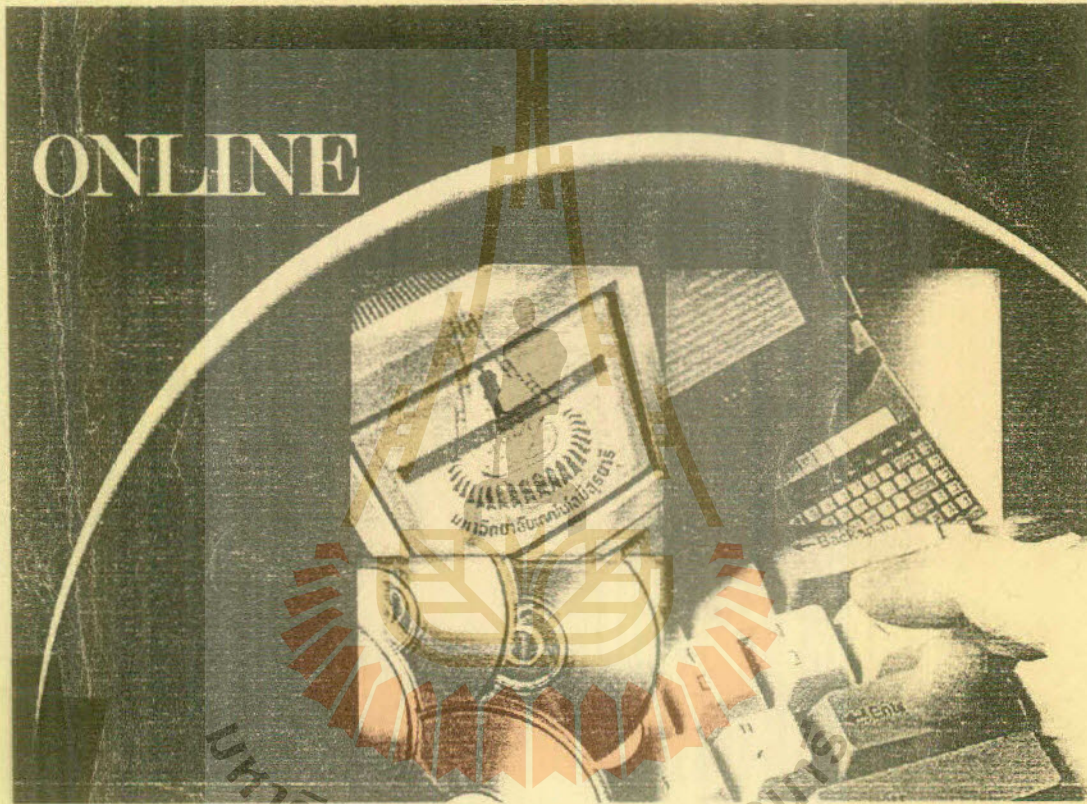


Information 5

204312

การจัดเก็บและการค้นคืนสารสนเทศ

Information Storage and Retrieval



มหาวิทยาลัยเทคโนโลยีสุรนารี

อาจารย์นิศาชล จำนงศรี

สาขาวิชาเทคโนโลยีสารสนเทศ สำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

2546

มทส

สวทส.ส64

น65ก64

2546

คำนำ

เอกสารฉบับนี้จัดทำขึ้นเพื่อให้ประกอบการเรียนการสอนรายวิชา 204312 การจัดเก็บและค้นคืนสารสนเทศ (204312 Information Storage and Retrieval) ทั้งนี้เพื่อให้นักศึกษาสามารถทำความเข้าใจกับเนื้อหาของรายวิชาได้ดียิ่งขึ้น และสามารถศึกษาเพิ่มเติมด้วยตนเองได้อย่างมีประสิทธิภาพ เนื้อหาของรายวิชาแบ่งเป็น 3 ส่วนคือ ส่วนแรกเป็นส่วนของกระบวนการวิเคราะห์เพื่อสร้างตัวแทนให้กับเอกสาร ซึ่งประกอบด้วยการจัดทำฐานข้อมูลบรรณานุกรม การทำดัชนี การทำสาระสังเขป และการสร้างธิซอร์ส โดยจะให้นักศึกษาได้เรียนรู้ทั้งในส่วนของทฤษฎีและการปฏิบัติ ส่วนที่สองจะเน้นที่กระบวนการค้นคืนสารสนเทศ ซึ่งประกอบด้วย การสัมภาษณ์ผู้ใช้ กระบวนการของการค้นคืนสารสนเทศซึ่งได้แก่ การสร้างกลยุทธ์ และการสร้างสูตรการค้น ซึ่งนักศึกษาจะได้ปฏิบัติจริง ทั้งนี้เพื่อให้เกิดทักษะในการทำงาน ส่วนที่สามได้แก่ ประเด็นเพิ่มเติมอื่น ๆ ที่เกี่ยวข้องกับการจัดเก็บและค้นคืนสารสนเทศ ประกอบด้วย (1) ฐานข้อมูลในการให้บริการสารสนเทศ ซึ่งจะกล่าวถึงฐานข้อมูลประเภทต่าง ๆ รวมทั้ง Search Engines (2) ระบบการเชื่อมประสานกับผู้ใช้สำหรับบริการค้นคืนสารสนเทศ ทั้งนี้เพื่อประโยชน์ในการออกแบบและพัฒนาระบบการค้นคืนสารสนเทศให้เป็นระบบที่ใช้ง่ายและมีประสิทธิภาพ และ (3) การประเมินผลการค้นและการประเมินระบบการค้นคืนสารสนเทศ โดยจะกล่าวถึงสาเหตุที่ทำให้การค้นคืนสารสนเทศไม่ประสบผลสำเร็จได้ด้วย ทั้งนี้เพื่อประโยชน์ในการพัฒนาระบบต่อไป

นิศาชล จำนงศรี

28 กุมภาพันธ์ 2546

สาขาวิชาเทคโนโลยีสารสนเทศ

สำนักวิชาเทคโนโลยีสังคม

มหาวิทยาลัยเทคโนโลยีสุรนารี

มหาวิทยาลัยเทคโนโลยีสุรนารี



มทส

สวทส.ส64

10.3.5

น65ก64

2546

Call No.

Bib No. H๕๕๖๑ 111๕19

ราคา

วัน เดือน ปี 26.0๘.2547

เลขทะเบียน B082092

มทส

สวทส.ส64

น65ก64

2546

หน้า

ก

ข-จ

สารบัญ

บทที่ 1 แนวคิดพื้นฐานเกี่ยวกับการจัดเก็บและการค้นคืน

สารสนเทศด้วยคอมพิวเตอร์ (Information Storage and Retrieval: Introduction) 1

การจัดเก็บและการค้นคืนสารสนเทศ 1

ความหมายของการค้นคืนสารสนเทศ 2

การดำเนินงานของระบบการจัดเก็บและค้นคืนสารสนเทศ 3

องค์ประกอบของระบบการจัดเก็บและค้นคืนสารสนเทศ 4

 กระบวนการของผู้ใช้ 5

 กระบวนการของผู้ผลิตฐานข้อมูล 8

 กระบวนการทำงานของคอมพิวเตอร์ 10

 สาเหตุที่กระบวนการค้นคืนไม่สมบูรณ์ 12

ผู้ใช้ระบบการค้นคืนสารสนเทศ 12

ปัญหาในการออกแบบและการใช้ระบบค้นคืนสารสนเทศ 14

ประเด็นการศึกษาด้านการจัดเก็บและการค้นคืนสารสนเทศ 15

บรรณานุกรม 16

บทที่ 2 ฐานข้อมูลบรรณานุกรม (Bibliographic Database) 17

 ความหมายของฐานข้อมูล 17

 ประเภทของฐานข้อมูล 17

 การสร้างฐานข้อมูลบรรณานุกรม 18

 ขั้นตอนการสร้างฐานข้อมูล 19

 บรรณานุกรม 29

บทที่ 3 การจัดทำสารระสังเขป (Abstracting) 30

 ความหมายของสารระสังเขป 30

 วัตถุประสงค์ของสารระสังเขป 30

 ประเภทของสารระสังเขป 31

 สารระสังเขปแบบพรรณนา (Descriptive Abstract) 31

 สารระสังเขปแบบให้ความรู้ (Informative Abstract) 31

 สารระสังเขปแบบพรรณนาและให้ความรู้ (Indicative-Informative

 Abstract) 32

 ปัจจัยที่มีผลต่อการกำหนดรูปแบบการเขียนสารระสังเขป 36

ลักษณะของสาระสังเขปที่มีคุณภาพ	37
ส่วนประกอบของสาระสังเขป	37
ขั้นตอนการเขียนสาระสังเขป	38
การประเมินสาระสังเขป	40
บรรณานุกรม	41
บทที่ 4 การจัดทำดรรชนี (Indexing)	42
ความหมายและวัตถุประสงค์ของดรรชนี	42
ความสำคัญของดรรชนี	43
หน้าที่ของดรรชนี	43
ประเภทของดรรชนี	44
ดรรชนีผู้แต่ง (Author Indexes)	44
ดรรชนีหัวเรื่อง (Subject Indexes)	44
ดรรชนีคำ (Word Indexes)	44
ดรรชนีประสานคำ (Coordinate Indexes)	46
ดรรชนีการอ้างถึง (Citation Indexes)	46
ภาษาดรรชนี (Indexing Language)	48
ประเภทของภาษาดรรชนี	48
ภาษาธรรมชาติ (Natural Language)	48
ภาษาควบคุม (Controlled Language)	49
การกำหนดคำดรรชนี	50
ข้อผิดพลาดที่อาจเกิดขึ้นในการกำหนดคำดรรชนี	51
บรรณานุกรม	53
บทที่ 5 ธิซอรัส (Thesauris)	54
ความหมายและวัตถุประสงค์ของธิซอรัส	54
ขอบเขตของธิซอรัส	55
บทบาทของธิซอรัสในการจัดเก็บและการค้นคืนสารสนเทศ	55
องค์ประกอบของธิซอรัส	56
ลักษณะความสัมพันธ์ของคำในธิซอรัส	58
การจัดแสดงคำสำคัญในธิซอรัส	61
ประโยชน์ของธิซอรัส	63
การพัฒนาธิซอรัส	63
ขั้นตอนการพัฒนาธิซอรัส	65
การประเมินธิซอรัส	68
บรรณานุกรม	69

บทที่ 1

แนวคิดพื้นฐานเกี่ยวกับการจัดเก็บและการค้นคืนสารสนเทศด้วยคอมพิวเตอร์

การจัดเก็บและการค้นคืนสารสนเทศ (Information Storage and Retrieval)

การจัดเก็บและการค้นคืนสารสนเทศ ได้แก่ กระบวนการในการรวบรวมสารสนเทศและทำรายการให้กับสารสนเทศที่รวบรวมไว้ ทั้งนี้เพื่อให้ทราบที่อยู่ของสารสนเทศและสามารถแสดงผลการค้นออกมาตามรูปแบบที่ต้องการ หลักสำคัญในการจัดเก็บและค้นคืนสารสนเทศคือ การกำหนดตัวแทนของเอกสาร (Information representation) และโครงสร้างในการจัดกลุ่มของเอกสาร (Information architecture) เพื่อใช้เป็นกรณีนำทางเข้าถึงเอกสาร ซึ่งการนำคอมพิวเตอร์เข้ามาใช้จะช่วยให้การเข้าถึงเอกสารในลักษณะ Full-text Searching* ทำได้ง่ายขึ้น โดยคอมพิวเตอร์จะจัดเก็บสารสนเทศตามโครงสร้างที่ออกแบบไว้ ซึ่งโครงสร้างระบบของการจัดเก็บและค้นคืนสารสนเทศ (Information storage and retrieval system architecture) ประกอบด้วยฐานข้อมูลเต็มรูปของเอกสาร (Full-text database) จัดเก็บสารสนเทศทุกรูปแบบแบบเต็มรูป ฐานข้อมูลดัชนี (Index database) จัดเก็บคำค้น ทั้งศัพท์อิสระและศัพท์ควบคุม ฐานข้อมูลรายการบรรณานุกรมของสารสนเทศ (Metadata database) จัดเก็บรายละเอียดของสารสนเทศแต่ละรายการ และระบบการเชื่อมประสานกับผู้ใช้ (User interface) ทำหน้าที่เป็นตัวติดต่อรับคำถามจากผู้ใช้และการแสดงผลจากการทำงานของระบบ เพื่อนำทางผู้ใช้ให้พบกับสารสนเทศที่ต้องการ

ในกระบวนการจัดเก็บและค้นคืนสารสนเทศด้วยคอมพิวเตอร์ คอมพิวเตอร์จะเป็นผู้ดำเนินกระบวนการดังกล่าวซึ่งช่วยให้การดำเนินงานทำได้อย่างสะดวกรวดเร็วขึ้น อย่างไรก็ตาม ในการออกแบบระบบเพื่อให้คอมพิวเตอร์สามารถดำเนินการทุกขั้นตอนได้อย่างอัตโนมัตินั้น ยกตัวอย่างเช่น search engines คุณภาพของการดำเนินงานยังไม่อยู่ในระดับที่น่าพึงพอใจ ดังจะเห็นได้จากผลการค้นที่ไม่ตรงกับความต้องการของผู้ค้น เนื่องจากยังคงประสบปัญหาในการกำหนดตัวแทนของเอกสาร (คำดัชนี สารสังเขป) และการกำหนดโครงสร้างของข้อมูล (โครงสร้างการจัดกลุ่มหรือโครงสร้างในการจัดหมวดหมู่) รวมทั้งการออกแบบระบบการเชื่อมประสานที่สามารถนำทางผู้ใช้ให้สามารถประสบความสำเร็จในการค้นคือ ตอบสนองความต้องการในการค้นคืนสารสนเทศ

ข้อควรพิจารณาในการออกแบบระบบการจัดเก็บและค้นคืนสารสนเทศคือ ในการค้นคืนสารสนเทศระบบควรจะสามารถตอบสนองความต้องการในการค้นคืนสารสนเทศของผู้ใช้ได้ ระบบที่ทำงานได้เร็วแต่ไม่ได้ให้ในสิ่งที่ผู้ใช้ต้องการไม่จัดว่าเป็นระบบที่ดี ผลการค้นที่เรียกออกมาได้จำนวนมากแต่ไม่ตรงกับความต้องการของผู้ค้นไม่ได้แสดงว่าระบบนั้นเป็นระบบที่ดี และระบบที่มีขั้นตอนในการทำงานยุ่งยากและต้องใช้เวลามากในการทำงานก็ไม่ใช่ว่าระบบที่พึงปรารถนาเช่นเดียวกัน และการที่ผู้ใช้ไม่สามารถเรียกเอกสารที่ต้องการออกมาจากระบบได้ไม่ได้แสดงว่าระบบล้มเหลวหรือเป็นระบบที่ไม่ดี ระบบสารสนเทศจะมีประสิทธิภาพก็ต่อเมื่อสามารถตอบสนองความต้องการสารสนเทศของผู้ใช้ได้ นั่นคือ ให้สารสนเทศที่ผู้ใช้ต้องการและสามารถนำสารสนเทศนั้นไปใช้ประโยชน์ได้ ซึ่งสารสนเทศจะมีประโยชน์และเรียกได้ว่าเป็นสารสนเทศที่มีคุณภาพก็ต่อเมื่อสารสนเทศนั้นสามารถตอบสนองความต้องการสารสนเทศในขณะนั้นของผู้ใช้ได้

* ความสามารถในการค้นหาเอกสารที่ต้องการโดยการนำคำค้นที่กำหนดไปเปรียบเทียบกับเอกสารทั้งหมด ซึ่งเป็นลักษณะการค้นคืนเอกสารที่จัดเก็บในรูปดิจิทัล (Digital format) หรือเอกสารอิเล็กทรอนิกส์ (Electronic document)

สารสนเทศคืออะไรสำหรับคอมพิวเตอร์

1. สิ่งแทนด้วยกลุ่มสัญลักษณ์

2. มีโครงสร้าง

3. และผู้ใช้สามารถอ่านและเข้าใจได้

* กลุ่มสัญลักษณ์คือ สิ่งที่กำหนดขึ้นแทนตัวสารสนเทศ เพื่อให้คอมพิวเตอร์รู้จักสารสนเทศดังกล่าว

* สัญลักษณ์ในความหมายโดยทั่วไปคือ สิ่งที่กำหนดขึ้นมาเพื่อแทนความหมายของสิ่งใดสิ่งหนึ่ง เช่น ธงชาติ คำพูด เครื่องหมาย ตัวพิมพ์ในเครื่องพิมพ์ดีด โดยสัญลักษณ์จะต้องมีคุณสมบัติ ดังนี้ มีความหมาย มีความแปลกใหม่ มีอยู่จริง และมีความน่าเชื่อถือใช้เป็นตัวแทนของสารสนเทศได้ เช่น

กลุ่มตัวเลข 12345 สื่อความหมายได้หลายอย่าง : ในกรณีสื่อความหมายแทนวันที่ สามารถอ่านค่าได้ ดังนี้

23 มกราคม 1945

3 ธันวาคม 1945

12 มีนาคม 1945

หรือกลุ่มตัวอักษร FOOD FISH, FISH FOOD

หากสลับตำแหน่งจะมีความหมายต่างกัน

ACT, CAT, ATC, TCA

หากสลับตำแหน่งจะมีความหมายต่างกัน

ดังนั้น คุณสมบัติที่สำคัญประการหนึ่งของสัญลักษณ์ก็คือจะต้องมีโครงสร้างที่สามารถสื่อสารให้เข้าใจได้ ซึ่งในการจัดเก็บและค้นหาสารสนเทศด้วยคอมพิวเตอร์จะเกี่ยวข้องกับแนวคิด 2 ประเด็น คือ

1. การกำหนดตัวแทนของสารสนเทศ (Information Representation)

2. การกำหนดโครงสร้างของสารสนเทศ (Information Architecture)

ซึ่งทั้ง 2 ประเด็นนี้เป็นปัญหาใหญ่สำหรับคอมพิวเตอร์ และเป็นสิ่งที่ท้าทายความสามารถของนักสร้างระบบสารสนเทศ ซึ่งจะต้องเป็นผู้กำหนดความหมายและความสำคัญของสัญลักษณ์ที่กำหนดขึ้นแทนตัวสารสนเทศ

การค้นหาสารสนเทศ (Information Retrieval : IR) คืออะไร

ความหมายของการค้นหาสารสนเทศ

Retrieval = Selective Searching = การค้นเพื่อให้ได้สิ่งที่ต้องการกลับคืนมา

การค้นหาสารสนเทศ (Information Retrieval : IR) คือ การค้นหาสารสนเทศที่ต้องการจากคลังสารสนเทศ หรือ ฐานข้อมูล และสามารถเรียกข้อมูลที่ต้องการออกมาแสดงผลบนจอภาพได้อย่างรวดเร็ว ซึ่งในการค้นหาสารสนเทศมักจะมีกระบวนการคัดเลือกเข้ามาเกี่ยวข้อง โดยการคัดเลือกจะดำเนินถึงปัจจัย 3 ประการคือ ราคาที่ต้องจ่าย เวลาที่ต้องใช้และเงินที่ต้องเสียไป (อาจพิจารณาเฉพาะบางปัจจัย หรือทุกปัจจัยก็ได้) พิจารณาจากตัวอย่างข้างล่าง

“การดูรายการข่าวในโทรทัศน์”

ตัวอย่างนี้ไม่ถือเป็นการค้นคืนสารสนเทศ เนื่องจาก ผู้ดูไม่สามารถเลือกเฉพาะรายการข่าวที่ต้องการออกมาดูได้ จะต้องดูไปเรื่อย ๆ จนกว่าจะถึงรายการข่าวที่ต้องการติดตาม หากข่าวนั้นรายงานผ่านไปแล้วก็ไม่สามารถเรียกกลับมาดูได้ การค้นคืนสารสนเทศผู้ค้นจะต้องสามารถเลือกดูเฉพาะรายการที่ต้องการได้โดยไม่ต้องมานั่งดูรายการทั้งหมดที่มีอยู่ (ทั้งนี้เพื่อประหยัดทั้งเวลา และค่าใช้จ่าย) ตัวอย่างที่ถือว่าเป็นการค้นคืนสารสนเทศ เช่น

“การเลือกอ่านหนังสือในห้องสมุด”

ทั้งนี้เนื่องจาก ผู้ใช้ห้องสมุดไม่จำเป็นต้องอ่านหนังสือทั้งหมดภายในห้องสมุด (เพื่อค้นหาสารสนเทศที่ต้องการ) แต่สามารถเลือกอ่านเฉพาะเล่มที่ต้องการได้

การดำเนินงานของระบบการจัดเก็บและค้นคืนสารสนเทศ

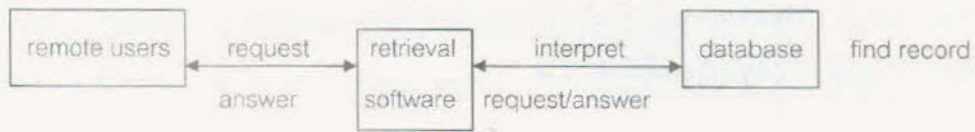
ระบบการจัดเก็บและค้นคืนสารสนเทศเป็นกระบวนการที่ได้รับการออกแบบมาเพื่อวัตถุประสงค์ในการจัดเก็บสารสนเทศเพื่อให้ง่ายต่อการค้นคืนสารสนเทศที่จัดเก็บอยู่ในระบบนั้น โดยจะทำหน้าที่ รวบรวม ทำรายการ จัดโครงสร้างของข้อมูล ช่วยค้น คัดเลือก และกลั่นกรองสารสนเทศที่เกี่ยวข้อง ซึ่งอาจดำเนินการโดยแรงงานคนหรือคอมพิวเตอร์ก็ได้ ระบบการค้นคืนสารสนเทศประกอบด้วยระบบการดำเนินงานย่อย ๆ 5 ระบบ คือ

1. การคัดเลือก (Selective) คือ การรวบรวมสารสนเทศตามนโยบายที่กำหนดไว้เพื่อให้สอดคล้องกับความต้องการของผู้ใช้
2. การวิเคราะห์เอกสาร (Document analysis) ได้แก่ การจัดหมวดหมู่ การทำรายการ การทำดัชนี และการทำสาระสังเขป การวิเคราะห์เอกสารมี 2 ขั้นตอน คือ 1) วิเคราะห์แนวคิด หรือเนื้อหาของเอกสาร 2) กำหนดศัพท์พรรณานี้ หรือ กำหนดหมวดหมู่ตามแนวคิดหรือเนื้อหาของเอกสาร
3. การจัดระเบียบแฟ้มข้อมูล (File organization) มี 2 ประเภท คือ 1) Linear file เป็นแฟ้มหลักที่จัดเก็บข้อมูลที่ใช้ค้นคืนเรียงตามลำดับเลขทะเบียน 2) Invert file หรือ Index file เป็นแฟ้มรายการศัพท์พรรณานี้ ซึ่งใช้ค้นข้อมูลใน Linear file
4. การค้นคืน (Retrieval) ได้แก่ การสร้างกลยุทธ์ในการค้น (Search strategy) เพื่อให้ได้ผลการค้นตาม que ผู้ใช้ต้องการ ประกอบด้วย 2 ขั้นตอนคือ 1) วิเคราะห์คำถามเพื่อทราบความต้องการที่แท้จริงของผู้ใช้ 2) กำหนดคำศัพท์ให้สอดคล้องกับแนวคิดตามความต้องการนั้น เพื่อนำคำศัพท์ที่กำหนดได้ไปดำเนินการค้นคืนจากฐานข้อมูล
5. การแจกจ่าย (Dissemination) ได้แก่ การนำผลการค้นคืนส่งให้กับผู้ใช้ โดยทั่วไป คือ รายการบรรณานุกรมและสาระสังเขป การแจกจ่ายสารสนเทศมี 2 รูปแบบ คือ 1) การจัดส่งเอกสาร (Document delivery) คือ การนำส่งตัวเอกสารที่มีเนื้อหาที่ผู้ใช้ต้องการ 2) การจัดส่งรายการเอกสาร (Notification of documents) คือ การจัดส่งรายการเอกสารที่คาดว่าจะอยู่ในความต้องการของผู้ใช้ อย่างไรก็ตาม กระบวนการการค้นคืนสารสนเทศจะสิ้นสุดลงเมื่อผู้ใช้ได้รับสารสนเทศที่ต้องการเท่านั้น

ระบบการค้นคืนสารสนเทศด้วยคอมพิวเตอร์ มี 2 ลักษณะ คือ

ระบบการค้นคืนสารสนเทศด้วยคอมพิวเตอร์ประกอบด้วย ฮาร์ดแวร์ ฐานข้อมูล และโปรแกรมการค้นคืน เพื่อให้ผู้ใช้และฐานข้อมูลสามารถโต้ตอบกันได้ มีลักษณะการให้บริการ 2 ลักษณะ คือ

- 1) Central service คือมี server เป็นตัวให้บริการซึ่งผู้ใช้จากภายนอกสามารถเข้ามาใช้บริการได้โดยอาศัยเครือข่ายโทรคมนาคม
- 2) Local system ทำงานในลักษณะ stand alone อ่านข้อมูลจากฐานข้อมูลซีดี-รอม



ซึ่งปัญหาสำคัญในการค้นคืนสารสนเทศคือ การกำหนดตัวแทนหรือคำแทนเนื้อหาของเอกสาร ได้แก่ ศัพท์ดัชนี (Index term) ต้องกำหนดให้ถูกต้องและเพียงพอกับเนื้อหาในแต่ละสาขาวิชา และเป็นสัญลักษณ์ที่สื่อความหมายตรงกันระหว่างผู้กำหนดคำศัพท์และผู้ใช้ ดังนั้น ในฐานข้อมูลจึงควรต้องประกอบด้วย คำศัพท์ที่เป็นภาษารธรรมชาติ (Natural language) ซึ่งเป็นคำที่ดึงออกมาจากเนื้อหาของเอกสารเพื่อช่วยเหลือผู้ใช้ในการค้นคืนสารสนเทศ ผู้ใช้จะต้องสร้าง query (คำถาม: สิ่งที่ต้องการให้เครื่องประมวลผลให้) เพื่อบอกสิ่งที่ต้องการให้คอมพิวเตอร์รู้ อาจใช้ภาษาอิสระหรือภาษาคอมพิวเตอร์โดยใช้กลวิธีการค้นของผู้เชี่ยวชาญซึ่งจะช่วยให้การค้นคืนสารสนเทศทำได้เร็วขึ้นและตรงกับความต้องการมากขึ้น ซึ่งการกำหนด query มักจะยุ่งยากซับซ้อน ปัญหาสำคัญในการกำหนด query คือ จะสร้าง query อย่างไรจึงจะถ่ายทอดความต้องการของผู้ใช้ออกมาได้ตรงที่สุด เพื่อจะนำไปบรรจบ (match) หรือ เปรียบเทียบ (compare) กับสารสนเทศที่มีอยู่ในฐานข้อมูล นอกจากนี้ผู้ใช้อาจยังสามารถกำหนดรูปแบบของเอกสารที่ต้องการไว้ใน query ได้ด้วย

องค์ประกอบของระบบการจัดเก็บและค้นคืนสารสนเทศ

การค้นคืนสารสนเทศจากฐานข้อมูลเกิดขึ้นจากคน 2 กลุ่ม คือ

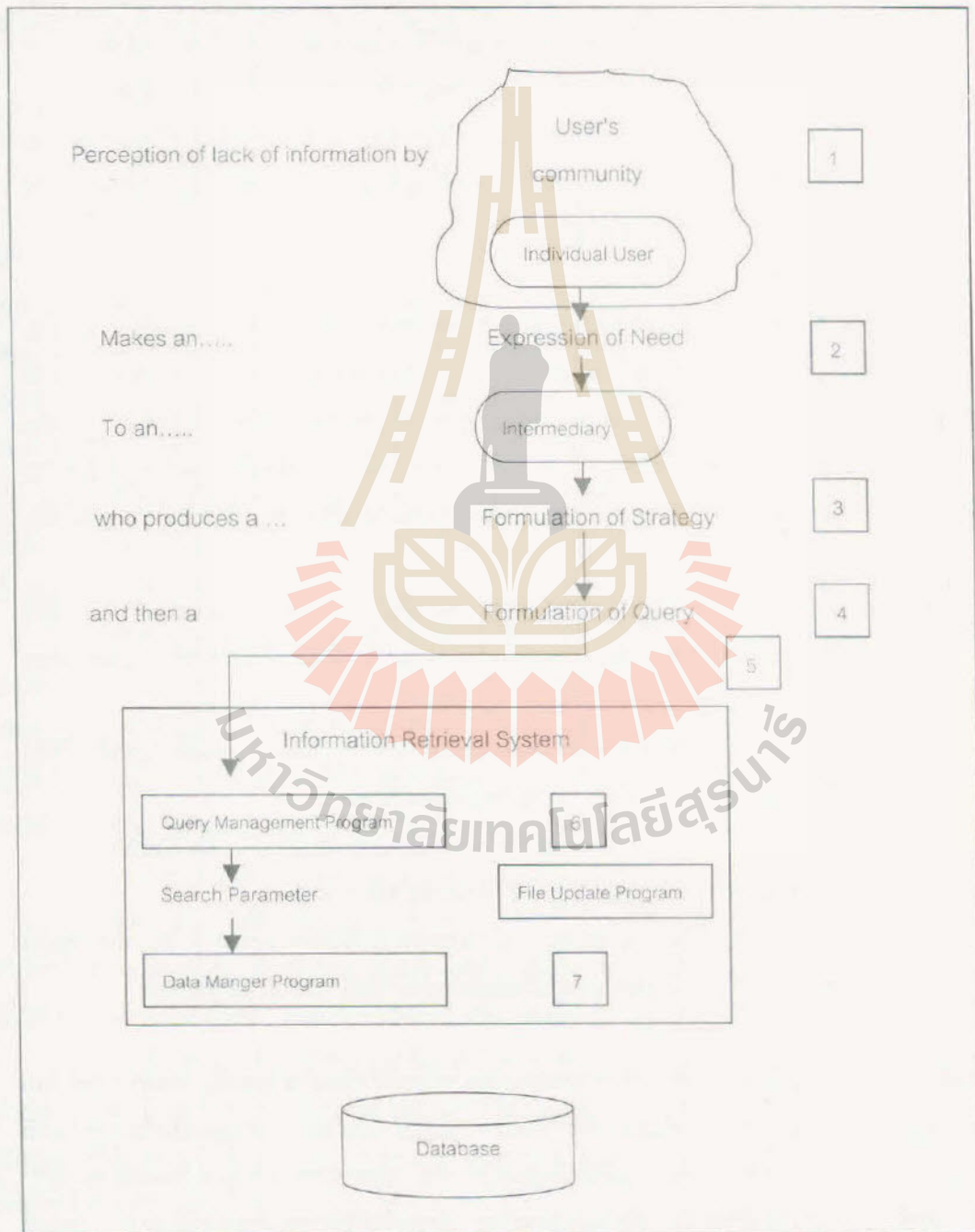
- 1) ผู้ใช้สารสนเทศ (Users of information / Prospective users / Information consumers) ซึ่งเป็นผู้ที่ต้องการสารสนเทศ และใช้ระบบสารสนเทศเพื่อให้ได้สารสนเทศที่ต้องการ เพื่อนำสารสนเทศไปใช้ในสถานการณ์ต่าง ๆ หรือเพื่อสร้างสรรสิ่งขึ้น ๆ ต่อไป
- 2) ผู้จัดระบบสารสนเทศ (Database producers) ทำหน้าที่รวบรวม จัดการ และจัดเก็บสารสนเทศอย่างเป็นระบบเพื่อประโยชน์ในการค้นคืนในภายหลัง ซึ่งมีขั้นตอนในการปฏิบัติ ดังนี้
 - 2.1) ศึกษาเพื่อทราบประเภท / รูปแบบของสารสนเทศที่ผลิตออกมาในปัจจุบัน
 - 2.2) ศึกษาความต้องการของผู้ใช้
 - 2.3) รวบรวมสารสนเทศเหล่านั้น นำมาจัดระบบ และจัดเก็บไว้ในฐานข้อมูล

ผู้ใช้สารสนเทศจะถ่ายทอดความต้องการผ่านทางผู้ให้บริการสารสนเทศ (Information intermediary) ซึ่งจะเป็นผู้ที่ให้ความช่วยเหลือแก่ผู้ใช้ เนื่องจากผู้สร้างฐานข้อมูลและผู้ใช้สารสนเทศมักจะมีมุมมองที่ต่างกัน สิ่งที่ผู้สร้างคิดว่าง่ายอาจกลายเป็นสิ่งที่ยากสำหรับผู้ใช้ ปัญหาที่มักเกิดขึ้นเสมอในการออกแบบระบบคือ ผู้ผลิตไม่ทราบความต้องการที่แท้จริงและพฤติกรรมในการค้นคืนสารสนเทศของผู้ใช้ ขณะที่ผู้ใช้ไม่ทราบหลักการใน

การจัดเก็บและกลวิธีที่ถูกต้องในการค้นคืนที่ผู้ผลิตได้ออกแบบมา ดังนั้น จึงต้องมีผู้ให้บริการทำหน้าที่เป็นตัวกลางสื่อสารความเข้าใจระหว่างคนทั้งสองกลุ่ม แปลงความต้องการของผู้ใช้สารสนเทศให้อยู่ในรูปแบบที่ระบบเข้าใจ เพื่อดึงสารสนเทศที่ผู้ใช้ต้องการออกมาจากฐานข้อมูลได้อย่างรวดเร็ว (แสดงความสัมพันธ์ได้ดังภาพ)



1. กระบวนการของผู้ใช้



ภาพที่ 1.1 การสื่อสารระหว่างผู้ใช้กับระบบจัดเก็บและค้นคืนสารสนเทศ

1.1 ความต้องการสารสนเทศของผู้ใช้

ในกรณีที่ผู้ใช้ได้รับมอบหมายงานชิ้นใหม่ ผู้ใช้มักจะรู้สึกว่ายังขาดความรู้ความเข้าใจในเรื่องดังกล่าว หรือมีความรู้สึกว่ามีข้อมูลที่มีอยู่ยังไม่เพียงพอและต้องการข้อมูลมากกว่าที่มีอยู่ ซึ่งความต้องการข้อมูลดังกล่าว เป็นจุดเริ่มต้นของกระบวนการค้นคืนสารสนเทศ (Information retrieval process)

อย่างไรก็ตาม ความต้องการสารสนเทศของผู้ใช้ ณ จุดนี้ ไม่ได้หมายความว่า " ผู้ใช้จะทราบว่า จริง ๆ แล้วเขาต้องการอะไร หรือยังขาดข้อมูลอะไรบ้าง " เช่น ในกรณีของแพทย์ที่ต้องรักษาผู้ป่วยโรคร้ายแรง หรือ ทนายความที่ว่าความในคดีที่ซับซ้อน ในกรณีนี้ทั้งแพทย์และทนายความทราบว่าเขาต้องการข้อมูลเกี่ยวกับโรค / คดีดังกล่าว แต่อย่างไรก็ตาม แพทย์ยังคงไม่ทราบว่า ควรจะใช้วิธีการใดทดสอบโรค หรือจะรักษาผู้ป่วยด้วยวิธีการใด ขณะที่ทนายความยังคงไม่ทราบว่า จะต้องอ่านเงื่อนไขกฎหมายข้อใด หรือควรศึกษาจากคดีใด

ดังนั้น สิ่งที่สำคัญที่สุดในกระบวนการทั้งหมดของการค้นคืนสารสนเทศ จึงอยู่ที่การวิเคราะห์ความต้องการของผู้ใช้ว่าแท้จริงแล้วผู้ใช้ต้องการอะไร เพื่อจะแปลงความต้องการนั้นเข้าไปสู่กระบวนการของการค้นคืนสารสนเทศ เพื่อให้ได้ผลการค้นที่ผู้ใช้พึงพอใจ

1.2 การแปล หรืออธิบายความต้องการ

ขั้นตอนต่อมา ผู้ใช้จะต้องอธิบายหรือแปลความต้องการสารสนเทศให้ Information agency (อาจเป็น บุคคล หรือเครื่องจักร) ทราบ เรียกบุคคลกลุ่มนี้ว่า Search intermediary หรือ Simply intermediary ทำหน้าที่ ถ่ายทอดความต้องการของผู้ใช้ให้คอมพิวเตอร์เข้าใจ ซึ่งหน้าที่นี้อาจเขียนโปรแกรมให้คอมพิวเตอร์ทำงานแทนได้ เนื่องจากการอธิบายความต้องการในขั้นตอนนี้ คือ การแปลความต้องการให้อยู่ในภาษาที่ระบบสามารถ เข้าใจได้ แต่เนื่องด้วยเหตุผลหลาย ๆ ประการ ในหลายครั้งการแปลความต้องการมักมีความผิดพลาดเกิดขึ้น เสมอ

ตัวอย่างเช่น ในกรณีของวิศวกรที่มีคำถามว่าต้องการข้อมูลเกี่ยวกับ "Processing of jewelry" ผู้ให้บริการต้องตีความว่า จริง ๆ แล้วเขาต้องการข้อมูลเกี่ยวกับอะไร ? สมมติว่าคำตอบที่ต้องการคือ "information about use of laser technology to drill holes in diamonds" แต่จากคำถาม

วิศวกรหรือความแตกต่างระหว่างสิ่งที่เขาถามกับสิ่งที่เขาต้องการหรือไม่ กรณีนี้ปัญหาคือผู้ใช้ไม่ ทราบความต้องการที่แท้จริงของตนเอง หรือผู้ใช้อยู่ยังขาดความรู้ในเรื่องที่ถาม

เขาคิดว่า JEWELRY PROCESSING กับ DIAMOND DRILLING เป็นเรื่องเดียวกันใช่หรือไม่ กรณีนี้ปัญหาคือ ผู้ให้บริการ และ/หรือผู้ใช้ ขาดความรู้หรือความเข้าใจในเรื่องที่ค้น

หรือที่วิศวกรใช้คำค้นนี้ (Processing of jewelry) เพราะเกรงว่าผู้ให้บริการจะไม่เข้าใจ เขาควรจะให้ คำอธิบายง่าย ๆ แก่ผู้ให้บริการ ? กรณีนี้ปัญหาอยู่ที่ ผู้ใช้ไม่มั่นใจในตัวผู้ให้บริการ (ไม่รู้ในเรื่องที่ค้น)

อีกตัวอย่างหนึ่ง ผู้ใช้ต้องการค้นหาคำอธิบายความหมายโดยทั่วไปของโรคโรคหนึ่ง โดยค้นจากชื่อโรค แต่ค้นคืนข้อมูลจากฐานข้อมูลที่มีเฉพาะข้อสนเทศทางการวิจัย ดังนั้น ผลที่ได้คือ รายละเอียดในการรักษาที่ ซับซ้อนและสาเหตุของโรค แต่ความต้องการที่แท้จริงของผู้ใช้ คือ "ต้องการทราบลักษณะโดยทั่วไปของโรค" ไม่ใช่ "รายละเอียดในการรักษาที่ซับซ้อน หรือสาเหตุของโรค" ข้อมูลที่ได้จึงไม่ตรงกับความต้องการ ขณะที่ แพทย์ที่ต้องการค้นหาข้อมูลเพื่อนำไปใช้ในการรักษา เมื่อค้นโดยใช้คำศัพท์ตัวเดียวกัน (ชื่อโรค) และด้วยเหตุผล เดียวกัน คือ เฉพาะชื่อโรคอาจยังไม่ชัดเจนพอ ระบบยังต้องการข้อมูลอื่น ๆ เพิ่มเติม เพื่อให้ได้ข้อสนเทศที่ตรง

กับความต้องการของผู้ใช้มากที่สุด แต่จากคำค้นเดียวกันทั้งสองคนจะได้ผลการค้นที่เหมือนกัน แต่จะมีเพียงคนเดียวเท่านั้นที่พึงพอใจกับผลการค้น สำหรับอีกคนข้อมูลที่ได้กลับไม่มีประโยชน์ จะเห็นได้ว่าทวิเคราะห์ความต้องการมีความสำคัญต่อการประมวลผลการค้นคืน ดังตัวอย่างที่วิเคราะห์ความต้องการผิด (กำหนดวัตถุประสงค์ผิด) ทำให้เลือกใช้ฐานข้อมูลที่ไม่ถูกต้อง (ดำเนินการผิด) ส่งผลให้ผลการค้นคืนที่ได้ไม่ตรงกับความต้องการของผู้ใช้ (ผลลัพธ์ผิด)

ในกระบวนการของการแปลงความต้องการ ผู้ให้บริการไม่จำเป็นต้องแปลงความต้องการของผู้ใช้ให้อยู่ในรูปสัญลักษณ์ เช่น ตัวเลข เลขหมู่ระบบ DDC หรือ ศัพท์สัมพันธ์ หรือ คำสำคัญ แต่อาจใช้ศัพท์อิสระแทนได้ อย่างไรก็ตาม ผู้ให้บริการจะต้องคำนึงถึงการใช้ศัพท์เทคนิคด้วย โดยเฉพาะในการค้นคืนสารสนเทศเฉพาะเรื่อง ซึ่งผู้ให้บริการจะต้องมีความรู้เกี่ยวกับศัพท์พื้นฐานที่จำเป็นในสาขาวิชาเหล่านั้น ในการแปลงความต้องการของผู้ใช้ ผู้ให้บริการจะต้องตระหนักถึงความสำคัญของขั้นตอนนี้ให้มาก เพราะหากแปลงความต้องการออกมาผิดหรือเพี้ยน จะส่งผลต่อการทำงานในขั้นต่อไปทำให้ได้ผลการค้นออกมาไม่ตรงกับความต้องการของผู้ใช้

1.3 การสร้างกลวิธีการค้น

ในกรณีที่ผู้ใช้ทำการค้นเองมักไม่ให้ความสำคัญกับการสร้างสูตรการค้นที่สมบูรณ์ ผู้ใช้ส่วนใหญ่มักจะคิดว่าการเขียนประโยคคำสั่งให้สมบูรณ์นั้นทำให้เสียเวลา จึงมักจะไม่ให้เวลากับการคิดคำค้นและสร้างสูตรการค้น เช่น ในกรณีการค้นหาหนังสือในห้องสมุด ผู้ใช้มักจะไม่นับถรรายการ หรือรายการของห้องสมุดก่อน แต่จะตรงไปที่ชั้นหนังสือที่คุ้นเคยและเริ่มค้นหาหนังสือด้วยความหวังว่าจะเจอหนังสือที่เขาต้องการ แต่กลับกลายเป็นว่าเสียเวลามากขึ้นเพราะไม่ทราบตำแหน่ง หรือที่อยู่ที่เหมาะสมของหนังสือ เมื่อไม่พบหนังสือที่ต้องการผู้ใช้อาจคิดว่า ห้องสมุดไม่มีหนังสือเล่มดังกล่าว ทำให้ไม่ประทับใจต่อบริการของห้องสมุด

กลวิธีในการค้น (Search strategy) เป็นการวางแผนเพื่อค้นคืนสารสนเทศที่ต้องการ ซึ่งนักสารสนเทศ (Information specialist) มักจะเป็นผู้กำหนดกลวิธีในการค้นคืน อย่างไรก็ตาม กิจกรรมนี้จะไม่เกิดขึ้นหากผู้ใช้ไม่ได้วางแผนจะไปพบผู้ให้บริการดังกล่าว

การกำหนดกลวิธีการค้นทำได้หลายระดับ บางคนอาจค้นโดยใช้ศัพท์กว้าง ๆ บางคนอาจใช้คำที่ชัดเจนและเจาะจงว่าต้องการเฉพาะเรื่องอะไร และต้องการเฉพาะสารสนเทศประเภทใด ต้องการเฉพาะที่ผลิตออกมาในช่วงปีใด และให้แสดงรายละเอียดด้วย เป็นต้น เช่น ในกรณีของวิศวกร (ช่างต้น) ที่ต้องการสารสนเทศเกี่ยวกับ "Drilling problem" ควรค้นคืนโดยใช้คำศัพท์เหล่านี้ (DIAMOND OR JEWELRY) AND DRILLING หรือ DIAMOND AND DRILLING และเลือกเฉพาะที่มีระบบควบคุม หรือระบบการวิจัย ช่วงปี 1995 - ปัจจุบันเท่านั้น

1.4 การสร้าง Query

ขั้นตอนนี้จะเกี่ยวข้องกับการเรียบเรียงความต้องการสารสนเทศให้อยู่ในรูปแบบที่เหมาะสมกับระบบการค้นคืนสารสนเทศที่ใช้ สำหรับระบบคอมพิวเตอร์การสร้าง query ก็เหมือนกับการเขียน equivalent ของโปรแกรมคอมพิวเตอร์ หรือ ความต่อเนื่องของคำสั่งที่จะส่งผลให้ได้สารสนเทศที่ต้องการ

query คือ รูปแบบที่ถูกต้องของการค้น โดยใช้คำที่เหมาะสมกับกลไกการค้นคืน ซึ่งมีองค์ประกอบที่สำคัญ 2 ประการ คือ 1) คำศัพท์ (term) คำค้น (words) รหัส (code) หรือ คำที่ใช้แทนระเบียบที่ผู้ใช้ต้องการค้นคืน 2) ตรรกะที่ใช้เชื่อมความสัมพันธ์ระหว่าง ชุดของคำ หรือ คำที่แตกต่างกัน เช่น DIAMOND AND DRILLING คือ คำศัพท์ที่เชื่อมด้วยตรรกะ AND

วัตถุประสงค์หลักของ query คือ ขยายศัพท์พรรณานี้ (Index terms) หรือสัญลักษณ์อื่น ๆ ให้ครอบคลุมเนื้อหาของเอกสารเพื่อให้ได้เอกสารที่ต้องการ แต่ปัจจุบันการสร้าง query หมายถึงสิ่งที่กำหนดขึ้นมาแทนตัวเอกสาร (แทบจะบ่งบอกถึงตัวเอกสารที่จัดเก็บในฐานข้อมูล)

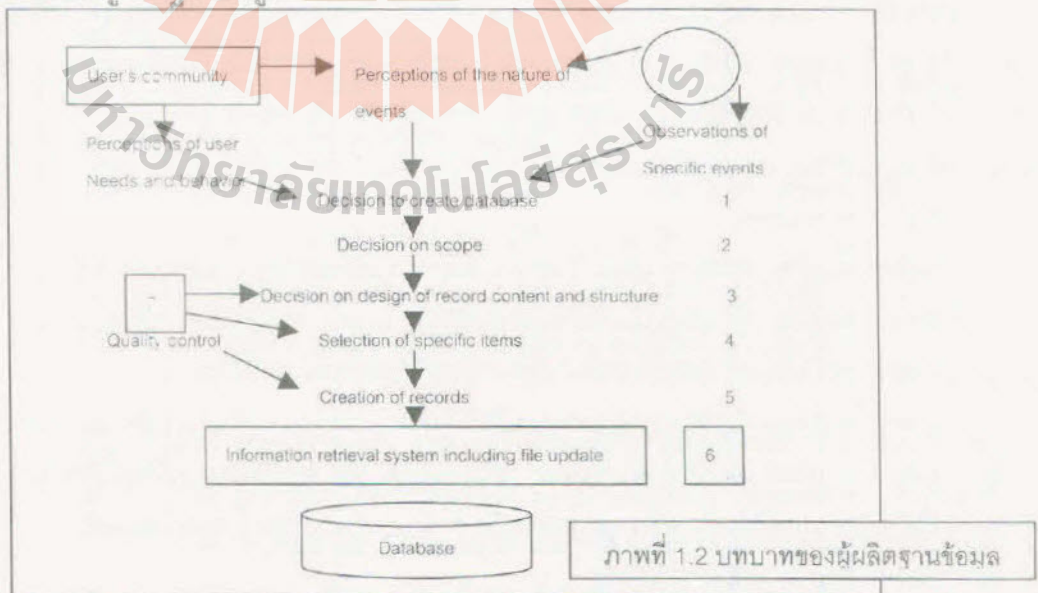
1.5 การประมวล Query ของคอมพิวเตอร์

คอมพิวเตอร์จะแปล query โดยอาศัยกระบวนการแปล บางระบบจะแปลอย่างตรงไปตรงมาและเที่ยงตรง คอมพิวเตอร์จะค้นคืนสิ่งที่ query ต้องการได้อย่างเที่ยงตรง ไม่มากไม่น้อยเกินไป ยกเว้นกรณีที่เกิดความผิดพลาดทางโปรแกรม การประมวลผลด้วยคอมพิวเตอร์ประกอบด้วย 2 ขั้นตอนใหญ่ ๆ คือ

- 1) การแปล query ให้อยู่ในรูปของ Search parameters set ซึ่งจะชี้เฉพาะลงไปว่ากำลังค้นหาเพิ่มข้อมูลใด หรือกำลังค้นหาคำใด และจะใช้เงื่อนไขใดในการค้นคืน ระบบการค้นคืนสารสนเทศเชิงพาณิชย์ส่วนใหญ่จะมีระบบการแปลที่ไม่ซับซ้อน เนื่องจากบริษัทให้ความสำคัญกับภาษาที่ใช้อธิบาย query ในกรณีที่ใช้ภาษารวมชาติการสร้าง query อาจต้องใช้การพิจารณามากขึ้น ซึ่งมีกระบวนการที่ค่อนข้างยุ่งยากและการตัดสินใจ อาจทำได้ไม่สมบูรณ์ เรียกโปรแกรมที่ทำงานนี้ว่า Query manager
- 2) ดำเนินการค้นตาม Search parameters อย่างไรก็ตาม หากทำได้อย่างรัดกุม กระบวนการนี้มักจะดำเนินไปตามที่กำหนดไว้ แทบจะไม่มีส่วนไหนที่คอมพิวเตอร์ต้องตัดสินใจอีก เรียกกระบวนการทำงานนี้ว่า Data manager

ผู้ใช้ต้องตระหนักว่า ผลการค้นคืนของคอมพิวเตอร์มักจะไม่ตรงกับความต้องการเลยทีเดียว ทั้งนี้อาจเนื่องจากการกำหนด query ไม่สมบูรณ์ หรือรายการนั้นไม่มีอยู่ในฐานข้อมูล และที่สำคัญ คือ ผู้ใช้ ผู้ให้บริการและผู้ผลิตมักจะมีมุมมองที่ต่างกันทำให้การกำหนดคำค้นหรือ query แตกต่างกันไป แต่คอมพิวเตอร์จะเปรียบเทียบ query กับระเบียบไปตามตัวอักษร จึงทำให้ผลการค้นที่ออกมาไม่เป็นไปตามที่คาดหวังไว้

2. กระบวนการของผู้ผลิตฐานข้อมูล



ความต้องการสารสนเทศของผู้ใช้อาจเกิดขึ้นจากภายในองค์กรหนึ่งหรือชุมชนหนึ่ง ขณะที่อีกหน่วยงานหนึ่งเป็นผู้สร้างฐานข้อมูลเพื่อตอบสนองความต้องการขององค์กรแรก เช่น สมุดโทรศัพท์ แฟ้มข้อมูลรายการอ้างอิงและบทความที่ตีพิมพ์ในวารสารทางการแพทย์ รายชื่อวารสารที่จัดจำหน่าย ข้อมูลตลาดหุ้น เป็นต้น

สิ่งที่ต้องพิจารณาก่อนที่จะเริ่มรวบรวมสารสนเทศ ได้แก่ ขอบเขตของเรื่องที่จะรวบรวม รายละเอียดที่ต้องจัดเก็บไว้ในระเบียบ นอกจากนี้ยังต้องคำนึงถึงรูปแบบในการนำเสนอ ตัวอย่างเช่น การสร้างฐานข้อมูลตลาดหุ้น สิ่งที่จะต้องพิจารณาคือ จะต้องรวบรวมหุ้นประเภทใดบ้าง ความถี่ในการปรับปรุงข้อมูล และจะนำเสนอรายละเอียดเหล่านี้อย่างไร stock symbol, number of shares, price, journal title, abstract, article title เป็นต้น

ขั้นตอนในการสร้างฐานข้อมูลประกอบด้วย

- 1) การตัดสินใจที่จะสร้างฐานข้อมูล เช่น การตัดสินใจออกหนังสือพิมพ์ฉบับใหม่
- 2) กำหนดขอบเขตของเรื่องที่จะรวบรวม เช่น จะต้องมีการรวบรวมอะไรบ้าง ข่าวประจำวัน ข่าวการเงิน ข่าวท้องถิ่น คอลัมน์พิเศษอื่น ๆ เป็นต้น
- 3) ออกแบบระเบียบของแต่ละรายการ เช่น ข่าวแต่ละประเภทต้องมีรายละเอียดอะไรบ้าง ใครทำอะไร ที่ไหน เมื่อไร อย่างไร
- 4) เลือกรายการที่จะนำมาบันทึกไว้ในฐานข้อมูล เช่น ในฉบับที่จะออกวันนี้ จะเลือกเรื่องใด หรือข่าวใดลงตีพิมพ์บ้าง
- 5) สร้างเนื้อหาของแต่ละระเบียบ เช่น การรวบรวมข่าว เขียนข่าว และตรวจสอบข่าว ตัวอย่างเช่น ในกรณีของข้อมูลตลาดหุ้น ตารางดัชนีราคาหุ้นซึ่งเป็นข้อมูลตัวเลขจะต้องถูกต้องและเที่ยงตรงที่สุด การกำหนดศัพท์หรือการจัดหมวดหมู่จะต้องทำโดยมนุษย์ ระบบจะต้องตอบคำถามเกี่ยวกับราคาล่าสุดของหุ้นได้ยกเว้นกรณีที่เกิดความผิดพลาดทางโปรแกรม จำนวนของรายการอ้างอิงที่ค้นได้ขึ้นอยู่กับปัจจัยต่าง ๆ ดังนี้

ก. บทความทั้งหมดในเรื่องที่ต้องการมีอยู่ในฐานข้อมูลหรือไม่ (Scope)

ข. กฎเกณฑ์และภาษาที่ระบบกำหนดให้ใช้เพื่ออธิบายเรื่องที่ต้องการ จะอธิบายลักษณะขององค์ประกอบแต่ละตัวอย่างไร (Record design)

ค. การกำหนดชื่อเรื่อง/ดัชนีของ Cataloger เช่น พิจารณาว่า จะให้ไว้ที่ หัวเรื่อง DRILLING DIAMONDS หรือ PROCESSING JEWELRY (Record content)

ง. การสร้าง query กำหนดขึ้นมาแทนความต้องการสารสนเทศได้ดีแค่ไหน (Query formulation)

จ. ระเบียบที่ค้นคืนได้ตรงกับความต้องการของผู้ใช้มากน้อยเพียงไร (Output evaluation)

- 6) ปรับปรุงฐานข้อมูล ปัจจุบันหนังสือพิมพ์ส่วนใหญ่ นักข่าวสามารถพิมพ์ข่าวลงในคอมพิวเตอร์ได้โดยตรง จะไม่มีกระดาษเกิดขึ้นจนกว่าหนังสือพิมพ์จะถูกตีพิมพ์ออกมา ในฐานข้อมูลเกือบทุกประเภทจะคำนึงถึง Keystroking operation โดยจะพยายามออกแบบระบบให้ผู้ใช้พิมพ์คำสั่งให้สั้นที่สุด หรือน้อยที่สุด ในการทำสำเนา (Copy) ระเบียบข้อมูลที่ต้องการ นอกจากนี้ ควรพิจารณาถึงวิธีการในการสร้างระเบียบ เมื่อเคาะ Enter นั้นหมายความว่าระเบียบได้ถูกนำไป

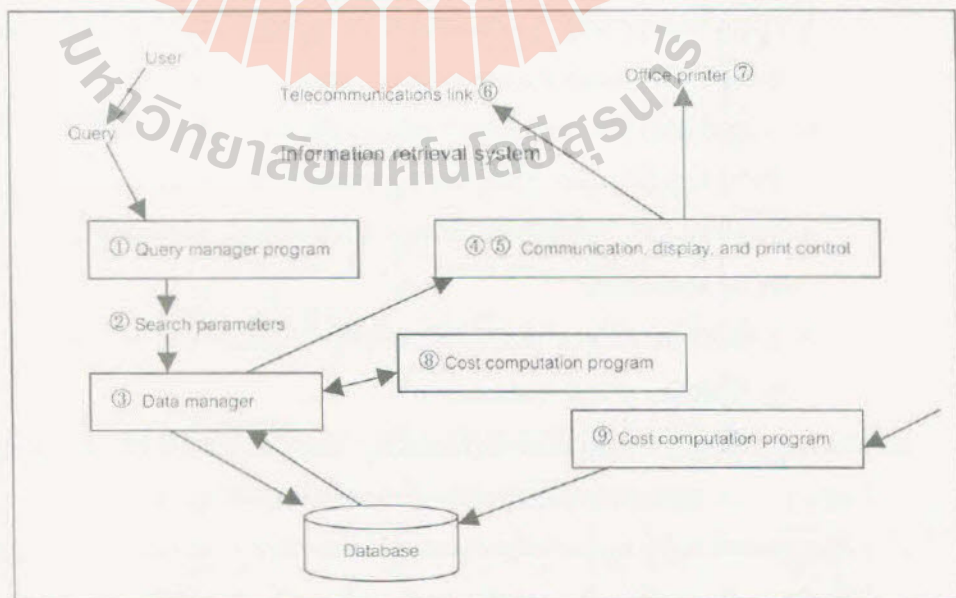
จัดเก็บในตำแหน่งที่เหมาะสมแล้ว เรียกกระบวนการนี้ว่า Updating หรือ Database maintenance การเพิ่มระเบียบไม่ใช่เป็นเพียงแค่การเพิ่มรายการในแฟ้มข้อมูลเท่านั้น แต่ยังหมายรวมถึงการสร้างดัชนีหรือการปรับปรุงแฟ้มดัชนีด้วย ซึ่งอาจทำโดยคอมพิวเตอร์ ทั้งนี้เพื่อเพิ่มประสิทธิภาพในการค้น การปรับปรุงแฟ้มดัชนีต้องทำอย่างสม่ำเสมอ ควรทำทุกครั้งที่มี ข้อมูลใหม่ หรือระเบียบใหม่เพิ่มเข้ามา หรือเมื่อมีการลบข้อมูลเก่าออกไป

- 7) การควบคุมคุณภาพ ต้องทำทั้งในส่วนของการสร้างระเบียบ และกระบวนการหลังจากตรวจสอบความถูกต้องของฐานข้อมูล การควบคุมคุณภาพจะเกี่ยวข้องกับความถูกต้องและความสม่ำเสมอของรายการที่ค้นคืนได้ โดยเปรียบเทียบกับคำสั่งที่ใช้ค้น ทั้งนี้เพื่อป้องกันความผิดพลาดที่อาจเกิดขึ้น

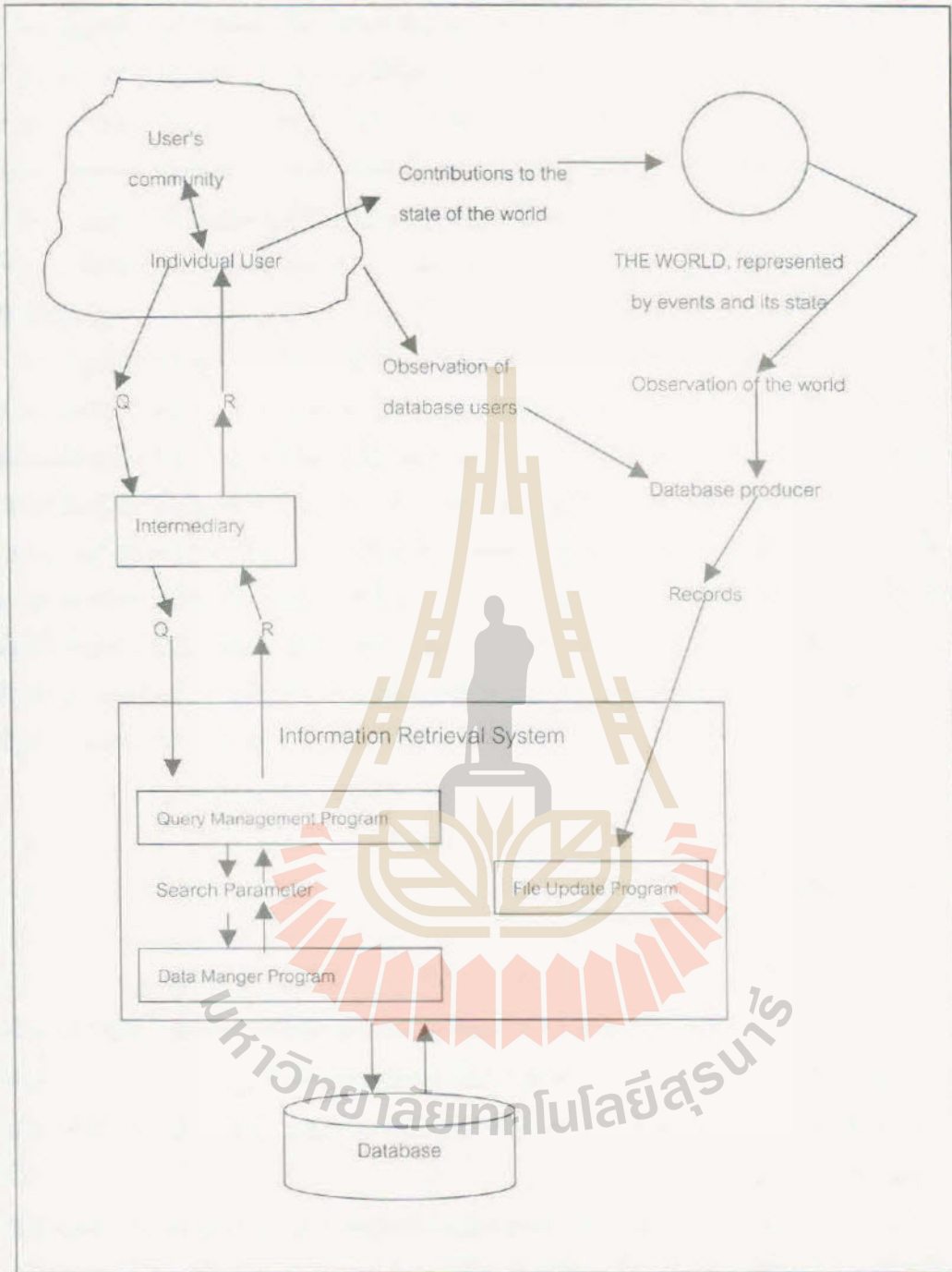
3. กระบวนการการทำงานของคอมพิวเตอร์

อธิบายกระบวนการทำงานของคอมพิวเตอร์ได้ง่าย ๆ ดังนี้

- 1) แปล Query
- 2) ดึงคำสำคัญ ออกมาและค้นหาในแฟ้มดัชนี
- 3) แจ้งจำนวนที่พบให้ผู้ใช้ทราบ จากนั้นให้ผู้ใช้ตรวจดูเพื่อกำหนดคำตรรกะที่จะใช้ค้นต่อไป ปรับเปลี่ยน query หากข้อมูลที่ค้นได้ในขั้นนี้ยังไม่เป็นที่พึงพอใจ หรือพิมพ์ผลการค้นหากผลการค้นที่ได้เป็นที่พึงพอใจ ระบบจะแสดงข้อความแจ้งผลการค้นพร้อมทั้งระเบียบที่ค้นได้
- 4) แสดงระเบียบที่ตรงกับ query
- 5) โอนถ่ายข้อมูลให้ผู้ใช้ หรือพิมพ์ผลการค้น
- 6) แจ้งผลการค้นและค่าใช้จ่ายให้ผู้ใช้ทราบทาง E-mail
- 7) แยกลำดับการทำงาน
- 8) เพิ่มรายการใหม่เข้าไปในฐานข้อมูล หรือตรวจสอบความถูกต้อง หรือลบทิ้ง



ภาพที่ 1.3 โปรแกรมหลักในการจัดเก็บและค้นคืนสารสนเทศ



ภาพที่ 1.4 วงจรของระบบสารสนเทศ

4. สาเหตุที่กระบวนการค้นคืนไม่สมบูรณ์

หากประเมินในฐานข้อมูลประกอบไปด้วย natural language text, graphics หรือ รูปแบบอื่น ๆ ที่ซับซ้อน หรือมีการกำหนดรหัส มีความเป็นไปได้ว่า ผู้ใช้มักจะไม่เข้าใจวิธีการค้น หรือไม่ได้สารสนเทศที่ต้องการในการค้นครั้งแรก นอกจากนี้ในการสร้างระเบียบควรจะกำหนดชื่อให้กับองค์ประกอบแต่ละส่วน เพื่อให้เกิดความเข้าใจตรงกันระหว่างผู้สร้างกับผู้ใช้ เนื่องจากมักมีปัญหว่า ผู้ใช้อ่านรายการนั้นไม่ออก เนื่องจากไม่ทราบว่ารหัสที่ใช้มีความหมายว่าอย่างไรหรือใช้แทนอะไร อีกประการหนึ่งคือ ผู้ใช้ไม่ทราบว่าควรตั้งคำถามอย่างไรเนื่องจากไม่ทราบความต้องการที่แท้จริงของตนเองและผู้ใช้ไม่ทราบว่าจำเป็นต้องใช้สัญลักษณ์ตัวใดบ้างในการค้นคืนข้อมูลจากระบบ โดยเฉพาะเมื่อต้องค้นข้อมูลจากระบบที่ไม่คุ้นเคย ผู้ใช้อาจพลาดข้อมูลที่ควรจะได้ ในการแก้ปัญหา ผู้ใช้มักจะเลือกการค้นแบบสุ่ม (Browse) ซึ่งมีข้อดีคือ ผู้ใช้สามารถเลือกดูได้ว่าในฐานข้อมูลที่กำลังใช้ ใช้คำศัพท์ หรือสัญลักษณ์อะไรบ้างแทนเรื่องที่ต้องการ และบางครั้งอาจพบศัพท์ใหม่ที่ไม่เคยรู้จักมาก่อน

