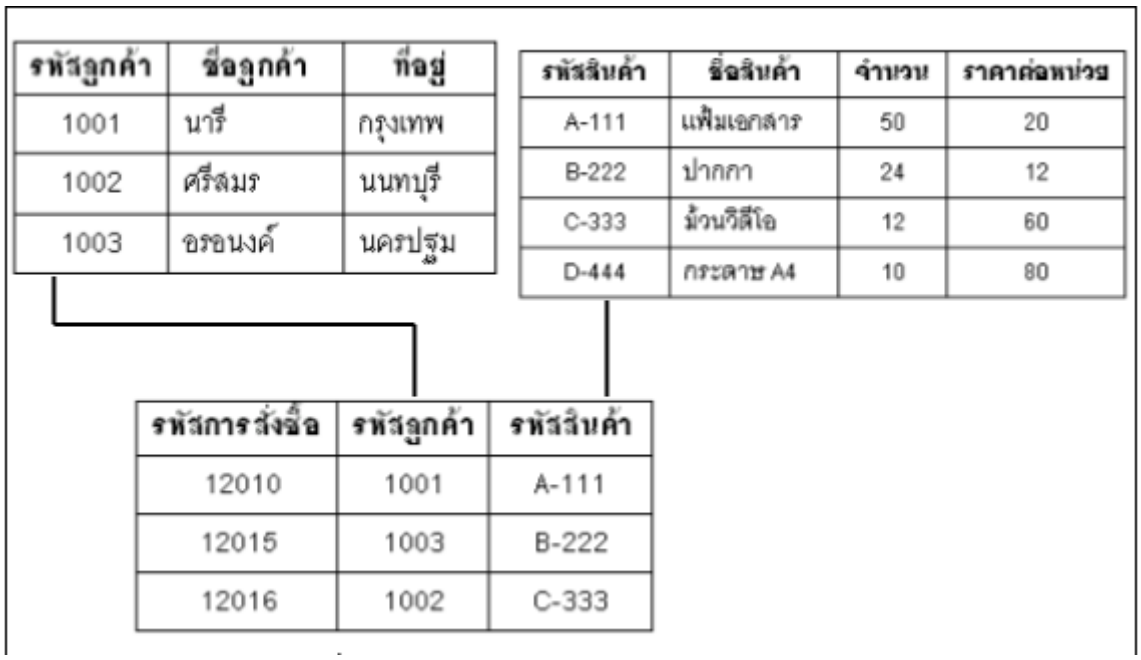
Conference on Data Systems Languages (CODASYL) เป็นหน่วยงานที่คิดค้นแบบจำลองฐานข้อมูลแบบเครือข่าย ขึ้นมาใช้แทนแบบจำลองฐานข้อมูลแบบลำดับชั้น โดยแก้ปัญหาที่ไม่สามารถรองรับความสัมพันธ์แบบกลุ่มต่อกลุ่มได้ ด้วยการนำทฤษฎีเซต ทางคณิตศาสตร์มาใช้ในแบบจำลองข้อมูลนี้ นั่นคือสมาชิกของเซตหนึ่งๆ สามารถเป็นสมาชิกของเซตอื่นได้อีกด้วยตัวอย่างเช่น คนที่ชอบเล่นกีฬาแบดมินตันอาจจะชอบเล่นปิงปองด้วยก็ได้ โดยโครงสร้างของแบบจำลองแบบเครือข่ายก็เป็น tree เช่นเดียวกับแบบจำลองแบบลำดับชั้น แต่จะเป็นแผนภูมิต้นไม้ที่ดูซับซ้อนมากขึ้นเพื่อรองรับความสัมพันธ์แบบกลุ่มต่อกลุ่มนั่นเอง



ถึงแม้ว่าการเก็บข้อมูลแบบเครือข่ายจะช่วยแก้ปัญหาความซ้ำซ้อนของข้อมูลให้หมดไปได้ก็ตาม แต่ความสัมพันธ์ของข้อมูลที่โยงกันไปโยงกันมาก็ทำให้ยากต่อการใช้งานและผู้ใช้ยังคงต้องเข้าใจโครงสร้างข้อมูลเพื่อให้สามารถนำข้อมูลมาใช้ได้เหมือนเดิม ดังนั้นแบบจำลองฐานข้อมูลแบบเครือข่ายจึงเหมาะสำหรับโปรแกรมเมอร์ที่คุ้นเคยเป็นอย่างดีกับโครงสร้างข้อมูลแบบต่างๆ ทั้งแบบง่ายๆ และแบบที่ซับซ้อนอย่างที่ใช้ในแบบเครือข่าย แต่ไม่เหมาะสมสำหรับผู้ใช้งานทั่วๆ ไป ซึ่งต้องการแบบจำลองข้อมูลที่สามารถทำความเข้าใจและใช้งานได้ง่าย

3. ฐานข้อมูลเชิงสัมพันธ์(The Relational Database Model)

เป็นแบบจำลองที่เราใช้กันอยู่ในปัจจุบัน โครงสร้างข้อมูลที่ใช้แสดงความสัมพันธ์ของข้อมูลนั้นเป็นตารางซึ่งเก็บข้อมูลที่มีลักษณะเหมือนกันไว้ ตารางข้อมูลที่เก็บข้อมูลอยู่นั้น เราไม่จำเป็นต้องรู้ว่าตารางเก็บข้อมูลอย่างไรและเก็บไว้ที่ไหนก็สามารถนำข้อมูลมาใช้ได้ โดยตารางจะมีชื่อเรียกเพื่อให้อ้างถึงเวลาต้องการข้อมูลในตารางนั้น และเมื่อเราต้องการข้อมูลในตารางเราก็จะใช้วิธีเปรียบเทียบค่าของข้อมูลแทน ด้วยความง่ายในการทำความเข้าใจและง่ายในการนำเข้าข้อมูล ทำให้แบบจำลองฐานข้อมูลเชิงสัมพันธ์มีผู้ใช้งานกันอย่างแพร่หลายในปัจจุบัน



รหัสลูกค้า	ชื่อลูกค้า	ที่อยู่
1001	นารี	กรุงเทพ
1002	ศรีสมร	นนทบุรี
1003	อรอนงค์	นครปฐม

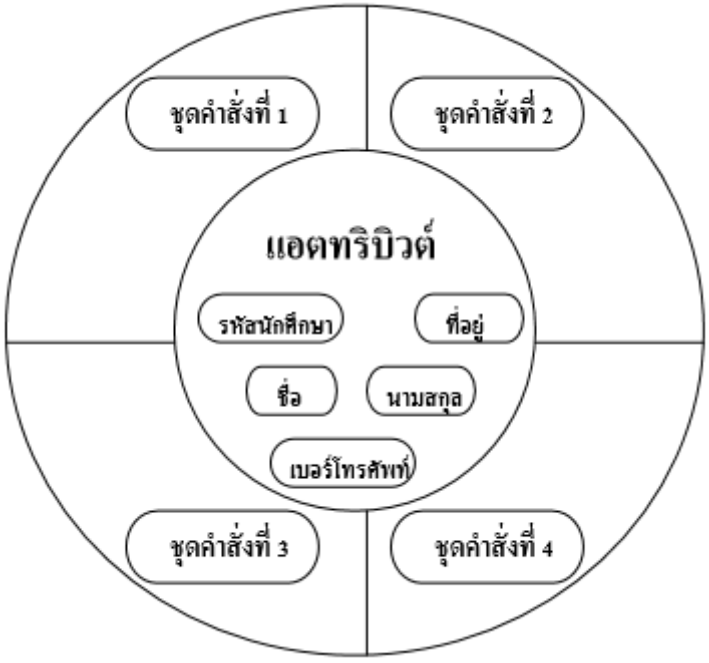
รหัสสินค้า	ชื่อสินค้า	จำนวน	ราคาต่อหน่วย
A-111	แฟ้มเอกสาร	50	20
B-222	ปากกา	24	12
C-333	ม้วนวิดีโอ	12	60
D-444	กระดาษ A4	10	80

รหัสการสั่งซื้อ	รหัสลูกค้า	รหัสสินค้า
12010	1001	A-111
12015	1003	B-222
12016	1002	C-333

4. ฐานข้อมูลเชิงวัตถุ (The Object-Oriented Database Model)

เป็นแบบจำลองที่นำมาใช้กับหน่วยงานขนาดใหญ่ จำเป็นต้องใช้บุคลากรที่มีความรู้ความชำนาญและประสบการณ์ในการจัดการกับข้อมูลที่มีความสลับซับซ้อน เนื่องจากข้อมูลในปัจจุบันไม่ได้มีเพียงแค่ข้อความเท่านั้น แต่ยังมีข้อมูลที่เป็นภาพและเสียงเพิ่มมากขึ้น นอกจากนั้นยังมีฐานข้อมูลด้านมัลติมีเดีย งานดีไซน์ที่ใช้คอมพิวเตอร์ช่วยในการออกแบบ (Computer Aid Design : CAD) และฐานข้อมูลไฮเปอร์เท็กซ์ (Hypertext) ซึ่งเป็นเอกสารที่มีลิงค์เชื่อมถึงกันได้ โดยเว็บเพ็จก็เป็นเอกสารไฮเปอร์เท็กซ์ชนิดหนึ่ง ทำให้ในบางครั้งแบบจำลองฐานข้อมูลเชิงสัมพันธ์ไม่เหมาะสมที่จะนำมาใช้กับงานดังกล่าว ถึงแม้ว่าแบบจำลองฐานข้อมูลเชิงสัมพันธ์จะสามารถเก็บข้อมูลภาพและเสียงได้ แต่จะต้องเพิ่มขั้นตอนเพื่อแปลงข้อมูลให้สามารถเก็บอยู่ในตารางได้ และก็ต้องแปลงกลับเมื่อจะนำมาใช้อีกด้วย ดังนั้นจึงมีการคิดแบบจำลองข้อมูลเชิงวัตถุ

ขึ้นมาเพื่อเพิ่มความสามารถของฐานข้อมูลให้กับภาษาคอมพิวเตอร์แบบวัตถุ (object-oriented) โดยมองและเก็บข้อมูลเป็นวัตถุ แต่ความสามารถในการค้นหาข้อมูลไม่ดีเท่ากับแบบจำลองฐานข้อมูลเชิงสัมพันธ์ แต่ก็มีประสิทธิภาพในการประมวลผลสูงกว่าเนื่องจากได้ความสามารถของภาษาคอมพิวเตอร์มาใช้ฐานข้อมูลแบบลำดับชั้น แบบเครือข่าย และฐานข้อมูลเชิงสัมพันธ์ ล้วนจัดเก็บเฉพาะข้อมูล (ข้อมูลและความสัมพันธ์ระหว่างข้อมูล) ไว้ในฐานข้อมูล ส่วนชุดคำสั่งที่ใช้ในการดำเนินการกับฐานข้อมูลจะจัดเก็บไว้ในซอฟแวร์ระบบจัดการฐานข้อมูลแยกต่างหาก แต่มีฐานข้อมูลชนิดหนึ่งที่จัดเก็บทั้งข้อมูลและชุดคำสั่งไว้ด้วยกัน ซึ่งการรวมข้อมูลและคำสั่งในการดำเนินการใดๆ เข้าด้วยกันจะเรียกสิ่งนั้นว่า “วัตถุ (object)” และฐานข้อมูลที่จะนำมาจัดการกับวัตถุก็จะถูกเรียกว่า “ฐานข้อมูลเชิงวัตถุ (object-oriented database)”



ฐานข้อมูลเชิงวัตถุ เป็นฐานข้อมูลที่สามารถจัดเก็บข้อมูลและวิธีการเสมือนวัตถุที่สามารถจะดึงไปใช้งานร่วมกันได้โดยอัตโนมัติ ระบบการจัดการฐานข้อมูลเชิงวัตถุสามารถจัดการข้อมูลในลักษณะต่างๆ เช่น แบบเขียน (drawing) รูปภาพ เสียงและภาพเคลื่อนไหวได้อย่างมีประสิทธิภาพ

ถึงแม้ว่าฐานข้อมูลชนิดนี้จะมีประสิทธิภาพในการจัดเก็บและจัดการ แต่ก็พบว่ายังไม่มีการนำมาใช้งานอย่างแพร่หลายเท่ากับฐานข้อมูลเชิงสัมพันธ์ ที่พบก็เพียงการนำแนวคิดเชิงวัตถุมาผสมผสานการทำงานกับแนวคิดฐานข้อมูลเชิงสัมพันธ์ จึงเรียกฐานข้อมูลชนิดนี้ว่า “object-relation database” ที่ผู้ดูแลฐานข้อมูลสามารถใช้ภาษาระบบจัดการฐานข้อมูลชนิดเดียวกับที่ใช้กับฐานข้อมูลเชิงสัมพันธ์ ดังนั้นจึงสามารถจัดหาระบบการจัดการฐานข้อมูลได้เช่นเดียวกับฐานข้อมูลเชิงสัมพันธ์เช่น โปรแกรมออราเคิล (Oracle), ดีบีทู (DB2) และ ไซเบส (SyBase) เป็นต้น

5. ฐานข้อมูลเชิงวัตถุ-สัมพันธ์(The Object-Relational Database Model)

ถูกสร้างขึ้นเพื่อให้ฐานข้อมูลเชิงสัมพันธ์ที่ใช้งานอยู่ สามารถเพิ่มคุณสมบัติของแบบจำลองเชิงวัตถุเข้าไปได้โดยไม่ต้องเสียค่าใช้จ่ายเพิ่มในด้านการออกแบบข้อมูลใหม่ หรือเปลี่ยนแปลงระบบฐานข้อมูลเดิม โดยสิ่งที่เพิ่มขึ้นมาจากแบบจำลองฐานข้อมูลเชิงสัมพันธ์ คือสามารถสร้างชนิดข้อมูลที่กำหนดเองได้ โดยชนิดข้อมูลที่สร้างเองนี้สามารถเก็บข้อมูลชนิดอื่นๆ ไว้ภายในได้อีก และยังสามารถสร้างวิธี(method) เพื่อจัดการกับข้อมูลภายในได้อีกด้วย ดังนั้นตารางจึงสามารถเก็บข้อมูลที่มีความซับซ้อนได้ตามชนิดข้อมูลที่เพิ่มขึ้นมา ซึ่งตัวอย่างของระบบการจัดการที่ได้รับความนิยมทั่วโลกก็คือโปรแกรมออราเคิล (Oracle)

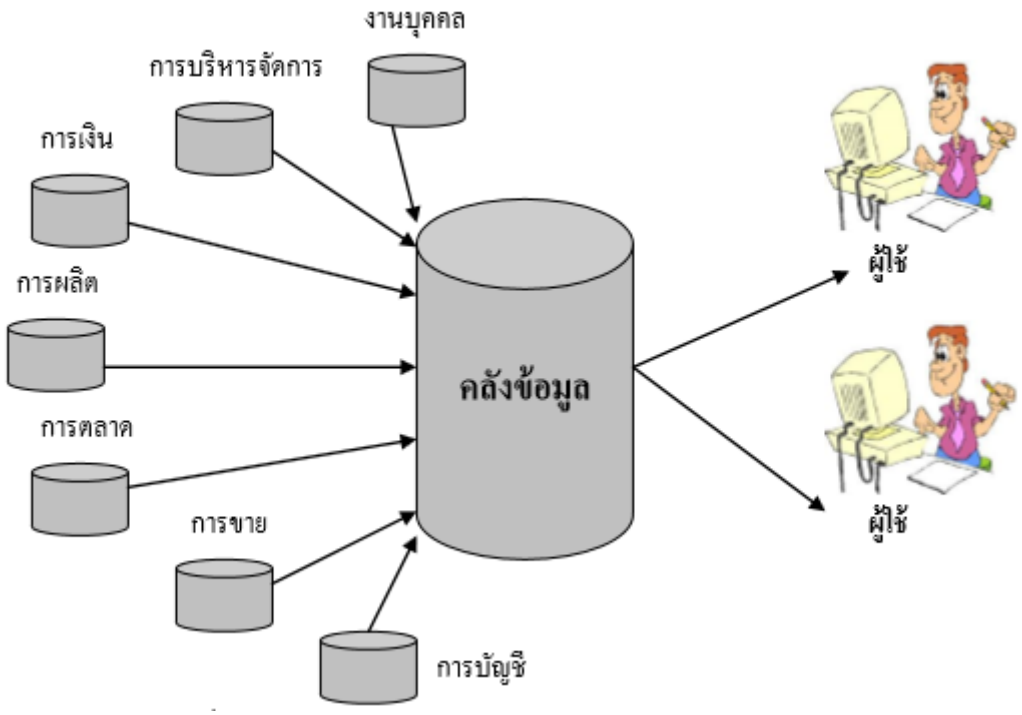
6. คลังข้อมูล (Data Warehouse)

ฐานข้อมูลส่วนใหญ่ในปัจจุบันใช้เพื่อจัดเก็บข้อมูลที่เป็นปัจจุบัน เพื่อนำออกมาใช้งานในด้านต่างๆ โดยเฉพาะด้านธุรกิจ เช่น คลังสินค้า การเงิน การผลิต และการขายสินค้า เป็นต้น แต่ในบางครั้งข้อมูลในอดีตก็มีประโยชน์ในการวางแผนการดำเนินงานในอนาคตได้ ซึ่งข้อมูลในอดีตเหล่านั้นอาจจัดเก็บในฐานข้อมูลอื่น จึงจำเป็นจะต้องมีการเชื่อมโยงฐานข้อมูลเพื่อเรียกใช้ข้อมูลที่สัมพันธ์กัน โดยเฉพาะอย่างยิ่งองค์กรขนาดใหญ่ที่มีฐานข้อมูลจำนวนมาก ข้อมูลในฐานข้อมูลที่หนึ่งมีความสัมพันธ์กับข้อมูลในฐานข้อมูลอีกที่หนึ่ง จึงต้องมีการเข้าถึงข้อมูลจากฐานข้อมูลที่แตกต่างกันแต่ข้อมูลสัมพันธ์กัน

ด้วยเหตุผลดังกล่าว ทำให้การพัฒนาฐานข้อมูลในรูปแบบเดิมมีความสามารถมากขึ้นเพื่อรองรับและตอบสนองความต้องการตามที่กล่าวไว้ข้างต้น กลายเป็นฐานข้อมูลที่มีชื่อเรียกว่า “คลังข้อมูล (Data Warehouse)”

คลังข้อมูล หมายถึง ฐานข้อมูลที่จัดเก็บข้อมูลที่ได้มาจากการสกัดข้อมูล (extract) จากฐานข้อมูลอื่น ซึ่งอาจจัดโครงสร้างแตกต่างกันหรืออยู่บนระบบปฏิบัติการแตกต่างกันก็ได้ ทั้งนี้เพื่อประโยชน์ในการวิเคราะห์และตัดสินใจเชิงธุรกิจ

ดังนั้นคลังข้อมูลจึงถือว่าเป็นแหล่งข้อมูลขนาดใหญ่ของทั้งองค์กร ซึ่งแต่ละองค์กรจะต้องมีการแบ่งส่วนการทำงานออกเป็นส่วนย่อยๆ ขึ้นอยู่กับหน้าที่การทำงาน เช่น การแบ่งส่วนออกเป็นการผลิต การเงิน การตลาด การขาย การบัญชี และการบริหารงานบุคคล เป็นต้น ซึ่งในบางกรณีต้องทำสำเนาข้อมูลที่มีความจำเป็นต้องใช้งานของแต่ละส่วนมาจัดเก็บไว้ภายในส่วนการทำงานย่อยๆเหล่านั้น ซึ่งจะเรียกแหล่งข้อมูลที่จัดเก็บข้อมูลที่ถูกสำเนามานี้ว่า “ดาต้า มาร์ท (data mart)” ดังรูป



สำหรับงานด้านการตัดสินใจแก้ปัญหานั้น คลังข้อมูลนับว่ามีส่วนช่วยในการจัดหาสารสนเทศในแต่ละด้านได้เป็นอย่างดี นอกจากนี้ คลังข้อมูลยังจะต้องมีความสามารถค้นหาข้อมูลเชิงวิเคราะห์เพื่อให้ได้ข้อมูลที่มีขอบเขตเหมาะสมกับปัญหาได้ชัดเจนขึ้น ซึ่งจะต้องอาศัยเทคนิคอย่างหนึ่งที่เรียกว่า “เหมืองข้อมูล (data mining)” เหมืองข้อมูล บางครั้งเรียกว่า “การค้นพบองค์ความรู้ (knowledge discovery)” เป็นเทคนิคที่ใช้ในการค้นหาข้อมูลเชิงวิเคราะห์ชั้นสูง สำหรับใช้จัดการเก็บข้อมูลที่มีอยู่ในฐานข้อมูลเป็นจำนวนมาก ซึ่งประกอบไปด้วยการค้นหา แยกแยะกลุ่มข้อมูล และคัดเลือกข้อมูลที่มีคุณค่า เป็นต้นเพื่อใช้อธิบายข้อมูลในอดีตและคาดการณ์ข้อมูลในอนาคต

7. ฐานข้อมูลแบบหลายมิติ(Multidimensional Database)

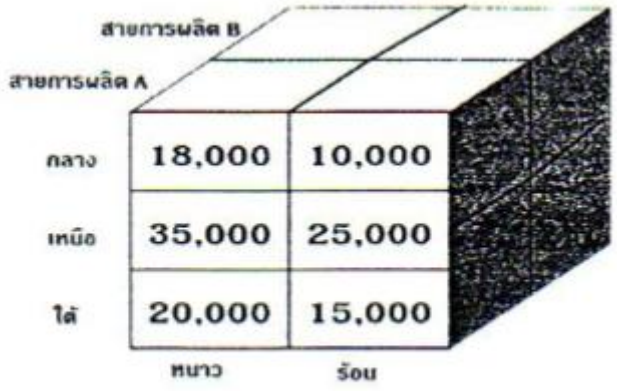
เป็นฐานข้อมูลที่พัฒนามาจากฐานข้อมูลเชิงสัมพันธ์ ซึ่งแต่ละแถวก็คือ เรคคอร์ดที่ทำให้ทราบคุณลักษณะของข้อมูลใด ๆ การมองข้อมูลที่เห็นเป็นเพียงเรคคอร์ดเท่านั้น เป็นการมองแบบมิติเดียว เช่น สมมติว่ามีตารางข้อมูลยอดขายสินค้าของพนักงานขายแต่ละคนในแต่ละภูมิภาค และในแต่ละภูมิภาคนั้นแบ่งออกเป็นยอดขายในแต่ละฤดูด้วยและในแต่ละฤดูแบ่งออกเป็นยอดขายในแต่ละสายการผลิต (production line) ดังตารางข้อมูลต่อไปนี้

ตารางที่ 2.1 ข้อมูลยอดขายสินค้าของพนักงานขายแต่ละคนในแต่ละภูมิภาค แต่ละฤดูกาล แต่ละสายการผลิต

ภูมิภาค	ฤดูกาล	สายการผลิต	ยอดขาย
กลาง	หนาว	A	18,000
กลาง	หนาว	B	15,000
กลาง	ร้อน	A	10,000
กลาง	ร้อน	B	12,500
เหนือ	หนาว	A	35,000
เหนือ	หนาว	B	37,500
เหนือ	ร้อน	A	25,000
เหนือ	ร้อน	B	20,000
ใต้	หนาว	A	20,000
ใต้	หนาว	B	18,000
ใต้	ร้อน	A	15,000
ใต้	ร้อน	B	12,500

จากตาราง จะเห็นว่าในแต่ละเรคคอร์ดนั้นแสดงยอดขาย 1 ค่าในภูมิภาค ฤดู และสายการผลิต ซึ่งคุณลักษณะข้อมูลทั้งหลายนี้ รวมกันอยู่ในเรคคอร์ดเดียว ปัญหาที่เกิดขึ้น ก็คือ การวิเคราะห์ข้อมูลในตารางข้อมูลนั้นทำได้ยาก เช่น ต้องการหายอดรวมในฤดูหนาวทุกสายการผลิตและทุกภูมิภาค จากตารางข้อมูลข้างต้น จะทำได้ยากมาก

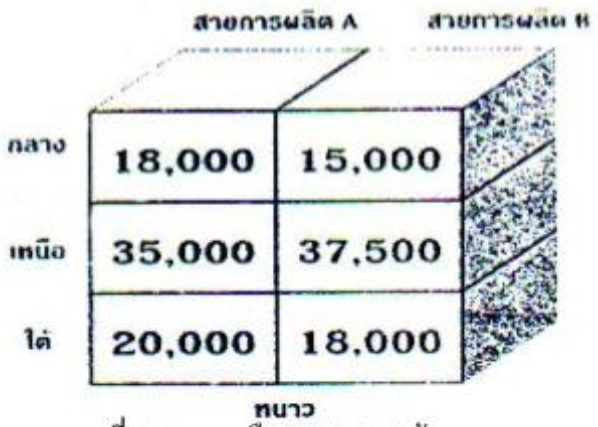
ดังนั้น จึงต้องมีการจัดเก็บข้อมูลที่มากกว่า 1 มิติ นั่นคือ การจัดเก็บข้อมูลไว้ในฐานข้อมูลแบบหลายมิติ ซึ่งจะเก็บข้อมูลได้มากกว่า 1 มิติ



