Abstract

Online Information Retrieval

Sumet Vacharachaisurapol

This article concentrates on the online information retrieval services which is the modern international information services system ensuing from the integration of communication technology and computerization. It states that this online searching service will be in existence provided that the telecommunication network is available. The data communication or teleprocessing is explained by distinguishing into connecting system, sender–receiver code, sender–receiver method, error inspection, sender–receiver speed, connecting patterns, and sender–receiver media. Hardware and software are clarified in the consideration of the computerized system as a main instrument in providing information retrieval services. The article is concluded with a brief discussion on database management system.

จุฬาลงกรณ์มหาวิทยาลัย

การสืบค้นข้อสนเทศระยะทางไกล

สุเมธ วัชระชัยสุรพล*

บทนำ

การสืบค้นข้อสนเทศระยะทางไกลเป็นการบริการก้าวใหม่ในระยะเวลาประมาณปีเศษที่ได้รับการกล่าวขวัญกันพอสมควรในด้านการให้บริการแก่ผู้ใช้ของห้องสมุดใหญ่ ๆ หลายแห่งในประเทศของเรา โดยเฉพาะอย่างยิ่งการสืบค้นข้อสนเทศจากฐานข้อมูลใหญ่ ๆ ที่มีข้อมูลจำนวนมากและทันสมัยสำหรับการใช้งานในสาขาต่าง ๆ กว่า 300 สาขาที่รู้จักกันในนาม "ไดอาลอค" (DIALOG) ซึ่งเป็นฐานข้อมูลของบริษัทเอกชนที่ให้บริการด้านข้อมูลข่าวสารต่าง ๆ โดยสถานที่ตั้งของบริษัทอยู่ในสหรัฐอเมริกา และให้บริการด้านการขายข้อมูลข่าวสารให้กับสมาชิกทั่วโลกรวมทั้งประเทศไทยด้วย สมาชิกจะต้องเสียค่าใช้จ่ายในการเป็นสมาชิกและการขอบริการแต่ละครั้งที่เรียกใช้บริการฐานข้อมูล การเสียค่าบริการเรียกใช้ข้อมูลก็จะแตกต่างกันออกไปตามประเภทของข้อมูลในแต่ละฐานข้อมูล ค่าบริการโดยประมาณอยู่ในช่วงระหว่าง 800–10,000 บาทต่อชั่วโมง สำหรับผู้ที่สนใจและอยากจะได้ข้อมูลข่าวสารที่ทันสมัยและอยู่ในความสนใจของตนนั้นนับว่าเป็นค่าใช้จ่ายที่ถูกมาก เพราะหากไม่มีบริการดังกล่าวแล้วจะต้องเสียค่าใช้จ่ายและเวลาเพิ่มขึ้นหลายสิบเท่า การบริการสืบค้นข้อสนเทศระยะทางไกลจึงนับว่าให้ประโยชน์มากต่อนักวิจัย นักวิชาการ นิสิต นักศึกษา นักธุรกิจ ตลอดจนประชาชนทั่วไป ในต่างประเทศได้มีการให้บริการดังกล่าวมานานพอสมควรแล้ว แต่ในประเทศไทยเพิ่งจะเริ่มต้นมาได้ไม่นานนัก ทั้งนี้มิใช่คนไทยมิได้ให้ความสนใจกับข้อมูลที่ทันสมัยหรือคนไทยไม่ได้มีการวิจัย แต่การบริการดังกล่าวจะเกิดขึ้นได้ต่อเมื่อเรามีเครือข่ายระบบสื่อสารที่เอื้ออำนวยเท่านั้น และขณะนี้เราทุกคนก็จะได้ประโยชน์ล้วนหน้าจากการสืบค้นข้อสนเทศระยะทางไกลทั้งทางตรงและทางอ้อม การสืบค้นข้อสนเทศระยะทางไกลเป็นผลจากประสมประสานระหว่างเทคโนโลยีการสื่อสารและคอมพิวเตอร์ ซึ่งเราจะกล่าวถึงรายละเอียดของแต่ละส่วนในบทความนี้ต่อไป

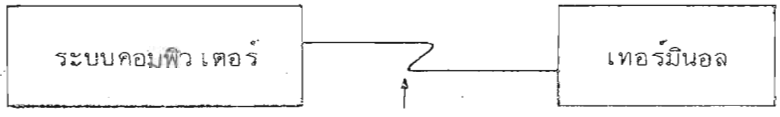
การสื่อสารข้อมูล

การสื่อสารข้อมูลหรือที่รู้จักกันในนาม "Data Communication" หรือ "Telecommunication" นั้นเป็นการรวมเอาเทคนิค 2 อย่างเข้าด้วยกัน อย่างแรกคือ Data Processing หรือ Information System ซึ่งเป็นกิจกรรมเกี่ยวกับข้อมูลเทคนิค อย่างที่สองคือ Telecommunication ซึ่งเป็นตัวกลางในการส่งข่าวสารจากจุดหนึ่งไปยังอีกจุดหนึ่งที่ห่างไกลออกไป เมื่อเอาเทคนิคทั้งสองประการมารวมกันเข้าจึงเรียกว่า "Data Communication" หรือ "Teleprocessing"

การสื่อสารข้อมูลเป็นเทคโนโลยีซึ่งรวมระบบช่ายสายโทรคมนาคมเข้ากับระบบคอมพิวเตอร์ทำให้ขอบข่ายการทำงานของการประมวลผลข้อมูลขยายออกไปอย่างกว้างขวาง อุปกรณ์รับส่งปลายทางที่เรียกกันว่าเทอร์มิ-