ตัวอย่าง เช่น ผู้ใช้ต้องการค้นหาประวัติของบริษัท IBM หรือพัฒนาการของอุตสาหกรรมคอมพิวเตอร์ หรือ ประวัติของ IBM เกี่ยวกับบทบาทด้านอุตสาหกรรมคอมพิวเตอร์ ผู้ใช้ไม่จำเป็นต้องทราบคำศัพท์ หรือ รหัสที่ใช้แทนเรื่อง "history of the computer industry, especially IBM" ในกรณีนี้ ผู้ใช้จะสุ่มหารายการที่ต้องการจากแฟ้มข้อมูล คำถามที่จะเกิดขึ้น ได้แก่ จะใช้คำใดค้นจึงจะได้ข้อมูลตรงกับความต้องการมากที่สุด และจากรายการที่ค้นได้รายการใดที่ตรงกับความต้องการมากที่สุด สิ่งที่ใช้ต้องการไม่ใช่เพียงแค่คำว่า "History of IBM" แต่ยังต้องการข้อมูลด้านบุคลากร อุปกรณ์ และเหตุการณ์ที่ทำให้เกิดเรื่องราวนั้น ๆ ขึ้นมา หนังสือที่ค้นได้เหล่านี้จะให้ข้อมูลด้านประวัติพัฒนาการ ผู้ใช้จะได้รายการบรรณานุกรม พร้อมกับ คำค้น (Descriptor) ซึ่งคำค้นนี้จะป็นสิ่งยืนยันให้ผู้ใช้ทราบว่าเอกสารที่ค้นได้เป็นเรื่องที่ต้องการ ซึ่งผู้ทำดัชนี (Indexer) จะเป็นผู้กำหนดคำค้น

ผู้ใช้ระบบการค้นคืนสารสนเทศ

อาจจำแนกผู้ใช้ออกเป็น 3 กลุ่ม ดังนี้

- 1) Information Specialists หรือ Reference Librarian
- 2) Professional persons หรือ End users ที่มีความต้องการสารสนเทศ และเป็นผู้ที่มีความรู้ในเรื่องนั้น ๆ อย่างมาก เช่น แพทย์ หนายความ นักเคมี และอาจารย์ เป็นต้น
- 3) Novices เป็นกลุ่มผู้ใช้ขาดความรู้เกี่ยวกับเรื่องที่ต้องการทราบ ได้แก่ นักศึกษา ผู้ใช้ทั่ว ๆ ไป

1. Information Specialists

ได้แก่ผู้เชี่ยวชาญด้านสารสนเทศ หรือบรรณารักษะงานบริการตอบคำถามและช่วยการค้นคว้าเป็นผู้ช่วยเหลือผู้ใช้ในการค้นคืนสารสนเทศ ซึ่งอาจมีชื่อเรียกแตกต่างกันไป นอกจากนี้ยังมีบุคคลอื่น ๆ ที่ทำงานลักษณะนี้แต่ไม่ถือว่าตนเองเป็นผู้ให้บริการสารสนเทศ เช่น สำนักงานส่งเสริมการเกษตร ผู้ที่ทำหน้าที่นี้ในเชิงธุรกิจจะเรียกตัวเองว่า Information Broker ซึ่งกลุ่มนี้จะไม่ได้สร้างสารสนเทศขึ้นมาเอง แต่ทำหน้าที่รวบรวมและแปลงข้อมูลให้กับลูกค้า

อาจกล่าวได้ว่า Information specialists คือผู้ที่ทำหน้าที่เป็นตัวกลาง (Intermediary) ระหว่างผู้ใช้กับระบบการค้นคืนสารสนเทศ ซึ่งจะต้องเป็นผู้ที่มีความรู้ ดังนี้

- 1) ต้องทราบประเภทของข้อสนเทศ ทั้งที่เป็นสหสาขาวิชาและวิชาเฉพาะ เช่น กฎหมาย การแพทย์ การธนาคารหรือต้องทราบว่าในสถาบันของตนมีทรัพยากรสารสนเทศอะไรบ้าง รวมทั้งทราบทรัพยากรฯ ของสถาบันอื่นด้วย
- 2) ทราบกลวิธีในการได้มาซึ่งสารสนเทศ ทั้งวิธีการเข้าถึงฐานข้อมูล การโอนย้ายข้อมูลจากซีดี-รอม มาไว้ในเครื่องคอมพิวเตอร์ การเข้าถึงฐานข้อมูลภายนอก
- 3) ทราบศัพท์เฉพาะที่จะใช้ค้นข้อมูลเฉพาะด้าน เช่น ด้านการแพทย์ กฎหมาย เคมี เป็นต้น
- 4) กลวิธีในการสัมภาษณ์ผู้ใช้ เพื่อทราบความต้องการสารสนเทศที่แท้จริง

2. End Users

ในหลายปีที่ผ่านมา การค้นคืนสารสนเทศจะถูกมองในลักษณะการจัดบริการให้กับผู้ใช้ผ่านทางผู้ให้บริการ เท่านั้น แต่หลังจากที่มีการใช้ PC อย่างแพร่หลาย ส่งผลให้ Professional end users เริ่มที่จะเข้ามามีบทบาทสำคัญในตลาดค้าสารสนเทศ ด้วยเหตุผล หลัก 2 ประการ คือ 1) PC ช่วยให้ผู้ใช้สามารถใช้บริการค้นคืนสารสนเทศออนไลน์ (Online retrieval) ได้โดยตรงไม่ต้องผ่านตัวกลาง 2) ตลาดห้องสมุดเริ่มอิ่มตัว (ในช่วงต้นปี ค.ศ.1980 ห้องสมุดส่วนใหญ่ยังคงเป็นกลุ่มเป้าหมายของบริการค้นคืนสารสนเทศออนไลน์ เนื่องจากยังมีห้องสมุดอีกหลายแห่งที่ไม่ได้เป็นลูกค้าของบริการสารสนเทศออนไลน์) ดังนั้น เพื่อเพิ่มจำนวนผู้ใช้กลุ่มผู้ให้บริการค้นคืนสารสนเทศออนไลน์จึงต้องหาลูกค้ากลุ่มใหม่ ซึ่งกลุ่มเป้าหมายก็คือ End users ที่สามารถเข้าใช้บริการผ่านทาง PC จากที่บ้าน หรือที่ทำงาน ดังนั้น ตลาดบริการค้นคืนสารสนเทศออนไลน์จึงขยายตัวออกไปอย่างรวดเร็ว

ลูกค้ากลุ่มใหญ่ ได้แก่ กลุ่มที่ต้องใช้สารสนเทศเป็นส่วนสำคัญในการดำเนินงาน เช่น แพทย์ นักกฎหมาย นักเคมี เป็นต้น นอกจากนี้ ยังมีผู้ใช้กลุ่มอื่น ๆ อีก เช่น นักวิจัยด้านวิทยาศาสตร์ หรือนักเทคโนโลยีด้านต่าง ๆ กลุ่มนี้ต้องการทราบความก้าวหน้าในสาขาวิชาของตน นักวิจัยทางธุรกิจต้องการทราบข้อมูลเกี่ยวกับการเปลี่ยนแปลงของตลาดและข้อมูลของบริษัทคู่แข่ง นักข่าวมีอิมโหมต้องการทราบข้อมูลเบื้องต้นของบุคคลสำคัญในแวดวงต่าง ๆ เหตุการณ์สำคัญร่วมสมัย หรือข่าวเด่นในรอบปี เป็นต้น

ผู้ใช้แต่ละกลุ่มจะอาศัยทักษะเฉพาะสาขาวิชาในการแปลงข้อมูลไปใช้ จึงทักษะเฉพาะสาขาวิชาคือสิ่งที่ผู้ให้บริการไม่ทราบ ผู้ให้บริการจะรับผิดชอบเพียงการช่วยค้น และช่วยประเมินสารสนเทศที่ค้นได้ หรือให้ความช่วยเหลือผู้ใช้ระหว่างทำการค้นคืน เพื่อให้สิ่งที่ไม่เกี่ยวข้องกับเรื่องที่ต้องการออกมาน้อยที่สุดและช่วยให้ผู้ใช้ได้ข้อมูลในเรื่องที่ต้องการทั้งหมด อย่างไรก็ตาม End users ส่วนใหญ่ยังคงขาดทักษะและไม่ทราบกลวิธีในการค้นคืนสารสนเทศ ซึ่งในส่วนนี้เป็นหน้าที่ของผู้ให้บริการที่จะต้องให้ความช่วยเหลือ

นอกจากขาดทักษะและไม่ทราบกลวิธีในการค้นคืนแล้ว ผู้ใช้บางกลุ่มยังขาดทักษะในการใช้คอมพิวเตอร์ด้วย ทำให้การค้นคืนข้อมูลด้วยตนเองทำได้ยากยิ่งขึ้น แต่สำหรับกลุ่มที่มีทักษะการใช้คอมพิวเตอร์ในระดับดีจะสามารถเรียนรู้กลวิธีการค้นคืนข้อมูลได้อย่างรวดเร็ว ทั้งการค้นคืนข้อมูลจากฐานข้อมูลภายในและภายนอก กลวิธีการค้นคืนเป็นสิ่งจำเป็นเนื่องจากภาษาที่ใช้ในการพูดและภาษาที่ใช้ในการค้นคืนข้อมูลนั้นมีรูปแบบที่ไม่เหมือนกัน และภาษาที่ใช้ในการค้นคืนสารสนเทศเป็นภาษาที่ต้องถ่ายทอดให้คอมพิวเตอร์เข้าใจ ซึ่งปัจจุบันคอมพิวเตอร์ส่วนใหญ่ยังมีขีดความสามารถด้านการเข้าใจภาษาพูดของมนุษย์ในระดับต่ำ

3. Novices

ผู้ใช้ในกลุ่มนี้ส่วนใหญ่ได้แก่ นักเรียนระดับมัธยม นักศึกษาระดับปริญญาตรี ซึ่งเป็นกลุ่มที่ยังขาดความรู้และความเชี่ยวชาญในเรื่องที่ต้องการค้น ไม่ทราบกลวิธีการทำงานของระบบ และยังขาดทักษะในการค้นคืนสารสนเทศ

ปัญหาในการออกแบบและการใช้ระบบค้นคืนสารสนเทศ

ปัญหาในการออกแบบและการใช้ระบบค้นคืนสารสนเทศจะเปลี่ยนแปลงอยู่ตลอดเวลาตามพัฒนาการของเทคโนโลยี ปัญหาที่เกิดขึ้นมักจำแนกไปตามพฤติกรรมของผู้ใช้และทักษะด้านต่าง ๆ ที่ผู้ใช้ต้องมีในการใช้งานระบบ ปัญหาเหล่านี้ ได้แก่

1. การออกแบบ (Design)

1.1 Representative of Information in the Database

ถ้ามีการกำหนดสัญลักษณ์แทนสารสนเทศ ผู้ใช้ทุกคนจะต้องทราบและเข้าใจสัญลักษณ์ที่เลือกใช้นั้น ๆ ข้อควรคำนึงคือ จะออกแบบระบบอย่างไรจึงจะประหยัดเวลาผู้ใช้มากที่สุดและเสนอทางเลือกให้ผู้ใช้สามารถทำความเข้าใจกับสัญลักษณ์ต่าง ๆ ที่กำหนดขึ้นได้ในระบบได้

1.2 Organization of Information within a Computer

ปัจจุบันปัญหาในการจัดการข้อมูลในคอมพิวเตอร์เป็นปัญหาที่กำลังได้รับความสนใจจากผู้สร้างระบบ นักออกแบบระบบจะต้องรับทราบปัญหานี้เพื่อป้องกันไม่ให้เกิดความล่าช้าในกระบวนการของการประเมินผล ซึ่งต้องพัฒนาทั้งในส่วนของ Organization และ Search Technique

1.3 User Interfaces

หากระบบการค้นคืนสารสนเทศได้รับการออกแบบมาอย่างดีโดยเฉพาะในส่วนของระบบการเชื่อมประสานกับผู้ใช้ (User Interface) ผู้ใช้ระบบจะไม่จำเป็นต้องมาเสียเวลาเพื่อเรียนรู้การใช้งานระบบ ผู้ใช้จะสามารถใช้งานระบบได้ด้วยตนเอง ใช้เวลาเพียงน้อยนิดผู้ใช้งานจะสามารถใช้งานระบบได้อย่างดีและเข้าใจขั้นตอนในการทำงานของระบบอย่างชัดเจน

1.4 Query-Management Programs

ทำหน้าที่แปลงคำถามของผู้ใช้ให้อยู่ในรูปของ Artificial Languages คือ ทำหน้าที่แปลงความต้องการของผู้ใช้ให้ระบบทราบว่าผู้ใช้ต้องการสารสนเทศอย่างไรได้ ซึ่งระบบจะตั้งคำถามกลับมายังผู้ใช้อีกทีเพื่อยืนยันความต้องการ ก่อนที่จะทำการประมวลผล

1.5 Data Management programs

เป็นหนึ่งในกระบวนการทางเทคนิคของระบบการค้นคืนสารสนเทศ ซึ่งการออกแบบระบบจะไม่ให้ผู้ใช้เห็นกระบวนการนี้ แม้ว่ากระบวนการนี้จะเป็นกระบวนการที่สำคัญกระบวนการหนึ่งก็ตาม เนื่องจากการแสดงกระบวนการนี้ให้ผู้ใช้เห็นระบบจะต้องเสียเวลามากขึ้นในการแสดงผลและผู้ใช้อาจต้องเสียค่าใช้จ่ายมากขึ้นตามเวลาที่เพิ่มขึ้น ดังนั้น หากกระบวนการนี้ทำงานได้เร็วจะช่วยลดค่าใช้จ่ายของผู้ใช้ได้มากขึ้น

1.6 Presentation (Display) of Information to User

การแสดงผลนั้นมีปัญหาทั้งในด้านเทคนิคและพฤติกรรม เช่น จะแสดงข้อมูลอะไรบ้าง จะใช้สัญลักษณ์อะไร จะจัดเรียงหน้าจอย่างไร และจะแสดงความสัมพันธ์กับหน้าจออื่น ๆ อย่างไร

2. พฤติกรรมของผู้ใช้และปัญหาในการฝึกอบรม

พฤติกรรมของผู้ใช้เป็นปัญหาที่ค่อนข้างแก้ไขได้ยากกว่าการออกแบบ

2.1 Understanding Information Representation and Organization

จะทำอย่างไรให้ผู้ใช้งานเข้าใจสัญลักษณ์ที่ใช้แทนเนื้อหาของสารสนเทศและโครงสร้างการจัดระบบสารสนเทศได้ง่ายที่สุด

2.2 Understanding How to Communicate Information Needs

ดังได้กล่าวไว้แล้วว่า ความสำเร็จของกระบวนการค้นคืนสารสนเทศทั้งหมดขึ้นอยู่กับความสามารถในการอธิบายความต้องการสารสนเทศของผู้ใช้ว่าจะทำได้ดีเพียงไร ซึ่งในส่วนนี้ระบบสารสนเทศอาจช่วยได้บ้างถ้าระบบได้ถูกออกแบบมาให้สามารถดึงข้อมูลออกมาจากผู้ใช้นั้นมากที่สุด (อาจโดยการตั้งคำถามหรือให้กรอกแบบฟอร์ม) มากกว่าที่จะรอให้ผู้ใช้งานให้ข้อมูลออกมาเอง

2.3 Understanding How to Formulate Query and Evaluate Results

เราจะต้องพิจารณาว่า การเข้าใจและการอธิบายความต้องการนั้นไม่เหมือนกับการแปลงความต้องการให้อยู่ในรูปแบบที่คอมพิวเตอร์เข้าใจ การทำให้คอมพิวเตอร์เข้าใจเป็นสิ่งที่ซับซ้อนกว่าและต้องใช้เวลามาก แต่หากผู้ใช้ไม่ให้ความสำคัญหรือไม่ต้องการเสียเวลาที่จะเรียนรู้วิธีการในการสื่อสารความต้องการให้คอมพิวเตอร์ทราบ การแก้ปัญหาในส่วนนี้จะยังทำได้ยาก

2.4 Understanding How to Evaluate Information and Systems

ถ้าผู้ใช้ทราบความต้องการที่แท้จริงของตัวเองได้มากเท่าไร ก็ยังประเมินผลการค้นคืนและระบบได้ดีเท่านั้น แม้ว่าพฤติกรรมในส่วนนี้จะยังคงเป็นปัญหาที่แก้ไขไม่ได้ก็ตาม แต่ผู้ให้บริการสารสนเทศจะต้องสอนวิธีการประเมินผลการค้นคืน และวิธีการที่จะสร้าง query ใหม่เพื่อให้ได้ผลการค้นที่ดียิ่งแก่ผู้ใช้ เพื่อให้ผู้ใช้สามารถประเมินได้ว่าระบบที่ดีควรจะทำอะไรได้บ้าง เช่นต้องมีระบบช่วยในการตัดสินใจ เป็นต้น

ประเด็นการศึกษาด้านการจัดเก็บและการค้นคืนสารสนเทศ

วัตถุประสงค์หลักในการศึกษาด้านการจัดเก็บและการค้นคืนสารสนเทศในช่วง 20 ปีที่ผ่านมาได้แก่ การศึกษาเพื่อหาวิธีที่จะช่วยให้ผู้ใช้สามารถเข้าถึงเอกสารที่จัดเก็บไว้ให้ได้มากที่สุด ปัจจุบันการศึกษาในด้านนี้ยังครอบคลุมถึง การสร้างแบบจำลองของระบบ (Modeling) การจัดหมวดหมู่และการจัดโครงสร้างการจัดเก็บเอกสาร (Document classification and categorization) โครงสร้างของระบบ (System architecture) ระบบการเชื่อมประสานกับผู้ใช้ (User interfaces) การแสดงข้อมูล (Data visualization) ภาษา (Language) ฯลฯ ที่ผ่านมากลุ่มผู้ที่ให้ความสนใจศึกษาเรื่องนี้จะจำกัดอยู่เฉพาะบรรณารักษ์ และนักสารสนเทศเท่านั้น แต่ในช่วงหลายปีที่ผ่านมา เมื่อคอมพิวเตอร์ส่วนบุคคลได้รับการพัฒนาให้สามารถทำงานกับข้อมูลลักษณะสื่อผสม (Multimedia) และไฮเปอร์เท็กซ์ (Hypertext) ซึ่งเป็นเครื่องมือสำคัญในการพัฒนาระบบการจัดเก็บและค้นคืนสารสนเทศ จากต้นปี ค.ศ. 1990 ปัจจัยสำคัญที่ทำให้เกิดการเปลี่ยนแปลงในประเด็นการจัดเก็บและการค้นคืนสารสนเทศคือ World Wide Web นั่นเอง

เมื่อ Web กลายเป็นแหล่งรวบรวมและเผยแพร่ความรู้ของมนุษยชาติ ผู้คนจากทั่วทุกมุมโลกสามารถแลกเปลี่ยนวัฒนธรรม แนวคิด และความรู้ต่าง ๆ ได้อย่างสะดวกรวดเร็ว นอกจากนี้ระบบการเชื่อมประสาน

ที่เป็นมาตรฐานเดียวกันยังทำให้คอมพิวเตอร์ต่างระบบสามารถสื่อสารกันได้ และที่สำคัญคือการทำทุกคนสามารถสร้าง Web site เพื่อเผยแพร่ข้อมูลข่าวสารไปยังคนอื่น ๆ ได้โดยไม่ต้องเสียค่าใช้จ่าย ทำให้ Web กลายเป็นสื่อสิ่งพิมพ์ชนิดใหม่ที่ทุกคนสามารถเข้าถึงได้โดยอาศัยคอมพิวเตอร์ซึ่งเป็นเครื่องมือที่ใช้ในการทำงานประจำวัน รวมทั้งกิจกรรมอื่น ๆ อาทิ การซื้อสินค้าและบริการผ่านทางอินเทอร์เน็ต ซึ่งเป็นธุรกิจที่ทำรายได้มหาศาลในปัจจุบัน

อย่างไรก็ตาม ยังคงมีปัญหามากมายประการในการค้นคืนเอกสารจาก Web ปกติครั้งที่ผู้ใช้พบว่าเอกสารที่ค้นได้มีมากเกินไป และส่วนใหญ่เป็นเอกสารที่ไม่ตรงกับความต้องการ และปัญหาอีกประการหนึ่งคือ การหลงทางบน Web ทั้งนี้เนื่องจากลักษณะการค้นข้อมูลบน Web จะมีการเชื่อมโยงข้อมูลในลักษณะ Hypertext ซึ่งผู้ใช้สามารถเชื่อมโยงเข้าไปดูข้อมูลจากเอกสารอื่น ๆ ได้ตลอดเวลา หากไม่มีการออกแบบระบบการนำทางที่ดี ผู้ใช้อาจเกิดความสับสนในการท่อง Web เนื่องจากไม่ทราบว่าขณะนี้ได้เข้ามาอยู่ที่ใดบน Web และไม่ทราบว่าเข้ามาอยู่ได้อย่างไรและจะต้องทำอย่างไรต่อไป ซึ่งปัญหานี้นำมาซึ่งการศึกษาในประเด็นการออกแบบโครงสร้างของข้อมูลบน Web และการกำหนดคัตวแทนของเอกสาร ทั้งนี้เพื่อหาแนวทางในการแก้ไขปัญหาดังกล่าว

เรียบเรียงจาก

Meadow, Charles T. 1992. "Introduction." in *Text Information Retrieval Systems*, p. 1-19.

San Diego : Academicpress.

Baeza-Yates, Ricardo ; Ribeiro-Neto, Bertier. 1999. "Introduction." In *Modern Information Retrieval*, p. 1-18. New York : Addison-Wesley.

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 2

ฐานข้อมูลบรรณานุกรม (Bibliographic Database)

ความหมายของฐานข้อมูล

ฐานข้อมูล (Database) เป็นแหล่งที่นำข้อมูลที่มีความหมายเกี่ยวข้องสัมพันธ์กัน มาจัดเก็บรวมไว้ด้วยกัน โดยมีวัตถุประสงค์เพื่อที่จะสามารถนำข้อมูลต่าง ๆ เหล่านี้มาใช้ร่วมกันได้ และข้อมูลเหล่านี้สามารถสืบค้น จัดเรียง ปรับปรุงให้ทันสมัย หรือเปลี่ยนแปลงได้โดยใช้โปรแกรมที่เรียกว่า ระบบจัดการฐานข้อมูล (Database Management System : DBMs) (พินิจวิภา, 2538: 3) โดยทั่วไปฐานข้อมูลมักจะถูกถือเป็นแหล่งกลางในการให้ข้อมูลเพื่อการประมวลผลต่าง ๆ วัตถุประสงค์ของการรวบรวมข้อมูลทีพิจารณาแล้วว่าเป็นที่น่าสนใจและเป็นประโยชน์เอาไว้ในฐานข้อมูลก็เพื่อลดภาระการเก็บข้อมูลที่ซ้ำซ้อนกัน และเป็นการใช้ข้อมูลร่วมกัน

นงลักษณ์ ไม่น้อยกิจ ได้กล่าวถึงความหมายของฐานข้อมูลตามรูปศัพท์ภาษาคือ Database ไว้ว่า Database มาจากคำ 2 คำ คือ Data และ Base

Data เป็นทรัพยากรที่ประกอบด้วยข้อเท็จจริง (facts) ในปริมาณมาก "data" เหล่านี้เป็นข้อเท็จจริงที่มนุษย์ต้องการใช้ และ data นี้ไม่ได้อยู่คงที่ data มีการเพิ่มขึ้น เปลี่ยนแปลง และปรับเปลี่ยนตลอดเวลา data ไม่ใช่ information แต่ data เป็นวัตถุดิบ และ data จะเป็น information ได้ เมื่อมีการนำมาจัดดำเนินการเพื่อแก้ไขปัญหา ส่วน Base คือ สิ่งที่เป็นฐานให้สร้างขึ้นไปได้ ฉะนั้น Database ก็คือ ทรัพยากรสารสนเทศซึ่งออกแบบมาเพื่อให้เราเริ่มต้นในการได้มาซึ่งความรู้ Database จะบอกให้ทราบว่า ใครเคยทำอะไรไว้แล้วบ้างก่อนหน้านี้ หรือได้มีใครเขียนอะไรไว้แล้วบ้าง ยิ่ง Base ใหญ่ขึ้นก็จะเป็นฐานให้เราสามารถสร้างอะไรได้มากขึ้น (นงลักษณ์, 2535: 82)

อาจสรุปได้ว่า ฐานข้อมูล คือ คอลเลกชันของข้อเท็จจริงในรูปแบบที่คอมพิวเตอร์อ่านได้ ประกอบด้วยแฟ้มข้อมูลหลายแฟ้มที่มีความเกี่ยวข้องสัมพันธ์กัน และสามารถนำข้อมูลต่าง ๆ ออกมาใช้ร่วมกันได้ ทั้งนี้เพื่อลดความซ้ำซ้อนในการจัดเก็บข้อมูล

ประเภทของฐานข้อมูล

ฐานข้อมูลที่สร้างขึ้นในปัจจุบันมีความหลากหลายทั้งในด้านเนื้อหาและการใช้งาน ดังนั้น การแบ่งประเภทของฐานข้อมูลจึงแบ่งได้หลายลักษณะ ในที่นี้จะแบ่งประเภทของฐานข้อมูลตามจุดมุ่งหมายในการจัดทำ ดังนี้

1. ฐานข้อมูลต้นแหล่ง (Source Database) จะให้ข้อมูลหรือเนื้อหาฉบับเต็มแก่ผู้ใช้ ข้อมูลที่ให้อาจเป็น ตัวเลข ข้อความ หรือทั้งตัวเลขและข้อความ ฐานข้อมูลประเภทนี้ ได้แก่
 - 1.1 ฐานข้อมูลข้อเท็จจริง (Factual Database) จะมีข้อมูลที่ไม่ใช่ตัวเลข เช่น นามานุกรมของหน่วยงาน โครงการวิจัย ซึ่งเป็นข้อมูลที่มีความจำเป็น และต้องการการปรับปรุงให้ทันสมัยอยู่เสมอ จัดเป็นแหล่งอ้างอิงปฐมภูมิ เพราะข้อมูลที่มีใช้ต้องการจะปรากฏในฐานข้อมูลโดยตรง

- 1.2 **ฐานข้อมูลเนื้อหาเต็ม (Full-text Database)** จะให้เนื้อหาทั้งหมดของเอกสาร เช่น บทความวารสารฉบับเต็ม รายงานการวิจัยฉบับเต็ม เป็นต้น จัดเป็นแหล่งอ้างอิงปฐมภูมิ เพราะผู้ใช้จะได้รับสารสนเทศที่ต้องการจากฐานข้อมูลโดยตรงเช่นกัน
- 1.3 **ฐานข้อมูลตัวเลข (Numeric Database)** ให้ข้อมูลตัวเลข สถิติต่าง ๆ เช่น สถิติการค้า การเงิน ราคาหุ้น ในบางกรณีอาจมีการให้บริการคำนวณให้กับผู้ใช้ฐานข้อมูลด้วย จัดเป็นแหล่งอ้างอิงปฐมภูมิ เพราะผู้ใช้จะได้รับสารสนเทศที่ต้องการจากฐานข้อมูลโดยตรง
2. **ฐานข้อมูลอ้างอิง (Reference Database)** เป็นฐานข้อมูลที่ไม่มีเนื้อหาฉบับเต็ม (full-text) แต่จะชี้แนะไปยังแหล่งที่สามารถให้ข้อมูลที่ต้องการได้ ฐานข้อมูลประเภทนี้ได้แก่ **ฐานข้อมูลบรรณานุกรม (Bibliographic Database)** ให้รายละเอียดเกี่ยวกับรายการบรรณานุกรมของเอกสาร คำกรณีย์ และสาระสังเขปของเอกสารดังกล่าว เพื่อให้ผู้ใช้ทราบข้อมูลคร่าว ๆ ของเอกสาร เพื่อประเมินความเกี่ยวข้องของเอกสาร และสามารถติดตามไปยังเอกสารฉบับเต็มได้โดยอาศัยรายการบรรณานุกรมของเอกสาร ซึ่งบรรณานุกรมเปรียบได้กับตัวชี้แนะไปยังแหล่งข้อมูลที่ต้องการ ฐานข้อมูลบรรณานุกรมถือเป็นแหล่งอ้างอิงทุติยภูมิที่จะนำผู้ใช้ไปสู่เอกสารต้นฉบับ

แม้ฐานข้อมูลจะมีอยู่หลายประเภทแต่ฐานข้อมูลที่เป็นที่นิยมจัดทำกันอย่างแพร่หลายได้แก่ ฐานข้อมูลบรรณานุกรม ซึ่งมีการเพิ่มขึ้นอย่างรวดเร็วตามอุตสาหกรรมฐานข้อมูล เนื่องจากบรรณานุกรมเป็นจุดเริ่มต้นของการค้นคว้า ปัจจุบันฐานข้อมูลบรรณานุกรมยังทำหน้าที่เป็น Gateway ที่จะนำผู้ใช้ไปสู่ข้อมูลฉบับเต็มของวารสาร นิตยสาร บรรณานุกรมของฐาน ข้อมูลอื่น รวมทั้งบริการอื่น ๆ ที่เกี่ยวข้อง เช่น การเชื่อมโยงไปยังฐานข้อมูลอื่นในลักษณะออนไลน์ บริการจัดส่งเอกสาร เป็นต้น

การสร้างฐานข้อมูลบรรณานุกรม

ฐานข้อมูลบรรณานุกรมในระยะแรก มักเป็นผลผลิตที่มีพื้นฐานจากเครื่องมืออ้างอิงที่เป็นสื่อสิ่งพิมพ์ เช่น วรรณคดีและสาระสังเขป ฐานข้อมูลบรรณานุกรมจึงเป็นที่รู้จักกันดีของนักสารสนเทศ นักวิจัย และนักวิชาการในสาขาต่าง ๆ โดยทั่วไปการจัดทำฐานข้อมูลบรรณานุกรมมีวัตถุประสงค์ ที่สำคัญ คือ

- ควบคุมสารสนเทศ โดยการจัดทำรายชื่อ หรือรายการบรรณานุกรมของเอกสารที่ได้รับการผลิตขึ้นมา เพื่อให้ผู้ใช้ทราบว่ามีการผลิตสารสนเทศรายการใดที่ได้รับการผลิตออกมาบ้าง
- ช่วยให้ผู้ใช้สามารถค้นหารายการเอกสารที่ต้องการได้อย่างรวดเร็ว โดยการจัดทำวรรณคดีสาระสังเขป และฟังก์ชัน (function) การช่วยค้นต่าง ๆ
- ขยายการใช้สารสนเทศออกไปในวงกว้าง และตอบสนองความต้องการของผู้ใช้เฉพาะกลุ่ม เนื่องจากผู้ใช้ทราบว่ามีสารสนเทศใดในเรื่องที่ตนสนใจ และทราบว่าจะสามารถติดตามได้จากที่ใด

ผู้ผลิตฐานข้อมูล

แต่เดิมการผลิตฐานข้อมูลมักเป็นกิจกรรมของสถาบันบริการสารสนเทศ แต่ปัจจุบันฐานข้อมูลส่วนใหญ่มีผลิตขึ้นโดยบริษัทเอกชนที่หวังผลกำไร และนำมาจำหน่ายให้กับสถาบันบริการสารสนเทศ ขณะที่สถาบันบริการสารสนเทศจะเน้นผลิตฐานข้อมูลทรัพยากรของสถาบันเป็นหลัก อาจจำแนกกลุ่มผู้ผลิตฐานข้อมูลได้เป็น 3 กลุ่มคือ กลุ่มที่ทำเพื่อหวังผลกำไร กลุ่มที่ทำไม่จริงจังและไม่หวังผลกำไร และกลุ่มที่ทำตามหน้าที่

1. กลุ่มที่ทำเพื่อหวังผลกำไร ได้แก่บริษัทเอกชนต่าง ๆ เช่น SilverPlatter Information ผลิตฐานข้อมูลทั่วไปกว่า 280 ฐาน Questel Orbit ผลิตฐานข้อมูลด้านวิทยาศาสตร์ เคมี การค้า และ ธุรกิจ Knig-Rider Information ให้บริการฐานข้อมูลของ Dialog, DataStar, KR Sciencebase เป็นต้น

2. กลุ่มที่ไม่หวังผลกำไร เช่น สมาคมวิชาชีพต่าง ๆ เช่น American Chemical Society ผลิตฐานข้อมูล Chemical Abstract (CAS)

3. กลุ่มที่ผลิตตามหน้าที่ ได้แก่ ภาครัฐและเอกชนที่ต้องผลิตเพราะมีกฎระเบียบ ข้อบังคับ หรือข้อตกลงต่าง ๆ ที่ต้องทำ เช่น US National Library of Medicine ผลิตฐานข้อมูล MEDLINE

ขั้นตอนการสร้างฐานข้อมูล

ในการสร้างฐานข้อมูลต้องมีการวางแผนที่ดี เพื่อให้การดำเนินงานเป็นไปโดยราบรื่น โดยมีขั้นตอนการทำงานดังต่อไปนี้

1. การศึกษาความต้องการหรือความเป็นไปได้
2. การจัดหาและรวบรวมข้อมูล
3. การออกแบบฐานข้อมูล
4. การเลือกซอฟต์แวร์
5. การจัดทำคู่มือ
6. การประเมินผล
7. การเผยแพร่และให้บริการ

1. การศึกษาความต้องการหรือความเป็นไปได้ มีวัตถุประสงค์เพื่อ

1) ระบุขอบเขตของฐานข้อมูลทั้งในด้านเนื้อหา (สาขาวิชา) ประเภทของเอกสาร ระยะเวลาของเอกสาร (ปีพิมพ์) และภาษาของเอกสารที่จะรวบรวมไว้ในฐานข้อมูล เพื่อเป็นกรอบการทำงานที่ชัดเจนในการกำหนดกลุ่มเป้าหมายและการรวบรวมข้อมูล

2) ระบุความต้องการของผู้ใช้ โดยการกำหนดกลุ่มผู้ใช้ที่เป็นเป้าหมายหลัก และกลุ่มที่คาดว่าจะเข้ามาใช้ ทั้งนี้เพื่อกำหนดระดับความลึก / กว้างของข้อมูลให้เหมาะสมกับผู้ใช้ รวมทั้งระดับความยากง่ายของคำตรรกะที่จะใช้เพื่อการสืบค้นสารสนเทศจากฐานข้อมูล และศึกษาถึงความต้องการอื่น ๆ ของผู้ใช้ ได้แก่ ลักษณะข้อมูลที่ใช้ต้องการเป็นบทความ รายงานการวิจัย รายงานการประชุม หนังสือ วิทยานิพนธ์ หรือ บทความวิเคราะห์ เป็นต้น รายละเอียดของข้อมูลที่ต้องการ อาทิ รายละเอียดทางบรรณานุกรม ภาษา ประเภทของทรัพยากร สาระสังเขป คำตรรกะ เป็นต้น รูปแบบการนำเสนอ การแสดงรายการบรรณานุกรมแบบย่อ แบบสมบูรณ์ การจัดลำดับข้อมูลที่น่าเสนอ การจัดเรียงผลการค้น ปริมาณข้อมูลที่ต้องการ ความทันสมัยของข้อมูล รวมทั้งศึกษาถึงความสามารถและพฤติกรรมการค้นคืนสารสนเทศของผู้ใช้ เพื่อออกแบบหน้าจอ คำสั่ง และข้อความต่าง ๆ ให้ง่ายต่อการใช้งาน

นอกจาก 2 ประเด็นข้างต้นแล้ว การศึกษาความเป็นไปได้ต้องพิจารณาถึง ความพร้อมด้าน ฮาร์ดแวร์และซอฟต์แวร์ที่จะใช้ในการพัฒนา รวมทั้งค่าใช้จ่ายที่จะเกิดขึ้นในการจัดทำ ผลของการศึกษาความเป็นไปได้อาจจัดทำในรูปของเอกสารเพื่อใช้ในการตัดสินใจต่อไป

ผู้ที่มีส่วนร่วมในการจัดทำฐานข้อมูลในหน่วยงาน ได้แก่ ผู้บริหาร ทำหน้าที่ในการบริหารสั่งการ ประธานงานในการดำเนินงาน ผู้แทนจากฝ่ายประมวลผลข้อมูลและฝ่ายจัดการสารสนเทศ เป็นผู้ออกแบบโครงสร้างของฐานข้อมูลและจัดทำโปรแกรมเพื่อการค้นคืน ผู้ใช้จากฝ่ายต่าง ๆ ในหน่วยงานอธิบายถึงความต้องการใช้งาน ผู้ใช้ที่เป็นตัวกลาง ทำหน้าที่วิเคราะห์ความต้องการของผู้ใช้ และออกแบบระบบการค้นคืนสารสนเทศและฟังก์ชันให้ความช่วยเหลือต่าง ๆ แก่ผู้ใช้ในการค้นคืนสารสนเทศ

2. การจัดหาและรวบรวมสารสนเทศ เป็นการดำเนินการตามผลที่ได้จากการศึกษาความเป็นไปได้ โดยจัดทำเป็นลายลักษณ์อักษรว่า ต้องการเอกสารประเภทใด จากแหล่งข้อมูลใด สิ่งที่ต้องพิจารณาคือ คุณภาพและความน่าเชื่อถือของแหล่งข้อมูล จึงต้องมีการกำหนดเกณฑ์ในการคัดเลือกเอกสาร เช่น บทความควรเป็นบทความที่ได้รับการตีพิมพ์ในวารสารทางวิชาการหรือวิชาชีพที่เป็นที่รู้จักแพร่หลาย หรือรายงานการวิจัยควรจัดทำโดยสถาบันหรือหน่วยงานที่เป็นที่ยอมรับในวงวิชาการนั้น ๆ ซึ่งในการคัดเลือกเอกสารต้องอาศัย วิจารณ์งานของผู้ทำการคัดเลือกเป็นสำคัญ

3. การออกแบบฐานข้อมูล กระบวนการที่ต้องเน้นในการออกแบบฐานข้อมูลได้แก่ การกำหนดเนื้อหาที่จะมีในฐานข้อมูล การบันทึกข้อมูลในฐานข้อมูล และรูปแบบการแสดงผลตามที่ใช้ต้องการ

ส่วนประกอบของฐานข้อมูลบรรณานุกรม ในฐานข้อมูลบรรณานุกรมอย่างน้อยจะต้องประกอบด้วย ส่วนประกอบ 2 ส่วน คือ 1) ข้อมูลทางบรรณานุกรม (Bibliographic data หรือ Metadata) เช่น ชื่อผู้แต่ง ชื่อเรื่อง แหล่งผลิตเอกสาร 2) คำสำคัญ (Key words) หรือบางครั้งใช้คำว่า ศัพท์บังคับ (Descriptors) เนื่องจากบางครั้งผู้ใช้ไม่ทราบชื่อเรื่อง หรือชื่อผู้แต่ง เพียงแต่ทราบประเด็นที่ต้องการจะค้นหาข้อมูลเท่านั้น การจัดทำคำสำคัญจะช่วยเหลือผู้ใช้ให้สามารถค้นหาเอกสารที่ต้องการได้ง่ายขึ้น

นอกจากองค์ประกอบหลัก 2 ส่วนข้างต้นแล้ว ฐานข้อมูลบรรณานุกรมในยุคหลัง ๆ ยังประกอบไปด้วย ส่วนประกอบที่สำคัญอีกบางส่วน คือ สารสังเขป (Abstract) เพื่อให้ผู้ใช้สามารถทราบเนื้อหาสาระที่สำคัญของเอกสารต้นฉบับ เพื่อประกอบการพิจารณาตัดสินใจในการเลือกใช้เอกสารแต่ละรายการ และในปัจจุบัน ฐานข้อมูลบรรณานุกรมที่ให้บริการในลักษณะออนไลน์ยังได้เพิ่มขีดความสามารถในการเชื่อมโยงไปยังเอกสารฉบับเต็ม และการเชื่อมโยงไปยังฐานข้อมูลและบริการอื่น ๆ ที่เกี่ยวข้องได้ด้วย

3.1 การออกแบบแฟ้มข้อมูลหลัก (Master file)

3.1.1 เนื้อหาที่จะมีในฐานข้อมูล ในการสร้างฐานข้อมูลบรรณานุกรมจะต้องทำการศึกษาถึงเนื้อหาที่ต้องการให้มีในฐานข้อมูล โดยเริ่มจากการพิจารณาขอบเขตของสาขาวิชา ภาษา ประเภทของเอกสาร ระยะเวลาที่ต้องการให้ฐานข้อมูลครอบคลุม ทั้งนี้เพื่อกำหนดว่าในระเบียบข้อมูลแต่ละระเบียบจะต้องประกอบด้วยเขตข้อมูลใดบ้าง องค์ประกอบที่ต้องพิจารณา ได้แก่

- มีเขตข้อมูลอะไรบ้าง (field name)
- ประเภทของข้อมูลที่จะบรรจุในแต่ละเขตข้อมูล (field type) ถ้าจะเป็นพยัญชนะ ตัวเลข ตัวอักษร หรือเป็นทั้งตัวเลขและตัวอักษร หรือเป็นรูปภาพ
- ความยาวของเขตข้อมูล (field length) กำหนดให้เหมาะสมกับความยาวข้อมูลที่ จะกรอก เช่น เขตข้อมูลปีพิมพ์ กำหนดเพียง 4 ตัวอักษร เนื่องจากปีมีตัวเลขเพียง 4 หลัก เขตข้อมูลชื่อผู้แต่งกำหนด 30 ตัวอักษร ตามความยาวของชื่อผู้แต่ง เป็นต้น
- เขตข้อมูลที่มีค่าซ้ำ (repeatable field) มีหรือไม่ เช่น เขตข้อมูลที่เป็นศัพท์ดรรชนี หรือคำสำคัญ จะเป็นเขตข้อมูลแบบมีค่าซ้ำ (คือ ค่าที่กรอกสามารถไปกรอกซ้ำได้ อีกในระเบียบอื่น ที่กำหนดให้ใช้คำสำคัญเหมือนกัน)
- มาตรฐานการลงรายการและคู่มือ ที่ใช้

เขตข้อมูลที่อยู่ในโครงสร้างของระเบียบมีความสำคัญมากในการสืบค้นและการแสดงผลการ สืบค้น ซึ่งจะช่วยให้ผู้ใช้สามารถสืบค้นหาเอกสารที่ต้องการจากเขตข้อมูลต่าง ๆ ได้ เช่น ต้องการสืบค้นเรื่อง การปลูกพืชไร่นาจากวารสาร และต้องตีพิมพ์ในช่วงปี พ.ศ. 2540-2545 เท่านั้น และเป็นเอกสารภาษาไทย ซึ่งหากไม่มีการกำหนดเขตข้อมูลดังกล่าว ผู้ใช้จะไม่สามารถกำหนดขอบเขตการค้นได้ตามต้องการได้ เขตข้อมูลที่สำคัญในโครงสร้างระเบียบของฐานข้อมูลบรรณานุกรม ได้แก่

- ชื่อผู้แต่ง (Author : AU) บางฐานข้อมูลจะให้ที่อยู่ของผู้แต่งไปด้วย (Address : AD) *
- ชื่อเรื่อง (Title : TI), ชื่อเรื่องต้นฉบับที่ไม่ใช่ภาษาอังกฤษ (Original Title : TO)
- แหล่งเอกสาร (Source : SO) ได้แก่ วารสาร (Journal Source : JS) หนังสือ (Monograph Source : MS) สถานที่พิมพ์ (Publication Area : PA) ประเภทของเอกสาร (Document Type : DT) ได้แก่ วารสาร (Journal : J) วิทยานิพนธ์ (Thesis & Dissertation : T)
- ภาษา (Language : LA)
- ปีพิมพ์ (Publication Year : PY)
- หมายเลขเอกสาร หรือ เลขทะเบียนเอกสาร (Accession Number : AN)
- บันทึกข้อความ (Note)
- เขตข้อมูลเพิ่มคุณค่า (Value-added fields) ได้แก่ คำสำคัญ (Descriptors / Keywords) และ สารสังเขป (Abstract)

3.1.2 การบันทึกข้อมูลในฐานข้อมูล ประกอบด้วย การออกแบบแผ่นงาน (Work sheet) เพื่อนำ ข้อมูลเข้า การสร้างระเบียบใหม่ การแก้ไขข้อมูล ในการออกแบบข้อมูลนำเข้า จะศึกษาจาก ผลลัพธ์ที่ผู้ใช้ต้องการให้นำเสนอข้อมูลโดยเฉพาะที่เป็นเอกสาร แล้วจึงมาพิจารณากำหนด ว่าวิธีการและรูปแบบการบันทึกข้อมูลควรเป็นเช่นไร เพื่อช่วยอำนวยความสะดวกในการ นำเข้าข้อมูล โดยทั่วไปการออกแบบเขตข้อมูลนำเข้ามักจะเริ่มจากการกำหนดโครงสร้างของ

* ชื่อของเขตข้อมูลและอักษรย่อที่มักใช้ในฐานข้อมูลบรรณานุกรม

ฐานข้อมูล การกำหนดโครงสร้างของฐานข้อมูลสามารถศึกษาได้จากรูปแบบที่ปรากฏในฐานข้อมูลบรรณานุกรมที่ให้บริการเชิงพาณิชย์ต่าง ๆ สิ่งที่ต้องศึกษาเพิ่มเติมในการออกแบบแผนงานสำหรับนำเข้าข้อมูล คือ มาตรฐานการลงรายการข้อมูลที่เป็นสากล ที่สำคัญ คือ MARC ซึ่งเป็นมาตรฐานในการจัดเก็บข้อมูลเพื่อให้เครื่องคอมพิวเตอร์สามารถอ่านได้ เป็นการสื่อสารข้อมูลบรรณานุกรมระหว่างผู้ใช้และเครื่องคอมพิวเตอร์ การเลือกเขตข้อมูลสำหรับฐานข้อมูลที่จัดสร้างขึ้นสามารถคัดเลือกเฉพาะเขตข้อมูลที่สำคัญ และมีความจำเป็นต้องใช้ในการลงรายการ วิธีการลงรายการในเขตข้อมูลที่ได้คัดเลือกไว้ในแผนงาน ควรจัดทำเป็นคู่มือเพื่อให้มีมาตรฐานในการลงรายการ โดยอาจยึดเกณฑ์การลงรายการของ MARC เป็นหลัก หรือเพิ่มเติมบางรายการให้เหมาะสมกับการปฏิบัติงาน

3.1.3 การออกแบบผลลัพธ์ หรือรูปแบบการแสดงผล เป็นการพิจารณาว่าผู้ใช้งานต้องการการนำเสนอสารสนเทศในรูปแบบใดทางจอภาพ เพื่อสั่งพิมพ์ผลการค้นหาเครื่องพิมพ์ ในฐานข้อมูลบรรณานุกรม การออกแบบผลลัพธ์จะเป็นการพิจารณาว่าเขตข้อมูลใดที่ผู้ใช้งานต้องการให้แสดงผล โดยทั่วไปการออกแบบผลลัพธ์จะแบ่งเป็น 2 รูปแบบ คือ รูปแบบสั้น และรูปแบบเต็ม บางฐานข้อมูลอาจเพิ่มทางเลือกในการแสดงผลรูปแบบ MARC ให้กับผู้ใช้ด้วย

- รูปแบบสั้น เป็นการแสดงเฉพาะส่วนที่เป็นข้อมูลบรรณานุกรม ในการออกแบบควรเปิดทางเลือกให้กับผู้ใช้ในการเลือกเขตข้อมูลที่ต้องการแสดงผล คือ สามารถเลือกแสดงเฉพาะเขตข้อมูลที่ต้องการ เช่น แสดงเฉพาะชื่อผู้แต่ง ชื่อเรื่อง ปีพิมพ์ คำสำคัญ และประเภทของเอกสาร นอกจากนี้ควรให้ผู้ใช้สามารถจัดลำดับเขตข้อมูลที่ต้องการแสดงผลการค้นหาได้ด้วย เช่น สามารถเรียงลำดับผลการค้นหาตามชื่อผู้แต่ง ชื่อเรื่อง หรือ ปีพิมพ์ตามแต่ความต้องการนำไปใช้ของแต่ละคน

- ควรแสดงผลรูปแบบเต็ม เป็นการแสดงผลโดยแสดงทุกเขตข้อมูลที่มีอยู่ในระเบียบตามโครงสร้างระเบียบของฐานข้อมูล และแสดงเขตข้อมูลของสารระสังเขปด้วย ซึ่งการออกแบบรูปแบบการแสดงผลรูปแบบเต็มควรเพิ่มทางเลือกในการแสดงผลเช่นเดียวกับการแสดงผลรูปแบบสั้น เพื่อเพิ่มความสะดวกในการแสดงผลให้กับผู้ใช้

ในการออกแบบ รูปแบบการแสดงผลมีสิ่งที่จะต้องพิจารณาเพิ่มเติม ได้แก่ การกำหนดตำแหน่งชื่อของเขตข้อมูลต่าง ๆ ควรกำหนดให้อ่านง่าย ไม่กำกวม หรือไม่ใกล้เคียงกับเขตข้อมูลอื่น เนื่องจากอาจสร้างความสับสนให้กับผู้ใช้ การเลือกใช้รูปแบบตัวอักษรที่ได้มาตรฐาน อ่านง่าย ขนาดเหมาะสมกับหน้าจอ รวมทั้งการเลือกใช้สีพื้นและตัวอักษรที่ตัดกันเหมาะสม เพื่อให้สามารถอ่านข้อความต่าง ๆ ได้ชัดเจน รวมทั้งการกำหนดตำแหน่งการจัดวางข้อความต่าง ๆ ทั้งนี้เพื่อความสะดวกในการเรียนรู้หน้าจอของผู้ใช้ คือ ผู้ใช้สามารถคาดเดาได้ว่าข้อความที่ต้องการจะดูนั้นจะปรากฏอยู่ที่ส่วนใดของหน้าจอ และข้อความพิเศษที่ปรากฏในตำแหน่งต่าง ๆ ของหน้าจอหมายถึงอะไร เช่น คำสั่งต่าง ๆ จะจัดวางไว้ที่ด้านล่างของหน้าจอเสมอ ขณะที่ข้อความแสดงความผิดพลาดในการทำงานจะปรากฏที่ตรงกลางของหน้าจอเสมอ เป็นต้น

ตัวอย่างรายการบรรณานุกรมแบบเต็ม (ฐานข้อมูล ERIC)

AN: EJ567665
CHN: IR537623
AU: Jones, -A.-James; Gardner, -Carrie; Zaenglein, -Judith-L.
TI: Desperately Seeking Standards: Bridging the Gap from Concept to Reality.
PY: 1998
SO: Knowledge-Quest; v26 n3 p38-42 Mar-Apr 1998.
ISSN: 1094-9046
DT: Journal-Articles (080); Reports-Descriptive (141)
LA: English
DEM: *Information-Literacy; *Learning-Resources-Centers; *Media Specialists; *National-Standards;
*Technological-Literacy
DER: Academic-Achievement; Benchmarking-; Case-Studies; Elementary-Secondary-Education
AB: Discussion of national standards for information-and-technology literacy focuses on experiences at one school where national standards were synthesized by library media specialists to develop local standards as well as a series of benchmarks by which student achievement could be measured. (Author/LRW)
CH: IR
FI: EJ
DTC: 080; 141
UD: 199902 (CIJE)

ตัวอย่างรายการบรรณานุกรมแบบย่อ (ฐานข้อมูล ERIC)

Record 1 of 1 in The ERIC Database (1992-2001/09)
AN: EJ567665
AU: Jones, -A.-James; Gardner, -Carrie; Zaenglein, -Judith-L.
TI: Desperately Seeking Standards: Bridging the Gap from Concept to Reality.
PY: 1998
SO: Knowledge-Quest; v26 n3 p38-42 Mar-Apr 1998.
AB: Discussion of national standards for information-and-technology literacy focuses on experiences at one school where national standards were synthesized by library media specialists to develop local standards as well as a series of benchmarks by which student achievement could be measured. (Author/LRW)
[View Complete Record](#)

Record 1 of 1 in The ERIC Database (1992-2001/09)
AN: EJ567665
AU: Jones, -A.-James; Gardner, -Carrie; Zaenglein, -Judith-L.
TI: Desperately Seeking Standards: Bridging the Gap from Concept to Reality.
PY: 1998
SO: Knowledge-Quest; v26 n3 p38-42 Mar-Apr 1998.
[View Complete Record](#)

3.2 การออกแบบแฟ้มตวรรษนี้ หรือแฟ้มข้อมูลผกผัน (Invert file) ทำหน้าที่ในการระบุตำแหน่งของคำข้อมูลในแฟ้มข้อมูลหลัก (master file) ที่ถูกกำหนดให้เป็นคำหลักในการสืบค้น และบันทึกในแฟ้มข้อมูลผกผันตามเงื่อนไขที่ผู้สร้างฐานข้อมูลกำหนดไว้ในตารางเลือกเขตข้อมูล (Field Select Table : FST) แฟ้มข้อมูลผกผันมีกระบวนการในการสร้าง ดังนี้

3.2.1 การพิจารณาในส่วนของเนื้อหา ซึ่งได้แก่ ลักษณะของคำที่จะกำหนดให้เป็นคำสำคัญหรือคำตวรรษนี้ ซึ่งมีประเด็นต้องพิจารณา ดังนี้

- การกำหนดคำที่จะเป็น Stop words หรือคำที่จะไม่กำหนดให้เป็นคำสำคัญในการค้น คำเหล่านี้ได้แก่ คำเชื่อม หรือคำนำหน้าต่าง ๆ เช่น with then by for form A An The กับ และ ที่ ซึ่ง อัน โดย เป็นต้น
- การพิจารณาเครื่องหมายวรรคตอน (Punctuation)ว่าจะกำหนดให้เป็นส่วนหนึ่งของคำสำคัญหรือไม่ เช่น On-line
- การพิจารณาความแตกต่างระหว่างตัวพิมพ์เล็กกับตัวพิมพ์ใหญ่ (Capitalization) จะกำหนดให้มีหรือไม่
- การกำหนดเขตข้อมูลที่จะใช้เป็นคำค้น
- การกำหนดวลี (Phrases) ให้เป็นคำตวรรษนี้เช่นเดียวกัน หรือจะกำหนดเฉพาะคำเดียวหรือคำผสมเท่านั้นที่สามารถเป็นคำตวรรษนี้ได้

3.2.2 จัดทำรายการของคำทุกคำที่ปรากฏอยู่ในระเบียบข้อมูล โดยทำการตัดเครื่องหมายต่าง ๆ ที่ปรากฏอยู่ในคำออก (หากพิจารณาในขั้นที่ 1 แล้วว่าจะไม่นำเครื่องหมายวรรคตอนที่ปรากฏอยู่ในคำต่าง ๆ มาคิดคำนวณ) จากนั้น ทำการกำหนดความต้องการเกี่ยวกับตัวพิมพ์ใหญ่ หากพิจารณาแล้วว่าจะไม่มีความแตกต่างระหว่างตัวพิมพ์ใหญ่และตัวพิมพ์เล็ก เช่น

Alexander's Feast = ALEXANDER S FEAST

3.2.3 กำหนดหมายเลขให้กับคำสำคัญทุกคำ รวมทั้งวลี ยกเว้นคำที่ไม่ใช้เป็นคำค้น โดยการกำหนดหมายเลขให้กับคำทุกคำที่อยู่ในบัญชีคำ ซึ่งหมายเลขที่กำหนดขึ้นจะเป็นตัวกำหนดตำแหน่งของคำว่าปรากฏอยู่ที่ระเบียบข้อมูลใด และอยู่ในเขตข้อมูลใด เช่น

DISCUSSES = 101 AB 1

หมายความว่า คำว่า DISCUSSES อยู่ในระเบียบข้อมูลหมายเลขที่ 101 อยู่ในเขตข้อมูล AB (Abstract) และเป็นคำในลำดับที่ 1 ของเขตข้อมูล AB

จากนั้น ทำการกำหนดหมายเลขให้คำวลี และคำผสม เช่น

MUSICAL ANALYSIS = 3,4

หมายความว่า กำหนดให้คำว่า MUSICAL เป็นคำลำดับที่ 3 และ ANALYSIS เป็นคำลำดับที่ 4 และเป็นคำที่อยู่ติดกัน

3.2.4 เรียงลำดับคำที่อยู่ในบัญชีคำตามลำดับตัวอักษร โดยเรียงตัวเลขไว้ก่อนตัวอักษร (เป็นการเรียงตามระบบ ASCII)

ตัวอย่างการสร้างแฟ้มคำดัชนี ในการค้นคืนแบบเนื้อหาเต็ม (พิมพ์วิภา, 2538: 61)

- TI A Plan for University Library Network Development in Thailand
- AB Investigates the feasibility of network of universities in Thailand, considering administrative system, library collections, involvement, in cooperative activities. Proposes a plan to establish a university library network, specifying the long-range goals, objectives, and functions of a network.