ภาพที่ 2.7 ยอดขายในฐานข้อมูลแบบหลายมิติ

จากภาพที่ 2.7 แสดงข้อมูลจากตารางข้อมูลในฐานข้อมูลเชิงสัมพันธ์ให้อยู่ในฐานข้อมูล 3 มิติ มีลักษณะรูปลูกบาศก์ (cube) มิติที่ 1 ของฐานข้อมูลนี้จะแสดงยอดขาย (sale) ในแต่ละภูมิภาค มิติที่ 2 แสดง

ฤดูกาล (season) และมิติที่ 3 แสดงสายการผลิต (product line) การแสดงข้อมูลหลายมิติ จะทำให้การวิเคราะห์ข้อมูล ยอดขายรวมทุกฤดูกาลและทุกภูมิภาคทำได้ง่ายขึ้น ด้วยวิธีการ “เฉือน (slicing)” ลูกบาศก์ออกเป็นส่วนๆ ตามที่ต้องการคำนวณเท่านั้น ดังภาพที่ 2.8



ภาพที่ 2.8 การเฉือน (slicing) ข้อมูล

นอกจากนี้หากผู้ใช้ต้องการวิเคราะห์ข้อมูลยอดขายในฤดูหนาวเพิ่มเติม เช่น หายอดขายรวมในฤดูหนาวเฉพาะสายการผลิต A และเฉพาะในภาคเหนือ ก็สามารถทำได้ด้วยการเฉือนเฉพาะยอดขายในฤดูหนาวในภาคเหนือ และเฉพาะสายการผลิต A ออกจากลูกบาศก์ที่เฉือนมาแล้วในขั้นตอนที่ผ่านมา และผลที่ได้ก็คือ “ลูกเต๋า (dicing)” ซึ่งในบางครั้งเรียกวิธีการดังกล่าวว่า “slice and dice”

วิธีการ “slice and dice” นั้นเป็นวิธีการที่จัดอยู่ในเทคนิคชนิดหนึ่งที่ใช้ในการวิเคราะห์ข้อมูลจากฐานข้อมูล ดาต้ามาร์ท และคลังข้อมูล เทคนิคดังกล่าวเรียกว่า “การประมวลผลเชิงวิเคราะห์แบบออนไลน์ (online analytical processing) ซึ่งเป็นการวิเคราะห์ข้อมูลจำนวนมากที่จัดเก็บอยู่ในฐานข้อมูลดาต้ามาร์ทและฐานข้อมูลแบบหลายมิติ โดยมีวัตถุประสงค์ก็เพื่อจัดหาสารสนเทศที่ตรงตามความต้องการของผู้ใช้ โดยที่ผู้ใช้ไม่จำเป็นต้องนำไปวิเคราะห์ต่อเลย

8. ฐานข้อมูลบนเว็บ (Web Database)

ฐานข้อมูลเว็บ คือ การสร้างเว็บเพจที่ผู้ใช้สามารถเลือกดูในสิ่งที่ต้องการได้ ในสมัยก่อนเว็บเพจมีความสามารถเพียงแสดงผลข้อมูลที่ได้มีการจัดเตรียมไว้เรียบร้อยแล้วหรือเรียกว่า “static web” เท่านั้น หากต้องการนำเสนอสิ่งอื่น เพิ่มเติมก็จะต้องทำการสร้างหน้าเว็บเพจขึ้นมาใหม่อยู่เสมอ อีกทั้งทำให้เว็บเพจขาดความน่าสนใจในการใช้งาน เนื่องจากผู้ใช้ทำได้เพียงเป็นฝ่ายรับข้อมูลเท่านั้น แต่เมื่อมีการนำฐานข้อมูลมา

ประยุกต์กับการสร้างเว็บเพจ ทำให้เว็บเพจมีความสามารถเพิ่มมากขึ้นเป็นอย่างมาก เช่น ผู้ใช้สามารถป้อนข้อความแสดงความคิดเห็นไว้บนเว็บเพจได้ หรือเรียกว่า “เว็บบอร์ด” เป็นต้น

หลักการของเว็บฐานข้อมูล คือ การสร้างฐานข้อมูลไว้ที่เครื่องเซิร์ฟเวอร์ที่ให้บริการเว็บไซต์ที่ต้องการสร้างเป็นเว็บฐานข้อมูล จากนั้นจึงเขียนสคริปต์ (script) ด้วยภาษาในการโปรแกรมมิ่งเว็บฐานข้อมูล เช่น ASP, PHP หรือ CGI เป็นต้น ไว้แทนการสร้างไฟล์ html โดยไฟล์สคริปต์เหล่านี้เมื่อมีผู้เข้ามาเรียกไปแสดงผลยังเครื่องของตนจะถูกประมวลผลโดยเครื่องเซิร์ฟเวอร์ก่อน ซึ่งเซิร์ฟเวอร์จะตรวจสอบคำสั่งที่มีในไฟล์สคริปต์แล้วทำการติดต่อหรืออ่านค่าข้อมูลตามคำสั่งที่มีในไฟล์สคริปต์จากฐานข้อมูล จากนั้นจึงจะน าไปแสดงผลยังเครื่องของผู้ใช้ที่เข้ามาเรียกดูเว็บ

จะเห็นว่าเว็บไซต์ที่สร้างขึ้นโดยใช้เว็บฐานข้อมูลมีมากมายในปัจจุบัน สามารถสังเกตได้จากเว็บเหล่านั้นมักจะมีแบบฟอร์มสำหรับให้ผู้ใช้กรอกข้อมูลหรือเลือกข้อมูลที่ต้องการให้แสดงผลได้เอง สามารถค้นหาข้อมูลต่าง ๆ ได้อย่างรวดเร็วและง่ายดาย ตัวอย่างเช่น www.google.co.th หรือ www.yahoo.com เป็นต้น นอกจากนี้ยังมีเว็บไซต์อื่น ๆ อีกมากมายที่ให้บริการในเรื่องของการค้นหาข้อมูลอย่างรวดเร็ว ด้วยกลไกในการค้นหาที่เรียกว่า “search engine” โดยขั้นแรกผู้ใช้ต้องพิมพ์คำค้น (keyword) ที่ต้องการค้นหาก่อน จากนั้นเว็บเซิร์ฟเวอร์ของ search engine จะค้นข้อมูลตามที่ผู้ใช้ต้องการ และแสดงผลการค้นหาออกมาในฐานข้อมูลบนเว็บจะประกอบไปด้วยฐานข้อมูลข้อความ และฐานข้อมูลรูปภาพ ซึ่งฐานข้อมูลข้อความ คือ กลุ่มของเอกสารที่ได้รับการเก็บไว้ในฐานข้อมูล โดยที่แต่ละเอกสารและ 27 เนื้อความในเอกสารสามารถถูกค้นหาได้โดยอาศัย Catalog หรือบัญชีรายชื่อของสิ่งต่างๆ ที่อยู่ภายในเอกสารนั้นๆ ปัจจุบันเทคนิคนี้ถูกนำไปใช้ในการค้นหาข้อมูลที่ต้องการบนเว็บ ส่วนฐานข้อมูลรูปภาพคือ ฐานข้อมูลที่ใช้เก็บรายละเอียดต่างๆ เกี่ยวกับรูปภาพ ซึ่งการใช้วิธีเก็บแบบเดิมจะทำได้ไม่ดีนัก ด้วยความสามารถนี้ทำให้สามารถแสดงภาพของสิ่งต่างๆ ได้ตามต้องการ เช่น การนำข้อมูลรูปภาพดังกล่าวไปใช้งานร่วมกับฐานข้อมูล โดยใช้ภาษาโปรแกรมมิ่งใดๆ ดึงรูปภาพจากฐานข้อมูลรูปภาพขึ้นมาแสดง จากนั้นจึงดึงข้อความออกมาจากฐานข้อมูลข้อความเพื่อนำมาใช้อธิบายรายละเอียดต่างๆของรูปภาพดังกล่าว เป็นต้น

9. ฐานข้อมูลแบบกระจาย (Distributed Database)

เป็นการกระจายการจัดเก็บข้อมูลไว้ในหลายๆ สถานที่ที่เรียกว่า “ไซต์ (site)” ซึ่งแต่ละไซต์จะมีเครื่องคอมพิวเตอร์และระบบฐานข้อมูลที่เป็นของตัวเอง เพื่อรองรับการใช้งานต่างๆ ของผู้ใช้ในไซต์นั้นๆ รวมทั้งมีการเชื่อมต่อกับเครื่องคอมพิวเตอร์ของไซต์อื่นด้วยเครือข่ายคอมพิวเตอร์ (Computer Network) เพื่อส่งถ่ายข้อมูล (Data Transmission) ไปมาระหว่างฐานข้อมูลของไซต์ต่างๆ ฐานข้อมูลชนิดนี้เหมาะกับธุรกิจขนาดใหญ่เช่น ธนาคาร และบริษัทประกันชีวิต เป็นต้น ที่มีสาขาตามที่ต่างๆ มากมายโครงสร้างหลักที่แยกความแตกต่างระหว่างระบบฐานข้อมูลศูนย์กลางออกจากระบบจัดการฐานข้อมูลแบบกระจายคือฮาร์ดแวร์ เพราะระบบจัดการฐานข้อมูลแบบกระจายต้องมีเครื่องคอมพิวเตอร์หลายเครื่องที่ ไซต์ หรือที่เรียกอีกอย่างหนึ่งว่าโหนด (node) ต่างๆ และแต่ละโหนดจะต้องติดต่อกับเครือข่ายสื่อสาร เพื่อส่งข้อมูลหรือคำสั่งต่างๆ ระหว่างโหนด โหนดต่างๆ ที่กล่าวถึงนี้อาจจะอยู่ในตึกเดียวกันหรืออยู่ห่างกันไม่มากนัก และติดต่อกันผ่านระบบเครือข่ายท้องถิ่น (LAN) หรืออาจอยู่ห่างไกลกันคนละอำเภอซึ่งจะต้องติดต่อผ่านเครือข่ายระไกลเช่น สายโทรศัพท์ก็ได้นอกจากนี้แต่ละโหนดอาจจะมีรูปแบบการติดต่อผ่านเครือข่ายที่แตกต่างกันก็ได้ เช่น โหนดที่

1 ติดต่อกับโหนดที่ 2 และโหนดที่ 2 ติดต่อกับโหนดที่ 3 แต่โหนดที่ 1 ไม่ได้ติดต่อกับโหนดที่ 3 เป็นต้นในกรณีนี้ถ้าโหนดที่ 1 และโหนดที่ 3 ต้องการติดต่อกันก็ต้องติดต่อผ่านโหนดที่ 2 ก่อนซึ่งอาจทำให้ประสิทธิภาพการทำงานลดลงบ้าง แต่ก็ไม่มีผลกับระบบฐานข้อมูลแต่อย่างใด

ฐานข้อมูลแบบกระจายทำให้การออกแบบระบบ และการจัดการดูแลฐานข้อมูลมีความยุ่งยากมากขึ้น รวมทั้งระบบจัดการฐานข้อมูลที่จะนำมาใช้กับฐานข้อมูลข้อมูลแบบกระจาย ก็จะมีความซับซ้อนมากขึ้นกว่าระบบจัดการฐานข้อมูลแบบศูนย์กลาง กล่าวคือต้องมีคุณสมบัติต่างๆ ดังนี้

— เรียกใช้ข้อมูลที่อยู่ห่างไกลออกไปบนโหนดต่างๆ โดยผ่านทางเครือข่ายสื่อสารได้

— เลือกตัวสินใจได้ว่าควรจะเรียกข้อมูลจากที่ใดมาใช้งานบ้าง

— จัดการดูแลการเรียกดูข้อมูลที่ต่างๆ ได้

— ถ้าระบบใดระบบหนึ่งมีปัญหาไม่ว่าจะด้วยสาเหตุใด เช่น เครื่องมีปัญหา หรือเกิดปัญหากับสารสื่อสาร ระบบจัดการฐานข้อมูลแบบกระจายจะต้องสามารถฟื้นฟูฐานข้อมูลได้เอง

การเลือกใช้งานฐานข้อมูลแบบใดแบบหนึ่ง จะขึ้นอยู่กับปัจจัยหลายประการ ซึ่งต้องทำการพิจารณาให้มีความเหมาะสมกับสภาพขององค์การ โดยปัจจัยที่ทำให้องค์กรใดองค์กรหนึ่งอาจพิจารณาเลือกใช้ฐานข้อมูลแบบกระจาย คือ

— ข้อมูลบางประเภทจะมีความเหมาะสมกับการการใช้ฐานข้อมูลแบบกระจาย เพราะโดยธรรมชาติของข้อมูลนั้นอยู่บนที่ต่าง ๆ กันอยู่แล้ว เช่น บริษัทที่มีสาขาอยู่ในจังหวัดต่างๆ หรือ 28 ธนาคารที่มีหลายๆ สาขา และแต่ละสาขาอยู่คนละแห่งในที่ไกลกันออกไป บริษัทหรือธนาคารดังกล่าวอาจต้องการเก็บข้อมูลของแต่ละสาขาไว้ที่สาขาเหล่านั้น เนื่องจากงานการเรียกใช้ข้อมูลส่วนมากจะเกิดขึ้นโดยพนักงาน หรือลูกค้าที่สาขานั้นๆ โดยมีโอกาสที่จะเรียกใช้ข้อมูลข้ามสาขาน้อยมาก แต่ขณะเดียวกันผู้บริหารที่สำนักงานใหญ่อาจต้องการเรียกใช้ข้อมูลจากหลายๆ สาขาเพื่อดูข้อมูลสรุปบางอย่างด้วย

— เพื่อเพิ่มความเชื่อถือได้ของข้อมูลและความมีประโยชน์ของข้อมูล เพราะการเรียกใช้ข้อมูลจะได้ข้อมูลที่ทันสมัยและถูกต้องอยู่เสมอ ไม่ว่าข้อมูลนั้นจะจัดเก็บอยู่ในที่ใดๆ ก็ตาม และหากเครื่องคอมพิวเตอร์หรือระบบจัดการฐานข้อมูลในโหนดอื่นจะยังคงใช้งานได้เหมือนเกิด แต่ถ้าใช้ระบบฐานข้อมูลศูนย์กลางแล้วหารเครื่องที่ศูนย์กลางเกิดปัญหา ก็จะทำให้ไม่สามาระเรียกดูข้อมูลได้ทั้งระบบ

— ทำให้สามารถใช้ข้อมูลร่วมกัน ในขณะที่การดูแลข้อมูลนั้นแยกจากกันในแต่ละโหนด

— เพิ่มประสิทธิภาพในการทำงาน เนื่องจากการกระจายฐานข้อมูลที่มีขนาดใหญ่ไปในหลาย ๆ แห่ง ทำให้การเรียกใช้ข้อมูลของแต่ละแห่งสามารถทำได้เร็วขึ้น เพราะข้อมูลได้ถูกแบ่งออกเป็นส่วนๆ นอกจากนี้การประมวลของแต่ละโหนดก็จะอิสระจากกัน ทำให้สามารถทำการประมวลผลพร้อมๆ กันได้

การประยุกต์ใช้ไมโครคอมพิวเตอร์และซอฟต์แวร์สำเร็จรูปในระบบการจัดการฐานข้อมูล

ระบบการจัดการฐานข้อมูลต้องอาศัยทั้งไมโครคอมพิวเตอร์และซอฟต์แวร์ในการจัดการฐานข้อมูล ซึ่งสามารถนำมาประยุกต์ใช้ได้ดังนี้

1. การประยุกต์ใช้ไมโครคอมพิวเตอร์ในระบบการจัดการฐานข้อมูล

ในระบบการจัดการฐานข้อมูลสามารถนำไมโครคอมพิวเตอร์มาประยุกต์ใช้งานได้ทั้งการจัดการฐานข้อมูลส่วนบุคคล และในระบบเครือข่าย ดังรายละเอียดต่อไปนี้

1.1 การจัดทำฐานข้อมูลส่วนบุคคล (personal database) เพื่อใช้งานฐานข้อมูลเพียงคนเดียว (single-user) หรืออาจเรียกว่า stand alone database โดยผู้ใช้งานจะทำงานกับฐานข้อมูลได้เฉพาะที่เครื่องนั้นเท่านั้น ซึ่งในการจัดการฐานข้อมูลแบบนี้จะใช้ไมโครคอมพิวเตอร์ได้ทั้งแบบตั้งโต๊ะและแบบเคลื่อนย้าย ที่ใช้สำหรับจัดเก็บฐานข้อมูลส่วนบุคคล ได้แก่

1.2 การจัดทำฐานข้อมูลในระบบเครือข่าย (network) เพื่อใช้งานฐานข้อมูลได้ครั้งละหลายคน (multi-user) ผ่านช่องทางการสื่อสารซึ่งอาจจะเป็นระบบเครือข่ายระยะใกล้ (Local Area Network : LAN) หรืออินเทอร์เน็ต (internet) ซึ่งส่วนใหญ่จะเป็นการจัดการฐานข้อมูลตามหน่วยงานหรือองค์กรต่างๆ เช่น สถาบันการศึกษา โรงพยาบาล ธนาคาร ร้านอาหาร และบริษัทต่างๆเป็นต้น โดยแบ่งฐานข้อมูลเป็น 3 ประเภทหลักๆ คือ

1.2.1 ฐานข้อมูลส่วนกลาง (central database หรือ host database) เป็นฐานข้อมูลสำหรับผู้ใช้หลายคน โดยที่ข้อมูลและโปรแกรมจะเก็บอยู่ที่เครื่องแม่ข่าย (host server) และมีเครื่องลูกข่าย (work terminal) ต่อเชื่อมผ่านระบบเครือข่าย การประมวลผลทั้งหมดจะทำที่เครื่องแม่ข่าย ดังนั้นเครื่องแม่ข่ายจำเป็นต้องมีขีดความสามารถสูงมาก เพื่อรองรับการประมวลผลพร้อมๆ กัน ดังนั้นเครื่องแม่ข่ายที่ใช้ส่วนใหญ่มักจะเป็นเครื่องเมนเฟรมคอมพิวเตอร์ (mainframe computer) หรือเครื่องมินิคอมพิวเตอร์ (minicomputer) ถ้าจะใช้เครื่องไมโครคอมพิวเตอร์ ก็ต้องเลือกที่มีประสิทธิภาพในการประมวลผลรวดเร็วและสามารถจัดเก็บข้อมูลได้มากกว่าเครื่องไมโครคอมพิวเตอร์ที่ใช้งานส่วนบุคคลทั่วๆ ไป

1.2.2ฐานข้อมูลไคลเอนต์/เซิร์ฟเวอร์(client/server database) เป็นฐานข้อมูลสำหรับผู้ใช้หลายคน โดยที่ข้อมูลจะเก็บที่เซิร์ฟเวอร์ และโปรแกรมจะอยู่ที่ไคลเอนต์การประมวลผลจะทำทั้งที่ไคลเอนต์และเซิร์ฟเวอร์ ดังนั้นวิธีนี้จะสามารถลดงานของเซิร์ฟเวอร์ลง เช่น ให้เซิร์ฟเวอร์ทำการอ่านข้อมูล และส่งให้ไคลเอนต์ทำการคำนวณผลลัพธ์เพื่อออกรายงาน เป็นต้น ดังนั้นไคลเอนต์มักจะหมายถึงเครื่องคอมพิวเตอร์ของผู้ใช้โดยเป็นการทำงานในส่วนของการให้บริการหรือติดต่อกับลูกค้า (frontend) ซึ่งมักจะใช้เป็นเครื่องไมโครคอมพิวเตอร์ที่ใช้งานทั่วๆ ไปหรือเครื่องคอมพิวเตอร์ส่วนบุคคล ส่วนเซิร์ฟเวอร์ก็คือเครื่องคอมพิวเตอร์ที่ติดตั้งระบบฐานข้อมูลไว้ (database server) และมีความสามารถสูงพอที่จะให้บริการผู้ใช้ได้พร้อมๆ กันหลายคน โดยจะเป็นส่วนของการบริหารงานภายในองค์กร (backend) ซึ่งควรจะเป็นเครื่องไมโครคอมพิวเตอร์ที่มีประสิทธิภาพสูงกว่าเครื่องที่ใช้งานทั่วๆ ไป หรืออาจจะเป็นเครื่องมินิคอมพิวเตอร์ก็ได้

1.2.3 ฐานข้อมูลแบบกระจาย (distributed database) หมายถึง ข้อมูลจะถูกเก็บแยกไว้คนละที่ เนื่องจากมีการใช้ข้อมูลจากระยะไกล (เช่น กรุงเทพฯ - เชียงใหม่) การเข้าถึงและการส่งข้อมูลจะทำผ่านระบบ

เครือข่าย (ตามสายเคเบิล) ซึ่งจะเกิดความล่าช้า ทำให้การตอบสนองไม่ทันต่อเหตุการณ์ ดังนั้นจึงแยกข้อมูลกระจายเก็บตามสาขา เช่น ที่กรุงเทพฯ ก็จะเก็บข้อมูลลูกค้ากรุงเทพฯ ที่เชียงใหม่ก็เก็บข้อมูลลูกค้าเชียงใหม่ เป็นต้น ซึ่งการควบคุมเรื่องความปลอดภัยจะทำได้ยากกว่าแบบศูนย์กลาง เนื่องจากต้องดูแลฐานข้อมูลทั้ง 2 ที่ แต่การจัดทำฐานข้อมูลแบบกระจายจะช่วยแก้ปัญหาที่เกิดขึ้นจากการทำงานแบบไคลเอนต์/เซิร์ฟเวอร์ เมื่อมีผู้ใช้จำนวนมากเกิน จนเซิร์ฟเวอร์ไม่สามารถทำงานได้รวดเร็วเหมือนเดิม เนื่องจากเกิดปัญหาคอขวด (bottleneck) ซึ่งวิธีการแก้ไขอาจจะเปลี่ยนเครื่องเซิร์ฟเวอร์ใหม่ให้มีความสามารถมากขึ้น หรือไม่ก็เพิ่มจำนวนเซิร์ฟเวอร์ขึ้นเพื่อให้แบ่งงานกันทำ หรืออาจจะทำการกระจาย database server ไปตามที่ต่างๆ ที่มีความต้องการใช้งานฐานข้อมูลเดียวกัน โดยมีการกระจายข้อมูลอยู่ 2 แบบ คือ การคัดลอกข้อมูลทั้งหมดไปไว้ทุกที่ และการแบ่งข้อมูลออกเป็นส่วนๆ ซึ่งอาจจะใช้ทั้ง 2 วิธีร่วมกันก็ได้ ทำให้ต้องใช้ค่าใช้จ่ายสูงกว่าแบบไคลเอนต์/เซิร์ฟเวอร์

2. การประยุกต์ใช้ซอฟต์แวร์สำเร็จรูปในระบบการจัดการฐานข้อมูล

ซอฟต์แวร์สำเร็จรูปในระบบการจัดการฐานข้อมูลเป็นซอฟต์แวร์ในด้านการประมวลผลและจัดการกับข้อมูล กลุ่มข้อมูลจำนวนมาก เช่น รายชื่อลูกค้า ผลิตภัณฑ์ที่จำหน่าย ซอฟต์แวร์ที่เด่นในกลุ่มนี้ได้แก่ ไมโครซอฟต์แอกเซส (Microsoft Access) ออราเคิล (Oracle) ดีเบส (dBase) ฟอกซ์โปร (FoxPro) อินฟอร์มิกซ์ (Informix) มายเอสคิวแอล (MySQL) และไซเบส (Sybase) เป็นต้น ซึ่งมีความสามารถในการสร้าง ดัดแปลง แก้ไขข้อมูลได้สะดวก รวดเร็ว เรียงลำดับข้อมูล ค้นหาข้อมูลได้ตามเงื่อนไข พิมพ์ไฟล์ หรือส่วนของไฟล์ตามที่ต้องการ สร้างหรือดัดแปลงโครงสร้างของไฟล์ข้อมูลได้นำข้อมูลบางส่วนที่ต้องการ มาสร้างเป็นไฟล์ใหม่ได้ทันทีและเก็บข้อมูลได้จำนวนมาก ตัวอย่างของซอฟต์แวร์สำเร็จรูปเพื่อใช้ในการจัดทำระบบการจัดการฐานข้อมูลที่นิยมใช้กันมีรายละเอียดดังต่อไปนี้