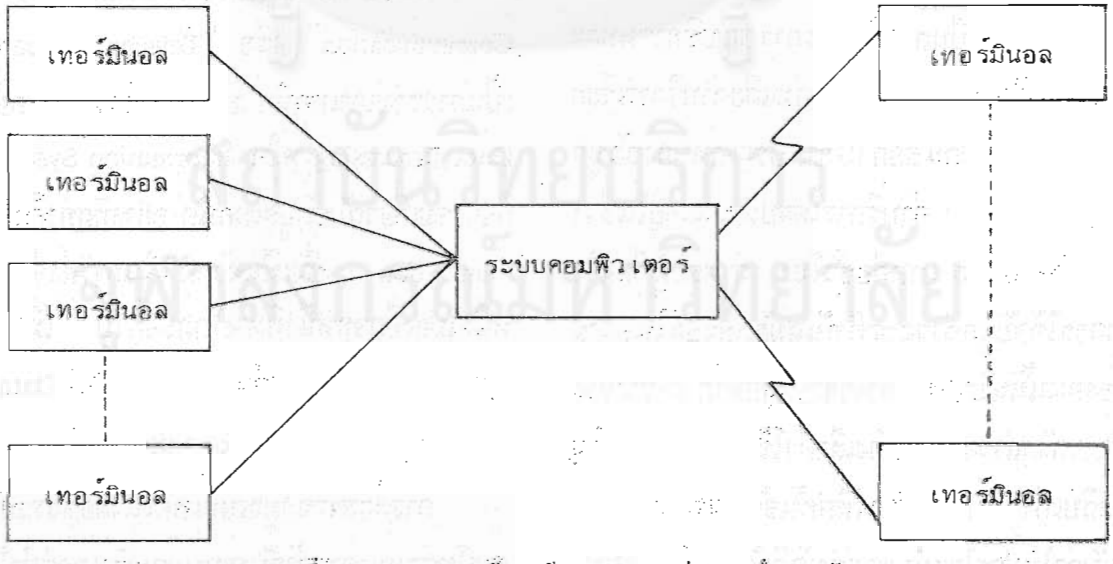
* สุเมธ วัชระชัยสุรพล วท.ม. ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

นอล (Terminal) นั้นสามารถติดต่อกับเครื่องคอมพิวเตอร์ที่อยู่ห่างไกลออกไปได้โดยอาศัยระบบเครือข่ายโทรคมนาคมทำให้ผู้ใช้ได้ประโยชน์มากมาย ลักษณะการเชื่อมต่อดังภาพที่ 1



ภาพที่ 1 การเชื่อมต่อสื่อสาร

ประโยชน์ที่ได้รับจากการสื่อสารข้อมูลที่จะเห็นได้ชัดเจนได้แก่ การประหยัดเวลาและการใช้ระบบคอมพิวเตอร์อย่างมีประสิทธิภาพ ทั้งนี้เพราะผู้ใช้ไม่ต้องเสียเวลากับการเดินทางไปยังศูนย์คอมพิวเตอร์ซึ่งเป็นที่ตั้งของเครื่องคอมพิวเตอร์ ผู้ใช้สามารถใช้เครื่องคอมพิวเตอร์ ณ ที่แห่งใดที่ระบบโทรคมนาคมไปถึง อาทิ เช่นที่ทำงาน ที่บ้าน หรือระหว่างไปทำธุรกิจในที่ต่าง ๆ โดยติดต่อกับคอมพิวเตอร์ผ่านทางเทอร์มินอล เครื่องคอมพิวเตอร์หนึ่งๆ อาจจะพ่วงต่อเทอร์มินอลได้หลายๆ ตัว ทำให้เครื่องคอมพิวเตอร์สามารถให้บริการกับผู้ใช้จำนวนมาก ทำให้เกิดการใช้คอมพิวเตอร์อย่างมีประสิทธิภาพมากยิ่งขึ้น และก่อให้เกิดการบริการในรูปแบบต่าง ๆ ที่อำนวยความสะดวกต่อผู้ใช้มากยิ่งขึ้นด้วย ดังภาพที่ 2



ภาพที่ 2 การประมวลข้อมูลด้วยคอมพิวเตอร์กับการสื่อสารข้อมูล

– หมายถึง การสื่อสารระยะใกล้ (การสื่อสารท้องถิ่น)

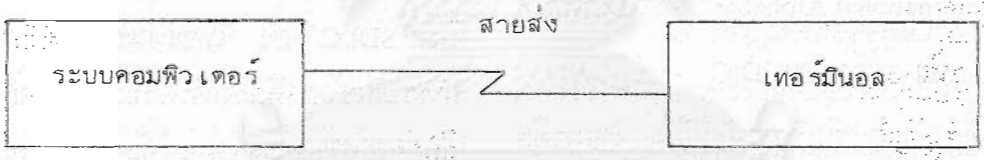
Z หมายถึง การสื่อสารระยะไกล

การสื่อสารข้อมูลขั้นพื้นฐาน คือ การเชื่อมต่ออุปกรณ์รับส่งปลายทางที่เรียกว่าเทอร์มินอลกับระบบคอมพิวเตอร์ ซึ่งอาจจะอยู่ในระยะห่างไกลด้วยระบบสื่อสาร การเชื่อมต่อดังกล่าวมีเทคนิคอยู่หลายอย่างที่น่าสนใจคือ

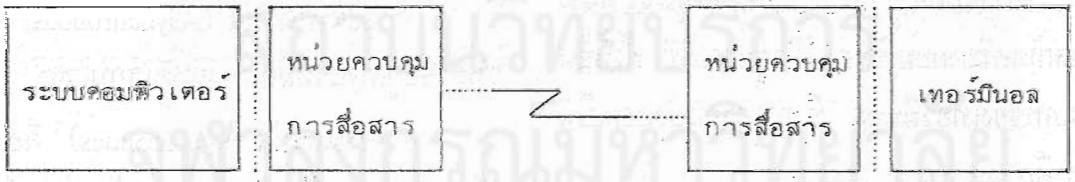
1. **ระบบการเชื่อมต่อ** ระบบการเชื่อมต่อระหว่างเทอร์มินอลและระบบคอมพิวเตอร์ในด้านต้นทางและปลายทางจะประกอบด้วยอุปกรณ์โทรคมนาคมและอุปกรณ์ประกอบอื่น ๆ ได้แก่

1.1 สายส่ง ในระบบการสื่อสารผ่านสายส่งนั้น สายส่งที่จะใช้ในการรับส่งข้อมูลและข่าวสารระหว่างด้านต้นทางและปลายทาง จะสามารถรับส่งได้ครั้งละหนึ่งหน่วยข้อมูล (Element) แต่ในระบบคอมพิวเตอร์ทั่ว ๆ ไป การรับส่งข้อมูลจะอยู่ในอีกลักษณะหนึ่งซึ่งแตกต่างกันออกไป โดยคอมพิวเตอร์จะรับส่งข้อมูลลักษณะของสัญญาณรูปแบบขนาน ในรูปของไบต์ (Byte) หรือรหัส 5 ถึง 8 บิตต่ออักขระ (Character) ดังนั้นการส่งแต่ละอักขระจะต้องทำให้รหัสของอักขระเหล่านี้อยู่ในรูปเรียงตัวกัน (Serialized) เสียก่อนจึงจะสามารถส่งผ่านสายส่งได้