วิธีการให้หมายเลขคำดัชนี



แฟ้มคำดัชนีของระเบียบ 1 ระเบียบ

คำ	หมายเลขระเบียบ	ชื่อเขตข้อมูล	หมายเลขประโยค	หมายเลขคำ
ESTABLISH	1	AB	2	3
LIBRARY	1	TI	1	3
NETWORK	1	TI	1	4
PLAN	1	TI	1	1
	1	AB	2	2
PROPOSES	1	AB	2	1
UNIVERSITY	1	AB	2	4

การสร้างแฟ้มคำดัชนีจะอาศัยกำหนดตำแหน่งและลำดับของคำแต่ละคำที่มีความหมาย เพื่อสามารถใช้ในการสืบค้นข้อมูลต่อไป

การจัดระเบียบแฟ้มข้อมูลหลัก

Document number	Index Term
0001	T3 T4 T6 T10
0002	T1 T3 T4 T7 T9
0003	T5 T10
0004	T10 T11
0005	T2 T3 T5 T7 T8 T10

การจัดระเบียบแฟ้มข้อมูลผกผัน

Terms	Document number			
T1	0002			
T2	0005			
T3	0001	0002	0005	
T4	0001	0002		
T5	0003	0005		
T6	0001			
T7	0002	0005		
T8	0005			
T9	0002			
T10	0001	0003	0004	0005

4. การเลือกซอฟต์แวร์ การเลือกซอฟต์แวร์สำหรับงานฐานข้อมูล อาจทำได้โดยการพัฒนาโปรแกรมขึ้นเอง หรือการซื้อโปรแกรมสำเร็จรูปมาใช้ในการพัฒนาฐานข้อมูล

4.1 พัฒนาโปรแกรมขึ้นเอง โดยใช้ระบบจัดการฐานข้อมูล (Database Management system :DBMs) เช่น FoxPro, dBase, Microsoft Access ฯลฯ ซึ่งมีข้อดี คือ ช่วยป้องกันความซ้ำซ้อนของโปรแกรมที่จัดเก็บ เนื่องจากข้อมูลจะถูกจัดเก็บเพียงครั้งเดียวแต่สามารถนำมาใช้ประโยชน์ร่วมกันได้ และมีการนิยามข้อมูลและความสัมพันธ์ของข้อมูล สามารถสร้าง Query เลือกข้อมูลที่ต้องการแสดงผลจากแฟ้มข้อมูลหลายๆ แฟ้มได้ มีระบบป้องกันความปลอดภัยของข้อมูล แต่มีข้อจำกัดในการกำหนดค่าความยาวของข้อมูล เนื่องจากต้องกำหนดขนาดของเขตข้อมูลไว้ตายตัว ซึ่งทำให้เสียพื้นที่ในการจัดเก็บข้อมูล

4.2 พัฒนาเองโดยใช้ซอฟต์แวร์สำหรับจัดเก็บและค้นคืนสารสนเทศ เช่น Mini-Micro CDS/ISIS และ BRS เป็นซอฟต์แวร์ที่ได้รับการออกแบบโดยเน้นขีดความสามารถในการค้นคืนสารสนเทศ สนับสนุนการจัดทำฐานข้อมูลโปรแกรมที่มีระเบียบและเขตข้อมูลที่มีความยาวผันแปร (variable length field) เขตข้อมูลที่มีค่าซ้ำ (repeatable field) และเขตข้อมูลย่อย (subfield) ที่สำคัญคือ ความสามารถในการสืบค้นข้อมูล ช่วยให้การสืบค้นข้อมูลทำได้หลากหลายวิธีและมีความสามารถสูง เช่น ค้นแบบเนื้อหาเต็ม (คือ สามารถสืบค้นจากคำที่มีความหมายทุกคำ ยกเว้นคำที่ไม่ใช่เป็นคำค้น) ใช้เทคนิคการตัดคำ (truncation) ใช้ตัวดำเนินการ and or not เพื่อสร้างสูตรการค้น สามารถค้นจากการใช้ตัวดำเนินการระบุตำแหน่งของคำ (positional operator) การทำคำเอกพจน์ให้เป็นพหูพจน์โดยอัตโนมัติ เป็นต้น

4.3 จัดซื้อโปรแกรมสำเร็จรูป เช่น โปรแกรมห้องสมุดอัตโนมัติ อาทิ DYNIX, INNOPAC, HORIZON, TINLIB, VTLS เป็นต้น มีข้อควรพิจารณา คือ ความน่าเชื่อถือและความมั่นคงของบริษัทผู้จัดจำหน่าย หรือผู้ผลิต เนื่องจากต้องมีการดูแลรักษาอย่างต่อเนื่อง ข้อกำหนดเกี่ยวกับ

ฮาร์ดแวร์ที่นำมาใช้ สามารถใช้ได้กับฮาร์ดแวร์ทั่วไปหรือไม่ การนำเข้าข้อมูล การสืบค้น และการแสดงผลข้อมูลทำได้ง่าย การรักษาความปลอดภัยของข้อมูล การฝึกอบรมและเอกสารคู่มือการใช้ และการให้บริการหลังขาย โดยก่อนเลือกซื้อควรทำการศึกษารีวิวฮาร์ดแวร์ก่อน โดยการติดต่อกับบริษัทจัดจำหน่ายเพื่อขอทดลองใช้ ศึกษาจากบทความความปริทัศน์ หรือบทความที่เขียนโดยผู้ใช้ฮาร์ดแวร์ หรือปรึกษาผู้ที่มีความเชี่ยวชาญ หรือผู้ที่มีประสบการณ์ในการใช้งานฮาร์ดแวร์มาก่อน โดยควรศึกษาเปรียบเทียบหลาย ๆ ตัว เพื่อค้นหาฮาร์ดแวร์ที่มีความเหมาะสมกับหน่วยงานมากที่สุด

คุณลักษณะที่ต้องพิจารณาของซอฟต์แวร์พัฒนาฐานข้อมูล

- ความสามารถในการจัดการกับข้อมูลที่มีค่าความยาวต่างกันได้ และมีเขตข้อมูลที่มีค่าความยาวผันแปร
- ความสามารถในการสื่อสารกับผู้ใช้หลายระดับ ทั้งผู้ที่มีประสบการณ์มากและผู้ที่มีประสบการณ์น้อยในการใช้ใช้งานฐานข้อมูล
- แก้ไขปรับปรุงง่าย
- ได้ตอบการทำงานได้อย่างรวดเร็ว
- ความสามารถในการสืบค้น
- ความสามารถเชิงการบริหาร การตรวจสอบการทำงาน การจัดทำสถิติ ความปลอดภัยของข้อมูล และค่าใช้จ่าย
- ความสามารถในการแสดงผล
- การปฏิบัติงานในระบบออนไลน์

5. การจัดทำคู่มือ การจัดทำคู่มือเป็นกิจกรรมสำคัญที่จะช่วยให้ฐานข้อมูลที่สร้างขึ้นได้รับการใช้งานและบำรุงรักษาต่อไปอย่างถูกต้อง โดยทั่วไปคู่มือที่เกี่ยวข้องกับฐานข้อมูลจะแบ่งเป็น 2 ประเภท คือ คู่มือสำหรับผู้ปฏิบัติงานฐานข้อมูล และคู่มือสำหรับผู้ใช้งานข้อมูล

5.1 คู่มือสำหรับผู้ปฏิบัติงานฐานข้อมูล ประกอบด้วยหัวข้อต่าง ๆ ดังต่อไปนี้

- บทนำ จะระบุวัตถุประสงค์ ขอบเขต และเนื้อหาของฐานข้อมูล และควรระบุถึงความจำเป็นในการรักษาฐานข้อมูล
- การสร้างระเบียบ ประกอบด้วย โครงสร้างระเบียบ ข้อความอธิบายเขตข้อมูล แบบบันทึกการนำเข้า และวิธีการลงรายการ
- การเพิ่มระเบียบและการปรับปรุงข้อมูลให้ทันสมัย ได้แก่ วิธีการบันทึกการนำเข้า การควบคุมคุณภาพการนำเข้า ซึ่งได้แก่การตรวจสอบความถูกต้องของข้อมูล
- การสืบค้นข้อมูล ได้แก่ ผลของการนำเข้าที่มีต่อการสืบค้น และวิธีการสืบค้น
- การแก้ไขระเบียบ ประกอบด้วย วิธีการระบุตำแหน่งระเบียบที่ต้องการแก้ไข วิธีการแก้ไขการลบระเบียบ การแก้ไขแฟ้มดัชนี
- การแสดงผลทางจอภาพและการจัดทำรายงาน ได้แก่ การกำหนดรูปแบบการแสดงผล และวิธีการจัดทำรายงาน

- วิธีการแก้ไขปัญหาด่าง ๆ อาทิ ปัญหาที่อาจเกิดขึ้น ข้อความระบุปัญหา/ข้อผิดพลาด (error messages) การตรวจสอบกับคู่มือระบบ (system manual)
- 5.2 คู่มือผู้ใช้ จัดทำขึ้นเพื่อให้ผู้ใช้สามารถใช้งานฐานข้อมูลได้ และอาจช่วยแก้ไขปัญหาย่างที่อาจเกิดขึ้นระหว่างใช้งานข้อมูล และอาจใช้เป็นเครื่องมือในการค้นหาคำตอบที่เฉพาะเจาะจงให้กับผู้ใช้ แต่ปัญหาสำคัญคือ ผู้ใช้มักจะไม่ชอบอ่านคู่มือ แต่ชอบลองทำจริงมากกว่า ดังนั้น คู่มือผู้ใช้จึงควรจัดทำขึ้นเพื่อผู้ที่ต้องการอ่าน และทำความเข้าใจกับคำแนะนำต่าง ๆ ก่อนที่จะปฏิบัติจริง และสิ่งสำคัญคือ ต้องชัดเจน อ่านเข้าใจง่าย จัดลำดับเนื้อหาตามขั้นตอนของการใช้งานฐานข้อมูล และควรจัดทำให้กับผู้ใช้ที่มีระดับความสามารถในการใช้งานฐานข้อมูลต่างกัน โดยทั่วไปจะจัดทำ 3 ระดับคือ ระดับผู้ที่ชำนาญหรือใช้งานบ่อย ผู้ที่มีประสบการณ์ปานกลาง และผู้ที่มีประสบการณ์น้อยหรือยังไม่เคยใช้งานข้อมูลมาก่อน ซึ่งทั้ง 3 ระดับต้องการคำแนะนำที่ต่างกัน เนื้อหาในคู่มือผู้ใช้ควรประกอบด้วย

- ความน่า ประกอบด้วย จัดรูปแบบองค์ ขอบเขตของฐานข้อมูล และการจัดลำดับเนื้อหาของคู่มือ
- วิธีการเข้าถึงข้อมูล
- การสืบค้นข้อมูล ได้แก่ คำสั่งที่ใช้ในการสืบค้น การสืบค้นโดยใช้ตัวดำเนินการต่าง ๆ การกำหนดข้อความการสืบค้น
- การแสดงผลและการจัดทำรายงานจากฐานข้อมูล ได้แก่ การแสดงผลทางจอภาพและการสั่งพิมพ์ผลการค้นทางเครื่องพิมพ์
- วิธีการแก้ไขปัญห ได้แก่ ปัญหาที่อาจเกิดขึ้น ข้อความระบุปัญหา/ข้อผิดพลาด (error messages)
- อภิธานศัพท์ (ถ้ามี)

ข้อควรพิจารณาในการจัดทำคู่มือ ขนาดของคู่มือไม่ควรใหญ่เกินไปโดยเฉพาะคู่มือผู้ใช้ ควรเป็นกระดาษ A4 รูปเล่มต้องแข็งแรง เนื่องจากมีการใช้งานบ่อย การจัดลำดับความสัมพันธ์ของเนื้อหา เช่น ง่าย มีภาพประกอบและตัวอย่างที่จำเป็นเพื่อเสริมสร้างความเข้าใจในเนื้อหามากขึ้น คุณภาพด้านพิมพ์ ทั้งคุณภาพของกระดาษที่จะรองรับหมึกพิมพ์ และคุณภาพของหมึกพิมพ์ที่สามารถดูชัดได้ดี เพื่อความชัดเจนของข้อความ

6. การประเมินฐานข้อมูล เพื่อพิจารณาว่าฐานข้อมูลที่ผลิตขึ้นมาเป็นที่ยอมรับและมีผู้ใช้งานหรือไม่ ทั้งนี้เพื่อสามารถแก้ไขปรับปรุงข้อบกพร่องต่าง ๆ ของฐานข้อมูลได้อย่างถูกต้องและเพื่อตัดสินใจในการดำเนินงานต่อไป โดยทั่วไปมักประเมินจากความพึงพอใจของผู้ใช้ ทั้งในด้าน ขอบเขต/เนื้อหาว่า มีความครอบคลุมถูกต้อง ทันสมัยมากน้อยเพียงใด การเข้าถึงฐานข้อมูล การค้นคืน และการโต้ตอบกับระบบทำได้ง่ายเพียงใด และใช้เวลาอย่างน้อยเท่าไร และประสิทธิภาพและความเที่ยงตรงของการค้น ในบางกรณีอาจประเมินจากค่าใช้จ่ายในการดำเนินงานฐานข้อมูลกับความคุ้มค่าที่มีฐานข้อมูล โดยอาจดูจากสถิติการใช้กับค่าใช้จ่ายในการลงทุน ในส่วนของผู้ใช้ อาจพิจารณาความคุ้มค่าระหว่างผลการค้นที่ได้กับค่าใช้จ่ายที่ต้องเสียไป

7. การเผยแพร่และให้บริการ ก่อนสร้างฐานข้อมูลผู้ผลิตต้องทำการศึกษาดตลาดและกลุ่มเป้าหมาย และหลังจากดำเนินการสร้างฐานข้อมูลและประเมินผลแล้ว ผู้ผลิตจะต้องทำการโฆษณาประชาสัมพันธ์ให้กลุ่มเป้าหมายได้ทราบและต้องพิจารณาถึงวิธีการในการเผยแพร่และให้บริการ ไม่ว่าจะเป็นการด้วยสิ่งพิมพ์ที่เป็นผลผลิตจากฐานข้อมูล ได้แก่ วรรณคดีและสาระสังเขป บรรณานุกรม บริการสารสนเทศเลือกสรรเฉพาะบุคคล หรือจะให้บริการในลักษณะ ออฟไลน์ หรือ ออนไลน์ ซึ่งต้องศึกษาจากความต้องการของกลุ่มเป้าหมายเป็นหลัก และความพร้อมของผู้ให้บริการ และอีกสิ่งหนึ่งที่ต้องพิจารณาคือ ช่องทางในการเผยแพร่ ผู้ผลิตฐานข้อมูลจะให้บริการด้วยตนเอง หรือจะให้บริการผ่านหน่วยงานอื่นที่มีความพร้อมด้านเทคโนโลยีและการบริหารจัดการมากกว่า นอกจากนี้ ยังต้องพิจารณาถึงการให้บริการด้านซอฟต์แวร์ การปรับปรุงข้อมูลให้ทันสมัย การบำรุงรักษาฐานข้อมูล และการจัดการด้านการตลาดของฐานข้อมูลซึ่งมีประเด็นที่สำคัญคือ การกำหนดราคาการเข้าใช้ที่เหมาะสม รวมทั้งบริการเพิ่มคุณค่าอื่น ๆ เช่น บริการจัดส่งเอกสารฉบับเต็ม บริการเชื่อมโยงไปยังตัวข้อมูลฉบับเต็มแบบออนไลน์ บริการสืบค้นข้อมูลจากฐานข้อมูลอื่น ๆ ที่เกี่ยวข้อง เป็นต้น

บรรณานุกรม

- การค้นคืนสารสนเทศออนไลน์. 2535. ภาควิชาบรรณารักษศาสตร์และสารนิเทศศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- พิมพ์วิภา ไพรมสมิทธิ์. 2538. ฐานข้อมูลบรรณานุกรม : การสร้างและการใช้. กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์และสารนิเทศศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- Meadow, Charles T. 1992. Text Information Retrieval System. San Diego, New York : Acedemic Press.

มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 3

การจัดทำสาระสังเขป (Abstracting)

ความหมายของสาระสังเขป

สาระสังเขป (Abstract) หมายถึง งานเขียนที่เขียนโดยการย่อสาระสำคัญของเอกสาร อย่างตรงจุดมุ่งหมาย ตามลำดับและวิธีเขียนของต้นฉบับเดิม โดยปราศจากความคิดเห็นของผู้เขียนสาระสังเขป เพื่อให้ได้สาระความรู้ที่แท้จริง อย่างเที่ยงตรงตามต้นฉบับเดิม มีความสมบูรณ์ในตัวเอง เพื่อให้ผู้อ่านตัดสินใจในการเลือกอ่านเอกสารต้นฉบับได้

เรื่องย่อที่มีลักษณะใกล้เคียงกับสาระสังเขป ได้แก่

บรรณนิทัศน์ (Annotation) เป็นข้อคิดเห็น หรือ คำอธิบาย สั้น ๆ เกี่ยวกับเนื้อหาของเอกสาร

การตัดตอน (Extract) เป็นการย่อเรื่องโดยการตัดตอนข้อความที่สำคัญของเอกสารมาเรียงต่อกัน เช่น ตัด 3 บรรทัดแรกจากบทนำ แล้วตัด 2 บรรทัดหลังจากสรุป แล้วนำทั้ง 5 บรรทัดมาเรียงต่อกัน

เรื่องย่อ (Summary) เป็นการย่อความที่เน้นเฉพาะสรุปผลที่สำคัญ

บทวิจารณ์ (Review) เป็นการแสดงความคิดเห็น การวิพากษ์วิจารณ์ รวมทั้งการประเมินเนื้อหาของเอกสาร

วัตถุประสงค์ของสาระสังเขป

1. เพื่อประหยัดเวลาให้กับผู้อ่านในการติดตามสารสนเทศที่ต้องการ ทั้งที่เป็นวรรณกรรมใหม่ ทันสมัย และวรรณกรรมย้อนหลังซึ่งมีอยู่หลายรูปแบบ และอยู่กระจัดกระจายอย่างกว้างขวางทั่วโลก เนื่องจากได้นำมารวบรวมไว้ในที่เดียวกัน ภายใต้หัวข้อเรื่องต่าง ๆ
2. เพื่อให้ผู้อ่านทราบถึงเนื้อหาของเอกสารได้อย่างรวดเร็ว เนื่องจากผู้อ่านไม่ต้องอ่านเอกสารฉบับเต็มที่มีเนื้อหาจำนวนมาก และต้องใช้เวลามากในการอ่าน แต่อ่านจากสาระสังเขปที่สั้นกว่า และมีเนื้อหาถูกต้องเที่ยงตรง
3. เพื่อช่วยให้ผู้อ่านตัดสินใจได้อย่างถูกต้องว่าควรเลือกเอกสารต้นฉบับนั้น ๆ ออกมาอ่านหรือไม่
4. เพื่อช่วยจัดปัญหาด้านภาษาให้กับผู้อ่าน เนื่องจากเอกสารต้นฉบับที่รวบรวมจากที่ต่าง ๆ ทั่วโลกนั้น จัดทำขึ้นจากหลายภาษา สาระสังเขปจะจัดทำออกมาในลักษณะภาษาที่ผู้อ่านสามารถเข้าใจเนื้อหาสาระของเอกสารต้นฉบับได้ อย่างไรก็ตาม สาระสังเขปที่ผลิตออกมาเพื่อการจำหน่ายส่วนใหญ่จะจัดทำเป็นภาษาอังกฤษ เนื่องจากเป็นภาษาสากลที่ทั่วโลกเข้าใจ
5. เพื่อช่วยให้การผลิตครั้งนี้ทำได้อย่างมีประสิทธิภาพ คือถูกต้องและรวดเร็ว เนื่องจาก เป็นการดึงคำสำคัญจากเนื้อหาที่มีความกระชับที่มิเฉพาะสาระสำคัญของเอกสาร จึงช่วยให้การดึงคำสำคัญจากเอกสารทำได้ง่ายและรวดเร็ว

ประเภทของสาระสังเขป

การจำแนกประเภทของสาระสังเขปอาจจำแนกได้หลายวิธี โดยอาศัยเกณฑ์ที่แตกต่างกัน เช่น

1. จำแนกตามขนาด หรือ ความยาวของสาระสังเขป เนื่องจากความยาวของสาระสังเขปสามารถเขียนได้ตั้งแต่ 12 - 100 คำ หรือมากกว่า 1000 คำ ขึ้นอยู่กับสาระสำคัญของเอกสาร
2. จำแนกตามรายละเอียดที่บันทึกในสาระสังเขป เนื่องจากการเขียนสาระสังเขปสามารถเขียนประเด็นสำคัญของเอกสารได้หลายลักษณะ เช่น การบันทึกเนื้อหาสำคัญของเอกสารในลักษณะที่เน้นประเด็นสำคัญ หรือให้ความรู้เกี่ยวกับประเด็นสำคัญของเอกสารอย่างครบถ้วน หรือเพิ่มเติมข้อคิดเห็นเชิงวิพากษ์วิจารณ์ได้ด้วย
3. จำแนกตามชนิดของผู้เขียนสาระสังเขป เช่น สาระสังเขปที่จัดทำโดยผู้แต่งเอกสารต้นฉบับ หรือจัดทำโดยผู้เชี่ยวชาญในสาขาวิชา หรือจัดทำโดยผู้เขียนสาระสังเขปอาชีพ

ลักษณะการเขียนสาระสังเขปจะแตกต่างกันไปตามนโยบายของแต่ละสถาบัน โดยพิจารณาให้สอดคล้องกับสภาพและจุดมุ่งหมายในการให้บริการแก่ผู้ใช้ของตนเป็นสำคัญ ซึ่งอาจจัดทำเพียงประเภทเดียวหรือหลายประเภทก็ได้ ดังนั้น ผู้จัดทำสาระสังเขปจึงควรมีความรู้ความเข้าใจเกี่ยวกับลักษณะโดยทั่วไปของสาระสังเขปแต่ละประเภทอย่างชัดเจน เพื่อจะได้สามารถกำหนดรูปแบบการเขียนเนื้อหาของสาระสังเขปได้อย่างเหมาะสม

อย่างไรก็ตาม การจำแนกประเภทของสาระสังเขปนิยมแบ่งตามรายละเอียดที่บันทึกในสาระสังเขป ดังนี้

1. สาระสังเขปแบบพรรณนา หรือ แบบบอกเล่า (Descriptive Or Indicative Abstract)
2. สาระสังเขปแบบให้ความรู้ (Informative Abstract)
3. สาระสังเขปแบบพรรณนาและให้ความรู้ (Indicative-Informative Abstract)

1. สาระสังเขปพรรณนา หรือ แบบบอกเล่า (Descriptive Or Indicative Abstract)

เป็นสาระสังเขปที่เขียนอย่างสั้น ๆ เพื่อชี้แนะให้ผู้อ่านเอกสารทราบว่า เอกสารต้นฉบับกล่าวถึงอะไรบ้าง เพื่อให้ผู้อ่านทราบเรื่องที่ต้องการค้นหาในเอกสารหรือไม่ เช่น " บทความกล่าวถึงปริมาณการปลูกหัวหอมที่เพิ่มขึ้นในรัฐแคลิฟอร์เนีย" โดยปราศจากการรายงานให้ทราบผลการค้นคว้าหรือบทสรุป หรือเรื่องราวความรู้ในเอกสาร ไม่มุ่งหวังให้ชี้แทนเอกสารต้นฉบับ หากต้องการทราบ รายละเอียดต้องศึกษาจากเอกสารต้นฉบับ มักจัดทำสำหรับ หนังสือ รายงานการประชุมสัมมนา รายงานที่ไม่มีข้อสรุป ความเรียง บรรณานุกรม และบทวิจารณ์ เป็นต้น วลีที่นิยมใช้เพื่อเริ่มต้นการเขียนคือ อภิปรายเกี่ยวกับ (.....is discussed) หรือ ศึกษาเกี่ยวกับ (.....has been investigated) เนื่องจากเป็นสาระสังเขปที่ไม่ต้องให้รายละเอียดเกี่ยวกับเนื้อหาของเอกสารต้นฉบับ ดังนั้น จึงเขียนง่าย ประหยัดเวลา และแรงงานในการเขียน ผู้เขียนไม่จำเป็นต้องเป็นผู้เชี่ยวชาญเฉพาะเรื่อง

2. สาระสังเขปแบบให้ความรู้ (Informative Abstract)

เป็นสาระสังเขปที่มีวัตถุประสงค์ให้ความรู้ข่าวสารที่เป็นประเด็นของเอกสารของเอกสารต้นฉบับอย่างครบถ้วน เช่น มีการชี้แจงวัตถุประสงค์ เหตุผลการวิจัย วิธีการและเครื่องมือที่ใช้ในการวิจัย ชี้แจงผลการค้นพบที่สำคัญ และบทสรุปที่ชี้ให้เห็นความสัมพันธ์ระหว่างผลการวิจัย วัตถุประสงค์ และการนำไปใช้ เช่น

“ บทความกล่าวถึงจำนวนการปลูกหัวหอมในรัฐแคลิฟอร์เนียที่เพิ่มขึ้นถึง 1,070,000,000 ล้านหัว ” สารระสังเขปแบบให้ความรู้จึงเป็นสารระสังเขปที่เสนอข้อมูลสำคัญของเอกสารทั้งในด้านปริมาณและคุณภาพ สามารถใช้แทนเอกสารต้นฉบับได้เนื่องจากการอธิบายประเด็นสำคัญของเอกสารอย่างครบถ้วน สารระสังเขปประเภทนี้จึงมักมีขนาดยาวกว่าประเภทอื่น ๆ แต่ทั้งนี้ ความยาวขึ้นอยู่กับเอกสารต้นฉบับเป็นสำคัญ ส่วนมากสารระสังเขปประเภทนี้มักเขียนสำหรับ งานวิจัย หรือบทความวารสารที่นำเสนอมูล ความรู้ การค้นคว้าทดลอง และมีแก่นเรื่อง (theme) เดียว

3. สารระสังเขปแบบพรรณนาและให้ความรู้ (Indicative-Informative Abstract)

แนวการเขียนของสารระสังเขปประเภทนี้จะเป็นแบบสารระสังเขปประเภทให้ความรู้ และเมื่อกล่าวถึงสาระสำคัญในระดับรองลงมาจะเขียนแบบสารระสังเขปแบบบอกเล่า สารระสังเขปประเภทนี้จะช่วยให้ผู้อ่านได้ข้อมูลที่สำคัญขณะที่ยังคงรักษาความสั้นกระชับของการเขียน เช่น “ กล่าวถึงการปลูกหัวหอมในแคลิฟอร์เนียที่เพิ่มขึ้นถึง 1,070,000,000 ล้านหัว ”

นอกจากนี้ ยังมีสารระสังเขปอีกประเภทหนึ่งที่มีลักษณะพิเศษแตกต่างไปจากสารระสังเขปโดยทั่วไป คือ สารระสังเขปประเภทวิพากษ์วิจารณ์ (Critical Abstract) เป็นสารระสังเขปประเภทที่ผู้เขียนสารระสังเขปสามารถแสดงความคิดเห็น หรือ ตีความหมายของเอกสารต้นฉบับ รวมทั้งวิธีการนำเสนอในสารระสังเขปได้ ซึ่งเป็นสิ่งที่ไม่สามารถทำได้ในสารระสังเขปประเภทอื่นซึ่งผู้เขียนจะต้องเขียนไปตามเนื้อหาที่แท้จริงของเอกสารเท่านั้น การเขียนแบบวิพากษ์วิจารณ์ เช่น “ บทความนี้กล่าวถึงปริมาณหัวหอมที่ปลูกเพิ่มขึ้นในแคลิฟอร์เนีย แต่ไม่ได้บอกบอกว่าเป็นสถิติของปีใด ซึ่งทำให้คุณค่าของสารสนเทศลดลง ” ดังนั้น การเขียนสารระสังเขปแบบวิพากษ์วิจารณ์ ผู้เขียนจะต้องมีความเข้าใจในเนื้อหาของเอกสารอย่างถ่องแท้ ต้องมีทั้งความเชี่ยวชาญในเนื้อหา และมีทักษะในการเขียนสารระสังเขปเป็นอย่างดี สารระสังเขปแบบวิพากษ์วิจารณ์มักเขียนสำหรับบทวิจารณ์และหนังสือ

ตัวอย่างสารระสังเขปประเภทต่าง ๆ (Rowley, 1988)

Rowley, Jenifer ; Butcher, D and Tuner, C. 1980. Consumer information and advice : the role of public libraries. *Aslib processing* 32 (11/2).

• INFORMATIVE ABSTRACT

- An examination of the work of Consumer Advice Centres and of the information sources and support activities that public libraries can offer, CACs have dealt with pre-shopping advice, education on consumers' rights and complaints about goods and services, advising the client and often obtaining expert assessments. They have drawn on a wide range of information sources including case records, trade literature, contact files and external links. The recent closure of many CACs has seriously affected the availability of consumer information and advice. Public libraries can make many kinds of information sources more widely available, both to the public and to the agencies now supplying consumer information and advice. Libraries can cooperate closely with advice agencies through local coordinating committees, shared premises, joint publicity, referral and the sharing of professional expertise.

- **INDICATIVE ABSTRACT**

- The work of Consumer Advice Centers is examined. The information sources used to support this work are reviewed. The recent closure of many CACs has seriously affected the availability of consumer information and advice. The contribution that public libraries can make in enhancing the availability of consumer information and advice both to the public and other agencies involved in consumer information and advice, is discussed.

- **INDICATIVE-INFORMATIVE ABSTRACT**

- The work of Consumer Advice Centers and the information sources used to support this work are reviewed. The recent closure of many CACs has seriously affected the availability of consumer information and advice. Public libraries can make many kinds of information sources more widely available, both to the public and to the agencies now supplying consumer information and advice. Libraries can cooperate closely with advice agencies through local coordinating committees, share premises, joint publicity, referral and the sharing of professional expertise.

- **EXTRACT**

- With the closure of many CACs, public libraries will need to reassess their existing provision of consumer information sources, and their cooperation with, and support for, the remaining consumer information and advice services in the community.

- **SHORT ABSTRACT**

- An example of the work of Consumer Advice Centers and of the information sources and support activities that public libraries can offer in the field of consumer information advice.
- อธิบายเฉพาะเนื้อหาที่สำคัญที่เป็นประเด็นหลักเพียง 2-3 ประโยค

- **KEYWORDS**

- CONSUMER INFORMATION; CONSUMER ADVICE CENTERS; INFORMATION SOURCES; ADVICE AGENCIES; PUBLIC LIBRARIES; LOCAL COORDINATING COMMITTEES.

มหาวิทยาลัยเทคโนโลยีสุรนารี

CRITICAL ABSTRACT (Lancaster, 1998: 99)

1989. Pao, Y. C., Dept. of Eng. Mech., Univ. of Nebr., Lincoln, Shy, D. S., et al., On relationship between bulk modulus and relative volume of lung during inflation-deflation maneuvers, p 136-142, *Journal of Biomechanical Engineering, Transactions of the ASME* v 104, n 2 (May 1982).

The paper presents an equation relating the bulk modulus of the lung to the relative volume during inflation and deflation. The average bulk modulus of the lung was obtained by injecting air via a 6-mm-i.d. cannula in the main lobar bronchus. "Regional lobe" volume changes were measured by roentgen-videographically determined placement of 25 metal markers implanted in the excised lower lobes of three dogs. Whole lobe volumes at various transpulmonary pressures were measured by water displacement. Pressure and volume measurements were used to calculate bulk modulus ($K = \Delta VP/\Delta V$). The "most satisfactory least squares curve-fit" of bulk modulus (K) vs. relative volume (V/V_{max}) was obtained with the equation $K = C/(1 - V/V_{max})^n$. Substituting for bulk modulus with the equation $K = VdP/dV$, and integrating enabled computer generated pressure-volume plots. This equation provided a better pressure-volume curve-fit than previously obtained, especially at low values of pressure and volume. Also, as expected, the bulk modulus was smaller at low volume, but the rate of change of modulus was greater during deflation than during inflation.

The authors assumed, without giving sufficient justification, that the "regional lobe" (the area bounded by the 25 markers) included a higher density of airways than the rest of the lobe. Using this assumption, the authors claimed that the modulus and rate of change of modulus were different for parenchyma tissue and the airways during both inflation and deflation. No mention, however, was made of paired t-tests or any other statistical tests. In fact, if they had done a paired t-test, they would have discovered that none of these differences were significant, even at the 90 percent confidence level.

Other sources of error which were not addressed include: the difference in the properties of excised lung and intact lung due to blood in the vessels, surrounding tissue, negative pressure, etc.; the effect of the markers on the pressure-volume relationship; the effect of strain rate on the modulus of lung tissue, which is a viscoelastic material; the time elapsed between regional volume measurement and whole volume measurements (this is important for viscoelastic material); the difference between the true regional ΔV and the measured ΔV ; and the differences between the mechanical properties of dog and human lung tissue.

Despite its limitations, the paper presents a step forward in the understanding of mechanical properties of the lung, and, thus, lung diseases. Therefore, it should be of benefit to researchers interested in respiratory mechanics and physiology.

D. S. Feldman, USA

ปัจจัยที่มีผลต่อการกำหนดรูปแบบการเขียนสาระสังเขป

การจะกำหนดว่าสาระสังเขปจะต้องมีขนาดสั้นยาวเท่าใดนั้น ไม่สามารถทำได้ เนื่องจากมีปัจจัยอื่น ๆ ที่เกี่ยวข้องอีกหลายประการ และเป็นปัจจัยที่มีผลกระทบต่อวิธีการเขียน ตลอดจนขนาดความสั้นยาวของสาระสังเขป ได้แก่

1. ลักษณะเฉพาะของเอกสารต้นฉบับ เช่น

- ขอบเขตของเนื้อหา (Scope) และความยาวของเอกสาร (Length)
- ลักษณะวิชาของเอกสาร (Subject)
- ภาษาที่เขียนในเอกสาร (Language)
- ความหาง่ายของเอกสาร (Availability)
- ลักษณะการเขียนและจุดมุ่งหมายของผู้เขียน (Purpose)

ลักษณะเหล่านี้มีผลต่อการกำหนดรูปแบบของสาระสังเขป และความยาวของสาระสังเขป เช่น หากเอกสารต้นฉบับเป็นเอกสารที่หายาก หรือเขียนเป็นภาษาอื่น ควรเขียนสาระสังเขปแบบให้ความรู้ เพื่อให้ผู้อ่านได้ข้อเท็จจริงเทียบเท่ากับการอ่านจากเอกสารต้นฉบับ ส่วนเอกสารที่เนื้อหามีความยาวมาก หรือเนื้อหามีความซับซ้อน ประกอบด้วยข้อคิดเห็น หรือเรื่องราวที่หลากหลาย เช่น บทวิจารณ์ รายงานการประชุมสัมมนา หนังสือ บรรณานุกรม มักเขียนเป็นสาระสังเขปแบบพรรณนา

2. จุดประสงค์ในการจัดทำสาระสังเขปของหน่วยผลิต หากมีจุดประสงค์เพื่อชี้แนะสารสนเทศเท่านั้น จะจัดทำสาระสังเขปแบบพรรณนา แต่หากต้องการจัดทำเพื่อประโยชน์ในการค้นคว้าอย่างถาวร จะทำแบบให้ความรู้

3. ความต้องการของผู้ใช้ เช่น หากผู้ใช้ต้องการติดตามความก้าวหน้าใหม่ ๆ ในสาขาวิชา จะต้องการสาระสังเขปที่มีขนาดสั้น และผ่านการเลือกสรรเป็นอย่างดี ส่วนผู้ใช้ที่ต้องการค้นคว้าเรื่องใดเรื่องหนึ่งอย่างลึกซึ้งจะต้องการสาระสังเขปที่มีเนื้อหามากเพียงพอ ในลักษณะของแบบให้ความรู้

4. การบริหารและการดำเนินงานของสาระสังเขป ได้แก่

- คุณสมบัติของบุคลากรที่ทำหน้าที่จัดทำสาระสังเขป ได้แก่ ความสามารถด้านวิชาการ ความสามารถในการเขียนสาระสังเขป เวลาในการทำงาน และความตั้งใจในการทำงานงบประมาณในการผลิต ซึ่งจะสัมพันธ์กับจำนวนบุคลากรที่จัดทำ ค่าจ้างและเวลาในการจัดทำ
- เทคนิคในการผลิตและสื่อที่ใช้ในการบันทึก หากใช้คอมพิวเตอร์ในการจัดทำมักนิยมเขียนให้มีเนื้อหาแบบให้ความรู้ ซึ่งนอกจากจะให้ความรู้ที่ชัดเจนขึ้นแล้วยังเป็นการเพิ่มคำสำคัญในส่วนของสาระสังเขป แต่หากบันทึกในบัตรรายการมักเขียนแบบสั้นเพื่อให้สามารถบรรจุสาระสำคัญทั้งหมดลงบนแผ่นกระดาษได้

ลักษณะของสาระสังเขปที่มีคุณภาพ

ในการเขียนสาระสังเขปเพื่อให้สามารถบรรลุตามวัตถุประสงค์ที่กำหนดไว้ ต้องคำนึงลักษณะสำคัญดังต่อไปนี้

1. สั้น กระชับ (Brevity or Conciseness) การเขียนสาระสังเขปต้องเขียนให้สั้นกระชับในขณะที่ได้ประเด็นสำคัญของเอกสารครบถ้วน โดยการหลีกเลี่ยงคำที่มีความหมายคลุมเครือ และการใช้คำฟุ่มเฟือย เนื่องจากความสั้นของสาระสังเขปจะช่วยประหยัดเวลาในการอ่าน และประหยัดค่าใช้จ่ายในการจัดทำ
2. ถูกต้องเที่ยงตรงตามข้อเท็จจริง (Accuracy or objectivity) สิ่งที่สำคัญที่สุดของการเขียนสาระสังเขปคือ ต้องถูกต้องเที่ยงตรงตามเอกสารต้นฉบับ ทั้งในส่วนของเนื้อหาและบรรณานุกรมของเอกสาร เนื่องจากหากผิดพลาดจะทำให้คุณค่าของเอกสารลดลง ดังนั้น ผู้เขียนต้องเรียบเรียงตาม ข้อเท็จจริงของเอกสารต้นฉบับ โดยไม่ใส่ความคิดเห็นส่วนตัวลงไปนอกเอกสาร
3. ความชัดเจน (Clarity) การเขียนสาระสังเขปควรเขียนให้ชัดเจนไม่คลุมเครือ ใช้ภาษาที่เข้าใจง่าย หลีกเลี่ยงประโยคที่มีความกำกวม หรือการใช้ศัพท์วิชาการที่เข้าใจยาก ไม่ใช้คำย่อหรือสัญลักษณ์ที่ไม่รู้จักโดยทั่วไป หากต้องใช้ควรอธิบายไว้ในวงเล็บเมื่อมีการใช้ครั้งแรก ควรใช้คำศัพท์ที่ผู้แต่งเอกสารใช้ ไม่ควรใช้คำศัพท์ตามผู้เขียนสาระสังเขป เนื่องจากอาจไม่ตรงกับความตั้งใจของผู้แต่ง และทำให้ความหมายผิดไป
4. ความคงที่ (Consistency) การเขียนสาระสังเขปควรเขียนให้มีความคงที่ทั้งในด้าน มาตรฐานที่ใช้ในการลงรายการบรรณานุกรม รูปแบบการเขียนสาระสังเขป มาตรฐานการสะกดคำ มาตรฐานการใช้คำย่อต่าง ๆ ซึ่งจะเห็นได้อย่างชัดเจนเมื่อนำสาระสังเขป 2 รายการมาเปรียบเทียบกัน

ส่วนประกอบของสาระสังเขป

สาระสังเขปที่สมบูรณ์ประกอบด้วยส่วนประกอบที่สำคัญ 3 ส่วน ดังนี้

1. ส่วนอ้างอิง (Reference Section) เป็นส่วนที่ให้รายละเอียดทางบรรณานุกรมของเอกสารที่นำมาจัดทำสาระสังเขป
2. ส่วนเนื้อหา (Body Section) เป็นส่วนของคำบรรยาย และสาระสังเขป คือส่วนที่อธิบายเนื้อหาสำคัญของเอกสาร
3. ส่วนชื่อผู้เขียนสาระสังเขป (Signature Section) เป็นการลงชื่อผู้เขียนสาระสังเขปบทนั้น ๆ เพื่อแสดงให้ทราบว่า เป็นผลงานของผู้เขียนคนใด และอาจใส่ชื่อย่อของหน่วยงานที่ผลิตด้วย

องค์ประกอบด้านเนื้อหาของสาระสังเขปที่เป็นที่ยอมรับของการเขียนสาระสังเขปของบทความทางวิชาการหรือรายงานการวิจัย มีดังต่อไปนี้

- วัตถุประสงค์ (Purpose) เพื่อบอกให้ทราบถึงเหตุผล หรือความสำคัญของการเขียนบทความหรือการทำวิจัย เพื่อให้ผู้อ่านทราบถึงสาระสำคัญของงานนั้นวัตถุประสงค์จะพบทั้งในสาระสังเขปประเภทบอกเล่าและให้ความรู้
- วิธีการ (Methodology) คือ ระเบียบวิธีที่จะทำให้งานเป็นไปตามวัตถุประสงค์ และทำให้เกิดผลลัพธ์ และข้อสรุป สำหรับงานวิจัยวิธีการก็คือ ระเบียบวิธีวิจัย เช่น การทดลอง การสำรวจ การสัมภาษณ์ ซึ่งข้อมูลนี้ผู้ใช้ควรได้ทราบ โดยเฉพาะถ้าเป็นวิธีการค้นคว้าใหม่ ๆ ที่ไม่มีผู้ใช้มาก่อน

- ผลลัพธ์ (Result) ควรเขียนระบุไว้อย่างสั้น ๆ และให้ความรู้กับผู้อ่าน ผลลัพธ์อาจเป็นผลการค้นคว้าทดลอง ข้อมูลที่เก็บรวบรวมมาได้ หรือได้จากการวิจัย ในสารระสังเขปประเภทให้ความรู้จะกล่าวถึงผลลัพธ์ ละเอียดมากกว่าสารระสังเขปแบบบอกเล่า
- บทสรุป (Conclusion) จะอภิปรายถึงผลลัพธ์ที่ปรากฏในงาน อาจพิจารณาว่าตรงกับวัตถุประสงค์ที่ตั้งใจหรือไม่ ในบทสรุปจะมีข้อเสนอแนะ การประเมิน และข้อคิดเห็นต่าง ๆ ของผู้เขียน
- สารสนเทศอื่น ๆ (Miscellaneous Information) นอกจากหัวข้อใหญ่ ๆ ที่กล่าวมาแล้ว สารสนเทศปลีกย่อยอื่น ๆ ก็อาจนำมาเขียนไว้ได้ เช่น ตาราง ภาพประกอบ แผนภูมิ แผนที่ เอกสารอ้างอิง องค์ประกอบด้านเนื้อหาเหล่านี้ อาจไม่ต้องมีครบในสารระสังเขปของเอกสารแต่ละชิ้น ทั้งนี้ขึ้นอยู่กับเอกสารต้นฉบับเป็นสำคัญ

ตัวอย่างองค์ประกอบด้านเนื้อหาของเอกสารประเภทต่าง

ประเภทเอกสาร

งานวิจัยและประเมินผล

กรณีศึกษา

บทความวิชาการ

รายงานการประชุมเชิงปฏิบัติการ/การประชุมและสัมมนา

เอกสารแนะนำ คู่มือ และเอกสารการปฏิบัติงาน

บรรณานุกรม

องค์ประกอบด้านเนื้อหา

วัตถุประสงค์/ขอบเขต/กลุ่มเป้าหมาย/วิธีการ/ผลลัพธ์/บทสรุป

ประเภท/วัตถุประสงค์/กลุ่มผู้รับสารและระดับการศึกษา/สถานที่และวัน เดือน ปี/โครงสร้าง/ราคา/การดำเนินการ/ประโยชน์และปัญหา/การประเมินผล/บทสรุป

ประเด็นหลัก/ประเด็นรอง/บทสรุป/ข้อเสนอแนะ วัตถุประสงค์/ผู้เข้าประชุม/โครงสร้างหรือกิจกรรมที่เกี่ยวข้อง

หัวเรื่อง/กลุ่มเป้าหมายและระดับการศึกษา/วัน เดือน ปี/วัตถุประสงค์/หัวข้อสำคัญ/การแบ่งส่วนต่าง ๆ ของโครงสร้าง/วิธีการ/ลักษณะพิเศษ

หัวเรื่อง/ขอบเขตด้านภูมิศาสตร์/หัวข้อ/ผู้ใช้/ระยะเวลา/จำนวนรายการ/บรรณนิทัศน์/ดรรชนี และแหล่งสารสนเทศ

ขั้นตอนการเขียนสารระสังเขป

1. พยายามอ่านเอกสารต้นฉบับอย่างมีจุดหมาย พยายามเข้าใจเนื้อหา ขอบเขต และประเด็นสำคัญของเอกสาร ในบางครั้งอาจต้องอ่านเอกสารมากกว่า 1 ครั้ง เพื่อจับใจความ การอ่านอาจไม่ต้องอ่านทุกคำทุกตัวอักษร แต่จะอ่านคร่าว ๆ เพื่อหาใจความสำคัญ โดยอ่านจากชื่อเรื่องย่อหน้าแรก ย่อหน้าสุดท้าย หัวข้อสำคัญ ประโยคแรกของแต่ละย่อหน้าตลอดจนคำนำ บทนำ

และภาคผนวกของเอกสารต้นฉบับ รวมทั้งบทบรรณาธิการและบทวิจารณ์ เพื่อสำรวจจุดมุ่งหมายของผู้แต่งในการอ่านเอกสารนั้น ๆ แล้วอ่านซ้ำอีกในประเด็นที่พิจารณาว่าเป็นส่วนสำคัญของเอกสาร ซึ่งจะแตกต่างกันไปตามชนิดของเอกสาร ดังตัวอย่างในข้างต้น

2. เขียน หรือจดประเด็นสำคัญของเอกสารออกมาได้ ซึ่งสามารถทำไปพร้อมกับขั้นตอนที่ 1 ได้
3. ร่างสารระสังเขปตามประเด็นสำคัญที่จดเอาไว้ในขั้นที่ 2 ต้องระวังไม่นำคำหรือข้อความที่เขียนย่อออกมาจากเอกสารต้นฉบับ เขียนตามรูปแบบการเขียนที่ดี
4. ตรวจร่างสารระสังเขป ได้แก่ การตรวจวรรคตอน ความถูกต้องของการสะกดคำ ความสั้นกระชับ ความถูกต้องของชื่อเฉพาะต่าง ๆ
5. เขียนสารระสังเขปสุดท้ายให้ถูกต้องสมบูรณ์

อย่างไรก็ตาม ภาระที่สำคัญในการเขียนสารระสังเขป คือ พยายามให้ข้อมูลมากที่สุดโดยใช้คำน้อยที่สุด ปัญหาที่สำคัญคือ จะเขียนอย่างไรให้สั้นกระชับแต่ได้ใจความชัดเจน แนวการเขียนที่ดีคือ พยายามสื่อความหมายที่จำเป็นที่สุดออกมาอย่างรวดเร็ว

วิธีการเขียนให้กระชับได้ใจความ

1. ขึ้นต้นประโยคแรกด้วยประโยคที่เป็นแก่นของเรื่อง และหลีกเลี่ยงการเขียนข้อความซ้ำกับชื่อเรื่อง โดยไม่จำเป็น
2. หลีกเลี่ยงประโยคยาว ๆ หรือประโยคที่ซ้ำซ้อน
3. สารระสังเขปควรอยู่ในย่อหน้าเดียว ทุกประโยคควรเป็นประโยคที่สมบูรณ์ อาจเขียนหลายย่อหน้าได้หากเอกสารต้นฉบับนั้นยาว
4. หลีกเลี่ยงการใช้คำที่มีความหมายคลุมเครือ ระมัดระวังคำย่อ ชื่อเฉพาะ หากจะใช้ควรอธิบายคำนั้น ๆ เพื่อให้เกิดความเข้าใจ
5. พยายามตัดคำ หรือวลีที่เขียนเยื้องออก เช่น ใช้ "ผู้เขียนสรุปว่า" แทน "ผู้เขียนบทความเรื่องนี้ได้กล่าวไว้ในตอนท้ายว่า" เป็นต้น
6. ระมัดระวังการใช้กาล (Tense) และวาก (Voice) ในการเขียนสารระสังเขปภาษาอังกฤษ สารระสังเขปที่ดีต้องใช้กาลเดียวกันตลอดทั้งเรื่อง ประเภทไวยากรณ์ที่มีข้อดีคือกาล (Past tense) และกรรมวาก (Active voice) ส่วนประเภทบอกเล่านิยมใช้ปัจจุบันกาล (Present tense) และกรรมวาก (Passive voice) เพื่อการเขียนที่สั้น และชัดเจนส่วนกรรมวากควรใช้ต่อเมื่อต้องการเน้นที่การกระทำนั้น ๆ
7. หลีกเลี่ยงการใช้ศัพท์วิชาการชั้นสูงที่ยากต่อการเข้าใจ
8. เขียนตารางสั้น ๆ สูตร สมการ หรือแผนภาพเท่าที่จำเป็น หรือเมื่อไม่มีทางเลือกอื่น ๆ

การประเมินสาระสังเขป

การประเมินคุณภาพสาระสังเขป อาจพิจารณาได้จากประเด็นต่อไปนี้ (ทิมพีราไพ, 2538 : 39-40)

1. หน้าที่ (Function) ของสาระสังเขป หน้าที่ของสาระสังเขปแต่ละประเภทนั้นแตกต่างกัน สาระสังเขปประเภทให้ความรู้จะเป็นการสรุปจับประเด็นสำคัญของเอกสาร ระบุถึงแนวคิดที่สำคัญ วิธีการและข้อมูลภายในเอกสาร สามารถใช้เป็นตัวแทนของเอกสาร ส่วนสาระสังเขปประเภทพรรณนาจะระบุถึงเนื้อหาของเอกสารนั้น ๆ แนะนำให้ผู้อ่านทราบว่าเป็นเรื่องนั้นกล่าวถึงอะไร ในการพิจารณาควรเป็นไปตามประเภทของสาระสังเขป ว่าผู้เขียนใช้การเขียนที่ถูกต้องเหมาะสมกับประเภทของเอกสารนั้น ๆ หรือไม่
2. เนื้อหา (Content) สาระสังเขปควรมีองค์ประกอบด้านเนื้อหาครบตามประเภทของเอกสาร ส่วนประกอบที่ผู้ใช้มักมองหาในสาระสังเขปคือ วัตถุประสงค์ ระเบียบวิธีวิจัย ผลของการศึกษา วิจัย และสรุป หากส่วนประกอบใดหายไป อาจหมายความว่าเกิดความบกพร่องในการจัดทำ สาระสังเขป
3. รูปแบบ (Form) จะครอบคลุมถึง แนวการเขียนและความยาวของสาระสังเขป ในส่วนของแนวการเขียนที่ดี จะต้องสามารถสื่อความหมายที่จำเป็นที่สุดออกมาให้เร็วที่สุด ซึ่งปัญหาที่สำคัญในการเขียนสาระสังเขปคือ เขียนอย่างไรให้มีความกระชับรัด โดยได้ใจความชัดเจน ดังนั้น หลังจากเขียนแล้วควรอ่านทบทวนอีกครั้งเพื่อตัดข้อความที่ไม่จำเป็นออก การเขียน สาระสังเขปภาษาอังกฤษต้องใช้กาลเดียวกันตลอด ส่วนประเด็นความยาวของสาระสังเขป โดยปกติจะไม่มีกำหนดความยาวที่แน่นอนของสาระสังเขป แต่ผู้เขียนสาระสังเขปควรเขียนให้สั้นกระชับที่สุดเพื่อช่วยผู้อ่าน ความยาวของสาระสังเขปอาจยืดหยุ่นได้ตามนโยบายของหน่วยงานหรือประเภทของเอกสาร

สาระสังเขปเป็นเขตข้อมูลเพิ่มคุณค่า (Value-added field) ของฐานข้อมูล เป็นสิ่งที่มีผลต่อคุณภาพของฐานข้อมูลโดยรวมว่าสามารถอำนวยความสะดวกให้กับผู้ใช้ได้มากน้อยเพียงใด ซึ่งหากต้องการให้ฐานข้อมูลสามารถตอบสนองความต้องการของผู้ใช้ได้อย่างเต็มที่ควรเพิ่มเขตข้อมูลเพิ่มคุณค่าสาระสังเขปเข้าไปในฐานข้อมูลด้วย

บรรณานุกรม

- การค้นคืนสารสนเทศออนไลน์. 2535. กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์และสารนิเทศศาสตร์
คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ทิพย์วัลย์ ทุลยะสุข. 2542. การจัดทำบรรณานุกรมและสารสังเขป. ขอนแก่น : ภาควิชาบรรณารักษศาสตร์และ
สารนิเทศศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น.
- พิมพ์ร่ำไท เปรมสมิทธิ์. 2538. ฐานข้อมูลบรรณานุกรม : การสร้างและการใช้. กรุงเทพฯ :
ภาควิชาบรรณารักษศาสตร์และสารนิเทศศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- Cleveland, Donald B. and Cleveland, Ana D. 2001. Introduction to Indexing and Abstracting.
3rd ed. Englewood, Colorado : Library Unlimited.
- Lancaster, F. W. 1998. Indexing and Abstracting in Theory and Practice. 2nd. ed. London :
Library Association.
- Meadow, Charles T. 1992. Text Information Retrieval System. San Diego, New York : Academic
Press.
- Rowley, Jenifer E. 1988. Abstracting and Indexing. 2nd. ed. London : Clive Bingley.



มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 4

การจัดทำดัชนี (Indexing)

เมื่อผู้ใช้ทำการค้นคืนสารสนเทศ มีหลายสถานการณ์ที่อาจเกิดขึ้นได้ ทั้งสถานการณ์ที่พึงประสงค์และไม่พึงประสงค์ ได้แก่ 1) ไม่พบเอกสารที่ต้องการแม้ว่าจะมีเอกสารนั้นในฐานข้อมูล 2) พบเอกสารแต่ยังไม่ใช่อะไรที่คิดเอาไว้ 3) พบเอกสารแต่ตรงกับความต้องการเพียงบางส่วน 4) พบเอกสารและเป็นเอกสารที่ตรงกับความต้องการ อะไรเป็นสาเหตุทำให้เกิดเหตุการณ์ที่ไม่พึงประสงค์ถึง 3 สถานการณ์ และทำอะไรจึงจะทำให้เกิดเฉพาะสถานการณ์ที่ 4 เท่านั้น

การทำดัชนีเป็นคำตอบของคำถามทั้งสองข้อ ดรรชนีที่ไม่มีคุณภาพนำมาซึ่งสถานการณ์ที่ 1-3 ขณะที่ดรรชนีที่มีคุณภาพนำมาซึ่งสถานการณ์ที่ 4 ดรรชนี คืออะไร มีบทบาทต่อการค้นคืนสารสนเทศอย่างไร และดรรชนีที่มีคุณภาพเป็นอย่างไร และมีวิธีการจัดทำอย่างไร เป็นสิ่งที่จะกล่าวถึงในบทนี้

ความหมายและวัตถุประสงค์ของดรรชนี

ดรรชนี (Index) เป็นเครื่องชี้หน้าที่จัดทำขึ้นอย่างเป็นระบบเพื่อชี้ไปยังตำแหน่งของคำหรือแนวคิดที่สำคัญในเอกสาร ประกอบด้วยรายการดรรชนีที่จัดเรียงอย่างเป็นระบบ โดยทั่วไปมักเรียงตามลำดับตัวอักษร เพื่ออำนวยความสะดวกแก่ผู้ใช้ในการค้นหาเอกสารที่ต้องการ ดรรชนีมักกำหนดจากสาระสำคัญของเอกสาร ชื่อบุคคล ชื่อสถานที่ ชื่อเฉพาะต่าง ๆ ที่เป็นแกนเรื่องของเอกสาร

การทำดรรชนี (Indexing) เป็นกระบวนการที่เกี่ยวข้องกับการวิเคราะห์เนื้อหาของเอกสาร โดยการอธิบายเนื้อหาของเอกสารออกมาเป็นคำ หรือวลีสั้น ๆ เพื่อชี้ไปยังตำแหน่ง ที่อยู่ของสารสนเทศในเอกสาร หรือเพื่อชี้ไปยังเอกสารที่เกี่ยวข้อง

ดรรชนีมีหลายประเภทดังจะได้กล่าวต่อไป แต่ในที่นี้จะเน้นเฉพาะ ดรรชนีเรื่อง (Subject Index) ซึ่งเป็นการกำหนดคำสำคัญเพื่อใช้แทนเนื้อหาของเอกสาร โดยการนำสาระสำคัญ หรือแนวคิดของเอกสารมาจัดทำเป็นภาษาในระบบการจัดทำดรรชนี เพื่อแสดงเนื้อหาของเอกสาร และเป็นเครื่องมือในการจัดเก็บและการค้นคืนสารสนเทศ

ดรรชนีเรื่องและสาระสังเขปจะทำหน้าที่เกี่ยวข้องกันในการเป็นตัวแทนของสาระในสารสนเทศ เพียงแต่สาระสังเขปจะเป็นการอธิบายสาระของสารสนเทศโดยย่อออกมาเป็นประโยค ขณะที่ดรรชนีอธิบายเนื้อหาโดยใช้คำดรรชนี ซึ่งมักจะเลือกใช้ตามรูปแบบของศัพท์ควบคุม

หน้าที่ของสาระสังเขปก็คือ ชี้ให้ทราบว่าเอกสารนั้นเป็นเรื่องเกี่ยวกับอะไร หรือย่อสาระสำคัญของเรื่อง กลุ่มของคำดรรชนีสามารถทำหน้าที่นี้ได้เช่นเดียวกัน ดังตัวอย่างของกลุ่มคำต่อไปนี้ ก็สามารถสื่อถึงสาระสำคัญของเรื่องว่าเกี่ยวข้องกับอะไรได้เช่นเดียวกัน

ศูนย์สารสนเทศ การใช้ทรัพยากรร่วมกัน สหบรรณานุกรม

การทำรายการร่วมกัน เครือข่ายออนไลน์ การยืมระหว่างสถาบัน

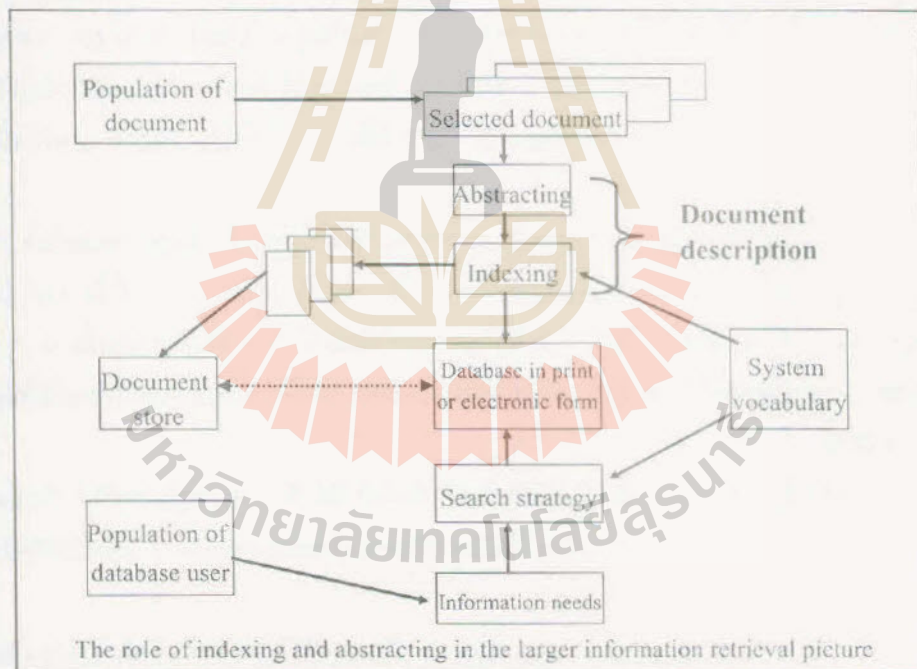
โดยความรู้สึกแล้วคำเหล่านี้สามารถใช้เป็นสาระสังเขปแบบย่อ (mini abstract) ได้ เนื่องจากสามารถตอบสนองวัตถุประสงค์ของสาระสังเขปได้ถ้าแสดงคำเหล่านี้ไว้ด้วยกันทั้งหมด

ความสำคัญของดรรชนี

เช่นเดียวกันกับทางด่วนสำหรับให้รถวิ่งที่ต้องการป้ายสัญญาณและแผนที่ในการบอกทาง ทางด่วนสารสนเทศ (Information superhighway) ก็ต้องการป้ายสัญญาณและแผนที่ในการบอกทางเพื่อให้ไปถึงสารสนเทศที่ต้องการได้อย่างรวดเร็วและถูกต้อง ซึ่งป้ายสัญญาณและแผนที่บนทางด่วนสารสนเทศก็คือ ดรรชนีและสาระสังเขปนั่นเอง ในยุคของข้อมูลข่าวสารเทคโนโลยีสารสนเทศช่วยเพิ่มศักยภาพให้กับการทำดรรชนี ปัจจุบันเราไม่เพียงแต่สร้างระบบในการทำดรรชนีเท่านั้น แต่ยังสร้าง Robot Spider และ meta-search tools เป็นตัวช่วยในการรวบรวมสารสนเทศ การจัดการทรัพยากรสารสนเทศได้พ้นออกมาจากกำแพงของห้องสมุดออกมาสู่สังคมภายนอก สู่เครือข่ายข้อมูลข่าวสารอย่างอินเทอร์เน็ต สิ่งที่ใช้ในยุคสารสนเทศต้องการคือการวิเคราะห์เนื้อหาของสารสนเทศอย่างเป็นระบบเพื่อที่จะสามารถชี้แนะพวกเขาไปสู่สารสนเทศที่ต้องการได้

ดรรชนีเป็นสิ่งที่มีจำเป็นสำหรับการค้นคืนสารสนเทศ เนื่องจากดรรชนีเป็นตัวชี้ไปยังตัวข้อมูลและแหล่งข้อมูล ผู้ใช้ดรรชนีสามารถค้นหาสารสนเทศที่ต้องการได้ตรงประเด็น และชี้แหล่งที่เก็บของสารสนเทศได้อย่างถูกต้อง รวดเร็ว ช่วยประหยัดเวลาและเพิ่มประสิทธิภาพในการค้น ซึ่งทั้งสองนี้จะเกิดขึ้นได้ก็ต่อเมื่อผู้ทำดรรชนีเลือกใช้คำที่ตรงกับภาษาของผู้ใช้มากที่สุด และมีรายการโยงไปยังสารสนเทศที่เกี่ยวข้องสัมพันธ์กัน

ดรรชนีและสาระสังเขปจะช่วยเพิ่มคุณค่าให้กับเอกสาร เนื่องจากดรรชนีจะช่วยให้ผู้ใช้พบเอกสาร นั้น หมายถึงเอกสารจะถูกนำออกมาใช้มากขึ้น ไม่เกิดความสูญเปล่าในการจัดทำเนื่องจากไม่ถูกนำออกมาใช้



หน้าที่ของดรรชนี

ดรรชนีเป็นเครื่องมืออำนวยความสะดวกในการค้นหาสารสนเทศ 2 ระดับ คือ

1. ดรรชนีค้นหาตัวเล่มเอกสาร (Guide to the hole works) เช่น รายการสารสนเทศในห้องสมุดเป็นเครื่องมือชี้แนะผู้ใช้ให้สามารถค้นหาหนังสือเล่มที่ต้องการจากหนังสือที่มีอยู่ในห้องสมุดจำนวนมาก โดยใช้ดรรชนีหัวเรื่อง ดรรชนีผู้แต่ง หรือ ดรรชนีชื่อเรื่อง หัวเรื่องที่กำหนดขึ้นจึงครอบคลุมเนื้อหาของเอกสารอย่างกว้าง ๆ

เพื่อให้ทราบว่าหนังสือเล่มใดบ้างที่มีเนื้อหาเกี่ยวกับเรื่องนั้น ๆ ไม่ได้มีจุดมุ่งหมายเพื่อให้เห็นความแตกต่างของเนื้อหาในหนังสือเล่มเดียวกัน ลักษณะหัวข้อเรื่องจึงมีลักษณะกว้าง (general) ตัวอย่างพรรณานี้ที่จัดทำขึ้นเพื่อค้นหาตัวเล่มเอกสารอีกประเภทหนึ่งคือ พรรณานี้ค้นหาบทความในวารสารพรรณานี้และสาระสังเขป (หรือฐานข้อมูล) ซึ่งต้องใช้พรรณานี้หัวข้อเรื่อง ชื่อผู้แต่ง และชื่อเรื่องค้นก่อนจึงจะได้รายการบรรณานุกรมของเอกสารที่จะนำไปสู่การค้นหาตัวเล่ม นอกจากนี้ ระบบการจำแนกหมวดหมู่สารสนเทศ เช่น ระบบทศนิยมดิวอี้ ระบบหอสมุดรัฐสภาอเมริกัน ซึ่งเป็นการจำแนกสารสนเทศออกเป็นกลุ่ม ๆ จัดเป็นพรรณานี้ค้นหาเอกสารอีกประเภทหนึ่ง เนื่องจากไม่ได้มีจุดมุ่งหมายในการค้นหารายละเอียดต่าง ๆ ภายในตัวเล่ม

2. พรรณานี้ค้นหาเนื้อหาภายในตัวเล่ม (Guide to content of work) เป็นพรรณานี้ที่จัดทำขึ้นเพื่อค้นหาสารสนเทศที่เฉพาะเจาะจงภายในเนื้อหาของเอกสารมากกว่าต้องการเพียงทราบชื่อเรื่องเท่านั้น พรรณานี้ที่จัดทำขึ้นเพื่อจุดประสงค์ดังกล่าว ได้แก่ พรรณานี้ท้ายเล่ม ซึ่งบอกให้ทราบว่าคำ หรือเรื่องที่ต้องการปรากฏอยู่ในหน้าใดของเอกสารเล่มนั้น

ประเภทของพรรณานี้

1. พรรณานี้ผู้แต่ง (Author Indexes) เป็นพรรณานี้ที่ใช้ชื่อผู้แต่งเป็นส่วนหลักของรายการพรรณานี้ จัดเรียงตามลำดับอักษรชื่อผู้แต่ง ซึ่งอาจเป็นชื่อบุคคล หรือนิติบุคคล จัดทำขึ้นเพื่อค้นหารายการเอกสารตามชื่อผู้แต่ง การทำรายการพรรณานี้ผู้แต่งควรใช้หลักเกณฑ์การลงรายการที่เป็นมาตรฐานเพื่ออำนวยความสะดวกในการค้นคืน และควรคำนึงถึงความถูกต้องของชื่อ เกณฑ์ในการจัดทำกรณที่มีผู้แต่งหลายคนจะกำหนดให้ทุกชื่อเป็นพรรณานี้หรือไม่ การทำพรรณานี้ให้กับผู้ที่ใช้ชื่อในการเขียนหลายชื่อ การจัดเรียงลำดับชื่อ และควรมีคู่มือในการปฏิบัติงาน

2. พรรณานี้หัวข้อเรื่อง (Subject Indexes) เป็นพรรณานี้ที่ใช้คำหรือหัวข้อเรื่องที่มีผู้กำหนดไว้แล้วเป็นหัวข้อเรื่องที่แสดงเนื้อหาหรือแนวคิดเหมือนกันจะจัดเรียงไว้ภายใต้ หัวเรื่องเดียวกัน มีการใช้รายการโยง (cross reference) โยงคำที่มีความหมายเดียวกัน หรือเกี่ยวข้องสัมพันธ์กัน ซึ่งกระจายอยู่ตามที่ตั้งต่าง ๆ เพื่อให้ผู้ใช้สามารถเข้าถึงรายการพรรณานี้ได้สะดวก อาจมีการอธิบายของเขตหรือความหมายของคำเพื่อให้เข้าใจความหมายชัดเจนยิ่งขึ้น

3. พรรณานี้คำ (Word Indexes) เป็นพรรณานี้ที่สร้างโดยใช้คำที่ปรากฏอยู่ในชื่อเรื่องหรือเนื้อหาของเอกสารนั้น ๆ มาเป็นคำค้นหรือคำหลัก และจัดเรียงตามลำดับอักษรของคำหลักนั้น ๆ พรรณานี้ที่จัดทำเป็นพรรณานี้คำ ได้แก่

3.1 อภิธานพรรณานี้ (Concordance) เป็นพรรณานี้ที่ใช้คำสำคัญทุกคำที่ปรากฏในเนื้อหาของเอกสารมาเป็นส่วนหลักหรือคำค้นแล้วจัดเรียงตามลำดับอักษร ผู้จัดทำไม่จำเป็นต้องมีทักษะในการจัดทำพรรณานี้ สามารถจัดทำได้อย่างรวดเร็วและใช้เครื่องคอมพิวเตอร์ทำได้ แต่จะมีข้อด้อยในการค้นหาคำที่มีความหมายเหมือนกันแต่เขียนต่างกัน เนื่องจากจะกระจายอยู่หลายที่ ทำให้หาเอกสารได้ไม่ครอบคลุมทั้งหมด

3.2 พรรณานี้เวียนคำ (Permuted Indexes) เป็นการสร้างรายการพรรณานี้หัวข้อเรื่องตามลำดับอักษร โดยการหมุนเวียนคำสำคัญภายในชื่อเรื่อง (ชื่อบทความหรือชื่อเอกสาร) แต่ละคำมาเป็นส่วนหลักหรือคำหลักของรายการพรรณานี้ บางครั้งเรียกว่า Permuted Title Indexes หรือ Permuted Word Indexes คำที่ไม่สำคัญจะถูกตัดทิ้งโดยรายการคำหยุด

(stop words) ที่กำหนดไว้ การสร้างดัชนีลักษณะนี้มาจากแนวคิดที่ว่า ชื่อเรื่องมักจะมี ความกระชับและสื่อถึงเนื้อหาของเอกสาร ดังนั้น ประสิทธิภาพของดัชนีประเภทนี้จึง ขึ้นอยู่กับความสามารถในการตั้งชื่อเรื่องของเอกสารให้สอดคล้องกับเนื้อหาของเอกสาร มักใช้กับงานด้านวิทยาศาสตร์ และมักใช้คอมพิวเตอร์จัดทำ แบ่งได้เป็น 2 ประเภท คือ

3.2.1 ดรรชนีควิก (KWIC Index : Keyword-In-Context Index) เป็นการนำคำ สำคัญจากชื่อเรื่องของเอกสารมาหมุนเวียนเป็นคำดรรชนี โดยเรียงตาม ลำดับอักษรของคำสำคัญที่ปรากฏในชื่อเรื่อง โดยคำที่ถูกกำหนดให้เป็น คำดรรชนีจะถูกจัดเรียงไว้ตรงกลาง

KWIC Index	
IN TEXAS.	<u>BLUE-EYED CATS</u>12
THE	<u>CAT</u> AND THE ECONOMY.....12
THE	<u>CAT</u> AND THE FIDDLE.....17
DOGS AND	<u>CATS</u> AND THEIR DISEASES.....3
BLUE-EYED	<u>CATS</u> IN TEXAS.....12
AND THEIR	<u>DOGS</u> , <u>DOGS</u> AND <u>CATS</u>3
THEIR DISEASES,	<u>DOGS</u> AND <u>CATS</u> AND.....3
AND THE	<u>ECONOMY</u> , THE <u>CAT</u>12
AND THE	<u>FIDDLE</u> , THE <u>CAT</u>17
IN	<u>TEXAS</u> , <u>BLUE-EYED</u> <u>CATS</u>12

3.2.2 ดรรชนีควอก (KWOC Index : Keyword-out-of-Context Index) มี ลักษณะเช่นเดียวกับ KWIC Index คือหมุนเวียนคำสำคัญจากชื่อเรื่อง มาเป็นคำดรรชนี แต่คำที่ถูกกำหนดให้เป็นคำดรรชนีในแต่ละครั้งจะถูก นำมาจัดเรียงไว้ทางด้านซ้ายของชื่อเรื่องแล้วพิมพ์ชื่อเต็มของเอกสารทาง ด้านขวามือ เพื่อสะดวกในการค้นหา

KWOC Index	
BLUE-EYED	BLUE-EYED CATS IN TEXAS.....12
CAT	THE CAT AND THE ECONOMY.....12
CAT	THE CAT AND THE FIDDLE.....17
CATS	DOGS AND CATS AND THEIR DISEASES.....3
CATS	BLUE-EYED CATS IN TEXAS.....12
DISEASES	DOGS AND CATS AND THEIR DISEASES.....3
DOGS	DOGS AND CATS AND THEIR DISEASES.....3
ECONOMY	THE CAT AND THE ECONOMY.....12
FIDDLE	THE CAT AND THE FIDDLE.....17
TEXAS	BLUE-EYED CATS IN TEXAS.....12

3.3 **ดรรชนียูนิเทอม (Uniterm Indexes)** เป็นดรรชนีที่ใช้คำหรือวลีที่แสดงแนวคิดเดียวจากเนื้อเรื่องของเอกสารมากำหนดเป็นหัวเรื่อง และจากหัวเรื่องจะสามารถทราบได้ทันทีว่าเอกสารนั้น ๆ มีเนื้อเรื่องเกี่ยวกับอะไร คำที่ใช้เป็นหัวเรื่องอาจเป็นตัวเลขหรือเครื่องหมายอื่นใดก็ได้ การกำหนดหัวเรื่องประเภทนี้ไม่จำเป็นต้องควบคุมการใช้คำศัพท์ คำที่เลือกจากเนื้อเรื่องจะมีมากหรือน้อยขึ้นอยู่กับเนื้อหาของเอกสาร

4. **ดรรชนีประสานคำ (Coordinate Indexes)** เป็นดรรชนีหัวเรื่องที่สร้างขึ้นโดยการผสมคำตั้งแต่ 2 คำขึ้นไป เพื่อสร้างหัวเรื่องใหม่ มี 2 ลักษณะ คือ

4.1 **ดรรชนีประสานคำล่วงหน้า (pre-coordinate indexes)** เป็นดรรชนีที่มีการประสานคำตั้งแต่ระยะการจัดทำดรรชนี มักพบในดรรชนีที่จัดทำในรูปสิ่งตีพิมพ์ เช่น บัญชีหัวเรื่อง ซึ่งมีการแสดงหัวเรื่องหลักและหัวเรื่องย่อย

4.2 **ดรรชนีประสานคำภายหลัง (post-coordinate indexes)** เป็นการประสานคำดรรชนีเพื่อสร้างคำใหม่ในระหว่างการค้น โดยการสร้างความสัมพันธ์ระหว่างคำโดยใช้ตรรกะบูลีน AND OR NOT

5. **ดรรชนีการอ้างถึง (Citation Indexes)** เป็นดรรชนีรายการงานเขียนที่มีผู้อื่นนำไปอ้างในรายการอ้างอิงท้ายเอกสาร โดยมีวัตถุประสงค์เพื่อให้เป็นเครื่องมือในการค้นหางานเขียนที่มีผู้อ้างถึง และรายการงานเขียนที่อ้างจากงานของผู้อื่น เพื่อความถี่เขียนคนใดที่อ้างถึงงานเขียนนั้น ๆ บ้าง ดรรชนีการอ้างถึงจึงเปรียบเสมือนดรรชนีหัวเรื่อง เพราะการที่ผู้แต่งอ้างถึงงานเขียนของผู้อื่นก็เนื่องมาจากมีความเกี่ยวข้องสัมพันธ์กับงานของตน จึงช่วยในการค้นหาเอกสารที่เกี่ยวข้องกัน ช่วยให้ทราบถึงพัฒนาการของสาขาวิชาหรือเรื่องดังกล่าว ดรรชนีการอ้างถึงที่รู้จักแพร่หลายได้แก่ Science Citation Index (SCI)

ส่วนที่ 1 **ดรรชนีการอ้างถึง (The citation index)** คือ ดรรชนีการอ้างถึงนั้น ประกอบด้วยชื่อผู้ที่ถูกอ้างถึงเป็นหัวเรื่องของดรรชนี ภายใต้อ้างถึงจะมีรายการงานที่อ้างถึงโดยย่อ คือ มีชื่อผู้เขียน ชื่อย่อของวารสาร ปีที่ หน้า ปีที่พิมพ์ แต่ไม่มีชื่อเรื่องของงานนั้น ถ้างานเขียนรายการใดมีผู้อ้างถึงหลายคนจะเรียงรายการผู้เขียนที่อ้างถึงนั้นตามลำดับอักษร

Citation Indexes			
BROADUS AD	VOL	PG	YR
69 CLIN RES 17 65			
FRANKLIN TJ NATURE-BIOL	246	119	73
MARCEL VL UN MED CAN	102	876	73
70 JCLIN INVEST 49 3222			
AUGUST GP JCLIN END	37	476	73
HO JK J PHARM EXP	185	336	73
KUCHELO UN MED CAR	102	2488	73
LATNER AL CLIN CHIM A	48	353	73
RAIJK SCJ-CLIN V	31	196	73
STEINER AL METABOLISM	22	1139	73
TURINSKY JP SOC EXP M	142	1189	73
WOO YT ARCH BIOCH	154	510	73

Science Citation Index, Part 1.