2.1 ไมโครซอฟต์แอกเซส (Microsoft Access) เป็นโปรแกรมที่มีความสามารถโดดเด่น การออกแบบเป็นไปตามหลักการของระบบจัดการฐานข้อมูลเชิงสัมพันธ์และอิงมาตรฐานและเตรียมการรองรับเพื่อเชื่อมต่อกับระบบเครือข่ายข้อมูล ทำให้เป็นที่น่าสนใจของผู้ใช้โปรแกรม จึงมีผู้นิยมใช้งานกันอย่างแพร่หลาย ทั้งนี้เพราะมีจุดเด่นหลายประการคือ

2.1.1 เป็นโปรแกรมระบบการจัดการฐานข้อมูลที่ง่ายในการใช้งาน มีเครื่องมือในการทำงานครบถ้วน ผู้ใช้ไม่จำเป็นต้องมีความรู้ในการเขียนโปรแกรม ก็สามารถสร้างและใช้งานระบบฐานข้อมูลได้มีระบบฐานข้อมูลภายใน ขณะเดียวกันจะต้องเข้าใช้ระบบฐานข้อมูลของโปรแกรมสำเร็จรูปอื่นๆ ที่ได้รับความนิยมได้เช่น ดีเบส (dBase), พาราดอกซ์ (Paradox) และฟอกซ์โปร (FoxPro) เป็นต้น ทั้งเรียกใช้และเปลี่ยนแปลงแก้ไขข้อมูลในแฟ้มข้อมูลตั้งต้นได้ด้วย

2.1.2 ระบบจัดการฐานข้อมูลเป็นไปตามแนวทางของระบบฐานข้อมูลแบบสัมพันธ์ (relational database management) อย่างถูกต้อง มีวิธีการสืบค้นที่เป็นมาตรฐาน อาทิเช่น ภาษาเอสคิวแอล (Structured Query Language : SQL) และคิวบีอี(Query By Example : QBE)

2.1.3 ใช้ความสามารถต่างๆ ของวินโดวส์ในการเชื่อมโยง หรือส่งผ่านวัตถุของโปรแกรมต่างๆ บนวินโดวส์ได้ครบถ้วน เช่น สามารถใช้งานโอแอลอี(Object Linking and Embedding : OLE) เป็นต้น

2.1.4 เป็นมาตรฐานเดียวกันสามารถเชื่อมต่อกับระบบการจัดการข้อมูลที่อยู่ในลักษณะเครือข่ายข้อมูลได้ เช่น เชื่อมกับไมโครซอฟต์เอสคิวแอลเซิร์ฟเวอร์ (Microsoft SQL server)และฐานข้อมูลเชิงสัมพันธ์ (relational database) ของมินิคอมพิวเตอร์ (minicomputer) และเมนเฟรมคอมพิวเตอร์ (mainframe computer)

จากคุณสมบัติที่โดดเด่นดังกล่าว ประกอบกับผู้ออกแบบโปรแกรมคือบริษัทไมโครซอฟต์ซึ่งเป็นเจ้าของระบบปฏิบัติการวินโดวส์(Windows) จึงทำให้โปรแกรมไมโครซอฟต์แอกเซสสามารถทำงานภายใต้สภาพแวดล้อมของวินโดวส์ได้อย่างมีประสิทธิภาพ และเป็นส่วนหนึ่งของโปรแกรมชุดไมโครซอฟต์ออฟฟิศ (Microsoft Office) ที่ได้รับความนิยมทั่วโลก

2.2 ออราเคิล (Oracle) เป็นระบบการจัดการฐานข้อมูลแบบวัตถุ-สัมพันธ์ (objectrelational) ที่รวมความสามารถของฐานข้อมูลเชิงสัมพันธ์กับความสามารถในการพัฒนาแอปพลิเคชันแบบเชิงวัตถุ (object-oriented) เข้าด้วยกัน มีประสิทธิภาพในการทำงานแบบไคลเอนต์/เซิร์ฟเวอร์ (client/server) หรือแบบอินเทอร์เน็ตเพื่อสนับสนุนการทำธุรกรรมทางอิเล็กทรอนิกส์ (e-business) และสามารถขยายระบบเพื่อรองรับงานให้เหมาะสมกับขนาดธุรกิจที่เปลี่ยนไป โดยสนับสนุนการทำงานบนหลายแพลตฟอร์ม (platform) ซึ่งข้อมูลที่อยู่ภายในฐานข้อมูลบนแพลตฟอร์มหนึ่งสามารถย้อยไปทำงานบนแพลตฟอร์มอื่นได้อย่างง่ายดาย ส่วนบริการต่างๆ ที่โปรแกรมออราเคิลจัดเตรียมไว้ให้ได้แก่ บริการด้านการจัดเก็บข้อมูล (storage) บริการค้นหาข้อมูล (query) การทำดัชนี (index) และการค้นหาข้อมูลที่เหมาะสม (query optimize) เป็นต้น สำหรับภาษาในการพัฒนานั้นสามารถใช้ได้ทั้งภาษา C, C++, PL/SQL หรือ Java ซึ่งหลักการในการพัฒนา ส่วนใหญ่จะยึดตามมาตรฐานของกลุ่มการจัดการวัตถุ (Object Management Group : OMG) เป็นหลัก

2.3 ดีเบส (dBase) เป็นโปรแกรมสำเร็จรูปในระบบการจัดการฐานข้อมูลที่มีความเก่าแก่ที่สุดสำหรับไมโครคอมพิวเตอร์ และเป็นโปรแกรมสำเร็จรูปตัวหนึ่งที่มีส่วนสำคัญในการช่วยผลักดันให้ไมโครคอมพิวเตอร์ได้รับความนิยม โปรแกรมสำเร็จรูปดีเบส ได้รับการออกแบบมาจากบริษัท แอสตัน เทต (Ashton Tate) แห่งสหรัฐอเมริกา โดยบริษัทแห่งนี้ได้ออกโปรแกรมดีเบสตัวแรกคือ ดีเบส ทู (dBase II) ต่อมาได้ปรับปรุงเป็นดีเบส ทรี(dBase III) ดีเบส ทรีพลัส (dBase III+) ดีเบสโฟร์(dBase IV) และดีเบสไฟว์ (dBase V) ตามลำดับ ปัจจุบันโปรแกรมดังกล่าวนี้พัฒนาและจำหน่ายโดยบริษัท บอร์แลนด์ อินเตอร์เนชั่นแนล (Boland International) ซึ่งเป็นบริษัทที่ได้ซื้อกิจการของบริษัทแอสตัน เทต ไว้โปรแกรมตัวนี้สามารถปฏิบัติการในระบบปฏิบัติการดอส (DOS) รวมกับไมโครซอฟต์วินโดวส์(Microsoft Windows) และยูนิกซ์(Unix) บางรุ่น สำหรับการใช้งานในประเทศไทย โปรแกรม ดีเบส ทรีพลัส นับเป็นโปรแกรมสำเร็จรูปในระบบการจัดการฐานข้อมูลที่มีการใช้งานมากที่สุดในระยะหนึ่ง เนื่องจากตัวแทนจำหน่ายโปรแกรมดังกล่าวได้ทำการพัฒนาความสามารถทางภาษาไทยเพิ่มเติมให้กับโปรแกรมตัวดังกล่าวนี้โดยเฉพาะในเรื่องของการจัดเรียงคำตามพจนานุกรมภาษาไทย ซึ่งมีปัญหาในการใช้งานโปรแกรมสำเร็จรูปในระบบการจัดการฐานข้อมูลส่วนใหญ่

2.4 ฟอกซ์โปร (FoxPro) เป็นโปรแกรมสำเร็จรูปในระบบการจัดการฐานข้อมูลอีกตัวหนึ่งที่ได้รับความนิยมสูงสุดต่อจากโปรแกรม ดีเบส ซึ่งในระยะหลังไม่ได้มีการพัฒนาความสามารถในภาษาไทยต่อมากนัก โปรแกรมสำเร็จรูปฟอกซ์โปร ได้รับการออกแบบและพัฒนาโดยบริษัทฟอกซ์เบส (Foxbase) แต่ต่อมาบริษัทไมโครซอฟต์คอปอเรชั่น ได้เล็งเห็นศักยภาพของโปรแกรมตัวนี้จึงขอซื้อกิจการต่อ ดังนั้นโปรแกรมดังกล่าวจึง

อยู่ภายใต้การพัฒนาของทีมงานในสำกัดบริษัทไมโครซอฟต์คอปอเรชั่น สำหรับการใช้งานในประเทศไทย โปรแกรมดังกล่าวสามารถเก็บข้อมูลสอบถามข้อมูลและพิมพ์รายงานที่เป็นภาษาไทยได้ อย่างไรก็ตามการจัดเรียงภาษาไทยยังไม่สามารถทำได้ตามพจนานุกรม จึงไม่ค่อยเป็นที่นิยมนัก

2.5 อินฟอร์มิกซ์(Informix) เป็นซอฟต์แวร์ระบบการจัดการฐานข้อมูล ที่ได้รับการพัฒนามาจากโพสเกรส (Postgres) โดยเพิ่มเติมเทคโนโลยีใหม่บางส่วนเข้าไป ได้แก่เทคโนโลยีที่เรียกว่าดาต้าเบลด (DataBlade) ที่เกิดจากการพัฒนาของบริษัทมอนเทจซอฟต์แวร์(Montage Software) ซึ่ง ดาต้าเบลดของอินฟอร์มิกซ์นี้เองที่เป็นส่วนในการสนับสนุนการทำงานกับข้อมูลที่เป็นวัตถุ (object) และเป็นจุดเด่นของผลิตภัณฑ์ของอินฟอร์มิกซ์ โดยดาต้าเบลดเป็นโปรแกรมโมดูล (software module) ที่ถูกใส่เข้าไปในระบบเพื่อเพิ่มขีดความสามารถในการปฏิบัติงานของระบบการจัดการฐานข้อมูล ลักษณะของดาต้าเบลดนั้นคล้ายคลาสไลบรารี(class library) ของ C++ ที่คอยบริหาร และจัดการคลาส (class) ของวัตถุ (object) ต่างๆ ซึ่งดาต้าเบลดนี้ ไม่ได้ใช้เพียงเพื่อดูแลชนิดข้อมูลแบบวัตถุ ที่เกิดขึ้นมาใหม่เท่านั้น หากแต่ยังช่วยทำให้เกิดการเข้าถึงข้อมูลต่างๆ เหล่านั้น ในรูปแบบใหม่ที่ดี และมีประสิทธิภาพกว่าเดิม เนื่องจากดาต้าเบลด ประกอบด้วยส่วนการกำหนดชนิดของข้อมูล (datatype definition) ในส่วนของการสร้างโครงสร้างของฐานข้อมูล การดูแลเมท็อด (method) หรือการดำเนินการ (operation) ของข้อมูลชนิดต่างๆ ว่าจะสามารถประมวลผลได้หรือไม่และทำหน้าที่ควบคุมกฎ หรือกฎการบูรณภาพ (integrity constrain) ของข้อมูลแต่ละชนิด

2.6 มายเอสคิวแอล (MySQL) เป็นระบบการจัดการฐานข้อมูลเชิงสัมพันธ์ ซึ่งเป็นที่นิยมมากสำหรับใช้งานในธุรกิจทั่วไป โดยเฉพาะอย่างยิ่งในโลกอินเทอร์เน็ต เนื่องจากเป็น freeware และเป็นซอฟต์แวร์ประเภท open source สามารถดาวน์โหลด source code ต้นฉบับได้จากอินเทอร์เน็ต โดยไม่เสียค่าใช้จ่ายใดๆ การแก้ไขก็สามารถทำได้ตามความต้องการ มายเอสคิวแอลเป็นโปรแกรมที่มีประสิทธิภาพสูง มีความรวดเร็วในการประมวลผล รองรับจำนวนผู้ใช้และขนาดของข้อมูลจำนวนมากได้ รวมทั้งยังสนับสนุนการใช้งานบนระบบปฏิบัติการมากมาย เช่น ยูนิกซ์ (Unix), โอเอสทู (OS/2), แมคโอเอส (Mac OS) และวินโดว์ (Windows) เป็นต้น นอกจากนี้ มายเอสคิวแอลยังสามารถใช้งานร่วมกับแพลตฟอร์มของการพัฒนาเว็บ (Web Development Platform) ได้เช่น C, C++, Java, Perl, PHP และ ASP เป็นต้น ทำให้มายเอสคิวแอลได้รับความนิยมอย่างมากในปัจจุบัน และมีแนวโน้มสูงยิ่งขึ้นต่อไปในอนาคต

2.7 ไซเบส (Sybase) เป็นระบบการจัดการฐานข้อมูลสำหรับอุปกรณ์พกพาไม่ว่าจะเป็น พ็อกเก็ตพีซี (PocketPC) หรือปาล์ม (Plam) ซึ่งในระบบของไซเบสจะมีสิ่งสำคัญที่เรียกว่าโมบายลิงค์ (MobiLink) ซึ่งทำหน้าที่เชื่อมโยงฐานข้อมูลสองตัว คือ ฐานข้อมูลที่อยู่บนอุปกรณ์พกพากับอีกฐานข้อมูลที่เป็นของส่วนกลางให้มีข้อมูลตรงกันด้วย โดยฐานข้อมูลกลางนั้นรองรับหลายยี่ห้อด้วยกัน เช่นไอบีเอ็ม ดีบีทู(IBM DB/2) ออราเคิล (Oracle) และเอสคิวเอลเซิร์ฟเวอร์(SQL Server) เป็นต้น ทำให้เทคโนโลยีนี้มีความสำคัญในการพัฒนาระบบโมบายแอพพลิเคชันอย่างมาก ด้วยคุณสมบัติการสื่อสารระยะไกลที่ให้ผู้ใช้จากทุกที่สามารถเข้าถึงข้อมูลส่วนกลางได้ด้วยความปลอดภัยและได้ข้อมูลที่ถูกต้องตรงกับฐานข้อมูลกลางที่สุด

การเลือกและการประเมินผลฮาร์ดแวร์และซอฟต์แวร์ในการจัดการฐานข้อมูล

การเลือกและประเมินผลฮาร์ดแวร์และซอฟต์แวร์ในการจัดการฐานข้อมูลมีความสำคัญยิ่งต่อการพิจารณาเพื่อเลือกใช้งานในระบบต่างๆ ได้อย่างเหมาะสม ซึ่งต้องอาศัยหลักเกณฑ์และปัจจัยในการพิจารณาที่แตกต่างกัน

1. การเลือกฮาร์ดแวร์และซอฟต์แวร์ในการจัดการฐานข้อมูล

การเลือกฮาร์ดแวร์และซอฟต์แวร์ที่เหมาะสมจะทำให้การจัดการฐานข้อมูลในระบบงานมีประสิทธิภาพและประหยัดค่าใช้จ่ายในการทำงาน ซึ่งมีรายละเอียดดังต่อไปนี้

1.1 การเลือกใช้ฮาร์ดแวร์ในระบบจัดการฐานข้อมูล

ในการเลือกใช้ฮาร์ดแวร์สำหรับระบบจัดการฐานข้อมูลต้องพิจารณาจากปัจจัยหลายประการประกอบกัน เช่น ขนาดของฐานข้อมูล การประมวลผล งบประมาณ จำนวนผู้ใช้ และซอฟต์แวร์เป็นต้น ดังรายละเอียดต่อไปนี้

1.1.1 ขนาดของฐานข้อมูล ซึ่งขึ้นอยู่กับจำนวนแอตทริบิวต์หรือคอลัมน์ และจำนวนทูเพิลหรือเรคคอร์ดของข้อมูลที่จะจัดเก็บในฐานข้อมูล จำนวนของข้อมูลจะทำให้สามารถกำหนดได้ว่าควรใช้สื่อจัดเก็บข้อมูลประเภทใด และคอมพิวเตอร์ที่จะใช้ควรมีประสิทธิภาพมากน้อยแค่ไหน เช่น ถ้าเป็นหน่วยงานขนาดเล็กอาจเลือกเครื่องคอมพิวเตอร์ที่มีประสิทธิภาพปานกลาง ซึ่งได้แก่ เครื่องไมโครคอมพิวเตอร์แต่ถ้าเป็นหน่วยงานขนาดใหญ่อาจต้องใช้เครื่องที่มีประสิทธิภาพสูงซึ่งได้แก่ เครื่องมินิคอมพิวเตอร์ขึ้นไป

1.1.2 การประมวลผล การประมวลผลฐานข้อมูลในเครื่องระดับไมโครคอมพิวเตอร์ สามารถทำการประมวลผลได้ 2 แบบ แบบแรกเป็นการประมวลผลฐานข้อมูลในเครื่องไมโครคอมพิวเตอร์เพียงเครื่องเดียว โดยมีผู้ใช้งานได้เพียงคนเดียวเท่านั้น (single user) ที่สามารถดึงข้อมูลหรือปรับปรุงข้อมูลภายในฐานข้อมูลได้ ส่วนแบบที่สองจะเป็นการนำไมโครคอมพิวเตอร์หลายตัวมาเชื่อมต่อกันในลักษณะของเครือข่ายระยะใกล้ (Local Area Network : LAN) ซึ่งเป็นรูปแบบของระบบเครือข่ายแบบไคลเอนต์/เซิร์ฟเวอร์(client / server network) โดยจะมีการเก็บฐานข้อมูลอยู่ที่เครื่องเซิร์ฟเวอร์ การประมวลผลต่างๆ จะกระทำที่เครื่องเซิร์ฟเวอร์ สำหรับเครื่องไคลเอนต์จะมีหน้าที่ดึงข้อมูลหรือส่งข้อมูลเข้ามาปรับปรุงในเครื่องเซิร์ฟเวอร์ หรือคอยรับผลลัพธ์จากการประมวลผลของเครื่องเซิร์ฟเวอร์ ซึ่งเครื่องเซิร์ฟเวอร์ต้องเป็นเครื่องไมโครคอมพิวเตอร์ที่มีประสิทธิภาพสูง ดังนั้นการประมวลผลแบบนี้จึงเป็นการเปิดโอกาสให้ผู้ใช้งานหลายคนสามารถใช้งานฐานข้อมูลร่วมกันได้ เป็นลักษณะของการทำงานแบบมัลติยูสเซอร์ (multi user)

1.1.3 งบประมาณ ในการเชื่อมต่อเครือข่ายเพื่อใช้ฐานข้อมูลร่วมกันนอกจากจะใช้เครื่องไมโครคอมพิวเตอร์ที่มีประสิทธิภาพสูงแล้ว ถ้าหน่วยงานใดมีงบประมาณมากยังสามารถใช้เครื่องเมนเฟรมคอมพิวเตอร์หรือมินิคอมพิวเตอร์เป็นเครื่องแม่ข่ายได้ซึ่งจะสามารถใช้ต่อกับเทอร์มินัลได้หลายเครื่อง เพื่อให้ผู้ใช้งานฐานข้อมูลหลายคน สามารถดึงข้อมูลหรือปรับปรุงข้อมูลภายในฐานข้อมูลเดียวกันพร้อมกันได้

1.1.4 จำนวนผู้ใช้งาน หากมีผู้เข้ามาใช้ฐานข้อมูลจำนวนมาก เครื่องคอมพิวเตอร์ที่ใช้ในการจัดการฐานข้อมูลต้องมีความสามารถในการเข้าถึงและประมวลผลได้อย่างรวดเร็วมากขึ้น

1.1.5 ซอฟต์แวร์ ในการเลือกฮาร์ดแวร์หรือเครื่องคอมพิวเตอร์เพื่อการจัดการฐานข้อมูลก็ต้องพิจารณาด้วยว่าเหมาะสมกับซอฟต์แวร์ที่จะใช้หรือไม่ คือในการติดตั้งซอฟต์แวร์แต่ละชนิดจะมีคุณสมบัติของเครื่องบอกว่าต้องมีคุณสมบัติอย่างน้อยเป็นอย่างไร

1.2 การเลือกใช้ซอฟต์แวร์ในระบบจัดการฐานข้อมูล

เนื่องจากระบบการจัดการฐานข้อมูลที่วางขายในปัจจุบันมาจากผู้ผลิตหลายราย ซึ่งแต่ละรายก็สร้างความโดดเด่นให้กับผลิตภัณฑ์ของตน การที่องค์กรจะเลือกใช้ระบบการจัดการฐานข้อมูลให้เหมาะสมกับการทำงานและให้เกิดประโยชน์สูงสุดนั้น เริ่มต้นจากการวิเคราะห์ความต้องการสารสนเทศและคุณลักษณะของฐานข้อมูล ซึ่งจะทำให้ทราบว่ามีข้อมูลประเภทใดบ้างที่จะต้องนำมาจัดเก็บเพื่อการประมวลผลให้เป็นสารสนเทศตามความต้องการ และควรจะให้ฐานข้อมูลประเภทใดจัดเก็บข้อมูลเหล่านั้น เมื่อจัดเก็บแล้วทำให้ฐานข้อมูลมีคุณลักษณะอย่างไรบ้าง ส่งผลให้ทราบได้ว่าควรจะใช้ระบบการจัดการฐานข้อมูลชนิดใด สำหรับคุณลักษณะของฐานข้อมูลที่ใช้ในการเลือกระบบการจัดการฐานข้อมูล มีดังนี้

1.2.1 ขนาดของฐานข้อมูล ซึ่งขึ้นอยู่กับจำนวนแอตทริบิวต์หรือคอลัมน์ และจำนวนทูเพิลหรือเรคคอร์ดของข้อมูลที่จะจัดเก็บในฐานข้อมูล จำนวนของข้อมูลจะทำให้สามารถกำหนดได้ว่าควรใช้ระบบการจัดการฐานข้อมูลผลิตภัณฑ์ใดที่สามารถรองรับกับจำนวนข้อมูลที่มีปริมาณมากได้ เช่น ถ้าเป็นหน่วยงานขนาดเล็กอาจเลือกระบบการจัดการฐานข้อมูลประสิทธิภาพปานกลางที่ใช้ในเครื่องระดับไมโครคอมพิวเตอร์ เช่น ไมโครซอฟต์แอกเซส (Microsoft Access), ดีเบส (dBase) และฟอกซ์โปร (FoxPro) เป็นต้น แต่ถ้าเป็นหน่วยงานขนาดใหญ่ที่ใช้เครื่องมินิคอมพิวเตอร์ขึ้นไป ควรใช้ระบบการจัดการฐานข้อมูลที่มีความสามารถสูง เช่น ออราเคิล (Oracle) อินฟอร์มิกซ์ (Informix) อินเกรส (Ingress) และดีบีทู(DB2) เป็นต้น

1.2.2 จำนวนผู้ใช้เมื่อเกิดภาวะพร้อมกัน จำนวนของผู้ใช้ที่จำเป็นต้องเข้าถึงข้อมูลตัวเดียวกันในเวลาเดียวกันและฐานข้อมูลเดียวกัน เป็นปัจจัยที่สำคัญอย่างมากในการเลือกใช้ระบบการจัดการฐานข้อมูล เนื่องจากเป็นปัจจัยที่ช่วยให้องค์กรสามารถคัดเลือกระบบการจัดการฐานข้อมูลที่มีประสิทธิภาพในการควบคุมการทำงานในภาวะพร้อมกันได้ กล่าวคือ หากระบบการจัดการฐานข้อมูลใดที่รองรับการใช้งานของผู้ใช้หลายคนได้ ก็จะต้องสามารถควบคุมการเกิดภาวะพร้อมกันได้ระบบการจัดการฐานข้อมูลชนิดใดที่ไม่สามารถควบคุมได้ องค์กรก็ไม่รับไว้พิจารณา คำว่า “scalability” จะช่วยพิจารณาระบบการจัดการฐานข้อมูลได้ โดยทำให้องค์กรทราบว่าระบบการจัดการฐานข้อมูลสามารถรองรับกับปริมาณข้อมูลได้มากเท่าใด หรือสามารถควบคุมการเกิดภาวะพร้อมกันในการใช้งานของผู้ใช้ได้จำนวนเท่าใด

1.2.3 ประสิทธิภาพ ประสิทธิภาพของระบบการจัดการฐานข้อมูลอาจหมายถึงคุณลักษณะหลายอย่าง ความเร็วในการเข้าถึงข้อมูล เป็นเงื่อนไขสำคัญสำหรับองค์กรหรือบางส่วนงานที่ต้องการความเร็วในการเข้าถึงข้อมูล เช่น ระบบการจองตั๋วเครื่องบินออนไลน์ การอนุมัติบัตรเครดิต เป็นต้น แต่สำหรับบางระบบหรือบางองค์กร “ความเร็ว” อาจไม่ใช่เงื่อนไขสำคัญ เช่น ระบบบัญชีเงินเดือน ที่จะมีการประมวลผลไม่บ่อยครั้งนัก ใน 1 เดือน เป็นต้น อีกคุณลักษณะที่สามารถใช้เป็นเงื่อนไขในการเลือกระบบการจัดการฐานข้อมูลได้ก็คือ “ความสามารถและความเร็วในการกู้คืนระบบ (system recovery)“ ซึ่งจะเกิดขึ้นก็ต่อเมื่อระบบล้มเหลว ซึ่งหากการกู้คืนระบบเป็นไปอย่างเชื่องช้า อาจก่อให้เกิดความเสียหายหรืออาจทำให้บางหน่วยงานไม่สามารถ