1.2 หน่วยควบคุมการสื่อสาร (DCCU ย่อมาจาก Data Communication Control Unit) เป็นอุปกรณ์ที่ใช้ในการจัดสัญญาณที่อยู่ในรูปแบบขนาน (Parallel Format) ให้อยู่ในรูปแบบเรียงตัวกัน (Serial Format) ทางด้านส่งและในด้านรับหน่วยควบคุมการสื่อสารจะแปลงสัญญาณในรูปแบบเรียงตัวกัน ให้กลับไปอยู่ในรูปแบบขนานเพื่อส่งต่อไปให้ระบบคอมพิวเตอร์ใช้งานต่อไป

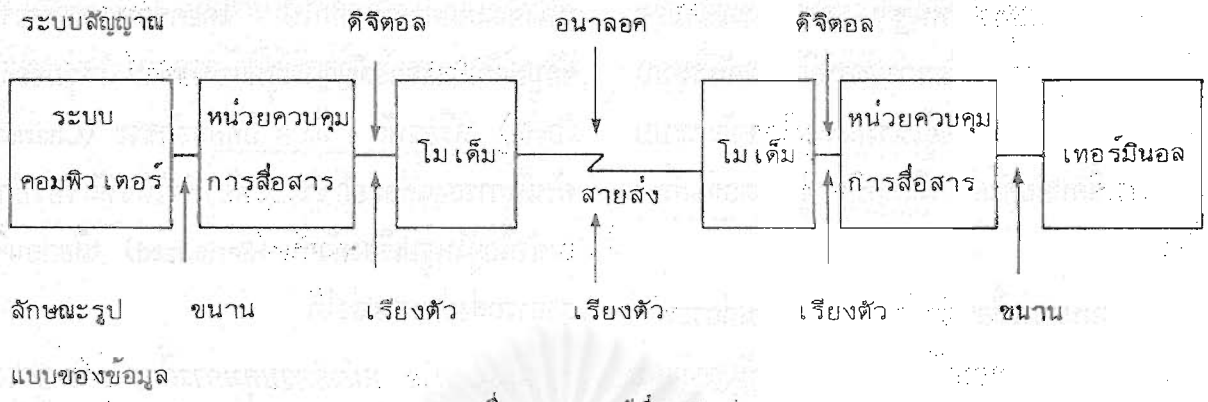


ภาพที่ 3 แสดงหน้าที่ของสายส่ง



ภาพที่ 4 แสดงหน้าที่ของหน่วยควบคุมการสื่อสาร

1.3 โมเด็ม (Modem) เป็นอุปกรณ์ที่ใช้ในการเปลี่ยนสัญญาณดิจิตอลให้อยู่ในรูปของสัญญาณอนาลอกซึ่งเป็นสัญญาณต่อเนื่องในด้านส่งและเปลี่ยนสัญญาณอนาลอกให้กลับเป็นสัญญาณดิจิตอลทางด้านรับ สาเหตุที่ต้องมีการเปลี่ยนสัญญาณดิจิตอลให้เป็นสัญญาณอนาลอกเนื่องจากสัญญาณดิจิตอลจะอยู่ในรูปของสัญญาณไฟฟ้ากระแสตรง ซึ่งไม่สามารถส่งไปในสายส่งได้ไกล จึงจำเป็นต้องเปลี่ยนให้อยู่ในรูปของสัญญาณอนาลอกซึ่งจะสามารถส่งไปในสายส่งได้ไกล ๆ



ภาพที่ 5 แสดงหน้าที่ของโมเด็ม

2. **รหัสการรับส่ง** ในระบบการสื่อสารข้อมูล ข่าวสารที่แลกเปลี่ยนกันระหว่างที่ทำการ 2 แห่ง จะประกอบด้วยอักขระข้อมูลและอักขระควบคุม ข่าวสารที่ใช้รับส่งกันจะถูกใส่รหัสด้วยส่วนผสมของบิตต่าง ๆ ซึ่งเป็นมาตรฐานแล้วด้านรับจะแปลงรหัสเหล่านั้นออกมาเป็นข่าวสารต่าง ๆ รหัสที่นิยมใช้กันทั่วไปมี 2 ชนิด คือ

– รหัส 7–บิต ISO บางครั้งเรียกว่า ASCII หรือ CCITT International Alphabet 5 หรือ ECMA 6

– รหัส 8–บิต EBCDIC

สิ่งที่สำคัญสำหรับเรื่องรหัสนี้ก็คือ ในการติดต่อกันจะต้องใช้รหัสประเภทเดียวกัน ระบบที่ใช้รหัส Ebcdic จะไม่สามารถใช้ติดต่อกับระบบที่ใช้รหัส Iso 7

3. **วิธีการรับส่ง** วิธีการรับส่งข้อมูลระหว่างเทอร์มินอลกับเครื่องคอมพิวเตอร์มีอยู่หลายวิธี ทั้งนี้ขึ้นอยู่กับประเภทของเทอร์มินอล ซึ่งสามารถแบ่งออกเป็น 3 ประเภท คือ

– เทอร์มินอลแบบ Start/Stop ทำงานโดยใช้หลักการส่งแบบทีละอักขระ โดยมีหน่วยแสดงจุดเริ่มต้นและจุดจบ เป็นตัวบอกจุดเริ่มต้นและจุดจบของแต่ละอักขระ โดยทั่วไปแล้วเทอร์มินอลประเภทนี้มักเป็นเทอร์มินอลความเร็วต่ำ เช่น เครื่องโทรพิมพ์

– เทอร์มินอล Synchronous ทำงานโดยใช้หลักการส่งข่าวสารเป็นกลุ่มของตัวอักษร ซึ่งจะมีตัวอักขระควบคุมเป็นตัวบอกจุดเริ่มต้นและจุดจบของแต่ละกลุ่มตัวอักขระเทอร์มินอลประเภทนี้มักเป็นเทอร์มินอลความเร็วสูง