ส่วนที่ 2 ดรรชนีแหล่งอ้างอิง (Source index) เป็นส่วนที่แสดงรายการบรรณานุกรมของงานที่อ้างอิงอย่างครบถ้วน รวมทั้งภาษาของบทความและจำนวนการอ้างอิง ตัวอย่างเช่น รายการบรรณานุกรมของเอกสารที่เขียนโดย Steiner A.L. ที่อ้างอิงงานเขียนของ Broadus ในส่วนที่ 1



ส่วนที่ 3 ดรรชนีหัวเรื่องเรียงคำ (Permuted subject index) เป็นรายการหัวเรื่องต่าง ๆ ที่ได้จากการนำคำสำคัญในชื่อเรื่อง (ชื่อบทความ) มาหมุนเวียนกันเรียงตามลำดับอักษร

6. ดรรชนีการจัดหมวดหมู่ (Classified Indexes) เป็นดรรชนีที่จัดเรียงคำดรรชนีตามลำดับความสัมพันธ์ของเรื่อง โดยเรียงจากเรื่องทั่วไปเข้าไปหาเรื่องที่มีความเฉพาะเจาะจงมากขึ้น สิ่งที่ต้องพิจารณาคือ การจัดกลุ่มและการจัดลำดับความสัมพันธ์ของเรื่องต้องชัดเจน และต้องมีดรรชนีเรียงตามลำดับอักษรเป็นเครื่องมือช่วยค้น
7. ดรรชนีท้ายเล่ม (Book Indexes) ได้แก่ รายการคำสำคัญที่ปรากฏในตัวเล่มของเอกสาร มักจัดทำไว้ท้ายเล่ม จัดเรียงตามลำดับอักษร คำที่ดึงมาทำดรรชนีมักมีความเกี่ยวข้องกับสาระสำคัญของเอกสาร เช่น หนังสือด้านการท่องเที่ยว คำที่นำมาทำเป็นดรรชนีท้ายเล่ม ได้แก่ ชื่อสถานที่ ชื่อเหตุการณ์ทางประวัติศาสตร์ ชื่อโรงแรม ชื่อแหล่งท่องเที่ยว เป็นต้น
8. ดรรชนีวารสาร (Periodical Indexes) เป็นรายการดรรชนีที่จัดทำขึ้นเพื่อค้นหาเนื้อหาหรือบทความที่ปรากฏอยู่ในวารสาร แบ่งเป็น 2 ประเภท คือ
 - 8.1 ดรรชนีวารสารเฉพาะชื่อ (Individual indexes) เป็นดรรชนีที่จัดทำขึ้นเพื่อค้นหาบทความที่ปรากฏอยู่ในวารสารชื่อใดชื่อหนึ่ง มักจัดทำเมื่อครบปี (Volume) โดยจัดทำ 3 แบบ คือ ดรรชนีชื่อผู้แต่ง ดรรชนีชื่อเรื่อง และดรรชนีหัวเรื่อง
 - 8.2 ดรรชนีวารสารเฉพาะวิชา (Broad Indexes) เป็นดรรชนีที่จัดทำขึ้นเพื่อค้นหาบทความในสาขาวิชาใดวิชาหนึ่ง โดยจัดทำ 3 แบบ คือ ดรรชนีชื่อผู้แต่ง ดรรชนีชื่อเรื่อง และดรรชนีหัวเรื่อง เช่นเดียวกับดรรชนีวารสารเฉพาะชื่อ

สรุป

ดรรชนี เป็นเครื่องมือช่วยค้นที่จัดทำขึ้นอย่างเป็นระบบ โดยทำหน้าที่ชี้ไปยังตำแหน่งของคำ หรือแนวคิดที่สำคัญในเอกสาร เป็นเครื่องมืออำนวยความสะดวกในการค้นหาสารสนเทศ 2 ระดับ คือ ดรรชนีช่วยค้นสารสนเทศภายในเล่มและดรรชนีช่วยค้นหาตัวเล่มเอกสาร ดรรชนีจะประกอบด้วยรายการดรรชนีที่จัดเรียงอย่างเป็นระบบ โดยทั่วไปมักเรียงตามลำดับอักษร ส่วนประกอบที่สำคัญของดรรชนีคือ ส่วนที่เป็นส่วนหลัก หรือ คำหลัก เนื่องจากเป็นส่วนที่ใช้เป็นหลักในการจัดเรียงรายการ

ดรรชนี เพื่อให้ผู้ใช้สามารถค้นหาสารสนเทศที่ต้องการได้ ส่วนหลักหรือคำหลักอาจเป็น หัวเรื่อง ชื่อผู้แต่ง สูตรโมเลกุล หมายเลขสิทธิบัตร หรือชื่ออื่น ๆ ก็ได้ การเรียกชื่อดรรชนีจึงมักเรียกตามส่วนหลัก เช่น หากใช้ชื่อผู้แต่งเป็นส่วนหลักจะเรียกดรรชนีผู้แต่ง เป็นต้น

ภาษาดรรชนี (Indexing Language)

ในส่วนต้นของบทนี้ ได้กล่าวถึง ดรรชนี (Index) ซึ่งเป็นเครื่องมือที่มีการสร้างไว้อย่างเป็นระบบเพื่อค้นหาสารสนเทศ โดยที่แนะนำผู้ใช้ไปสู่เนื้อหาของเอกสารที่ต้องการ รวมถึงชี้แนะไปยังแหล่งเอกสาร โดยทั่วไปในระบบการจัดเก็บและค้นคืนสารสนเทศ ดรรชนีมีส่วนประกอบที่สำคัญ 2 ส่วน คือ 1) รายละเอียดทางบรรณานุกรม ซึ่งแสดงข้อมูลทางกายภาพของเอกสาร และสัญลักษณ์ซึ่งแสดงที่อยู่ของเอกสาร และ 2) คำแทนสาระของเอกสาร ซึ่งอาจสร้างขึ้นโดยมนุษย์ หรือเครื่องจักรก็ได้ ขึ้นอยู่กับภาษาดรรชนีที่หน่วยงานเลือกใช้

ในส่วนที่สองนี้จะกล่าวถึงภาษาดรรชนี ซึ่งเป็นส่วนสำคัญในการกำหนดคำแทนสาระให้กับเอกสาร

ภาษาดรรชนี หรือภาษาเอกสาร (Indexing Language or Document language)

ภาษาดรรชนีเป็นภาษาที่หน่วยผลิตและบริการสารสนเทศกำหนดขึ้นสำหรับใช้ในการอธิบายเนื้อหาของเอกสารเพื่อสะดวกในการจัดเก็บและค้นคืน ภาษาดรรชนีมีหลายประเภทแตกต่างกันไปตามรายละเอียด ขอบเขต โครงสร้าง และการใช้ประโยชน์ ฯลฯ เช่น การจัดหมวดหมู่ (Classification) หัวเรื่อง (Subject heading) คำสำคัญ (Keyword) ศัพท์บ่งชี้ (Descriptor) ศัพท์สัมพันธ์ (Thesaurus) พจนานุกรมศัพท์เฉพาะ (Lexicons) เป็นต้น การจะเลือกใช้ภาษาดรรชนีประเภทใดจึงขึ้นอยู่กับวัตถุประสงค์ นโยบายและสภาพแวดล้อมของหน่วยผลิตและบริการสารสนเทศแต่ละแห่งเป็นสำคัญ

ประเภทของภาษาดรรชนี

โดยทั่วไปภาษาดรรชนีแบ่งได้เป็น 2 ประเภท คือ

1. ภาษาธรรมชาติ (Natural Language) หรือภาษาดรรชนีแบบไม่ควบคุมคำศัพท์ (Uncontrolled vocabulary) เป็นการทำดรรชนีโดยการดึงคำสำคัญที่พบในเนื้อหาของเอกสารทั้งฉบับ หรือในชื่อเรื่อง หรือในสาระสังเขป ฯลฯ เป็นศัพท์ดรรชนี (index terms) ที่ใช้ในการค้นคืนเอกสารนั้น ตัวอย่างดรรชนีประเภทนี้ เช่น ดรรชนีชื่อผู้แต่ง ดรรชนีชื่อเรื่อง ดรรชนีการอ้างถึง ดรรชนีเรื่อง (เฉพาะกรณีที่น่าคำศัพท์ที่ปรากฏในเนื้อเรื่องมาเป็นคำดรรชนี) เนื่องจากดรรชนีภาษาธรรมชาติเป็นการดึงคำสำคัญที่ปรากฏอยู่ในเอกสารออกมาเป็นคำศัพท์ดรรชนีโดยตรง ดังนั้น จึงสามารถใช้คอมพิวเตอร์ในการจัดทำ โดยคอมพิวเตอร์จะอ่านและวิเคราะห์ค่าความถี่ของคำที่ปรากฏในเอกสาร คำใดที่มีความถี่อยู่ในกรอบทฤษฎีที่กำหนดจะถูกดึงออกมาเป็นคำดรรชนี โดยตัดคำที่เป็นคำหยุดออก ตัวอย่างของดรรชนีภาษาธรรมชาติที่ใช้คอมพิวเตอร์ในการจัดทำเช่น KWIC Index และ KWOC Index เหมาะสำหรับการจัดทำดรรชนีให้กับเอกสารที่มีปริมาณมาก

ข้อดีของการทำดัชนีแบบภาษาอิสระ

- คำศัพท์ที่ใช้เป็นดัชนีมีความใหม่ ทันกับพัฒนาการของความรู้ในเรื่องนั้น ๆ อยู่เสมอ
- คำศัพท์ที่ผู้เขียนเอกสารใช้กับคำที่ใช้เอกสารใช้ตรงกัน
- คำศัพท์ที่มีความเฉพาะเจาะจงตรงกับเนื้อหา
- ผู้ใช้จดจำได้ง่าย
- ใช้คอมพิวเตอร์จัดทำได้ซึ่งประหยัดกว่าการใช้คนจัดทำ

ข้อด้อยของการทำดัชนีแบบภาษาอิสระ

- กรณีที่เอกสารมีเนื้อหาเดียวกันแต่ผู้เขียนใช้คำศัพท์ต่างกัน (คำพ้องความหมาย) ผู้ใช้จะต้องใช้คำศัพท์หลายคำเพื่อค้นหาเอกสารเรื่องเดียวกันแต่ใช้คำศัพท์ต่างกันในการกำหนดค่าแทนสาระให้ได้ทุกรายการ ซึ่งในความเป็นจริงผู้ใช้ไม่สามารถทราบคำศัพท์เหล่านั้นได้ทุกคำ ทำให้ผู้ใช้ไม่สามารถค้นพบเอกสารทั้งหมด ซึ่งนับเป็นการเสียโอกาสในการใช้เอกสารดังกล่าว
- คำศัพท์ที่สะกดเหมือนกันอาจมีความหมายต่างกัน (คำพ้องรูป) ดังนั้น จึงทำให้สารสนเทศที่ไม่เกี่ยวข้องออกมาด้วย เนื่องจากภาษาราชการไม่มีกลไกในการควบคุมคำศัพท์
- กรณีของคำศัพท์ที่สะกดได้หลายแบบแต่มีความหมายเดียวกัน ผู้ใช้ต้องทำการค้นหลายครั้งจึงจะได้เอกสารทั้งหมด
- ประเด็นที่แฝงอยู่ในเอกสารแต่ไม่ได้ปรากฏออกมาเป็นคำดัชนี จะไม่ถูกดึงออกมาทำให้ไม่สามารถค้นเอกสารเรื่องดังกล่าวออกมาได้
- ภาษาราชการไม่มีรายการโยงที่แสดงความสัมพันธ์ระหว่างคำศัพท์ ผู้ใช้จึงต้องใช้ความรู้ความสามารถในการคิดคำต่าง ๆ ที่สัมพันธ์กันเหล่านั้นด้วยตนเอง ซึ่งอาจคิดได้ไม่ครบถ้วนสมบูรณ์ ทำให้การค้นคืนสารสนเทศทำได้ไม่สมบูรณ์ตามไปด้วย

2. ภาษาควบคุม หรือ ศัพท์ควบคุม (Controlled Vocabulary) เป็นภาษาดัชนี ที่มีคนเป็นผู้ควบคุมการสร้างความที่จะใช้แทนเนื้อหาของเอกสาร และมีผู้ควบคุมวิธีการกำหนดคำสำหรับเอกสารแต่ละรายการ ซึ่งโดยทั่วไปจะมีบัญชีศัพท์ควบคุม (Authority list) สำหรับใช้ในการกำหนดคำดัชนี และผู้จัดทำดัชนีจะแปลแนวความคิดสำคัญที่ได้จากเอกสารไปเป็นคำดัชนีตามที่กำหนดไว้แล้วในบัญชีศัพท์ควบคุมอย่างเหมาะสม เพื่อให้สามารถค้นคืนสารสนเทศได้อย่างมีประสิทธิภาพและมีมาตรฐานในการจัดทำดัชนี การใช้ศัพท์ควบคุมจะช่วยให้การกำหนดค่าแทนสาระของเอกสารทำได้ถูกต้องและคงที่ (consistency) มีมาตรฐาน เนื่องจากจะควบคุมให้มีการใช้คำที่เหมาะสมที่สุดเพียงคำเดียว การควบคุมในขั้นนี้ได้แก่

2.1 การควบคุมเชิงภาษา เช่น การกำหนดวิธีการทางไวยากรณ์ในการเขียนคำศัพท์ การวางเกณฑ์ให้คำพ้องความหมายทุกคำโยงไปสู่คำใดคำหนึ่งที่กำหนดให้เป็นศัพท์ดัชนี การวางแนวปฏิบัติต่อคำที่มีความหมายใกล้เคียงกัน การทำคำอธิบายให้แก่คำศัพท์ที่มีความหมายไม่ชัดเจน การนำศัพท์ที่สัมพันธ์กันแสดงไว้ด้วยกัน และการพยายามกำหนดให้

ใช้คำที่สอดคล้องกับการใช้ของคนทั่วไป เป็นต้น ทั้งนี้มีจุดมุ่งหมายเพื่อลดความกำกวม และความซ้ำซ้อนของคำศัพท์ เพื่ออำนวยความสะดวกในการเข้าถึงสารสนเทศของผู้ใช้โดยการให้คำศัพท์ที่คุ้นชิน

2.2 การควบคุมจำนวน หมายถึงการควบคุมให้เอกสารแต่ละรายการมีจำนวนศัพท์ดรรชนีในปริมาณหนึ่ง ซึ่งพอเพียงต่อการเข้าถึงเนื้อหาของเอกสาร

ข้อดีของภาษาควบคุม ได้แก่

- กระบวนการเปรียบเทียบคำตามกับข้อมูลในระบบทำได้ง่าย และเป็นกระบวนการที่สั้น เนื่องจากผู้ทำดรรชนีและผู้ใช้ซึ่งเป็นผู้ค้นคืนสารสนเทศใช้คำศัพท์ตรงกัน
- ประหยัดเวลาในการสืบค้นเรื่องที่มีคำพ้องความหมายหลาย ๆ คำ คำที่มีความหมายใกล้เคียงกันและคำที่สะกดต่างกัน เนื่องจากทั้งเอกสารและคำศัพท์ถูกกำหนดให้ใช้ศัพท์ดรรชนีคำใดคำหนึ่งเพียงคำเดียวเป็นคำแทนสาระ และมีการโยงคำที่ไม่ใช่ไปสู่คำที่ใช้ได้ด้วย
- สามารถขยายขอบเขตการสืบค้นไปสู่แนวคิดที่เกี่ยวข้องสัมพันธ์กันได้ตามต้องการ โดยอาศัยศัพท์ในชุดความสัมพันธ์หนึ่ง ๆ เป็นเครื่องนำทาง
- คำศัพท์ในบัญชีคำแทนสาระไม่ซ้ำซ้อน ทำให้บัญชีมีความกระชับและประหยัดพื้นที่

ข้อด้อยของภาษาควบคุม

- ลงทุนสูง ทั้งในด้านสติปัญญา เวลา และ แรงงาน ในการสร้างบัญชีคำศัพท์ดรรชนี
- ต้องใช้ความสามารถของคนในการวิเคราะห์เนื้อหาของเอกสาร ซึ่งใช้เวลามากกว่าให้เครื่องคอมพิวเตอร์ทำ หากมีเอกสารจำนวนมากอาจจัดทำไม่ได้ทันกับความต้องการของผู้ใช้
- ความถูกต้องของการกำหนดคำศัพท์ขึ้นอยู่กับความสามารถของจัดทำดรรชนี ในการวิเคราะห์เนื้อหาของเอกสารและการเลือกคำศัพท์ที่เหมาะสมที่สุดเป็นคำแทนสาระของเอกสาร

การกำหนดคำดรรชนี

1. อ่านเอกสารเพื่อพิจารณาเนื้อหาสาระที่เอกสารกล่าวถึง พยายามเก็บรวบรวมแนวคิดที่สำคัญของเอกสาร
2. ถ่ายทอดแนวคิด เนื้อหาสาระออกมาเป็นศัพท์อิสระในภาษารวมชาติ หรือใช้คำที่พบในเอกสารนั้น
3. ใช้คู่มือเพื่อค้นหาศัพท์บังคับ ซึ่งในบางครั้งอาจตรงกับศัพท์อิสระ หากไม่ตรงจะต้องพิจารณาว่าจะเลือกคำใดจากคู่มือจึงจะตรงกับเนื้อหามากที่สุด ผู้จัดทำคำดรรชนีควรพิจารณาดังนี้

3.1 เลือกคำที่เหมือนกับศัพท์อิสระ

ศัพท์อิสระ	:	ศัพท์ควบคุม
DOGS	:	DOGS

3.2 เลือกคำที่มีความหมายเหมือน หรือใกล้เคียงกัน

ศัพท์อิสระ ศัพท์ควบคุม
PUPILS : STUDENTS

3.3 มีคำที่มีความหมายแคบกว่าที่จะใช้แทนได้หรือไม่

ศัพท์อิสระ ศัพท์ควบคุม
COMPUTERS : MICROCOMPUTERS

3.4 ควรใช้คำที่มีความหมายกว้างกว่าหรือไม่

ศัพท์อิสระ ศัพท์ควบคุม
INFORMATION FLOW : INFORMATION FLOW
BT COMMUNICATION

3.5 ควรใช้คำที่เกี่ยวข้องกันหรือไม่

ศัพท์อิสระ ศัพท์ควบคุม
INFORMATION CENTER : INFORMATION SERVICES
RT INFROMATION MANAGEMENT

อย่างไรก็ตาม การที่จะเลือกคำตรรกะนี้มักจะต้องคำนึงถึงเนื้อหาของเอกสารเป็นสำคัญ เช่น เรื่องมลภาวะ (pollution)

- POLLUTION
- UF Environment Pollution
- BT Environment Degradation
- NT Air Pollution
- Industrial Pollution
- Marine Pollution
- Noise Pollution
- Radioactive Pollution
- Thermal Pollution
- Water Pollution

ในกรณีนี้เอกสารกล่าวถึงเฉพาะมลภาวะทางเสียงเท่านั้น ก็จะต้องเลือกเฉพาะ " Noise Pollution " แต่หากเอกสารกล่าวถึงมลภาวะโดยรวมและกล่าวถึงมลภาวะทางเสียงด้วยก็ต้องเลือก ทั้ง Pollution และ Noise Pollution

ข้อผิดพลาดที่อาจเกิดขึ้นในการกำหนดคำตรรกะนี้

1. การเลือกคำที่ใช้ผิด ทั้งนี้ อาจเนื่องมาจากไม่เข้าใจเนื้อหาของเอกสาร หรือ ไม่รอบคอบในการเนื้อหาของเอกสาร
2. การไม่ให้คำที่ใช้แทนแนวคิดที่สำคัญ เนื่องจากในขั้นตอนการวิเคราะห์เอกสารผู้วิเคราะห์อ่านข้ามตอนสำคัญของเรื่องหรืออ่านแล้วไม่เข้าใจว่ามีแนวคิดที่สำคัญอะไรบ้าง จึงไม่ได้ให้คำศัพท์ที่เกี่ยวข้องกับเรื่องนั้น

3. การใช้คำที่กว้างหรือแคบเกินไป การใช้คำที่กว้างเกินไปจะทำให้ได้เอกสารที่ไม่ตรงความต้องการ เช่น เอกสารฉบับหนึ่งกล่าวถึงพัฒนาการทางเศรษฐกิจ "Economic development" เท่านั้น แต่ผู้จัดทำคำตรรกะนี้ให้คำศัพท์ว่า "Economy" ซึ่งเป็นคำที่มีความหมายกว้างกว่ามากทำให้เมื่อผู้ใช้ได้รับเอกสารรายการนี้แล้วคาดหวังว่าจะได้ภาพกว้าง ๆ เกี่ยวกับเศรษฐกิจ แต่กลับได้เอกสารที่กล่าวถึงเฉพาะ "Economic development" เท่านั้น แต่ในทางตรงข้ามหากกำหนดคำตรรกะที่มีความหมายแคบเกินไป อาจทำให้เอกสารฉบับนั้นหลุดหายไป เช่น เอกสารกล่าวถึงปัญหาการจ้างงาน "Employment" แต่ผู้จัดทำเห็นว่าเอกสารฉบับนี้เน้นการจ้างงานในชนบท "Rural Employment" จึงกำหนดศัพท์ตรรกะที่มีความเฉพาะกับเรื่องมากขึ้นคือ "Rural Employment" เท่านั้น ทั้ง ๆ ที่เอกสารกล่าวถึงเรื่องการจ้างงานโดยทั่ว ๆ ไปด้วย ดังนั้น หากที่ผู้ที่ต้องการเอกสารเกี่ยวกับการจ้างงานก็จะได้เอกสารฉบับนี้ เนื่องจากไม่ได้กำหนดคำตรรกะ "Employment" ไว้ด้วย ดังนั้น ในการกำหนดคำตรรกะจึงควรเลือกใช้คำที่มีความหมายชัดเจน และตรงกับเนื้อหามากที่สุด และควรหลีกเลี่ยงศัพท์เทคนิค ศัพท์วิชาการ และหลีกเลี่ยงการใช้คำที่ไม่เหมาะสม เช่น ควรใช้คำ "เด็กอ่อนแอ" แทนคำว่า "เด็กจรจัด" หรือใช้คำว่า "ชุมชนแออัด" แทนคำว่า "สลัม" เป็นต้น

ดังนั้น การกำหนดคำตรรกะจะต้องอาศัยความเข้าใจและทักษะของผู้จัดทำเป็นสิ่งสำคัญ ตั้งแต่การวิเคราะห์เนื้อหาของเอกสาร และกำหนดลำดับขั้นเป็นตัวแทนแนวคิดที่สำคัญ โดยอาจใช้คำที่ปรากฏในเอกสาร ใช้ศัพท์ควบคุม หรือใช้ทั้ง 2 แบบรวมกัน โดยเลือกคำจากเอกสาร ชื่อเฉพาะ คำใหม่ หรือคำที่ไม่มีปรากฏในคู่มือเพื่อช่วยให้ผลการสืบค้นตรงตามความต้องการของผู้ใช้มากยิ่งขึ้น

โดนัลด์ บี เครฟแลนด์ และ เอนา ดี เครฟแลนด์ กล่าวว่า การทำคำตรรกะไม่ใช่เรื่องยาก คนที่อ่านออกเขียนได้ รู้จักตัวอักษร ก็สามารถจัดทำคำตรรกะนี้ได้ เพราะคำสำคัญทุกคำปรากฏอยู่ในเอกสาร สิ่งที่เราต้องทำก็เพียงแค่ตั้งคำต่าง ๆ เหล่านั้นออกมา เช่นเดียวกับที่ มาร์ค ทเวน (Mark Twain) กล่าวว่า การเขียนไม่ใช่เรื่องยาก เพียงแค่นำคำต่าง ๆ มาเรียงต่อกันเป็นแถวบนหน้ากระดาษ และนั่นหมายความว่า ทุกคนสามารถเขียนได้ เพราะคำทุกคำมีอยู่ในพจนานุกรม แต่ทั้งสองอย่างเป็นการปฏิบัติที่ต้องใช้ความตั้งใจในการศึกษาและปฏิบัติ เราจะนำคำมาเรียงต่อกันเป็นแถวอย่างไร เพื่อทำให้เกิดให้เกิดอรรถาธิบายในกรณีอื่น จะตั้งคำออกมาอย่างไรจึงจะสามารถเป็นตัวแทนที่ดีของเอกสารและนำไปสู่การค้นคืนที่ถูกต้องได้ การทำคำตรรกะไม่ได้อาศัยเพียงแต่ความสามารถในการอ่านออกเขียนได้ และความรู้สึกรู้สึกว่าถูกหรือผิด ควรจะเป็นหรือไม่เป็น แต่ต้องอาศัยความรู้ความเข้าใจเกี่ยวกับหลักเกณฑ์ในการทำคำตรรกะซึ่งมีพัฒนาการมากกว่าศตวรรษ นอกจากนี้ยังต้องเข้าใจถึงประโยชน์ของคำตรรกะ รู้จักผู้ที่จะนำคำตรรกะไปใช้ วิธีการและวัตถุประสงค์ของการนำคำตรรกะไปใช้ (Cleveland and Cleveland, 2001 : 1)

บรรณานุกรม

- การค้นคืนสารสนเทศออนไลน์. 2535. กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์ คณะอักษรศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย .
- ทิพย์วัล ตุลยะสุข. 2542. การจัดทำบรรณานุกรมและสาระสังเขป. ขอนแก่น : ภาควิชาบรรณารักษศาสตร์และ
สารนิเทศศาสตร์ คณะมนุษยศาสตร์ มหาวิทยาลัยขอนแก่น.
- นฤมล ปราชญ์โยธิน, ทวีศักดิ์ กอนันตกุล, เปรมิน จินดาวิมลเสิต. 2536. อิซอร์สกับระบบสารสนเทศ.
กรุงเทพฯ : ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ.
- พิมพ์ฟ้าไพ เปรมลมิทธิ์. 2538. ฐานข้อมูลบรรณานุกรม : การสร้างและการใช้. กรุงเทพฯ : ภาควิชา
บรรณารักษศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย .
- Cleveland, Donald B. and Cleveland, Anna D. 2001. Introduction to Indexing and Abstracting.
3 rd ed. Colorado : Library Unlimited.
- Lancaster, F. W., 1998. Indexing and Abstracting in Theory and Practice. London : Library Association.



มหาวิทยาลัยเทคโนโลยีสุรนารี

บทที่ 5

ธิซอรัส (Thesaurus)

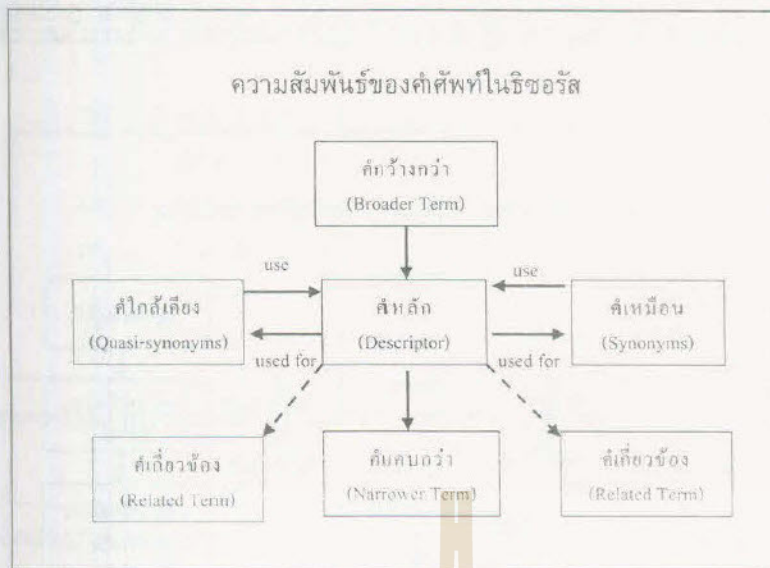
ความหมายและวัตถุประสงค์ของธิซอรัส

ธิซอรัส เป็นคำทับเสียงคำภาษาอังกฤษว่า Thesaurus ซึ่งมีรากศัพท์มาจากภาษากรีก มีความหมายดั้งเดิมว่า ชุมคลัง หรือ ชุมสมบัติ ของสรรพสิ่ง ต่อมาเมื่อผู้ใช้นี้ในความหมายของพจนานุกรม ในภาษาไทยมีคำเรียก Thesaurus อยู่หลายคำ อาทิ คลังคำ ชุมทรัพย์คำ ศัพท์สัมพันธ์ บัญชีศัพท์สัมพันธ์ ทะเบียนศัพท์สัมพันธ์ และธิซอรัส ซึ่งเป็นคำที่ต่อมาจากคำภาษาอังกฤษ แต่คำที่นิยมใช้มากที่สุดคือ ศัพท์สัมพันธ์ อย่างไรก็ตาม คลังคำ นับว่าเป็นคำแปลที่พยายามรักษาความหมายของศัพท์ดั้งเดิมไว้ โดยกลุ่มฯ รักษาชนิดได้อธิบายความหมายของคลังคำว่า คลังคำเป็นหนังสือที่รวบรวมคำศัพท์ที่มีความหมายใกล้เคียงกัน หรืออยู่ในชุดคำเดียวกันมาได้ด้วยกัน แล้วจัดหมวดหมู่เพื่อให้ผู้ใช้สามารถค้นหาได้ง่ายขึ้น เป็นหนังสือที่ใช้เสริมพจนานุกรมอีกทีหนึ่ง โดยใช้ส่วนทางกับพจนานุกรม กล่าวคือ พจนานุกรมตั้งต้นที่คำไปสู่ความหมาย คลังคำตั้งต้นที่ความหมายไปสู่คำ (นฤมล, 2536 : 49)

เมื่อนำธิซอรัส มาใช้ในงานสารสนเทศ ธิซอรัสนี้ยังคงความหมายของการเป็นคลังคำ แต่เป็นคลังคำของศัพท์ควบคุม คือ จะกำหนดหรือแนะนำให้ใช้ศัพท์คำใดคำหนึ่ง เป็นตัวแทนของกลุ่มคำที่มีรูปต่างกัน หรือเขียนต่างกันแต่มีความหมายเดียวกัน และจะประมวลศัพท์ที่อยู่ในกลุ่มเดียวกันไว้ด้วยกัน โดยมีการแสดงโครงสร้างความสัมพันธ์ของคำในกลุ่ม และมีสัญลักษณ์แสดงความสัมพันธ์กับได้ภายใต้รูปแบบที่ได้มาตรฐาน โดยสัญลักษณ์ที่ใช้แทนความสัมพันธ์จะมีรูปแบบเฉพาะตัวมากกว่าบัญญัติคำศัพท์ประเภทอื่นๆ

วัตถุประสงค์ของธิซอรัส คือ จัดทำขึ้นเพื่อเป็นเครื่องมือที่มีประสิทธิภาพในการจัดเก็บและค้นคืนสารสนเทศด้วยวิธีการเข้าถึงสารสนเทศผ่านทางเนื้อหา (subject access) คำว่าประสิทธิภาพในที่นี้ได้แก่ ความสามารถในการสืบค้นข้อมูลเรื่องใดเรื่องหนึ่งได้อย่างครบถ้วน โดยไม่ต้องทราบคำศัพท์ที่เกี่ยวข้องทั้งหมด และความสามารถในการช่วยเหลือให้ผู้จัดทำบรรณานุกรมและผู้สืบค้นใช้ศัพท์ที่เป็นมาตรฐานเดียวกัน ทำให้การกำหนดคำแทนสาระของเอกสารที่มีความเที่ยงตรง





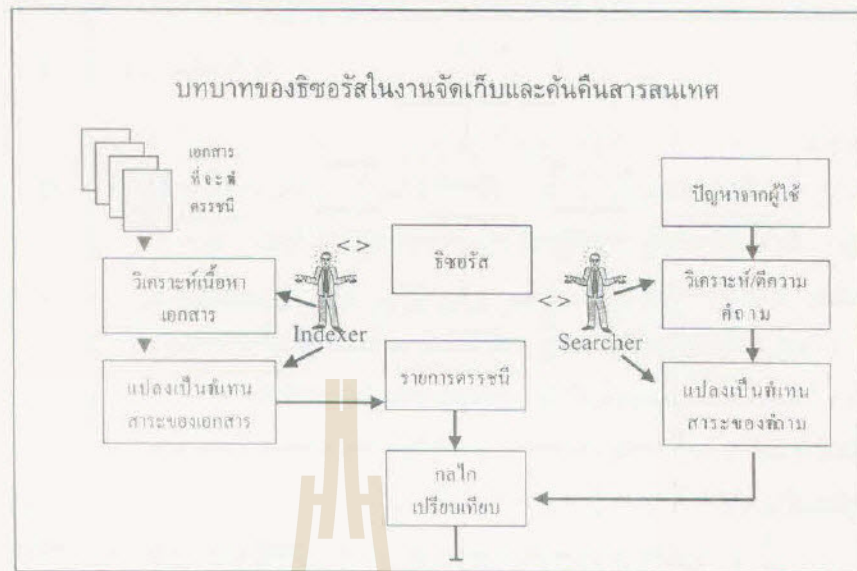
ขอบเขตของธิซอร์ส

- 1) โน้มนำเนื้อหา ธิซอร์สมักมีขอบเขตครอบคลุมเฉพาะคำศัพท์ในแขนงวิชาใดวิชาหนึ่ง อย่างไรก็ตาม มีธิซอร์สบางชื่อที่มีนโยบายในการสร้างเพื่อใช้ในการทำดัชนีให้กับเอกสารไม่จำกัดสาขาวิชา ได้แก่ UNBIS Thesaurus ของสหประชาชาติ เน้นด้าน กิจกรรมมนุษย์ การบริหาร การเมือง การศึกษา เศรษฐกิจ สังคม และวัฒนธรรม และ Root Thesaurus ซึ่งพัฒนาโดย The British Standard Institution เน้นด้านมาตรฐานเกี่ยวกับวัสดุ และกระบวนการในแง่คุณภาพ ปริมาณ การนำไปใช้ และเทคโนโลยี โดยธิซอร์สทั้งสองชื่อมีวัตถุประสงค์เพื่อการแลกเปลี่ยนสารสนเทศ ด้านการมาตรฐานและเอกสารอื่น ๆ และออกแบบมาเพื่อเป็นคู่มือในการทำดัชนีเอกสารจาก ทุก ๆ ประเทศ และในสาขาวิชาต่าง ๆ อย่างกว้างขวาง แต่มีจุดเน้นต่างกัน และในแต่ละสาขาวิชา จะให้น้ำหนักไม่เท่ากัน เช่น Root Thesaurus จะให้ด้านสังคมศาสตร์และมนุษยศาสตร์น้อยกว่า สาขาอื่น ๆ ส่วน UNBIS Thesaurus จะให้น้ำหนักด้านวิทยาศาสตร์และเทคโนโลยีน้อยกว่าสาขาอื่น
- 2) ในแง่ภาษา ธิซอร์สทั้งที่เป็นภาษาใดภาษาหนึ่ง (Monolingual Thesaurus) และธิซอร์สหลายภาษา (Multilingual Thesaurus)

บทบาทของธิซอร์สในการจัดเก็บและค้นคืนสารสนเทศ

1. บทบาทในขั้นตอนของการจัดเก็บ ธิซอร์สจะทำหน้าที่เป็นคู่มือของนักทำดัชนีในระบบภาษา ควบคุม กล่าวคือ นักทำดัชนีจะวิเคราะห์เนื้อหาของเอกสาร แล้วจำแนกแนวคิดของเอกสาร ออกมาเป็นคำดัชนี โดยอาศัยธิซอร์สเป็นหลักในการเลือกใช้คำศัพท์ให้มีความหมายเหมาะสมกับ เนื้อหาที่แท้จริงของเอกสาร และควบคุมความคงเส้นคงวาในการใช้คำศัพท์ให้ตรงกันทุกครั้งที่ต้อง กำหนดคำแทนสาระให้กับเอกสารที่มีเนื้อหาในแง่มุมมองเดียวกัน ทั้งนี้เพื่อให้เกิดประสิทธิภาพและ ประสิทธิภาพในขั้นตอนการค้นคืนต่อไป
2. บทบาทในขั้นตอนของการค้นคืน ในกระบวนการค้นคืน ผู้สืบค้นจะอาศัยธิซอร์สเป็นเครื่องมือในการ ค้นหาคำที่ตรงกับความหมายของเรื่องที่จะสืบค้น โดยจะเริ่มต้นด้วยการวิเคราะห์และตีความ

ของคำถาม แล้วแปลงคำถามออกมาเป็นคำตรรกะนี้ ซึ่งเป็นหลักการเดียวกับการวิเคราะห์เอกสารเพื่อการจัดเก็บ



องค์ประกอบของอิซอร์ส

อิซอร์สเป็นคลังของศัพท์ควบคุม ซึ่งประกอบด้วย ศัพท์เป็นชุด ๆ แต่ละชุดมีส่วนประกอบดังนี้

1. คำหลักของชุด (Descriptor)
2. คำที่มีความหมายพ้องกับคำหลัก หรือคำที่มีความหมายใกล้เคียงกับคำหลัก แต่ระบบไม่ใช้และกำหนดให้ใช้คำหลักแทน (used for)
3. ข้อความอธิบายคำหลัก (scope note) เพื่อความชัดเจนของคำศัพท์ แต่ไม่จำเป็นต้องทำให้กับคำศัพท์ทุกคำ
4. คำที่มีความสัมพันธ์กับคำหลัก ได้แก่
 - คำที่เป็นต้นสกุลของคำหลัก (top term)
 - คำที่มีความหมายกว้างกว่าคำหลัก (broader term)
 - คำที่มีความหมายแคบกว่าคำหลัก (narrow term)
 - คำที่มีความหมายเกี่ยวข้องกับคำหลัก (related term) แต่ไม่ใช่ในลักษณะที่เป็นความหมายกว้างกว่า หรือแคบกว่า

สัญลักษณ์ในอิซอร์ส ที่ใช้ส่วนใหญ่มี ดังนี้ (สัญลักษณ์เหล่านี้จะวางอยู่หน้าคำศัพท์)

SN	=	Scope note	ใช้นำหน้าข้อความอธิบายคำศัพท์
U	=	Use	ใช้นำหน้าคำศัพท์ที่กำหนดให้ใช้
UF	=	Used for	ใช้นำหน้าคำศัพท์ที่ไม่ใช้
TT	=	Top term	ใช้นำหน้าคำศัพท์ต้นสกุล หรือคำศัพท์กว้างที่สุด หรือคำรวม
BT	=	Broader term	ใช้นำหน้าคำศัพท์ที่มีความหมายกว้างกว่าคำหลัก
NT	=	Narrower term	ใช้นำหน้าคำศัพท์ที่มีความหมายแคบกว่าคำหลัก
RT	=	Related term	ใช้นำหน้าคำศัพท์ที่มีความหมายเกี่ยวข้องกับคำหลัก

ตัวอย่างการใช้สัญลักษณ์

R & D	
SN	Abbreviation for RESEARCH AND EXPERIMENTAL DEVELOPMENT
UF	RESEARCH AND DEVELOPMENT
NT	EXPERIMENTAL DEVELOPMENT RESEARCH
RT	CREATIVITY ENGINEERING

ชุดความสัมพันธ์ของคำศัพท์ที่เป็นคำกว้างสุดหรือต้นสกุล

ACADEMIC ACHIEVEMENT	
UF	Gradepoint Average Scholastic Achievement School Achievement
BT	ACHIEVEMENT
NT	ACADEMIC OVERACHIEVEMENT ACADEMIC UNDERACHIEVEMENT
RT	ACADEMIC ACHIEVEMENT MOTIVATION ACADEMIC ACHIEVEMENT PREDICTION ACADEMIC APTITUDE ACADEMIC FAILURE

ชุดความสัมพันธ์ของคำศัพท์ที่มีสถานภาพอยู่ในช่วงกลาง ๆ

JOURNALISTIC ETHICS	
BT	PROFESSIONAL ETHICS
RT	POLITICAL ETHICS
TT	ETHICS

ชุดความสัมพันธ์ของคำศัพท์ที่เป็นคำแคบสุด

นอกจากนี้อาจพบสัญลักษณ์บางอย่างในธิอรรถบางชื่อ เช่น

ACR	=	Acronym	คำย่อ
ANT	=	Antonym	คำศัพท์ที่มีความหมายตรงข้าม
BUF	=	Broadly used for	เมื่อนำหน้าคำใดหมายความว่าไม่ให้ใช้คำนั้น แต่ให้ใช้คำที่มาข้างหน้าแทน (ซึ่งเป็นคำที่มีความหมายกว้างกว่าคำที่ไม่ให้ใช้นั้น ๆ)
DT	=	Date of input	วันที่ป้อนข้อมูล

EQ	=	Equivalence/ Quasi synonym	คำเหมือน หรือคำใกล้เคียง
HN	=	History note	เป็นการแจ้งประวัติการใช้คำบางคำให้ผู้ทราบ เพื่อให้สามารถสืบค้นข้อมูลภายใต้คำสืบค้นที่ถูกต้อง
HOM	=	Homonym	คำพ้องเสียง
JRG	=	Jargon	ภาษาเฉพาะกลุ่ม
PN	=	Posting note	จำนวนที่ปรากฏของคำนั้น
PT	=	Prior term	ความหมายทำนองเดียวกับ HN
SC	=	Subject code	รหัสแสดงหมวดหมู่
SEE			ใช้แสดงโอกาสทางเลือก กรณีที่คำนั้นมีหลายบริบท
SYN	=	Synonym	คำพ้องความหมาย

ลักษณะความสัมพันธ์ของคำในธิขอรัส

ความสัมพันธ์ของคำศัพท์ในธิขอรัส จำแนกได้เป็น 3 ลักษณะ ดังนี้

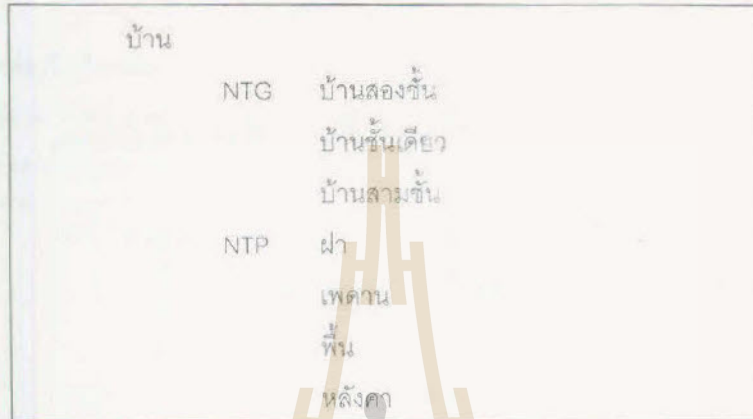
1. ความสัมพันธ์ในลักษณะเท่าเทียมกัน (Equivalence relationship) ได้แก่ คำพ้องความหมาย คำย่อ อักษรย่อ คำบัญญัติ คำยืมจากภาษาอื่น คำแปล คำภาษาถิ่น คำที่สะกดได้หลายแบบ เป็นต้น
- สัญลักษณ์ที่ใช้ ได้แก่ UF = Used for และ U = Use ตัวอย่าง เช่น

ชุมชนแออัด		สลัม		U	ชุมชนแออัด
UF	สลัม		U		ชุมชนแออัด
	แหล่งเสื่อมโทรม				แหล่งเสื่อมโทรม
					U
					ชุมชนแออัด
UNESCO		UNESCO			UNESCO
UF	UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION				UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION
					U
					UNESCO
SULPHUR		SULFUR			SULFUR
	UF	SULFUR			U
					SULPHUR
นปฐ.		คณะกรรมการประสานงาน			คณะกรรมการประสานงาน
	UF	คณะกรรมการประสานงาน			โครงการอันเนื่องมาจากพระราชดำริ
		โครงการอันเนื่องมาจากพระราชดำริ			U
					นปฐ.
ธง		ธง			ธง
	UF	ธง			U
					ธง

2. ความสัมพันธ์ในลักษณะลดหลั่นตามลำดับชั้น (Hierarchical relationship) ได้แก่ ความสัมพันธ์ในลำดับชั้นที่สูงกว่าหรือต่ำกว่า แบ่งย่อยเป็น 3 ประเภท คือ

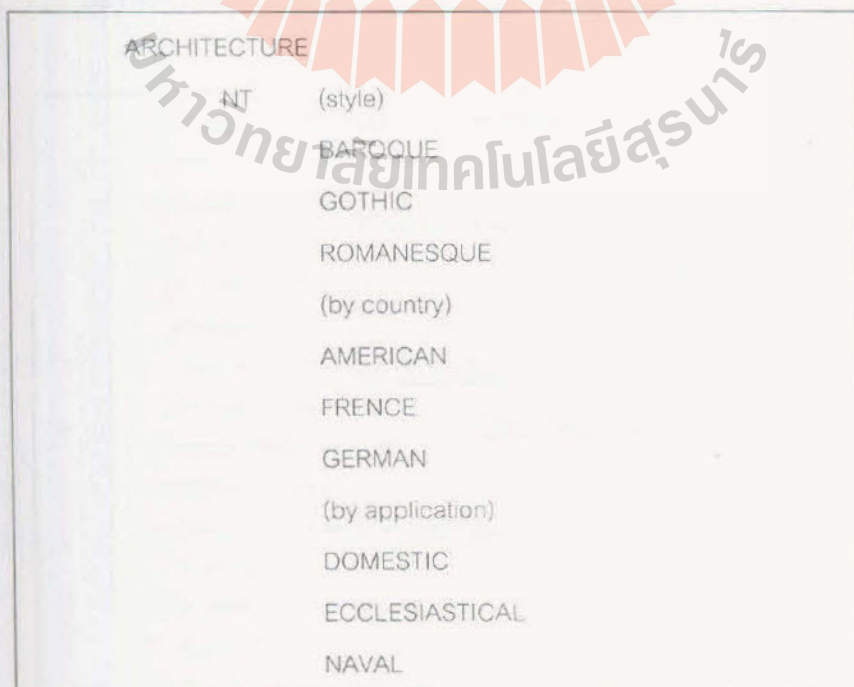
2.1 ความสัมพันธ์ระหว่างสกุล (Genus) กับชนิด (Species) สัญลักษณ์ที่เฉพาะ คือ BTG (Broader term generic) และ NTG (Narrower term generic) หรืออาจใช้ BT และ NT ตามปกติก็ได้

2.2 ความสัมพันธ์ระหว่างองค์รวม (Whole) กับองค์ย่อย (Part) เรียกว่า partitive relationship อาจใช้สัญลักษณ์ที่เฉพาะเป็น BTP (Broader term partitive) และ NTP (Narrower term partitive) ก็ได้ ตัวอย่างเช่น



ข้อสังเกต สมาชิกในกลุ่ม NT ด้วยกันของศัพท์คำหนึ่ง ๆ นั้น เป็นความสัมพันธ์ในระดับ เดียวกัน และเป็นคำที่มีคุณสมบัติบางประการร่วมกัน เช่น พื้น ฝา เพดาน หลังคา ล้วนมีคุณสมบัติเป็นองค์ประกอบของบ้านเช่นเดียวกัน เป็นต้น

2.3 ความสัมพันธ์ระหว่างสิ่งใดสิ่งหนึ่ง (Thing) กับชื่อเฉพาะของสิ่งนั้น (Proper name) เรียกว่า Instance relationship สัญลักษณ์ที่ใช้คือ NT ตัวอย่าง เช่น



3. ความสัมพันธ์ในลักษณะซึ่งเกี่ยวข้องกัน หรือความหมายคาบเกี่ยวกัน (Associative Relationship) เป็นความสัมพันธ์ในแนวระดับ หรือแนวนอน (Horizontal relationship) ซึ่งต่างจากความสัมพันธ์ระหว่าง BT กับ NT ซึ่งเป็นความสัมพันธ์ในแนวตั้ง (Vertical relationship) ความสัมพันธ์ในเชิงนี้ยังไม่มีความแน่นอน หลักกว้าง ๆ มีอยู่ว่าศัพท์คำใดก็ตามที่มีความหมายเกี่ยวข้องกัน หรือ เมื่อผู้สืบค้นนึกถึงศัพท์คำใด ๆ แล้วมักจะนึกไปถึงอีกคำหนึ่งเสมอ และถ้าศัพท์เหล่านั้นไม่ได้มีความสัมพันธ์เชิงลดหลั่นกับศัพท์หลัก ให้ถือว่าเป็นศัพท์สัมพันธ์ในลักษณะเกี่ยวข้องกันได้ กรณีที่จะเกิดขึ้นได้ เช่น

3.1 ความสัมพันธ์ระหว่างสาขาวิชากับวัตถุ หรือสิ่งที่ศึกษา เช่น

FORESTRY
RT FORESTS

3.2 สิ่งใดสิ่งหนึ่งกับสิ่งตรงข้าม หรือแนวคิดที่ตรงข้ามกันแต่สัมพันธ์กัน เช่น

AIRCRAFT
RT ANTI - AIRCRAFT WEAPONS

3.3 สิ่งใดสิ่งหนึ่งกับคุณสมบัติที่เกี่ยวข้องเด่นชัด เช่น

RUBBER
RT ELASTICITY

3.4 วัตถุดิบกับผลผลิตจากวัตถุดิบ เช่น

WOOD
RT FURNITURE

3.5 สิ่งใดสิ่งหนึ่งกับการประยุกต์ หรือนำไปใช้ เช่น

WATER SUPPLY
RT IRRIGATION

3.6 การกระทำกับผลของการกระทำนั้น เช่น

WEAVING
RT CLOTH

3.7 กระบวนการกับเครื่องมือของกระบวนการ เช่น

TEMPERATURE CONTROL
RT THERMOSTATS

3.8 ผลกับเหตุ เช่น

DISASTERS
RT EARTHQUAKES

3.9 กิจกรรมที่เสริมซึ่งกันและกัน เช่น

TEACHING
RT LEARNING

โดยปกติแล้ว ศัพท์ที่อยู่ในกลุ่มเดียวกัน (เช่น กลุ่ม BT ด้วยกัน หรือกลุ่ม NT ด้วยกัน) จะไม่นำมาจัดความสัมพันธ์เป็นแบบ RT อีก นอกจากนี้โดยความเป็นจริงแล้วศัพท์ที่มีความสัมพันธ์กันแบบ RT อาจไม่จำเป็นต้องมีความสัมพันธ์กันแบบกันและกันเสมอไป (คือ X RT Y ไม่จำเป็นว่า Y จะต้องมี X เป็น RT แต่ในทางปฏิบัติ F.W. Lancaster แนะนำให้จัดความสัมพันธ์เป็นแบบซึ่งกันและกันเสมอ อย่างน้อยก็เพื่อไม่ให้เกิดปัญหาในการบำรุงรักษาอิซอร์ส เพราะหากกำหนดให้ X RT Y แล้ว Y ไม่มี X เป็น RT แล้ว หากต่อมา Y ถูกถอนออกจากระบบ จะเกิดปัญหายังคงมีค่า Y อยู่ในรายการโยงความสัมพันธ์ของ X ซึ่งไม่สอดคล้องกับความเป็นจริงเพราะ Y ถูกถอดออกจากระบบแล้ว

การจัดแสดงคำสำคัญในอิซอร์ส

การจัดแสดงคำศัพท์ในอิซอร์ส สามารถทำได้หลายรูปแบบ เช่น

1. แสดงคำศัพท์ตามลำดับอักษร (Alphabetical display)

CINE CAMERAS	
BT	MOVING PICTURE CAMERAS
NT	UNDERWATER CINE CAMERAS
RT	CINEMA
CINEMA	
RT	CINE CAMERAS
DIVING	
RT	UNDERWATER CINE CAMERAS
INSTANT PICTURE CAMERAS	
SN	Cameras which produce a finished print directly
BT	STILL CAMERAS

2. แสดงคำศัพท์ตามลำดับชั้นการลดหลั่นของของคำศัพท์ (Hierarchical display)

- ใช้เครื่องหมาย ". " หรือ "- " หรือ "BT, NT" แสดงลำดับการลดหลั่นของคำศัพท์

COMMUNICATIONS
. POSTAL SERVICES
. . AIR MAIL SERVICES
. . . PARCELS POST
. TELECOMMUNICATION
. . AERONAUTICAL TELECOMMUNICATION
. . . AERONAUTICAL RADIO SERVICES
. . . . RADIO BROADCASTING
. RADIO COMMUNICATIONS
. REMOTE SENSING
. SATELLITE COMMUNICATIONS
. TELEGRAPH
. TELEPHONE
. TELETYPE
. TELEVISION BROADCASTING

LIBRARIES	
NT1	ACADEMIC LIBRARIES
NT2	COLLEGE LIBRARIES
NT3	JUNIOR COLLEGE LIBRARIES
NT2	UNIVERSITY LIBRARIES
NT1	NATIONAL LIBRARIES
NT1	PUBLIC LIBRARIES
NT2	URBAN LIBRARIES
NT2	RURAL LIBRARIES
NT1	SCHOOL LIBRARIES
NT2	PRIMARY SCHOOL LIBRARIES
NT2	SECONDARY SCHOOL LIBRARIES
NT1	SPECIAL LIBRARIES
NT2	ACADEMIC SPECIAL LIBRARIES
NT2	INDUSTRIAL LIBRARIES

3. แสดงคำศัพท์ตามลำดับหมวดหมู่ (Systematic display)

- ต้องมีเครื่องหมายแสดงรายการคำศัพท์เรียงตามลำดับอักษรไว้ด้วยเพื่อเป็นเครื่องมือช่วยค้น

301	OPTICAL EQUIPMENT
302	CAMERAS
	RT: Photography 824
303	MOVING PICTURE CAMERAS
	By medium
304	CINE CAMERAS
	RT: Cinema 895
305	UNDERWATER CINE CAMERAS
306	TELEVISION CAMERAS
	RT: Television 897
307	STEREO CAMERAS
308	STILL CAMERAS
309	INSTANT PICTURE CAMERAS
	SN: Cameras with produce a finished print directly

CAMERAS	302
RT	Photography 824
CINE CAMERAS	304
RT	Cinema 895
CINEMA	895
RT	Cine cameras 304
DIVING	931
RT	Underwater cameras 316

4. การจัดเรียงแบบเปลี่ยนลำดับคำ (Permuted index)

Child Abuse	
Drug Abuse	
Inhalant Abuse	
Academic Achievement	
Academic Achievement Motivation	
Academic Achievement prediction	
Academic Aptitude	
Academic Environment	
Academic Failure	
Academic Over achievement	
Academic Specialization	
Academic Under achievement	
College Academic Achievement	
Acalcula	

ประโยชน์ของอิซอร์ส

1. อิซอร์สเป็นเครื่องมือที่ช่วยให้เกิดความคงเส้นคงวาในการกำหนดค่าแทนสาระให้แก่เอกสาร และทำให้หน่วยงานสามารถกำหนดค่าแทนสาระของเอกสารได้อย่างมีเอกภาพโดยยึดอิซอร์สเป็นหลัก
2. ผู้สืบค้นข้อมูลสามารถเข้าถึงสารสนเทศที่ต้องการได้อย่างสะดวกรวดเร็ว เนื่องจากมีการควบคุมการใช้คำศัพท์เพื่อแทนสาระของเอกสารในเรื่องเดียวกัน และผู้สืบค้นสามารถใช้โครงสร้างความสัมพันธ์ของคำศัพท์เพื่อขยายผลการค้นไปยังเอกสารอื่นๆ ที่เกี่ยวข้องได้
3. เป็นคู่มือในการใช้คำศัพท์ในการเขียนสาระสังเขป เพื่อให้ใช้คำศัพท์ตรงกัน ทั้งนี้เพื่อให้เกิดความเที่ยงตรงในการใช้คอมพิวเตอร์เพื่อกำหนดค่าสำคัญ
4. เกิดมาตรฐานในการกำหนดค่าดรรชนีระหว่างสถาบันซึ่งจะอำนวยความสะดวกในการสืบค้นสารสนเทศแก่ผู้ใช้ เนื่องจากทุกสถาบันใช้มาตรฐานเดียวกันในการกำหนดค่าดรรชนี
5. นักวิเคราะห์หมวดหมู่สามารถใช้อิซอร์สเป็นแนวทางในการวิเคราะห์หมวดหมู่ของเอกสาร
6. สามารถนำไปใช้ในการกำหนดค่าดรรชนีในระบบการจัดเก็บและค้นคืนสารสนเทศทั้งระบบที่สร้างคำค้นด้วยมือและระบบที่ใช้คอมพิวเตอร์ในการกำหนดค่าดรรชนี ได้อย่างมีประสิทธิภาพ

การพัฒนาอิซอร์ส

เพื่อให้อิซอร์ส สามารถทำหน้าที่เป็นคู่มือในการกำหนดค่าดรรชนีได้อย่างมีประสิทธิภาพ อิซอร์สจะต้องมีคุณสมบัติหลายประการ ๆ ได้แก่ ขอบเขตเนื้อหาวิชาที่ครอบคลุมต้องเหมาะสม การจัดความสัมพันธ์ของคำศัพท์จะต้องถูกต้องและมีมาตรฐาน คำศัพท์ควรมีชีวิต และเป็นคลังคำที่เจริญงอกงามได้ นั่นคือ คำศัพท์ต้องได้รับการปรับปรุงให้ทันสมัย ทันต่อการเปลี่ยนแปลงขององค์ความรู้ในแต่ละสาขาวิชาอยู่เสมอ และสามารถ

ขยายออกไปได้เมื่อความรู้ในสาขาวิชามีการแตกแขนงออกไป เนื่องจากวิทยาการ ภาษา และพฤติกรรมของมนุษย์เปลี่ยนแปลงเสมอ

นักพัฒนาอิซอร์สจึงควรมีคุณสมบัติ ดังนี้

1. มีความรู้อย่างลึกซึ้งในสาขาวิชาที่จะสร้างอิซอร์สเนื่องจาก หากขาดความรู้อย่างลึกซึ้งในสาขาวิชาอาจทำให้อิซอร์สขาดความถูกต้อง และทำให้การลำดับความสัมพันธ์ของคำศัพท์ไม่ถูกต้อง
2. มีความรู้ด้านภาษาเนื่องจากให้ภาษาเป็นเครื่องมือในการสื่อสารข้อมูลระหว่างระบบสารสนเทศกับมนุษย์
3. มีความรู้ความเชี่ยวชาญในการสร้างอิซอร์สทั้งด้านทฤษฎีและปฏิบัติ

แนวทางการพัฒนาอิซอร์ส

ในการพัฒนาอิซอร์สต้องใช้การลงทุนสูงทั้งในด้านกำลังความคิด กำลังกาย และกำลังใจ ดังนั้น หลายหน่วยงานจึงไม่ได้พัฒนาอิซอร์สขึ้นมาใช้เอง แต่นำอิซอร์สที่มีอยู่ในสาขาที่เกี่ยวข้องมาปรับใช้ให้เหมาะกับหน่วยงานของตน แนวทางการพัฒนาอิซอร์ส โดยทั่วไปมี 3 วิธีคือ 1). ปรับใช้อิซอร์สที่มีอยู่แล้ว (มีการจัดทำแล้วในสาขาดังกล่าว) ให้เหมาะสมกับการใช้ของหน่วยงาน 2) แปลอิซอร์สที่มีอยู่แล้วแต่เป็นภาษาอื่น และ 3) สร้างอิซอร์สขึ้นมาใหม่ หากไม่มีอิซอร์สในสาขาที่ต้องการอยู่เลย

แนวทางปฏิบัติในการสร้างอิซอร์ส

ในการสร้างอิซอร์ส มีแนวปฏิบัติอยู่ 3 วิธี ซึ่งผู้ปฏิบัติอาจยึดแนวทางใดแนวทางหนึ่งในการปฏิบัติ ดังนี้