ทำงานได้ นอกจากนี้ประสิทธิภาพยังรวมถึง ความสามารถในการควบคุมการเกิดพร้อมกัน ดังที่ได้กล่าวไปแล้วด้วย

1.2.4 การประสานการทำงาน เกณฑ์การเลือกใช้ระบบการจัดการฐานข้อมูลอีกประการหนึ่งที่สำคัญก็คือ “ความสามารถในการประสานการทำงานได้กับโปรแกรมประยุกต์และฐานข้อมูลอื่น” สิ่งที่จะบ่งชี้ได้ว่าระบบการจัดการฐานข้อมูลจะใช้งานร่วมกันได้หรือไม่ นอกจากจะเป็นชนิดของฐานข้อมูลแล้ว ยังมีอีกสิ่งหนึ่งที่สำคัญก็คือ “ระบบปฏิบัติการ (operation System)” ที่จะใช้ดำเนินการระบบการจัดการฐานข้อมูลเช่น ยูนิกซ์ (UNIX), วินโดวส์ (Windows) หรือ ลินุกซ์ (Linux) เป็นต้น หากเป็นระบบปฏิบัติการเดียวดันก็จะใช้ระบบการจัดการฐานข้อมูลและฐานข้อมูลร่วมกันได้ง่าย แต่หากเป็นระบบปฏิบัติการต่างชนิดกัน ก็ต้องพิจารณาว่าระบบการจัดการฐานข้อมูลชนิดใดที่มีเครื่องมือช่วยให้สามารถใช้ฐานข้อมูลร่วมกันได้ ทั้งจากภายในและภายนอกองค์กร

1.2.5 ลักษณะสำคัญ ไม่ว่าจะเป็นการทำงาน หรือรูปลักษณะของซอฟแวร์ที่เป็นเอกลักษณ์เฉพาะในระบบการจัดการฐานข้อมูลแต่ละผลิตภัณฑ์ ใช้เป็นเงื่อนไขในการพิจารณาเลือกใช้ระบบการจัดการฐานข้อมูลขององค์กรได้เป็นอย่างดีระบบการจัดการฐานข้อมูลโดยส่วนใหญ่มักจะมีเครื่องมือเพื่อช่วยในการกำหนดระบบรักษาความปลอดภัย การกำหนดสิทธิ์ในการเข้าถึงข้อมูลที่มาพร้อมกับระบบการจัดการฐานข้อมูลนอกจากนี้ลักษณะสำคัญอาจหมายความรวมถึง การใช้งานที่ง่าย มีคู่มือประกอบการใช้งานที่เข้าใจง่ายและชัดเจน

1.2.6 ผู้ผลิตและจัดจำหน่าย สำหรับการพิจารณาผู้ผลิตและจัดจำหน่ายระบบการจัดการฐานข้อมูลอาจพิจารณาได้จากขนาดขององค์กร ความมีชื่อเสียง และสถานะทางการเงิน หากผู้ผลิตและจัดจำหน่ายรายใดที่อยู่ในวงการทำงานด้านเทคโนโลยีสารสนเทศมานาน ก็จะมีประสบการณ์มาก สามารถให้คำปรึกษาและแก้ปัญหาที่เกิดขึ้นได้ดี หรือหากผู้ผลิตและจัดจำหน่ายรายใด เป็นองค์กรที่มีขนาดใหญ่ ก็อาจมีแผนกให้บริการหลังการขายเป็นสัดส่วนชัดเจน สามารถช่วยเหลือลูกค้าได้อย่างเต็มที่ตลอด 24 ชั่วโมง ซึ่งประเด็นการบริการหลังการขาย นับว่าเป็นปัจจัยสำคัญในการพิจารณาเลือกระบบการจัดการฐานข้อมูล

1.2.7 งบประมาณ เงื่อนไขในการเลือกใช้ระบบการจัดการฐานข้อมูลที่สำคัญมากถึงมากที่สุดก็คือ “ต้นทุน (Cost)” หรือราคาของระบบการจัดการฐานข้อมูลซึ่งจะแตกต่างกันไปตามผู้ผลิตและจัดจำหน่ายแต่ละรายระบบการจัดการฐานข้อมูลสำหรับเครื่องไมโครคอมพิวเตอร์จะมีต้นทุนต่ำกว่าระบบการจัดการฐานข้อมูลสำหรับเครื่องเมนเฟรมคอมพิวเตอร์ บางองค์กรเลือกที่จะเช่าระบบการจัดการฐานข้อมูลเป็นรายเดือน หรือรายปี ทั้งนี้ขึ้นอยู่กับสถานะทางการเงิน นโยบาย และเงื่อนไขอื่นๆ ในขณะนั้น

2. การประเมินผลฮาร์ดแวร์และซอฟต์แวร์ในการจัดการฐานข้อมูล

ฮาร์ดแวร์และซอฟต์แวร์ในการจัดการฐานข้อมูลมีหลายประเภท เมื่อตัดสินใจเลือกใช้งานแล้วจำเป็นต้องมีการประเมินการใช้งานเพื่อให้ได้ฮาร์ดแวร์และซอฟต์แวร์ที่มีสมรรถนะ และราคาที่เหมาะสม ตรงกับความต้องการ

2.1 การประเมินผลฮาร์ดแวร์ควรพิจารณาจากปัจจัยดังต่อไปนี้

2.1.1 สมรรถนะ ความเร็ว ความจุ และปริมาณงานเป็นเท่าใด

2.1.2 ต้นทุน ราคาเช่าซื้อหรือราคาซื้อเท่าไร ต้นทุนของการดำเนินงานและการบำรุงรักษาเท่าไร

2.1.3 ความเชื่อถือ อะไรคือความเสี่ยงในการขัดข้องของงานและความต้องการในเรื่องการบำรุงรักษา

2.1.4 สภาพพร้อมใช้งาน เมื่อไรที่เริ่มใช้งานได้

2.1.5 ความเข้ากันได้ เข้ากันได้กับฮาร์ดแวร์และซอฟต์แวร์ที่มีอยู่เดิมหรือไม่ใช้ได้กับฮาร์ดแวร์และซอฟต์แวร์ที่ขายโดยผู้ขายที่มีการแข่งขันหรือไม่

2.1.6 ภาวะเชื่อมต่อ เชื่อมต่อได้โดยง่ายกับทั้งเครือข่ายท้องถิ่นและเครือข่ายบริเวณกว้างของคอมพิวเตอร์ประเภทต่างๆและอุปกรณ์ต่อพ่วงหรือไม่

2.1.7 การปรับขนาดได้สามารถจัดการความต้องการประมวลผลของผู้ใช้จำนวนมาก รายการเปลี่ยนแปลง การสอบถาม และความต้องการการประมวลผลสารสนเทศหรือไม่

2.1.8 ซอฟต์แวร์ระบบและโปรแกรมประยุกต์ซอฟต์แวร์สามารถใช้ได้ดีที่สุดกับฮาร์ดแวร์นี้หรือไม่

2.1.9 การสนับสนุน มีบริการที่ต้องการสำหรับการสนับสนุนและบำรุงรักษาหรือไม่

2.2 การประเมินผลซอฟต์แวร์ควรพิจารณาจากปัจจัยดังต่อไปนี้

2.2.1 ประสิทธิภาพ เป็นซอฟต์แวร์ที่มีการพัฒนาคำสั่งคอมพิวเตอร์หรืออ็อบเจ็กต์ที่ดีซึ่งใช้หน่วยความจำหรือเวลาของหน่วยประมวลผลกลางมากหรือไม่

2.2.2 ความยืดหยุ่น สามารถจัดการงานประมวลผลที่ได้รับมอบหมายได้โดยง่ายโดยไม่ต้องมีการแก้ไขมากมายหรือไม่

2.2.3 ความปลอดภัย มีกระบวนคำสั่งควบคุมสำหรับข้อผิดพลาด การทำหน้าที่ผิดปกติและการใช้ที่ไม่เหมาะสมหรือไม่

2.2.4 ภาวะเชื่อมต่อ มีความสามารถด้านเครือข่ายซึ่งสามารถเข้าถึงอินเทอร์เน็ตอินทราเน็ต เอ็กซ์ทราเน็ต และเครือข่ายอื่นๆได้โดยง่าย หรือสามารถทำงานกับบราวเซอร์เครือข่ายหรือซอฟต์แวร์เครือข่ายอื่นๆได้หรือไม่

2.2.5 เอกสาร ซอฟต์แวร์มีเอกสารที่ดีหรือไม่ รวมทั้งมีวิธีการใช้ที่ช่วยเหลือผู้ใช้หรือไม่

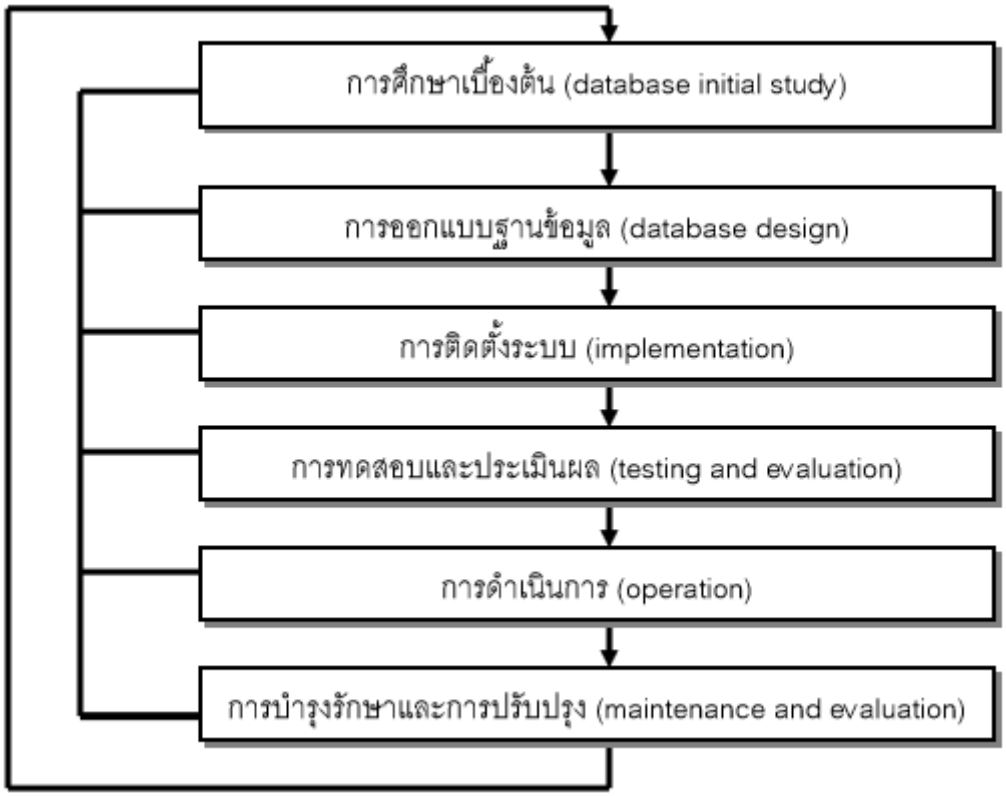
2.2.6 ฮาร์ดแวร์ ฮาร์ดแวร์ที่มีอยู่มีคุณสมบัติที่ต้องการซึ่งสามารถใช้ซอฟต์แวร์นี้ได้ดีที่สุดหรือไม่

<u>การออกแบบฐานข้อมูล</u>

ฐานข้อมูลที่มีประสิทธิภาพและตรงกับความต้องการของผู้ใช้นั้น ต้องอาศัยการพัฒนาฐานข้อมูลที่มีการวางแผนอย่างเป็นระบบและมีขั้นตอนที่ถูกต้อง โดยในบทนี้จะศึกษาถึงขั้นตอนการพัฒนาระบบฐานข้อมูล ซึ่งจะเน้นขั้นตอนการออกแบบฐานข้อมูลเป็นหลัก ตั้งแต่การพัฒนาแบบจ าลองอี-อาร์และการท านอร์มัลไลเซชัน เพื่อให้ได้ฐานข้อมูลที่สมบูรณ์ที่สุด

ขั้นตอนการพัฒนาระบบฐานข้อมูล

วัฏจักรฐานข้อมูล (The Database Life Cycle: DBLC) เป็นขั้นตอนในการพัฒนาหรือจัดทำระบบฐานข้อมูล ซึ่งประกอบด้วย 6 ขั้นตอน ดังภาพ



1. การศึกษาเบื้องต้น

การศึกษาเบื้องต้นมีวัตถุประสงค์เพื่อวิเคราะห์สถานการณ์ขององค์กร กำหนดปัญหาและข้อจำกัด กำหนดวัตถุประสงค์และขอบเขตของระบบ ดังรายละเอียดต่อไปนี้

1.1 การวิเคราะห์สถานการณ์ขององค์กร

เพื่อศึกษาสภาพแวดล้อมในการทำงานขององค์กร ความต้องการในการปฏิบัติงานโดยควรรู้ว่าโครงสร้างขององค์กรเป็นอย่างไร ใครเป็นผู้ควบคุมอะไร และใครทำรายงานให้ใครเป็นต้น

1.2 กำหนดปัญหาและข้อจำกัด

โดยการศึกษาว่า ระบบที่มีอยู่มีการทำงานอย่างไร ข้อมูลที่ป้อนเข้าสู่ระบบมีอะไรบ้าง และระบบสร้างรายงานอะไร มีการใช้รายงานเหล่านี้อย่างไรและใครเป็นผู้ใช้ เพื่อให้ทราบถึงปัญหาและข้อจำกัดในการป้อนข้อมูล หรือการค้นหาข้อมูลเพื่อการทำรายงาน

1.3 กำหนดวัตถุประสงค์และขอบเขตของระบบในการกำหนดวัตถุประสงค์ของระบบฐานข้อมูลควรสอดคล้องกับความต้องการของผู้ใช้ จากคำถามเหล่านี้

- วัตถุประสงค์แรกเริ่มของระบบที่นำเสนอคืออะไร

- ระบบนี้ต้องเชื่อมต่อกับระบบอื่นๆ ที่มีอยู่ในองค์กรหรือไม่

- ระบบนี้จะมีการใช้ข้อมูลร่วมกันกับระบบหรือผู้ใช้อื่นหรือไม่

เมื่อทราบวัตถุประสงค์แล้วก็ทำการกำหนดขอบเขตของระบบโดยการออกแบบตามความต้องการในการปฏิบัติงาน เพื่อใช้ในการออกแบบฐานข้อมูลต่อไป

2. การออกแบบฐานข้อมูล

เมื่อผู้ออกแบบฐานข้อมูลมีความเข้าใจลักษณะขององค์กร ปัญหาและข้อจำกัด รวมทั้งวัตถุประสงค์และขอบเขตของระบบแล้ว ก็ทำการออกแบบฐานข้อมูลดังต่อไปนี้

2.1 การออกแบบเชิงแนวคิด

โดยการพัฒนาแบบจำลองอี-อาร์(E-R Model) ที่ใช้อธิบายถึงความสัมพันธ์ระหว่างสิ่งที่เราสนใจจะจัดเก็บ ที่เรียกว่า เอนทิตี(entity) และรายละเอียดหรือคุณสมบัติ (attribute) ของสิ่งที่จะจัดเก็บ แล้วทำการแปลงแบบจำลองอี-อาร์ เป็นโครงสร้างตารางฐานข้อมูล จากนั้นก็ทำการนอร์มัลไลเซชัน (normalization) เพื่อให้ได้โครงสร้างของตารางที่ดี สามารถควบคุมความซ้ำซ้อนของข้อมูล หลีกเลี่ยงความผิดปกติของข้อมูล

2.2 การเลือกโปรแกรมจัดการฐานข้อมูลในการตัดสินใจเลือกซื้อโปรแกรมจัดการฐานข้อมูลขององค์กรใด ควรพิจารณาถึงสิ่งต่อไปนี้

2.2.1 ค่าใช้จ่ายต่างๆ เช่น ราคาการซ่อมบำรุง การปฏิบัติงาน ลิขสิทธิ์ การติดตั้งการฝึกอบรม และค่าใช้จ่ายในการเปลี่ยนไปใช้ระบบใหม่

2.2.2 คุณลักษณะและเครื่องมือของระบบจัดการฐานข้อมูล โปรแกรมฐานข้อมูลบางตัวจะรวมเอาเครื่องมือต่างๆ ที่ให้ความสะดวกในงานการพัฒนาโปรแกรมประยุกต์ ตัวอย่างเช่นการออกแบบหน้าจอ การสร้างรายงาน การสร้างโปรแกรมประยุกต์ และพจนานุกรมข้อมูล เป็นต้นทำให้สะดวกในการบริหารฐานข้อมูลใช้ง่าย มีความสามารถในการรักษาความปลอดภัย และการควบคุมการใช้งานพร้อมกัน เป็นต้น

2.2.3 ความสามารถในการใช้ข้าม platforms ข้ามระบบและภาษา

2.2.4 ความต้องการด้านฮาร์ดแวร์ หน่วยความจำ และเนื้อที่ที่ใช้ในการจัดเก็บ

2.3 การออกแบบทางตรรกะ

จะเกี่ยวข้องกับการตัดสินใจใช้รูปแบบเฉพาะของฐานข้อมูล (แบบลำดับชั้น แบบเครือข่าย และแบบเชิงสัมพันธ์ เป็นต้น) การกำหนดรูปแบบของฐานข้อมูล ซึ่งการออกแบบเชิงตรรกะจะเป็นการแปลงการออกแบบระดับเชิงแนวคิด ให้เป็นแบบจำลองของฐานข้อมูลในระดับภายใน (internal model) ตามระบบการจัดการฐานข้อมูล (DBMS) เช่น MS-Access และ Oracle โดยการสร้างตาราง ฟอร์ม คิวรี และรายงาน เป็นต้น

2.4 การออกแบบทางกายภาพ

การออกแบบทางกายภาพ คือ กระบวนการในการเลือกหน่วยจัดเก็บข้อมูล และลักษณะการเข้าถึงข้อมูลของฐานข้อมูล การสร้างดรรชนี (index) การจัดทำคลัสเตอร์ (clustering) ซึ่งเป็นการจัดเก็บข้อมูลที่มีการใช้งานบ่อยๆ ไว้ในหน่วยเก็บข้อมูลเดียวกัน หรือการใช้เทคนิคแฮชชิง (hashing technique) ในการจัดตำแหน่งที่อยู่ของข้อมูลภายในหน่วยเก็บข้อมูล เป็นต้น

3. การติดตั้งระบบ

ขึ้นอยู่กับระบบจัดการฐานข้อมูลที่ใช้ โดยเริ่มต้นจากการสร้างฐานข้อมูล กำหนดผู้จัดการฐานข้อมูล กำหนดพื้นที่ๆ ที่ต้องการใช้ และการสร้างตารางต่างๆ ในระบบ

4. การทดสอบและประเมินผล

เพื่อการตรวจสอบดูว่าระบบที่พัฒนามาสามารถทำงานได้ตามที่ต้องการหรือไม่ซึ่งควรมีการเตรียมข้อมูลทดสอบไว้ล่วงหน้า

5. การดำเนินการ

เมื่อฐานข้อมูลผ่านขั้นตอนการทดสอบและประเมินผล ต่อไปก็เป็นขั้นตอนการดำเนินการ หรือการติดตั้งระบบ ซึ่งต้องเป็นระบบที่สมบูรณ์พร้อมให้ผู้ใช้ได้ใช้งานนั่นเอง ซึ่งอาจรวมไปถึงการฝึกอบรมให้แก่ผู้ใช้ ที่เป็นพนักงานที่ต้องใช้งานจริงด้วย

6. การบำรุงรักษาและการปรับปรุง

หลังจากระบบได้เริ่มดำเนินการ ผู้จัดการฐานข้อมูลจะต้องเตรียมการบำรุงรักษาฐานข้อมูลโดยการสำรองข้อมูลไว้ เพื่อสะดวกในการกู้คืนข้อมูล เมื่อระบบมีปัญหา และหากมีการใช้งานไปนานๆ อาจต้องทำการปรับปรุงแก้ไขโปรแกรมให้เหมาะสมกับเหตุการณ์ และความต้องการของผู้ใช้ที่เปลี่ยนแปลงไป

เมื่อเราทราบขั้นตอนการพัฒนาระบบฐานข้อมูลทั้งหมดแล้ว ในบทนี้จะเน้นถึงรายละเอียดของขั้นตอนการออกแบบฐานข้อมูลเชิงแนวคิด ตั้งแต่การพัฒนาแบบจ าลองอี-อาร์ (E-R Model) และการทำนอร์มัลไลเซชัน (normalization) ดังรายละเอียดที่จะกล่าวต่อไป

แบบจำลองอี-อาร์

แบบจำลองอี-อาร์ (Entity-Relationship Model: E-R Model) เป็นแบบจำลองข้อมูลที่ประยุกต์มาจากแนวคิดเรื่อง Semantic Model และมีการพัฒนามาเป็น E-R Model โดย Peter Pin Shan Chen จาก Massachusetts Institute of Technology ในปี ค.ศ. 1976 และได้รับความนิยมมาจนถึงปัจจุบัน

1. ความหมายและความสำคัญของแบบจำลองอี-อาร์

แบบจำลองอี-อาร์ เป็นเครื่องมือที่ใช้ในการออกแบบฐานข้อมูล ที่แสดงความสัมพันธ์ระหว่างเอนทิตี หรือสิ่งที่เราต้องการจะจัดเก็บไว้ในฐานข้อมูล โดยนำเสนอในรูปของของแผนภาพ ที่เรียกว่า อี-อาร์ไดอะแกรม (E-R Diagram) ด้วยการใช้สัญลักษณ์ต่างๆ

แบบจำลองอี-อาร์ มีความสำคัญในการเป็นสื่อกลางเพื่อสื่อสารกับบุคลากรต่างๆที่เกี่ยวข้องกับระบบฐานข้อมูล ไม่ว่าจะเป็นในระดับผู้บริหาร นักเขียนโปรแกรม และผู้ใช้ในระดับปฏิบัติการ เป็นต้น ทำให้เข้าใจระบบได้อย่างถูกต้องตรงกัน เนื่องจากมีการแสดงภาพรวมของระบบในลักษณะของรูปภาพหรือแผนภาพ ทำให้เข้าใจง่าย ดังนั้นระบบที่ออกแบบมาจึงมีความถูกต้องและเป็นไปตามวัตถุประสงค์ขององค์กร

2. องค์ประกอบของแบบจำลองอี-อาร์

แบบจำลองอี-อาร์ ประกอบด้วย เอนทิตีแอตทริบิวต์ คีย์ และความสัมพันธ์ดังรายละเอียดต่อไปนี้

2.1 เอนทิตี

เอนทิตี(entity) คือ สิ่งต่างๆ ที่ผู้ใช้งานฐานข้อมูลต้องการจะจัดเก็บ ซึ่งมีลักษณะเป็นคำนาม ทั้งรูปธรรมและนามธรรม เช่น บุคคล สถานที่ วัตถุสิ่งของ และเหตุการณ์ต่างๆ เป็นต้นตัวอย่างของเอนทิตีใน “ระบบการลงทะเบียนเรียนของนักศึกษา” ประกอบด้วย รายวิชา นักศึกษาการลงทะเบียน ผลการเรียนประจำเทอม สาขาวิชา คณะ และหลักสูตร เป็นต้น

เอนทิตี้ที่รวบรวมได้จากระบบสามารถแยกแยะและจัดเป็นหมวดหมู่ได้ตามชนิดของเอนทิตีได้ดังต่อไปนี้

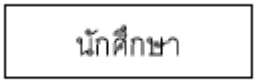
- หมวดบุคคล ได้แก่ เอนทิตี้ -> นักศึกษา พนักงาน ประชาชน ผู้ป่วย และลูกค้าเป็นต้น

- หมวดสถานที่ ได้แก่ เอนทิตี้ -> รัฐ ประเทศ จังหวัด ภาค สาขา และวิทยาเขตเป็นต้น

- หมวดวัตถุ ได้แก่ เอนทิตี้ -> อาคาร เครื่องจักร ผลผลิต หนังสือ วัตถุดิบ และรถยนต์ เป็นต้น

- หมวดเหตุการณ์ ได้แก่ เอนทิตี้-> การขาย การลงทะเบียน การเดินทางการสั่งซื้อของ การออกใบเสร็จรับเงิน และการให้รางวัล เป็นต้น