– เทอร์มินอล PACKET ทำงานโดยใช้หลักการส่งข่าวสารเป็นกลุ่มของบิตเรียกว่า "PACKET" บิตต่าง ๆ เหล่านี้จะอยู่ระหว่างชุดของบิตควบคุมที่เรียกว่า "CONTROL FLAGS" โครงสร้างของ PACKET นี้จะขึ้นอยู่กับ PROTOCOL ซึ่งอาจเป็น HDLC หรือ SDLC ฯลฯ เทอร์มินอลประเภทนี้จะใช้สำหรับการรับส่งข้อมูลที่ต้องการความเร็วสูงมาก ๆ เช่น การติดต่อระหว่างคอมพิวเตอร์กับคอมพิวเตอร์

สำหรับการรับส่งข้อมูลจะมีการรับส่งได้ใน 2 ลักษณะ คือ

– อะซินโครนัส (Asynchronous) คือ การรับส่งข้อมูลโดยไม่มีความสัมพันธ์กับเวลา

– ซินโครนัส (Synchronous) คือ การรับส่งข้อมูลโดยทางด้านรับและทางด้านส่งจะมีระบบการสัมพันธ์ทางด้านเวลาเข้ามาเกี่ยวข้อง

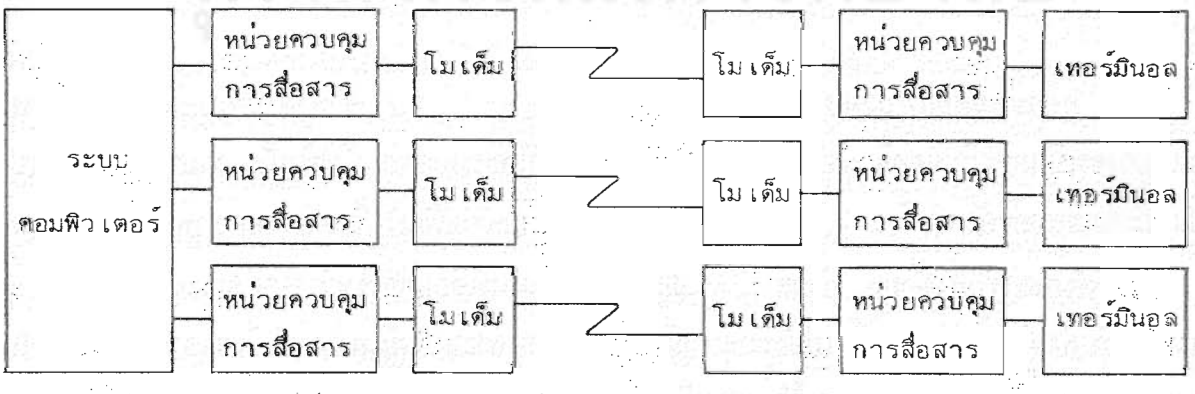
4. **การตรวจสอบความผิดพลาด** ในการรับส่งข้อมูลข่าวสารผ่านระบบโทรคมนาคมมักจะประสบปัญหาข้อมูลผิดพลาดเนื่องจากสิ่งรบกวนต่างๆ ยิ่งข้อมูลต้องเดินทางไกลเท่าไรความเสี่ยงย่อมสูงมากขึ้น แต่ในระบบโทรคมนาคมได้มีระบบการตรวจสอบความผิดพลาดของข้อมูลที่ส่งไปตามสายเมื่อข้อมูลเหล่านั้นไปถึง

ปลายทางแล้ว เป็นการประกันความผิดพลาดที่อาจเกิดขึ้นระหว่างข้อมูลเดินทางด้วยกันหลายวิธีการ เช่น การตรวจสอบความถูกต้องของข้อมูลแต่ละอักขระที่ส่งไปที่เรียกกันว่า Vertical Redundancy Check (VRC) และการตรวจสอบความถูกต้องของชุดข้อมูลที่ส่งว่ามีความถูกต้องหรือไม่ดังที่เรียกกันว่า Logitudinal Redundancy Check (LRC) หรือมีการตรวจสอบความผิดพลาดอีกลักษณะหนึ่งเรียกว่า Cyclic Redundandy Check (CRC) นอกจากนี้ยังมีการตรวจสอบความผิดพลาดอื่น ๆ อีกหลายวิธี ทั้งนี้เพื่อประกันมิให้เกิดความผิดพลาดขึ้นได้ และเมื่อตรวจจับความผิดพลาดได้ระบบจะมีการร้องขอให้มีการส่งข้อมูลนั้นซ้ำอีกหนหนึ่ง และจะทำจนกว่าได้ข้อมูลที่ถูกต้อง แต่อย่างไรก็ตามการส่งข้อมูลซ้ำ ๆ เมื่อพบข้อผิดพลาดจะกระทำในขอบเขตหนึ่งที่กำหนด อาทิเช่น ส่งซ้ำไม่เกิน 10 ครั้ง ถ้าหากยังพบว่าข้อมูลที่ส่งซ้ำถึง 10 หน ยังผิดอยู่ ระบบจะหยุดทำการส่งข้อมูล ซึ่งถือว่าระบบการสื่อสารข้อมูลล้มเหลว ดังนั้นถ้าระบบเครือข่ายโทรคมนาคมไม่อยู่ในสภาพที่ดีย่อมก่อให้เกิดปัญหาการสื่อสารข้อมูลอย่างแน่นอน

5. **ความเร็วในการรับส่ง** ความเร็วในการรับส่งข้อมูลขึ้นกับคุณภาพของวงจรอีเล็กทรอนิกส์ที่ได้ออกแบบไว้ โดยความเร็วนี้มักจะกำหนดเป็นจำนวนบิตที่สามารถส่งได้ใน 1 วินาที ปกติจะใช้ความเร็วต่าง ๆ กัน คือ 110, 3008, 600, 1200, 2400, 4800 และ 9600 บิตต่อวินาที การเลือกใช้ความเร็วในการรับส่งข้อมูลนั้นขึ้นอยู่กับปัจจัยหลาย ๆ ประการ อาทิ คุณภาพของวงจรอิเล็กทรอนิกส์ที่ออกแบบคุณภาพสายส่ง ชนิดของอุปกรณ์รับส่ง ทั้งนี้เพราะการเลือกความเร็วในการรับส่งข้อมูลที่มีประสิทธิภาพดีเช่น 9600 บิตต่อวินาทีนั้น คุณภาพสายส่งจะต้องดีเพียงพอและคุณภาพขององค์ประกอบอื่น ๆ จะต้องดีตามไปด้วย นั่นย่อมหมายความว่าจะต้องลงทุนที่ค่อนข้างสูง ยิ่งต้องส่งข้อมูลระยะไกลเพียงใดค่าลงทุนย่อมจะสูงตามไปด้วย