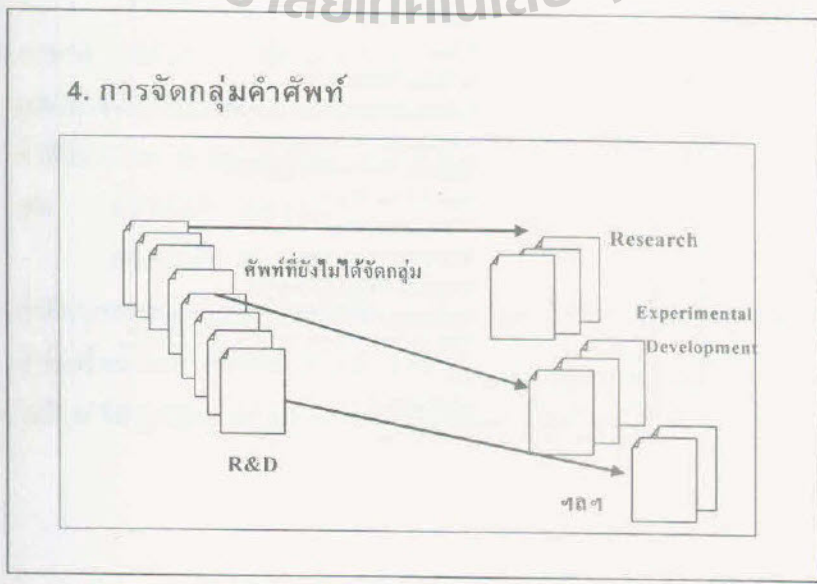
1. ระบบทฤษฎี เป็นระบบที่อาศัยคณะกรรมการ หรือคณะผู้เชี่ยวชาญ จัดสร้างอิซอร์สให้เสร็จสมบูรณ์ระดับหนึ่งแล้วนำมาให้ คณะผู้สร้างจะรวบรวม คิด พิจารณาคำศัพท์ จำแนกคำศัพท์หลัก ๆ ไว้ แล้วกำหนดคำศัพท์ที่เกี่ยวข้องสัมพันธ์กับศัพท์เหล่านั้นจนครบ แต่อาจเกิดปัญหา คือ การจะสร้างให้ครอบคลุมกลุ่มคำศัพท์ทุกคำทำได้ยาก และอาจมีปัญหาในการจัดลำดับความสัมพันธ์แบบสลดหลั่นของคำ เนื่องจากศัพท์บางคำสามารถจำแนก NT ได้หลายกลุ่ม ดังนั้น หากขาดข้อมูลที่ครบถ้วนการจัดกลุ่มอาจขาดหรือเกินได้ ตัวอย่างเช่น

TOY
NT (by materials)
CLOTH TOYS
METAL TOYS
PLASTIC TOYS
RUBBER TOYS
WOODEN TOYS
NT (by age groups)
ADULT TOYS
BABY TOYS
CHILDREN'S TOYS
NT (by locomotion)
CLOCKWORK TOYS
ELECTRIC TOYS
PULL TOYS
PUSH TOYS

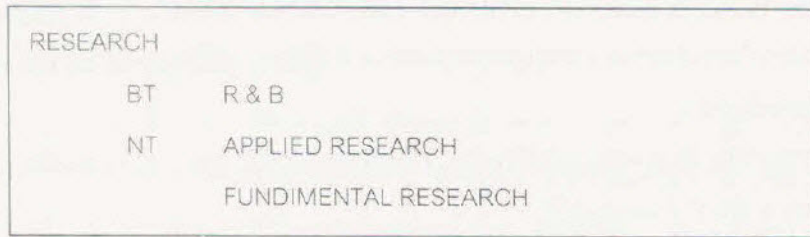
2. ระบบปฏิบัติ เป็นการสร้างอิซอร์สจากพื้นฐานของการปฏิบัติงานจริง คำศัพท์ต่าง ๆ จะได้จาก การวิเคราะห์เนื้อหาของเอกสารและจากการศึกษาผู้ใช้ ตลอดจนแหล่งคำศัพท์อื่น ๆ วิธีนี้มีข้อดีคือ คำศัพท์จะ ตรงกับความต้องการในการใช้งานของระบบสารสนเทศมากกว่าวิธีแรก แต่มีปัญหาคือ ในการกำหนดคำศัพท์จะ ยึดผู้ใช้หรือเอกสารเป็นหลัก
3. ระบบผสม เป็นการผนวกข้อดีของระบบที่ 1 และระบบที่ 2 ได้แก่ การพึ่งพาผู้เชี่ยวชาญโดยทำงาน บนพื้นฐานของความเป็นจริงในทางปฏิบัติ

ขั้นตอนการพัฒนาอิซอร์ส

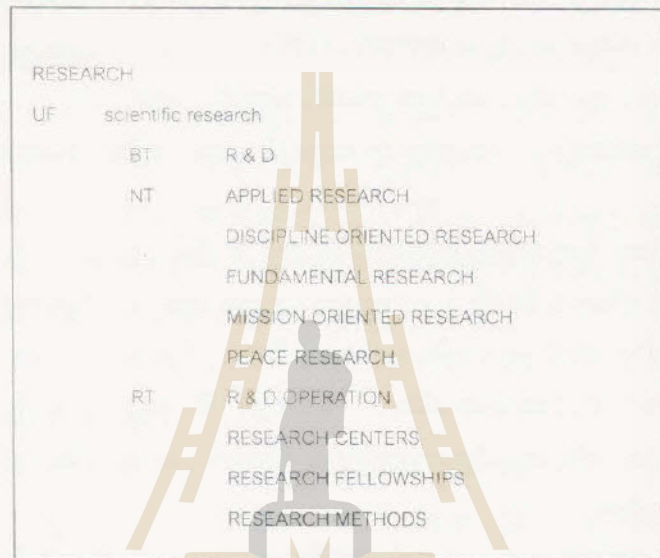
1. วางแผนการดำเนินงาน ครอบคลุมการทำงาน 2 ด้านคือ 1) ด้านการกำหนดเนื้อหาและรูปแบบ ได้แก่ การกำหนดขอบเขตของสาขาวิชาที่จะจัดทำ การกำหนดขอบเขตการใช้ การกำหนด มาตรฐานการควบคุมคำศัพท์ และการกำหนดวิธีการทางไวยากรณ์ และ 2) ด้านการดำเนินงาน ได้แก่ การจัดสรรกำลังคน การแสวงหาผู้ทรงคุณวุฒิ และการเตรียมวัสดุอุปกรณ์เพื่อการดำเนิน งาน เป็นต้น
2. การสะสมคำศัพท์ ซึ่งสามารถสะสมได้จากแหล่งต่าง ๆ ดังนี้ อิซอร์สที่มีการเผยแพร่ในขณะนั้น จากสิ่งพิมพ์ทั่วไปในสาขาวิชานั้น ๆ จากบรรณานุกรมเฉพาะวิชา จากบัญชีศัพท์พรรณานี้ จากบัญชี หัวเรื่อง จากอภิธานศัพท์ และจากคำแนะนำของผู้เชี่ยวชาญในสาขาวิชา โดยการรวบรวมคำศัพท์ จะบันทึกคำศัพท์ลงในบัตรบัตรละหนึ่งคำ หากพบคำที่มีการอธิบายไว้เหมือนกันควรบันทึกไว้ ทันที และควรบันทึกที่มาและบริบทของคำไว้ด้วย เพื่อเพิ่มประสิทธิภาพในการจัดโครงสร้าง ความสัมพันธ์ของคำศัพท์
3. การประมวลคำศัพท์ คือ การรวบรวมคำศัพท์ที่ได้จากการสะสมคำศัพท์เข้าไว้ด้วยกัน อย่างเป็นระบบโดยมีวัตถุประสงค์เพื่อ ขจัดคำซ้ำที่ได้มาจากแหล่งต่าง ๆ และเพื่อรวมคำพ้องความหมาย หรือคำใกล้เคียงกันไว้ด้วยกัน ผลลัพธ์ในขั้นตอนนี้จะได้คำศัพท์ 2 กลุ่มคือ 1) คำศัพท์ที่จะใช้ในการ สร้างอิซอร์ส (active or preferred terms or descriptors) และ 2) คำศัพท์ที่จะไม่ใช้ (non-preferred terms or non-descriptors)
4. การจัดกลุ่มคำศัพท์ ตัวอย่าง เช่น



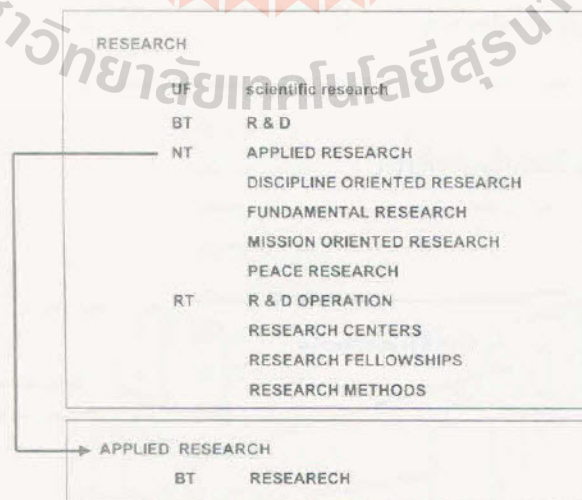
5. การจัดลำดับความสัมพันธ์ของคำศัพท์ในแต่ละกลุ่ม เช่น



6. การปรับ หรือเพิ่มศัพท์ที่เกี่ยวข้องในลักษณะต่าง ๆ



7. การตรวจสอบเพื่อบันทึกคำศัพท์ในลำดับความสัมพันธ์ถัดไปที่ยังไม่ได้อยู่ในฐานะคำหลักลงในบัตรเพิ่ม (added card) เพื่อให้ศัพท์ทุกคำมีโอกาสเรียงอยู่ในตำแหน่งของคำหลัก ทั้งนี้จะต้องบันทึกคำที่มีความสัมพันธ์กับคำศัพท์นั้นลงไปด้วย



8. ปฏิบัติตามขั้นตอนที่ 6-7 จนครบถ้วนทุกคำ เช่น

APPLIED RESEARCH	
BT	RESEARCH
NT	PROCESS RESEARCH
	PRODUCT RESEARCH
RT	EXPERIMENTAL RESEARCH

9. การตรวจตรา ปรับแต่งรายการคำศัพท์แต่ละชุดให้ถูกต้องสมบูรณ์และมีความสอดคล้องกัน ระหว่างคำศัพท์ที่สัมพันธ์กันในลักษณะใดลักษณะหนึ่ง เช่น BT NT เป็นต้น และต้องตรวจสอบ การกำหนดรายการโยงให้ครอบคลุมคำศัพท์ทุกคำที่ผู้ใช้อาจเลือกใช้เป็นคำสืบค้นให้ครบถ้วน ที่สุดเท่าที่จะทำได้ และทำคำอธิบาย (scope note) ให้กับคำศัพท์ทุกคำที่มีความหมายกำกวม คำที่ต้องแนะนำการใช้ หรือคำที่จำกัดขอบเขตหรือบริบทในการใช้
10. ตรวจสอบเรียงลำดับอักษรของคำศัพท์ที่อยู่ภายใต้ BT NT RT และจัดแสดงรายการคำศัพท์ ตามรูปแบบที่กำหนด การแสดงคำศัพท์ควรแสดงความแตกต่างระหว่างคำศัพท์ที่กำหนดให้ใช้ และคำศัพท์ที่ไม่ให้ใช้อย่างชัดเจน เช่น ใช้ ตัวหนา/บาง หรือ พิมพ์เล็ก/ใหญ่ เป็นต้น

วิธีการทางไวยากรณ์

คำศัพท์ ประกอบ ศัพท์แสดงเนื้อหา (subject descriptors) และชื่อเฉพาะ (proper name) อาจรวมไว้ เป็นบัญญัติเดียวกัน หรือแยกบัญญัติชื่อเฉพาะออกไปต่างหาก

- ชื่อเฉพาะ (proper name) ได้แก่ ชื่อบุคคล (personal name) ชื่อนิติบุคคล (corporate name) ชื่อทางภูมิศาสตร์ (geographic name) ชื่อการประชุม ชื่อโครงการ ชื่ออุปกรณ์ ชื่อสูตรเคมี ฯลฯ ใช้มาตรฐาน AACR II เป็นแนวปฏิบัติในการกำหนดคำและการเขียน
- คำศัพท์แสดงเนื้อหา (subject descriptors) ควรมีลักษณะ ดังต่อไปนี้
 - เป็นคำโดดหรือคำผสมก็ได้
 - นิยมใช้คำนาม หลีกเลียงคำกริยา คำคุณศัพท์ คำเชื่อมด้วยบุพบท เช่น
Transfer of technology ควรใช้ TECHNOLOGY TRANSFER
 - ภาษาต่างประเทศนิยมใช้คำนามพหูพจน์ ส่วนคำนามเอกพจน์ใช้เฉพาะกรณีที่เป็นชื่อวิชาเฉพาะ เช่น CHEMISTRY BOTANY
 - คำที่มีความหมายได้หลายอย่าง หรือคำที่สื่อความหมายไม่ชัดเจน ควรมีคำขยายไว้ในวงเล็บ เช่น
MERCURY (METAL)
MERCURY (PLANET)
 - คำที่ห้องความหมายให้คัดเลือกใช้เพียง 1 คำ โดยคำนึงถึงความชัดเจนของคำและผู้ใช้ ส่วนคำที่เหลือทั้งหมดให้ทำรายการโยงไปสู่คำที่ใช้
 - ไม่นิยมใช้อักษรย่อหรือคำย่อ ยกเว้นกรณีอักษรย่อ หรือคำย่อที่เป็นมาตรฐานสากล

2.7 ไม่ควรใส่เครื่องหมายวรรคตอนในคำศัพท์ ยกเว้นกรณีใช้วงเล็บเพื่อขยายความ และกรณีคำภาษาต่างประเทศบางภาษา ส่วนคำอื่น ๆ อาจหลีกเลี่ยงเครื่องหมายโดยวิธีต่อไปนี้

- ใช้การเว้น 1 ระยะแทน เช่น X RAY
- ทิมพ์ติดต่อกันไป เช่น MICROECONOMIC ANALYSIS
MULTIPURPOSE PROJECT

2.8 การสะกด การแปล การถ่ายทออักขร ให้ยึดมาตรฐานที่เป็นที่ยอมรับทั่วไป

2.9 ไม่นิยมใช้การกลับคำ (inverted entry) เช่น

ใช้ HIGHER EDUCATION ไม่ใช้ EDUCATION, HIGHER

2.10 กรณีชื่อหรือหลายภาษาควรเคารพกฎเกณฑ์ของภาษานั้น ๆ

2.11 ไม่ควรเชื่อมคำศัพท์ไว้ล่วงหน้า (Pre-coordinate terms) โดยเฉพาะการเชื่อมคำด้วยบุพบท "และ" และ "กับ" เนื่องจากสามารถใช้ตรรกะบูลีนเชื่อมคำในภายหลังได้

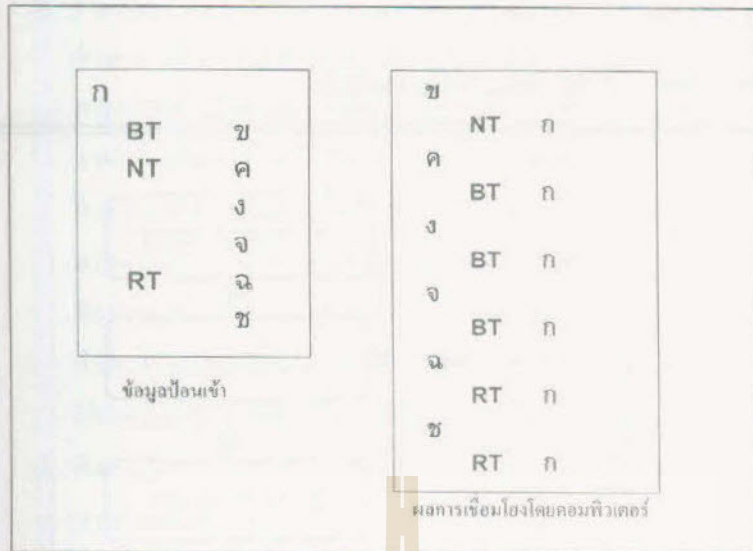
การประเมินอิซอร์ส

การประเมินอิซอร์สมิ 2 ระดับ คือ

1. การประเมินเนื้อหาและรูปแบบของอิซอร์สที่สร้างขึ้น มีสิ่งที่จะต้องประเมินอย่างน้อย 3 ประการ คือ
 - 1.1 ขอบเขตของอิซอร์ส ได้แก่ การครอบคลุมคำศัพท์ในสาขาวิชาได้อย่างครบถ้วน เพียงพอต่อการจัดทำบรรณนิและการค้นคืน
 - 1.2 ลักษณะการจัดความสัมพันธ์เชิงโครงสร้างของคำศัพท์ (BT NT RT) ถูกต้อง เหมาะสม
 - 1.3 ลักษณะการจัดแสดงคำศัพท์ ช่วยให้ผู้ใช้เข้าใจและใช้ได้ง่ายหรือรวดเร็วซึ่งการประเมินในระดับที่ 1 สามารถให้ผู้เชี่ยวชาญในสาขาวิชา ผู้ใช้ และผู้เชี่ยวชาญในการพัฒนาอิซอร์สเป็นผู้ประเมินได้
2. การประเมินในฐานะเป็นปัจจัยของระบบการจัดเก็บและค้นคืนสารสนเทศ การประเมินในระดับนี้จะอาศัยการวัดผลการเรียกคืน (recall) และความตรงต่อความต้องการจำเพาะ (precision or relevant) ซึ่งจะได้กล่าวถึงรายละเอียดในบทของการประเมินระบบการจัดเก็บและค้นคืนสารสนเทศ

บทบาทของคอมพิวเตอร์ในการพัฒนาอิซอร์ส

1. ระยะการสะสมคำ จะใช้คอมพิวเตอร์ในการอ่านและการคัดเลือกคำศัพท์ออกมาประมวลได้เหมือนการใช้คอมพิวเตอร์ในการทำบรรณนิภาษาธรรมชาติ
2. ระยะการสร้างอิซอร์ส จะใช้คอมพิวเตอร์ใน
 - การบันทึกและป้อนคำศัพท์เข้าสู่ระบบ
 - เชื่อมโยงความสัมพันธ์และสะท้อนกลับอัตโนมัติ เมื่อศัพท์ชุดหนึ่ง ๆ ได้รับการป้อนเข้าสู่ระบบ ดังตัวอย่าง



- ตรวจสอบ เพิ่ม ลด หรือเปลี่ยนแปลงแก้ไขคำศัพท์ในระบบ
- เรียงคำศัพท์ตามลำดับอักษร ได้แก่ 1) เรียงตามลำดับอักษรของคำศัพท์ภายใต้ความสัมพันธ์ BT NT RT และ 2) เรียงลำดับอักษรของคำทั้งบัญชี
- ค้นคืนคำศัพท์และรายการที่ได้จากการค้นคืน
- สร้างดรรชนีช่วยค้นรูปแบบต่าง ๆ เช่น ดรรชนีเวียนคำ (permuted index)
- พิมพ์ออร์วัล ได้แก่ การพิมพ์ระหว่างปฏิบัติการและการพิมพ์เป็นรูปเล่ม
- ประมวลผลและแสดงสถิติที่จำเป็น



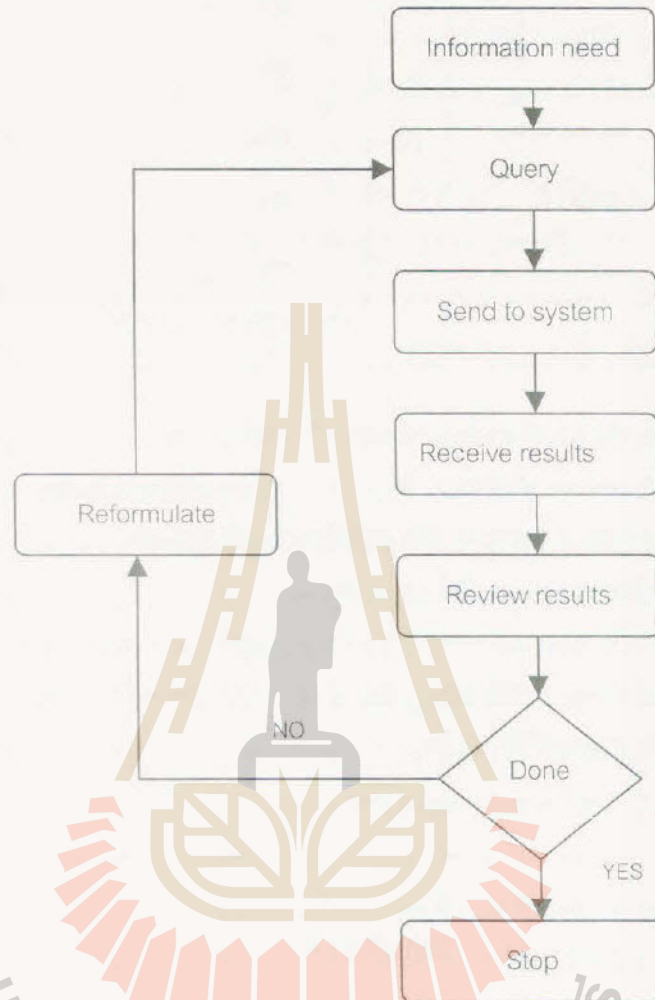
เรียบเรียงจาก

นฤมล ปราชญ์โยธิน, ศึกษาศาสตร์ ภาควิชาเทคโนโลยีสารสนเทศ, เปรมิน จินดาวิมลเลิศ, 2536. ออร์วัลกับระบบสารสนเทศ.

กรุงเทพฯ : ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

บทที่ 6

การค้นคืนสารสนเทศ (Information Retrieval)



The standard retrieval interaction model

กระบวนการค้นคืนสารสนเทศ

การค้นคืนสารสนเทศประกอบด้วย 2 กระบวนการหลัก คือ

1. Question Asking หรือ การตั้งคำถาม จะเกิดขึ้นเมื่อผู้ใช้เกิดความต้องการสารสนเทศ ซึ่งความต้องการสารสนเทศจะเกิดขึ้นเมื่อผู้ใช้รู้สึกขาดความรู้ในเรื่องใดเรื่องหนึ่ง แล้วตระหนักว่าความรู้ที่มีอยู่ไม่เพียงพอที่จะใช้ในการแก้ปัญหา และต้องการค้นหาคำตอบเพื่อนำมาแก้ไขปัญหานั้น

กระบวนการในการวิเคราะห์ปัญหาของผู้ใช้

- 1) Information problem หรือ การรู้ปัญหา เกิดขึ้นเมื่อผู้ใช้รู้สึกขาดความรู้ในเรื่องใดเรื่องหนึ่ง
- 2) Information need หรือ การรู้สึกต้องการคำตอบ เกิดขึ้นเมื่อต้องการนำคำตอบนั้นไปใช้

- 3) Question หรือ การตั้งคำถาม ได้แก่ การถ่ายทอดความคิด หรือปัญหาออกมาเป็นคำพูด เป็นประโยค หรือภาษาเขียน ซึ่งจะเป็นจุดเริ่มต้นของกระบวนการค้นคืนสารสนเทศในฝั่งของระบบจัดเก็บและค้นคืน อย่างไรก็ตาม คำถามที่เกิดขึ้นในขั้นนี้อาจไม่ใช่ตัวแทนที่แท้จริงของปัญหาสารสนเทศเสมอไป ทั้งนี้เนื่องจากผู้ใช้อาจกำหนดคำถามกว้าง หรือแคบเกินไป และอาจนึกถึงเพียงประเด็นข้างเคียงไม่ใช่ประเด็นหลักของคำถาม เนื่องจากยังไม่ทราบความต้องการที่แท้จริงของตนเอง หรือทราบความต้องการแต่ไม่สามารถถ่ายทอดออกมาเป็นคำพูด เนื่องจากยังขาดความรู้ในเรื่องที่กำลังค้นหาคำตอบ จึงยังไม่สามารถถ่ายทอดออกมาเป็นคำพูด หรือประโยคที่ชัดเจนได้ หรือใช้ศัพท์เทคนิค หรือศัพท์เฉพาะที่ผู้อื่นไม่เข้าใจ หรือไม่ให้ความสำคัญกับการตั้งคำถาม
- 4) Request or Query หรือ การกำหนดคำถามเพื่อป้อนเข้าไปในระบบการค้นคืนสารสนเทศเพื่อค้นหาสารสนเทศที่ต้องการ

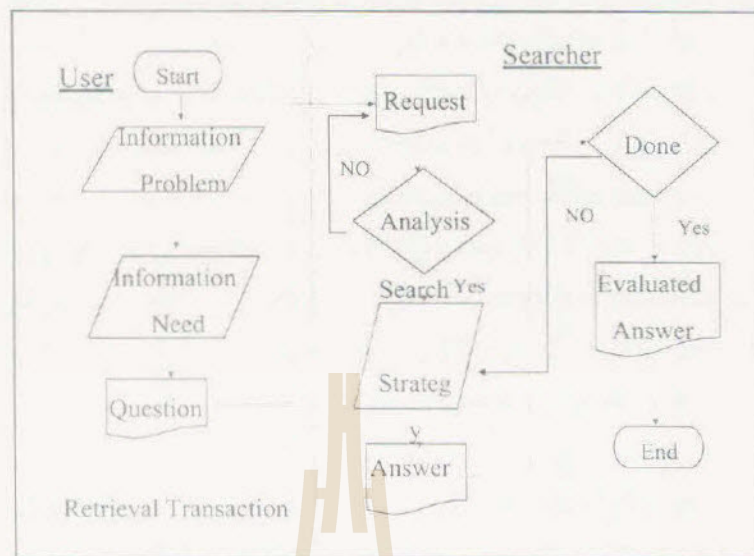
2. Question Answering หรือ การหาคำตอบ ทำได้ 2 วิธี คือ

- 1) โดยการที่ผู้ใช้เข้ามาที่ระบบการจัดเก็บและค้นคืนสารสนเทศเพื่อค้นหาคำตอบด้วยตนเอง กรณีนี้ผู้ใช้ต้องสามารถเข้าใจความต้องการที่แท้จริงของตนเอง
- 2) โดยการติดต่อให้ผู้ให้บริการสารสนเทศเป็นผู้ค้นหาคำตอบให้ กรณีนี้ ผู้ใช้และผู้ให้บริการสารสนเทศจะร่วมกันวิเคราะห์ความต้องการสารสนเทศของผู้ใช้ โดยผู้ให้บริการสารสนเทศจะต้องใช้ทักษะในการสัมภาษณ์และการสื่อสารระหว่างบุคคลเพื่อให้ได้ความต้องการที่แท้จริงของผู้ใช้ออกมา ในขณะที่ผู้จะเป็นฝ่ายถ่ายทอดความต้องการให้กับผู้ให้บริการรับทราบ รวมทั้งแจ้งคำศัพท์เฉพาะ และอธิบายคำศัพท์ให้ผู้ให้บริการเกิดความเข้าใจในขอบเขตเนื้อหาของคำศัพท์และประเด็นของเรื่องที่ต้องการค้นหาคำตอบ

อาจสรุปขั้นตอนการค้นคืนสารสนเทศ ได้ดังนี้

1. ผู้ใช้ติดต่อกับระบบจัดเก็บและค้นคืนสารสนเทศ หรือผู้ให้บริการค้นคืนสารสนเทศ
2. ผู้ใช้วิเคราะห์คำถามหรือความต้องการสารสนเทศของตนเองออกมาเป็นคำพูด เพื่อถ่ายทอดแก่ผู้ค้น
3. ผู้ค้นวิเคราะห์คำถามของผู้ใช้เพื่อทราบความต้องการที่แท้จริง
4. ซักถามผู้ใช้เกี่ยวกับประเด็นหลัก และคำศัพท์ของเรื่องที่ต้องการค้น เพื่อความเข้าใจที่ตรงกัน และยืนยันความต้องการของผู้ใช้
5. ยืนยันคำถามครั้งสุดท้ายก่อนทำการค้น
6. เลือกระบบฐานข้อมูลที่เกี่ยวข้องกับเรื่องที่ต้องการคำตอบ
7. ผู้ค้นแปลงแนวคิดและกำหนดความสัมพันธ์ระหว่างแนวคิด
8. สร้างสูตรการค้นโดยใช้คำค้นตามรูปแบบของระบบฐานข้อมูลที่ใช้
9. วางแผนกลยุทธ์ในการค้น หรือการกำหนดทางเลือกอื่น หรือสูตรการค้นอื่น หากการค้นในครั้งแรกล้มเหลว
10. ดำเนินการค้นคืนสารสนเทศ

11. ประเมินผลการค้นตามวัตถุประสงค์ของการค้น จนกว่าจะได้ผลการค้นที่พอใจ
12. ส่งผลการค้นให้กับผู้ใช้



ปัจจัยที่มีผลต่อความสำเร็จของการค้นหาคำตอบ

1. กระบวนการของการวิเคราะห์คำถาม ได้แก่

- พฤติกรรมการค้นคืนสารสนเทศของผู้ใช้
- การสัมภาษณ์ผู้ใช้

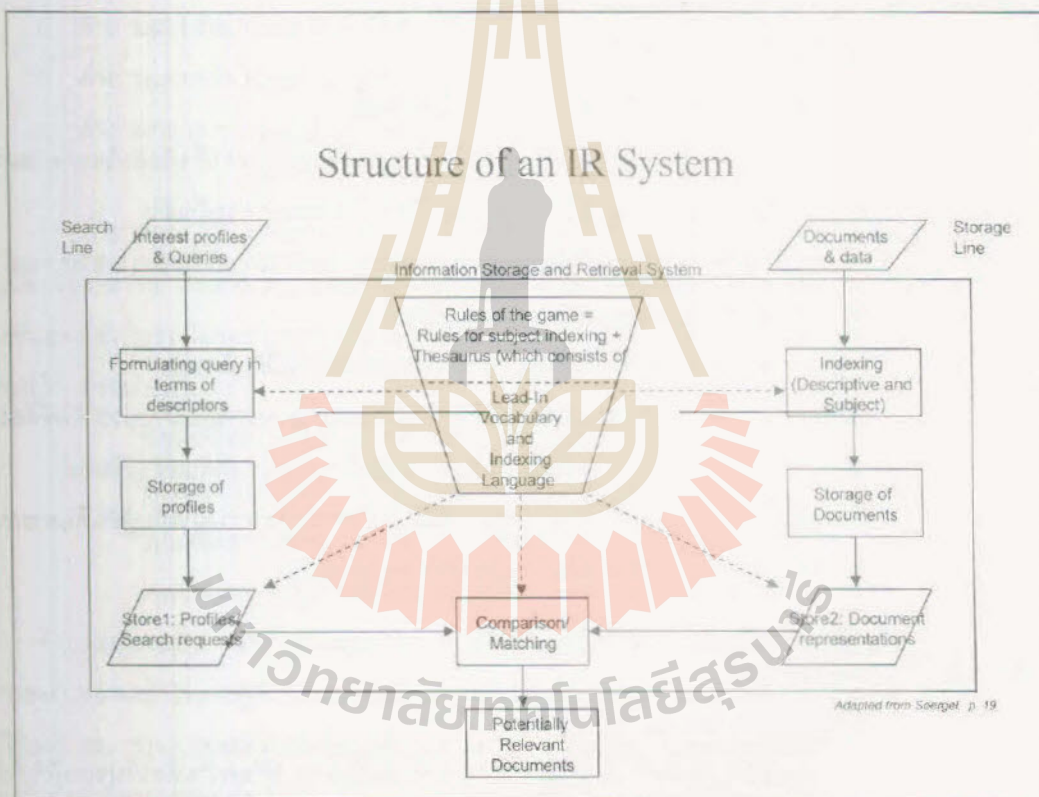
กระบวนการของการวิเคราะห์คำถามมีผลอย่างยิ่งต่อความสำเร็จหรือความล้มเหลวในการค้นคืนสารสนเทศ เนื่องจากเป็นจุดเริ่มต้นของการกำหนดขอบเขตของเรื่องที่ต้องการหาคำตอบ และวัตถุประสงค์ในการค้นหรือการนำไปใช้ ซึ่งหากวิเคราะห์ผิดพลาดจะส่งผลกระทบต่อกระบวนการทั้งหมด เพราะแม้จะดำเนินการค้นได้ถูกต้อง คือได้เอกสารที่ตรงกับคำค้นออกมาเป็นจำนวนที่สมควรนำไปใช้ได้ แต่ผลการค้นไม่ตรงกับความต้องการที่แท้จริงของผู้ใช้เนื่องจากวิเคราะห์ความต้องการผิด ทำให้กำหนดคำค้นผิด และสร้างสูตรการค้นผิด การค้นในครั้งนั้นก็ถือได้ว่าล้มเหลวเช่นกัน ดังนั้น เพื่อให้ได้คำตอบที่ตรงกับความต้องการ จึงควรให้ความสำคัญกับการวิเคราะห์ความต้องการสารสนเทศของผู้ใช้ ซึ่งทำได้โดยการสัมภาษณ์และการศึกษาพฤติกรรมของผู้ใช้ประกอบกัน

โดยทั่วไปการวิเคราะห์คำถามมีวัตถุประสงค์เพื่อ ทราบข้อมูลเบื้องต้นเกี่ยวกับสารสนเทศที่ผู้ใช้ต้องการ เช่น ชื่อผู้แต่ง ประเด็นเรื่องที่ต้องการค้น วัตถุประสงค์ของการค้น เพื่อกำหนดลักษณะของเอกสาร เช่น ช่วงเวลา ภาษา ประเภทของเอกสาร ฯลฯ เพื่อกำหนดคำตรรกะและสร้างสูตรการค้น

2. การออกแบบระบบการจัดเก็บ ได้แก่

- การกำหนดโครงสร้างการจัดเก็บสารสนเทศ ได้แก่ การออกแบบโครงสร้างระเบียบในการจัดเก็บข้อมูลของเอกสารแต่ละรายการ ควรประกอบด้วยเขตข้อมูลใดบ้าง และเขตข้อมูลใดที่ต้องกำหนดให้เป็นคำค้น

- การกำหนดตัวแทนของเอกสาร ได้แก่ การจัดโครงสร้างหมวดหมู่หรือจัดกลุ่มของเอกสารและการกำหนดชื่อที่ชัดเจนให้กับแต่ละหมวดหมู่ และการกำหนดคำค้นให้กับเอกสารแต่ละรายการ โดยการกำหนดเป็นภาษาธรรมชาติ และ/หรือ ภาษาควบคุม
3. การออกแบบระบบการเชื่อมประสานกับผู้ใช้ ได้แก่
- การควบคุมการทำงานของระบบ เช่น การเรียกใช้งานฟังก์ชันต่าง ๆ การพิมพ์และส่งคำค้นเข้าสู่ระบบ การเรียกใช้เมนูต่าง ๆ เป็นต้น
 - การออกแบบหน้าจอ ทั้งหน้าจอการค้นและหน้าจอแสดงผลแบบต่าง ๆ
 - การค้นคืน ได้แก่ การออกแบบฟังก์ชันต่าง ๆ เพื่อช่วยในการค้นคืนสารสนเทศ เช่น ตรรกะบูลีน การตัดคำ การค้นจากเขตข้อมูลต่าง ๆ เป็นต้น
 - การช่วยเหลือในการใช้งานระบบ เช่น หน้าจอให้ความช่วยเหลือ ข้อความแสดงความผิดพลาด ข้อความให้คำแนะนำช่วยเหลือ เป็นต้น



การสัมภาษณ์เพื่อการค้นคืนสารสนเทศ (Reference Interview)

Users has the best understanding of what information is needed. And you, as the search specialist, have the best understanding of where the information is.

การสัมภาษณ์เพื่อการค้นคืนสารสนเทศ เป็นการสนทนาระหว่างผู้ใช้และผู้ที่จะทำหน้าที่ค้นคืนสารสนเทศ ซึ่งเป็นผู้ที่เชี่ยวชาญการค้น (Search specialist) เพื่อทำความเข้าใจที่ชัดเจนเกี่ยวกับความต้องการสารสนเทศให้หรือคำถามของผู้ใช้ เป็นกระบวนการสื่อสารระหว่างบุคคล โดยที่บุคคลหนึ่งบรรยายให้อีกบุคคลหนึ่งรับทราบปัญหาหรือความต้องการ ซึ่งอาจเป็นเรื่องที่บุคคลผู้รับฟังอาจไม่มีความรู้ในเรื่องนั้นเลยก็ได้ ดังนั้นการสัมภาษณ์ผู้ใช้จึงเป็นศิลปะในการสนทนาอย่างหนึ่งเพื่อให้ได้ข้อมูลที่จะเป็นประโยชน์ต่อขั้นตอนต่อไปของกระบวนการค้นคืนสารสนเทศ

วัตถุประสงค์ของการสัมภาษณ์เพื่อค้นคืนสารสนเทศ

การสัมภาษณ์ผู้ใช้ มีวัตถุประสงค์เพื่อ

- ทราบขอบเขตของเรื่องที่ต้องการคำตอบ กว้างหรือลึก
- ทราบวัตถุประสงค์ของการนำสารสนเทศไปใช้ เช่น เพื่อแก้ปัญหา เพื่อตัดสินใจ เพื่อติดตามความทันสมัย เพื่อขยายความรู้ ทำรายงาน หรือทำวิจัยเฉพาะทาง เป็นต้น
- ทราบรูปแบบของสารสนเทศที่ต้องการ เช่น บทความ รายงานการวิจัย ข่าว บทความวิเคราะห์ ฯลฯ
- ทราบความทันสมัย หรือช่วงเวลาของข้อมูลที่ต้องการ
- ภูมิหลังของผู้ใช้ เช่น ระดับการศึกษา อาชีพ ตำแหน่งหน้าที่ที่มีรับผิดชอบ
- พฤติกรรมการค้นหาสารสนเทศ และประวัติการค้นที่ผ่านมาเกี่ยวกับประเด็นที่ต้องการคำตอบ เช่น แหล่งสารสนเทศที่ใช้เป็นประจำ แหล่งข้อมูลใดบ้างที่ค้นแล้ว เป็นต้น
- คำศัพท์ที่ใช้ในการค้นคืน เพื่อยืนยันความเข้าใจที่ตรงกันระหว่างผู้ใช้กับผู้ค้นคืนสารสนเทศ และทราบเรื่องที่เกี่ยวข้องสัมพันธ์กับเรื่องที่ต้องการคำตอบ

ข้อควรปฏิบัติในการสัมภาษณ์

การสัมภาษณ์ควรทำในลักษณะ Active Listening ดังนี้

- ให้ความสำคัญกับสิ่งที่ผู้ใช้พูด โดยพยายามเข้าใจในมุมมอง หรือความคิดเห็นของผู้ใช้เกี่ยวกับเรื่องที่ผู้ใช้สนใจ
- ให้ความสำคัญและสนใจกับผู้ใช้ ไม่ขัดจังหวะในการพูด หรือไม่พยายามเปลี่ยนเรื่อง
- ตั้งคำถามกลับเมื่อไม่แน่ใจ เช่น อาจให้ผู้ใช้ยกตัวอย่างที่เป็นรูปธรรม หรือการนำไปใช้ที่ชัดเจน
- ถามกลับผู้ให้ด้วยคำศัพท์ที่ผู้ใช้ถามเพื่อยืนยันความถูกต้อง
- พยายามทำความเข้าใจกับคำถามก่อนนำเสนอทางเลือกแก่ผู้ใช้
- ถามผู้ใช้ด้วยประโยคที่สื่อความหมายชัดเจน หลีกเลี่ยงการใช้ศัพท์วิชาการด้านสารสนเทศ
- รอคำตอบจากผู้ใช้ อาจโดยใช้คำพูด หรือ ท่าทาง

- แสดงท่าทีสนใจคำถามและแสดงท่าทางเต็มใจให้ความช่วยเหลือ
- แสดงท่าทีให้ผู้ใช้นั้นเห็นว่าผู้สัมภาษณ์เข้าใจในสิ่งที่ผู้ใช้พูด
- ไม่ใช้อารมณ์และความคิดเห็นส่วนตัว ของผู้สัมภาษณ์
- เคารพในสิทธิของผู้ใช้ โดยไม่เปิดเผยข้อมูลของผู้ใช้แก่บุคคลอื่น
- เคารพในคำถามของผู้ใช้
- ควรเริ่มสัมภาษณ์ด้วยคำถามแบบเปิด จากนั้นอาจใช้คำถามแบบปิดเพื่อจำกัดขอบเขตเฉพาะเรื่องที่ใช้สนใจ
- หรืออาจเริ่มต้นการสัมภาษณ์ด้วยคำถามกลาง ๆ กรณีที่ผู้ใช้พอจะทราบความต้องการที่แท้จริงของตนเอง

ประเภทของคำถาม

คำถามที่ใช้ในการสัมภาษณ์เพื่อการค้นคืนสารสนเทศจำแนกได้เป็น 3 ประเภท คือ

1. คำถามแบบปิด (Close question)
2. คำถามแบบเปิด (Open question)
3. คำถามกลาง ๆ (Neutral question)

1. คำถามแบบปิด (Close questions) ได้แก่ คำถามประเภทที่ต้องการคำตอบ ใช่ หรือไม่ใช่ สิ่งนี้หรือสิ่งนั้น Yes or No, That or this เป็นคำถามที่กำหนดคำตอบไว้ในตัวเพื่อให้ผู้ใช้เลือกตอบ หรือเป็นลักษณะคำถามที่แสดงให้เห็นว่า ผู้ถามพยายามจับคู่ความต้องการของผู้ใช้ตามสารสนเทศที่มีอยู่ เป็นคำถามที่จำกัดการตอบของผู้ใช้ ตัวอย่างเช่น

- "ข้อมูลนี้จะนำไปใช้ในการทำโครงการใช่หรือไม่"
- "คุณต้องการงานของนักเขียนอเมริกาหรือแคนาดาคะ"
- "คุณต้องการบทความจากวารสารหรืออุตสาหกรรมคะ"
- "คุณต้องการข้อมูลเกี่ยวกับโครงการอวกาศของอเมริกาใช่หรือไม่"

คำถามปลายเปิด เหมาะสำหรับการถามก่อนปิดการสัมภาษณ์ เพื่อเป็นการยืนยันความต้องการของผู้ใช้ว่าตรงกับความเข้าใจของผู้สัมภาษณ์ใช่หรือไม่ หลังจากที่ผ่านมาขั้นตอนการสัมภาษณ์จนทราบความต้องการของผู้ใช้แล้ว เช่น "สรุปว่าคุณต้องการข้อมูลที่เป็นรายงานการวิจัยเกี่ยวกับการปลูกถั่วไร่ค่างใช่หรือไม่" เพื่อเป็นการยืนยันความต้องการก่อนเริ่มต้นการค้น คำถามประเภทนี้ไม่เหมาะที่จะใช้ในการเริ่มต้นการสัมภาษณ์เพราะจะทำให้ต้องใช้เวลามากในการสัมภาษณ์ผู้ใช้ เนื่องจากเป็นการคาดเดาความต้องการของผู้ใช้โดยไม่เปิดโอกาสให้ผู้ใช้ได้ออกถึงความต้องการ

2. คำถามแบบเปิด (Open questions) ได้แก่ คำถามประเภท Who What When Where Why How

ซึ่งเป็นลักษณะคำถามที่เปิดโอกาสให้ผู้ใช้ได้แสดงความคิดเห็นอย่างเปิดเผย โดยใช้คำพูดของตนเอง เป็นคำถามที่ชวนให้ตอบ ไม่มีการจำกัดคำตอบไว้ล่วงหน้า ตัวอย่างเช่น

"ขอข้อมูลเพิ่มเติมเกี่ยวกับโครงการอวกาศของอเมริกาอีกนิดได้นะคะ"

? ประเด็นใดของโครงการ

"คุณต้องการข้อมูลประเภทไหนคะ"

"คุณต้องการข้อมูลอะไรคะ"

"คุณต้องการข้อมูลเกี่ยวกับอะไรคะ"

? ตอบได้หลายแบบ เช่น รูปแบบการเขียน สื่อที่ใช้บันทึก เนื้อหาของข้อมูล ฯลฯ

เนื่องจากคำถามแบบเปิด เป็นคำถามที่เปิดโอกาสให้ผู้ผู้ใช้ได้แสดงความคิดเห็นอย่างเปิดกว้างเกี่ยวกับสารสนเทศที่ต้องการโดยไม่จำกัดคำตอบ จึงเหมาะใช้เป็นคำถามเพื่อเริ่มการสนทนา หรือเริ่มการสัมภาษณ์ เพื่อชักชวนให้ผู้ผู้ใช้แสดงความคิดเห็น หรือให้ข้อมูลเกี่ยวกับสารสนเทศที่ต้องการ อย่างไรก็ตาม คำถามประเภทนี้อาจนำมาทั้งคำตอบที่เกี่ยวข้องและไม่เกี่ยวข้องกันเรื่องที่ผู้ใช้กำลังหาคำตอบ ผู้ใช้จะหลุดออกนอกประเด็นได้ง่าย เนื่องจากไม่มีการติดกรอบของคำถาม เช่น "มีอะไรให้ช่วยหรือเปล่าคะ" "ต้องการข้อมูลอะไรคะ" ซึ่งคำตอบที่เป็นไปได้มีหลายคำตอบ และอาจไม่ใช่คำตอบที่จะนำมาใช้ในการค้นคืนสารสนเทศ

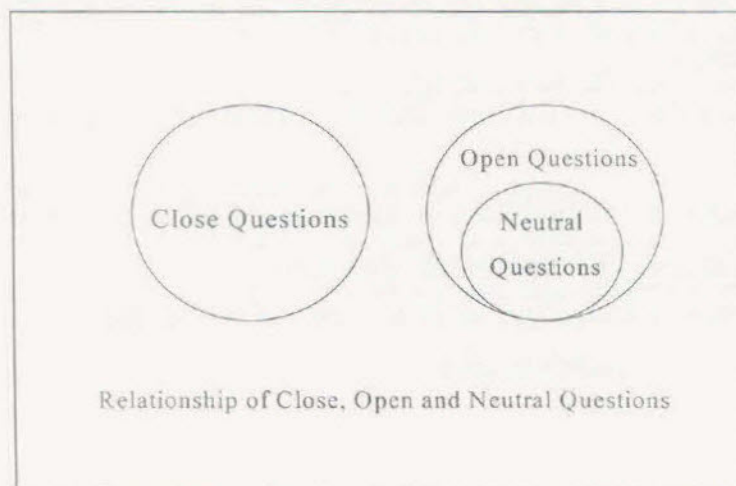
3. คำถามกลาง ๆ (Neutral questions) เป็นลักษณะของคำถามแบบเปิด แต่เปิดอยู่ในแนวของประเด็นของเรื่องที่ใช้ต้องการจะค้นหาคำตอบ โดยใช้คำศัพท์ที่เกี่ยวข้องเพื่อชักชวนให้ผู้ผู้ใช้พูดในประเด็นที่เฉพาะเจาะจงขึ้น รวมทั้งวัตถุประสงค์ของการนำไปใช้ เป็นการเปิดโอกาสให้ผู้ผู้ใช้ได้ตั้งคำถามในมุมมองตามความคิดของผู้ใช้ที่มีต่อเรื่องที่สนใจ เป็นคำถามที่เกิดขึ้นระหว่างการสัมภาษณ์ เพื่อค้นหาความต้องการที่แท้จริงของผู้ใช้ คำถามแบบกลาง ๆ จัดเป็นเซตย่อย (subset) ของคำถามแบบเปิด ตัวอย่างเช่น

คำถามเพื่อเข้าใจสถานการณ์

"ช่วยเล่าให้ฟังนิดนึงนะคะว่าปัญหาของเรื่องที่กำลังศึกษาเกิดขึ้นได้อย่างไร มีที่มาที่ไปอย่างไร
"จะต้องทำอย่างไรจึงจะแก้ปัญหานี้ได้" หรือ "จะต้องใช้วิธีการศึกษารูปแบบใดจึงจะหาคำตอบ
"คุณมีวิธีการศึกษาอย่างไร" หรือ "คุณใช้การทดลองวิธีใด" หรือ "คุณเก็บข้อมูลอย่างไรคะ"

คำถามเพื่อปิดช่องว่าง

"คุณต้องการทราบข้อมูลประเด็นใดของโครงการอวกาศของสหรัฐอเมริกาคะ"
"คุณกำลังศึกษาประเด็นใดของการปลูกพืชไร้ดินคะ"
"ข้อมูลอะไรบ้างคะที่คุณยังได้ไม่ครบเกี่ยวกับการป้องกันโรคเอดส์"
"คุณกำลังพยายามทำความเข้าใจในประเด็นใดคะ"
"คุณคิดว่าข้อมูลใดที่คุณยังได้ไม่ครบ ทำให้ยังเข้าใจเรื่องการ GMO ไม่ชัดเจนคะ"



ตัวอย่างคำถามทั้ง 3 ประเภท

ผู้ใช้ : คุณมีข้อมูลเกี่ยวกับบริษัทขนาดใหญ่หรือเปล่าครับ

คำถามชุดที่ 1

คำถามแบบปิด : คุณต้องการเป็นรายงานประจำปีใช่หรือไม่

คำถามแบบเปิด : คุณต้องการทรัพยากรประเภทใด

คำถามแบบกลาง : ช่วยกรุณาบอกประเด็นที่คุณกำลังศึกษาด้วยคะ เพื่อดิฉันจะสามารถหา
ข้อมูลที่ตรงกับประเด็นที่คุณกำลังศึกษามากที่สุด

คำถามชุดที่ 2

คำถามแบบปิด : ต้องการข้อมูลของบริษัทในประเทศหรือบริษัทระดับนานาชาติคะ

คำถามแบบเปิด : บริษัทขนาดใหญ่นี้ใหญ่แค่ไหนคะ

คำถามแบบกลาง : คุณช่วยบอกได้หรือไม่ว่าข้อมูลนี้จะเอาไปทำอะไรคะ

การค้นหาคำตอบ

การค้นหาคำตอบโดยผู้ให้บริการสารสนเทศ สิ่งที่ต้องแจ้งให้ผู้ใช้ทราบ ได้แก่

- กำหนดเวลาที่จะได้รับสารสนเทศ
- สถานที่ หรือวิธีจัดส่งสารสนเทศให้กับผู้ใช้
- รูปแบบของสารสนเทศที่จะได้รับ
- ค่าใช้จ่าย ทั้งค่าดำเนินการและค่าบริการ
- ส่งให้ผู้ใช้ประเมินผลการค้น
- ชี้แจงให้ผู้ใช้ทราบหากไม่พบข้อมูลที่ต้องการ
- แนะนำแหล่งที่คาดว่าจะมีสารสนเทศดังกล่าว

การค้นหาคำตอบที่ดำเนินการโดยผู้ใช้สารสนเทศร่วมกับผู้ใช้

- ให้คำแนะนำในการใช้เครื่องมือช่วยค้นต่าง ๆ
- หลีกเลี่ยงการใช้ศัพท์ด้านทฤษฎีและคำย่อ
- อธิบายกลวิธีการค้น
- อำนวยความสะดวกในการรับผลการค้น

การทำแบบสัมภาษณ์

ลักษณะของแบบสัมภาษณ์ที่ดี

- สั้น กระชับ
- สามารถเก็บข้อมูลที่ต้องการได้ทั้งหมด
- ควรจบในหนึ่งหน้า A 4 เพื่อประหยัดเวลาในการกรอกข้อมูล และป้องกันการลืมตอบคำถาม
บางข้อที่อยู่ด้านหลัง

คำถามในแบบสัมภาษณ์

1. ข้อมูลส่วนบุคคล ได้แก่ ชื่อ-สกุล อาชีพ ตำแหน่ง สถานที่ติดต่อ หมายเลขโทรศัพท์ โทรสาร E-mail
2. วัตถุประสงค์การใช้สารสนเทศ อาทิ เพื่อ การวิจัย การศึกษา การทำธุรกิจ
3. วิธีการจัดส่งข้อมูล เช่น มารับเอง ส่งทางไปรษณีย์ E-mail
4. ช่วงเวลาที่ต้องการรับสารสนเทศ
5. ลักษณะของสารสนเทศที่ต้องการ อาทิ สารระสังเขป บรรณานุกรม บรรณานุกรมและสารระสังเขป หรือเอกสารฉบับเต็ม เป็นต้น
6. ประเภททรัพยากรสารสนเทศที่ต้องการ เช่น หนังสือ บทความสารสาร รายงานการวิจัย เป็นต้น
7. ช่วงระยะเวลาของสารสนเทศที่ต้องการ
8. ภาษาของเอกสารที่ต้องการ
9. ชื่อเรื่อง หัวข้อการศึกษา หรือวิจัย
10. คำศัพท์ที่ใช้ หรือที่เกี่ยวข้อง
11. ขอบเขตของเรื่องที่ต้องการ
12. ข้อมูลเพิ่มเติมอื่น ๆ อาทิ ชื่อผู้แต่ง ประวัติการค้นที่ผ่านมา เอกสารที่ใช้เป็นประจำ

สรุป

คำถามที่ได้จากการวิเคราะห์คำถาม (Question Analysis) หรือความต้องการของผู้ใช้ อาทิ

- แนวคิดหลักหรือคำสำคัญที่ต้องปรากฏในบทความที่ต้องการคืออะไร และแนวคิดนี้กำหนดคำแทนสาระได้อย่างไรบ้าง มีคำพ้องความหมายคำอื่นๆ อีกหรือไม่
- ต้องการข้อมูลแบบครอบคลุมทั้งหมด หรือต้องการบทความที่เกี่ยวข้องไม่กับบทความ
- ต้องการเฉพาะบทความใหม่ ๆ หรือต้องการข้อมูลจากทรัพยากรทุกประเภทที่มีอยู่
- ต้องการเฉพาะงานวิจัยที่เกี่ยวข้องกับมนุษย์ใช้หรือไม่
- ต้องการระบุกลุ่มประชากรที่ศึกษาหรือไม่ เช่น เพศ กลุ่มอายุ
- ต้องการเฉพาะข้อมูลที่เป็นภาษาอังกฤษเท่านั้น (ฐานข้อมูลส่วนใหญ่ เช่น MEDLINE, IPA จะจัดทำเป็นภาษาอังกฤษ รวมทั้งจัดที่สารระสังเขปเป็นภาษาอังกฤษให้ แต่ตัวจริงของเอกสารอาจเป็นภาษาอื่น)
- ลักษณะข้อมูลที่ต้องการเป็นแบบไหน เช่น บทความปริทัศน์ ผลการทดลองในห้องปฏิบัติการ หรือข้อมูลเชิงวิเคราะห์

ตัวอย่างแบบสัมภาษณ์

TIAC Your best information service

Search, Request Form
Search Number.....
DD/MM/YY.....

USER NAME: _____ USER ID: _____

PURPOSE: Research Business planing Thesis Other:

OUTPUT DELIVERY: Mail Personal Pick-up

OUTPUT PRINT: On-line Off-line

SEARCH SPECIFICATION:

(a) Need Information	(b) Type	(c) Language
<input type="checkbox"/> From Year.....	<input type="checkbox"/> Reference Only	<input type="checkbox"/> English
<input type="checkbox"/> To.....	<input type="checkbox"/> Reference, abstracts	<input type="checkbox"/> All languages
<input type="checkbox"/>	<input type="checkbox"/> Reference, index terms abstracts	<input type="checkbox"/>

TITLE(Or Topic) OF YOUR SEARCH: _____

KEYWORDS: Synonyms, trade names, scientific names, process, countries etc. _____

REQUEST STATEMENT: Describe the exact subject of the search. Indicate subject discipline, aspects of particular interest, related areas but should be excluded. Please use precise terminology wherever possible. _____



"The contents of databases is protected by copyright. I agree Not to duplicate the output of database searches whether in print or electronic form, for commercial purposes, for further publication, or for distribution outside my organization".

SIGNATURE _____

บรรณานุกรม

- Dervin, Brenda and Dewdney, Patricia. (1986). Neutral Question: A New Approach to the Reference Interview. *RQ* 25(4): 506-513.
- Parus, Dale J. (1996). The Reference Interview : Communication and Patron. [Online]. Available: <http://www.lis.uiuc.edu/review/winter1996/parus.html>.
- Wright, Melanie and Guy, Luara. (1997). Where do I find it and what do I do with it : practical problem-solving in the data library present at the 1997 lassist/FIDO conference in Odense, Denmark in May, 1997. [Online]. Available: http://dpls.dacc.wisc.edu/types/data_reference.htm
- The San Joaquin Valley Information Service. (1999). Reference inter view. CORE Correspondence Course [Online]. Available: <http://www.sjvls.lib.ca.us/sjvls/corr/ref/CCREF.htm>.



บทที่ 7

กลยุทธ์การค้นคืนสารสนเทศ (Search Strategy)

"When you design a search strategy you are planning how you will look for information. The more care and thought you put into your search strategy, the more relevant your search results will be."

กลยุทธ์การค้นคืนสารสนเทศ ได้แก่ การวางแผนการค้นหาเพื่อให้ได้มาซึ่งสารสนเทศที่ต้องการ การวางแผนการค้นหาที่ดีจะช่วยประหยัดเวลาในการค้นคืนสารสนเทศ ทำให้ผู้ค้นสามารถค้นคืนสารสนเทศได้หลากหลายวิธี เพื่อให้ได้สารสนเทศที่ตรงและเกี่ยวข้องกับเรื่องที่ต้องการ กลยุทธ์ในการค้นคืนสารสนเทศของผู้ค้นแต่ละคนอาจแตกต่างกันไปขึ้นอยู่กับความรู้ในเรื่องที่ค้น และประสบการณ์ในการค้นคืนสารสนเทศของผู้ค้นแต่ละคน

การสร้างกลยุทธ์ในการค้น

การสร้างกลยุทธ์ในการค้น จะเริ่มตั้งแต่ การวิเคราะห์แนวคิดและความต้องการสารสนเทศ และการแปลงแนวคิดที่ได้เป็นชุดคำศัพท์ อาจจำแนกขั้นตอนของการพัฒนากลยุทธ์การค้น ได้ 5 ขั้นตอน ดังนี้

1. กำหนดประเด็นของคำถาม (Finding the focus of your question)
 2. วิเคราะห์คำถาม และจำแนกแนวคิดที่อยู่ในคำถามออกมาเป็นคำศัพท์ (Finding key concept)
 3. กำหนดคำศัพท์อื่น ๆ ที่ใกล้เคียงกับเรื่องที่ต้องการค้น (Finding alternative terms)
 4. สร้างสูตรการค้น (Use basic search technique)
 5. กำหนดลักษณะสารสนเทศที่ต้องการ (Think about sort of information)
1. กำหนดประเด็นของคำถาม (Finding the focus of your question)
 - 1.1 กำหนดประเด็นของคำถามให้ชัดเจน โดยการตั้งคำถามที่เกี่ยวข้อง ชัดใจหยกที่กำหนดมาอย่างไม่ชัดเจน เช่น

How has Australia's relationship with England changed since 1945?

(ตั้งแต่ ปี 1945 ความสัมพันธ์ระหว่างออสเตรเลียกับอังกฤษมีการเปลี่ยนแปลงอย่างไร)

คำถามที่เกี่ยวข้อง

- ความสัมพันธ์ที่ต้องการศึกษาเป็นความสัมพันธ์ด้านใด : การเมือง การทหาร กฎหมาย การค้า หรือ มุมมองของคนส่วนใหญ่
- ความสัมพันธ์ช่วงปี 1945 เป็นอย่างไร และตอนนี้เป็นอย่างไร
- ความสัมพันธ์ของทั้งสองประเทศได้รับอิทธิพลจากประเทศอื่น หรือมีประเทศอื่นเข้ามาเกี่ยวข้องด้วยหรือไม่

- 1.2 หาคำตอบเบื้องต้นจากแหล่งอ้างอิง เช่น สารานุกรม หรืองานเขียนประเภทอื่นที่ให้ความรู้ทั่วไป คำตอบที่ได้ขึ้นอยู่กับสถานการณ์ หรือวัตถุประสงค์การนำไปใช้ เช่น คำตอบที่ต้องการอาจเป็นประเด็นด้านกฎหมาย เนื่องจาก ผู้ใช้กำลังศึกษาด้านกฎหมาย หรือ อาจต้องการความรู้ในเชิงประวัติศาสตร์ หากผู้ใช้เป็นศึกษาศาสาประวัติศาสตร์ เป็นต้น

2. ทำความเข้าใจกับแนวคิดหลัก (Define Key concepts)

- 2.1 วิเคราะห์คำถามออกเป็นส่วน ๆ ตามแนวคิดหลักของเรื่อง โดยการดึงคำสำคัญออกมา ซึ่งแนวคิดหลัก หรือคำสำคัญดังกล่าว จะเป็นประโยชน์ในการนำไปค้นคืนสารสนเทศจากฐานข้อมูลต่อไป และสามารถให้ตรรกะบูลีน ในการสร้างสูตรการค้น เช่น

How has Australia's relationship with England changed since 1945?

ประเด็นที่สำคัญได้แก่ : Australia, Relationship, English

- 2.2 ถามตัวเองว่าเข้าใจคำถาม และคำศัพท์ที่จะใช้ในการค้นหาลักษณะอย่างถ่องแท้หรือยัง ซึ่งการทำความเข้าใจกับเรื่องที่กำลังค้นหาและคำศัพท์ที่ใช้ในการค้น ซึ่งเป็นสิ่งสำคัญที่จะมีผลต่อการดำเนินการค้นและการประเมินผลการค้น เนื่องจากวิเคราะห์ผิดก็จะมีผลต่อการค้นที่ผิดพลาดได้ ผู้ค้นสามารถตรวจสอบความหมายของคำศัพท์ได้จากพจนานุกรมเฉพาะ

- 2.3 สำหรับแนวคิดพื้นฐานเกี่ยวกับประเด็นหลัก สามารถหาได้จาก สารานุกรม ตำรา หรือบริการตอบคำถามและช่วยการค้นคว้า

3. กำหนดคำศัพท์อื่น ๆ ที่ใกล้เคียงกัน (Finding alternative terms)

3.1 สิ่งที่ต้องคำนึงถึงได้แก่

- คำที่มีความหมายเหมือนกัน (synonyms) ตัวอย่าง เช่น mobile telephones, cellular telephones
- รูปคำเอกพจน์ / พหูพจน์ (plural / singular forms) ตัวอย่าง เช่น women, woman
- คำที่สะกดได้หลายแบบ (spelling variations) ตัวอย่าง เช่น behaviour, behavior
- คำที่มาจากศัพท์ดั้งเดิมอื่น (variations of a root word) ตัวอย่าง เช่น feminism, feminist, feminine
- คำย่อ และคำย่อ (acronyms) ตัวอย่าง เช่น chief executive officer, CEO

- 3.2 พิจารณาการใช้ภาษาอังกฤษแบบอเมริกัน เนื่องจาก หนังสือ วารสาร เอกสารเว็บ และฐานข้อมูลส่วนใหญ่ผลิตในสหรัฐอเมริกา ดังนั้น คำศัพท์ที่ใช้และรูปแบบการสะกดคำ จะใช้แบบอเมริกันเห็น ดังนั้น ในการกำหนดคำศัพท์จึงควรพิจารณาถึงรูปแบบการใช้คำศัพท์แบบอเมริกันด้วย

- 3.3 ใช้ริชชอร์ด หรือ พจนานุกรม หากไม่เข้าใจคำศัพท์เฉพาะ หรือการสะกดคำ เช่น

How has Australia's relationship with England changed since 1945?

(ตั้งแต่ ปี 1945 ความสัมพันธ์ระหว่างออสเตรเลียกับอังกฤษมีการเปลี่ยนแปลงอย่างไร)

Concept	Alternative
AUSTRALIA'S	Australian Australians Australia
RELATIONSHIP	Relation
ENGLISH	English British Great Britain Britain British Empire

4. ใช้เทคนิคการค้นพื้นฐาน ที่ไม่ซับซ้อน (Use basic search)

- สร้างสูตรการค้น หลังจากที่ได้จำแนกคำถามออกมาเป็นแนวคิด และกำหนดคำศัพท์สำคัญแล้ว ให้นำคำศัพท์เหล่านั้นมาสร้างสูตรการค้น หรือ Query เพื่อให้ระบบสารสนเทศเข้าใจคำถาม โดยใช้ตรรกะบูลีน (Boolean) ในการเชื่อมประเด็นของการค้น ใช้เทคนิคการตัดคำ (Truncation) ในการลดคำ และใช้คุณลักษณะด้านการค้นอื่น ๆ (Search feature) ช่วยในการกำหนดความเฉพาะเจาะจงของการค้น

ตัวอย่าง 1

How has Australia's relationship with England changed since 1945?

(ตั้งแต่ ปี 1945 ความสัมพันธ์ระหว่างออสเตรเลียกับอังกฤษมีการเปลี่ยนแปลงอย่างไร)

(Australia OR Australians) AND Relation AND (English OR British OR Great Britain OR Britain OR British Empire)

ตัวอย่าง 2

The use of microcomputers to teach reading skills to minority high school students.