ในอี-อาร์ไดอะแกรม ใช้สัญลักษณ์รูปสี่เหลี่ยมผืนผ้า แทนหนึ่งเอนทิตีโดยใช้ชื่อของเอนทิตีนั้นๆ ก ากับอยู่ภายใน เช่น



แทน เอนทิตี**นักศึกษา**

2.2 แอตทริบิวต์

แอตทริบิวต์ (attribute) คือ คุณสมบัติต่างๆ ของเอนทิตีที่เราต้องการจัดเก็บในฐานข้อมูล ตัวอย่างเช่น

- เอนทิตีบัตรประชาชน ประกอบด้วยแอตทริบิวต์ หรือสิ่งที่บ่งบอกคุณสมบัติของประชาชนแต่ละคน ได้แก่ หมายเลขบัตรประชาชน ชื่อ นามสกุล วันเดือนปีเกิด ภูมิลำเนา วันที่ออกบัตร วันที่บัตรหมดอายุ ส่วนสูง น้ำหนัก และกรุ๊ปเลือด เป็นต้น

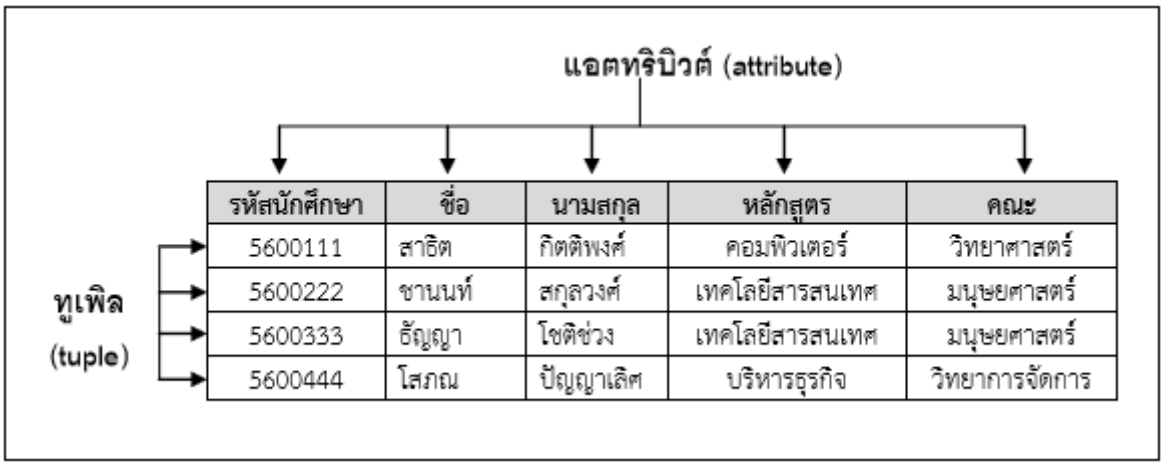
- เอนทิตีพนักงาน ประกอบด้วยแอตทริบิวต์ ได้แก่ รหัสพนักงาน ชื่อ นามสกุล ที่อยู่ เบอร์โทรศัพท์ สถานภาพสมรส และเงินเดือน เป็นต้น

- เอนทิตีสินค้า ประกอบด้วยแอตทริบิวต์ ได้แก่ รหัสสินค้า ชื่อสินค้า ราคา และจำนวน เป็นต้น

- เอนทิตีนักศึกษา ประกอบด้วยแอตทริบิวต์ ได้แก่ รหัสนักศึกษา ชื่อ นามสกุลเพศ วันเดือนปีเกิด ที่อยู่ และเบอร์โทรศัพท์ เป็นต้น

- เอนทิตีวิชา ประกอบด้วยแอตทริบิวต์ได้แก่ รหัสวิชา ชื่อวิชา และจำนวนหน่วยกิต เป็นต้น

ค่าของข้อมูลในแต่ละแอตทริบิวต์ประกอบกัน เรียกว่า ทูเพิล (tuple) ซึ่งเป็นแถวของข้อมูลในตาราง โดยแต่ละแถวหรือแต่ละทูเพิลจะประกอบด้วยหลายแอตทริบิวต์หรือหลายคอลัมน์ของข้อมูล จำนวนแถวของข้อมูลในตารางเรียกว่า Cardinality และจำนวนแอตทริบิวต์ทั้งหมดในตารางเรียกว่า Degree อย่างเช่น จากภาพที่ 6.2 มี 4 Cardinality 5 Degree



รหัสนักศึกษา	ชื่อ	นามสกุล	หลักสูตร	คณะ
5600111	สาธิต	กิตติพงศ์	คอมพิวเตอร์	วิทยาศาสตร์
5600222	ชานนท์	สกุลวงศ์	เทคโนโลยีสารสนเทศ	มนุษยศาสตร์
5600333	ธัญญา	โชติช่วง	เทคโนโลยีสารสนเทศ	มนุษยศาสตร์
5600444	โสภณ	ปัญญาเลิศ	บริหารธุรกิจ	วิทยาการจัดการ

ในอี-อาร์ไดอะแกรม ใช้สัญลักษณ์รูปวงรีแทนหนึ่งแอตทริบิวต์ โดยใช้ชื่อของแอตทริบิวต์นั้นๆ ก ากับอยู่ภายใน เช่น

แทนแอตทริบิวต์ของ**ชื่อ**

2.3 คีย์

คีย์ (key) คือ แอตทริบิวต์ที่สามารถใช้บ่งบอกความแตกต่างของแต่ละทูเพิลได้ อาจเป็นแอตทริบิวต์เดี่ยวๆ หรือ กลุ่มของแอตทริบิวต์ก็ได้

ประเภทของคีย์ประกอบด้วย

2.3.1 ซุปเปอร์คีย์ (super key) คือ แอตทริบิวต์หรือกลุ่มของแอตทริบิวต์ที่สามารถบ่งบอกความแตกต่างของแต่ละทูเพิลได้

ตารางที่ 6.1 ข้อมูลในเอนทิตีนักศึกษา

รหัสนักศึกษา	ชื่อ	นามสกุล	เลขที่บัตรประชาชน
56001	สามารถ	ประเสริฐกุล	3120100475991
56002	วิชา	ปัญญาเลิศ	3120100475992
56003	น้ำทิพย์	วิเศษศิริ	3120100475993

จากตารางที่ 6.1 ประกอบไปด้วยซุปเปอร์คีย์ดังต่อไปนี้

- รหัสนักศึกษา

- รหัสนักศึกษา, ชื่อ

- รหัสนักศึกษา, ชื่อ, นามสกุล

- เลขที่บัตรประชาชน

2.3.2 คีย์คู่แข่ง (candidate key) คือ ซุปเปอร์คีย์ที่น้อยที่สุด ที่สามารถบ่งบอกความแตกต่างของแต่ละทูเพิลได้จากตารางที่ 6.1 ประกอบไปด้วยคีย์คู่แข่งดังต่อไปนี้

– รหัสนักศึกษา

– เลขที่บัตรประชาชน

2.3.2 คีย์หลัก (primary key) คือ คีย์คู่แข่งที่ถูกเลือก เพื่อใช้บ่งบอกความแตกต่างของแต่ละทูเพิล

จากตารางที่ 6.1 คีย์หลัก คือ รหัสนักศึกษา หรือเลขที่บัตรประชาชนอย่างใดอย่างหนึ่ง

คุณสมบัติของคีย์หลัก

1) คีย์หลักซ้ำกันไม่ได้

2) คีย์หลักอาจเป็นแค่หนึ่งแอตทริบิวต์หรือกลุ่มของแอตทริบิวต์ก็ได้อย่างเช่นในตารางที่ 6.1 มีแอตทริบิวต์เดียวที่เป็นคีย์หลัก ซึ่งอาจจะเป็น “รหัสนักศึกษา” หรือ “เลขที่บัตรประชาชน” ก็ได้ แต่ข้อมูลบางตารางอาจต้องอาศัยแอตทริบิวต์ตั้งแต่ 2 ตัวขึ้นไปมาประกอบกันเป็นคีย์หลัก เพื่อให้เกิดความแตกต่างระหว่างทูเพิล ดังเช่นในตารางที่ 6.2

ตารางที่ 6.2 ข้อมูลในเอนทิตีการลงทะเบียนเรียนของนักศึกษา ปีการศึกษา 1/56

รหัสนักศึกษา	**ชื่อ**	**รหัสวิชา**	**ชื่อวิชา**	**ปีการศึกษา**
5620249001	ปรีชา	111	คอมพิวเตอร์	1/56
5620249001	ปรีชา	222	ภาษาไทย	1/56
5620249002	เกรียงไกร	111	คอมพิวเตอร์	1/56
5620249003	ฉัตรชัย	333	ภาษาอังกฤษ	1/56
5620249003	ฉัตรชัย	222	ภาษาไทย	1/56
5620249003	ฉัตรชัย	444	สิ่งแวดล้อม	1/56

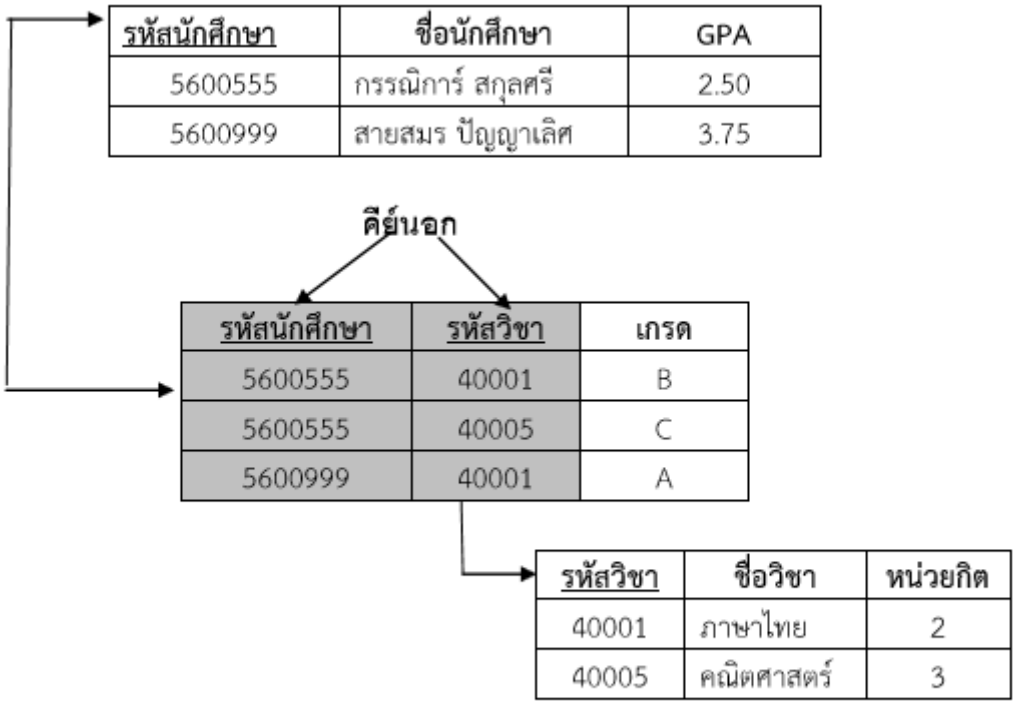
จากตารางที่ 6.2 ไม่สามารถให้แอตทริบิวต์รหัสนักศึกษา เป็นคีย์หลักเพียงแอตทริบิวต์เดียวได้ เพราะจะเห็นว่า รหัสนักศึกษา 5620249001 ของทูเพิลหรือแถวที่ 1 จะไปซ้ำกับแถวที่ 2 แต่ถ้าให้แอตทริบิวต์ “รหัสนักศึกษา” และ “รหัสวิชา” เป็นคีย์หลัก แล้วพิจารณาข้อมูลของ 2 แอตทริบิวต์นี้ จะเห็นว่าข้อมูลไม่ซ้ำกันแล้ว ดังนั้นตารางที่ 6.2 จึงมีคีย์หลักซึ่งประกอบด้วยแอตทริบิวต์ 2 ตัวประกอบกัน คือ “รหัสนักศึกษา” และ “รหัสวิชา”

3) คีย์หลักจะเป็นค่าว่าง (null) ไม่ได้ เพราะฉะนั้นในการกรอกข้อมูลต่างๆ ลงในตาราง แอตทริบิวต์ใดที่เรากำหนดให้เป็นคีย์หลักต้องกรอกข้อมูลให้ครบ คือ จะไม่มีค่าไม่ได้ แต่แอตทริบิวต์อื่นอาจจะปล่อยเว้นว่างไว้ก็ได้ถ้าไม่ทราบค่า

2.3.4 คีย์นอก (foreign key) คือ แอตทริบิวต์ที่ใช้ในการเชื่อมต่อกับเอนทิตีอื่นๆ เพื่อแสดงความสัมพันธ์

คุณสมบัติของคีย์นอก คือ

- คีย์นอกสามารถมีค่าซ้ำกันได้
- คีย์นอกสามารถเป็นค่าว่างได้
- คีย์นอกที่ไม่เป็นค่าว่างจะเป็นค่าที่ชี้ไปยังคีย์หลักของเอนทิตีที่สัมพันธ์กัน



ภาพที่ 6.3 แอตทริบิวต์ที่เป็นคีย์รองที่ใช้ในการเชื่อมต่อกับเอนทิตีอื่น

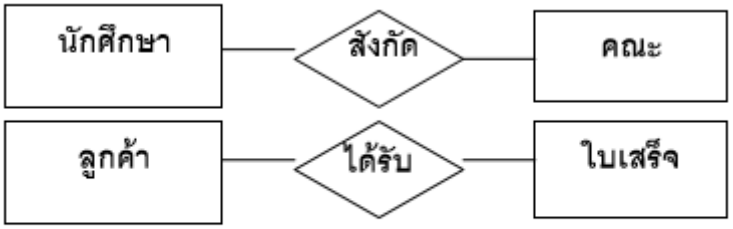
2.3.5 คีย์รอง (secondary key) คือ แอตทริบิวต์ที่ไม่เป็น key หลัก แต่สามารถใช้ในการค้นหาข้อมูลนั้นๆ ได้ โดยคีย์รองจะมีค่าซ้ำกันได้ตัวอย่างเช่น ในตารางที่ 6.3 มีรหัสนักศึกษาเป็นคีย์หลัก แต่หากต้องการค้นหาข้อมูลจากชื่อนักศึกษา แอตทริบิวต์ชื่อก็จะเป็นคีย์รอง หรือถ้าต้องการค้นหาข้อมูลจากนามสกุลนักศึกษา แอตทริบิวต์นามสกุลก็จะเป็นคีย์รอง เป็นต้น

ตารางที่ 6.3 คีย์รองที่ใช้ในการค้นหาข้อมูล

รหัสนักศึกษา	ชื่อ	นามสกุล	เบอร์โทรศัพท์
56111	สามารถ	ประเสริฐกุล	0-1111-1111
56112	วิชา	ปัญญาเลิศ	0-2222-2222
56113	น้ำทิพย์	วิเศษศิริ	0-5555-5555
56114	สมจิตร์	สมสกุลวงศ์	0-6666-6666
56115	วิชา	รักศักดิ์ศรี	0-9999-9999

2.4 ความสัมพันธ์

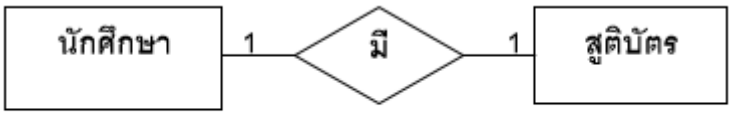
ความสัมพันธ์ (relationship) เป็นการอธิบายความสัมพันธ์ระหว่างเอนทิตีที่มีความความสัมพันธ์กันว่ามีความสัมพันธ์กันอย่างไร โดยในอี-อาร์ไดอะแกรมใช้สัญลักษณ์รูปสี่เหลี่ยมข้าวหลามตัด ที่มีชื่อของความสัมพันธ์นั้นกำกับอยู่ภายใน และเชื่อมต่อกับเอนทิตีที่เกี่ยวข้องกับความสัมพันธ์ด้วยเส้นตรง ดังตัวอย่างด้านล่าง



ภาพที่ 6.4 ตัวอย่างความสัมพันธ์ระหว่างเอนทิตี

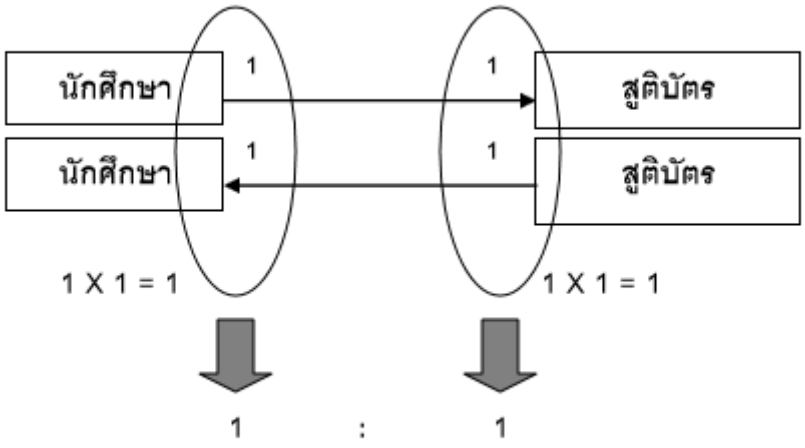
ความสัมพันธ์ระหว่างเอนทิตีแบ่งเป็น 3 ประเภท คือ ความสัมพันธ์แบบหนึ่งต่อหนึ่ง ความสัมพันธ์แบบหนึ่งต่อกลุ่ม และความสัมพันธ์แบบกลุ่มต่อกลุ่ม ดังรายละเอียดต่อไปนี้

2.4.1 ความสัมพันธ์แบบหนึ่งต่อหนึ่ง (one to one relationship หรือ 1:1) หมายถึง ข้อมูลในเอนทิตีหนึ่ง มีความสัมพันธ์กับข้อมูลในอีกหนึ่งเอนทิตีเพียงข้อมูลเดียว ตัวอย่างเช่นนักศึกษาแต่ละคนจะมีสูติบัตรได้เพียงใบเดียวเท่านั้น และสูติบัตรหนึ่งใบก็เป็นของนักศึกษาได้เพียงคนเดียวเท่านั้นเช่นกัน



ภาพที่ 6.5 ความสัมพันธ์ของข้อมูลแบบหนึ่งต่อหนึ่ง

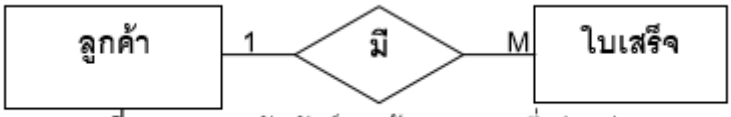
ในการพิจารณาความสัมพันธ์ระหว่างเอนทิตีแบบหนึ่งต่อหนึ่ง ต้องมองสองทิศ คือ มองจากซ้ายไปขวา และก็ต้องมองจากขวาไปซ้าย แล้วจึงนำความสัมพันธ์ทั้งสองทิศ มาพิจารณารวมกัน ดังภาพที่ 6.6



ภาพที่ 6.6 วิธีการพิจารณาความสัมพันธ์แบบหนึ่งต่อหนึ่ง

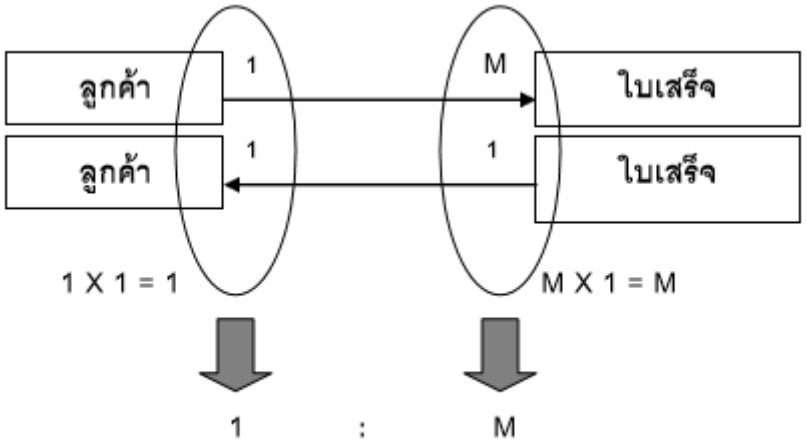
2.4.2 ความสัมพันธ์แบบหนึ่งต่อกลุ่ม (one to many relationship หรือ 1:M) หมายถึง ข้อมูลในเอนทิตีหนึ่ง มีความสัมพันธ์กับข้อมูลในอีกหนึ่งเอนทิตีมากกว่าหนึ่งข้อมูลตัวอย่างเช่น ลูกค้าหนึ่งคนมีใบเสร็จได้

หลายใบ เนื่องจากลูกค้าหนึ่งคนอาจมาซื้อสินค้าหลายครั้ง แต่ใบเสร็จหนึ่งใบต้องเป็นของลูกค้าเพียงคนเดียวเท่านั้น



ภาพที่ 6.7 ความสัมพันธ์ของข้อมูลแบบหนึ่งต่อกลุ่ม

ในการพิจารณาความสัมพันธ์ระหว่างเอนทิตีแบบหนึ่งต่อกลุ่ม ต้องมองสองทิศ คือ มองจากซ้ายไปขวาและก็ต้องมองจากขวาไปซ้าย แล้วจึงน าความสัมพันธ์ทั้งสองทิศ มาพิจารณารวมกัน ดังภาพที่ 6.8

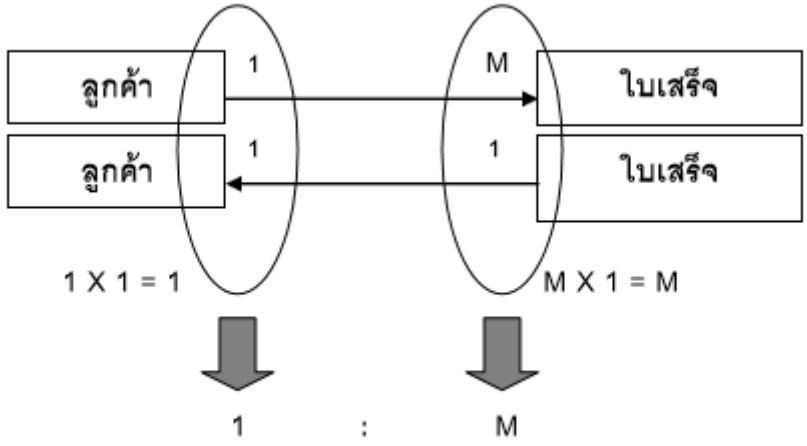


ภาพที่ 6.8 วิธีการพิจารณาความสัมพันธ์แบบหนึ่งต่อกลุ่ม

2.4.2 ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (many to many relationship หรือ M:M) หมายถึง ข้อมูลมากกว่าหนึ่งข้อมูลในเอนทิตีหนึ่ง มีความสัมพันธ์กับข้อมูลในอีกหนึ่งเอนทิตีมากกว่าหนึ่งข้อมูล ตัวอย่างเช่น นักศึกษาหนึ่งคนสามารถลงทะเบียนเรียนได้หลายวิชา และวิชาแต่ละวิชามีนักศึกษาลงทะเบียนเรียนได้หลายคน

ภาพที่ 6.7 ความสัมพันธ์ของข้อมูลแบบหนึ่งต่อกลุ่ม

ในการพิจารณาความสัมพันธ์ระหว่างเอนทิตีแบบหนึ่งต่อกลุ่ม ต้องมองสองทิศ คือ มองจากซ้ายไปขวาและก็ต้องมองจากขวาไปซ้าย แล้วจึงนำความสัมพันธ์ทั้งสองทิศ มาพิจารณารวมกัน ดังภาพที่ 6.8

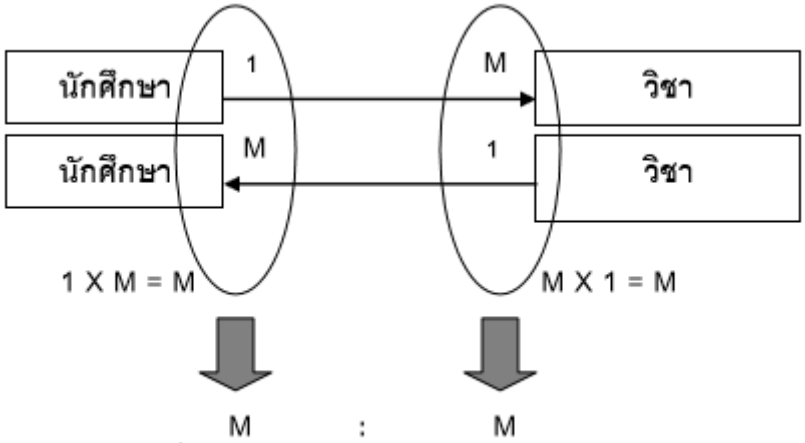


ภาพที่ 6.8 วิธีการพิจารณาความสัมพันธ์แบบหนึ่งต่อกลุ่ม

2.4.3 ความสัมพันธ์แบบกลุ่มต่อกลุ่ม (many to many relationship หรือ M:M) หมายถึง ข้อมูลมากกว่าหนึ่งข้อมูลในเอนทิตีหนึ่ง มีความสัมพันธ์กับข้อมูลในอีกหนึ่งเอนทิตีมากกว่าหนึ่งข้อมูล ตัวอย่างเช่น นักศึกษาหนึ่งคนสามารถลงทะเบียนเรียนได้หลายวิชา และวิชาแต่ละวิชามีนักศึกษาลงทะเบียนเรียนได้หลายคน