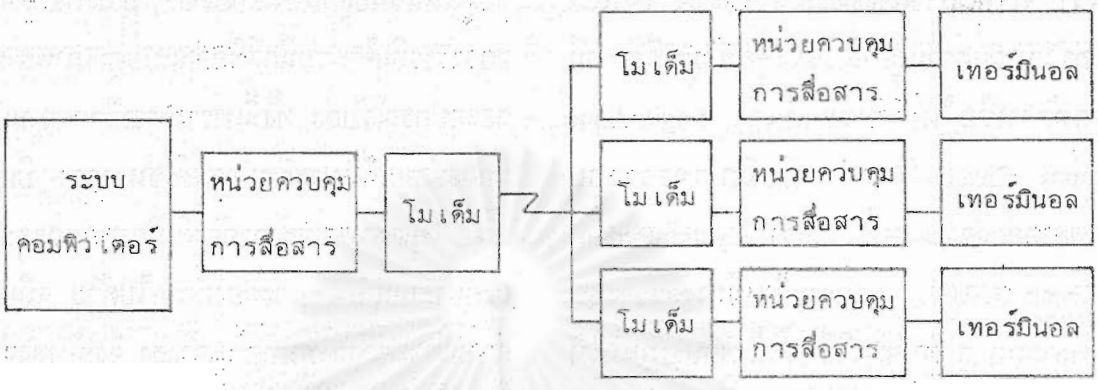
6. **รูปแบบการเชื่อมต่อ** การเชื่อมต่อระหว่างเทอร์มินอลกับคอมพิวเตอร์ หรือการเชื่อมต่อระหว่างคอมพิวเตอร์กับคอมพิวเตอร์ผ่านสายส่งสามารถกระทำได้ในหลายรูปแบบ แต่ละรูปแบบมีข้อดีข้อเสียแตกต่างกัน ค่าใช้จ่ายก็แตกต่างกันออกไปด้วย เช่น

6.1 การเชื่อมต่อจุดต่อจุด (Point to Point) การเชื่อมต่อลักษณะจุดต่อจุดจะได้รับความสะดวกและรวดเร็วในการใช้งานดีที่สุด แต่การลงทุนค่อนข้างสูงโดยเฉพาะอย่างยิ่งกรณีที่เชื่อมต่อด้วยสาย ซึ่งถ้าหากเราต้องการต่อเทอร์มินอลไปใช้งาน 4 เทอร์มินอลจะต้องใช้สาย 4 คู่ โมเด็ม 4 ชุด และทำการต่อดังภาพที่ 6



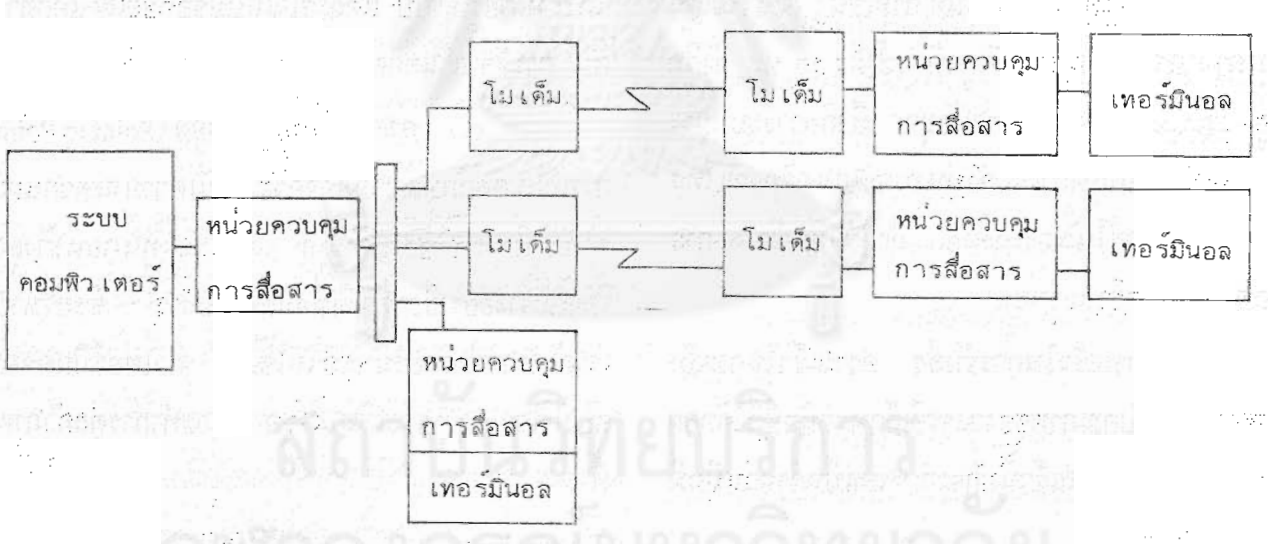
ภาพที่ 6 แสดงรูปแบบการเชื่อมต่อจุดต่อจุด (Point to Point)

6.2 การเชื่อมต่อแบบหลายจุด (Multipoint) การเชื่อมต่อแบบหลายจุดจะช่วยให้ประหยัดสายส่งไปได้มาก ถ้าหากต้องต่อไปในระยะทางไกล ๆ ซึ่งจะทำให้ประหยัดค่าใช้จ่ายเกี่ยวกับค่าสายส่งไปได้มาก ลักษณะการเชื่อมต่อมีลักษณะดังภาพที่ 7



ภาพที่ 7 แสดงรูปแบบการเชื่อมต่อแบบหลายจุด (Multipoint)

6.3 การเชื่อมต่อแบบหลายลักษณะ (Multidrop) การเชื่อมต่อแบบหลายลักษณะนี้ช่วยให้มีความคล่องตัวในการต่อเทอร์มินอลอีกรูปแบบหนึ่ง ดังแสดงด้วยภาพที่ 8