Keywords: microcomputers, reading skills, minority, high school students

Natural Language :

- Microcomputers: personal computers, PC
- Reading Skills: reading ability, literacy,
- Minority: ethnic groups
- High school students: young adults, teenagers, adolescent

Thesaurus:

- Microcomputers:
- Reading Skills: reading comprehension, reading ability, language skills, literacy, reading improvement, vocabulary skills
- Minority: minority group children
- High school students: college bound students, non college bound students
high school freshmen, high school seniors

Search Formula

Microcomputers AND (reading skills OR reading ability) AND minority groups AND high school students

5. กำหนดลักษณะสารสนเทศที่ต้องการ หลังจากที่เราทราบประเด็นของเรื่องที่ต้องการแล้ว สิ่งที่ต้องทำต่อมาคือ พิจารณาว่าข้อมูลที่ต้องการมีลักษณะอย่างไรและจะมีอยู่ในแหล่งสารสนเทศประเภทใด

Type of information	Sources	Where to look
ข้อมูลที่ทันสมัย เป็นปัจจุบัน (Current or up to date information)	วารสาร อินเทอร์เน็ต	ฐานข้อมูล ครดรีนและสาระสังเขป Search engines Gateways
ข้อมูลที่ให้ภาพรวมของเรื่อง (Comprehensive information)	หนังสือ	รายการสารสนเทศของห้องสมุด ชั้นหนังสือ
ข้อมูลพื้นฐาน	หนังสือ สารานุกรม	รายการสารสนเทศของห้องสมุด

การปรับปรุงผลการค้น

หลังจากที่ทำการค้นแล้ว ในบางครั้งผลการค้นที่ได้ อาจได้จำนวนสารสนเทศที่มากเกินไป หรือน้อยเกินไป จนไม่สามารถใช้ตอบคำถามได้ ผู้ค้นอาจต้องกลับมาพิจารณาคำค้นที่ใช้อีกครั้ง รวมทั้งพิจารณาสูตรการค้นที่ใช้อีกครั้ง หากกำหนดไว้กว้างเกินไปต้องกำหนดให้แคบเข้ามา แต่หากกำหนดไว้แคบเกินไปจะต้องขยายให้กว้างขึ้น

1. การขยายผลการค้นให้กว้างขึ้น

การขยายผลการค้น ได้แก่ การเพิ่มจำนวนผลการค้นที่ตรงกับความต้องการให้มากขึ้นกว่าการค้นในครั้งแรก ทำได้ดังนี้

- ลดจำนวนแนวคิด หรือคำศัพท์ที่ใช้ลง
- ใช้ ตรรกะ OR ในการเชื่อมคำค้น หรือเชื่อมประเด็น
- นำคำพ้องอื่น ที่มีความหมายใกล้เคียงกันเข้ามารวมในสูตรการค้น
- ใช้คำศัพท์ที่กว้าง ๆ (general term)

- ใช้หัวเรื่องเป็นคำค้น (ศัพท์ในบัญชีหัวเรื่องมีความหมายกว้างกว่าในริซอร์ส)
- ตรวจสอบความถูกต้องของการใช้ ตรรกะบูลีนในการเชื่อมประเด็นการค้นแต่ละประเด็น
- ใช้การตัดคำเพื่อเพิ่มความหลากหลาย หรือความเป็นไปได้ของคำศัพท์ที่จะปรากฏในเอกสาร หรือสำหรับคำที่สะกดได้หลายแบบ

2. การจำกัดผลการค้นให้แคบลง

การจำกัดผลการค้น ได้แก่ การลดจำนวนผลการค้นที่ได้ให้น้อยลง

- ใช้ตัวเชื่อม AND หรือ NOT
- ใช้คำศัพท์ที่มีความเฉพาะเจาะจงยิ่งขึ้น
- ใช้คำศัพท์จากบัญชีหัวเรื่องเป็นคำค้น
- ตรวจสอบความถูกต้องของการใช้ ตรรกะบูลีนในการเชื่อมประเด็นการค้นแต่ละประเด็น
- ใช้คำศัพท์ที่ละเอียดขึ้น ตรงยิ่งขึ้น
- ยกเลิกเทคนิคการตัดคำ

3. ใช้แหล่งข้อมูลที่พบเป็นจุดเริ่มต้นของการค้นหาข้อมูลอื่นเพิ่มเติม

เมื่อได้สารสนเทศที่เกี่ยวข้อง หรือ ข้อมูลที่จะสามารถนำไปใช้ในการตอบคำถามได้แล้ว ควรใช้ข้อมูลที่ค้นพบเหล่านี้ เป็นจุดเริ่มต้นในการค้นหาเอกสาร อื่น ๆ ต่อไป โดยสามารถใช้ประโยชน์จากชื่อผู้แต่ง คำศัพท์ที่ปรากฏในเรื่อง ประเด็นสำคัญที่เอกสารกล่าวถึง และรายการบรรณานุกรมท้ายเอกสาร ซึ่งจะแสดงถึงรายการเอกสารอื่น ๆ ที่มีความเกี่ยวข้องกัน ซึ่งหลักนี้ มีลักษณะเช่นเดียวกับการค้นหาสารสนเทศในห้องสมุด โดยที่หนังสือที่มีเนื้อหาเรื่องเดียวกัน หรือเกี่ยวข้องสัมพันธ์จะจัดเรียงไว้บนชั้นเดียวกันภายใต้หมวดหมู่เดียวกัน ดังนั้นเมื่อพบว่า หนังสือเล่มหนึ่งมีเนื้อหาตรงกับความต้องการ นั้นหมายความว่าหนังสือเล่มอื่น ๆ บนชั้นเดียวกัน ที่จัดเรียงไว้ใกล้กันย่อมมีเนื้อหาคล้ายกัน หรือใกล้เคียงกัน น่าจะนำไปใช้ประโยชน์ได้

4. ใช้คำศัพท์ หรือคำค้นจากฐานข้อมูลของห้องสมุด หรือฐานข้อมูลที่เกี่ยวข้องเป็นแนวทางในการสืบค้นครั้งต่อไป
5. ใช้บรรณานุกรมของเอกสาร เพื่อค้นหาเอกสารอื่น ๆ เนื่องจากในบรรณานุกรมจะบอกถึงรายการเอกสารที่เกี่ยวข้องกับกับเอกสารที่อ่าน (ปัจจุบัน) ที่ผู้เขียนใช้ในการศึกษาเพื่อเขียนเอกสารปัจจุบัน

อาจสรุปกลยุทธ์ในการค้นคืนสารสนเทศได้เป็น 10 ขั้นตอนดังนี้

1. กำหนดแนวคิดสำคัญของเรื่องที่ต้องการค้น
2. กำหนดคำศัพท์ที่สามารถอธิบาย หรือใช้เป็นตัวแทนของแนวคิดดังกล่าวได้
3. กำหนดคำศัพท์อื่นที่มีความหมายใกล้เคียงกัน หรือเกี่ยวข้องสัมพันธ์กัน กับเรื่องที่ค้น
4. กำหนดเทคนิคการค้นที่เหมาะสม เช่น การตัดคำ ตรรกะบูลีน การกำหนดตำแหน่งคำค้น เป็นต้น
5. เครื่องมือช่วยค้นเช่น Search Engine หรือฐานข้อมูลที่คาดว่าจะมีคำตอบในเรื่องที่ต้องการ
6. อ่านคำแนะนำในการค้น ศึกษา ส่วนแสดงความช่วยเหลือ การค้นในขั้นสูง คำถามที่ถามบ่อย ๆ ฯลฯ
7. สร้างประโยค หรือสูตรการค้น ตามรูปแบบของเครื่องมือช่วยค้นหรือฐานข้อมูลที่เลือกใช้
8. ประเมินผลการค้นว่าตรงกับสูตรการค้นหรือไม่

9. ปรับปรุงสูตรการค้น หากผลการค้นที่ได้ยังไม่เป็นไปตามที่ต้องการ โดยกลับไปขั้นที่ 2-4 อีกครั้ง เพื่อแก้ไขสูตรการค้น
10. ค้นจากฐานข้อมูลอื่น ๆ โดยทำตามกระบวนการที่ 6-9 หากข้อมูลที่ได้ยังไม่เป็นที่พอใจ

สรุป ประเด็นสำคัญ ในการพัฒนากลยุทธ์ในการค้นคืนสารสนเทศ ได้แก่

- ระบุคำถามให้ชัดเจน
- ทำความเข้าใจกับเรื่องที่ค้น (หาความรู้เบื้องต้นเกี่ยวกับเรื่องที่ค้น)
- จำแนกคำถามออกเป็นแนวคิด/ คำศัพท์
- พิจารณาคำศัพท์ที่มีความหมายใกล้เคียงกัน
- พิจารณาแผนการค้นและผลการค้น
- ใช้สิ่งที่พบเป็นแนวทางในการค้นหาข้อมูลอื่น ๆ

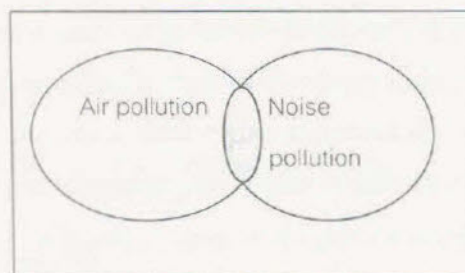
เทคนิคการค้นคืนสารสนเทศ (Search Technique)

เทคนิคการค้นคืนสารสนเทศ ได้แก่ วิธีการสร้างคำถาม (Query) หรือ สูตรการค้น (Search formula) ในการค้นคืนสารสนเทศ โดยการเชื่อมประเด็นการค้นหรือแนวคิดของเรื่อง รวมทั้งวิธีการกำหนดคำศัพท์ที่จะใช้ในการค้น ซึ่งเทคนิคการค้นได้แก่ การใช้ตรรกะบูลีน (Boolean logic) เทคนิคการตัดคำ (Truncation) การกำหนดตำแหน่งของคำค้น (Position operators)

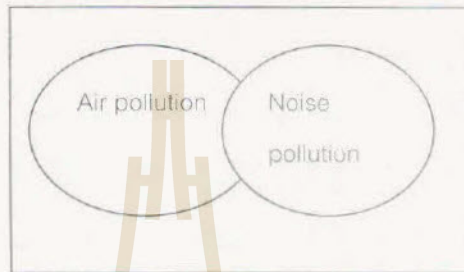
1. ตรรกะบูลีน (Boolean logic)

การแปลงแนวคิดออกมาเป็นคำถามของการสืบค้น (Query) ในบางครั้ง การใช้คำศัพท์เพียงคำเดียว อาจไม่สามารถครอบคลุมแนวคิดทั้งหมดของสารสนเทศที่ต้องการได้ ดังนั้น จึงต้องใช้ตรรกะบูลีนในการเชื่อมประเด็นต่าง ๆ เข้าด้วยกัน ตัวอย่างในการใช้ตรรกะบูลีนได้แก่ AND OR NOT

- 1.1 AND / และ ใช้ในกรณีที่ต้องการจำกัดประเด็นการค้น เพื่อดึงเฉพาะเอกสารที่มีคำศัพท์ทุกคำที่ใช้ในการค้นออกมา เช่น Air pollution AND Noise pollution เอกสารที่ถูกเรียกออกมาจะต้องมีแนวคิดทั้งสองประเด็นอยู่ในเอกสาร ดังภาพตัวอย่าง เอกสารที่มีทั้งเรื่องของ Air pollution และ Noise pollution คือเอกสารที่อยู่ในพื้นที่สีเทา



- 1.2 OR / หรือ เป็นตัวดำเนินการที่ใช้เพื่อต้องการขยายผลการค้น โดยจะดึงเอกสารทุกรายการที่มีแนวคิดหรือคำศัพท์ที่ใช้ในการค้นออกมา โดยเอกสารที่ถูกดึงออกมาไม่จำเป็นต้องมีคำศัพท์ทุกคำที่ใช้ในการค้นปรากฏอยู่ อาจมีเพียงคำใดคำหนึ่งเท่านั้นก็ได้ เช่น noise pollution OR air pollution หมายความว่า เอกสารที่ค้นออกมาได้แต่ละรายการไม่จำเป็นต้องมีทั้งเรื่องของมลภาวะทางเสียงและมลภาวะทางอากาศอยู่ในเอกสารเดียวกัน แต่มีเพียงเรื่องของมลภาวะทางเสียงเรื่องเดียว หรือ มลภาวะทางอากาศเรื่องเดียวก็จะถูกดึงออกมา OR มักจะใช้เพื่อต้องการขยายผลการค้นเมื่อผลการค้นที่ได้มีจำนวนน้อยเกินไป การใช้ OR จะทำให้ได้เอกสารทุกรายการที่มีประเด็นที่ต้องการค้น



อาจจำแนกการใช้ OR ได้ ดังนี้ ในกรณีที่ต้องการค้นด้วยคำศัพท์ที่สะกดได้หลายแบบ ศัพท์ที่มีความหมายเหมือนกันแต่เขียนต่างกัน หรือ การเชื่อมคำศัพท์ที่มีเนื้อหาเกี่ยวข้องกัน โดยมีวัตถุประสงค์ให้การค้นครอบคลุมคำศัพท์ทุกคำ เช่น

- คำศัพท์ที่สะกดได้หลายแบบ

Honor OR Honour

Organization OR Organisation

Judgement OR Judgment

- คำศัพท์ที่มีความหมายเหมือนกัน

Woman OR Female

Student OR Pupils

Occupation OR jobs OR Careers

- คำศัพท์ที่มีเนื้อหาเกี่ยวข้องกัน

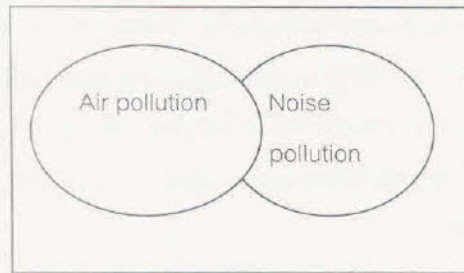
Children OR Adolescent OR Pediatric

Closed head injury OR Traumatic brain injury

Psychosocial OR Neuropsychological OR Behavioral

- 1.3 NOT / ไม่ เป็นตัวดำเนินการที่ใช้สำหรับตัดแนวคิดที่ไม่เกี่ยวข้องออกไป โดยจะเลือกเฉพาะรายการที่เกี่ยวข้อง ส่วนแนวคิดที่มาข้างหลัง NOT จะถูกตัดออก ดังนั้น เอกสารใดที่มีคำศัพท์ที่อยู่หลัง NOT ปรากฏอยู่จะไม่ถูกดึงออกมา เช่น Air pollution NOT Noise pollution หมายความว่าเอกสารที่จะถูกดึงออกมาต้องมีเฉพาะประเด็นของ Air pollution เท่านั้น เอกสารที่มีเรื่องของ Noise pollution อยู่ด้วยจะไม่ถูกดึงออกมาเช่นกัน เอกสารที่มีเรื่อง Noise pollution

เพียงประเด็นเดียวหรือ มี Air pollution และ Noise pollution อยู่ด้วยกันก็ตาม ก็จะไม่ถูกดึงออกมา ประเด็นสำคัญคือ ต้องมี Air pollution แต่ต้องไม่มี Noise pollution ดังภาพ



ลำดับการทำงานของตรรกบูลีน

- คำที่อยู่ในเครื่องหมายวงเล็บจะถูกดึงมาประมวลผลก่อน เช่น (a OR b) AND (c OR d) ระบบ จะทำการคำนวณในส่วนของ a OR b และ c OR d ก่อน จากนั้น จึงนำคำทั้งสองกลุ่มมา เชื่อมกันด้วย AND
- ระบบจะทำงานจากซ้ายไปขวา เช่น จะทำงานที่ a OR b ก่อน c OR d
- ในกรณีที่ไม่มีเครื่องหมายวงเล็บ ระบบจะทำงานตามลำดับตัวปฏิบัติ ดังนี้ 1) จะคำนวณ NOT ก่อน AND 2) จะคำนวณ AND ก่อน OR

2. เทคนิคการตัดคำ (Truncation)

เป็นเทคนิคการค้นที่ช่วยอำนวยความสะดวกแก่ผู้ใช้ในกรณีที่ผู้ใช้ จำรูปแบบการสะกดที่ถูกต้องของ คำศัพท์ที่ใช้ค้นไม่ได้ หรือไม่แน่ใจ หรือไม่ต้องการพิมพ์คำค้นดังกล่าวจนจบคำ หรือใช้ในกรณีที่คำดังกล่าว สะกดได้หลายแบบ มักใช้ในกรณีค้นจากรากศัพท์ เช่น

agricultural?
agricultural area
agricultural banks
agricultural cooperative
agricultural economy
etc.

เครื่องหมายที่ใช้แทนการตัดคำที่มักพบในฐานข้อมูลต่าง ๆ เช่น \$, ?, *, # ซึ่งก่อนใช้ ผู้ค้นควร ตรวจสอบก่อนว่าในฐานข้อมูลที่เลือกใช้เลือกใช้เครื่องหมายใด เพื่อป้องกันความผิดพลาดในการค้น การใช้ เทคนิคการตัดคำทำได้ ดังนี้

- 2.1 การตัดคำด้านหลังคำ มักใช้ในกรณีค้นจากรากคำ ดังตัวอย่างข้างต้น หรือ ต้องการค้นทั้งคำที่เป็นเอกพจน์และพหูพจน์ (กรณีเติม s หรือ es)

เช่น Comput? จะได้ Computer, Computers
Computing, Computation

- 2.2 การตัดกลางคำ มักใช้เพื่อแก้ปัญหาด้านการสะกดคำ อาทิ คำที่สะกดได้หลายแบบ หรือไม่แน่ใจว่าสะกดคำถูกต้อง

เช่น Colo?r จะได้ทั้ง Color และ Colour

Wom?n จะได้ทั้ง Woman และ Women

- 2.3 การตัดคำด้านหน้าคำ ใช้ในกรณีที่ไม่คำที่อยู่ด้านหลังเป็นคำเดียวกัน หรือสะกดเหมือนกัน

เช่น ?plankton จะได้ทั้ง Zoo Plankton

Cry Plankton

3. การกำหนดตำแหน่งของคำค้น (Positional Operator)

- 3.1 กำหนดการค้นตามเขตข้อมูล ลักษณะ รูปแบบการใช้คำสั่งจะขึ้นอยู่กับกรอบแบบของแต่ละฐานข้อมูล เช่น

IT = in title หมายถึง กำหนดให้ค้นที่เขตข้อมูลชื่อเรื่อง

transformation in ti เป็นรูปแบบที่ใช้ในฐานข้อมูลของ SilverPlatter

transformation. ti เป็นรูปแบบที่ใช้ในฐานข้อมูลของ OVID

PY = year ranges หมายถึง กำหนดให้ค้นที่เขตข้อมูลปีพิมพ์

:#2 and py = 1999

transformation and py >1999

DE = descriptor หมายถึง กำหนดให้ค้นที่เขตข้อมูลคำสำคัญ หรือ คำค้น

Transformation in de

Neoplasm. de

- 3.2 การกำหนดตำแหน่งของคำค้นในเขตข้อมูล (Proximity searches)

เพื่อกำหนดตำแหน่งของคำค้นในเขตข้อมูลให้มีความเฉพาะยิ่งขึ้น ใช้ในกรณีที่คำค้นมีมากกว่า 1 คำขึ้นไป เช่น กำหนดให้คำค้นที่ต้องการอยู่ติดกัน หรือวางอยู่ในตำแหน่งที่ใกล้กันโดยกำหนดให้มีคำอื่นมาอยู่คั่นได้ 2-3 คำ หรือกำหนดให้คำค้นที่ต้องอยู่ในบรรทัดเดียวกันเท่านั้น

เช่น ต้องการค้นคำว่า Information retrieval แต่ในบางครั้งผลการค้นที่ได้ จะมีทั้งรายการที่คำว่า Information retrieval อยู่ติดกัน ซึ่งเป็นเหตุการณ์ที่ต้องการ และ รายการที่ คำว่า Information retrieval ไม่อยู่ติดกัน เช่นอาจมีคำอื่นมาคั่นกลาง 2-3 คำ หรือคำแรกอยู่ต้นประโยค คำหลังอยู่ท้ายประโยค หรือ มีการสลับตำแหน่งของคำ เช่น นำ retrieval มาวางอยู่หน้า information โดยมีคำอื่น ๆ มาคั่น ซึ่งเป็นเหตุการณ์ที่ไม่ต้องการ ดังนั้น เพื่อป้องกันปัญหาดังกล่าว ในฐานข้อมูลจึงได้กำหนดให้มี การค้นแบบ Proximity searches

ตัวดำเนินการที่กำหนดให้ใช้ในฐานข้อมูลต่าง ๆ เช่น near, adjacency, with, (W), (nw), (F) ฯลฯ

ตัวอย่างเช่น

SilverPlatter :	transformation near dna	อยู่ใกล้กัน
OVID :	(transformation adj3 dna)	มีคำค้นได้ 3 คำ
DIALOG :	plant (W) production	อยู่ติดกัน
	Production (W1) plant	มีคำค้นได้ 1 คำ
	Economic (F) recover	คำแรกแล้วตามด้วยคำหลัง
BRS :	plant adj production	อยู่ใกล้กัน
	Product with plant	อยู่ติดกัน
	Aspirin same headache	อยู่ในบรรทัดเดียวกัน

การประเมินคุณค่าด้านเนื้อหาของเอกสาร

เป็นการประเมินคุณค่าของเอกสารที่ค้นได้ว่า มีความเหมาะสม มีคุณค่า มีความน่าเชื่อถือที่จะนำมาใช้ประกอบการค้นคว้าได้หรือไม่ การประเมินคุณค่าของเอกสารสามารถประเมินได้จาก

1. ชื่อเรื่อง มีความเฉพาะเจาะจงเพียงพอ หรือ มีชื่อเรื่องรองที่บ่งบอกถึงรายละเอียดของเนื้อหา
2. ชื่อผู้แต่ง เป็นผู้มีชื่อเสียงในสาขาวิชาที่เขียน เป็นสมาชิกขององค์กร หรือสมาคมวิชาชีพ มีคุณวุฒิในสาขาที่เขียน ที่สำคัญ คือ ต้องมีชื่อผู้แต่ง หรือชื่อผู้รับผิดชอบเนื้อหา
3. ปีพิมพ์ บอกรถึงความทันสมัย หรือช่วงเวลาของสารสนเทศที่เขียน โดยเฉลี่ยหนังสือแต่ละเล่มใช้เวลาในการจัดทำประมาณ 2 ปี ดังนั้น ความทันสมัยของเนื้อหาอาจไม่เท่ากับปีตีพิมพ์
4. ปีที่ หรือ วันที่ของการปรับปรุงแก้ไข จะบ่งบอกถึงความทันสมัยของข้อมูล วารสารบางชื่ออาจออกซ้ำกว่ากำหนด โดยเฉพาะวารสารวิชาการ เนื่องจากยังรวบรวมเนื้อหาที่ต้องการได้ไม่ครบ
5. ครั้งที่พิมพ์ จำนวนครั้งที่พิมพ์สามารถบอกรถึงคุณภาพหรือความนิยมของงาน
6. สำนักพิมพ์ เป็นองค์กรที่ถูกต้อง น่าเชื่อถือ ความเชี่ยวชาญในสาขาวิชาที่ตีพิมพ์
7. ชื่อวารสาร เป็นวารสารทางวิชาการ หรือบันเทิง หรือ เป็นวารสารที่ได้รับความนิยม
8. กลุ่มเป้าหมายคือใคร เนื่องจากความยากง่ายของเนื้อหาจะแตกต่างกันไปตามกลุ่มผู้ใช้
9. เจ้าของเว็บไซต์ ควรเป็นหน่วยงานที่น่าเชื่อถือ เช่น หน่วยงานของรัฐ (.gov) สถาบันการศึกษา (.ac, .edu) หรือสมาคมวิชาชีพต่าง ๆ เป็นต้น
10. มีคำอธิบายเนื้อหาคร่าวๆ หรือ ทดลองค้นหาที่ Search engine จัดทำ หรือ เจ้าของเว็บเป็นคนทำ โดยทั่วไป หาก Search engine เป็นผู้จัดทำ มักไม่ใช่คำอธิบายเนื้อหาที่แท้จริง เนื่องจากเป็นข้อความที่คัดลอกมาจาก 2-3 บรรทัดแรกของหน้าเอกสารเว็บ

บรรณานุกรม

- การค้นคืนสารสนเทศออนไลน์. (2535). กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย .
- พิมพ์ร่ำไพ เปรมสมิทธิ์. (2538). ฐานข้อมูลบรรณานุกรม : การสร้างและการใช้. กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย .
- มหาวิทยาลัยสุโขทัยธรรมมาธิราช สาขาวิชาศิลปศาสตร์. (2541). เอกสารการสอนชุดวิชาเทคโนโลยีสารสนเทศเบื้องต้น หน่วยที่ 9-15. พิมพ์ครั้งที่ 2. นนทบุรี : มหาวิทยาลัย.
- Monash Univeysity. (2000). How do I develop my search strategy. [Online]. Available: <http://www.lib.monash.edu.au/vl/ssstrat/ssstrcon.htm>.
- Walker, Geraldene and Janes, Joseph. (1999). Online Retrieval : A Dialog of theory and practice. Englewood, Colorado : Library Unlimited.

แบบฝึกหัด

- I'm looking for information about dan transformation in grampositive bacteria.
Keywords: _____
Alternative terms: _____
Search formula: _____
- The use of hypertext for information storage and retrieval : A case study in environment law.
Keywords: _____
Alternative terms: _____
Search formula: _____
- Graduate school of information management in USA.
Keyword: _____
Alternative terms: _____
Search formula: _____

4. The study of user satisfaction of user interface of DYNIX OPAC.

Keywords: _____

Alternative terms: _____

Search formula: _____

5. The impact of information architecture on academic web site usability

Keywords: _____

Alternative terms: _____

Search formula: _____

6. A study on result of information retrieval using human-weighted indexing technique.

Keywords: _____

Alternative terms: _____

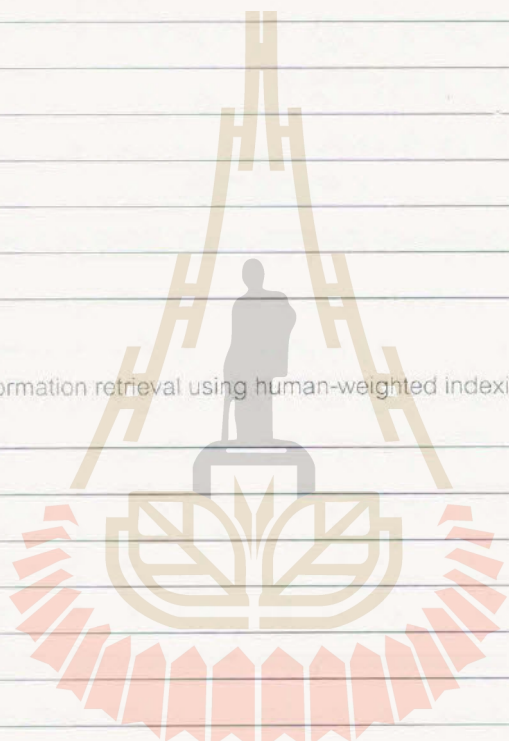
Search formula: _____

7. Metadata and Searching from the internet.

Keywords: _____

Alternative terms: _____

Search formula: _____



มหาวิทยาลัยราชภัฏนครราชสีมา

บทที่ 8

ฐานข้อมูลในการให้บริการค้นคืนสารสนเทศ (Database For Information Services)

ประเภทของฐานข้อมูลที่ใช้ในการบริการสารสนเทศ

ฐานข้อมูลที่ใช้ในการบริการค้นคืนสารสนเทศในปัจจุบันอาจแบ่งได้เป็น 4 ประเภทใหญ่ ๆ ดังนี้

1. ฐานข้อมูลที่พัฒนาขึ้นเอง (In-house Database)
2. ฐานข้อมูลออนไลน์ (Online Database)
3. ฐานข้อมูลซีดี-รอม (CD-Rom Datacase)
4. Search Engines

1. ฐานข้อมูลที่พัฒนาขึ้นเอง (In-house Database)

เป็นฐานข้อมูลที่หน่วยงานบริการสารสนเทศพัฒนาขึ้นมาใช้เอง เพื่อจัดเก็บสารสนเทศภายในของหน่วยงาน มักได้แก่ ฐานข้อมูลรายการทรัพยากรสารสนเทศของหน่วยงาน (Library catalog) และฐานข้อมูลเนื้อหาเต็มที่สร้างจากทรัพยากรสารสนเทศของหน่วยงาน โปรแกรมที่ใช้ในการจัดทำ อาทิ CDS/ISIS ชื่อโปรแกรมสำเร็จรูปสำหรับห้องสมุดอัตโนมัติ เช่น DYNIX, VTLS, INNOPAC, TINLIB, HORIZON เป็นต้น

2. ฐานข้อมูลออนไลน์ (Online Database)

เป็นฐานข้อมูลที่ต้องใช้วิธีการค้นคืนสารสนเทศจากระยะไกลไปยังหน่วยงานที่ให้บริการคือ Online search service การติดต่อเพื่อใช้บริการ ทำได้ 2 วิธีคือ ติดต่อโดยตรงผ่านสายโทรศัพท์ ในระบบ Package Switching ผ่านเครือข่าย ThaiPack ของ กสท. หรือ ติดต่อผ่านทางเครือข่ายอินเทอร์เน็ต ผู้ให้บริการฐานข้อมูลออนไลน์ อาจแบ่งได้เป็น 2 กลุ่ม คือ ผู้ให้บริการฐานข้อมูลทั่วไป และผู้ให้บริการฐานข้อมูลเฉพาะด้าน ตัวอย่าง เช่น

2.1 ผู้ให้บริการฐานข้อมูลทั่วไป เช่น

- DIALOG Information Services ให้บริการฐานข้อมูลกว่า 600 ฐาน มีบริการจัดส่งเอกสารชื่อ Knight Ridder Information
- OBIT Search Service ให้บริการกว่า 100 ฐาน ในสาขาวิชาเคมี พลังงาน วิศวกรรมและอิเล็กทรอนิกส์ สุขภาพ ความปลอดภัยและสิ่งแวดล้อม วัสดุศาสตร์ สิทธิบัตร 1 ใน 3 ของฐานข้อมูลจะไม่ซ้ำกับบริษัทอื่น
- BRS Information Technologies ให้บริการกว่า 120 ฐานข้อมูลในสาขาการแพทย์ การศึกษา วิทยาศาสตร์สุขภาพ ธุรกิจการเมือง สังคมศาสตร์ และมนุษยศาสตร์
- First Search ให้บริการฐานข้อมูลกว่า 70 ฐาน ด้าน ศิลปะ มนุษยศาสตร์ การศึกษา วิทยาศาสตร์สุขภาพ สังคมศาสตร์ ธุรกิจ เศรษฐศาสตร์ วิทยาศาสตร์ทั่วไป การแพทย์ และกฎหมาย

2.2 ผู้ให้บริการฐานข้อมูลเฉพาะด้าน

- STN (Scientific and Technical Information Network) เน้นข้อมูลด้านวิทยาศาสตร์ ได้แก่ เคมี ชีวะ ฟิสิกส์ วัสดุศาสตร์ วิศวกรรมศาสตร์

- Mead Data Central ให้บริการฐานข้อมูลด้านกฎหมาย ฐานข้อมูลที่ให้บริการคือ NEXIS และ LEXIS
- MEDLINE ของ NLM : National Library of Medicine เน้นด้านการแพทย์และสาธารณสุข

3. ฐานข้อมูลซีดี-รอม (CD-Rom Datacase)

ฐานข้อมูลซีดี-รอม จำแนกเป็น 2 ประเภทคือ ฐานข้อมูลบรรณานุกรม (Bibliographic database) ซึ่งมักจัดทำแยกตามสาขาวิชา และ ฐานข้อมูลเนื้อหาเต็ม (Full-text database) ซึ่งแยกจัดทำตามลักษณะเนื้อหาของทรัพยากรสารสนเทศ เช่น หนังสือพิมพ์ สารานุกรม พจนานุกรม แผนที่ เป็นต้น ตัวอย่างฐานข้อมูล ซีดี-รอม เช่น

3.1 ฐานข้อมูลบรรณานุกรม

- ABI/INFORM เป็นฐานข้อมูลด้านธุรกิจ
- AGRICOLA เป็นฐานข้อมูลด้าน เกษตรศาสตร์
- ERIC เป็นฐานข้อมูลด้าน การศึกษา
- LISA เป็นฐานข้อมูลด้าน บรรณารักษศาสตร์ และสารสนเทศศาสตร์
- Information Science เป็นฐานข้อมูลด้าน สารสนเทศศาสตร์
- COMPENDEX เป็นฐานข้อมูลด้าน วิศวกรรมศาสตร์
- MEDLINE เป็นฐานข้อมูลด้าน การแพทย์
- SSCI (Social Science Citation Index) เป็นฐานข้อมูลด้าน สังคมศาสตร์
- Life Science Collection เป็นฐานข้อมูลด้าน วิทยาศาสตร์ชีวภาพ
- DAO (Dissertation Abstract on Disc) เป็นฐานข้อมูล วิทยานิพนธ์

3.2 ฐานข้อมูลเนื้อหาเต็มรูป

- CHEM-BANK เป็นฐานข้อมูลด้าน เคมี
- CONSULT เป็นฐานข้อมูลด้าน การวินิจฉัยโรค และการรักษาโรค
- Facts On File เป็นฐานข้อมูลด้าน สรุปข่าวจาก นสพ. ทั่วโลก
- Internal Medicine เป็นฐานข้อมูลด้าน การแพทย์
- Patent Images เป็นฐานข้อมูลด้าน สิทธิบัตร
- CROSS CULTURAL เป็นฐานข้อมูลด้าน จิตวิทยา สังคมวิทยา และมานุษยวิทยา

ลักษณะการให้บริการของฐานข้อมูลซีดี-รอม มักจัดทำใน 2 ลักษณะ คือ

- 1) Stand alone Workstation เป็นการให้บริการในลักษณะเครื่องเดี่ยว โดยไม่มีการเชื่อมต่อกับเครือข่าย โดยการนำแผ่นซีดี-รอมไปอ่านเมื่อต้องการเรียกใช้ข้อมูล หรือ โดยการ กำหนดแยกตายตัวเป็น คอมพิวเตอร์ 1 เครื่อง ต่อ 1 ฐานข้อมูล
- 2) 2) CD-Net เป็นการให้บริการในลักษณะเครือข่าย ในระดับ LAN (Local Area Network) หรือ อินทราเน็ตขององค์กร โดยการนำฐานข้อมูลที่ให้บริการทั้งหมดใส่ไว้ในเซิร์ฟเวอร์ เพื่อให้บริการข้อมูลแก่ลูกข่าย โดยที่เครื่องลูกข่าย จะมีเมนูแสดงรายการฐานข้อมูลที่มีให้บริการ ผู้ใช้สามารถเรียกใช้ฐานข้อมูลที่ต้องการได้ผ่านทางเมนูดังกล่าว การให้บริการในลักษณะ

ของเครือข่าย ผู้ให้บริการจะต้องจ่ายค่าลิขสิทธิ์ในการเข้าใช้ตามจำนวนผู้ใช้ที่สามารถเข้าใช้งานได้ (User License) แก่ ผู้จัดจำหน่ายฐานข้อมูล

ประเภทของผู้ผลิตและผู้จำหน่ายฐานข้อมูล

ผู้ผลิตและผู้จำหน่ายฐานข้อมูลสามารถแบ่งประเภทตามลักษณะงานที่ทำ ได้ ดังนี้

1. Supermarket CD Publishers ได้แก่ ผู้ที่ทำหน้าที่เป็นตัวแทนจำหน่ายฐานข้อมูลในหลายสาขาวิชา โดยไม่ได้เป็นผู้ผลิตฐานข้อมูลเอง ตัวอย่าง เช่น Technology Supply Limited SilverPlatter และ Online Search Service เป็นต้น
2. Database Producers ได้แก่ กลุ่มบริษัทที่ผลิตฐานข้อมูลเอง และ ให้บริการผ่านทางออนไลน์ และ Web เช่น CA (Chemical Abstract) On CD Wilson Business Abstract Mintel CD and Niosis Genref on Compact Disc เป็นต้น
3. Publishers ได้แก่ สำนักพิมพ์ต่าง ๆ โดยการนำสิ่งพิมพ์ที่จัดทำและจัดจำหน่ายในรูปสิ่งพิมพ์มาจัดทำในรูปฐานข้อมูล และจัดจำหน่ายเอง เช่น Blackwell และ Wiley and Chadwyck Healey เป็นต้น
4. Document Supply Centers ได้แก่ ผู้ที่ทำหน้าที่รวบรวมและจัดเก็บสิ่งพิมพ์ต่อเนื่อง และรายงานการประชุมในรูปซีดีรอม เช่น British Library เป็นต้น

ประเด็นที่ต้องพิจารณาในการเลือกใช้ฐานข้อมูลออนไลน์และซีดีรอม

1. ความเหมาะสมและสภาพแวดล้อม ในกรณีของฐานข้อมูลออนไลน์ ได้แก่ ข้อมูลต้องทันสมัย มีการปรับปรุงอยู่เสมอ สามารถค้นได้ครั้งละหลายฐานข้อมูล ความครอบคลุมของสาขาวิชาที่ให้บริการ และระยะเวลาที่ครอบคลุม คือ สามารถค้นย้อนหลังได้หลายปี ในส่วนของฐานข้อมูลซีดีรอม สิ่งที่ต้องคำนึงถึงคือ โดยส่วนใหญ่ในหนึ่งแผ่นจะให้ข้อมูลเพียง 1 ปี ควรสอบถามถึงความถี่ของระยะเวลาในการปรับปรุง (Update) ฐานข้อมูล และระยะเวลาในการจัดส่งข้อมูลที่ปรับปรุงแล้วมาให้ อีกประเด็นหนึ่งคือ ซอฟต์แวร์ที่ใช้ในการค้น หากเป็นฐานข้อมูลที่มาจากบริษัทจะใช้ซอฟต์แวร์ที่ต่างกัน อาจเป็นกรณีที่มีการให้กับผู้ใช้ในการเรียนรู้การใช้ฐานข้อมูลเหล่านี้ ในส่วนของผู้ใช้บริการต้องจัดเตรียมฮาร์ดแวร์และระบบปฏิบัติการที่สามารถรองรับการทำงานของซอฟต์แวร์ดังกล่าวได้
2. ค่าใช้จ่าย อาจเปรียบเทียบค่าใช้จ่ายของฐานข้อมูลออนไลน์และซีดี-รอม ได้ดังนี้
 - ค่าการสื่อสาร การใช้งานฐานข้อมูลออนไลน์จะเกิดใช้จ่ายด้านการสื่อสาร ทั้งค่าโทรศัพท์ และอินเทอร์เน็ต
 - ค่าการใช้งานข้อมูล การเข้าใช้ฐานข้อมูลออนไลน์จะต้องเสียค่าบริการเข้าใช้บริการ อาจเสียเป็นรายครั้ง หรือ โดยการที่สถาบันบริการสารสนเทศสมัครสมาชิกเป็นรายปี ซึ่งอาจมีการจำกัดจำนวนครั้งที่สามารถเข้าค้นได้ หรืออาจมีการคิดค่าใช้จ่ายตามจำนวนคำค้น (query) ที่ใช้
 - ค่าสมาชิก กรณีที่เข้าใช้ฐานข้อมูลหนึ่งเป็นประจำควรสมัครสมาชิกเนื่องจากจะประหยัดค่าใช้จ่ายในการเข้าใช้ได้มากกว่า เนื่องจากมีส่วนลด และอาจได้รับบริการพิเศษอื่น ๆ

การเข้าใช้เป็นรายครั้งจะเสียค่าใช้จ่ายสูงกว่า ในส่วนของซีดี-รอมจะมีค่าใช้จ่ายในการบอกรับเป็นรายปีเช่นเดียวกับการบอกรับวารสาร

- มีบริการเสริมอื่น ๆ ผู้ให้บริการฐานข้อมูลออนไลน์มักจัดให้มีบริการพิเศษอื่น ๆ รวมไว้ด้วยเพื่ออำนวยความสะดวกให้กับผู้ใช้ เช่น บริการจัดส่งเอกสารทั้งทางอิเล็กทรอนิกส์ และทางไปรษณีย์ กรณีที่ผู้ใช้ต้องการสารสนเทศแบบเต็มรูป บริการเชื่อมโยงไปยังแหล่งข้อมูลอื่น ๆ ที่เกี่ยวข้อง บริการข้อมูลเต็มรูป บริการสอนการใช้ฐานข้อมูล ซึ่งบริการพิเศษเหล่านี้ (ยกเว้นอันหลังสุด) ผู้ใช้จะต้องเสียค่าใช้จ่ายเพิ่มในการขอรับบริการซึ่งจ่ายเป็นรายครั้ง ผู้ให้บริการฐานข้อมูลเพียงแต่ รวบรวมบริการเหล่านี้ไว้เพื่ออำนวยความสะดวกแก่ผู้ใช้ ไม่เสียเวลาต้องไปหาบริการจากที่อื่น ส่วนใหญ่มักเป็นบริษัทลูก หรือบริษัทที่มีความร่วมมืออยู่กับผู้ให้บริการฐานข้อมูล

3. วิธีการค้น ได้แก่ การเตรียมฟังก์ชันและความช่วยเหลือในการค้นคืนสารสนเทศ ในกรณีของฐานข้อมูลออนไลน์จะเพิ่มเติมในส่วนของการเตรียมความพร้อมของฐานข้อมูลให้สามารถพร้อมใช้ได้ตลอดเวลา ป้องกันปัญหากรณีสายไม่ว่าง กรณีเข้าใช้หลายคนพร้อมกัน ปัญหาทางเทคนิคอื่น ๆ ที่ทำให้การค้นต้องหยุดชะงัก เช่น สายหลุด เครือข่ายล่ม เป็นต้น
4. เวลาที่ใช้ในการค้น ในกรณีของการใช้บริจกอร์ออนไลน์ เนื่องจากการค้นผ่านทางเครือข่าย และมีค่าใช้จ่ายด้านการสื่อสาร ดังนั้นหากใช้เวลาในการค้นนาน ค่าใช้จ่ายด้านการสื่อสารก็จะเพิ่มขึ้นด้วย ดังนั้น ก่อนดำเนินการค้น ผู้ค้นจะต้องมีการเตรียมตัวมาก่อน และมีการวางแผนการค้นมาเป็นอย่างดี

4. Search Engine

Search Engine เป็นคำที่ใช้เรียกเครื่องช่วยค้นหาสารสนเทศจากอินเทอร์เน็ต ซึ่งมีวิธีการจัดทำ 2 แบบ คือ โดยการเขียนโปรแกรมให้คอมพิวเตอร์เป็นผู้จัดทำ (Crawler-based search engines) และ การจัดทำโดยใช้มนุษย์เป็นผู้จัดทำ (Human-powered directories) ในการค้นคืนสารสนเทศด้วย Search Engine ผู้ค้นพึงตระหนักว่า ผลการค้นที่ได้คือข้อมูลที่มาจาก ฐานข้อมูลของ Search Engine ไม่ใช่ข้อมูลที่เรียกออกมาจากเว็บต่าง ๆ โดยตรง โดย Search Engine จะทำการสำรวจเอกสารเว็บแล้วนำมารวบรวมไว้ในฐานข้อมูลของตน ดังนั้น ผลการค้นที่ได้จาก Search Engine ที่ต่างกัน จะให้ผลการค้นต่างกัน ขึ้นอยู่กับ ขนาดของฐานข้อมูล ความถี่ในการปรับปรุงข้อมูล ความสามารถของโปรแกรมการค้น การกำหนดครรรชนีและการจัดกลุ่มข้อมูล

วิธีการทำงานของ Search Engine

โดยทั่วไป Search engines มีกระบวนการทำงาน ดังนี้

การสร้าง search engines โดยใช้คอมพิวเตอร์เป็นผู้จัดทำ แบ่งการทำงานเป็น 3 ส่วน คือ

- 1) โปรแกรมรวบรวมเอกสารเว็บ คือ spider หรือ crawler จะทำหน้าที่ออกไปสำรวจเว็บจากโดเมนต่าง ๆ โดย spider จะเข้าไปสำรวจและอ่านหน้าเว็บ และหากพบ links ก็จะทำตาม links ภายใน site จนครบ ซึ่งจากการทำงานในลักษณะเียงเอนนี้ จึงเป็นที่มาของคำว่า spider หรือ crawler จากนั้น spider จะนำข้อมูลเว็บดังกล่าวไปเก็บไว้ในฐานข้อมูลของ Search Engine และ

spider จะกลับไปตรวจสอบข้อมูลในเว็บ นั้น ๆ อย่างสม่ำเสมอ เช่น ทุก 1 หรือ 2 เดือน เพื่อสำรวจความเปลี่ยนแปลง

- 2) รายการบรรณานุกรม ข้อมูลที่ spider พบจะถูกส่งต่อมาที่กระบวนการที่ 2 คือ การทำบรรณานุกรม ข้อมูลที่ spider พบจะถูกทำสำเนาและ ส่งมาจัดเก็บที่รายการบรรณานุกรม (index or catalog) ซึ่งเปรียบเสมือนเป็นสมุดเล่มใหญ่จัดเก็บสำเนาของเอกสารเว็บที่ spider พบ หากข้อมูลที่เว็บต้นฉบับมีการเปลี่ยนแปลงข้อมูลในสมุดบรรณานุกรมนี้จะเปลี่ยนแปลงด้วย เอกสารเว็บจะถูกทำบรรณานุกรมและจัดเก็บจะเก็บตามบัญญัติบรรณานุกรมที่กำหนดไว้
- 3) ส่วนประกอบที่ คือ Search engine software จะทำหน้าที่ ในการค้นคืนข้อมูลจากฐานข้อมูลของ search engine จะเริ่มต้นการทำงานเมื่อ ผู้ใช้ป้อนคำค้นหาเข้ามา โปรแกรมจะทำหน้าที่ นำคำค้นหาของผู้ใช้ไปจับคู่กับบรรณานุกรมในฐานข้อมูล แล้วทำการดึงข้อมูล (เอกสารเว็บ) ที่ตรงกับคำค้นหาออกมา และจัดลำดับผลการค้นตามระดับความเกี่ยวข้องที่โปรแกรมประเมินได้ ซึ่ง Search engine แต่ละตัวจะใช้ตรรกะที่แตกต่างกันไป ตัวอย่างการสร้างเกณฑ์ในการกำหนดค่าน้ำหนักให้กับคำบรรณานุกรมของ Search engine อาจจัดลำดับ ดังนี้ (ลำดับที่ 1 จะให้ค่าน้ำหนักความเกี่ยวข้องกับคำค้นหามากที่สุด)
 - ลำดับที่ 1 คำ หรือวลี ที่ปรากฏใน Metatag elements
 - ลำดับที่ 2 คำ หรือวลี ที่ปรากฏใน Title tag
 - ลำดับที่ 3 คำ หรือวลี ที่ปรากฏใน Main heading และ Sub heading
 - ลำดับที่ 4 จำนวนครั้งที่ คำ หรือวลี ปรากฏในส่วนเนื้อหาของเอกสาร
 - ลำดับที่ 5 ความถี่ที่เว็บอื่น ๆ เชื่อมโยงเข้ามา

ประเภทของ Search engines

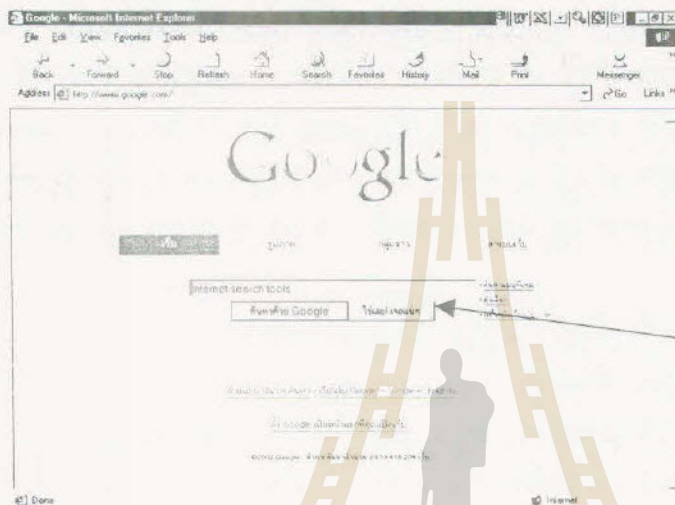
Search engines จำแนกได้เป็น 5 ประเภทใหญ่ ๆ ดังนี้ (Bradley, 2002: p. 19)

1. Free text Search Engines
2. Index or Directory -based Search Engines
3. Multi or Meta Search Engines
4. Natural-language Search Engines
5. Resource or Site-Specific Search Engines

1. Free text Search Engines

เป็น Search Engines ที่สามารถค้นได้โดยใช้คำค้นเพียงคำเดียว หรือ หลาย ๆ คำ หรือ ค้นด้วยวลีได้ เช่น อาจค้นด้วย ชื่อบริษัท วรรณคดีที่ยกมาจากบทกลอน หมายเลข ชื่อนักคิด หรือ คำภาษาต่างประเทศ เป็นต้น เหมาะกับการค้นหาสารสนเทศเฉพาะเรื่องมากกว่าการค้นหาเรื่องทั่ว ๆ ไป ที่ผู้ค้นยังไม่คุ้นเคยกับประเด็นเรื่อง เนื่องจากจะมีปัญหาในการกำหนดคำศัพท์ที่เกี่ยวข้องเพื่อนำไปใช้ในการค้น ดังนั้น คุณภาพของผลการค้นจะขึ้นอยู่กับความสามารถของผู้ค้นในการกำหนดคำค้น เป็นสำคัญ การใช้ Search Engines ประเภทนี้ จะให้ผลการค้นที่ดีหากผู้ค้นมีความคุ้นเคยกับเนื้อหาของเรื่องที่ค้น หรือ ผู้ค้นทราบแน่ชัดว่าต้องการสารสนเทศเรื่องใด หรือ การค้นที่ใช้ศัพท์เฉพาะจำนวนมาก ตัวอย่าง Search Engines ประเภทนี้ เช่น

- All the Web (<http://www.alltheweb.com>)
- Alta Vista (<http://www.altavista.com>)
- Excite (<http://www.excite.com>)
- Google (<http://www.google.com>)
- Hotbot (<http://www.hotbot.com>)
- Lycos (<http://www.lycos.com>)
- Northernlight (<http://www.northernlight.com>)



หน้าจอรับคำสั่ง
ของ Google

Search Box
และปุ่มรับคำสั่ง

2. Index or Directory -based Search Engines

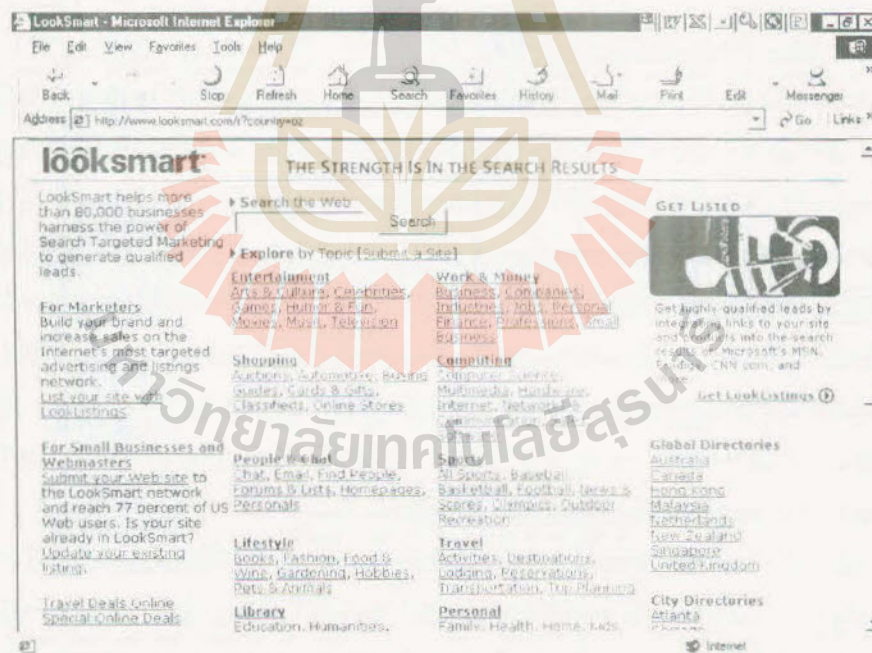
ลักษณะการจัดทำจะแตกต่างจาก Search Engines โดยทั่วไป เนื่องจากเป็นการจัดทำโดยมนุษย์ เป็นผู้จัดทำหน้านี้ โดยจะจัดกลุ่มข้อมูลตามหัวเรื่อง และภายใต้หัวเรื่องจะทำการแบ่งย่อยหัวเรื่องตามลำดับจากเรื่องทั่วไป ไปสู่เรื่องที่มีความเฉพาะเจาะจงมากขึ้น ซึ่งโครงสร้างของการจัดหมวดหมู่จะถูกเตรียมไว้ก่อน จากนั้นจึงนำเว็บไซต์ต่าง ๆ ที่รวบรวมมาไปจัดเก็บตามหมวดหมู่ที่จัดทำไว้ ข้อดีของ Directory คือ การแบ่งหมวดหมู่ที่ชัดเจนจะช่วยนำทางผู้ใช้ในการเข้าถึงข้อมูล จากประเด็นกว้าง ๆ ที่ยังไม่ชัดเจนไปสู่ ประเด็นเรื่องที่ชัดเจน หรือมีความเฉพาะยิ่งขึ้น ตามลำดับการจัดความสัมพันธ์ของหมวดหมู่ที่แบ่งไว้ ผู้ใช้จะเริ่มคาดเดาได้ว่า เรื่องที่ต้องการจะจัดเก็บได้ภายใต้หมวดหมู่ย่อยใด ดังนั้น ในการค้น ผู้ใช้ไม่จำเป็นต้องมีความรู้ในเรื่องที่ค้น หรือไม่จำเป็นต้องรู้คำศัพท์เฉพาะมาก่อน

อย่างไรก็ตาม Directory อาจนำมาซึ่งปัญหาในการค้นได้ หากการกำหนดโครงสร้างความสัมพันธ์ของหมวดหมู่ไม่ชัดเจน ไม่สมเหตุสมผล เช่น โครงสร้างหมวดหมู่ใหญ่ หมวดย่อยมีความซ้ำซ้อน หรือคาบเกี่ยวกัน หรือเรื่องเดียวกันแต่อยู่ได้หลายที่หลายระดับ อาจสร้างความสับสนให้กับผู้ค้นได้ และปัญหาที่สำคัญอีกประการหนึ่ง คือ การกำหนดคำที่ใช้แทนหมวดหมู่ ต้องมีความชัดเจน และสามารถครอบคลุมเนื้อหาภายใต้หมวดหมู่ ความเข้าใจในการจัดหมวดหมู่อาจขึ้นอยู่กับ ความรู้เดิมของผู้จัดทำและผู้ใช้ วัฒนธรรมของของแต่ละพื้นที่ ทักษะคติของของผู้จัดทำและผู้ใช้ ดังนั้น Directory บางตัว เช่น Yahoo! จึงแก้ปัญหาโดยการจัดทำออกมาหลายแบบตามการใช้งานในแต่ละพื้นที่

ข้อด้อยอีกประการหนึ่งของ Directory คือ ในกรณีที่ผู้ค้นมีความคุ้นเคยกับประเด็นเรื่องที่ต้องการสารสนเทศแล้ว และทราบแล้วว่าสารสนเทศที่ต้องการจัดเก็บไว้ที่หมวดหมู่ใด โดยเฉพาะหมวดหมู่ย่อยที่อยู่ในระดับลึก ผู้ใช้ต้องเสียเวลาในการเลือกเข้าไปที่ระดับชั้น จนถึงระดับที่ต้องการ ไม่สามารถที่จะเข้าถึงระดับของหมวดหมู่ที่ต้องการได้โดยตรง ผู้จัดทำจึงแก้ปัญหาด้วยการเพิ่ม Free-text search box เพื่อให้ผู้ค้นสามารถพิมพ์คำค้นเพื่อที่จะสามารถเข้าถึงหมวดหมู่ย่อยที่ต้องการได้โดยตรง

และเนื่องจาก Directory จัดทำดัชนีโดยมนุษย์ ดังนั้น ขนาดของฐานข้อมูลโดยส่วนใหญ่จะมีขนาดเล็กกว่า Search Engines ทั่วไป และในการจัดทำดัชนีมักจัดทำในระดับกว้าง คือจะให้ค่าแทนเนื้อหาทั้งหมดของเว็บไซต์ และเมื่อเรียกดูข้อมูลมักจะเรียกไปที่หน้า Home page ขณะที่ Search Engine ทั่วไป จะจัดทำดัชนีในระดับลึก คือจัดทำให้กับทุกหน้าของเว็บไซต์ และการเรียกจะเรียกไปที่หน้าที่มีคำดัชนีนั้นโดยตรง และเนื่องจากการจัดทำดัชนีเป็นการจัดทำในระดับกว้าง ดังนั้นจึงเหมาะในการค้นหาเรื่องทั่ว ๆ ไป จากกว่าเรื่องที่มีความเฉพาะเจาะจง ตัวอย่าง Search Engines ประเภทนี้ เช่น

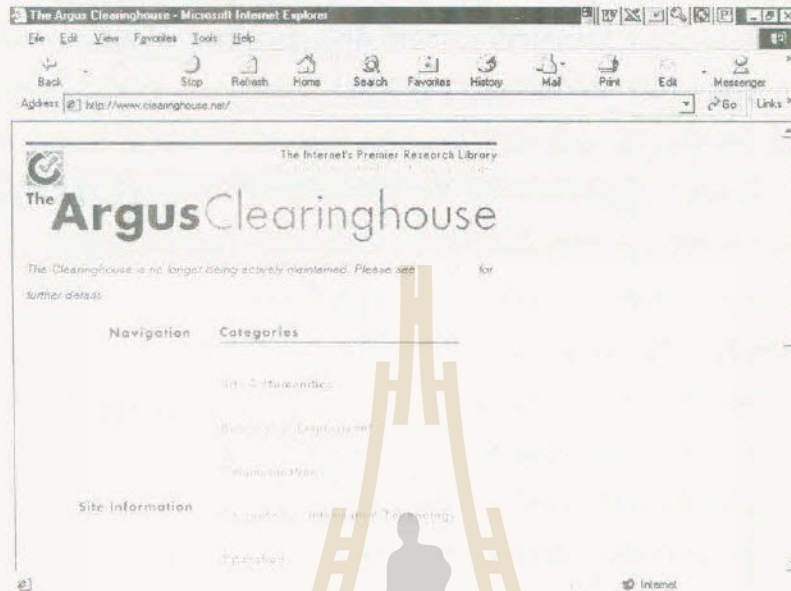
- Yahoo! (<http://www.yahoo.com>)
- Galaxy (<http://www.einet.net>)
- Webcrawler (<http://www.webcrawler.com/info.wbcnw>)
- Look Smart (<http://www.looksmart.com>)
- Open Directory (<http://www.dmoz.org>)



การแสดงผล Directory ของ Look Smart

นอกจากนี้ในกลุ่มของ Directory ยังแบ่งเป็นกลุ่มย่อยได้อีกกลุ่มหนึ่งคือ Specialized Subject Directory ซึ่งได้แก่ กลุ่มของ Directory ที่ทำหน้าที่รวบรวมข้อมูลเฉพาะเรื่องในลักษณะของ Clearinghouses โดยผู้ที่ทำหน้าที่รวบรวมข้อมูลจะเป็นผู้เชี่ยวชาญในแต่ละด้าน หรือสาขาวิชา และในการแสดงผลการค้นของ Directory ประเภทนี้ จะให้บรรณนิทัศน์ของเอกสารเว็บได้ด้วย ตัวอย่าง Directory กลุ่มนี้ เช่น

- Argus Clearinghouses (<http://www.clearinghouse.net>)
- About.com (<http://www.about.com>)
- WWW Virtual Library (<http://www.vlib.org>)



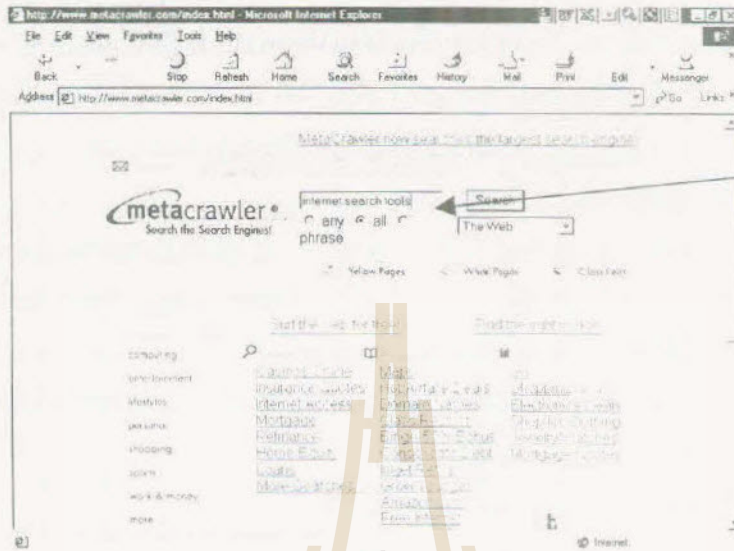
3. Multi- or Meta-Search Engines

Search Engines ประเภทนี้ อาจจัดได้ว่าไม่ใช่ Search Engines ที่แท้จริง เนื่องจากไม่ได้ทำการสืบค้นข้อมูลเอง แต่จะส่งต่อคำถามจากผู้ใช้ (query) ไปให้ Search Engines ตัวอื่น ซึ่งเป็นกลุ่ม Search Engines ที่อยู่ในความร่วมมือ ดังนั้น ผลการค้นที่ได้จึงเป็นผลการค้นจาก Search Engines ในกลุ่มนี้ ไม่ใช่จากฐานข้อมูลของ Multi-Search Engines โดยจะแสดงที่มา (ชื่อของ Search Engines) ที่เป็นเจ้าของข้อมูลไว้ต่อท้ายรายการที่ค้นได้แต่ละรายการ ข้อดีของ Search Engines ประเภทนี้คือ ผู้ใช้สามารถค้นเรื่องที่ต้องการได้จากแหล่งเดียว ไม่ต้องเสียเวลาไปค้นจากหลายที่ โดย Search Engines จะตัดข้อมูลที่มีความซ้ำซ้อนกันออกไป เหมาะที่จะใช้ในกรณีที่ต้องการรวบรวมข้อมูลที่ต้องการให้ครอบคลุมมากที่สุด เนื่องจาก Search Engines เพียงตัวเดียวอาจรวบรวมข้อมูลได้ไม่ครอบคลุมทั้งหมด หรือ Search Engines อาจไม่ได้ทำดรรชนีให้ และที่สำคัญ Multi-Search Engines ช่วยประหยัดเวลาในการค้นให้กับผู้ใช้

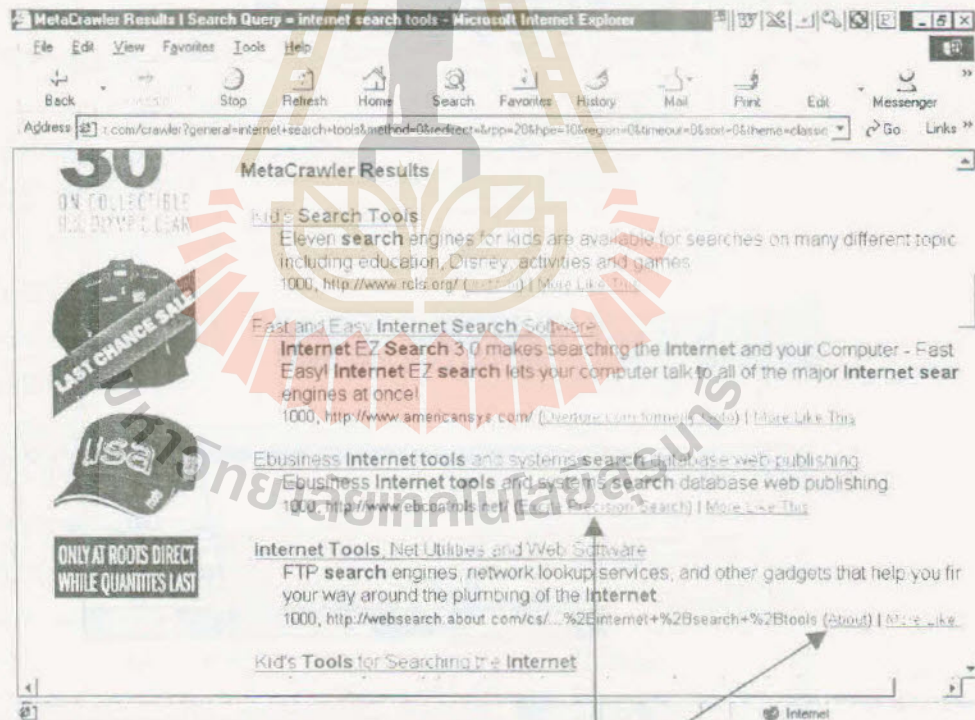
อย่างไรก็ตาม Multi-Search Engines มีข้อด้อยประการสำคัญที่ผู้ค้นต้องคำนึงถึงคือ รูปแบบคำค้นที่ Search Engines แต่ละตัวใช้ เนื่องจากการค้นด้วย Multi Search Engines เป็นการค้นจากคำค้นเดียว (Query) จากหน้าจอของ Multi-Search Engines แต่ในความเป็นจริง Search Engines แต่ละตัวมีโครงสร้างประโยค (Syntax) ของตนเองซึ่งแตกต่างกันไปตามแต่จะกำหนด ดังนั้น ในบางกรณี คำค้น มีการสร้างสูตรการค้นที่ซับซ้อน หรือ ใช้ภาษาอื่น ๆ ที่ไม่ใช่ภาษาอังกฤษ อาจให้ผลการค้นที่ไม่เที่ยงตรงได้ เนื่องจากไม่เข้าใจคำสั่งที่แท้จริง ตัวอย่าง Search Engines ประเภทนี้ เช่น

- Dogpile (<http://www.dogpile.com>)
- Metacrawler (<http://www.metacrawler.com>)
- ProFusion (<http://www.profusion.com>)

- Search (<http://www.search.com>)
- SurfWax (<http://www.surfwax.com>)
- Ixquick (<http://www.ixquick.com>)



ค้นหาด้วย
คำว่า internet
search tools



ผลการค้นแสดง รายชื่อของ Search Engines ที่ให้ข้อมูล
เพื่อทำการเชื่อมโยงไปดู่ข้อมูลจาก Search Engines
ดังกล่าว

4. Natural-language Search Engines

Search Engines กลุ่มนี้มีจำนวนไม่มาก เป็นพัฒนาการใหม่ของ Search Engines โดยในการสืบค้นข้อมูล นอกจากจะค้นตามคำค้นที่ผู้ใช้พิมพ์เข้ามาแล้ว Search Engines ประเภทนี้จะนำคำค้นของผู้ใช้ไปเปรียบเทียบกับคำศัพท์อื่น ๆ ที่มีลักษณะใกล้เคียงกัน หรือ มีความเกี่ยวข้องกัน เช่น เมื่อผู้ใช้ค้นด้วยคำว่า "tax revenue" Search Engines จะทราบว่าเรื่องที่ค้นเกี่ยวข้องกับเรื่องใดบ้างและจะทำการดึงข้อมูลจากประเด็นที่คิดว่าเกี่ยวข้องกับเรื่องที่ค้นออกมาด้วย เช่น ในตัวอย่างนี้ Search Engines จะมองหาประเด็นต่อไปนี้ด้วย Financial, Business and Economic เป็นต้น ที่เป็นที่ยอมรับคือ Ask Jeeves (<http://www.aj.com>) หรืออีกรูปแบบหนึ่ง ที่จัดทำแบบ British version คือ <http://www.ask.co.uk>

Search Engines ประเภทนี้จะมีประโยชน์อย่างมาก เมื่อผู้ใช้ไม่ทราบว่าจะใช้คำศัพท์คำใดในการค้นหาสารสนเทศที่ต้องการ เช่น ต้องการค้นหาข้อมูลเกี่ยวกับก๊าซที่สามารถละลายในน้ำ โดยที่ผู้ค้นไม่มีความรู้ในเรื่องดังกล่าวมาก่อน ผู้ค้นสามารถทำการค้นโดยการถามคำถามว่า "Which gases are partially soluble in water?" Search Engines จะทำการค้นหาคำตอบให้ได้ ตัวอย่าง Search Engines ประเภทนี้ เช่น

- Ask Jeeves (<http://www.aj.com>)
- Ask Jeeves UK (<http://www.ask.co.uk/>)
- Albert (<http://www.albert.com/demo.php>)

The image shows two screenshots of the Ask Jeeves search engine interface. The top screenshot shows the search page with a search bar containing the question "Which gases are partially soluble in water?" and a search button. The bottom screenshot shows the search results page with a list of links and titles related to water treatment and ammonia. Arrows point from text boxes to the search bar and the results list.