ภาพที่ 6.9 ความสัมพันธ์ของข้อมูลแบบกลุ่มต่อกลุ่ม

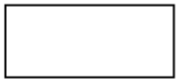
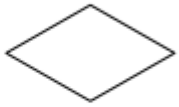
ในการพิจารณาความสัมพันธ์ระหว่างเอนทิตีแบบกลุ่มต่อกลุ่ม ต้องมองสองทิศ คือ มองจากซ้ายไปขวา และก็ต้องมองจากขวาไปซ้าย แล้วจึงนำความสัมพันธ์ทั้งสองทิศ มาพิจารณารวมกัน ดังภาพที่ 6.10



ภาพที่ 6.10 วิธีการพิจารณาความสัมพันธ์แบบกลุ่มต่อกลุ่ม

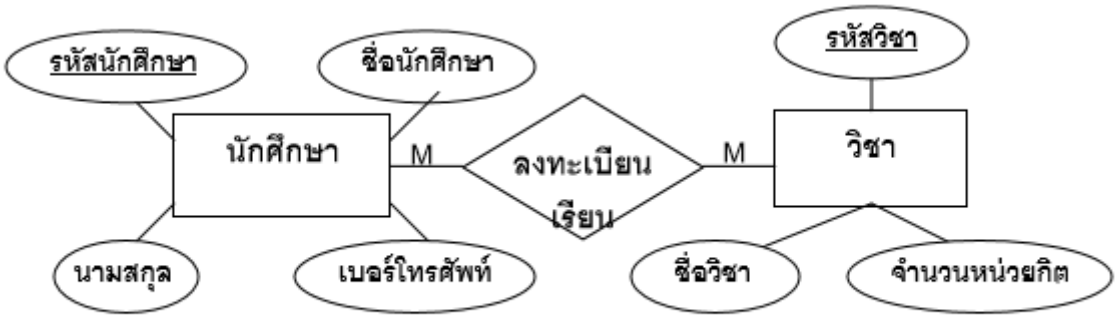
3. สัญลักษณ์ในแบบจำลองอี-อาร์

ตารางที่ 6.4 สัญลักษณ์ที่สำคัญๆ ในแบบจำลองอี-อาร์

สัญลักษณ์	ความหมาย
	เอนทิตี
	ความสัมพันธ์
	แอตทริบิวต์
	คีย์หลัก

ตัวอย่าง

นักศึกษา และ วิชา เป็นเอนทิตีที่เราสนใจจะจัดเก็บ ซึ่งเอนทิตีนักศึกษาจะประกอบด้วยแอตทริบิวต์ ได้แก่ รหัสนักศึกษา ชื่อนักศึกษา นามสกุล และเบอร์โทรศัพท์ เป็นต้น โดยมีรหัสนักศึกษาเป็นคีย์หลัก ส่วนเอนทิตีวิชาจะประกอบด้วยแอตทริบิวต์ ได้แก่ รหัสวิชา ชื่อวิชา และจำนวนหน่วยกิต เป็นต้น โดยมีรหัสวิชาเป็นคีย์หลัก ซึ่งความสัมพันธ์ระหว่างเอนทิตีนักศึกษาและเอนทิตีวิชา เป็นแบบกลุ่มต่อกลุ่ม คือ นักศึกษาหนึ่งคนสามารถลงทะเบียนเรียนได้หลายวิชา และวิชาแต่ละวิชามีนักศึกษาลงทะเบียนเรียนได้หลายคน ดังนั้นเราสามารถนำเสนอในรูปของของแผนภาพ ที่เรียกว่า อี-อาร์ไดอะแกรม (E-R Diagram) ด้วยการใช้สัญลักษณ์ต่างๆ ดังนี้



ภาพที่ 6.11 ตัวอย่างของ อี-อาร์ไดอะแกรม (E-R Diagram)

3. การแปลงแบบจำลองอี-อาร์เป็นโครงสร้างตารางฐานข้อมูล

ขั้นตอนในการแปลงแบบจำลองอี-อาร์เป็นโครงสร้างของตารางในฐานข้อมูลมีขั้นตอนดังต่อไปนี้

4.1 แปลงเอนทิตีปกติในแบบจำลองอี-อาร์เป็น 1 ตาราง ซึ่งประกอบด้วยแอตทริบิวต์ของเอนทิตีนั้นๆ โดยชื่อของตารางก็คือชื่อของเอนทิตี และแอตทริบิวต์ของเอนทิตี ก็คือแอตทริบิวต์ของตาราง สำหรับแอตทริ

บิวต์ที่เป็นคีย์หลักของตาราง ให้ขีดเส้นใต้ที่แอตทริบิวต์นั้นเช่นเดียวกับในแบบจำลองอี-อาร์ซึ่งจากภาพที่ 6.11 นำมาแปลงเป็นตารางได้ 2 ตาราง คือ

นักศึกษา

รหัสนักศึกษา	ชื่อ	นามสกุล	เบอร์โทรศัพท์
56111	สามารถ	ประเสริฐกุล	0-1111-1111
56112	วิชา	ปัญญาเลิศ	0-2222-2222

วิชา

รหัสวิชา	ชื่อวิชา	หน่วยกิต
40001	ภาษาไทย	2
40005	คณิตศาสตร์	3

4.2 แปลงความสัมพันธ์เป็นตาราง

4.2.1 แปลงความสัมพันธ์แบบ 1 : M นั้นไม่ต้องสร้างตารางใหม่ แต่ให้น าแอตทริบิวต์ ที่เป็นคีย์หลักของเอนทิตีที่อยู่ด้านความสัมพันธ์ที่เป็น 1 ไปเพิ่มเป็นแอตทริบิวต์ของตารางด้านที่มีความสัมพันธ์เป็น M

4.2.2 แปลงความสัมพันธ์แบบ M : M จะได้ตารางใหม่ 1 ตาราง ซึ่งประกอบด้วยแอตทริบิวต์ของความสัมพันธ์นั้นรวมกับแอตทริบิวต์ที่เป็นคีย์หลักของ 2 เอนทิตีที่มีความสัมพันธ์ แบบ M : M

จากแบบจำลองอี-อาร์ในภาพที่ 6.11 สามารถสร้างตารางตามขั้นตอนนี้ได้อีก 1 ตาราง คือ ตารางการลงทะเบียน ซึ่งประกอบด้วยแอตทริบิวต์ รหัสนักศึกษา (คีย์หลักของเอนทิตีนักศึกษา) และ รหัสวิชา (คีย์หลักของเอนทิตีวิชา) ฉะนั้นตารางใหม่ที่เกิดขึ้น ซึ่งก็คือ ตารางการลงทะเบียน มี รหัสนักศึกษาและรหัสวิชา เป็นคีย์หลัก ดังนี้

การลงทะเบียน

รหัสนักศึกษา	รหัสวิชา
56111	40001
56111	40005
56112	40001

จากการแปลงแบบจำลองอี-อาร์ตามขั้นตอนข้างต้น สรุปตารางที่ได้ทั้งหมด 3 ตาราง ดังต่อไปนี้

นักศึกษา

รหัสนักศึกษา	ชื่อ	นามสกุล	เบอร์โทรศัพท์
56111	สามารถ	ประเสริฐกุล	0-1111-1111
56112	วิชา	ปัญญาเลิศ	0-2222-2222

วิชา

รหัสวิชา	ชื่อวิชา	หน่วยกิต
40001	ภาษาไทย	2
40005	คณิตศาสตร์	3

การลงทะเบียน

รหัสนักศึกษา	รหัสวิชา
56111	40001
56111	40005
56112	40001

ทั้งนี้โครงสร้างฐานข้อมูลที่ได้จากการแปลงแบบจำลองอี-อาร์นั้นจะอยู่ใน 1NF ดังนั้นจึงจำเป็นต้องนำมาทำนอร์มัลไลเซชันต่อ เพื่อให้ได้ฐานข้อมูลที่ปราศจากความซ้ำซ้อนหรือซ้ำซ้อนน้อยที่สุด แต่ถ้าได้ทำการออกแบบฐานข้อมูลโดยการใช้แบบจำลองอี-อาร์ มาอย่างถูกต้องแล้ว เมื่อแปลงเป็นโครงสร้างฐานข้อมูลแบบสัมพันธ์ จะได้โครงสร้างความสัมพันธ์ที่จัดกลุ่มของแอตทริบิวต์มาเป็นอย่างดี และบางทีโครงสร้างของความสัมพันธ์ที่ได้นั้นอาจอยู่ในนอร์มัลฟอร์มที่สูงกว่านอร์มัลฟอร์มที่ 1 แล้ว อย่างไรก็ตามขั้นตอนถัดมาจำเป็นต้องวิเคราะห์ความสัมพันธ์ระหว่างแอตทริบิวต์ ซึ่งก็คือ วิธีนอร์มัลไลเซชันที่จะกล่าวต่อไป

นอร์มัลไลเซชัน (Normalization)

การออกแบบฐานข้อมูลที่ดีต้องมีความซ้ำซ้อนในการจัดเก็บข้อมูลน้อยที่สุด หรือไม่มีความซ้ำซ้อนเลย ซึ่งต้องอาศัยหลักการในการทำนอร์มัลไลเซชันดังรายละเอียดต่อไปนี้

1. แนวคิดเกี่ยวกับนอร์มัลไลเซชัน

นอร์มัลไลเซชันเป็นวิธีการที่ใช้ในการตรวจสอบและแก้ไขปัญหาเกี่ยวกับความซ้ำซ้อนของข้อมูล โดยดำเนินการให้ข้อมูล ในแต่ละรีเลชั่น (relation) อยู่ในรูปที่เป็นหน่วยที่เล็กที่สุดที่ไม่สามารถแตกออกเป็นหน่วยย่อยๆ ได้อีก โดยยังคงความสัมพันธ์ระหว่างข้อมูลในรีเลชั่นต่างๆ ไว้ตามหลักการที่กำหนดไว้ใน relational model การทำนอร์มัลไลเซชันนี้ เป็นการด าเนินการอย่างเป็นลำดับ ที่กำหนดไว้ด้วยกันเป็นขั้นตอน ตามปัญหาที่เกิดขึ้นในขั้นตอนนั้นๆ ซึ่งแต่ละขั้นตอนจะมีชื่อตามโครงสร้างข้อมูลที่กำหนดไว้ดังนี้ 1. First Normal Form (1NF) 2. Second Normal Form (2NF) 3. Third Normal Form (3NF) 4. Boyce-Codd Normal

Form (BCNF) 5. Fourth Normal Form (4NF) และ 6. Fifth Normal Form (5NF) ในการออกแบบฐานข้อมูลเพื่อลดความซ้ำซ้อนในการจัดเก็บข้อมูลอย่างน้อยต้องมีคุณสมบัติเป็น 3 NF เพราะจริงๆ แล้ว ในการทำงานทั่วๆ ไป แค่ 3 NF ก็สามารถใช้งานได้แล้ว แต่สำหรับ BCNF ไปจนถึง 5NF เป็นฐานข้อมูลชนิดพิเศษจริงๆ ที่แทบจะไม่มีในชีวิตประจำวัน โอกาสพบประมาณ 0.01 % ดังนั้นในที่นี้จะศึกษาเพียงแค่ 1NF 2NF และ 3NF เท่านั้น เพื่อใช้เป็นแนวทางในการศึกษาระดับอื่นต่อไป

2. รูปแบบของนอร์มัลฟอร์ม (Normal Form : NF)

2.1 First Normal Form (1NF) ตารางที่ผ่านการทำนอร์มัลไลเซชันระดับที่ 1 หรือ First Normal Form ต้องมีคุณสมบัติดังนี้

ไม่มีคอลัมน์ใดในตารางที่มีค่ามากกว่า 1 ค่า คือ ค่าในแต่ละคอลัมน์ต้องเป็น atomic หรือไม่อยู่ในรูปของ repeating group

หมายความว่าข้อมูลที่เก็บในแต่ละคอลัมน์จะต้องมีลักษณะเป็นค่าเดียว (single valued) ไม่สามารถแบ่งย่อยได้อีกในการทำนอร์มัลไลเซชันจะต้องดูข้อมูลในตารางเป็นหลัก ตัวอย่างเช่น ข้อมูลในตารางที่ 6.5 แสดงการเก็บข้อมูลเกี่ยวกับนักศึกษา ซึ่งแต่ละคนสามารถอยู่ชมรมและมีงานอดิเรกได้มากกว่า 1 อย่าง

ตารางที่ 6.5 ข้อมูลนักศึกษา

รหัสนักศึกษา	ชื่อ	ที่อยู่	ชมรม	งานอดิเรก
56001	นารี ศิริพร	กรุงเทพ	ดนตรี อาสาพัฒนาชนบท อนุรักษ์สิ่งแวดล้อม	เล่นกีต้าร์
56002	ศรีสมร อมรชัย	นนทบุรี	พระพุทธศาสนา	สะสมพระเครื่อง ร้องเพลง
56003	อรอนงค์ สมประสงค์	กรุงเทพ	ดนตรี	ตีกลอง

จากตารางที่ 6.5 จะเห็นว่าข้อมูลในคอลัมน์ชมรมและงานอดิเรกมีค่ามากกว่า 1 ค่า แสดงว่าไม่เป็น atomic หรืออยู่ในรูปของ repeating group ดังนั้นตารางที่ 6.5 จึงไม่เป็น 1NF โดยเราจะเรียกตารางที่ยังไม่ผ่านแม้แต่ 1NF ว่า Unnormalized Form (UNF) ซึ่งมีวิธีการที่จะท าให้เป็น 1NF คือ

1) แยกคอลัมน์ที่มีค่ามากกว่า 1 ค่าออกเป็นแถวใหม่

2) เพิ่มข้อมูลที่เหมาะสมเข้าไปในคอลัมน์ที่ว่างอยู่ของแถวที่เกิดขึ้นใหม่

จากตารางที่ 6.5 ที่ไม่มีคุณสมบัติเป็น 1NF สามารถทำให้มีคุณสมบัติเป็น 1NF ได้ดังตารางที่ 6.6 ซึ่งมีรหัสนักศึกษา ชมรม และงานอดิเรก เป็นคีย์หลัก

ตารางที่ 6.6 ข้อมูลนักศึกษาที่ผ่านการทำนอร์มัลไลเซชัน ระดับที่ 1 แล้ว

รหัสนักศึกษา	ชื่อ	ที่อยู่	ชมรม	งานอดิเรก
56001	นารี ศิริพร	กรุงเทพ	ดนตรี	เล่นกีต้าร์
56001	นารี ศิริพร	กรุงเทพ	อาสาพัฒนาชนบท	เล่นกีต้าร์
56001	นารี ศิริพร	กรุงเทพ	อนุรักษ์สิ่งแวดล้อม	เล่นกีต้าร์
56002	ศรีสมร อมรชัย	นนทบุรี	พระพุทธศาสนา	สะสมพระเครื่อง
56002	ศรีสมร อมรชัย	นนทบุรี	พระพุทธศาสนา	ร้องเพลง
56003	อรอนงค์ สมประสงค์	กรุงเทพ	ดนตรี	ตีกลอง

สาเหตุที่แยกคอลัมน์ที่มีค่ามากกว่า 1 ค่าออกเป็นแถวใหม่ เนื่องจากไม่รู้จำนวนที่แน่นอนของค่าที่มีอยู่ในคอลัมน์นั้น เช่น ไม่ทราบว่านักศึกษาแต่ละคนจะมีงานอดิเรกกันคนละไม่เกินกี่อย่าง แต่ถ้าเราทราบจำนวนที่แน่นอนของคอลัมน์ที่มีหลายค่านั้น เราอาจแยกเป็นคอลัมน์ใหม่ได้เลย ตัวอย่างเช่น การเก็บชื่อผู้แต่งของหนังสือในห้องสมุด ซึ่งหนังสือเล่มหนึ่งอาจจะมีผู้แต่งหลายคนแต่ในการเก็บชื่อผู้แต่งจะเก็บเพียง 3 คนเท่านั้น ในกรณีนี้ควรจะแบ่งคอลัมน์ซึ่งเก็บชื่อผู้แต่งออกเป็นหลายคอลัมน์โดยขึ้นกับจำนวนผู้แต่งที่มากที่สุดที่มีอยู่หรือเราต้องการเก็บข้อมูลเอาไว้ ซึ่งจะทำให้ค่าของแต่ละคอลัมน์ เป็น Atomic ดังตารางต่อไปนี้

ตารางที่ 6.7 การเก็บข้อมูลหนังสือ ที่มีคุณสมบัติเป็น 1NF

ISBN	ชื่อหนังสือ	ผู้แต่ง 1	ผู้แต่ง 2	ผู้แต่ง 3
9749151001	การจัดการฐานข้อมูล	วาสนา ทรัพย์แก้ว		
9749151002	ระบบสารสนเทศ	วิเชียร เธียรชัย	นฤมล สมสกุล	
9749151003	เทคโนโลยีสารสนเทศ	สุขุม เฉลยทรัพย์	ปริศนา มัชฌิมา	กาญจนา เผือกคง

ถึงแม้ว่าตารางที่ 6.6 จะได้รับการออกแบบให้อยู่ในรูป 1NF แล้ว แต่ลักษณะของข้อมูลภายในอาจก่อให้เกิดปัญหาขึ้นได้อีก เช่น ข้อมูลที่เกี่ยวกับนักศึกษารหัส 56001 ถูกจัดเก็บไว้ในแถวที่ 1, 2 และ 3 ได้แก่ ชื่อ และที่อยู่ โดยจะเห็นว่าเป็นการเก็บข้อมูลที่ซ้ำซ้อนกัน ทำให้สิ้นเปลืองเนื้อที่ในการจัดเก็บ และก่อให้เกิดปัญหาในการเปลี่ยนแปลงข้อมูลด้วย เช่น ถ้านักศึกษารหัส 56001 มีการเปลี่ยนชื่อ หรือที่อยู่ ก็ต้องทำการแก้ไขข้อมูลหลายแถว ซึ่งถ้ามีการแก้ไขข้อมูลไม่ครบ ก็อาจทำให้ข้อมูลภายในตารางเกิดความขัดแย้งกันได้ ดังนั้นจึงต้องมีการนอร์มัลไลเซชันระดับที่ 2 ต่อไป

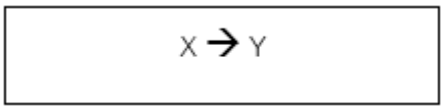
2.2 Second Normal Form (2NF) ตารางที่ผ่านการทำนอร์มัลไลเซชันระดับที่ 2 หรือ Second Normal Form ต้องมีคุณสมบัติดังนี้

1) **ต้องมีคุณสมบัติของ 1NF**
2) **ทุก nonprime attribute จะต้องขึ้นกับ prime (primary key) ทุกตัว**

นั่นคือแอตทริบิวต์ที่ไม่ใช่คีย์หลักจะต้องมีค่าขึ้นอยู่กับคีย์หลักเท่านั้น โดยถ้าคีย์หลักประกอบด้วยแอตทริบิวต์ที่มากกว่า 1 ตัวก็จะต้องขึ้นอยู่กับแอตทริบิวต์ทั้งหมดที่เป็นคีย์หลัก ไม่ใช่ขึ้นอยู่กับบางตัว

การที่จะรู้ว่าแอตทริบิวต์ใดขึ้นอยู่กับแอตทริบิวต์ใดนั้น ต้องใช้ความรู้ในเรื่องฟังก์ชันการขึ้นต่อกัน หรือ functional dependency ซึ่งเป็นสิ่งที่ใช้แสดงความสัมพันธ์ระหว่างแอตทริบิวต์ ในรูปแบบฟังก์ชัน เพื่อช่วยในการตัดสินใจว่าแอตทริบิวต์ที่ไม่ใช่คีย์หลักควรจะปรากฏเป็นคอลัมน์อยู่ในตารางหรือควรจะแยกออกมาสร้างเป็นตารางใหม่

สมมุติว่า X และ Y เป็นแอตทริบิวต์ในตารางหนึ่ง ถ้า Y ขึ้นอยู่กับ X จะสามารถเขียนฟังก์ชันการขึ้นต่อกันได้ดังนี้



การที่ Y ขึ้นอยู่กับ X หมายความว่า ทุกๆค่าของ X ที่เราเลือกขึ้นมา จะสามารถหาค่าของ Y มา 1 ค่าที่สอดคล้องกับค่าของ X ได้เสมอ เช่น จากตารางที่ 6.8 เป็นตารางที่เก็บข้อมูลเกี่ยวกับนักศึกษา ถ้าถามว่านักศึกษาคนใดที่มีรหัสนักศึกษาเป็น 56111 ก็สามารถตอบได้ทันทีว่าคือนักศึกษ าที่ชื่อว่ า สาม ารถ ป ระเสริฐกุล ดังนั้นชื่อนักศึกษ าจึงขึ้นอยู่กับ รหัสนักศึกษ าซึ่งเขียนเป็นฟังก์ชันการขึ้นต่อกันได้ว่า รหัสนักศึกษา-> ชื่อ นั่นเอง

ตารางที่ 6.8 ตารางนักศึกษา

รหัสนักศึกษา (คีย์หลัก)	**ชื่อ**	**นามสกุล**	**เบอร์โทรศัพท์**
56111	สามารถ	ประเสริฐกุล	0-1111-1111
56112	วิชา	ปัญญาเลิศ	0-2222-2222
56113	น้ำทิพย์	ปัญญาเลิศ	0-2222-2222
56114	สมจิตร์	สมสกุลวงศ์	0-6666-6666
56115	วิชา	รักศักดิ์ศรี	0-9999-9999

เมื่อเข้าใจในเรื่องฟังก์ชันการขึ้นต่อกันแล้ว เราลองมาพิจารณาว่าตารางที่ 6.9 จะมีคุณสมบัติเป็น 2NF หรือไม่

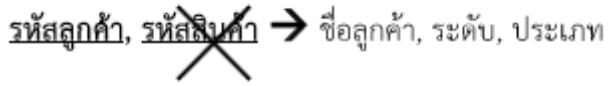
ตารางที่ 6.9 ข้อมูลการสั่งซื้อสินค้าของลูกค้า

รหัสลูกค้า	**ชื่อลูกค้า**	**ระดับ**	**ประเภท**	**รหัสสินค้า**	**ชื่อสินค้า**	**จำนวนสินค้า**
001	นารี	A	ชั้นดี	P111	ปากกา	10
001	นารี	A	ชั้นดี	P222	ดินสอ	12
001	นารี	A	ชั้นดี	P333	ยางลบ	10
002	ศรีสมร	B	ปานกลาง	P222	ดินสอ	15
003	อรอนงค์	C	พอใช้	P333	ยางลบ	16

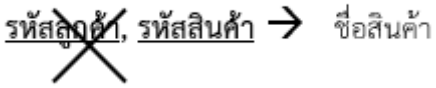
ก่อนอื่นเราต้องพิจารณาว่าตารางที่ 6.9 มีคุณสมบัติเป็น 1NF หรือไม่ จากข้อมูลในตารางจะเห็นว่าไม่มีคอลัมน์ใดในตารางที่มีค่ามากกว่า 1 ค่า แสดงว่าผ่านคุณสมบัติเป็น 1NF จากนั้นต้องพิจารณาต่อว่ามีแอตทริ

บิวต์ใดเป็น prime หรือคีย์หลัก ส่วนที่เหลือก็จะเรียกว่า nonprime จากตารางที่ 6.9 จะมีแอตทริบิวต์ รหัสลูกค้า และ รหัสสินค้า เป็น prime ส่วน ชื่อลูกค้า ที่อยู่ ชื่อสินค้าและจำนวนสินค้า เป็น nonprime

จากคุณสมบัติของ 2NF คือ nonprime ต้องขึ้นกับ prime ทุกตัว ในการพิจารณาว่าเป็น 2NF หรือไม่ จะต้องพิจารณา nonprime ทีละตัว ซึ่งมีผลสรุปการขึ้นต่อกันดังนี้