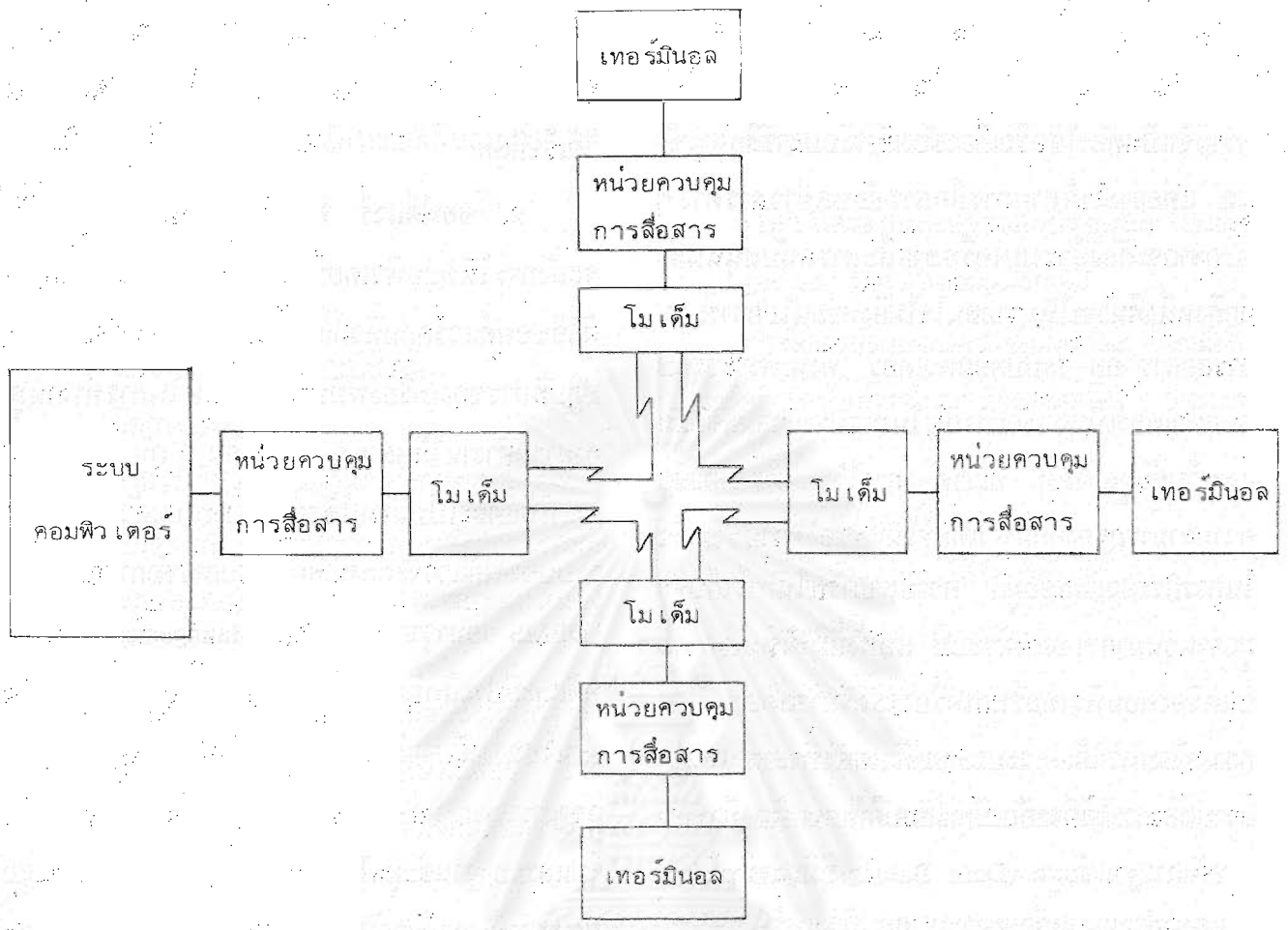


ภาพที่ 8 แสดงรูปแบบการเชื่อมต่อแบบหลายลักษณะ (Multidrop)

6.4 การเชื่อมต่อแบบวงจร (Loop) การเชื่อมต่อแบบวงจรเป็นการเชื่อมต่ออีกลักษณะหนึ่งที่ใช้กัน มีรูปแบบการเชื่อมต่อดังภาพที่ 9

7. **ตัวกลางรับส่งข้อมูล** ตัวกลาง (Media) ที่ทำหน้าที่รับส่งข้อมูลจากจุดหนึ่งไปยังอีกจุดหนึ่งนั้น สามารถแยกออกเป็นประเภทใหญ่ ๆ ได้หลายชนิด คือ

7.1 ตัวพาที่เป็นกระแสไฟฟ้า ตัวพาข้อมูลที่ใช้ระบบกระแสไฟฟ้านั้น สามารถหาข้อมูลไปได้ไกลมากน้อยเพียงใดขึ้นอยู่กับคุณภาพของสายที่เป็นตัวกลางในการเชื่อมต่อระหว่างจุดส่งและจุดรับข้อมูล คุณภาพของสายแปรผันตามราคาซึ่งราคาของสายมีตั้งแต่เมตรละ 3–70 บาท ซึ่งหมายความถึงค่าใช้จ่ายเกี่ยวกับสาย



ภาพที่ 9 แสดงรูปแบบการเชื่อมต่อแบบวงจร (Loop)

จะมากน้อยเพียงใดขึ้นอยู่กับ ราคาสายและระยะทาง แต่อย่างใดก็ตามประเทศไทยมีกฎหมายว่าด้วยการสื่อสาร ซึ่งควบคุมการเดินสายออกนอกตัวอาคารโดยเฉพาะอย่างยิ่งสายเพื่อการสื่อสารจะอยู่ในการควบคุมขององค์การโทรศัพท์แห่งประเทศไทย หน่วยงานต่าง ๆ หรือประชาชนจะเดินสายใช้เองไม่ได้

7.2 ตัวพาที่เป็นคลื่นวิทยุ ตัวพาที่เป็นคลื่นวิทยุหรือที่รู้จักกันในนามของคลื่นไมโครเวฟก็เป็นตัวกลางอีกประเภทหนึ่งที่นิยมใช้กัน ซึ่งบางครั้งได้รับความสะดวกและถูกกว่าการเดินสาย และการส่งข้อมูลไปในระยะทางไกล ๆ ออกไป ตัวที่จะเข้ามาช่วยเสริมคือระบบดาวเทียมเพื่อการสื่อสาร ทั้งนี้เพราะการใช้คลื่นไมโครเวฟนั้นจุดส่งและจุดรับจะต้องไม่มีสิ่งกีดขวาง ถ้าหากมีสิ่งกีดขวางจะไม่สามารถส่งข้อมูลถึงกันได้ ดังนั้นดาวเทียมสื่อสารจึงเข้ามามีบทบาทอย่างมาก และช่วยให้เราสามารถติดต่อสื่อสารกันอีกซีกโลกหนึ่งในระยะเวลาเพียงสั้น ๆ จนเกือบจะไม่มีความรู้สึกในความแตกต่างเรื่องเวลาที่เสียไปเนื่องจากการเดินทางของข้อมูล