คำถาม

คำตอบจากคำถาม

"Which gases are partially soluble in water?"

You may find my search results helpful:

1. Venturi Aeration: Innovative Technologies in Liquids Treatment. "Water Treatment" Page: Non-chemical pH Adjustment, Corrosion Control, Radon Removal and VOC Stripping Summary Non-chemical pH Adjustment, Corrosion Control, Radon... From: http://www.venturinaeration.com/venturi_control.htm
2. Island Well Drillers Water Problems When millions of vapor particles unite, they form droplets of moisture. As these increase in ... Cover Page The Hydrologic Cycle Meteric Water... From: http://www.islandwelldrillers.com/wp_mete_water.htm
3. Glossary of Terms ... Water Chemistry, Water Treatment & Fuel Oil Technology Kor-Chem Inc. KCT is the Trade Mark for Kor-Chem, Inc. Our Plant and office is located at ... Kor-Chem Inc. KCT is the Trade Mark for... From: <http://www.kor-chem.com/texture/glossary.html>
4. MSDS Anhydrous Ammonia Sherritt Granular Ammonium Sulphate 19-20-22 Sherritt Crystalline Ammonium.

5. Resource or Site-Specific Search Engines

อาจเรียกได้ว่าเป็นกลุ่มของ Search Engines ที่มีจำนวนมากที่สุด เป็น Search Engines ที่จัดทำขึ้นเพื่อใช้ค้นหาข้อมูลในเรื่องใดเรื่องหนึ่งโดยเฉพาะ หรือจัดทำขึ้นเพื่อค้นหาข้อมูลจากทรัพยากรประเภทใดประเภทหนึ่งเท่านั้น หรือเพื่อค้นหาข้อมูลจากเฉพาะพื้นที่ โดยจะเน้นการรวบรวมทั้งในระดับกว้างและลึกของสาขาวิชาที่รับผิดชอบ ข้อมูลที่จัดเก็บมักไม่มีใน Search Engines ทั่ว ๆ ไป (Tyner, 2002) เช่น

Regional (Canada)

- AltaVista Canada
(<http://www.altavistacanada.com>)
- Excite Canada (<http://www.excite.ca/>)
- Yahoo! Canada (<http://ca.yahoo.com/>)

Regional (Other)

- Geographically specific search engines (Beaucoup):
 - Americas
 - Asia/Australia/Middle East/Africa
 - Europe

Companies

- Advice for Investors
(<http://www.adviceforinvestors.com>)
- Hoover's Online (<http://www.hoovers.com/>)
- InfoSpace Canada
(http://www.infospace.com/canada/index_yl_w_ca.htm)
- Wall Street Research Net
(<http://www.wsrn.com/>)
- WorldPages (<http://www.worldpages.com/>)

People (E-mail addresses)

- Bigfoot (<http://bigfoot.com/>)
- InfoSpace Canada Email Search
(<http://www.infospace.com/canada/email1.htm>)
- WhoWhere?
(<http://www.whowhere.lycos.com>)
- Yahoo! People Search
(<http://people.yahoo.com/>)

People (Postal addresses & telephone numbers)

- Bigfoot (<http://bigfoot.com/>)

- Canada 411
(<http://canada411.sympatico.ca>)
- InfoSpace Canada People Finder
(http://www.infospace.com/canada/index_pl_ca.htm)

- Switchboard.Com
(<http://www.switchboard.com>)

Images

- The Amazing Picture Machine
(<http://www.ncrtec.org/picture.htm>)
- Lycos Multimedia
(<http://multimedia.lycos.com/>)
- WebSEEk
(<http://www.ctr.columbia.edu/webseek/>)
- Yahoo! Picture Gallery
(<http://gallery.yahoo.com/>)

Jobs

- +Jobs Canada
- Monster.ca (<http://www.monster.ca/>)
- Monster.com (<http://www.monster.com/>)
- Canada Job Bank
(<http://www.jobbank.gc.ca/>)
- The Riley Guide
(<http://www.dbm.com/jobguide/>)

Games

- Games Domain
(<http://www.gamesdomain.com/>)
- GameSpot
(<http://gamespot.com/gamespot/>)

Software

- Jumbo (<http://www.jumbo.com>)
- Shareware.com
(<http://shareware.cnet.com/>)

- [ZDNet Downloads](http://www.zdnet.com/downloads/)
(<http://www.zdnet.com/downloads/>)

Health/Medicine

- [Achoo](http://www.achoo.com/) (<http://www.achoo.com/>)
- [BioMedNet](http://www.bmn.com/) (<http://www.bmn.com/>)
- [Combined Health Information Database](http://chid.nih.gov/)
(<http://chid.nih.gov/>)
- [MayoClinic.com](http://www.mayohealth.org/)
(<http://www.mayohealth.org/>)
- [MEDLINEplus](http://www.nlm.nih.gov/medlineplus/)
(<http://www.nlm.nih.gov/medlineplus/>)

Education/Children's Sites

- [AOL NetFind Kids Only](http://www.aol.com/netfind/kids/)
(<http://www.aol.com/netfind/kids/>)
- [Blue Web'n](http://www.kn.pacbell.com/wired/bluewebn/)
(<http://www.kn.pacbell.com/wired/bluewebn/>)
- [Education World](http://www.education-world.com/) (<http://www.education-world.com/>)
- [Kids Domain](http://www.kidsdomain.com/) (<http://www.kidsdomain.com/>)
- [KidsClick!](http://www.kidsclick.org/) (<http://www.kidsclick.org/>)
- [Yahooligans!](http://www.yahooligans.com/) (<http://www.yahooligans.com/>)

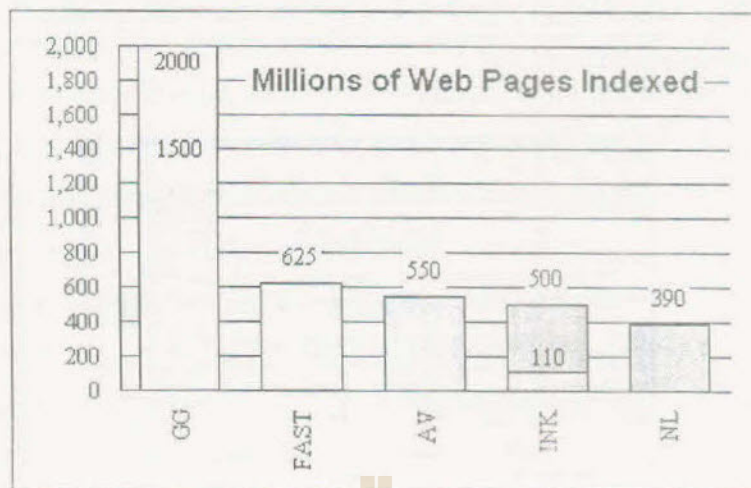
การเปรียบเทียบ Search Engine

การเปรียบเทียบ Search Engine มีเกณฑ์ในการพิจารณา (Tyner, 2002) ดังนี้

1. ขนาดของฐานข้อมูล (Database Size)
2. ความทันสมัยของฐานข้อมูล (Currency)
3. คุณลักษณะด้านการติดต่อกับผู้ใช้ ในการค้นข้อมูล (Search Interfaces)
4. คุณลักษณะด้านการค้น (Search Features)
5. คุณลักษณะด้านการแสดงผลการค้น (Results List Display Features)
6. คุณลักษณะพิเศษอื่น ๆ (Other Features of Note)

1. ขนาดของฐานข้อมูล (Database Size)

ขนาดของฐานข้อมูลแสดงถึง ปริมาณของเอกสารเว็บที่ Search Engines รวบรวมและทำดัชนีได้ ฐานข้อมูลขนาดใหญ่ย่อมหมายถึงปริมาณเอกสารที่มีมาก และน่าจะครอบคลุมสารสนเทศที่ต้องการมากกว่า ฐานข้อมูลขนาดเล็ก อย่างไรก็ตาม ขนาดของฐานข้อมูลขึ้นอยู่กับความสามารถของฮาร์ดแวร์และซอฟต์แวร์ที่ใช้ ในการดำเนินการ อย่างไรก็ตามคุณภาพของ Search Engines ไม่ได้ขึ้นอยู่กับขนาดของฐานข้อมูลเพียง อย่างเดียว หากเปรียบเทียบผลการค้นที่ได้จากเครื่องมือค้นที่ใช้ คนกับคอมพิวเตอร์ทำ แน่นอนว่าปริมาณ ผลการค้นที่ได้ ระบบที่ใช้คอมพิวเตอร์ทำดัชนีย่อมได้ผลการค้นที่มากกว่า แต่หากมองในแง่คุณภาพของ การทำดัชนี ระบบที่ใช้คนทำดัชนีจะให้ผลการค้นที่เที่ยงตรงกว่า นั่นคือ ปริมาณผลการค้นที่มากไม่ได้ หมายความว่า Search engine ตัวนั้นจะมีประสิทธิภาพสูงกว่า Search engine ตัวอื่น ๆ



KEY: GG=Google, FAST=FAST, AV=AltaVista, INK=Inktomi, NL=Northern Light.

ข้อมูล ณ วันที่ 11 ธันวาคม 2544 (<http://www.searchenginewatch.com/reports/sizes.html>)

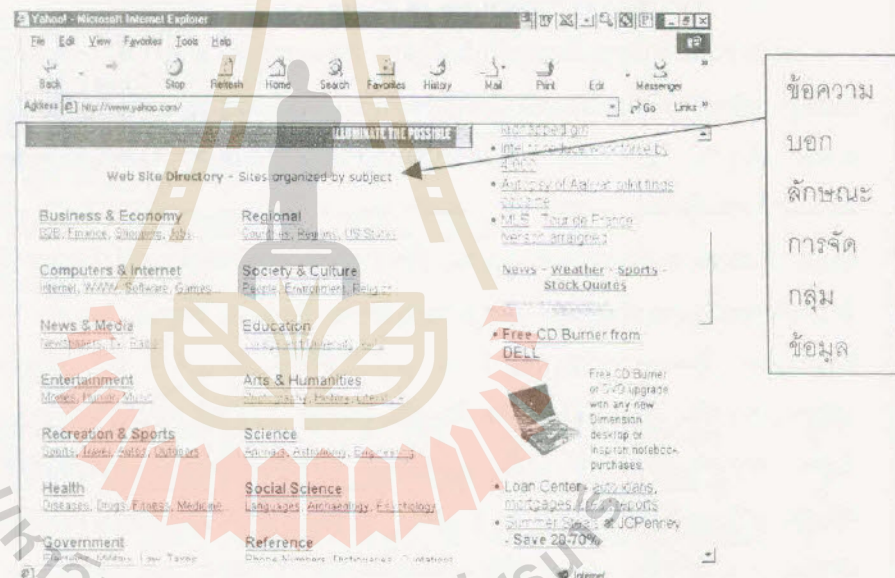
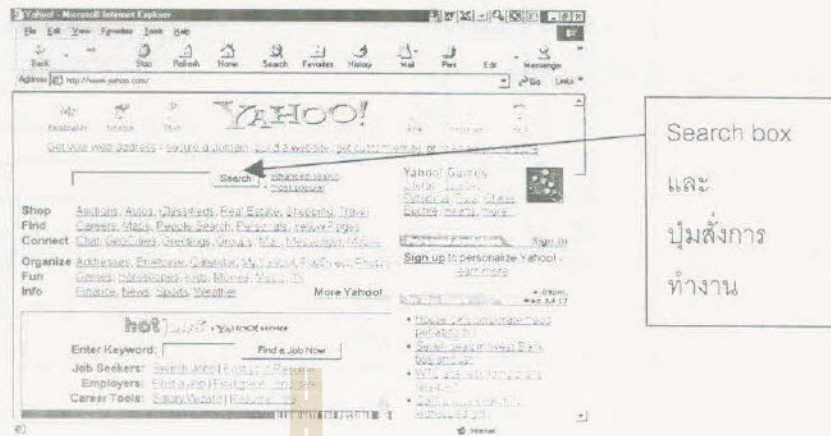
2. ความทันสมัยของฐานข้อมูล (Currency)

ความทันสมัยของข้อมูลเป็นสิ่งสำคัญ โดยเฉพาะเอกสารเว็บที่มีการผลิตออกมาอย่างรวดเร็ว ผู้ใช้แทบทุกคนต้องการข้อมูลที่ทันสมัย ต้องการทราบความเคลื่อนไหวล่าสุดในเรื่องที่สนใจ ผู้จัดทำ Search Engine ควรวางนโยบาย และกำหนดการในการปรับปรุงข้อมูลให้ทันสมัย ซึ่งความทันสมัยเป็นสิ่งที่จะสามารถดึงดูดความสนใจของผู้ใช้ได้ โดยส่วนใหญ่ Search engine ที่ใช้ Spider หรือ Crawler เป็นตัวรวบรวมข้อมูล ข้อมูลจะค่อนข้างทันสมัย เนื่องจาก ทั้ง Spider หรือ Crawler จะเข้าไปเยี่ยมเว็บที่ทำดัชนีให้แล้วอยู่เป็นประจำเพื่อทำการตรวจสอบและปรับปรุงหากข้อมูลในเว็บมีการเปลี่ยนแปลงข้อมูลที่ Search Engines ก็เปลี่ยนแปลงตามไปด้วย ซึ่งต่างจากข้อมูลใน Directory ซึ่งใช้คนเป็นผู้ทำดัชนี หากข้อมูลในหน้าเอกสารมีการเปลี่ยนแปลง หรือ มีการเปลี่ยน URL ข้อมูลใน Directory จะไม่มีการแก้ไข อีกประเด็นหนึ่งที่ต้องพิจารณาในเรื่องของความทันสมัยของข้อมูลคือ ความถี่ในการทำดัชนีให้กับข้อมูลใหม่ที่เกิดขึ้น อย่างไรก็ตามผู้ใช้ควรตระหนักว่าข้อมูลที่ได้จาก Search Engines ไม่ใช่ข้อมูลที่มีความเป็นปัจจุบันเหมือนข่าวที่มีการรายงานทันทีเมื่อมีเหตุการณ์เกิดขึ้น แต่จะเป็นข้อมูลที่เกิดขึ้นมาแล้วในช่วงเวลานั้น ซึ่งอาจจะมีมากกว่า 1 วันขึ้นไป เมื่อ Search Engines ไปพบ และสำเนากลับมาไว้ในฐานข้อมูลและกำหนดดัชนีให้ ผู้ค้นข้อมูลจาก Search Engines จึงจะสามารถพบเอกสารดังกล่าว ซึ่ง ความถี่ในการปรับปรุงฐานข้อมูลของ Search Engine แต่ละตัวจะแตกต่างกันไปตามการโปรแกรมให้กับ Spider หรือ Crawler

3. คุณลักษณะด้านการติดต่อกับผู้ใช้ ในการค้นข้อมูล (Search Interfaces)

คุณลักษณะด้านการติดต่อกับผู้ใช้ ในการค้นข้อมูล ได้แก่ การอำนวยความสะดวกในการใช้งานเครื่องมือต่าง ๆ ตลอดช่วงระยะเวลาดำเนินการสืบค้นข้อมูล เริ่มตั้งแต่การพิมพ์คำค้น หรือการเลือกกลุ่มข้อมูลควรที่จะต้องมีความชัดเจน ว่าผู้ค้นจะต้องเริ่มการทำงานอย่างไร จะต้องไปที่ไหนก่อน โดยการแสดง prompt command หรือ ข้อความที่บอกว่าจะเริ่มการค้นอย่างไร เช่น ใน Yahoo! จะแสดง Search box ที่ชัดเจน เพื่อให้

ผู้ใช้พิมพ์คำค้นที่ต้องการในช่องที่เหลื่อม พร้อมทั้งปุ่มสั่งการทำงาน หลังจากที่พิมพ์คำค้นแล้ว และข้อความบอกลักษณะการจัดกลุ่มของข้อมูล ดังภาพ



รวมทั้งการเตรียมข้อความให้ความช่วยเหลือ และการจัดทำหน้าให้ความช่วยเหลืออื่น ๆ การเตรียมตัวนำทาง (Navigators) ไปยังขั้นตอนต่าง ๆ ของการค้นหา การจัดเตรียมวิธีการค้นทั้งแบบง่าย (Simple Search) และแบบที่สามารถสร้างสูตรการค้นหาได้เฉพาะมากขึ้น (Advanced Search) ความสะดวกในการแก้ไขสูตรการค้นหา การเลือกค้นเฉพาะข้อมูลบางประเภท เช่น ภาพ หรือ เสียง เป็นต้น

4. คุณลักษณะด้านการค้น (Search Features)

คุณลักษณะด้านการค้น ได้แก่ การเตรียมฟังก์ชัน และเทคนิคการค้นหาต่าง ๆ เพื่อช่วยในการค้นข้อมูล เช่น ตรรกะบูลีน AND OR NOT เพื่อจำกัดและขยายแนวคิดในการค้น การใช้เทคนิคการตัดคำ (Truncation) การกำหนดตำแหน่งของคำค้นจากเขตข้อมูลต่างและการกำหนดตำแหน่งของคำค้นในเนื้อหาส่วนต่าง ๆ เช่น กำหนดให้อยู่ใกล้กัน หรือติดกัน เป็นต้น (Positional and Proximity search) การจำกัดการค้นหาด้วยภาษา

การกำหนด Case sensitive (การใช้ตัวพิมพ์ใหญ่และพิมพ์เล็กในการพิมพ์คำค้น มีผลต่อการค้น) ซึ่งใน Search Engines แต่ละตัวจะจัดเตรียมไว้ไม่เหมือนกัน อย่างไรก็ตาม Search Engines ทุกตัว ควรมีฟังก์ชันการค้นพื้นฐาน คือ AND OR NOT เพื่อที่ผู้ใช้จะสามารถขยาย หรือ จำกัดผลการค้นได้ กรณีที่ผลการค้นในครั้งแรกยังไม่เป็นที่พอใจ และสัญลักษณ์ต่าง ๆ ที่ใช้ในการค้น ควรมีความเป็นสากล และใช้ง่าย ซึ่งผู้ใช้สามารถศึกษาฟังก์ชันการค้น และการทำงานของ Search Engines ได้จาก ส่วนของการให้ความช่วยเหลือ (Help, Instruction ฯลฯ) ของ Search Engine หรือจากเว็บไซต์ที่ให้ข้อมูลเกี่ยวกับ Search Engines เช่น

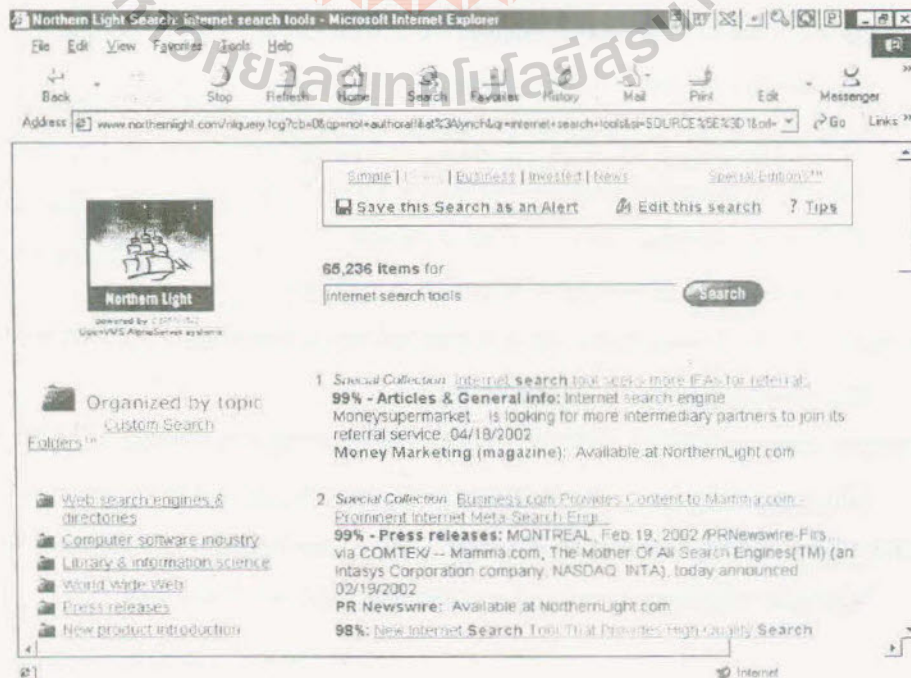
- Search Engine Watch (<http://www.searchenginewatch.com>)
- Search Engine Showdown (<http://www.searchengineshowdown.com>)
- Search Engine Guide (<http://www.searchengineguide.com>)
- Info People (<http://www.infopeople.com>)

ตัวอย่าง ตารางแสดงฟังก์ชันของ Search Engines ดูที่ภาคผนวก

5. คุณลักษณะด้านการแสดงผลการค้น (Results List Display Features)

ประเด็นที่ต้องพิจารณาในส่วนของการแสดงผลการค้น ได้แก่

- รายละเอียดที่แสดง การแสดงรายละเอียดของผลการค้น ควรที่จะให้ข้อมูลเพียงพอแก่การตัดสินใจของผู้ใช้ในการ เลือกเข้าไปดูเนื้อหาเต็มของเว็บนั้น ๆ และควรแสดงอย่างเป็นระบบ และมีมาตรฐานในการจัดทำ เช่น ใช้มาตรฐานของ Dublin Core Metadata Element Sets รายละเอียดที่ควรแสดง อาทิ ชื่อเอกสาร ชื่อ URL บีบีพิมพ์ หรือปีที่จัดทำเอกสาร ภาษา รูปแบบของแฟ้มข้อมูล ขนาดของแฟ้มข้อมูล ซอฟต์แวร์ที่ต้องใช้ในการแสดงผลข้อมูล บรรณวิทศน์ การแสดงการเชื่อมโยงไปยัง ข้อมูลอื่นที่เกี่ยวข้องสัมพันธ์กัน การแสดงเปอร์เซ็นต์ความเกี่ยวข้องของเอกสารกับคำค้น จำนวนผลการค้นที่ได้ เป็นต้น
- การเลือกแสดงผลข้อมูลตามรูปแบบที่ต้องการ เช่น การเลือกจำนวนผลการค้นที่ต้องการแสดงใน 1 หนึ่งหน้า การเรียงลำดับผลการค้น เช่น เรียงตามลำดับอักษรชื่อ เอกสาร เรียงตามการแสดงผลเปอร์เซ็นต์ความเกี่ยวข้องของเอกสารกับคำค้น หรือเรียงตามปีที่จัดทำ เป็นต้น



ตัวอย่างหน้าจอการแสดงผลของ Northern Light ที่ด้านบนก่อนแสดงรายการที่ค้นได้จะแสดงจำนวนผลการค้นหาได้ แสดง Search Box เพื่อให้ผู้ใช้สามารถสร้างสูตรการค้นใหม่ได้ ที่ด้านซ้ายของหน้าจอ จะแสดง Folder ของการแบ่งกลุ่มข้อมูลให้มีความเฉพาะยิ่งขึ้น หรือ แบ่งตามประเด็นที่เห็นว่าเกี่ยวข้อง ที่ด้านขวาของหน้าจอจะแสดงรายการข้อมูลที่ค้นได้ ตามลำดับเปอร์เซ็นต์ความเกี่ยวข้องกับคำค้นหาที่ใช้ โดยตัวเลขเปอร์เซ็นต์ จะแสดงเป็นตัวหนาไว้ด้านหน้า ชื่อเรื่อง ปีที่ผลิต (ถ้ามี) รายละเอียดโดยย่อ และ highlight ข้อความที่ตรงกับคำศัพท์ที่ใช้ค้น

6. คุณลักษณะพิเศษอื่น ๆ (Other Features of Note)

คุณลักษณะพิเศษของ Search Engines ได้แก่ คุณลักษณะพิเศษ อื่น ๆ ที่ไม่มีในข้อ 1-5 หรือ บริการพิเศษอื่น ๆ เช่น การแสดง Directory เพื่อช่วยเหลือผู้ใช้ในการกำหนดหัวข้อ หรือประเด็นที่ชัดเจนขึ้น การจัดกลุ่มให้กับผลการค้นโดยแยกตามประเด็นความเกี่ยวข้องของเรื่องที่มีความเฉพาะเจาะจงขึ้น การสร้างสูตร การค้นใหม่จากสูตรการค้นหรือผลการค้นเดิม หรือผลการค้นปัจจุบัน การแสดงรายชื่อหนังสือที่เกี่ยวข้องของสำนักพิมพ์ต่าง ๆ การแสดงรายการเชื่อมโยงไปยังกลุ่มของผลการค้นที่มีลักษณะคล้ายกัน เป็นต้น

บรรณานุกรม

- กาญจนา ใจกว้าง. (2541). "หน่วยที่ 11 คอมพิวเตอร์กับงานบริการค้นคืนสารสนเทศ." ใน เอกสารประกอบ การสอนชุดวิชาเทคโนโลยีสารสนเทศเบื้องต้น หน่วยที่ 9-15. หน้า 103-148. นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- Bar-Ilan, Judit. (2002). Methods for Measuring Search Engine Performance Over Time. *Journal of American Society For Information Science And Technology* 53(4): 308-319.
- Bradley, Phil. (2002). *The Advanced internet searcher's handbook*. 2nd. ed. London: Library Association.
- Infopeople. (2002). Search Tools Chart. [Online]. Available: <http://www.infopeople.org/search/chart.html>
- Sullivan, Danny. (2001). How Search Engines Rank Web Pages. [Online]. Available: <http://searchenginewatch.com/webmasters/rank.html>
- _____. (2001). How Search Engines Wrok. [Online]. Available: <http://searchenginewatch.com/webmasters/work.html>.
- _____. (2001). Search Engine Sizes. [Online]. Available: <http://www.searchenginewatch.com/reports/sizes.html>.
- Tyner, Ross. (2002). Sink or Swim: Internet Search Tools & Techniques. [Online]. Available: <http://www.ouc.bc.ca/libr/connect96/search.htm>.
- University of California Berkeley Library. (2002). The BEST Search Engines. [Online]. Available: <http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/SearchEngines.html>.
- Westera, Gillian. (2002). Comparison of Search Engine User Interface Capabilities. [Online]. Available: <http://lisweb.curtin.edu.au/staff/gwpersonal/compare.html>.

แบบฝึกหัด

แบบฝึกหัดที่ 1 : Search Engines

เลือกหัวข้อจากรายการกำหนดให้ด้านล่าง 1 หัวข้อ ใช้ Search Engines อย่างน้อย 2 ตัวเพื่อค้นเรื่องดังกล่าว แล้วเปรียบเทียบผลการค้นหาที่ได้จาก Search Engines ทั้งหมดที่ใช้ค้น โดยเปรียบเทียบในประเด็นต่อไปนี้

- ความยากง่ายในการค้น
- มีการให้ความช่วยเหลือในการสร้างสูตรการค้นหาหรือไม่ เช่น ข้อความแนะนำ Help เป็นต้น
- จำนวนผลการค้นหาได้
- สัดส่วนของผลการค้นหาที่ได้กับเอกสารที่ตรงกับความต้องการความต้องการ คิดเป็นกี่เปอร์เซ็นต์
- ความทันสมัยของข้อมูล
- รายละเอียดที่ให้ในแต่ละรายการที่ค้นหาได้เพียงพอหรือไม่
- ลำดับการจัดเรียงข้อมูลใช้วิธีใด มีหลักในการจัดเรียงหรือไม่ ถ้ามีจัดเรียงแบบใด เช่น เรียงตามวันที่ หรือ จัดเรียงตามเปอร์เซ็นต์ความเกี่ยวข้องของเอกสารกับคำค้น เป็นต้น
- ฟังก์ชันพิเศษที่ช่วยในการค้นหา Search Engines อื่นไม่มี

ประเด็นในการค้น

Alqueida	Human Clonning	Internet Search Tools
----------	----------------	-----------------------

Search Engines

Altavista	Excite	Alltheweb	Google	Hotbot
-----------	--------	-----------	--------	--------

แบบฝึกหัดที่ 2 : Multi-Search Engines

ใช้ Multi-Search Engines เพื่อค้นหาสารสนเทศในประเด็นเดียวกับแบบฝึกหัดที่ 1 จากนั้นเปรียบเทียบผลการค้น เช่นเดียวกับแบบฝึกหัดที่ 1

Multi-Search Engines

Dogpile	Ixquick	Metacrawler	ProFusion	SurfWax
---------	---------	-------------	-----------	---------

แบบฝึกหัดที่ 3 : Subject Directories

3.1 Browse ดูการจัดกลุ่มข้อมูลของ Yahoo (โดยไม่ใช้ฟังก์ชันในการค้น) เพื่อค้นหาข้อมูลในเรื่องเดียวกับข้อ 1 หลังจากได้ผลการค้นแล้ว (สมมติว่าทุกคนค้นข้อมูลออกมาได้) ให้ประเมินในประเด็นต่อไปนี้

- ระหว่างการค้นด้วย Search Engine กับ Subject Directory นักศึกษาคิดว่า วิธีใดให้ผลการค้นที่ดีกว่า เพราะเหตุใด
- ข้อดีและข้อด้อยของแต่ละวิธีมีอะไรบ้าง

3.2 ใช้ Search Box ของ Yahoo ค้นเรื่องเดียวกับข้อข้อ 3.1 แล้วตอบคำถามด้านล่าง

- วิธีการการค้นข้อมูลจาก Yahoo วิธีการใดประสบผลสำเร็จในการค้นมากกว่ากัน ทำไมจึงเป็นเช่นนั้น
- การค้นข้อมูลจาก Yahoo แตกต่างจากการค้นด้วย Search Engines ทั่วไปอย่างไร

3.3 Browse และ Search ข้อมูลจาก Specialized Search Directory ในหัวข้อเดียวกับ 3.1 โดยเลือก Search Engine ที่เหมาะสมกับเรื่องที่ค้น สังเกตผลการค้นที่ได้ ตรงกับความต้องการของนักศึกษามากกว่า Subject directory ทั่วไปหรือไม่

Subject Directory

Argus Clearinghouse	About.com	WWW Virtual Library
---------------------	-----------	---------------------

แบบฝึกหัดที่ 4 : ทดสอบทักษะการค้นของตนเอง

เลือกหัวข้อที่นักศึกษาสนใจจากรายการด้านล่างมา 1 หัวข้อ ใช้เครื่องมือช่วยค้นในข้อ 1-3 เพื่อค้นหาข้อมูลในประเด็นดังกล่าว โดยพยายามทำสิ่งต่อไปนี้ในการค้น

- ใช้การค้นทั้งแบบ Simple Search และ Advanced Search
- อ่านคำแนะนำในการค้นก่อนดำเนินการค้น
- ใช้เทคนิคต่อไปนี้ในการค้น Boolean logic, Proximity operators, Wildcard characters และ Phrase searching
- ใช้ "more like this" จากผลการค้นที่ได้ เพื่อสร้างการค้นครั้งใหม่
- ทดลองปรับเปลี่ยนการแสดงผลการค้นหลาย ๆ แบบ รวมทั้งการจัดเรียงผลการค้น

ประเด็นในการค้น

1. Mad Cow Disease กับ Creutzfeldt-Jakob Syndrome มีความสัมพันธ์กันอย่างไร
2. เป็นไปได้หรือไม่ที่ Hard drive ของคอมพิวเตอร์ของคุณจะติดไวรัสจากการ Run โปรแกรมที่ download จากอินเทอร์เน็ต
3. มีข้อมูลเกี่ยวกับการทำสาโทบนอินเทอร์เน็ตมีหรือไม่
4. นักศึกษากำลังจะไปดูภาพยนตร์เรื่อง Lord of the ring ภาค 2 และต้องการข้อมูลเกี่ยวกับการวิจารณ์ภาพยนตร์เรื่องนี้ นวนิยายต้นฉบับ ชีวิตประวัติของผู้กำกับ

บทที่ 9

ระบบการเชื่อมประสานกับผู้ใช้สำหรับการค้นคืนสารสนเทศ (User Interfaces for Information Retrieval)

"Well designed, effective computer systems generate positive feelings of success, competence, mastery, and clarity in the user community. When an interactive system is well-designed, the interface almost disappears, enabling users to concentrate on their work, exploration, or pleasure." (Shneiderman, 1997: p.10)

การศึกษาระบบการเชื่อมประสานกับผู้ใช้ มีวัตถุประสงค์หลักเพื่อ ออกแบบระบบการค้นคืนสารสนเทศให้ง่ายต่อการใช้งานมากที่สุด และเป็นระบบที่เรียนรู้ได้ง่าย เพื่อให้ผู้ใช้สารสนเทศทุกคนสามารถค้นคืนสารสนเทศได้ด้วยตนเอง และสามารถแก้ไขปัญหาที่เกิดขึ้นได้ด้วยตนเอง ซึ่งระบบการเชื่อมประสานที่ดีจะช่วยให้ผู้ใช้สามารถควบคุมการทำงานของระบบได้อย่างราบรื่น ดังคำกล่าวของชไนเดอร์แมนข้างต้น

ระบบการเชื่อมประสานกับผู้ใช้ มาจากคำภาษาอังกฤษว่า User Interfaces หมายถึง การติดต่อสื่อสารระหว่างคอมพิวเตอร์กับผู้ใช้ ขณะใช้งานระบบ หรือระหว่างการค้นคืนข้อมูลจากระบบคอมพิวเตอร์ ประกอบด้วยองค์ประกอบหลัก 2 ประการ คือ มนุษย์ซึ่งเป็นผู้ใช้งานระบบ และคอมพิวเตอร์ซึ่งทำหน้าที่ตอบสนองความต้องการของผู้ใช้ โดยติดต่อสื่อสารผ่านอุปกรณ์รับและแสดงผล เช่น แป้นพิมพ์ หน้าจอ จอสัมผัส เป็นต้น

แกรี เอส. ลอว์เรนซ์ (Garry S. Lawrence) โจเซฟ อาร์ แมตทิวส์ (Joseph R. Mathews) และ ชาร์ลส อี. มิลเลอร์ (Charls E. Miller) ได้ให้คำจำกัดความที่ค่อนข้างชัดเจนของระบบการเชื่อมประสานกับผู้ใช้ในความหมายของการค้นคืนสารสนเทศแบบออนไลน์ว่า หมายถึง การติดต่อสื่อสารระหว่างผู้ใช้กับระบบรายการสารสนเทศออนไลน์ ประกอบด้วยวิธีการค้น รูปแบบการแสดงผล การให้คำแนะนำและการสอนและหน้าจอให้คำแนะนำช่วยเหลือ ระบบการเชื่อมประสานจะเป็นจุดเชื่อมต่อระหว่างผู้ใช้กับกลไกการค้น โดยมีโปรแกรมการเชื่อมประสาน (Interface Software) เป็นตัวกลางนำคำสั่งจากผู้ใช้ไปสู่กระบวนการทำงานของระบบคอมพิวเตอร์ และแปลงผลการค้นออกมาเป็นรูปแบบที่ผู้ใช้สามารถเข้าใจได้ อย่างไรก็ตาม การเปลี่ยนแปลงแก้ไขระบบการเชื่อมประสานกับผู้ใช้จะไม่ส่งผลต่อโครงสร้างพื้นฐานของระบบสารสนเทศออนไลน์ (Lawrence, Mathews and Miller, 1983: p. 413)

การศึกษาในเรื่อง ระบบการเชื่อมประสานกับผู้ใช้ (User Interfaces : UI) มีความใกล้เคียงกับการศึกษาเรื่อง Human-Computer Interaction : HCI ซึ่งมุ่งศึกษาเพื่อพัฒนาระบบคอมพิวเตอร์เพื่อให้มนุษย์สามารถใช้งานคอมพิวเตอร์ได้ง่าย ขณะที่ UI จะเน้นที่การติดต่อสื่อสารกับคอมพิวเตอร์ ผ่านทางการมองเห็น การได้ยิน การใช้คำสั่งเพื่อควบคุมการทำงาน ขณะที่ใช้งานระบบคอมพิวเตอร์

ซึ่งการศึกษาด้าน UI และ HCI มีความครอบคลุมในหลายสาขาวิชาด้วยกัน ได้แก่ วิทยาการคอมพิวเตอร์ (Computer science), จิตวิทยา (Psychology), สารสนเทศศาสตร์ (Information science), สังคมวิทยา (Sociology), การออกแบบกราฟิก (Graphic design), การจัดการ (Management), และการศึกษา (education) โดยการออกแบบระบบจะเน้นที่ผู้ใช้เป็นศูนย์กลาง (User-centered systems design) ทั้งนี้เพื่อให้

ผู้ใช้ทุกกลุ่มสามารถใช้งานคอมพิวเตอร์และระบบสารสนเทศเพื่อตอบสนองความต้องการใช้งานของตนเองได้ ประเภทของผู้ใช้ ได้แก่ ผู้ที่มีความเชี่ยวชาญในการใช้งานคอมพิวเตอร์หรือระบบสารสนเทศ (Expert Users) กลุ่มที่มีความคุ้นเคยกับการใช้งานคอมพิวเตอร์หรือระบบสารสนเทศในระดับปานกลาง (Experienced Users) และกลุ่มที่ไม่มีประสบการณ์ในการใช้ หรือไม่เคยใช้งานระบบเลย (Novice Users)

ประเภทของระบบการเชื่อมประสานกับผู้ใช้

เบน ไชนเดอร์แมน (Ben Shneiderman) ได้จำแนกประเภทของระบบการเชื่อมประสานกับผู้ใช้ ออกเป็น 5 ประเภท (Shneiderman, 1987: p. 57) ดังนี้

1. ระบบเมนูทางเลือก (Menu Selection)
2. ระบบเติมคำ (Form Fill-in)
3. ระบบภาษาคำสั่ง (Command Language)
4. ระบบภาษารธรรมชาติ (Natural Language)
5. ระบบการปรับแต่งโดยตรง (Direct Manipulation)

1. ระบบเมนูทางเลือก (Menu Selection)

ระบบเมนูทางเลือก ออกแบบมาเพื่อแก้ปัญหาการจดจำภาษาคำสั่ง ซึ่งบางคำมีรูปประโยคที่ยาว และจำยาก ในการใช้ระบบนี้ผู้ใช้สามารถเลือกวิธีการทำงานที่เหมาะสมที่สุดจากรายการทางเลือกที่ระบบนำเสนอ การแสดงรายการทางเลือกอาจใช้ข้อความสั้น ๆ รูปภาพ หรือ ไอคอน (Icons) ดังนั้น ผู้ใช้จึงไม่จำเป็นต้องมีความรู้เกี่ยวกับความสามารถของระบบ การเลือกรายการจากเมนู อาจใช้กรพุ่มพ์รหัสบางตัว (เช่น หมายเลขหรือตัวอักษร) ที่สัมพันธ์กับรายการที่ต้องการ โดยใช้เคอร์เซอร์คีย์ (Cursor Keys) เมาส์ (Mouse) จอยสติค (Joystick) จอสัมผัส (Touch screen) หรืออุปกรณ์ชี้ (Pointing Devices) อื่น ๆ เพื่อเลือกรายการที่ต้องการ ตัวเลือกในเมนูควรใช้คำที่มีความหมายชัดเจน เพื่อประหยัดเวลาในการทำความเข้าใจ รูปแบบโครงสร้างของเมนูควรมลักษณะที่เหมือนกันทั้งระบบ และไม่ควรนำเสนอทางเลือกน้อย หรือมากเกินไป โดยปกติในหนึ่งหน้าจอไม่ควรนำเสนอทางเลือกน้อยกว่า 4 ทางเลือก และไม่ควรเกิน 12 ทางเลือก เพื่อประหยัดเวลาในการค้น และเพื่อสามารถนำเสนอบนหน้าจอได้อย่างชัดเจน (Downton, 1991: p. 89) เมนูทางเลือกจึงเป็นระบบการเชื่อมประสานกับผู้ใช้ จึงเหมาะกับผู้ใช้ที่ไม่มีประสบการณ์หรือมีประสบการณ์น้อยในการใช้งานระบบ

2. ระบบเติมคำ (Form Fill-in หรือ Form-Filling หรือ Fill-in-the-blank)

หน้าจอของระบบคอมพิวเตอร์จะแสดงแบบกรรายการ เพื่อให้ผู้ใช้ป้อนข้อมูลที่จำเป็นในการทำงาน ในหน้าจอจะประกอบด้วย ชื่อเขตข้อมูลและช่องว่างให้ผู้ใช้พิมพ์ข้อมูล ดังนั้นในระบบการเชื่อมประสานแบบเติมคำนี้ ผู้ใช้ต้องเข้าใจชื่อเขตข้อมูล ทราบค่าที่จะสามารถพิมพ์เข้าไปได้ และทราบวิธีการใส่ข้อมูล เพื่อที่จะสามารถตอบได้กับการทำงานของระบบได้ ผู้ใช้จึงต้องได้รับการอบรมการใช้งานระบบมาก่อน หรือต้องมีความรู้เกี่ยวกับการทำงานของระบบมาบ้าง การเชื่อมประสานแบบเติมคำทำงานได้เร็วกว่าระบบเมนูเนื่องจากไม่ถูกจำกัดขั้นตอนการทำงาน แต่จะตอบสนองการทำงานได้ช้ากว่าและใช้พื้นที่บนหน้าจอเปลืองกว่าระบบภาษาคำสั่ง ข้อควรระวังในการออกแบบระบบคือ ควรป้องกันไม่ให้เกิดการเขียนทับชื่อเขตข้อมูล ระบบนี้เหมาะกับผู้ใช้ทุกระดับประสบการณ์ที่ได้รับการอบรมการใช้งานระบบมาแล้ว

3. ระบบภาษาคำสั่ง (Command Language)

เป็นระบบที่ผู้ใช้สามารถเลือกใช้คำสั่งต่าง ๆ เพื่อให้ระบบปฏิบัติตามคำสั่งได้โดยตรง ผู้ใช้สามารถค้นข้อมูลได้ทั้งในระดับกว้างและระดับลึกที่มีความเฉพาะเจาะจงได้อย่างรวดเร็ว ให้ความต่อเนื่องในการทำงาน ทำให้ผู้ใช้รู้สึกว่าเป็นผู้ควบคุมการทำงานของระบบ ระบบจะไม่แสดงคำสั่ง แต่อาจให้ความช่วยเหลือในลักษณะของการแสดงข้อความพร้อมรับ (Prompt) ดังนั้น ในระบบภาษาคำสั่งผู้ใช้ต้องเรียนรู้และจดจำการใช้คำสั่ง ระบบนี้จึงเหมาะกับผู้ใช้ที่มีความชำนาญหรือผู้ใช้ประจำ อย่างไรก็ตาม ระบบการเชื่อมประสานแบบภาษาคำสั่งมักเกิดความผิดพลาดได้สูง ผู้ใช้จึงจำเป็นต้องได้รับการฝึกหัดการใช้งานระบบ และระบบควรจัดเตรียมข้อความแสดงผิดพลาดและให้คำแนะนำช่วยเหลือทางออนไลน์ไว้ให้ครอบคลุมมากที่สุด ทั้งนี้เนื่องจากความผิดพลาดสามารถเกิดขึ้นได้หลายลักษณะ

4. ระบบภาษาธรรมชาติ (Natural Language)

เป็นระบบที่ผู้ใช้ติดต่อกับระบบโดยใช้ภาษาอังกฤษที่ใช้พูดในชีวิตประจำวัน แนวคิดในการใช้ภาษาธรรมชาติเกิดจากปัญหาในการใช้ภาษาคำสั่งที่มีความยุ่งยาก ดังนั้น ผู้ใช้ที่ไม่มีประสบการณ์จะสามารถติดต่อกับคอมพิวเตอร์ได้อย่างรวดเร็ว เนื่องจากไม่ต้องใช้รูปแบบของประโยคที่กำหนดไว้ตายตัว และผู้ใช้จะได้รับคำตอบที่สามารถเข้าใจได้โดยไม่ต้องได้รับการฝึกหัดการใช้งานระบบเป็นพิเศษ อย่างไรก็ตาม แม้ว่าระบบภาษาธรรมชาติได้แสดงให้เห็นถึงความเป็นไปได้ในการติดต่อกับคอมพิวเตอร์โดยใช้ภาษาที่ใช้อยู่ในชีวิตประจำวัน แต่ในระบบคอมพิวเตอร์ การที่จะให้คอมพิวเตอร์เลียนแบบความไม่แน่นอนและความคลุมเครือของรูปแบบการใช้คำและรูปแบบของประโยคที่ใช้อยู่ในภาษาประจำวันยังคงเป็นสิ่งที่ต้องพัฒนาต่อไป การออกแบบระบบภาษาธรรมชาติจะครอบคลุมถึง

4.1 การเชื่อมประสานแบบระบบผู้เชี่ยวชาญและปัญญาประดิษฐ์ (Expert and Intelligent Interface) ซึ่งเป็นระบบที่ผู้ใช้ไม่จำเป็นต้องบอกข้อมูลที่ถูกต้อง หรือไม่จำเป็นต้องอธิบายรายละเอียดให้ทราบ แต่ระบบจะสามารถดำเนินการตามความต้องการได้

4.2 ระบบการเชื่อมประสานแบบใช้เสียงพูด (Voice-base interface) ซึ่งระบบการเชื่อมประสานจะออกมาในลักษณะที่คอมพิวเตอร์กับมนุษย์โต้ตอบกันโดยใช้เสียงพูด (Voice to Voice) หรือคอมพิวเตอร์เป็นฝ่ายพูดขณะที่มนุษย์ได้ตอบทางเครื่องพิมพ์ (Voice to Keyboard)

5. ระบบการปรับแต่งโดยตรง (Direct Manipulation)

เป็นระบบการเชื่อมประสานที่ผู้ใช้สามารถเข้าไปปรับแต่งสิ่งที่ปรากฏบนหน้าจอได้โดยตรง การทำงานหรือคำสั่งที่แสดงบนหน้าจออาจแทนด้วยไอคอนและใช้เมาส์เป็นอุปกรณ์เลือกชี้ไปยังไอคอนที่ต้องการปรับแต่งข้อมูล ทำให้ผู้ใช้สามารถทำงานได้ง่ายและสะดวกรวดเร็วขึ้นและสามารถดูผลการทำงานได้ทันทีรูปแบบของผลการทำงานที่ได้รับจะเหมือนกับรูปแบบที่แสดงทางหน้าจอ ข้อจำกัดของระบบการเชื่อมประสานแบบนี้คือ การกำหนดไอคอนที่ชัดเจนเพื่อสื่อความหมายแทนการทำงาน โดยเฉพาะการทำงานที่มีลักษณะเป็นนามธรรม เช่น คำสั่งให้ตรวจสอบตัวสะกด เป็นต้น เทคนิคที่นำมาใช้กับระบบการเชื่อมประสานแบบนี้ อาทิ WIMP (Window, Icon, Menu, Pointer), GUI (Graphic User Interface), Hypertext, Hypermedia เป็นต้น ระบบการเชื่อมประสานแบบปรับแต่งโดยตรง เหมาะกับผู้ใช้ที่ไม่มีประสบการณ์ในการใช้งานระบบ โดยผู้ใช้สามารถเรียนรู้การใช้งานระบบได้อย่างรวดเร็วจากการทดลองใช้ สามารถพิจารณาข้อผิดพลาดที่เกิดขึ้นและสามารถแก้ไขได้ด้วยตนเอง

จากระบบการเชื่อมประสานกับผู้ใช้ทั้ง 5 ประเภทสามารถจำแนกข้อดีข้อด้อยของระบบการเชื่อมประสานกับผู้ใช้แต่ละประเภท (Rowley, 1990: p. 104) ได้ดังนี้

ประเภท	ข้อดี	ข้อด้อย
1. ระบบเมนูทางเลือก (Menu Selection)	<ul style="list-style-type: none"> - เรียนรู้ได้เร็ว - ไม่ต้องจำคำสั่ง - ลดจำนวนการกดแป้นพิมพ์ - มีการจัดโครงสร้างคำสั่ง - มีโครงสร้าง หรือรูปแบบในการตัดสินใจ - แก้ไขข้อผิดพลาดได้ง่าย - เหมาะกับผู้เริ่มใช้ 	<ul style="list-style-type: none"> - อาจทำให้ผู้ใช้ที่มีประสบการณ์มากหรือเชี่ยวชาญในการใช้งานระบบทำงานได้ช้า - ใช้พื้นที่บนหน้าจอมาก - ต้องการอัตราการแสดงผลที่รวดเร็ว - ระบบขนาดใหญ่การทำงานช้า - ต้องจำกัดตัวเลือกในแต่ละเมนู
2. ระบบเติมคำ (Form Fill-in)	<ul style="list-style-type: none"> - มีการจัดโครงสร้าง - ไม่ต้องจำคำสั่งมาก - ใส่ข้อมูลง่าย - เรียนรู้ได้ง่าย - เหมาะกับผู้ใช้ทุกประเภท 	<ul style="list-style-type: none"> - ใช้พื้นที่บนหน้าจอมาก - ต้องมีตัวควบคุมการเคลื่อนที่ (Cursor Control) - ต้องการการอบรมการใช้ - แบบฟอร์มเหมาะกับการบันทึกข้อมูล
3. ระบบภาษาคำสั่ง (Command Language)	<ul style="list-style-type: none"> - มีความยืดหยุ่น - เข้าถึงกิจกรรมที่ต้องการได้โดยตรง - ทำงานได้เร็ว - เชื้อต่อการตัดสินใจของผู้ใช้ - เหมาะกับการทำงานที่ซับซ้อน - เหมาะกับผู้ใช้ที่มีทักษะและประสบการณ์ 	<ul style="list-style-type: none"> - ต้องจดจำคำสั่ง - มักเกิดข้อผิดพลาดในการค้น - ต้องใช้เวลาฝึกการใช้ - ต้องอาศัยทักษะที่เกิดจากการฝึกการใช้ อย่างสม่ำเสมอสำหรับผู้เริ่มใช้
4. ระบบภาษาธรรมชาติ (Natural Language)	<ul style="list-style-type: none"> - ลดภาระของผู้ใช้ในการเรียนรู้รูปประโยค - ไม่ต้องใช้รูปประโยคเฉพาะ - มีความยืดหยุ่น - สื่อสารกันได้อย่างเป็นธรรมชาติ - เหมาะกับผู้เริ่มใช้ 	<ul style="list-style-type: none"> - ต้องพิมพ์ตัวอักษรหลายตัว - ความกำกวมของภาษาและ คำศัพท์เขียนโปรแกรมยาก - ต้องการพื้นฐานความรู้ภาษาอังกฤษ - ฮาร์ดแวร์ต้องมีประสิทธิภาพสูง
5. ระบบการปรับแต่งโดยตรง (Direct Manipulation)	<ul style="list-style-type: none"> - เรียนรู้ง่าย - หลีกเลี่ยงข้อผิดพลาดได้ง่าย - นำเสนอแนวคิดการทำงานแบบเห็นภาพ - ดึงดูดความสนใจจากผู้ใช้ได้ดี 	<ul style="list-style-type: none"> - เขียนโปรแกรมยาก - กำหนดภาพสัญลักษณ์ที่ชัดเจนได้ยาก - ฮาร์ดแวร์ต้องมีประสิทธิภาพสูง

คุณลักษณะของระบบการเชื่อมประสานที่ใช้งานง่าย

การจะพิจารณาระบบการเชื่อมต่อเป็นระบบที่ใช้งานได้ง่าย ขึ้นอยู่กับผู้ใช้เป็นผู้พิจารณา เนื่องจากสิ่งที่ผู้สร้างระบบคิดว่าง่ายแต่สำหรับผู้ใช้อาจเป็นสิ่งที่ยาก อย่างไรก็ตามมาตรฐานของระบบที่ใช้งานง่ายไม่อาจกำหนดได้แน่นอน เนื่องจากจะปรับเปลี่ยนไปตามแนวความคิดของมนุษย์ และพัฒนาการของเทคโนโลยี คอมพิวเตอร์ทั้งฮาร์ดแวร์และซอฟต์แวร์ ณ เวลานั้น แต่ในการออกแบบระบบการเชื่อมประสานนั้น นักออกแบบมักคำนึงถึงหลักการพื้นฐาน 2 ประการ คือ ใช้งานและดึงดูดใจ (Shiao-Feng, 1994: p. 144) และคำนึงถึงความสำเร็จในการตอบคำถามมากกว่าความสำเร็จในการค้นหาข้อมูลที่ตรงกับคำค้น (Borgman, 1996: p. 500) โดยทั่วไประบบการเชื่อมประสานที่ดีควรออกแบบให้ผู้ใช้ที่มีทักษะ และประสบการณ์แตกต่างกันสามารถใช้งานได้ อย่างมีประสิทธิภาพใกล้เคียงกัน เบน ชไนเดอร์แมน ได้เสนอเกณฑ์ 5 ประการในการพิจารณาความเป็นระบบที่ใช้งานได้ง่าย (Shneiderman, 1987:p. 14-15) ไว้ดังนี้

1. เวลาที่ใช้ในการเรียนรู้ ได้แก่ เวลาที่ใช้เสียไปกับการเรียนรู้การใช้คำสั่งต่าง ๆ ที่เกี่ยวข้องกับการทำงาน หรือเวลาที่ใช้ในการเรียนรู้การใช้งานระบบ
2. ความเร็วในการทำงาน ได้แก่ เวลาที่ระบบใช้ในการทำงานมาตรฐาน 1 ชุด
3. อัตราความผิดพลาดที่เกิดจากผู้ใช้ ได้แก่ จำนวนและประเภทของความผิดพลาดที่เกิดขึ้นในการทำงานมาตรฐาน 1 ชุด
4. ความพึงพอใจของผู้ใช้ ได้แก่ ระดับความชอบของผู้ใช้ที่มีต่อการใช้งานระบบ สามารถทดสอบได้จากแบบสัมภาษณ์ หรือแบบสำรวจ ซึ่งประกอบด้วยมาตรวัดความพึงพอใจและการให้เขียนแสดงความคิดเห็น
5. ระยะเวลาที่จำรูปแบบการใช้งานได้ หลังจากเสร็จสิ้นการค้นคืนข้อมูล นั่นคือ ใช้จำรูปแบบการใช้งานระบบได้นานเพียงใด อาจเป็น 1 ชั่วโมง 1 วัน หรือ 1 สัปดาห์ เป็นต้น

คุณลักษณะของความเป็นระบบที่ใช้งานได้ง่ายสามารถสรุปเป็นตาราง (Aluri, Kemp and Boll, 1991:p. 227-228) ได้ดังนี้

คุณลักษณะ	คำอธิบาย
1. ใช้งานง่าย (Simplicity)	
1.1 มีคำสั่งน้อย (Few keywords)	- ใช้คำสั่งน้อย ใช้คำสั่งสั้นๆ หลีกเลี่ยงการใช้คำศัพท์เฉพาะกรณี
1.2 ใส่ข้อมูลง่าย (Simplicity of input)	- ใส่ข้อมูลได้เร็ว (ขึ้นอยู่กับลักษณะแป้นพิมพ์ และทักษะการพิมพ์ของผู้ใช้) มีลักษณะอักษรช่วยความจำ มีโครงสร้างการใส่ข้อมูลอย่างง่าย
1.3 ใช้คำสั่งสั้น ๆ (Short commands)	- ไม่ใช้คำซ้ำมาก หลีกเลี่ยงการใส่ข้อมูลหลาย ๆ ตัว หรือการใช้คำสั่งยาว ๆ
1.4 ใช้คำสั่งง่าย ๆ (Simple commands)	- มีโครงสร้างคำสั่งที่ง่าย ใช้ประโยคคำสั่งง่าย ๆ สามารถโต้ตอบกับระบบได้ง่าย ไม่ใช้คำฟุ่มเฟือย

คุณลักษณะ	คำอธิบาย
<p>2. ความชัดเจน (Clarity)</p> <p>2.1 โครงสร้างมีลำดับชั้น (Hierarchical structure)</p> <p>2.2 ฟังก์ชันการทำงาน (Functionality)</p> <p>2.3 มีความเป็นอันหนึ่งอันเดียวกัน (Homogeneity)</p> <p>2.4 การยอมรับปัญหา (Problem Orientation)</p>	<ul style="list-style-type: none"> - มีโครงสร้างการใช้คำสั่งแบบมีลำดับชั้น (มีลำดับคำสั่งหลัก คำสั่งย่อย) - มีการแยกประเภทคำสั่ง ไม่ใช่คำสั่งหลายคำสั่งในการทำงานลักษณะเดียวกัน และไม่ใช้คำสั่งเพียงคำสั่งเดียวกับการทำงานหลายหน้าที่ และมีคำสั่งที่ชัดเจนในการทำงานเฉพาะ - ทุกคำสั่งมีโครงสร้างเหมือนกัน คำสั่งที่ใช้มีความหมายเหมือนกันในบริบทเดียวกัน <p>สามารถยอมรับข้อจำกัด หรือข้อยกเว้นทางเทคนิคบางประการที่มีสาเหตุจากโครงสร้างข้อมูล การเขียนโปรแกรม ฯลฯ ทุกคำสั่งสามารถปฏิบัติการในระหว่างการทำงานได้</p>
<p>3. มีลักษณะเฉพาะตัว (Uniqueness)</p> <p>3.1 มีการกำหนดแน่นอน (Determinism)</p> <p>3.2 ไม่มีสถานการณ์คลุมเครือ (No undefined state)</p>	<ul style="list-style-type: none"> - ทุกคำสั่งมีความสมบูรณ์ในตัวเองสามารถปฏิบัติงานได้โดยตัวปฏิบัติการหรือทางเลือกที่ได้กำหนดไว้ - มีการกำหนดทางเลือกไว้จนกว่าผู้ใช้จะเปลี่ยนทางเลือกใหม่
<p>4. ใช้ภาษาง่าย ๆ (Comfortable language)</p> <p>4.1 คำสั่งมีประสิทธิภาพ (Powerful commands)</p> <p>4.2 มีความยืดหยุ่น (Flexibility)</p> <p>4.3 ใช้บทสนทนาสั้น ๆ (Short Dialog)</p> <p>4.4 ใช้ประโยชน์จากโครงสร้างได้เต็มที่ (Full use of data structure)</p>	<ul style="list-style-type: none"> - แสดงคำสั่งที่สามารถทำงานได้ภายใน 1 ขั้นตอน - คำสำคัญ (Keywords) ที่ใช้มีทั้งแบบสั้นและแบบยาว อาจมีได้หลายภาษา ทำงานได้ทั้งโดยตรงและโดยอ้อม ระบบสามารถปรับตัวเองเข้ากับความรู้อและประสบการณ์ของผู้ใช้ ทั้งผู้ใช้ชาคร ผู้ที่นาน ๆ มาใช้ครั้ง และผู้เชี่ยวชาญ ทำให้ผู้ใช้รู้สึกว่าจะสามารถ ควบคุมการทำงานของระบบได้ - คำสั่งมีความสมบูรณ์ในตัวเอง (ทั้งคำสั่งหลักและคำสั่งย่อย) สามารถใส่คำสั่งที่มีความต่อเนื่องกันได้ในคราวเดียว สามารถเข้าใจคำสั่งใหม่ ๆ ได้ - สามารถแสดงโครงสร้างข้อมูลและสามารถนำไปใช้ประโยชน์ในการค้นและตรวจดู (browsing) ได้

คุณลักษณะ	คำอธิบาย
<p>5. ความช่วยเหลืออื่น ๆ (Other comfort)</p> <p>5.1 การใส่ข้อมูล (Input comfort)</p> <p>5.2 การหยุดชั่วคราว (Interruption)</p> <p>5.3 ภาษาของผลการค้น (Output language)</p> <p>5.4 ความช่วยเหลือเพิ่มเติมอื่น ๆ (Additional comfort)</p>	<ul style="list-style-type: none"> - สามารถกลับไปแสดงข้อมูลนำเข้า หรือผลการค้นก่อนที่จะแก้ไขข้อผิดพลาด โดยใช้วิธีการเมนูหรือคำสั่ง - สามารถหยุดการทำงานไว้ชั่วคราวได้ตลอดเวลา - ชัดเจน สั้น เข้าใจข้อความของระบบ สามารถคาดเดาผลการค้นได้จากข้อมูลที่นำเข้า และสามารถนำผลการค้นไปเป็นข้อมูลนำเข้าได้ในกรณีที่เหมาะสม - มีฟังก์ชันคีย์ มีการให้สัญญาณเสียงหลังจากส่งผ่านข้อมูลแล้ว การเน้น และ/หรือ การขีดเส้นใต้ การจัดเรียงทางเลือกอย่างชัดเจน จัดแสดงผลการค้นได้ชัดเจน แสดงผล หรือตรวจดูไปข้างหน้าและย้อนหลังได้หลายวิธี มีเพิ่มข้อมูลของผู้ใช้ (User's notebook)
<p>6. การแจ้งสถานการณ์และการนำกลับมาใช้ใหม่ (Evidence and reusability)</p> <p>6.1 การแจ้งสถานการณ์ (Evidence)</p> <p>6.2 การให้ความช่วยเหลือ (Help function)</p> <p>6.3 การนำกลับมาใช้ (Reusability)</p>	<ul style="list-style-type: none"> - แจ้งขั้นตอนการทำงานของระบบ (เช่น Waiting for Input, Input, Waiting for output, Output) แสดงการรับคำสั่งไปปฏิบัติ ข้อความแจ้งความล่าช้าที่เกิดขึ้น การเตือนเกี่ยวกับการใช้คำสั่งที่ยู้งยาก หรือคำสั่งที่ต้องใช้ความสามารถมากขึ้นในการปฏิบัติ - ให้ข้อมูลเกี่ยวกับการทำงานของระบบ แสดงการทำงานที่กำลังใช้อยู่ แสดงฟังก์ชันการทำงานทั้งหมด แสดงโครงสร้างเนื้อหาของฐานข้อมูล แสดงการทำงานที่ผ่านมา แสดงคำสั่งที่สามารถดำเนินการได้ต่อไป แสดงโปรโตคอลที่ใช้ในการติดต่อ - สามารถนำคำสั่งเดิม หรือผลการค้นมาเป็นข้อมูลนำเข้า หรือคำสั่งใหม่ได้ สามารถเก็บคำสั่งไว้ใช้ในคราวต่อไปได้
<p>7. ความแน่นอนสม่ำเสมอ (Stability)</p> <p>7.1 การควบคุมความผิดพลาด (Error handling)</p> <p>7.2 ไม่มีสถานการณ์บังคับ (No compulsory situations)</p>	<ul style="list-style-type: none"> - แสดงข้อความที่ชัดเจนเพื่อป้องกันการใส่ข้อมูลผิดพลาด แก้ไขความผิดพลาดเล็ก ๆ น้อย ๆ ที่สามารถเกิดขึ้นได้ มีรูปแบบของการแสดงความผิดพลาด - ไม่บังคับให้ต้องดำเนินการไปตามวิธีการที่ถูกต้องกำหนดไว้สามารถหยุดการทำงานได้ตลอดเวลา
<p>8. ความปลอดภัยของข้อมูล (Data security)</p>	<ul style="list-style-type: none"> - ใช้ Password ที่แตกต่างกันในการเข้าถึงข้อมูลแต่ละระดับ การรักษาความปลอดภัยของข้อมูลบางส่วนต้องไม่ขัดขวางการใช้ข้อมูลส่วนที่เปิดให้ใช้ได้ (Open data)

ประเด็นการศึกษาระบบการเชื่อมประสานกับผู้ใช้

ประเด็นการศึกษาระบบการเชื่อมประสานกับผู้ใช้เพื่อการค้นคืนสารสนเทศในระบบออนไลน์ จะจำแนกการศึกษาออกเป็น 4 ประเด็น ดังนี้

1. การควบคุมการทำงาน (Operation Control)
2. การควบคุมสูตรการค้น (Search Formulation Control)
3. การควบคุมการแสดงผลการค้น (Output Control) และการจัดแสดงหน้าจอการแสดงผล
4. การช่วยเหลือในการใช้งานระบบ (Support Function)

1. การควบคุมการทำงาน (Operation Control)

การควบคุมการทำงาน ได้แก่ การเลือกใช้ ฟังก์ชันการทำงานต่าง ๆ เพื่ออำนวยความสะดวกระหว่าง การค้นคืน และเมื่อสิ้นสุดการค้นคืน รายการที่ควรตรวจสอบ ได้แก่

- การ login / logout
- การใช้ password เพื่อแสดงสิทธิการเข้าใช้
- วิธีการเลือกใช้คำสั่ง เช่น การเลือกใช้ตัวเลือกจากเมนูหลักและเมนูย่อย
- วิธีการพิมพ์คำสั่ง
- การเปลี่ยนตัวเลือกก่อนส่งคำสั่งเข้าไปในระบบ
- การเปลี่ยนภาษาในการค้น
- มีลำดับขั้นของการทำงานระบบ
- การใช้ฟังก์ชันคีย์ต่าง ๆ
- การหยุดการทำงานชั่วคราว
- การยกเลิกการทำงาน
- การย้อนการทำงานกลับไปยังขั้นตอนก่อนหน้า หรือการกระโดดไปยังขั้นตอนต่าง ๆ
- การแสดงสถานภาพการทำงาน
- เวลาที่ใช้ในการค้น / การแสดงผล
- เวลาที่ระบบใช้ในการโต้ตอบกับผู้ใช้

2. การควบคุมสูตรการค้น (Search Formulation Control)

การควบคุมสูตรการค้น ได้แก่ การใช้คำสั่งและกลวิธีต่าง ๆ ในการค้นคืน ทั้งเพื่อขยายผลการค้น และจำกัดผลการค้น รายการที่ควรตรวจสอบ ได้แก่

- การใช้ตรรกะบูลีน (Boolean Logic)
- การค้นคืนแบบการตัดคำ (Truncation) หรือการใช้ Wild Card
- การค้นคืนด้วยคำสำคัญ (Keyword search)
- การค้นคืนแบบสุ่ม (Browsing)
- การจำกัดการค้น
- การค้นคืนด้วยภาษาธรรมชาติ (Natural Language)

- การค้นคืนโดยใช้จุดเข้าถึงข้อมูลต่าง ๆ (Other access points)
- การเปลี่ยนคำค้นโดยไม่ต้องกลับไปเมนูหลัก (กรณีค้นคืนจากจุดเข้าถึงข้อมูลตัวเดียวกัน)

3. การควบคุมการแสดงผลการค้น (Output Control) และการจัดแสดงหน้าจอการแสดงผล

การควบคุมการแสดงผลการค้น ได้แก่ การเลือกรูปแบบหรือวิธีการในการแสดงผลการค้นคืน ประเด็นที่ควรตรวจสอบ ได้แก่

- การเลือกรูปแบบการแสดงผล แบบสมบูรณ์ แบบย่อ แบบMARC แบบ Dublin Core หรือรูปแบบพิเศษอื่น ๆ
- การเลือกแสดงผลเฉพาะเขตข้อมูลที่ต้องการ
- การจัดเรียงผลการค้นตามเขตข้อมูลต่าง ๆ
- การรวมการค้นแต่ละครั้งเป็นเพิ่มข้อมูลเดียวกัน เพื่อพิมพ์ผลการค้นออกทางเครื่องพิมพ์ หรือ save ลง diskette
- การแสดงผลไปข้างหน้าและย้อนหลัง
- การย้อนกลับไปยังตำแหน่งแรกที่แสดงผล
- การออกแบบหน้าจอการแสดงผลการค้น เพื่อความชัดเจนและรวดเร็วในการดูผลการค้น ประกอบด้วย
 - ◆ การจัดวางตำแหน่งของข้อความต่าง ๆ เช่น ชื่อของหน้าจอ คำค้นที่ใช้ จำนวนผลการค้น ผลการค้น ข้อความแสดงการสิ้นสุดผลการค้น ตำแหน่งการวางบรรทัดรายการคำสั่ง (Commands line) หรือเมนู เป็นต้น
 - ◆ ลำดับการจัดวางเขตข้อมูลในหน้าจอการแสดงผลแบบต่าง ๆ
 - ◆ การใช้สี และแถบเน้นข้อความ
 - ◆ รูปแบบและขนาดตัวอักษรที่แสดงชื่อของเขตข้อมูล (Label) และรายละเอียดข้อมูลในแต่ละเขตข้อมูล
 - ◆ ความสะอาด และ อ่านง่ายของหน้าจอการแสดงผล

4. การช่วยเหลือในการใช้งานระบบ (Support-Function)

การช่วยเหลือในการใช้งานระบบ ได้แก่ การเตรียมความช่วยเหลือขณะผู้ใช้งานกำลังใช้งานระบบ กรณีเกิดปัญหาในการใช้งาน โดยจะจัดทำในลักษณะของหน้าจอให้ความช่วยเหลือ (Help Screen) และ การแสดงข้อความให้ความช่วยเหลือต่าง ๆ เช่น Suggestion / Help Messages, Error Messages, Online Tutorials เป็นต้น สิ่งที่ต้องพิจารณาคือ

- ความสั้น กระชับ และชัดเจน ของข้อความต่าง ๆ
- การใช้น้ำเสียงที่สุภาพ ไม่แสดงถึงการตำหนิผู้ใช้
- รูปแบบการแสดงผล Help / Error Messages
- การจัดเตรียมความช่วยเหลือไว้หลายระดับ ตามระดับความเชี่ยวชาญของผู้ใช้
- การแสดงตัวอย่างในหน้าจอให้ความช่วยเหลือแต่ละประเภท
- การเปลี่ยนภาษาในหน้าจอให้ความช่วยเหลือ เมนู และบรรทัดแสดงคำสั่ง

บรรณานุกรม

- Aluri, Rao; Kemp, D. Alasdair and Ball, John J. (1991). *Subject Analysis in Online Catalogs*. Englewood, Colorado: Libraries Unlimited.
- Babu, B. Ramesh. (2000). Web OPAC interfaces: an overview. *The Electronic Library* 18(5).
- Borgman, Christine L. (1996). Why are Online Catalog Still Hard to Use?. *Journal of American Society for Information Science* 47(July): 493-503.
- Downton, Andy, ed. (1991). *Engineering the Human Computer Interface*. London: McGraw-Hill Book.
- Lawrence, Gary S.; Matthews, Joseph R. and Miller, Charles E. (1983). Cost and Features of Online Catalogs : The State of the Art. *Information Technology and Libraries* 2 (September): 409-449.
- Meadow, Charles T. (1992). *Text Information Retrieval Systems*. San Diego, California: Academic Press.
- Rowley, Jennifer E. (1990). *The Basic of System Analysis and Design for Information managers*. London: Clive Bingly.
- Shiao-Feng, Su. (1994). Dialog with an OPAC : How Visionary was Swanson in 1964?. *Library Quarterly* 64(April): 130-161.
- Shneiderman, Ben. (1987). *Designing the User Interface: Strategies for Effective Human-computer Interaction*. Massachusetts: Addison-Wesley.
- Shneiderman, Ben. (1997). *Designing the User Interface: Strategies for Effective Human-computer Interaction*. 3rd. Massachusetts: Addison-Wesley.
- Shneiderman, Ben; Byrd, Don; and Croft, W. Bruce. (1997). Clarifying Search : A User-Interface Framework for Text Searches. *D-Lib Magazine*. [Online]. Available: <http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html>.

บทที่ 10 การประเมินระบบการค้นคืนสารสนเทศ (Information Retrieval System: Measurement and Evaluation)

การประเมิน เป็นกระบวนการที่สำคัญในการจัดทำโครงการหรือกิจกรรมใดกิจกรรมหนึ่ง ทั้งนี้เพื่อทราบถึงสิ่งที่จะต้องแก้ไขปรับปรุง อุปสรรคปัญหาในการดำเนินงาน โดยผลจากการประเมินจะนำไปใช้เพื่อการตัดสินใจในการลงทุน หรือจัดทรัพยากรให้กับกิจกรรมนั้น ๆ สำหรับงานบริการ การประเมินมักจะมีจุดประสงค์ในการระบุจุดแข็งจุดอ่อน รวมทั้งความล้มเหลวของการดำเนินงาน และเพื่อเป็นการหาวิธีการที่จะพัฒนาปรับปรุงบริการต่าง ๆ เหล่านั้น (พิมพ์วิภาไพ, 2538 หน้า 84)

วัตถุประสงค์ของการประเมิน

โดยทั่วไปการประเมินมักมีวัตถุประสงค์หลัก 3 ข้อ คือ

- เพื่อตรวจสอบประสิทธิภาพของระบบสารสนเทศในปัจจุบัน
- เพื่อทราบแนวทางในการปรับปรุงระบบสารสนเทศในอนาคต
- เพื่อทราบความเป็นไปได้ในการปรับปรุงระบบสารสนเทศนั้นให้ดีที่สุด เพื่อที่จะสามารถเลือกอำนวยความสะดวกสูงสุดต่อผู้ใช้

รูปแบบการประเมินระบบการค้นคืนสารสนเทศ

1. การประเมินประสิทธิภาพของระบบ (Effectiveness) เป็นการประเมินความพึงพอใจของผู้ใช้ต่อการปฏิบัติงานเพื่อให้ได้สารสนเทศที่ต้องการ เช่นการประเมินความพึงพอใจของผู้ใช้ต่อความสามารถด้านการค้น การทำงานของระบบ และจำนวนผลการค้น เป็นต้น
2. การประเมินประสิทธิภาพคุ้มกับค่าใช้จ่าย (Cost-Effectiveness) เป็นการประเมินความพึงพอใจต่อผลการค้นกับค่าใช้จ่ายที่เกิดขึ้น ว่าประสิทธิภาพของงาน หรือผลของการค้นมีประสิทธิภาพคุ้มกับค่าใช้จ่ายที่เสียไปหรือไม่
3. การประเมินประโยชน์คุ้มค่า หรือ คุณค่าของบริการ (Cost-Benefits) เป็นการประเมินคุณค่าของระบบการค้นคืนสารสนเทศ โดยเป็นการพิจารณาระหว่างประโยชน์ที่ได้รับกับการลงทุนหรือการจัดให้มีบริการ ว่าประโยชน์ที่ได้รับคุ้มค่างกับการลงทุนหรือไม่ สมควรแก่การมีไว้ให้บริการหรือไม่ ซึ่งต้นทุนจะพิจารณาทั้งในส่วนของต้นทุนคงที่ (Fix cost) และ ต้นทุนแปรผัน (Variable cost)

ข้อมูลพื้นฐานที่จำเป็นต่อการประเมินผล

การประเมินจะกระทำได้อีกต่อเมื่อ ได้ข้อมูลเพียงพอแก่การวิเคราะห์ถึงความล้มเหลวหรือล้มเหลวของระบบการค้นคืนสารสนเทศนั้น ๆ ข้อมูลพื้นฐานที่จำเป็นต่อการประเมินผล ได้แก่

1. จุดประสงค์ในการค้นคืนสารสนเทศ ซึ่งจะมีผลต่อความพึงพอใจต่อผลการค้นที่ได้รับ
2. เวลาที่ใช้ในการปฏิบัติการ ได้แก่ ความเร็วในการทำงาน โดยเฉพาะในกรณีที่ต้องเสียค่าโทรคมนาคม
3. เวลาที่ใช้ในการเตรียมการก่อนการปฏิบัติการ ได้แก่ การวิเคราะห์ความต้องการ การคิดคำค้น การสร้างสูตรการค้น / กลวิธีการค้น การเลือกฐานข้อมูล

4. ค่าใช้จ่าย ได้แก่ ค่าใช้จ่ายคงที่ ค่าอุปกรณ์ ค่าสมาชิก และ ค่าใช้จ่ายไม่คงที่ : ค่าการสื่อสาร ค่าการจัดส่งเอกสาร
5. ข้อดี ข้อด้อย ของระบบการค้นคืนสารสนเทศ หรือ ฐานข้อมูล เช่น ฟังก์ชันในการค้น (Search features) บริการพิเศษ ต่าง ๆ เช่น : บริการสารสนเทศทันสมัย บริการจัดส่งเอกสาร เป็นต้น
6. การครอบคลุมหัวข้อเรื่องที่ทำการค้นทั้งในระดับกว้างและระดับลึก
7. จำนวนข้อมูลที่เรียกขึ้นมาได้ในหัวข้อเรื่องที่ทำการค้น (%)
8. จำนวนข้อมูลที่เรียกขึ้นมาได้ในหัวข้อเรื่องที่ทำการค้น (%) และตรงกับความต้องการของผู้ใช้
9. จำนวนเอกสารที่เรียกขึ้นมาและให้ความรู้ใหม่แก่ผู้ใช้
10. สรุปคุณค่าในการค้นคืนสารสนเทศแต่ละครั้งโดยทั่วไป

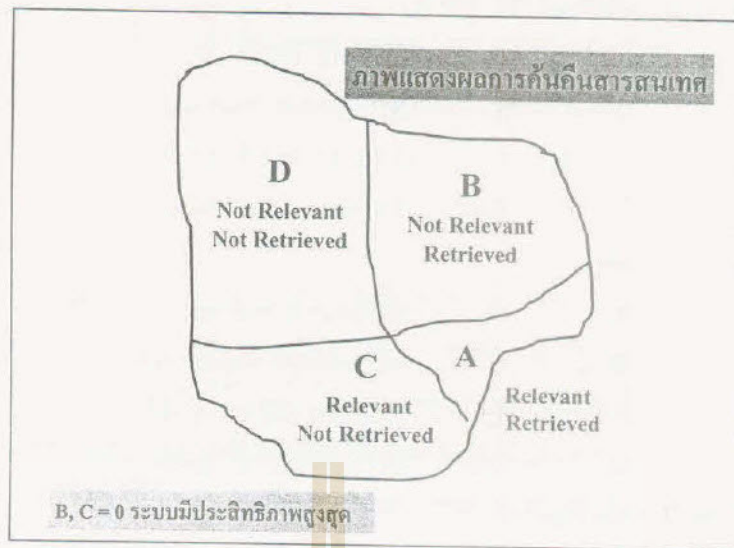
ส่วนใหญ่ผลการค้นคืนสารสนเทศมักจะมองจากภายนอก เช่น คำถามถึงความพอใจของผู้ใช้มักให้เลือกตอบ ดังนี้

ก. พอใจมาก	5	คะแนน
ข. พอใจ	4	คะแนน
ค. พอใจพอสมควร	3	คะแนน
ง. ผิดหวัง	2	คะแนน
จ. ผิดหวังมาก	1	คะแนน

คำถามในการประเมินผลลักษณะนี้ เป็นเพียงการประเมินผลในระดับต้น ซึ่งให้ข้อมูลด้านระดับความสำเร็จจากภายนอก คือ จากความพึงพอใจของผู้ใช้ ซึ่งยังไม่ใช่ว่าผลที่ถูกต้อง 100 เปอร์เซ็นต์ เนื่องจากความพึงพอใจของผู้ใช้แต่ละคนวัดได้ยาก และมีปัจจัยที่เกี่ยวข้องหลายตัว อาทิ ระดับความรู้ ความสามารถในการใช้งานระบบ อารมณ์ ความรู้สึก ซึ่งจะแตกต่างกันไปตามสถานการณ์ นอกจากนี้ยังรวมไปถึงทัศนคติและอคติส่วนบุคคล หากต้องการประเมินผลให้ได้ข้อมูลมีประโยชน์ต่อการปรับปรุงระบบต่อไป ควรต้องมีการควบคุมปัจจัยต่าง ๆ ให้รัดกุม ซึ่งอาจต้องใช้การสังเกตพฤติกรรมของผู้ใช้ควบคู่ไปด้วย

การประเมินประสิทธิภาพของระบบ (Effectiveness)

ในการประเมินผลแง่ประสิทธิภาพของการค้นคืนสารสนเทศ คำถามที่จำเป็นที่สุดจะเกี่ยวกับจำนวนข้อมูลที่เรียกขึ้นมาได้ในการปฏิบัติการค้นคืนสารสนเทศในหัวข้อหนึ่ง ๆ ว่าจำนวนที่ค้นคืนมาได้นั้นตรงกับความต้องการของผู้ใช้มากน้อยเพียงใด ซึ่งในระบบการค้นคืนสารสนเทศทั่วไป ผลการค้นคืนสารสนเทศมักออกมาในลักษณะดังภาพ



จากภาพ

- A คือ Hits ข้อมูลที่เรียกขึ้นมาได้และเกี่ยวข้องกับเรื่องที่ต้องการ (Relevant and Retrieved)
- B คือ Noise ข้อมูลที่เรียกขึ้นมาได้แต่ไม่เกี่ยวข้องกับเรื่องที่ต้องการ (Not relevant but Retrieved)
- C คือ Misses ข้อมูลที่เรียกขึ้นมาไม่ได้แต่เกี่ยวข้องกับเรื่องที่ต้องการ (Relevant but Not retrieved)
- D ระบบทำงานถูกต้อง ข้อมูลที่ไม่เกี่ยวข้องกับเรื่องที่ต้องการและไม่ถูกเรียกขึ้นมา (Not relevant and Not retrieved)

สิ่งที่ผู้ใช้ต้องการที่สุดคือ A และ D เท่ากับ 100 เปอร์เซ็นต์ และลดปริมาณของ B และ C ให้มากที่สุด แต่ในความเป็นจริงเหตุการณ์ที่เกิดขึ้นมากที่สุด คือ D และเหตุการณ์ที่เกิดขึ้นน้อยที่สุด คือ A

ในการวิจัยด้านประสิทธิภาพ หากจำนวน B และ C = "0" จะแสดงถึงประสิทธิภาพสูงสุด แต่ในความเป็นจริง หากจำนวน B และ C ต่ำ ก็คือจากการค้นคืนสารสนเทศนั้นมีประสิทธิภาพพอควร

ตัวอย่าง

ผลการค้นที่ถูกต้องสมบูรณ์ สำหรับการค้นคืนเอกสารที่เกี่ยวข้องซึ่งมีทั้งหมด 30 รายการ จากเอกสารที่มีอยู่ในฐานข้อมูลทั้งหมด 1000 รายการ จะต้องได้ผล ดังนี้

- A = 30 (Relevant, Retrieved)
- B = 0 (Not relevant, Retrieved)
- C = 0 (Relevant, Not Retrieved)
- D = 970 (Not relevant, Not retrieved)

แต่ในความเป็นจริงผลที่มักเกิดขึ้น คือ

A	=	25	(Relevant, Retrieved)
B	=	5	(Not relevant, Retrieved)
C	=	5	(Relevant, Not Retrieved)
D	=	965	(Not relevant, Not retrieved)

และในกรณีที่ย่ำ คือ

A	=	15	(Relevant, Retrieved)
B	=	15	(Not relevant, Retrieved)
C	=	15	(Relevant, Not Retrieved)
D	=	955	(Not relevant, Not retrieved)

หรือในกรณีที่เลวร้ายที่สุด คือ

A	=	5	(Relevant, Retrieved)
B	=	25	(Not relevant, Retrieved)
C	=	25	(Relevant, Not Retrieved)
D	=	945	(Not relevant, Not retrieved)

จะเห็นว่า ในทุกกรณี มีเอกสารถูกเรียกออกมาไม่น้อยกว่า 30 รายการ แต่จำนวนเอกสารที่เกี่ยวข้องที่เรียกออกมาได้แตกต่างกัน ซึ่งเราสามารถนำตัวเลขเหล่านี้มาคำนวณหาเปอร์เซ็นต์ของเอกสารที่เกี่ยวข้องที่สามารถเรียกค้นออกมาได้ เพื่อประเมินประสิทธิภาพของระบบการค้นคืนสารสนเทศ เนื่องจากวัตถุประสงค์หลักของระบบการค้นคืนสารสนเทศคือ การค้นคืนสารสนเทศ ดังนั้น การวัดความสามารถในการเรียกค้นเอกสารที่เกี่ยวข้องออกมาได้ จึงเป็นสิ่งสำคัญในการประเมินระบบการค้นคืนสารสนเทศ ทั้งนี้ เพื่อทราบสัดส่วนของเอกสารที่เกี่ยวข้องที่สามารถเรียกค้นออกมาได้จริง

การคำนวณหาเปอร์เซ็นต์จำนวนเอกสารที่เรียกขึ้นมาได้ เรียกว่า Recall หรือ Recall ratio ใช้ตัวย่อว่า R เป็นการประเมินความสมบูรณ์ของระบบในการเรียกค้นเอกสารที่เกี่ยวข้อง (Relevant documents) ออกมาได้ คำนวณได้จาก

$$\text{Recall ratio} = \frac{\text{จำนวนรายการที่ค้นได้และตรงกับเรื่องที่ต้องการ}}{\text{จำนวนรายการที่ตรงกับเรื่องที่ต้องการทั้งหมด}} \times 100$$

แทนค่าด้วยสูตร

$$R = \frac{a}{a+c} \times 100$$

เมื่อ a = เอกสารที่เรียกขึ้นมาได้และเกี่ยวข้องกับเรื่องที่ต้องการ

c = เอกสารที่เรียกขึ้นมาไม่ได้แต่เกี่ยวข้องกับเรื่องที่ต้องการ

$$\begin{aligned} \text{ดังนั้น เมื่อ } a = 25 \quad c = 5 \quad \text{จำนวนเปอร์เซ็นต์ของ Recall} &= \frac{25}{25 + 5} \times 100 \\ &= 83.3\% \end{aligned}$$

อย่างไรก็ตาม การได้มาซึ่งค่า Recall เป็นเรื่องที่ยากยิ่งกว่าที่คิด เนื่องจากจะทราบว่ามีฐานข้อมูลมีเอกสารที่ตรงกับความต้องการอยู่เท่าใดนั้นทำได้ยาก โดยเฉพาะถ้าฐานข้อมูลนั้นมีขนาดใหญ่มาก นั่นคือ เราไม่สามารถทราบค่า C ที่แท้จริง วิธีการที่ใช้ในการพยายามหาค่า Recall คือ การให้ผู้ค้นหลายคน ค้นคืนสารสนเทศในเรื่องเดียวกัน โดยมีผู้ค้นหลักหนึ่งคนที่ให้ผลลัพธ์แก่ผู้ใช้ และหากผู้ค้นคนอื่น ๆ ได้รับรายการที่ตรงเพิ่มขึ้น ก็จะนำเอาจำนวนนั้นมารวมกันเพื่อเป็นจำนวน รายการที่ตรงกับเรื่องที่ต้องการทั้งหมดที่มีในฐานข้อมูล เช่น

ผู้สืบค้นหลัก ค้นได้ 28 รายการ
 ผู้ค้นคนที่สอง 30 รายการ (เพิ่มกว่าผู้สืบค้นหลัก 2 รายการ)
 ผู้ค้นคนที่สาม 32 รายการ (เพิ่มกว่าผู้สืบค้นหลัก 4 รายการ)
 ดังนั้น มีรายการที่ผู้สืบค้นหลัก ค้นออกมาไม่ได้ 4 รายการ

$$\begin{aligned} \text{ค่า Recall} &= \frac{28}{28 + 4} \times 100 \\ &= 87\% \end{aligned}$$

อย่างไรก็ตาม เป็นไปได้ว่า ผู้สืบค้นทั้ง 3 คน อาจไม่ได้รายการที่ตรงกับเรื่องที่ต้องการทั้งหมด เพราะความผิดพลาดที่เกิดขึ้นจากการจัดทำคำสำคัญเพื่อช่วยการสืบค้น ทำให้บางรายการหลุดไปจากการสืบค้น การใช้ค่า Recall นี้ จึงมักไม่ค่อยปรากฏในแบบประเมินผลการสืบค้น เพราะผู้ใช้ย่อมไม่ทราบแน่นอนว่ามีรายการที่ตรงกับเรื่องที่ต้องการทั้งหมดเท่าไร (พิมพ์วิภา, 2538: หน้า 88) ดังนั้น ในการทดลองเพื่อประเมินประสิทธิภาพของระบบจากค่า Recall จึงมักทำในลักษณะจำลอง โดยการสร้างฐานข้อมูลตัวอย่าง ซึ่งผู้ประเมินจะทราบปริมาณเอกสารที่เกี่ยวข้องที่แท้จริง

การวัดประสิทธิภาพของระบบการค้นคืนสารสนเทศนอกจากคำนวณหาเปอร์เซ็นต์จำนวนเอกสารที่เรียกขึ้นมาได้ หรือค่า Recall แล้ว ยังต้องคำนวณเปอร์เซ็นต์ของจำนวนเอกสารที่เรียกขึ้นมาได้และตรงกับความต้องการ หรือค่า Precision ทั้งนี้ เนื่องจาก บางครั้งเอกสารที่เรียกค้นขึ้นมาได้ อาจไม่ใช่เอกสารที่เกี่ยวข้อง หรือตรงกับความต้องการของผู้ใช้ทั้งหมด ทั้งนี้ เนื่องจากการเรียกค้นข้อมูลจากฐานข้อมูลคอมพิวเตอร์เป็นเพียงการจับคู่ คำค้นที่ใช้กับเอกสารที่มีอยู่ในฐานข้อมูล ซึ่งหากกำหนดคำค้น หรือ คำแทนสาระของเอกสารผิดพลาดจะทำให้ไม่สามารถเรียกเอกสารที่ถูกต้องออกมาได้ เป็นต้น ดังนั้น การคำนวณหาค่า Precision จึงเป็นการวัดความสามารถของระบบการค้นคืนสารสนเทศในการที่จะไม่เรียกค้นเอกสารที่ไม่เกี่ยวข้องหรือไม่ตรงกับความต้องการออกมา

การคำนวณหาค่า Precision ทำได้ ดังนี้

$$\text{Precision ratio} = \frac{\text{จำนวนรายการที่ค้นได้และตรงกับเรื่องที่ต้องการ}}{\text{จำนวนรายการที่ค้นได้ทั้งหมด}} \times 100$$

$$\text{แทนค่าด้วยสูตร} \quad P = \frac{a}{a+b} \times 100$$

เมื่อ a = เอกสารที่เรียกขึ้นมาได้และเกี่ยวข้องกับเรื่องที่ต้องการ

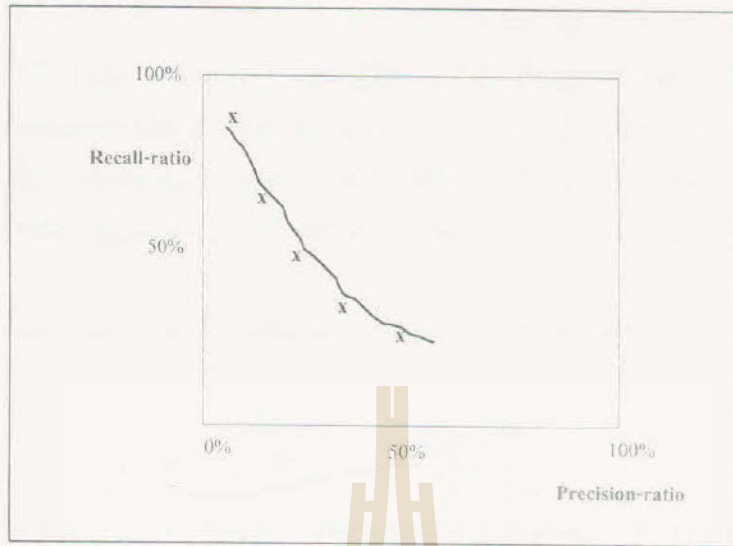
b = เอกสารที่เรียกขึ้นมาได้แต่ไม่เกี่ยวข้องกับเรื่องที่ต้องการ

$$\begin{aligned} \text{ดังนั้น เมื่อ } a = 25 \quad b = 15 \quad \text{จำนวนเปอร์เซ็นต์ของ Precision} &= \frac{25}{25+15} \times 100 \\ &= 62.5\% \end{aligned}$$

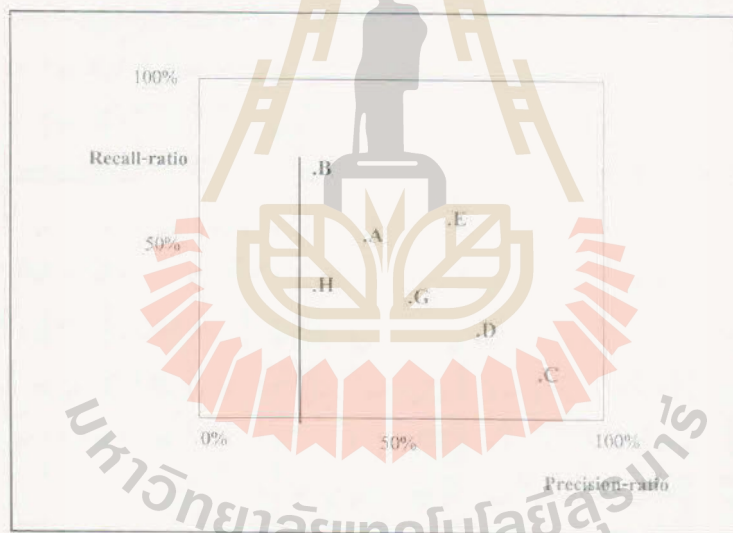
ทั้ง Recall และ Precision เป็นที่นิยมในการประเมินผลประสิทธิภาพการค้นคืนสารสนเทศ อย่างไรก็ตาม หากจะศึกษาเพื่อให้ทราบปัญหาที่แท้จริงของระบบจำเป็นต้องใช้ข้อมูลทั้ง Recall และ Precision ควบคู่กัน โดยใช้ข้อมูล R และ P ของการค้นคืนสารสนเทศหลาย ๆ ครั้ง เปรียบเทียบกัน ดังภาพที่ 1 และภาพที่ 2 ในหน้าถัดไป ผลการค้นที่สมบูรณ์ หรือระบบที่ดีเยี่ยม ทั้งค่า Recall และ Precision จะต้องเท่ากับ 100% ดังตัวอย่าง

A	=	30	(Relevant, Retrieved)
B	=	0	(Not relevant, Retrieved)
C	=	0	(Relevant, Not Retrieved)
D	=	970	(Not relevant, Not retrieved)

แต่ เนื่องจากปัญหาด้านความซับซ้อนขององค์ความรู้ ภาษา และระบบการค้นคืนสารสนเทศ ผลการค้นในข้างต้นจะเกิดขึ้นได้ยาก อย่างไรก็ตาม ผู้ใช้ส่วนใหญ่จะพึงพอใจที่ระดับ Recall และ Precision เท่ากับ 75% (McCarthy, 2000) ขึ้นไป อย่างไรก็ตาม ความสัมพันธ์ของ Recall และ Precision มักจะมีความสัมพันธ์ในลักษณะผกผัน คือ ถ้า Recall เพิ่มขึ้น Precision จะลดลง แต่หาก Recall ลดลง Precision จะเพิ่มขึ้น (McCarthy, 2000) นั่นคือ ยิ่งผู้ค้นพยายามที่จะเพิ่มปริมาณเอกสารที่เกี่ยวข้องที่ค้นออกมาได้ จำนวนเอกสารที่ไม่เกี่ยวข้องที่ค้นออกมาได้ก็ยิ่งจะเพิ่มจำนวนตามไปด้วย



ภาพที่ 1



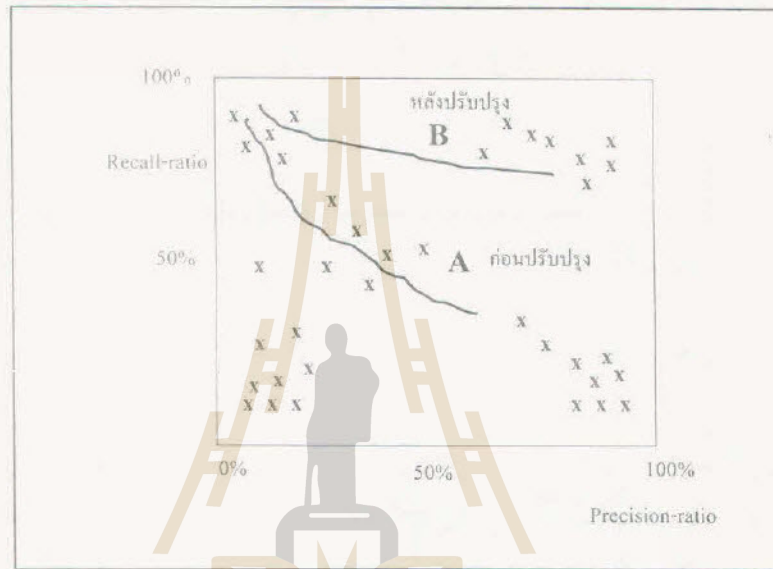
ภาพที่ 2

ภาพที่ 1 และภาพที่ 2 แสดงการทดลองค้นหลาย ๆ ครั้ง เพื่อเปรียบเทียบผลการค้น จะเห็นว่าความสัมพันธ์ระหว่าง Recall และ Precision มักเป็นไปในลักษณะผกผัน นั่น คือเมื่อ R สูง P จะต่ำ และเมื่อ R ต่ำ P จะสูง แต่หากกล่าวในแง่ของประสิทธิภาพการค้น อาจเป็นระบบที่ไม่ค่อยดีนัก เนื่องจาก ค่า P และ R ค่อนข้างแปรผัน อย่างไรก็ตาม ประสิทธิภาพในแง่ความพึงพอใจของผู้ใช้ขึ้นอยู่กับต้องการ Recall และ Precision สูงต่ำเพียงใด เช่น ถ้าผู้ใช้ต้องการข้อมูลที่มีลักษณะเฉพาะ P จะสูง R จะต่ำ แต่หากผู้ใช้ต้องการรวบรวมข้อมูลในระดับกว้าง P จะต่ำ R จะสูง

สำหรับการประเมินเพื่อปรับปรุงระบบ ข้อมูลเฉลี่ยของค่า Recall และ Precision จะถูกนำมาพิจารณา ศึกษาเพื่อปรับปรุงประสิทธิภาพของระบบ ข้อมูลเฉลี่ยของ Recall และ Precision นี้ ได้จากการค้นคืน

สารสนเทศหลาย ๆ รูปแบบด้วยกัน ในภาพที่ 3 เส้น A แสดงข้อมูลเฉลี่ยตัว X ซึ่งเป็นข้อมูลที่ได้จากการค้นคืนสารสนเทศหลาย ๆ ครั้ง แต่แตกต่างกันในรูปกลวิธีในการค้น จะเห็นว่า X จะกระจายเป็นกลุ่ม ๆ แสดงถึงผลการค้นที่ดีและไม่ดี

เมื่อทำการค้นด้วยวิธีที่ต่างกัน จะพบว่า วิธีใด หรือ อะไรที่ทำให้การค้นคืนมีประสิทธิภาพ และอะไรที่เป็นปัญหา หรือทำให้การค้นคืนไม่มีประสิทธิภาพ ดังนั้น เส้น B จะแสดงถึงการพัฒนาปรับปรุงระบบการค้นคืนสารสนเทศให้มีประสิทธิภาพดีขึ้น ซึ่งปัจจัยที่มีผลให้ค่า P และ R แตกต่างกัน เช่น การใช้คำค้น วิธีการค้น ระบบดรรรชนี ระบบการกำหนดคำศัพท์ ประสบการณ์ ความสามารถและบุคลากรที่ทำการค้นคืนข้อมูล เป็นต้น



ภาพที่ 3

นอกจากค่า Recall และ Precision แล้ว ยังมีค่าที่จะนำมาใช้เป็นเกณฑ์ในการวัดความพึงพอใจของผู้ใช้ต่อผลการสืบค้น คือ ค่า Novelty ซึ่งเป็นอัตราร้อยละของเอกสารที่ค้นได้และตรงกับเรื่องที่ต้องการ และมีความ "ใหม่" สำหรับผู้ใช้ กล่าวคือ ผู้ใช้ไม่เคยเห็น หรือทราบเกี่ยวกับเอกสารรายการนี้มาก่อน ในกรณีที่ผู้ใช้ไม่เคยทำการค้นคืนข้อมูลในเรื่องนั้นมาก่อน ค่า Novelty จะสูง

วิธีการคำนวณค่า Novelty ratio หรือ E ทำได้ดังนี้

$$E = \frac{\text{จำนวนรายการใหม่ที่ค้นได้และตรงกับเรื่องที่ต้องการ} \times 100}{\text{จำนวนรายการที่ค้นได้ทั้งหมด}}$$

$$E = \frac{e}{a + b} \times 100$$

e = จำนวนรายการใหม่ที่ค้นได้และตรงกับเรื่องที่ต้องการ

เมื่อ a = 25, b = 15, e = 10 ดังนั้น $E = \frac{10}{25 + 15} \times 100 = 25\%$

หรือในบางกรณีจะคำนวณโดยใช้สูตร

$$E = \frac{\text{จำนวนรายการใหม่ที่ค้นได้และตรงกับเรื่องที่ต้องการ}}{\text{จำนวนรายการที่ค้นได้และตรงกับเรื่องที่ต้องการ}} \times 100$$

ดังนั้น $E = \frac{e}{a} \times 100$

เมื่อ $a = 25$, $e = 10$ ดังนั้น $E = \frac{10}{25} \times 100$
 $= 40\%$

การประเมินโดยใช้ค่า Recall Precision และ Novelty เป็นการประเมินโดยเปรียบเทียบกับความต้องการสารสนเทศของผู้ใช้ ซึ่งความไม่พึงพอใจของผู้ใช้อาจเกิดขึ้นได้จากหลายกรณี เช่น หากผู้ใช้ต้องการผลการค้นที่ครอบคลุมประเด็นที่ต้องการอย่างกว้างขวาง และตรงกับความต้องการ แต่การค้นคืนไม่สามารถค้นได้มากเท่าที่ต้องการ นั่นหมายความว่า Recall ต่ำ หรือผลการค้นที่ได้มีเป็นจำนวนมาก แต่ไม่เกี่ยวข้องกับเรื่องที่ต้องการ ซึ่งหมายความว่า Precision ต่ำ หรือผู้ใช้ได้รับรายการที่เกี่ยวข้องแต่ไม่มีเรื่องที่ใหม่ ผู้ใช้อาจรู้สึกว่าเป็นการเสียเวลาที่ค้นคืนข้อมูลจากฐานข้อมูลนี้ ซึ่งกรณีต่าง ๆ เหล่านี้ล้วนสามารถทำให้ผู้ใช้ไม่พึงพอใจกับการค้นคืนข้อมูลจากฐานข้อมูลได้ทั้งสิ้น

การประเมินประสิทธิภาพคุ้มกับค่าใช้จ่าย (Cost-Effectiveness) และ การประเมินประโยชน์คุ้มค่า หรือ การประเมินคุณค่าของบริการ (Cost-Benefits)

การประเมินผลในแง่ประสิทธิภาพคุ้มค่าใช้จ่าย จะแสดงถึงความสัมพันธ์ระหว่างระดับความสำเร็จหรือประสิทธิภาพการค้นคืนสารสนเทศ กับค่าใช้จ่ายที่ทำให้เกิดความสำเร็จนั้น ส่วนการประเมินผลแง่มีประโยชน์คุ้มค่า หรือการประเมินคุณค่าของบริการ จะแสดงความสัมพันธ์ระหว่างประโยชน์ที่ได้รับจากการจัดให้มีบริการสารสนเทศนั้นมากับค่าใช้จ่ายที่เกิดขึ้น ซึ่งการประเมินทั้งสองประเภทนี้อาจเห็นผลสมจริงได้ยากกว่าการประเมินประสิทธิภาพของระบบซึ่งประเมินจากผู้ใช้โดยตรง โดยการประเมินประสิทธิภาพคุ้มค่าใช้จ่ายจะเกี่ยวข้องกับผู้บริหาร แต่ถ้าเป็นการประเมินคุณค่าของบริการจะเกี่ยวข้องกับผู้บริหารสูงสุด ซึ่งในการจัดให้มีบริการใดก็ตามจะต้องทำการประเมินคุณค่าของบริการนั้นเป็นอันดับแรกก่อนที่จะจัดให้มีบริการนั้นขึ้น

ค่าใช้จ่ายที่ควรนำมาพิจารณาในการประเมินมี 2 ประเภท ได้แก่

1. ค่าใช้จ่ายคงที่ (Fixed Cost) เช่น ค่าอุปกรณ์ ค่าเช่าสถานที่ ค่าอุปกรณ์เป็นสมาชิกฐานข้อมูล เงินเดือนพนักงาน เป็นต้น
2. ค่าใช้จ่ายไม่คงที่ หรือค่าใช้จ่ายแปรผัน (Variable Cost) เช่น ค่าจัดส่งเอกสารทางไปรษณีย์ ค่าโทรศัพท์ ค่าโทรคมนาคม เป็นต้น

ในการวิจัยเชิงประสิทธิภาพคุ้มค่าใช้จ่าย สามารถปรับปรุงคุณค่าการค้นคืนสารสนเทศได้ ดังนี้

1. คงระดับประสิทธิภาพการให้บริการค้นคืนสารสนเทศ แต่ลดระดับค่าใช้จ่าย

2. งบประมาณค่าใช้จ่าย แต่พยายามปรับปรุงประสิทธิภาพการค้นคืนสารสนเทศ

ปกติการวิจัยเชิงประสิทธิภาพคุ้มค่าใช้จ่ายควรจัดเตรียมบริการหลาย ๆ วิธี หรือ หลายลักษณะ และคิดค่าใช้จ่ายเปรียบเทียบกัน ซึ่งวิธีการที่ดีที่สุด คือ วิธีที่จ่ายน้อยกว่าแต่ผลการค้นคืนอยู่ในระดับที่เหมาะสม คือไม่น้อยไปกว่าวิธีอื่น ๆ

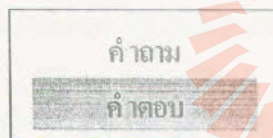
ในการประเมินผลในแง่ประโยชน์คุ้มค่า หรือคุณค่าของบริการ มีหลักในการพิจารณากว้าง ๆ คือ

1. ในการค้นคืนสารสนเทศนั้น ประหยัดค่าใช้จ่ายกว่าค่าใช้จ่ายในการค้นหาข้อมูลที่ต้องการกว่าวิธีอื่น
2. หากไม่มีการค้นคืนสารสนเทศวิธีดังกล่าว จะหาข้อมูลไม่ได้เลย
3. สามารถใช้การค้นคืนสารสนเทศเพื่อช่วยในการตัดสินใจ หรือช่วยลดปัญหาในการตัดสินใจที่จำเป็นต้องทำแม้จะมีข้อมูลไม่เพียงพอ
4. ลดความซ้ำซ้อนในการค้นคว้าวิจัย และปรับปรุงการทำงานให้ดีขึ้น
5. การค้นคืนสารสนเทศช่วยกระตุ้นการค้นพบสิ่งที่มีประโยชน์ใหม่ ๆ

ความสำเร็จและความล้มเหลวของการค้นคืนสารสนเทศ

โดยปกติผู้ใช้ย่อมปรารถนาที่จะได้คำตอบที่ตรงกับความต้องการจากระบบ และสามารถนำไปใช้ประโยชน์ได้ทันที และสารสนเทศนั้นจะต้องสอดคล้องกับความรู้ ความสามารถของผู้ใช้ทั้งในด้านเนื้อหาและรูปแบบ ซึ่งเป็นภาวะที่ระบบก็ต้องการเช่นกัน แต่ในทางปฏิบัติความสอดคล้องระหว่างคำถามกับคำตอบอาจปรากฏในลักษณะใดลักษณะหนึ่ง ดังต่อไปนี้

1. คำถามและคำตอบมีความสอดคล้องกันพอดี (Exact match or Whole match)



2. ข้อมูลที่ต้องการปรากฏในส่วนใดส่วนหนึ่งของผลลัพธ์ / คำตอบที่ได้ (Inclusive match)



3. ผลลัพธ์ที่ได้ตอบคำถามได้เพียงบางส่วน (Partial match)



4. คำตอบไม่ตรงกับความต้องการโดยสิ้นเชิง (No match)



ซึ่งในความเป็นจริง โอกาสที่จะได้ผลการค้นดังข้อ 1 นั้นน้อยมาก ดังนั้น ผลการค้นที่พึงปรารถนาคือ ข้อมูลที่ต้องการสามารถเรียกค้นออกมาได้เกือบครบถ้วน ดังได้กล่าวมาแล้วในส่วนของ Recall และ Precision ซึ่งปัจจัยที่ทำให้การค้นคืนได้ผลไม่เต็มที่มีหลายประการ เช่น คำถามไม่ชัดเจน ความสามารถในการวิเคราะห์ คำถาม การตีความปัญหาของผู้ใช้ การกำหนดคำค้นและกลวิธีการค้น วิธีการค้นคืนยังไม่เหมาะสม ระบบดรรชนีด้อยประสิทธิภาพ กลไกการทำงานของระบบการจัดเก็บและค้นคืนผิดพลาด คุณภาพของสารสนเทศที่เลือก / ประเมินเข้ามาจัดเก็บ ซึ่งระบบการค้นคืนสารสนเทศที่ดีควรสามารถสะท้อนการวิเคราะห์ปัญหาของผู้ใช้ออกมาเป็นความต้องการใช้สารสนเทศได้อย่างชัดเจน แม้ว่าบางครั้งผู้ใช้ที่มีปัญหาเกิดขึ้นบางคน อาจจะไม่ได้ตระหนักถึงความต้องการใช้สารสนเทศของตนก็ตาม และเมื่อผสมผสานกับกลไกในการจัดเก็บและค้นคืนสารสนเทศที่มีประสิทธิภาพแล้ว ควรจะได้สารสนเทศที่มีคุณค่าต่อการนำไปใช้เพื่อจัดการกับปัญหานั้น ๆ และอยู่ในรูปที่สอดคล้องกับความรู้และภูมิหลังของผู้ใช้ด้วย

ปัจจัยที่มีผลต่อความสำเร็จและความล้มเหลวของการค้นคืนสารสนเทศ

อาจสรุปปัจจัยที่มีผลต่อความสำเร็จและความล้มเหลวของการค้นคืนสารสนเทศ ได้ 3 กลุ่ม ดังนี้

1. ผู้ใช้ ปัจจัยในส่วนของผู้ใช้ ได้แก่
 - การวิเคราะห์ความต้องการสารสนเทศ
 - ทักษะและประสบการณ์ในการค้นคืนสารสนเทศ
 - ความรู้ในเรื่องที่ค้น
 - ความรู้เกี่ยวกับระบบที่ใช้ เช่น การใช้ Function หรือ Command ความรู้เกี่ยวกับความสามารถในการทำงานของระบบ
2. ผู้ให้บริการ ปัจจัยในส่วนของผู้ให้บริการ ได้แก่
 - ทักษะในการสัมภาษณ์ เพื่อทราบความต้องการที่แท้จริงของผู้ใช้
 - การเลือกคำค้น หรือ กลวิธีการค้น
 - ความรู้ในเรื่องที่ค้น
3. ระบบสารสนเทศ ปัจจัยในส่วนของระบบสารสนเทศ ได้แก่
 - ระบบการจัดเก็บ ได้แก่ การเลือกหรือการประเมินสารสนเทศ การจำแนกหรือการวิเคราะห์เนื้อหาของสารสนเทศ การกำหนดค่าแทนสาระ การสร้างระบบคำดรรชนี
 - ระบบการค้นคืน ได้แก่ การจัดการเพิ่มดรรชนี การกำหนดเขตข้อมูล (Fields) ที่ใช้เป็นจุดเข้าถึงข้อมูล (Access points) คุณลักษณะด้านการค้น (Search Features)
 - ระบบการเชื่อมประสานกับผู้ใช้ (User Interfaces) ได้แก่ ความเป็น User Friendly คือ ใช้งานง่าย เรียนรู้ง่าย ดึงดูดใจ

ปัญหาในการค้นคืนสารสนเทศออนไลน์ของผู้ใช้

ปัญหาในการค้นคืนสารสนเทศออนไลน์ของผู้ใช้ จำแนกได้ ดังนี้

1. ไม่ทราบคำศัพท์ที่จะใช้ค้น
2. ไม่ทราบวิธีขยายผลการค้น

3. ไม่ทราบวิธีจำกัดผลการค้น เช่น มักใช้เฉพาะ คำเดียวในการค้น (Single keyword search)
4. จำกัดการค้นมากเกินไป (Over specification) เช่น ใช้คำศัพท์มากเกินไปในการค้นหนึ่งครั้ง (Too many terms input in a single search)
5. ในบางกรณีที่การค้นด้วยวลี (Phrase) หรือ คำสมเหมาะสมกว่า แต่ผู้ใช้อักค้นด้วยคำสำคัญ (Keyword) เช่น การค้นคืนเอกสารเกี่ยวกับห้องสมุดอัตโนมัติ ผู้ใช้มักใช้คำว่า "Library AND Automation" แทนที่จะใช้คำว่า "Library Automation" ซึ่งเป็นคำที่ตรงกว่า
6. ไม่เข้าใจสาเหตุที่ทำให้เกิดความผิดพลาด เช่น ไม่เข้าใจ Error messages
7. ใช้คำค้นที่ไม่ชัดเจน
8. ใช้ Stop words ในการค้น
9. ยุ่งยากในการสำรวจ หรือ คู่มือการค้นที่ยาวมาก หรือมีรายละเอียดมาก
10. รูปแบบการแสดงผลการค้นไม่ตรงกับความต้องการ
11. พิมพ์ผิด
12. กลุ่มที่มาใช้ไม่บ่อยครั้ง ต้องเรียนรู้ระบบใหม่ทุกครั้งที่มาใช้
13. ไม่ทราบความแตกต่างระหว่างการค้นโดยใช้ศัพท์อิสระ กับ ศัพท์ควบคุม ค้นผิด file ค้นผิด field เป็นต้น
14. การใช้คำย่อในฐานข้อมูล
15. ไม่ทราบขอบเขตของฐานข้อมูล
16. ลืมเหลอในการใช้ Search features
17. ไม่เข้าใจรหัส/คำย่อที่ใช้ในหน้าจอการแสดงผล
18. ข้อมูลที่แสดงในการแสดงผลแบบสั้น (Brief display) ไม่เพียงพอ
19. ไม่เข้าใจ Help Message / Help Screen

วิธีการประเมินระบบโดยผู้ใช้อย่างผู้ประเมิน

การประเมินระบบโดยนำข้อมูลจากผู้ใช้งานในการวิเคราะห์ความสามารถของระบบ อาจทำการรวบรวมข้อมูลได้ ดังนี้

1. ใช้การสังเกตขณะผู้ใช้งานการค้นคืนสารสนเทศจากระบบ ข้อมูลที่เก็บ ได้แก่ เรื่องที่ค้น กลวิธีในการค้น คำค้นที่ใช้ ตรรกะในการค้น รวมทั้งผลการค้นที่ได้ วิธีการนี้ค่อนข้างจะยุ่งยากและอาจก่อให้เกิดความรำคาญแก่ผู้ใช้ ที่สำคัญคือ สิ้นเปลืองเวลาและแรงงานมาก
2. ใช้แบบฟอร์ม ซึ่งได้แก่ แบบสอบถามและแบบสัมภาษณ์ หัวข้อที่สอบถามได้แก่ เรื่องที่สืบค้นและกลยุทธ์ในการค้น อาทิ การเลือกใช้คำศัพท์ สูตรการค้น จำนวนรายการที่ได้รับ ความพึงพอใจต่อผลการสืบค้น และ ปัญหาที่พบ ข้อควรพิจารณาคือ แบบประเมินที่ดีควรสั้นและกระชับ หากเป็นแบบฟอร์มที่ยาวและต้องใช้เวลาในการตอบผู้ใช้อักไม่ให้ความร่วมมือ เนื่องจากผู้ใบบางคนไม่มีเวลามากพอที่จะคิดทบทวนหาเหตุผล หรือข้อความยาว ๆ คำถามในลักษณะนี้ เช่น "ท่านเห็นว่าผลการค้นคืนในครั้งนี้มีประโยชน์สำหรับท่านมากน้อยเพียงไร และโปรดให้เหตุผลประกอบว่าทำไมจึงเป็นเช่นนั้น" เป็นต้น คำถามนี้ ผู้ใช้ต้องคิดว่า ตนเองควรจะระบุว่าผลการค้นคืนเป็นประโยชน์มากที่สุด มาก น้อย หรือ ไม่เป็นประโยชน์เลย แล้วยังต้องพยายามให้เหตุผล

ด้วย ซึ่งคงต้องใช้เวลาพอสมควร และหากนำคำถามที่ถามไว้เป็นคำถามข้อแรก ๆ ในแบบฟอร์ม ก็ยังจะทำให้ผู้ใช้หมดความตั้งใจที่จะตอบมากขึ้น อีกประการหนึ่งที่ต้องพิจารณาคือการจัดรูปแบบของคำถามและคำตอบ คำถามที่ยาวและเป็นคำถามปลายเปิด (แบบ ก.) อาจจะทำให้กระชับและสะดวกกับผู้ใช้ยิ่งขึ้น โดยปรับเป็นคำถามสั้น ๆ ที่มีคำตอบให้เลือก (แบบ ข.) ดังตัวอย่าง

การสอบถามวัตถุประสงค์ของการค้น

แบบ ก. ท่านมีวัตถุประสงค์เช่นไรจากการสืบค้นฐานข้อมูล
หมายความว่า เมื่อท่านแสดงความจำนงค์ในการสืบค้นข้อมูล ท่านจะมีแผนที่จะนำสารสนเทศที่ได้ไปใช้ทำอะไร

แบบ ข.

ท่านจะใช้สารสนเทศที่ได้จากฐานข้อมูลเพื่อ

- | | |
|---|---|
| <input type="checkbox"/> การเขียน | <input type="checkbox"/> การสอน |
| <input type="checkbox"/> การทำวิจัย | <input type="checkbox"/> การทำวิทยานิพนธ์ |
| <input type="checkbox"/> อื่นๆ โปรดระบุ | |

นอกจากนี้ ผู้ประเมินควรพิจารณาด้วยว่า ต้องการได้รับคำตอบอะไรจากการสอบถามผู้ใช้ หรือ จะใช้ผลที่ได้จากการประเมินสำหรับแก้ไขปรับปรุงในเรื่องใดเป็นพิเศษ รวมทั้งความพึงพอใจของผู้ใช้ต่อการสืบค้น อาจใช้คำถามดังนี้

การสอบถามความพึงพอใจของผู้ใช้

ท่านมีความพึงพอใจอย่างไรต่อ

1. ความเข้าใจของผู้สืบค้นกับคำสืบค้นของท่าน (ดูแนวคิดสำคัญที่ใช้ในการวางกลยุทธ์สืบค้น)

ไม่พึงพอใจ 1 2 3 4 5 พึงพอใจมาก

2. ความละเอียดถี่ถ้วนของผู้สืบค้นในการเลือกรายข้อมูลที่เหมาะสม และใช้คำหรือวลีที่ถูกต้องในการค้น

ไม่พึงพอใจ 1 2 3 4 5 พึงพอใจมาก

แบบประเมินการสืบค้นสารสนเทศจากฐานข้อมูล มักประกอบด้วยคำถามเกี่ยวกับ เรื่องดังต่อไปนี้

- ข้อมูลเกี่ยวกับผู้ใช้
- วัตถุประสงค์ในการค้นคืน
- ผลลัพธ์ของการสืบค้น ได้แก่ จำนวนรายการที่ค้นได้ทั้งหมด จำนวนรายการที่ค้นได้และตรงกับความต้องการ และ จำนวนรายการที่ใหม่และตรงกับความต้องการ
- ความเห็นเกี่ยวกับค่าใช้จ่ายในการค้นคืนสารสนเทศ
- ความคิดเห็นเกี่ยวกับเวลาสนองตอบ (Response time) นับตั้งแต่เริ่มยื่นคำขอในการค้นคืนสารสนเทศ จนกระทั่งได้ผลการค้น (บรรณานุกรม)
- ความเห็นและข้อเสนอแนะเกี่ยวกับบริการสืบค้นจากฐานข้อมูล

แบบประเมินที่จัดทำขึ้นควรทำการทดสอบก่อนว่าสามารถจะสื่อให้ผู้กรอกเข้าใจได้ตรงกันกับที่ผู้ประเมินต้องการ และมีจำนวนข้อคำถามที่เหมาะสม ไม่มากจนผู้ใช้ไม่ยอมตอบ หรือน้อยเกินไปจนไม่สามารถได้ผลที่จะนำมาประเมินได้ครบถ้วน แบบประเมินควรจัดวางไว้ที่เทอร์มินัลที่ใช้ในการค้นคืน และผู้ให้บริการควรพยายามเชิญชวนให้ผู้ใช้ได้ช่วยประเมิน ในกรณีที่ผู้ให้บริการทำการค้นคืนข้อมูลให้ แบบประเมินควรแจกไปพร้อม ๆ กับผลการค้นคืนสารสนเทศ และอาจให้ผู้ใช้ช่วยประเมินขณะทำการค้นคืนสารสนเทศให้กับผู้ใช้

ในแบบประเมินเกี่ยวกับการค้นคืนสารสนเทศมักจะมีคำถามที่จะพิจารณาว่า ผลการค้นที่ได้มีเอกสารที่เกี่ยวข้องกับความต้องการของผู้ใช้หรือไม่ ในกรณีที่ผู้ประเมินต้องการหาค่า Recall Precision และ Novelty เอง ก็มักจะให้ผู้ใช้ระบุว่า มีรายการที่ตรงกับความต้องการที่รายการ และมีรายการใหม่ที่รายการ เช่น

- | |
|--|
| - ใน การค้น ท่านได้รายการบรรณานุกรมทั้งสิ้น _____ รายการ |
| - จำนวนรายการที่ตรงกับเรื่องที่ต้องการ _____ รายการ |
| - จำนวนรายการที่ท่านไม่เคยเห็น / ทราบมาก่อน _____ รายการ |

ในบางครั้งถ้าผู้ประเมินต้องการให้ผู้ใช้ระบุค่าร้อยละโดยตรง ก็จะต้องตั้งคำถามโดยกำหนดคำตอบไว้ ดังตัวอย่าง

- | |
|--|
| 1. จำนวนรายการที่ตรงกับความต้องการ คิดเป็นกี่เปอร์เซ็นต์ของรายการที่ค้นได้จากการสืบค้นครั้งนี้ |
| _____ 0% |
| _____ 1-25% |
| _____ 26-50% |
| _____ 51-75% |
| _____ 76-100% |
| 2. ในบรรดารายการที่ตรงกับความต้องการ มีรายการที่ "ใหม่" หรือเป็นรายการที่ไม่เคยทราบมาก่อนที่จะ ตรวจสอบ |
| ผลลัพธ์การสืบค้น |
| _____ 0% ของรายการที่ค้นได้เป็นรายการใหม่ |
| _____ 1-25% ของรายการที่ค้นได้เป็นรายการใหม่ |
| _____ 26-50% ของรายการที่ค้นได้เป็นรายการใหม่ |
| _____ 51-75% ของรายการที่ค้นได้เป็นรายการใหม่ |
| _____ 76-100% ของรายการที่ค้นได้เป็นรายการใหม่ |

คำถามและคำตอบเช่นนี้ จะช่วยทวนเวลาของผู้ประเมินในการที่จะค้นหาค่าต่าง ๆ ที่ต้องการและคำตอบที่ได้รับก็จะนำไปใช้ประโยชน์ได้ทันที แต่มีข้อสังเกตว่าคำตอบที่เป็นการกำหนดของ บางครั้งอาจทำให้ไม่ได้คำตอบที่เฉพาะเจาะจง เป็นการประมาณการเท่านั้น (พิมพรีาไพ, 2538: หน้า 94)

การประเมินระบบการจัดเก็บและค้นคืนสารสนเทศเป็นเรื่องจำเป็น เนื่องจากจะช่วยในการแก้ไขปรับปรุงระบบให้ดียิ่งขึ้น แม้กระบวนการของการประเมินจะยุ่งยากและใช้เวลา อย่างไรก็ตาม หน่วยงานควรพิจารณาเลือกใช้วิธีการที่เหมาะสมกับสภาพแวดล้อมของการดำเนินงานของตน การใช้แบบฟอร์มในการประเมินเพียงอย่างเดียวอาจไม่เพียงพอที่จะให้ภาพว่าระบบการจัดเก็บและค้นคืนสารสนเทศที่จัดทำขึ้นมีคุณภาพเพียงพอหรือไม่ อาจต้องให้การสัมภาษณ์ผู้ใช้ประกอบกัน หรือมีการทดสอบการใช้ระบบของผู้ใช้กลุ่มต่าง ๆ ที่มีพื้นความรู้ และความต้องการที่แตกต่างกัน เพื่อความเที่ยงตรงของข้อมูล และควรทำการทดสอบหลาย ๆ ครั้ง เพื่อให้ได้ข้อมูลรูปที่ชัดเจนเกี่ยวกับปัญหาที่เกิดขึ้น หรือประสิทธิภาพในการทำงานของระบบโดยรวม

บรรณานุกรม

การค้นคืนสารสนเทศออนไลน์. 2535. กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย .

นฤมล ปราชญ์โยธิน, ทวีศักดิ์ กอนันตกุล, เปรมิณ จินดาภิรมลเลิศ. 2536. อิซอร์สกับระบบสารสนเทศ. กรุงเทพฯ : ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ.

พิมพรีาไพ เปรมสมิทธิ์. 2538. ฐานข้อมูลบรรณานุกรม : การสร้างและการใช้. กรุงเทพฯ : ภาควิชาบรรณารักษศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย .

MacCarthy, Cavan. (2000). Relevant and Recall : Online Information Systems. School of Library and Information Science University of Iowa. [Online]. Available: <http://mingo.info-science.uiowa.edu/mccarthy/online17relevancerecall/classnotes.html>

Maglaughlin, Kelly L. and Sonnenwald, Diane H. (2002). User Perspective on Relevance Criteria: A Comparison among Relevant, Partially Relevant, and Not-Relevant Judgments. *Journal of The American Society For Information Science And Technology* 53(5): 327-342.

Meadow, Charles T. 1992. *Text Information Retrieval Systems*. San Diego, California : Academic press.