จากข้อมูลในตารางที่ 6.9 จะสังเกตว่าชื่อลูกค้า, ระดับ และประเภทจะขึ้นกับรหัสลูกค้าเพียงอย่างเดียว ไม่ขึ้นกับรหัสสินค้าเลย ทำให้ไม่เป็นไปตามคุณสมบัติของ 2NF และนอกจากนี้ยังมีกรณีอื่นอีกที่ทำให้ตารางที่ 6.9 ไม่เป็นไปตามคุณสมบัติของ 2NF ได้แก่



จากข้อมูลในตารางที่ 6.9 จะสังเกตว่าชื่อสินค้าจะขึ้นกับรหัสสินค้าเพียงอย่างเดียว ไม่ขึ้นกับรหัสลูกค้าเลย ทำให้ไม่เป็นไปตามคุณสมบัติของ 2NF มีเพียงจำนวนสินค้าอย่างเดียวที่ขึ้นกับรหัสลูกค้า และรหัสสินค้า สรุปว่า

ดังนั้นถ้าต้องการให้ตารางที่ 6.9 มีคุณสมบัติเป็น 2NF จะต้องท าการแตกตารางออกมา ตามความสัมพันธ์ของฟังก์ชันการขึ้นต่อกัน เป็น 3 ตาราง ดังนี้

ตารางที่ 6.10 ข้อมูลลูกค้า

รหัสลูกค้า	ชื่อลูกค้า	ระดับ	ประเภท
001	นารี	A	ชั้นดี
002	ศรีสมร	B	ปานกลาง
003	อรอนงค์	C	พอใช้

ตารางที่ 6.11 ข้อมูลสินค้า

รหัสสินค้า	ชื่อสินค้า
P111	ปากกา
P222	ดินสอ
P333	ยางลบ

ตารางที่ 6.12 ข้อมูลการสั่งซื้อสินค้า

รหัสลูกค้า	รหัสสินค้า	จำนวนสินค้า
001	P111	10
001	P222	12
001	P333	10
002	P222	15
003	P333	16

จากตารางที่ 6.10 หากต้องการเพิ่มข้อมูลลูกค้าขึ้นมาอีก 1 คน อาจมีปัญหาเกิดขึ้นคือ ถ้าลูกค้าที่จะเพิ่มเข้าไปเป็นลูกค้าในระดับ A แสดงว่าต้องเป็นลูกค้าชั้นดี ในช่องประเภทต้องใส่ว่า “ชั้นดี” เท่านั้น ถ้าหากใส่ว่าปานกลาง หรือพอใช้ ข้อมูลก็จะขัดแย้งกัน เพราะฉะนั้นตารางที่ผ่าน 2NF บางตารางอาจเกิดปัญหาในเรื่องของการเพิ่มข้อมูลได้ ดังนั้นจึงต้องมีการนอร์มัลไลเซชันระดับที่ 3 ต่อไป

2.3 Third Normal Form (3NF)

ตารางที่ผ่านการท านอร์มัลไลเซชันระดับที่ 3 หรือ Third Normal Form ต้องมีคุณสมบัติดังนี้

1) **ต้องมีคุณสมบัติของ 2NF**
2) **nonprime ต้องไม่ขึ้นกับ nonprime**

ตารางที่ 6.13 ข้อมูลใบเสร็จรับเงินจากลูกค้า

เลขที่ใบเสร็จ	รหัสลูกค้า	ชื่อลูกค้า	ที่อยู่
1001	43	กฤษณา	121 อ.เมือง จ.นนทบุรี
1002	55	ศักดิ์สิทธิ์	222 อ.ปากเกร็ด จ.นนทบุรี
1003	43	กฤษณา	121 อ.เมือง จ.นนทบุรี

จากตารางที่ 6.13 จะมีแอตทริบิวต์ เลขที่ใบเสร็จ เป็น prime ส่วน รหัสลูกค้าชื่อลูกค้า และที่อยู่ เป็น nonprime นักศึกษาต้องพิจารณาก่อนว่าตารางที่ 6.13 มีคุณสมบัติเป็น 2NF หรือไม่ ซึ่งจากข้อมูลจะสังเกตว่ามีคุณสมบัติเป็น 2NF เนื่องจากไม่อยู่ในรูป repeating group และ nonprime ทุกตัวขึ้นกับ prime ทุกตัว คือ รหัสลูกค้า ชื่อลูกค้า และที่อยู่ ขึ้นกับ เลขที่ใบเสร็จ ซึ่งเป็น prime เพียงตัวเดียว

<u>เลขที่ใบเสร็จ</u> → รหัสลูกค้า, ชื่อลูกค้า, ที่อยู่

เมื่อมีคุณสมบัติเป็น 2NF แล้ว ก็พิจารณาต่อว่า nonprime ขึ้นกับ nonprime หรือไม่ ถ้าไม่มี nonprime ตัวใดขึ้นต่อกัน ก็แสดงว่ามีคุณสมบัติเป็น 3NF แต่จากข้อมูลในตารางที่ 6.13 มี nonprime บางตัวที่ขึ้นต่อกัน ได้แก่ ชื่อลูกค้าและที่อยู่ ขึ้นกับรหัสลูกค้า ท าให้ไม่เป็นไปตามคุณสมบัติของ 2NF

รหัสลูกค้า→ ชื่อลูกค้า, ที่อยู่

ดังนั้นจึงต้องมีการแตกตารางที่ 6.13 ออกมาเป็น 2 ตาราง ตามความสัมพันธ์ของฟังก์ชันการขึ้นต่อกันดังนี้

ตารางที่ 6.14 ข้อมูลลูกค้าจากตารางที่ 6.13

<u>รหัสลูกค้า</u>	ชื่อลูกค้า	ที่อยู่
43	กฤษณา	121 อ.เมือง จ.นนทบุรี
55	ศักดิ์สิทธิ์	222 อ.ปากเกร็ด จ.นนทบุรี

ตารางที่ 6.15 ข้อมูลใบเสร็จ

<u>เลขที่ใบเสร็จ</u>	รหัสลูกค้า
1001	43
1002	55
1003	43

อีกตัวอย่างหนึ่ง จากตารางที่ 6.10 ซึ่งผ่านการนอร์มัลไลเซชันระดับที่ 2 แล้วแต่ยังมีปัญหาการเพิ่มข้อมูลอยู่ ดังที่กล่าวมาแล้วข้างต้น จึงต้องมีการท าให้เป็น 3NF เนื่องจาก nonprime บางตัวขึ้นกับ nonprime คือ ระดับ-> ประเภท ดังนั้นจึงต้องแยก เป็น 2 ตาราง ดังต่อไปนี้

ตารางที่ 6.10 ข้อมูลลูกค้า

<u>รหัสลูกค้า</u>	ชื่อลูกค้า	ระดับ	ประเภท
001	นารี	A	ชั้นดี
002	ศรีสมร	B	ปานกลาง
003	อรอนงค์	C	พอใช้

ตารางที่ 6.16 ข้อมูลลูกค้าจากตารางที่ 6.10

<u>รหัสลูกค้า</u>	ชื่อลูกค้า	ระดับ
001	นารี	A
002	ศรีสมร	B
003	อรอนงค์	C

ตารางที่ 6.17 ข้อมูลระดับลูกค้า

<u>ระดับ</u>	ประเภท
A	ชั้นดี
B	ปานกลาง
C	พอใช้

เนื้อหา

เครือข่ายคอมพิวเตอร์

เครือข่ายคอมพิวเตอร์

เมื่อการใช้งานคอมพิวเตอร์หรืออุปกรณ์ต่าง ๆ มากขึ้น จึงเกิดความต้องการที่จะเชื่อมต่อคอมพิวเตอร์เหล่านั้นถึงกันเพื่อเพิ่มขีดความสามารถของระบบให้สูงขึ้น เพิ่มการใช้งานด้านต่าง ๆ และลดต้นทุนระบบโดยรวมลง มีการแบ่งใช้งานอุปกรณ์และข้อมูลต่าง ๆ ตลอดจนสามารถทำงานร่วมกันได้

สิ่งสำคัญที่ทำให้ระบบข้อมูลมีขีดความสามารถเพิ่มขึ้น คือ การโอนย้ายข้อมูลระหว่างกัน และการเชื่อมต่อหรือการสื่อสาร การโอนย้ายข้อมูลหมายถึงการนำข้อมูลมาแบ่งกันใช้งาน หรือการนำข้อมูลไปใช้ประมวลผลในลักษณะแบ่งกันใช้ทรัพยากร เช่น แบ่งกันใช้ซีพียู แบ่งกันใช้ฮาร์ดดิสก์ แบ่งกันใช้โปรแกรม และแบ่งกันใช้อุปกรณ์อื่น ๆ ที่มีราคาแพงหรือไม่สามารถจัดหาให้ทุกคนได้ การเชื่อมต่อคอมพิวเตอร์เป็นเครือข่ายจึงเป็นการเพิ่มประสิทธิภาพการใช้งานให้กว้างขวางและมากขึ้นจากเดิม

การเชื่อมต่อในความหมายของระบบเครือข่าย ไม่ได้จำกัดอยู่ที่การเชื่อมต่อระหว่างเครื่องไมโครคอมพิวเตอร์ แต่ยังรวมไปถึงการเชื่อมต่ออุปกรณ์รอบข้าง เทคโนโลยีที่ก้าวหน้าทำให้การทำงานเฉพาะมีขอบเขตกว้างขวางยิ่งขึ้น มีการใช้เครื่องบริการแฟ้มข้อมูลเป็นที่เก็บรวบควมแฟ้มข้อมูลต่างๆ มีการทำฐานข้อมูลกลาง มีหน่วยจัดการระบบสื่อสารหน่วยบริการใช้เครื่องพิมพ์ หน่วยบริการการใช้ซีดี หน่วยบริการปลายทาง และอุปกรณ์ประกอบสำหรับต่อเข้าในระบบเครือข่ายเพื่อจะทำงานเฉพาะเจาะจงอย่างใดอย่างหนึ่ง ในรูป เป็นตัวอย่างเครือข่ายคอมพิวเตอร์ที่จัดกลุ่มเชื่อมโยงเป็นระบบ

เครือข่ายคอมพิวเตอร์ก่อให้เกิดความสามารถในการปฏิบัติการร่วมกัน ซึ่งหมายถึงการให้อุปกรณ์ทุกชิ้นที่ต่ออยู่บนเครือข่ายทำงานร่วมกันได้ทั้งหมดในลักษณะที่ประสานรวมกัน โดยผู้ใช้เห็นเสมือนใช้งานในอุปกรณ์เดียวกัน จึงเป็นวิธีการในการนำเอาอุปกรณ์ต่างชนิดจำนวนมาก มารวมกันเป็นเสมือนระบบเดียวกัน ทั้ง ๆ ที่อุปกรณ์เหล่านั้นอาจจะมาจากต่างยี่ห้อ ต่างบริษัท ก็ได้

เครือข่ายแลนหนึ่งเครือข่ายจะมีการทำงานรวมกันเป็นกลุ่ม ที่เรียกว่า กลุ่มงาน (workgroup) แต่เมื่อเชื่อมโยงหลายๆ กลุ่มงานเข้าด้วยกันก็จะเป็นเครือข่ายขององค์กร และถ้าเชื่อมโยงระหว่างองค์กรผ่าน

เครือข่ายแวน ก็จะได้เครือข่ายขนาดใหญ่ขึ้น

การประยุกต์ใช้งานเครือข่ายคอมพิวเตอร์เป็นไปอย่างกว้างขวางและสามารถใช้ประโยชน์ได้มากมาย ทั้งนี้เพราะเครือข่ายคอมพิวเตอร์ทำให้เกิดการเชื่อมโยงอุปกรณ์ต่างๆ เข้าด้วยกัน และสื่อสารข้อมูลระหว่างกันได้

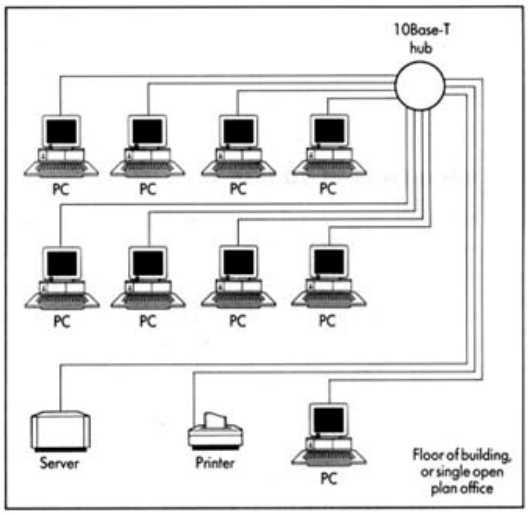
ประเภทของเครือข่ายคอมพิวเตอร์

แบ่งตามขนาดและขอบเขตของพื้นที่เชื่อมโยง

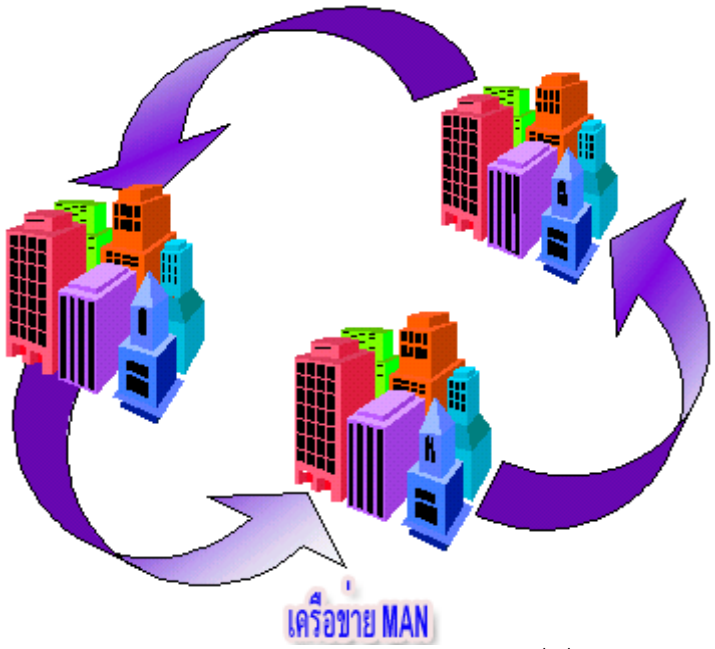
แบ่งออกเป็น 3 ชนิด คือ

1. เครือข่ายแลน (Local Area Network : LAN) หรือเครือข่ายคอมพิวเตอร์ท้องถิ่น เป็นเครือข่ายคอมพิวเตอร์ ซึ่งเชื่อมโยงคอมพิวเตอร์และอุปกรณ์สื่อสารที่อยู่ในท้องที่บริเวณที่ไม่ไกลเข้าด้วยกัน เช่น ภายในอาคาร หรือภายในองค์กรที่มีระยะทางไม่ไกลมากนัก เครือข่ายแลนจัดได้ว่าเป็นเครือข่ายเฉพาะขององค์กร การสร้างเครือข่ายแลนนี้องค์กรสามารถทำเองได้ โดยวางสายสัญญาณสื่อสารภายในอาคารหรือภายในพื้นที่ของตนเอง เครือข่ายแลนมีตั้งแต่เครือข่ายขนาดเล็กที่เชื่อมโยงคอมพิวเตอร์ตั้งแต่สองเครื่องขึ้นไปภายในห้องเดียวกันจนถึงเชื่อมโยงคอมพิวเตอร์ที่อยู่ระหว่างห้องหรืออาคาร เช่น มหาวิทยาลัยที่มีการวางเครือข่ายเชื่อมโยงคอมพิวเตอร์ระหว่างอาคารภายในมหาวิทยาลัย เครือข่ายแลนจึงเป็นเครือข่ายที่รับผิดชอบโดยองค์กรที่เป็นเจ้าของเครือข่ายแลน

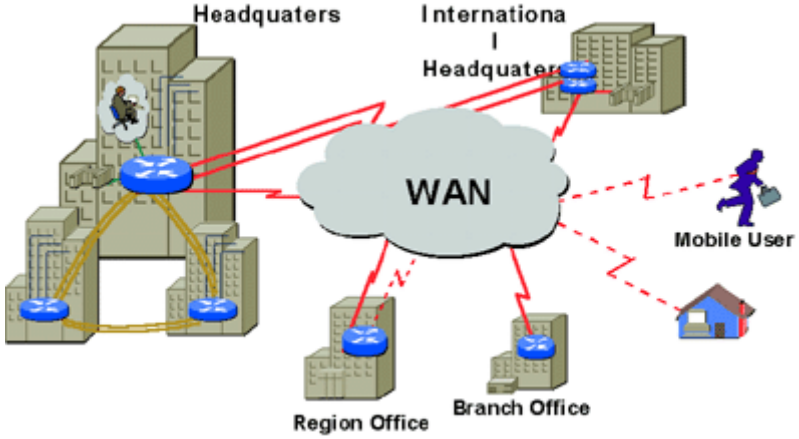
ลักษณะสำคัญของเครือข่ายแลน คือ อุปกรณ์ที่ประกอบภายในเครือข่ายสามารถส่งรับสัญญาณกันด้วยความเร็วสูงมาก โดยทั่วไปมีความเร็วตั้งแต่หลายสิบล้านบิตต่อวินาที จนถึงกว่าพันล้านบิตต่อวินาที การสื่อสารในระยะใกล้จะมีความเร็วในการสื่อสารสูง ทำให้การรับส่งข้อมูลมีความผิดพลาดน้อย และสามารถรับส่งข้อมูลจำนวนมากในเวลาจำกัดได้ เครือข่ายแลนจึงเป็นเครือข่ายที่มีบทบาทสำคัญในการเชื่อมโยงระบบคอมพิวเตอร์ในองค์กรและมีแนวโน้มที่จะทำให้ทรัพยากรและการประมวลผลในองค์กรเชื่อมโยงเป็นระบบเดียวทำให้ใช้งานร่วมกันได้ทั้งองค์กร



2. เครือข่ายระดับเมือง (MAN : Metropolitan Area Network) เป็นการเชื่อมต่อเครือข่ายระหว่างอาคาร ซึ่งอาจจะอยู่ในพื้นที่เดียวกันหรือนอกพื้นที่ที่อาจจะอยู่กันคนละมุมเมืองก็ได้ แต่พื้นที่การเชื่อมต่อนั้นจะกำหนดให้อยู่ภายในเมืองหนึ่ง อาจจะเป็นเมืองที่มีขนาดใหญ่ระดับเมืองหลวง หรือมหานครก็ได้ บางครั้งอาจเป็นการรวมเอาเครือข่ายในเมืองขนาดเล็กเข้าด้วยกันก็ได้ ปกติเครือข่าย MAN นี้มักจะใช้บริการของระบบสื่อสารสาธารณะที่จัดบริการโดยหน่วยงานต่าง ๆ



3. เครือข่ายพื้นที่กว้าง (WAN : Wide Area Network) คอมพิวเตอร์ที่เชื่อมโยงระบบคอมพิวเตอร์ในระยะห่างไกล เช่น เชื่อมโยงระหว่างจังหวัด ระหว่างประเทศ การสร้างเครือข่ายระยะไกลจึงต้องพึ่งพาระบบบริการเครือข่ายสาธารณะ เช่น ใช้สายวงจรเช่าจากองค์การโทรศัพท์แห่งประเทศไทย หรือการสื่อสารแห่งประเทศไทย ใช้วงจรสื่อสารผ่านดาวเทียม ใช้วงจรสื่อสารเฉพาะกิจที่มีให้บริการแบบสาธารณะ เครือข่ายแวนจึงเป็นเครือข่ายที่ใช้กับองค์กรที่มีสาขาห่างไกลและต้องการเชื่อมสาขาเหล่านั้นเข้าด้วยกัน เช่น ธนาคาร มีสาขาทั่วประเทศ มีบริการรับฝากถอนเงินผ่านตู้เอทีเอ็ม



เครือข่ายแวนเชื่อมระยะทางไกลมาก จึงมีความเร็วในการสื่อสารไม่สูง เนื่องจากจะมีสัญญาณรบกวนในสาย และการเชื่อมโยงระยะไกลจำเป็นต้องใช้เทคนิคพิเศษในการลดปัญหาข้อผิดพลาดของการรับส่งข้อมูล

เครือข่ายแวน เป็นเครือข่ายที่ทำให้เครือข่ายแลนหลายๆ เครือข่ายเชื่อมถึงกันได้เช่นที่ทำการสาขาทุกแห่งของธนาคารมีเครือข่ายแลน เพื่อใช้ทำงานภายในสาขานั้น และมีการเชื่อมโยงเครือข่ายแลนทุกสาขาให้เป็นระบบเดียวด้วยเครือข่ายแวน

เครือข่ายอินเทอร์เน็ตเป็นเครือข่ายบริการสาธารณะที่มีการใช้งานได้ทั่วโลก เครือข่ายอินเทอร์เน็ตเป็นตัวอย่างการใช้เครือข่ายแวนเชื่อมโยงเครือข่ายแลนเข้าด้วยกัน

บทบาทของเครือข่ายแวนจะทำให้ทุกบริษัท ทุกองค์กร ทุกหน่วยงานเชื่อมโยงเครือข่ายคอมพิวเตอร์ของตนเองเข้าสู่เครือข่ายกลาง เพื่อแลกเปลี่ยนข้อมูลระหว่างกัน และทำงานร่วมกันในระบบที่

ต้องติดต่อสื่อสารระหว่างกัน

เทคโนโลยีที่ใช้กับเครือข่ายแวนมีความหลากหลาย มีการเชื่อมโยงระหว่างประเทศด้วยช่องสัญญาณดาวเทียม เส้นใยนำแสง คลื่นไมโครเวฟ คลื่นวิทยุ สายเคเบิลทั้งที่วางไปตามถนนและวางใต้น้ำ เทคโนโลยีของการเชื่อมโยงได้รับการพัฒนาไปมาก แต่ก็ยังไม่พอเพียงกับความต้องการที่เพิ่มมากขึ้นอย่างรวดเร็ว

แบ่งตามรูปแบบการเชื่อมต่อ

ในการเชื่อมต่อคอมพิวเตอร์หรืออุปกรณ์ต่าง ๆ ทางคอมพิวเตอร์เข้าเป็นเครือข่ายนั้น เราสามารถเชื่อมต่อกันโดยมีรูปแบบการเชื่อมต่อ (Topology) ที่แตกต่างกันได้ 3 แบบ คือ

1. การเชื่อมต่อแบบดาว (Star Topology)
2. การเชื่อมต่อแบบบัส (Bus Topology)
3. การเชื่อมต่อแบบวงแหวน (Ring Topology)

1. การเชื่อมต่อแบบดาว (Star Topology) เป็นรูปแบบที่ เครื่องคอมพิวเตอร์ทุกเครื่องที่เชื่อมต่อเข้าด้วยกันในเครือข่าย จะต้องเชื่อมต่อกับอุปกรณ์ตัวกลางตัวหนึ่งที่เรียกว่า ฮับ (HUB) หรือเครื่อง ๆ หนึ่ง ซึ่งทำหน้าที่เป็นศูนย์กลางของการเชื่อมต่อสายสัญญาณที่มาจากเครื่องต่าง ๆ ในเครือข่าย และควบคุมเส้นทางการสื่อสาร ทั้งหมด เมื่อมีเครื่องที่ต้องการส่งข้อมูลไปยังเครื่องอื่น ๆ ที่ต้องการในเครือข่าย เครื่องนั้นก็จะต้องส่งข้อมูลมายัง HUB หรือเครื่องศูนย์กลางก่อน แล้ว HUB ก็จะทำหน้าที่กระจายข้อมูลนั้นไปในเครือข่ายต่อไป

ข้อดี

- การติดตั้งเครือข่ายและการดูแลรักษาทำ ได้ง่าย หากมีเครื่องใดเกิดความเสียหาย ก็สามารถตรวจสอบได้ง่าย และศูนย์ กลางสามารถตัดเครื่องที่เสียหายนั้นออกจากการสื่อสาร ในเครือข่ายได้เลย โดยไม่มีผลกระทบกับระบบเครือข่าย

ข้อเสีย

- เสียค่าใช้จ่ายมาก ทั้งในด้านของเครื่องที่จะใช้เป็น เครื่องศูนย์กลาง หรือตัว HUB เอง และค่าใช้จ่ายในการติดตั้งสายเคเบิลในเครื่องอื่น ๆ ทุกเครื่อง การขยายระบบให้ใหญ่ขึ้นทำได้ยาก เพราะการขยายแต่ละครั้งจะต้องเกี่ยวเนื่องกับเครื่องอื่นๆ ทั้งระบบ

2. การเชื่อมต่อแบบบัส (Bus Topology) เป็นรูปแบบที่ เครื่องคอมพิวเตอร์จะถูกเชื่อมต่อกันโดยผ่านสายสัญญาณแกนหลัก ที่เรียกว่า BUS หรือ แบ็คโบน (Backbone) คือ สายรับส่งสัญญาณข้อมูลหลัก ใช้เป็น

ทางเดินข้อมูลของทุกเครื่องภายในระบบเครือข่าย และจะมีสายแยกย่อยออกไปในแต่ละจุด เพื่อเชื่อมต่อเข้ากับคอมพิวเตอร์เครื่องอื่น ๆ ซึ่งเรียกว่าโหนด (Node) ข้อมูลจากโหนดผู้ส่งจะถูกส่งเข้าสู่สายบัสในรูปของแพ็กเกจ ซึ่งแต่ละแพ็กเกจจะประกอบไปด้วยข้อมูลของผู้ส่ง, ผู้รับ และข้อมูลที่จะส่ง การสื่อสารภายในสายบัสจะเป็นแบบ 2 ทิศทางแยกไปยังปลายทั้ง 2 ด้านของ บัส โดยตรงปลายทั้ง 2 ด้านของบัส จะมีเทอร์มิเนเตอร์ (Terminator) ทำหน้าที่ลบล้างสัญญาณที่ส่งมาถึง เพื่อป้องกันไม่ให้สัญญาณข้อมูลนั้นสะท้อนกลับ เข้ามายังบัสอีก เพื่อเป็นการป้องกันการชนกันของข้อมูลอื่น ๆ ที่เดินทางอยู่บนบัสในขณะนั้น

สัญญาณข้อมูลจากโหนดผู้ส่งเมื่อเข้าสู่บัส ข้อมูลจะไหลผ่านไปยังปลายทั้ง 2 ด้านของบัส แต่ละโหนดที่เชื่อมต่อเข้ากับบัส จะคอยตรวจดูว่า ตำแหน่งปลายทางที่มากับแพ็กเกจข้อมูลนั้นตรงกับตำแหน่งของตนหรือไม่ ถ้าตรง ก็จะรับข้อมูลนั้นเข้ามาสู่โหนด ตน แต่ถ้าไม่ใช่ ก็จะปล่อยให้สัญญาณข้อมูลนั้นผ่านไป จะเห็นว่าทุก ๆ โหนดภายในเครือข่ายแบบ BUS นั้นสามารถรับรู้สัญญาณข้อมูลได้ แต่จะมีเพียงโหนดปลายทางเพียงโหนดเดียวเท่านั้นที่จะรับข้อมูลนั้นไปได้

ข้อดี

- ไม่ต้องเสียค่าใช้จ่ายในการวางสายสัญญาณมากนัก สามารถขยายระบบได้ง่าย เสียค่าใช้จ่ายน้อย ซึ่งถือว่าระบบบัสนี้เป็นแบบโทโปโลยีที่ได้รับความนิยมใช้กันมากที่สุดมา ตั้งแต่อดีตจนถึงปัจจุบัน เหตุผลอย่างหนึ่งก็คือสามารถติดตั้งระบบ ดูแลรักษา และติดตั้งอุปกรณ์เพิ่มเติมได้ง่าย ไม่ต้องใช้เทคนิคที่ยุ่งยากซับซ้อนมากนัก

ข้อเสีย

- อาจเกิดข้อผิดพลาดง่าย เนื่องจากทุกเครื่องคอมพิวเตอร์ ต่อยู่บนสายสัญญาณเพียงเส้นเดียว ดังนั้นหากมี สัญญาณขาดที่ตำแหน่งใดตำแหน่งหนึ่ง ก็จะทำให้เครื่องบางเครื่อง หรือทั้งหมดในระบบไม่สามารถใช้งานได้ตามไปด้วย

- การตรวจหาโหนดเสีย ทำได้ยาก เนื่องจากขณะใดขณะหนึ่ง จะมีคอมพิวเตอร์เพียงเครื่องเดียวเท่านั้น ที่สามารถส่งข้อความ ออกมาบนสายสัญญาณ ดังนั้นถ้ามีเครื่องคอมพิวเตอร์จำนวนมากๆ อาจทำให้เกิดการคับคั่งของเน็ตเวิร์ค ซึ่งจะทำให้ระบบช้าลงได้

3. การเชื่อมต่อแบบวงแหวน (Ring Topology) เป็นรูปแบบที่ เครื่องคอมพิวเตอร์ทุกเครื่องในระบบเครือข่าย ทั้งเครื่องที่เป็นผู้ให้บริการ(Server) และ เครื่องที่เป็นผู้ขอใช้บริการ(Client) ทุกเครื่องถูกเชื่อมต่อกันเป็นวงกลม ข้อมูลข่าวสารที่ส่งระหว่างกัน จะไหลวนอยู่ในเครือข่ายไปใน ทิศทางเดียวกัน โดยไม่มีจุดปลายหรือเทอร์มิเนเตอร์เช่นเดียวกับเครือข่ายแบบ BUS ในแต่ละโหนดหรือแต่ละเครื่อง จะมีรีพีตเตอร์ (Repeater) ประจำแต่ละเครื่อง 1 ตัว ซึ่งจะทำหน้าที่เพิ่มเติมข้อมูลที่จำเป็นต่อการติดต่อสื่อสารเข้าในส่วนหัวของแพ็กเกจที่ส่ง และตรวจสอบข้อมูลจากส่วนหัวของ Packet ที่ส่งมาถึง ว่าเป็นข้อมูลของตนหรือไม่ แต่ถ้าไม่ใช่ก็จะปล่อยข้อมูลนั้นไปยัง Repeater ของเครื่องถัดไป

ข้อดี

- ผู้ส่งสามารถส่งข้อมูลไปยังผู้รับได้หลาย ๆ เครื่องพร้อม ๆ กัน โดยกำหนดตำแหน่งปลายทางเหล่านั้นลงในส่วนหัวของแพ็กเกจข้อมูล Repeaterของแต่ละเครื่องจะทำการตรวจสอบเองว่า ข้อมูลที่ส่งมาให้นั้น เป็นตนเองหรือไม่
- การส่งผ่านข้อมูลในเครือข่ายแบบ RING จะเป็นไปในทิศทางเดียวจากเครื่องสู่เครื่อง จึงไม่มีการชนกันของสัญญาณ ข้อมูลที่ส่งออกไป
- คอมพิวเตอร์ทุกเครื่องในเน็ตเวิร์กมีโอกาสที่จะส่งข้อมูลได้อย่างทัดเทียมกัน

ข้อเสีย

- ถ้ามีเครื่องใดเครื่องหนึ่งในเครือข่ายเสียหาย ข้อมูลจะไม่สามารถส่งผ่านไปยังเครื่องต่อ ๆ ไปได้ และจะทำให้เครือข่ายทั้งเครือข่าย หยุดชะงักได้
- ขณะที่ข้อมูลถูกส่งผ่านแต่ละเครื่อง เวลาส่วนหนึ่งจะสูญเสียไปกับการที่ทุก ๆ Repeater จะต้องทำการตรวจสอบตำแหน่งปลายทางของข้อมูลนั้น ๆ ทุก ข้อมูลที่ส่งผ่านมาถึง

แบ่งโดยใช้ลักษณะหน้าที่การทำงานของคอมพิวเตอร์ในเครือข่ายเป็นเกณฑ์

สามารถแบ่งได้เป็น 2 ประเภทดังนี้

1 Peer-to-Peer Network หรือเครือข่ายแบบเท่าเทียม

เป็นการเชื่อมต่อเครื่องคอมพิวเตอร์เข้าด้วยกัน โดยเครื่องคอมพิวเตอร์ แต่ละเครื่อง จะสามารถแบ่งทรัพยากรต่างๆ ไม่ว่าจะเป็นไฟล์หรือเครื่องพิมพ์ซึ่งกันและกันภายในเครือข่ายได้ เครื่องแต่ละเครื่องจะทำงานในลักษณะที่ทัดเทียมกัน ไม่มีเครื่องใดเครื่องเครื่องหนึ่งเป็นเครื่องหลักเหมือนแบบ Client / Server แต่ก็ยังคงคุณสมบัติพื้นฐานของระบบเครือข่ายไว้เหมือนเดิม การเชื่อมต่อแบบนี้มักทำในระบบที่มีขนาดเล็กๆ เช่น หน่วยงานขนาดเล็กที่มีเครื่องใช้ไม่เกิน 10 เครื่อง การเชื่อมต่อแบบนี้มีจุดอ่อนในเรื่องของระบบรักษาความปลอดภัย แต่ถ้าเป็นเครือข่ายขนาดเล็ก และเป็นงานที่ไม่มีข้อมูลที่เป็นความลับมากนัก เครือข่ายแบบนี้ ก็เป็นรูปแบบที่น่าเลือกนำมาใช้ได้เป็นอย่างดี

2 Client-Server Network หรือเครือข่ายแบบผู้ใช้บริการและผู้ให้บริการ

เป็นระบบที่มีเครื่องคอมพิวเตอร์ทุกเครื่องมีฐานะการทำงานที่เหมือน ๆ กัน เท่าเทียมกันภายในระบบเครือข่าย แต่จะมีเครื่องคอมพิวเตอร์เครื่องหนึ่ง ที่ทำหน้าที่เป็นเครื่อง Server ที่ทำหน้าที่ให้บริการทรัพยากรต่าง ๆ ให้กับ เครื่อง Client หรือเครื่องที่ขอใช้บริการ ซึ่งอาจจะต้องเป็นเครื่องที่มีประสิทธิภาพที่ค่อนข้างสูง ถึงจะทำให้การให้บริการมีประสิทธิภาพตามไปด้วย ข้อดีของระบบเครือข่าย Client - Server เป็นระบบที่มี

การรักษาความปลอดภัยสูงกว่า ระบบแบบ Peer To Peer เพราะว่าการจัดการในด้านรักษาความปลอดภัยนั้น จะทำกันบนเครื่อง Server เพียงเครื่องเดียว ทำให้ดูแลรักษาง่าย และสะดวก มีการกำหนดสิทธิการเข้าใช้ทรัพยากรต่าง ๆให้กับเครื่องผู้ขอใช้บริการ หรือเครื่องClient

แบ่งโดย ใช้ระดับความปลอดภัยของข้อมูลเป็นเกณฑ์

การแบ่งประเภทเครือข่ายตามระดับความปลอดภัยของข้อมูล ซึ่งจะแบ่งออกได้เป็น 3 ประเภทคือ อินเทอร์เน็ต (Internet) อินทราเน็ต (Intranet) และ เอ็กส์ทราเน็ต (Extranet) อินเทอร์เน็ตเป็นเครือข่ายสาธารณะที่ทุกคนสามารถเชื่อมต่อเข้าได้ เครือข่ายนี้จะไม่มีความปลอดภัยของข้อมูลเลย ถ้าทุกคนสามารถเข้าถึงข้อมูลที่แชร์ไว้บนอินเทอร์เน็ตได้ ในทางตรงกันข้าม อินทราเน็ตเป็นเครือข่ายส่วนบุคคล ข้อมูลจะถูกแชร์เฉพาะผู้ที่ใช้อยู่ข้างในเท่านั้น หรือผู้ใช้อินเทอร์เน็ตไม่สามารถเข้ามาดูข้อมูลในอินทราเน็ตได้ ถึงแม้ว่าทั้งสองเครือข่ายจะมีการเชื่อมต่อกันอยู่ก็ตาม ส่วนเอ็กทราเน็ตนั้นเป็นเครือข่ายแบบกึ่งอินเทอร์เน็ตและอินทราเน็ตกล่าวคือ การเข้าใช้เอ็กส์ทราเน็ตนั้นมีการควบคุม เอ็กส์ทราเน็ตส่วนใหญ่จะเป็นเครือข่ายที่เชื่อมต่อระหว่างองค์กรเพื่อแลกเปลี่ยนข้อมูลบางอย่างซึ่งกันและกัน ในการแลกเปลี่ยนข้อมูลนี้ต้องมีการควบคุมเพราะเฉพาะข้อมูลบางอย่างเท่านั้นที่ต้องการแลกเปลี่ยน

1 อินเทอร์เน็ต (Internet) เครือข่ายสาธารณะ

อินเทอร์เน็ตเป็นเครือข่ายที่ครอบคลุมทั่วโลก ซึ่งมีคอมพิวเตอร์เป็นล้านๆเครื่องเชื่อมต่อเข้ากับระบบและยังขยายตัวขึ้นเรื่อย ๆ ทุกปี อินเทอร์เน็ตมีผู้ใช้ทั่วโลกหลายร้อยล้านคน และผู้ใช้เหล่านี้สามารถแลกเปลี่ยนข้อมูลข่าวสารกันได้อย่างอิสระ โดยที่ระยะทางและเวลาไม่เป็นอุปสรรค นอกจากนี้ผู้ใช้ยังสามารถเข้าดูข้อมูลต่าง ๆ ที่ถูกตีพิมพ์ในอินเทอร์เน็ตได้ อินเทอร์เน็ตเชื่อมแหล่งข้อมูลต่าง ๆ เข้าด้วยกันไม่ว่าจะเป็นองค์กรธุรกิจ มหาวิทยาลัย หน่วยงานของรัฐบาล หรือแม้กระทั่งแหล่งข้อมูลบุคคล องค์กรธุรกิจหลายองค์กรได้ใช้อินเทอร์เน็ตช่วยในการทำการค้า เช่น การติดต่อซื้อขายผ่านอินเทอร์เน็ตหรืออีคอมเมิร์ช (E-Commerce) ซึ่งเป็นอีกช่องทางหนึ่งสำหรับการทำธุรกิจที่กำลังเป็นที่นิยม เนื่องจากมีต้นทุนที่ถูกกว่าและมีฐานลูกค้าที่ใหญ่มากส่วนข้อเสียของอินเทอร์เน็ตคือ ความปลอดภัยของข้อมูลเนื่องจากทุกคนสามารถเข้าถึงข้อมูลทุกอย่างที่แลกเปลี่ยนผ่านอินเทอร์เน็ตได้

อินเทอร์เน็ตใช้โปรโตคอลที่เรียกว่า “TCP/IP (Transport Connection Protocol/Internet Protocol)” ในการสื่อสารข้อมูลผ่านเครือข่ายซึ่งโปรโตคอลนี้เป็นผลจากโครงการหนึ่งของกระทรวงกลาโหมสหรัฐฯ โครงการนี้มีชื่อว่า ARPANET (Advanced Research Projects Agency Network) ในปี ค.ศ.1975 จุดประสงค์ของโครงการนี้เพื่อเชื่อมต่อคอมพิวเตอร์ที่อยู่ห่างไกลกัน และภายหลังจึงได้กำหนดให้เป็นโปรโตคอลมาตรฐานในเครือข่ายอินเทอร์เน็ต

ในปัจจุบันอินเทอร์เน็ตได้กลายเป็นเครือข่ายสาธารณะ ซึ่งไม่มีผู้ใดหรือองค์กรใดองค์กรหนึ่งเป็นเจ้าของอย่างแท้จริง การเชื่อมต่อเข้ากับอินเทอร์เน็ตต้องเชื่อมต่อผ่านองค์กรที่เรียกว่า “ISP (Internet Service Provider)” ซึ่งจะทำหน้าที่ให้บริการในการเชื่อมต่อเข้ากับอินเทอร์เน็ต นั่นคือ ข้อมูลทุกอย่างที่ส่งผ่านเครือข่าย ทุกคนสามารถดูได้ นอกเสียจากจะมีการเข้ารหัสลับซึ่งผู้ใช้ต้องทำเอง

2 อินทราเน็ต (Intranet) หรือเครือข่ายส่วนบุคคล

ตรงกันข้ามกับอินเทอร์เน็ต อินทราเน็ตเป็นเครือข่ายส่วนบุคคลที่ใช้เทคโนโลยีอินเทอร์เน็ต เช่น เว็บ, อีเมล, FTP เป็นต้น อินทราเน็ตใช้โปรโตคอล TCP/IP สำหรับการรับส่งข้อมูลเช่นเดียวกับอินเทอร์เน็ต ซึ่งโปรโตคอลนี้สามารถใช้ได้กับฮาร์ดแวร์หลายประเภท และสายสัญญาณหลายประเภท ฮาร์ดแวร์ที่ใช้สร้างเครือข่ายไม่ใช่ปัจจัยหลักของอินทราเน็ต แต่เป็นซอฟต์แวร์ที่ทำให้อินทราเน็ตทำงานได้ อินทราเน็ตเป็น

เครือข่ายที่องค์กรสร้างขึ้นสำหรับให้พนักงานขององค์กรใช้เท่านั้น การแชร์ข้อมูลจะอยู่เฉพาะในอินทราเน็ตเท่านั้น หรือถ้ามีการแลกเปลี่ยนข้อมูลกับโลกภายนอกหรืออินเทอร์เน็ต องค์กรนั้นสามารถที่จะกำหนดนโยบายได้ ในขณะที่การแชร์ข้อมูลอินเทอร์เน็ตนั้นยังไม่มีองค์กรใดที่สามารถควบคุมการแลกเปลี่ยนข้อมูลได้ เมื่อเชื่อมต่อเข้ากับอินเทอร์เน็ต พนักงานบริษัทของบริษัทสามารถติดต่อสื่อสารกับโลกภายนอกเพื่อการค้นหาข้อมูลหรือทำธุรกิจต่าง ๆ การใช้โปรโตคอล TCP/IP ทำให้ผู้ใช้สามารถเข้าใช้เครือข่ายจากที่ห่างไกลได้ (Remote Access) เช่น จากที่บ้าน หรือในเวลาที่ต้องเดินทางเพื่อติดต่อธุรกิจ การเชื่อมต่อเข้ากับอินทราเน็ตโดยการใช้โมเด็มและสายโทรศัพท์ ก็เหมือนกับการเชื่อมต่อเข้ากับอินเทอร์เน็ต แต่แตกต่างกันที่เป็นการเชื่อมต่อเข้ากับเครือข่ายส่วนบุคคลแทนที่จะเป็นเครือข่ายสาธารณะอย่างเช่นอินเทอร์เน็ต การเชื่อมต่อกันได้ระหว่างอินทราเน็ตกับอินเทอร์เน็ตถือเป็นประโยชน์ที่สำคัญอย่างหนึ่ง

ระบบการรักษาความปลอดภัยเป็นสิ่งที่แยกอินทราเน็ตออกจากอินเทอร์เน็ต เครือข่ายอินทราเน็ตขององค์กรจะถูกปกป้องโดยไฟร์วอลล์ (Firewall) ซึ่งอาจจะเป็นได้ทั้งฮาร์ดแวร์และซอฟต์แวร์ที่ทำหน้าที่กรองข้อมูลที่แลกเปลี่ยนกันระหว่างอินทราเน็ตและอินเทอร์เน็ตเมื่อทั้งสองระบบมีการเชื่อมต่อกัน ดังนั้นองค์กรสามารถกำหนดนโยบายเพื่อควบคุมการเข้าใช้งานอินทราเน็ตได้

อินทราเน็ตสามารถสนองความต้องการของผู้ใช้ในองค์กรได้หลายอย่าง ความง่ายในการตีพิมพ์บนเว็บทำให้เป็นที่นิยมในการประกาศข่าวสารขององค์กร เช่น ข่าวภายในองค์กร กฎ ระเบียบ และมาตรฐาน การปฏิบัติงานต่าง ๆ เป็นต้น หรือแม้กระทั่งการเข้าถึงฐานข้อมูลขององค์กรก็ง่ายเช่นกัน ผู้ใช้สามารถทำงานร่วมกันได้ง่าย และมีประสิทธิภาพมากขึ้น

3 เอ็กส์ทราเน็ต (Extranet) หรือเครือข่ายร่วม

เอ็กส์ทราเน็ต (Extranet) เป็นเครือข่ายกึ่งอินเทอร์เน็ตกึ่งอินทราเน็ต กล่าวคือ เอ็กส์ทราเน็ตคือเครือข่ายที่เชื่อมต่อระหว่างอินทราเน็ตของสององค์กร ดังนั้นจะมีบางส่วนของเครือข่ายที่เป็นเจ้าของร่วมกันระหว่างสององค์กรหรือบริษัท การสร้างอินทราเน็ตจะไม่จำกัดด้วยเทคโนโลยี แต่จะยากตรงนโยบายที่เกี่ยวกับการรักษาความปลอดภัยของข้อมูลที่ทั้งสององค์กรจะต้องตกลงกัน เช่น องค์กรหนึ่งอาจจะอนุญาตให้ผู้ใช้ขององอีกองค์กรหนึ่งล็อกอินเข้าระบบอินทราเน็ตของตัวเองหรือไม่ เป็นต้น การสร้างเอ็กส์ทราเน็ตจะเน้นที่ระบบการรักษาความปลอดภัยข้อมูล รวมถึงการติดตั้งไฟร์วอลล์หรือระหว่างอินทราเน็ตและการเข้ารหัสข้อมูลและสิ่งที่สำคัญที่สุดก็คือ นโยบายการรักษาความปลอดภัยข้อมูลและการบังคับใช้

เครือข่ายคอมพิวเตอร์ในงานสารสนเทศ

ปัจจุบันระบบเครือข่ายคอมพิวเตอร์ได้กลายป็นส่วนประกอบสำคัญส่วนหนึ่งในงานสารสนเทศ ซึ่งประโยชน์ของระบบเครือข่ายต่องานสารสนเทศนั้น ยกตัวอย่างได้ดังนี้

1. การใช้ฐานข้อมูลร่วมกัน บนเครือข่ายมีสถานีที่เป็นเครื่องให้บริการ ซึ่งเป็นที่เก็บข้อมูลข่าวสารหรือข้อมูลใช้งาน แล้วให้ผู้ใช้ซึ่งเป็นเครื่องขอใช้บริการเรียกใช้ข้อมูล การเรียกใช้ฐานข้อมูลร่วมกันทำให้การปรับปรุงข้อมูล การขอดู และการเรียกค้นกระทำได้ทันที เช่น เมื่อฝ่ายขายขายสินค้า ก็มีการลดจำนวนสินค้าออกจากบัญชีสินค้าคงคลัง เมื่อฝ่ายผลิตขอดูข้อมูลก็ได้ทราบข้อมูลที่เป็นปัจจุบันได้ทันทีว่ามีสินค้าเหลือเท่าไร เนื่องจากเราสามารถใช้ข้อมูลจากฐานข้อมูลร่วมกัน

2. การแบ่งปันทรัพยากรในเครือข่าย นอกจากที่เราจะใช้ฐานข้อมูลร่วมกันได้แล้ว เพื่อเกิดประโยชน์สูงสุด และเป็นการประหยัดทรัพยากร เราสามารถใช้อุปกรณ์ร่วมกันได้

3. การติดต่อสื่อสารระหว่างกันบนเครือข่าย เมื่อมีการเชื่อมโยงสถานีงานหรือเครื่องคอมพิวเตอร์เข้าด้วยกัน ผู้ใช้คอมพิวเตอร์ทุกคนที่อยู่บนเครือข่าย จะสามารถใช้คอมพิวเตอร์ติดต่อสื่อสารระหว่างกัน สามารถ

ส่งไปรษณีย์อิเล็กทรอนิกส์ระหว่างกัน ตลอดจนสามารถโอนย้ายข้อมูลระหว่างกันได้ แต่การดำเนินการต่างๆ ควรเป็นไปตามกฎเกณฑ์ที่ผู้บริหารเครือข่ายหรือองค์กรกำหนดไว้

4. สำนักงานอัตโนมัติ แนวคิดของสำนักงานสมัยใหม่คือการลดปริมาณการใช้กระดาษ โดยการหันมาใช้ระบบการทำงานด้วยคอมพิวเตอร์ที่สามารถแลกเปลี่ยนข้อมูลระหว่างกันได้ทันทีทันใด ปัจจุบันสำนักงานส่วนใหญ่ใช้ซอฟต์แวร์ประมวลคำพิมพ์งานเอกสาร ดังนั้นถ้ามีเครือข่ายคอมพิวเตอร์ภายในองค์กร การสื่อสารส่งงานระหว่างกันที่เป็นกระดาษก็สามารถใช้สัญญาณอิเล็กทรอนิกส์แทนได้ ระบบสำนักงานอัตโนมัติจึงเป็นระบบการทำงานที่ทุกสถานีงานเปรียบเสมือนโต๊ะทำงาน การทำงานแบบสำนักงานอัตโนมัติทำให้เกิดความคล่องตัว และรวดเร็ว

การใช้งานเครือข่ายคอมพิวเตอร์ยังมีอีกมาก มีการประยุกต์กันได้หลายอย่างตั้งแต่การโอนย้ายแฟ้มข้อมูลระหว่างกัน การทำงานเป็นกลุ่ม การใช้ทรัพยากรร่วมกัน การนัดหมาย การส่งงาน แม้แต่ในสถานการศึกษาก็ใช้เครือข่ายเพื่อการเรียนการสอน ใช้เป็นแหล่งความรู้ให้เรียกค้นข้อมูล เป็นต้น