7.3 ตัวพาที่เป็นลำแสง ตัวพาที่เป็นลำแสงหรือที่รู้จักกันในนามของการใช้ Optic Fiber ซึ่งเป็นตัวกลางในการรับส่งข้อมูลหรือลำแสงเลเซอร์ ซึ่งสามารถช่วยพาข้อมูลไปได้เช่นเดียวกัน ระบบที่พบว่าดีคือระบบการใช้เส้นใยแสง (Optic Fiber) พบว่ามีประสิทธิภาพในการทำงานสูงมาก และไม่ถูกรบกวนด้วยสิ่งรบกวนอื่น ๆ ดังที่ได้ประสบในวิธีอื่น ๆ แต่อย่างไรก็ตามระบบการใช้เส้นใยยังมีราคาค่อนข้างสูง และใช้เทคนิคที่สลับซับซ้อนพอควร

การทราบถึงกระบวนการรับส่งข้อมูล จะช่วยให้เราเข้าใจ มั่นใจ และใช้ประโยชน์ได้อย่างเต็มที่เมื่อมีความจำเป็นต้องใช้หรือเกี่ยวข้องกับระบบการสื่อสารข้อมูล แต่อย่างไรก็ตามการสื่อสารข้อมูลข่าวสารต่าง ๆ นอกจากจะต้องมีระบบเครือข่ายสื่อสารสนับสนุนแล้ว อีกสิ่งหนึ่งที่น่าจะให้ความสนใจไม่ยิ่งหย่อนไปกว่าระบบการสื่อสาร คือ ระบบคอมพิวเตอร์ ทั้งนี้เพราะระบบคอมพิวเตอร์เป็นตัวจักรสำคัญในการประมวลผลข้อมูลและเก็บข้อมูลต่าง ๆ ที่เราต้องการไว้ การอาศัยขีดความสามารถของคอมพิวเตอร์ในแง่ของความรวดเร็วในการประมวลผลข้อมูล ความสามารถในการเก็บข้อมูลจำนวนมาก ๆ อย่างมีระบบ แต่ทั้งนี้มิได้หมายความว่าเครื่องคอมพิวเตอร์ทุกเครื่องจะสามารถตอบสนองความต้องการได้ ระบบคอมพิวเตอร์ที่จะตอบสนองความต้องการได้ดีจะต้องมีคุณสมบัติพิเศษ คือ มีระบบการจัดการฐานข้อมูล (Data Base) ที่มีประสิทธิภาพเป็นองค์ประกอบหลักของระบบคอมพิวเตอร์

ระบบคอมพิวเตอร์

เมื่อพูดถึงระบบคอมพิวเตอร์ซึ่งเป็นตัวจักรสำคัญในการให้บริการด้านการสืบค้นข้อสนเทศแล้ว เราจะต้องมองระบบคอมพิวเตอร์ออกเป็น 2 ส่วนใหญ่ ๆ คือ ฮาร์ดแวร์และซอฟต์แวร์

1. ฮาร์ดแวร์ คือ ระบบตัวเครื่องคอมพิวเตอร์ซึ่งมีองค์ประกอบหลัก ๆ คือ

- หน่วยรับข้อมูล (Input Unit)
- หน่วยแสดงผลลัพธ์ (Output Unit)
- หน่วยประมวลผลกลาง (Central Processing Unit)
- หน่วยความจำ (Memory)
- หน่วยความจำสำรอง (Secondary Storage)

โดยเฉพาะอย่างยิ่งจะต้องมีหน่วยความจำสำรองที่ใหญ่พอที่จะเก็บข้อมูลต่าง ๆ ที่จะใช้บริการ ปกตินิยมใช้จานแม่เหล็กขนาดใหญ่ที่มีความจุข้อมูลสูงจำนวนหลาย ๆ หน่วย เพื่อให้สามารถเข้าถึงข้อมูลที่ต้องการได้เร็วที่สุด

2. ซอฟต์แวร์ คือ โปรแกรมคอมพิวเตอร์ที่จะสั่งการให้คอมพิวเตอร์ทำงานตามวัตถุประสงค์ของผู้ใช้ซอฟต์แวร์คอมพิวเตอร์นั้นนอกจากจะต้องมีระบบปฏิบัติการของเครื่องที่มีประสิทธิภาพในการทำงานสูง สามารถทำงานได้หลาย ๆ งานพร้อม ๆ กัน และแบ่งเวลาให้แต่ละโปรแกรมใช้อย่างเหมาะสมแล้ว สิ่งหนึ่งที่ระบบซอฟต์แวร์จะต้องมี คือ ระบบการจัดการฐานข้อมูล (DBMS ย่อมาจาก Data Base Management System) ที่ดีและมีประสิทธิภาพ จึงจะสามารถเอื้อให้การบริการเป็นไปอย่างมีประสิทธิภาพ เมื่อพูดถึงฐานข้อมูลนักวิชาการหลาย ๆ ท่านได้ให้ความหมายไว้แตกต่างกันออกไปและมองฐานข้อมูลในทัศนะต่าง ๆ กัน ในที่นี้จะขอนำเอาคำนิยามเกี่ยวกับฐานข้อมูลของ James Martin ซึ่งได้ให้คำนิยามของฐานข้อมูลในหนังสือชื่อ "Principles of Data-base Management" มาเพื่อเป็นแนวทางในการทำความเข้าใจเกี่ยวกับฐานข้อมูลดังนี้

"A collection of data designed to be used by different programmers is called a database. We will define it as a collection of interrelated data stored together with controlled redundancy to serve one or more applications in an optimal fashion; the data are stored so that they are independent of programs which use the data; a common and controlled approach is used in adding new data and modifying and retrieving existing data within the database. One system is said to contain a collection of databases if they are each entirely seperate in structure."

จากลักษณะนิยามของฐานข้อมูลที่ยกเป็นตัวอย่างในที่นี้จะเห็นว่า การจัดการฐานข้อมูลนั้น มุ่งหมายเพื่อให้ข้อมูลในฐานข้อมูลมีลักษณะที่ดีหลายประการ เช่น

- กะทัดรัด
- รวดเร็วในการใช้งาน
- มีความคล่องตัวในการใช้งาน
- มีความทันสมัย
- สามารถใช้งานร่วมงานกันได้หลาย ๆ งาน

โดยปกติแล้วระบบการจัดการฐานข้อมูลของคอมพิวเตอร์จะถูกออกแบบมาเพื่อให้สามารถตอบสนองลักษณะการใช้งานในด้านต่าง ๆ คือ

- สร้างแฟ้มข้อมูลเพื่อเก็บระเบียบข้อมูลต่าง ๆ
- บำรุงรักษาแฟ้มข้อมูลต่างๆให้ทันสมัย โดยการปรับปรุงลบและเพิ่มระเบียนในแฟ้มข้อมูลที่สร้างขึ้นแล้ว
- สามารถดึงข้อมูลจากแฟ้มข้อมูล ในลักษณะต่าง ๆ ที่กำหนดได้
- สามารถแสดงผลข้อมูลที่ดึงมาได้ในรูปแบบที่กำหนด

ดังนั้นระบบการจัดการฐานข้อมูลต่าง ๆ จึงมักจะต้องมีภาษาสำหรับผู้ใช้ในการโต้ตอบและประมวลผลข้อมูลในลักษณะที่ต้องการ

ภาษาโต้ตอบและการจัดการข้อมูล (Query Languages and Data Manipulation)

การจัดการข้อมูลของฐานข้อมูลต่าง ๆ โดยทั่วไปแล้วจะมีคำสั่งเพื่อการจัดการแยกออกเป็น 4 ประเภทใหญ่ ๆ คือ

- การดึงข้อมูล (Retrieval)
- การเปลี่ยนแปลงแก้ไข (Modification)
- การลบข้อมูล (Deletion)
- การเพิ่มข้อมูล (Insertion)

โดยการจัดการข้อมูลในลักษณะดังที่กล่าวแล้วสามารถแสดงด้วยภาษาที่โต้ตอบกับคอมพิวเตอร์ในรูปของคำสั่งต่าง ๆ เช่น

Retrieval	SELECT
Modification	UPDATE
Deletion	DELETE
Insertion	INSERT

ลักษณะคำสั่งที่ใช้อาจแปรเปลี่ยนไปตามภาษาของแต่ละระบบฐานข้อมูลที่บริษัทผู้ผลิตจะกำหนด ดังตัวอย่างภาษาต่าง ๆ อาจใช้คำสั่งเหมือนกันและต่างกันดังนี้

Function	Languages			
	SQL	QBE	NOMAD	dBASE
Retrieval	SELECT	P	LIST	DISPLAY
Modification	UPDATE	U	CHANGE	REPLACE
Deletion	DELETE	D	DELETE	DELETE
Insertion	INSERT	I	INSERT	APPEN

หมายเหตุ

SQL ย่อมาจาก SAMPLE QUERIES LANGUAGES

QBE ย่อมาจาก QUERY BY EXAMPLE

ภาษาที่ใช้สำหรับติดต่อกับระบบจัดการฐานข้อมูลนับได้ว่าเป็นภาษาระดับสูงที่เรียนรู้ได้ง่ายและกะทัดรัดมากเมื่อเทียบกับภาษาคอมพิวเตอร์อื่น ๆ

นอกจากฟังก์ชั่นหลัก ๆ ดังที่กล่าวแล้วนั้น ระบบจัดการฐานข้อมูลยังมีฟังก์ชั่นอื่น ๆ อีกหลายอย่าง เพื่ออำนวยความสะดวกกับผู้ใช้ คือ

- การใช้ข้อมูลร่วมกันของผู้ใช้หลาย ๆ คน
- มีระบบการควบคุมการใช้ข้อมูลร่วมกันที่ดี
- มีระบบควบคุมความถูกต้องของข้อมูล
- มีระบบควบคุมความปลอดภัยของข้อมูล
- และฟังก์ชั่นอื่น ๆ อีกมากมาย

สรุป

การสืบค้นข้อสนเทศระยะทางไกล เป็นเทคโนโลยีใหม่ทางด้านการบริการสารสนเทศที่เกิดขึ้นจากการนำเอาศาสตร์ 2 แขนงมาประยุกต์ร่วมกัน คือ ศาสตร์ว่าด้วยการสื่อสารซึ่งสามารถเชื่อมต่อการสื่อสารระหว่างจุดสองจุดไม่ว่าจะอยู่ ณ ไกลสักเพียงใด และศาสตร์ทางด้านคอมพิวเตอร์ ซึ่งสามารถเก็บข้อมูลจำนวนมาก ๆ และทำงานได้รวดเร็วมาประสมประสานกันอย่างสอดคล้อง ทำให้มนุษยชาติได้รับข่าวสารต่าง ๆ ด้วยความรวดเร็วและถูกต้อง และในปัจจุบันนี้การแข่งขันทางด้านต่าง ๆ ทวีความรุนแรงและเข้มข้นมากยิ่งขึ้น ข่าวสารต่าง ๆ ที่ทันเหตุการณ์จะช่วยให้การตัดสินใจการกระทำต่าง ๆ มีความถูกต้องและรวดเร็วยิ่งขึ้น และนับวันระบบสารสนเทศจะมีบทบาทมากยิ่งขึ้นทุกวัน

การเข้าใจระบบการทำงานของการสืบค้นข้อสนเทศระยะทางไกลจึงนับว่าเป็นประโยชน์อย่างมากต่อผู้ใช้ ในแง่คิดต่าง ๆ และสามารถใช้ประโยชน์ได้อย่างถูกต้องและเหมาะสม ฐานข้อมูล DIALOG เป็นฐานข้อมูลหนึ่งในหลาย ๆ ฐานข้อมูลที่ได้ให้บริการการสืบค้นข้อสนเทศ และแต่ละฐานข้อมูลก็จะมีลักษณะข้อมูลที่ให้บริการซ้ำกันบ้างแตกต่างกันออกไปบ้าง ดังนั้นผู้ใช้บริการจึงควรศึกษาดูว่าฐานข้อมูลใดมีลักษณะเด่นทางด้านใด จึงจะได้ประโยชน์จากการใช้ฐานข้อมูลนั้น ๆ สูงสุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย