บทที่ 5
การจัดทำบันทึกข้อมูลศัพท์

เมื่อรวบรวมข้อมูลจากคลังข้อมูลภาษามาบันทึกไว้ในบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้นแล้ว (Conceptual network and extraction records) ขั้นตอนต่อมาในการจัดทำประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาลก็คือการจัดทำบันทึกข้อมูลศัพท์ (Terminological Records) ซึ่งเป็นขั้นตอนที่สำคัญมากอีกขั้นตอนหนึ่งในการจัดทำประมวลศัพท์ การทำงานในขั้นตอนนี้ต้องอาศัยการประมวลข้อมูลที่ได้จากขั้นตอนต่างๆ มารวบรวมวิเคราะห์ ซึ่งผลลัพธ์ที่ได้จะไปปรากฏเป็นบันทึกข้อมูลศัพท์ในประมวลศัพท์ฉบับสมบูรณ์

ว่าด้วยเรื่องบันทึกข้อมูลศัพท์

ในการจัดทำประมวลศัพท์แต่ละครั้ง ผู้จัดทำย่อมกำหนดรูปแบบของบันทึกข้อมูลศัพท์แตกต่างกันไปตามแต่วัตถุประสงค์ใช้งาน กลุ่มเป้าหมาย และข้อจำกัดต่างๆ แต่สิ่งที่บันทึกข้อมูลศัพท์จะต้องมีร่วมกันก็คือ ลักษณะเด่นที่ว่า ข้อมูลที่บันทึกไว้เป็นข้อมูลที่ได้จากการวิเคราะห์และประมวลข้อมูลที่รวบรวมมาได้จากกระบวนการขั้นต้นซึ่งมีความซับซ้อนและต้องอาศัยเวลา ดังนั้นการบันทึกข้อมูลศัพท์จึงเปรียบเสมือนการสรุปผลงานที่ผู้จัดทำต้องใช้ความรู้ทั้งหมดที่ได้รวบรวมมาเพื่อให้มีความสมบูรณ์และถูกต้องมากที่สุด

เมื่อกล่าวถึงการจัดทำบันทึกข้อมูลศัพท์โดยทั่วไปแล้ว ส่วนที่สำคัญที่สุดมักจะอยู่ที่การให้นิยามของศัพท์ หรือที่เรียกกันว่า Terminological Definition เพราะเป็นส่วนที่ผู้ใช้งานประมวลศัพท์จะได้รับประโยชน์โดยตรง แต่นอกจากส่วนที่เป็นนิยามแล้ว ข้อมูลต่างๆที่บันทึกไว้ในข้อมูลประมวลศัพท์ต่างก็เป็นข้อมูลที่สำคัญและมีความหมายต่อประโยชน์ใช้งานของประมวลศัพท์เป็นอย่างยิ่ง Sager (1990) ได้แบ่งประเภทของข้อมูลพื้นฐาน (Basic Category) ที่มักจะมีการบันทึกไว้ในประมวลศัพท์ (Sager, 1990: 143) ดังนี้

(1) ข้อมูลอ้างอิง (Source Information) ได้แก่ รายการอ้างอิงของศัพท์ นิยาม ตัวอย่าง และอื่นๆ

(2) ศัพท์ (Entry Terms)

(3) รูปภาษาและมโนทัศน์ (Semantic and conceptual specification) ได้แก่ นิยาม ประเภทศัพท์ บันทึกการใช้งาน

(4) ข้อมูลทางภาษาศาสตร์ (Linguistic Specification) เช่น ตัวย่อ (abbreviation) คำเหมือน (Synonym) เป็นต้น

(5) ข้อมูลการใช้งานจริง (Pragmatic Specification) เช่น ตัวอย่าง หรือ ข้อสังเกตการใช้งาน เป็นต้น

(6) ข้อมูลจัดระบบ (Administrative Information) เช่น หมายเลขบันทึก (Record Number) ผู้แต่ง (Author) วันที่บันทึก (Record Date)

(7) ศัพท์ภาษาต่างประเทศที่มีความหมายเช่นเดียวกัน (Foreign Language Equivalent)

ซึ่งจากการพิจารณาองค์ประกอบข้างต้น ช่วยให้ผู้ทำวิจัยกำหนดวิธีวิทยาในการบันทึกข้อมูลศัพท์ได้ รายละเอียดดูได้จากวิธีวิทยาในหัวข้อถัดไป

วิธีวิทยาในการจัดทำบันทึกข้อมูลศัพท์

ในการจัดทำประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาล ผู้ทำวิจัยได้รวบรวมบันทึกข้อมูลพื้นฐานทั้งหมดที่ได้จัดทำไว้ในขั้นตอนต่างๆ และได้พิจารณาถึงความเหมาะสมของกลุ่มเป้าหมายและวัตถุประสงค์ในการทำวิจัยครั้งนี้ แล้วจึงได้กำหนดรูปแบบของบันทึกประมวลศัพท์ให้สอดคล้องกับความต้องการและความเหมาะสมดังกล่าว โดยได้คำนึงถึงองค์ประกอบพื้นฐานซึ่งควรมีอยู่ในบันทึกข้อมูลศัพท์ ดังตัวอย่างข้อมูลที่ได้อ้างถึงไว้ในหัวข้อก่อนหน้านี้ และได้ข้อสรุปว่าองค์ประกอบที่ควรจัดให้มีในบันทึกข้อมูลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาล มีดังนี้

ขั้นตอนที่ 1 การกำหนดประเภทข้อมูล

ข้อมูลต่างๆ ที่รวบรวมได้จากคลังข้อมูลภาษา และบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น มีอยู่หลากหลาย และยังสามารถนำมาวิเคราะห์หาข้อมูลต่างๆ เกี่ยวกับศัพท์นั้นๆ ได้อีกมาก แต่ประเภทของข้อมูลที่ควรจะเลือกมานำเสนอเป็นผลการวิจัยไม่ใช้ข้อมูลทั้งหมดที่จะหาได้ แต่เป็นข้อมูลทั้งหมดที่จะเป็นประโยชน์ต่อผู้ใช้งานและสนองตอบต่อจุดมุ่งหมายในการเป็นคู่มือใช้งานศัพท์เฉพาะด้านและเครื่องมือสำหรับนักแปลนักวิชาการได้ด้วย เมื่อพิจารณาถึงปัจจัยเหล่านี้แล้ว จึงกำหนดประเภทของข้อมูลที่จะนำมาบันทึกไว้ในบันทึกประมวลศัพท์ได้ดังนี้

ศัพท์หลัก (Entry Term/Term) ข้อมูลนี้เป็นสิ่งที่ขาดไม่ได้สำหรับประมวลศัพท์ นอกจากจะเป็นสัญลักษณ์ที่ใช้แทนมโนทัศน์แล้ว ศัพท์ที่เลือกมาเป็นศัพท์หลักจะต้องมีลักษณะที่ผู้ใช้งานคุ้นเคยกว่าศัพท์อื่นที่มีความหมายแบบเดียวกัน เพราะศัพท์หลักคือศัพท์ที่ผู้ใช้งานจะใช้เป็นตัวหลักในการค้นหาความหมายของศัพท์หรือมโนทัศน์ที่ต้องการ ดังนั้นในกรณีที่มโนทัศน์หนึ่งมีศัพท์ที่เป็นคำเหมือน (Synonym) อยู่มากกว่าหนึ่ง ก็จำเป็นต้องอาศัยกฎเกณฑ์ในการเลือกศัพท์ใดศัพท์หนึ่งขึ้นมาเป็นศัพท์หลัก ซึ่งรายละเอียดจะได้กล่าวถึงต่อไปในขั้นตอนที่ 2

นิยาม (Definitions) ดังที่ได้กล่าวไว้ในหัวข้อก่อนแล้วว่า นิยามจัดเป็นสิ่งที่สำคัญมากในบันทึกประมวลศัพท์ Sager (1990) ได้กล่าวไว้ว่า 'The definition is the first item that links the entry term to the concept which it represents. As such it is the bridge between concept and term.'(Sager, 1990: 145)

ในฐานะของสะพานเชื่อมระหว่างมโนทัศน์และศัพท์ นิยามเป็นสิ่งที่ผู้จัดทำประมวลศัพท์ต้องให้ความสำคัญอย่างมาก และเป็นส่วนที่จะแสดงมโนทัศน์ของศัพท์นั้นๆ ให้ผู้ใช้งานทราบ นับว่านิยามเป็นองค์ประกอบที่สำคัญและขาดไม่ได้มากที่สุดอย่างหนึ่งในบันทึกข้อมูลศัพท์

ความสัมพันธ์กับศัพท์อื่น (Cross-reference) ส่วนนี้เป็นการนำเสนอความสัมพันธ์ระหว่างศัพท์นี้กับศัพท์อื่นๆ ที่อาจอยู่ในกลุ่มเดียวกัน เพื่อแสดงว่าศัพท์ในชุดหรือกลุ่มนี้ประกอบไปด้วยศัพท์ใดบ้าง ข้อมูลนี้จะช่วยให้ผู้ใช้งานประมวลศัพท์ทราบว่าศัพท์นี้มีศัพท์อื่นใดบ้างที่เกี่ยวข้องกันหรือเกี่ยวเนื่องอยู่ในกลุ่มเดียวกัน ซึ่งอาจจะใช้ในการค้นคว้าอ้างอิงต่อไปได้

หมวดเรื่อง (Subject Field) หมวดเรื่องเป็นการแจกแจงกลุ่มของศัพท์ในประมวลศัพท์ โดยมากจะแบ่งตามมิติที่ใช้ในการจัดมโนทัศน์สัมพันธ์ เป็นข้อมูลที่ชี้ให้เห็นว่าศัพท์แต่ละศัพท์อยู่ในหมวดเรื่องใดบ้าง เพื่อความสะดวก และเป็นการยืนยันความถูกต้องในการใช้งาน

ข้อมูลการใช้ศัพท์ (Scope Note) หัวข้อนี้รวมถึง ข้อจำกัด ข้อยกเว้น ลักษณะพิเศษต่างๆ ในการใช้ศัพท์นั้นๆ ในบริบทต่างๆ ซึ่งลักษณะเหล่านี้อาจจะมีเฉพาะในบางบริบท หรือเกิดเฉพาะกับศัพท์นี้เท่านั้น หรือเกิดเฉพาะในศาสตร์เฉพาะด้านเท่านั้น

ข้อมูลลักษณะเฉพาะของศัพท์ (Usage Note) หัวข้อนี้แสดงรายละเอียดลักษณะของศัพท์ว่า เป็นภาษาพูด (Colloquial) เป็นศัพท์แสลง (Slang) หรือเป็นศัพท์แปล (Translation) ข้อมูลประเภทนี้จะช่วยให้ผู้ใช้งานรู้สถานการณ์ที่เหมาะสมในการใช้ศัพท์ หรือจะพบศัพท์นั้น

ตัวอย่างการใช้ศัพท์ (Illustration) ใช้ช่วยในการแสดงวิธีการใช้งานของศัพท์ ข้อมูลตัวอย่างนี้มักได้มาจากคลังข้อมูลภาษา บางครั้งการให้ตัวอย่างที่เหมาะสมจะช่วยให้ผู้ใช้งานประมวลศัพท์เกิดความเข้าใจที่มากขึ้นเกี่ยวกับศัพท์นั้นๆ

แหล่งอ้างอิง (Sources) แหล่งอ้างอิงนี้ หมายถึงแหล่งอ้างอิงของ ศัพท์หลัก นิยาม ตัวอย่าง และศัพท์ไทย ซึ่งอาจใช้ตัวย่อในการแสดงข้อมูลอ้างอิงนี้ได้เพื่อไม่ให้เป็นการเปลืองเนื้อที่ในการทำวิจัยครั้งนี้ ได้ใช้รหัสย่อในการแสดงที่มาของข้อมูล เช่น CEP01 ซึ่ง C หมายถึงมาจากข้อมูลในคลังข้อมูลภาษา (Corpus based data) E หมายความว่าเป็นข้อมูลภาษาอังกฤษ (English) และ P แสดงว่ามีลักษณะการสื่อสารระหว่าง Expert - Uninitiated สำหรับหมายเลขที่ต่อท้ายแสดงว่าข้อมูลนี้นับเป็นไฟล์ลำดับที่ 01 สำหรับข้อมูลประเภทนี้ เป็นต้น รหัสเช่นนี้จะวงเล็บไว้หลังข้อมูลที่ต้องการอาศัยการอ้างอิงเพื่อให้ผู้ใช้งานสามารถทราบแหล่งที่มาของข้อมูลได้ โดยการนำรหัสดังกล่าวมาดูเปรียบเทียบกับแหล่งข้อมูลจากตารางรายการอ้างอิง (ดูตารางข้อมูลอ้างอิงได้ในภาคผนวก ก) การแสดงแหล่งอ้างอิงนี้ก็เพื่อให้ผู้ใช้งานสามารถค้นคว้าต่อเองได้ และเพื่อให้งานมีความน่าเชื่อถือเพราะมีแหล่งข้อมูลอ้างอิงที่แน่นอน

รูปศัพท์อื่นๆ (Other linguistic Specification) เช่น ตัวย่อ (Abbreviation) คำที่มีความหมายเดียวกัน (Variants) คำตรงข้าม (Antonym) เป็นต้น ข้อมูลตรงนี้จะช่วยให้ผู้ใช้งานรู้เรื่องศัพท์ดีขึ้นและอาจช่วยในการทำความเข้าใจศัพท์และการใช้ศัพท์ได้มากขึ้น

ข้อมูลเชิงไวยากรณ์ (Grammatical Category) ข้อมูลนี้จะระบุว่าศัพท์หลักจัดอยู่ในรูปไวยากรณ์ใด เช่น เป็นนาม (noun) กิริยา (verb) วิเศษณ์ (Adjective) เป็นต้น นอกจากจะมีประโยชน์ในแง่การใช้งานแล้วข้อมูลในส่วนนี้ยังช่วยแยกแยะความแตกต่างระหว่างศัพท์ที่อยู่ในรูปเดียวกันแต่มีการใช้เป็นทั้งสองประเภท เช่น Recharge ซึ่งเป็นได้ทั้งนามและกิริยา เป็นต้น

ศัพท์ไทย (Thai Equivalent) ข้อมูลศัพท์ไทยได้จากข้อมูลที่คัดออกมาจากคลังข้อมูลภาษามาเก็บไว้ในบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น ศัพท์ใดที่ยังไม่มีศัพท์ไทย หรือ

มีใช้อยู่เดิมแต่ไม่เหมาะสม ผู้วิจัยอาจเสนอแนะคำใหม่ได้ วิธีการในส่วนนี้จะได้พูดถึงต่อไปในบทที่ 6 (การกำหนดศัพท์ภาษาไทยในประมวลศัพท์)

ข้อมูลเชิงจัดการ (Administrative Information) หมายถึงข้อมูลที่ใช้ช่วยในการจัดระบบการทำงานของบันทึกประมวลศัพท์นี้ ตัวอย่างของข้อมูลเชิงจัดการ ได้แก่ หมายเลขบันทึก (Record Number) ซึ่งในการทำวิจัยครั้งนี้ ผู้ทำวิจัยก็ได้จัดทำหมายเลขบันทึกไว้ทั้งสำหรับบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น และบันทึกข้อมูลศัพท์ เพื่อให้ง่ายต่อการอ้างอิงและลดความสับสนที่อาจเกิดขึ้นได้

เมื่อกำหนดองค์ประกอบของบันทึกประมวลศัพท์ได้แล้ว ก็มาถึงขั้นตอนการร่างแบบฟอร์มไว้กรอกข้อมูลเพื่อเป็นบรรทัดฐานในการจัดทำต่อไป โดยแบบฟอร์มดังกล่าวมีลักษณะดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 4 แบบฟอร์มบันทึกข้อมูลศัพท์

1	2	3
4	5	
6		
7		
8		
9		

1 English Term (ศัพท์อังกฤษ)*

2 Thai Term (ศัพท์ไทย)*

3 Number of Record (หมายเลขบันทึก)

4 Grammatical Category (ประเภททางไวยากรณ์ศัพท์)

5 Subject Field (หมวดเรื่อง)

6 Definition (นิยาม)*

7 Illustration (ตัวอย่าง)*

8 Note (ข้อสังเกตการใช้งานศัพท์ในบางบริบท)

9 Linguistic Specification (รูปศัพท์อื่น)*

10 Cross-reference (ศัพท์อื่นในกลุ่มที่เกี่ยวข้องทางด้านมโนทัศน์สัมพันธ์)*

* หมายถึงข้อมูลที่มีการใส่รหัสแสดงแหล่งข้อมูลอ้างอิงไว้ด้วย

ขั้นตอนที่ 2 การบันทึกข้อมูลลงในบันทึกข้อมูลศัพท์

เมื่อได้แบบฟอร์มของบันทึกข้อมูลศัพท์แล้ว ขั้นตอนต่อไปคือ การนำข้อมูลที่มีอยู่มารวบรวม วิเคราะห์ และนำผลที่ได้ไปบันทึกไว้ในแบบฟอร์มดังกล่าวตามหัวข้อที่ได้กล่าวถึงไว้แล้วในขั้นตอนที่ 1 โดยใช้คอมพิวเตอร์ช่วยในการเก็บข้อมูล ซึ่งทำให้สามารถบันทึกข้อมูลได้มาก และหาข้อมูลที่ต้องการได้รวดเร็ว

ข้อมูลที่นำมาใช้ในการรวบรวมนี้ ส่วนหนึ่งเป็นข้อมูลจากคลังข้อมูลภาษาโดยตรง ส่วนหนึ่งเป็นข้อมูลจากบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น (Conceptual Relation &

Extraction Records) และส่วนหนึ่งเป็นข้อมูลจากแหล่งภายนอก เช่น จากผู้เชี่ยวชาญ พจนานุกรม สารานุกรม เป็นต้น การจะเลือกใช้ข้อมูลใดบ้างนั้นขึ้นอยู่กับหัวข้อที่ต้องการนำไปใช้ ซึ่งจะได้กล่าวในรายละเอียดไล่ตามลำดับที่ปรากฏในแบบฟอร์มบันทึกข้อมูลศัพท์ ดังนี้

1) English Term (**ศัพท์อังกฤษ**) ตามที่ได้กล่าวไว้ในขั้นตอนที่ 1 แล้วว่า การจัดชุดศัพท์ตามมโนทัศน์ อาจมีข้อเสียในแง่ที่ต้องเลือกศัพท์จากหลายศัพท์ที่สื่อมโนทัศน์เดียวกัน ในการทำวิจัยครั้งนี้ ผู้ทำวิจัยได้กำหนดเกณฑ์ความถี่มาใช้ในการคัดเลือกศัพท์หลัก โดยมีสมมติฐานว่า ข้อมูลศัพท์ในคลังข้อมูลภาษาเป็นตัวแทนการใช้งานของศัพท์ในสถานการณ์การสื่อสารที่หลากหลายและครอบคลุมเรื่องการเติมน้ำลงชั้นน้ำบาดาลในทุกด้าน ดังนั้น ศัพท์ที่สื่อมโนทัศน์เดียวกัน ศัพท์ใดพบว่ามีการใช้มากที่สุดในคลังข้อมูลภาษา ศัพท์นั้นก็น่าจะเป็นศัพท์ที่ผู้ใช้งานคุ้นเคยมากที่สุดและมีการใช้งานมากที่สุดในบรรดาศัพท์ที่สื่อมโนทัศน์เดียวกัน ตัวอย่างเช่น ศัพท์ที่สื่อถึงบ่อน้ำบาดาลที่ใช้ตรวจวัดปริมาณน้ำในบ่อเติมน้ำ มีอยู่สามศัพท์ คือ Observation Well, Monitoring Well และ Collector Well ทั้งหมดต่างก็สื่อถึงมโนทัศน์เดียวกัน การที่จะเลือกศัพท์ใดมาเป็นศัพท์หลักในบันทึกข้อมูลศัพท์ ทำได้โดยการดูความถี่จากคลังข้อมูลภาษา พบว่า Monitoring Well มีการใช้มากที่สุด จึงสันนิษฐานว่าถ้านำ Monitoring well น่าจะมีประโยชน์ในการใช้งานมากกว่า เพราะเป็นศัพท์ที่ผู้ใช้งานมีแนวโน้มจะพบมากกว่าศัพท์อื่น ส่วน Observation Well และ Collector Well ก็ให้เป็นศัพท์ที่มีความหมายเดียวกัน (Variant, Synonym) ซึ่งอยู่ในหัวข้อรูปศัพท์อื่น (Linguistic Specification)

รูปแบบของศัพท์หลักเมื่อนำมาบันทึกไว้ ควรจะอยู่ในรูปที่เป็นกลาง เช่นเดียวกับที่ใช้ในพจนานุกรม เช่น นามอยู่ในรูปเอกพจน์ ใช้ตัวอักษรเล็กนำหน้าศัพท์ (ยกเว้นกรณีที่เป็นชื่อเฉพาะ) หรือคำกิริยาต้องอยู่ในรูป active infinitive เป็นต้น

2) Thai Term (**ศัพท์ไทย**) วิธีวิทยาในการบันทึกข้อมูลศัพท์ไทยมีรายละเอียดค่อนข้างมาก ดังนั้นจึงยกไปกล่าวไว้ต่างหากในบทถัดไป (บทที่ 6)

3) Number of Record (**หมายเลขบันทึก**) ใส่เรียงไปตามลำดับ โดยให้ขึ้นต้นด้วย TR (Terminological Records) แล้วตามด้วยตัวเลข เช่น TR001 = บันทึกข้อมูลศัพท์ลำดับที่ 1 (รายละเอียดดูได้ในภาคผนวก ค)

4) Grammatical Category (**ประเภททางไวยากรณ์ศัพท์**) บันทึกตามข้อมูลที่ปรากฏในคลังข้อมูลภาษาและบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น (Conceptual Relation & Extraction Records)

5) Subject Field (**หมวดเรื่อง**) บันทึกตามข้อมูลจากบันทึกมโนทัศน์สัมพันธ์ ตัวอย่างเช่น ศัพท์ 'Recharge Well' จะมีหมวดเรื่องดังนี้

Groundwater Recharge > Recharge Method > Well Recharge > Recharge Well

ซึ่งข้อมูลนี้ก็ทำให้ผู้ใช้งานทราบว่า Recharge Well เป็นศัพท์หนึ่งที่เกี่ยวข้องอยู่ในเรื่อง Well Recharge ซึ่งเป็นวิธีหนึ่งของ Recharge Method ในศาสตร์เรื่องการเติมน้ำลงชั้นน้ำบาดาล

6) Definition (**นิยาม**) วิธีการได้มาซึ่งนิยามสำหรับประมวลศัพท์ (Terminological Definition) อาจมีได้หลายรูปแบบ เช่น รวบรวมจากหนังสืออ้างอิงประเภทพจนานุกรมหรือสารานุกรม จากความรู้ศาสตร์เฉพาะด้านของผู้จัดทำประมวลศัพท์เอง หรือได้จากคลังข้อมูลภาษาเป็นต้น ในการจัดทำประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาลครั้งนี้ ผู้ทำวิจัยได้เลือกใช้วิธีการหานิยามศัพท์จากคลังข้อมูลภาษาเป็นหลัก และใช้วิธีการอื่นๆ เป็นวิธีการเสริมเมื่อไม่สามารถหานิยามจากคลังข้อมูลภาษาได้ หรือเมื่อข้อมูลที่ได้จากคลังข้อมูลภาษาไม่เพียงพอหรือไม่ชัดเจน

หลักเกณฑ์การเลือกนิยามให้กับศัพท์ในประมวลศัพท์ กรณีที่มีนิยามศัพท์ในข้อมูลภาษาให้เลือกใช้มากกว่าหนึ่งนิยาม การเลือกให้พิจารณาดูว่า นิยามใดมีความชัดเจนมากที่สุด เนื่องจากกลุ่มเป้าหมายเป็นบุคคลทั่วไปที่ไม่มีความรู้เฉพาะด้านมาก่อน การให้นิยามจึงควรเป็นลักษณะของการอธิบาย (Descriptive) และใช้ภาษาที่ง่ายต่อความเข้าใจ หลีกเลี่ยงการใช้ศัพท์เฉพาะด้านอื่นๆ ให้น้อยที่สุด ถ้าจำเป็นต้องใช้ให้ใส่แหล่งข้อมูลอ้างอิงเพื่อให้ผู้ใช้งานสามารถตามไปดูความหมายของศัพท์ดังกล่าวได้ด้วย ลักษณะของนิยามที่อาจพบในคลังข้อมูลภาษามีรูปแบบใหญ่ๆ อยู่สามรูปแบบด้วยกันตามนิยามของ Trimble (1996) ที่ Pearson(1998) ได้รวบรวมไว้ (Pearson, 1998: 98)

(1) Formal Definition นิยามนี้จะให้ข้อมูลสามแบบ คือ ศัพท์ (Name of Term) ชนิด (Class) และความแตกต่างจากศัพท์อื่นๆ ในกลุ่มเดียวกัน เช่น An anemometer is a

meteorological instrument that registers the speed of wind on a dial or gage. ซึ่งประโยคนี้ให้ข้อมูลศัพท์ = anemometer ชนิดของศัพท์ = meteorological instrument และความแตกต่างที่ศัพท์นี้มีเมื่อเทียบกับศัพท์อื่น =เป็น meteorological instrument ที่ registers the speed of wind on a dial or gage ได้ เป็นต้น

(2) Semi-formal Definition นิยามกลุ่มนี้มีข้อมูลเหมือน Formal Definition ทุกอย่าง ยกเว้นแต่ไม่ระบุชนิดศัพท์ (Class) เท่านั้น เช่น An anemometer registers the speed of the wind on a dial or gage.

(3) Non-formal Definition นิยามแบบนี้จะให้ข้อมูล คือ ศัพท์ กับคำหรือวลีที่มีความหมายใกล้เคียงกับศัพท์นั้น หรือให้ข้อมูลลักษณะเด่นๆ ของศัพท์นั้น เช่น An arachnid is a spider.

ซึ่งในการเลือกประโยคที่ให้นิยามจากคลังข้อมูลภาษามาใช้ จะเน้นคุณภาพของนิยามเรียงลำดับจาก Formal Definition, Semi-formal Definition จนถึง Non-formal Definition ซึ่งในกรณีที่พบแต่ Semi-formal Definition และ Non-formal Definition อาจต้องใช้ประกอบกับวิธีการอื่นด้วย เช่น การค้นหานิยามจากหนังสืออ้างอิงประเภทพจนานุกรมหรือประมวลศัพท์ หรือจากการสอบถามผู้เชี่ยวชาญ เป็นต้น อย่างไรก็ตาม นิยามที่จะนำมาใช้บันทึกจะต้องเป็นนิยามที่อยู่บนพื้นฐานที่ว่ากลุ่มเป้าหมายต้องเข้าใจและนำไปใช้งานได้ไม่ยาก ไม่มีความคลุมเครือ และใช้ภาษาให้ง่ายที่สุด

เมื่อกำหนดลักษณะของนิยามที่จะรวบรวมจากคลังข้อมูลภาษาแล้ว สิ่งที่ต้องคำนึงถึงต่อมาก็คือ จะรวบรวมบริบทที่แสดงนิยามดังกล่าวได้อย่างไร Pearson(1998) ได้หยิบยกเอางานวิจัยของ Flowerdew (1996) มาสรุปไว้ว่า เราสามารถสังเกตหาบริบทที่แสดงนิยามของศัพท์ได้โดยสังเกตจากรูปศัพท์หรือไวยากรณ์ (syntactic or lexical devices) โดยส่วนใหญ่แล้วจะปรากฏคำสำคัญ (Key word) อย่าง We call/ is called/ are called/ called หรือ or, know as เป็นต้น ซึ่งประโยคหรือวลีที่ตามหลังตัวคำหรือวลีเหล่านี้มักจะเป็นนิยามของศัพท์ที่อยู่ข้างหน้าคำหรือวลีดังกล่าว เช่น Now a photo that we take through a microscope ***we call*** a micrograph. (Cited in Pearson, 1998: 103) ซึ่งจากประโยคนี้เราทราบว่า 'photo that one take through a microscope' เป็นนิยามของ 'micrograph' เพราะมีการใช้ 'we call' นำหน้าศัพท์ดังกล่าว เป็นต้น

เมื่อได้บริบทที่แสดงนิยามของศัพท์ออกมาแล้ว ก็นำมาพิจารณาแบ่งประเภทตามที่ได้กล่าวมาแล้ว จากนั้นในกรณีที่สมควรหรือจำเป็นต้องใช้วิธีการอื่นๆ มาช่วยในการหานิยามเพิ่มเติม ก็ให้พิจารณาเป็นรายศัพท์ไป ขั้นตอนสุดท้ายก็จะได้นิยามศัพท์ทั้งชุดที่มีลักษณะตรงตามที่ได้ตั้งเป้าหมายไว้

7) Illustration **(ตัวอย่าง)** การเลือกตัวอย่างสำหรับประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาลเน้นการเลือกข้อมูลจากคลังข้อมูลภาษา โดยตัวอย่างอาจไม่ใช่ Defining Context แต่อาจจะเป็น Metalinguistic contextหรือ Testimonial Context ก็ได้ เพราะตัวอย่างในบันทึกข้อมูลศัพท์นี้ไม่ได้นำมาใช้เป็นบริบทในการหานิยามเช่นในบันทึกข้อมูลศัพท์เบื้องต้น แต่จะเน้นว่าเป็นตัวอย่างที่แสดงการใช้งานได้ชัดเจนที่สุดมากกว่า ดังนั้น เกณฑ์ในการเลือกบริบทข้อมูลตัวอย่าง จึงเน้นว่าเป็นบริบทการใช้ศัพท์ที่แสดงลักษณะการใช้ศัพท์ได้ชัดเจนที่สุด เข้าใจง่ายที่สุด และแสดงลักษณะเด่นออกมาได้มากที่สุด แต่อย่างไรก็ตาม ข้อมูลดังกล่าวเน้นว่าควรเป็นข้อมูลจากคลังข้อมูลภาษาเพื่อสื่อให้เห็นว่าเป็นตัวอย่างการใช้งานจริง หรืออาจนำตัวอย่างนี้มาจากบันทึกข้อมูลศัพท์เบื้องต้นที่รวมอยู่ในบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้นก็ได้ เพราะเป็นข้อมูลที่นำมาจากคลังข้อมูลภาษาเช่นเดียวกัน และการใช้ข้อมูลในบันทึกข้อมูลศัพท์เบื้องต้นก็จะทำให้ประหยัดเวลามากขึ้นด้วย

8) Note **(ข้อมูลการใช้ศัพท์)** ข้อมูลนี้ได้มาจากคลังข้อมูลภาษาโดยตรง โดยประมวลจากบริบทที่ศัพท์ปรากฏว่า ศัพท์นี้มีลักษณะการใช้งานที่แตกต่างจากศัพท์อื่นเป็นพิเศษอย่างไรหรือเปล่า เช่น มีข้อจำกัดการใช้งานหรือไม่ มีลักษณะพิเศษอย่างไรหรือไม่ เป็นต้น ซึ่งข้อมูลส่วนนี้อาจไม่มีในศัพท์ทุกศัพท์ ดังนั้นจึงเป็นส่วนพิเศษของข้อมูลประมวลศัพท์ ซึ่งอาจมีหรือไม่มีก็ได้

9) Linguistic Specification **(รูปศัพท์อื่น)** ข้อมูลที่บันทึกไว้ในส่วนนี้คือ รูปศัพท์อื่นๆ ที่สื่อถึงตัวศัพท์หลัก เช่น ตัวย่อ (Abbreviation) คำเหมือน (Synonym) คำตรงข้าม (Antonym) เป็นต้น ซึ่งข้อมูลเหล่านี้ก็ได้มาจากคลังข้อมูลภาษาและบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น หรืออาจได้จากแหล่งข้อมูลอื่น เช่น พจนานุกรม หรือสารานุกรม เป็นต้น

10) Cross-reference **(ศัพท์อื่นในกลุ่มที่เกี่ยวข้องทางด้านมโนทัศน์สัมพันธ์)** ข้อมูลส่วนนี้ได้จากบันทึกมโนทัศน์สัมพันธ์โดยตรง โดยจะแสดงรายชื่อศัพท์ที่เกี่ยวข้องพร้อมข้อมูลอ้างอิงไปยังศัพท์นั้นๆ (ดูตัวอย่างจากภาพที่ 5)

นอกจากองค์ประกอบที่กล่าวมาทั้งหมดแล้ว ข้อมูลอีกแบบที่ใช้กำกับศัพท์เพื่อระบุลักษณะเฉพาะของศัพท์ หรือที่เรียกว่า ข้อมูลลักษณะเฉพาะของศัพท์ (Usage Note) เนื่องจากข้อมูลกลุ่มนี้เป็นข้อมูลย่อยดังนั้นจึงอาจใช้อักษรย่อแทนลักษณะแต่ละแบบแล้วนำไปใส่เป็นวงเล็บไว้หลังศัพท์นั้นๆ ได้เลย โดยกำหนดให้ col. = colloquial , slang = slang และ tran. = translation เป็นต้น และไม่จำเป็นต้องมีข้อมูลนี้กำกับอยู่หลังศัพท์ทุกคำ ยกเว้นเฉพาะศัพท์ที่มีลักษณะพิเศษสามประเภทนี้ ต่อไปนี้คือตัวอย่างบันทึกข้อมูลศัพท์และข้อสังเกตในการใช้งานบันทึกข้อมูลศัพท์ซึ่งอยู่ในภาคผนวก ค

ภาพที่ 5 ตัวอย่างบันทึกข้อมูลศัพท์

TR001	Eng: Recharge (CEB01.txt)	Thai: การเติมน้ำ (CTB01.txt)
Grammatical Category: noun	Subject Field: Artificial Recharge > Basic Elements	
Definition: การเพิ่มเติมน้ำบาดาลลงในชั้นน้ำบาดาล อาจเกิดขึ้นโดยกระบวนการธรรมชาติ ที่น้ำฝนตกลงมาแล้วไหลลงสู่ชั้นน้ำบาดาล หรือโดยการกระทำของมนุษย์ ผ่านทางการสูบน้ำลงไปเติมในชั้นน้ำบาดาล การเติมน้ำแบ่งออกเป็นสองประเภทคือ การเพิ่มเติมน้ำตามธรรมชาติ และการเติมน้ำลงชั้นน้ำบาดาล (R01)		
Illustration: Recharge can significantly increase the sustainable yield of an aquifer. (CEP04.txt)		
Note: -		
Linguistic Specification: -		
Cross-reference: Artificial Recharge (การเติมน้ำลงชั้นน้ำบาดาล) (TR006), Natural Recharge (การเพิ่มเติมน้ำตามธรรมชาติ) (TR007)		

ข้อสังเกตในการใช้งานบันทึกข้อมูลศัพท์

1. ข้อมูลที่บันทึกในบันทึกข้อมูลศัพท์ส่วนใหญ่เป็นข้อมูลที่ได้คลังข้อมูลภาษา และจากบันทึกมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น (Conceptual Network and Extraction Records) โดยแต่ละบันทึกข้อมูลศัพท์จะใช้ข้อมูลจากบันทึกมโนทัศน์สัมพันธ์และข้อมูลศัพท์ที่มีหมายเลขบันทึกตรงกัน เช่น TR001 ก็จะใช้ข้อมูลจาก CR001 เป็นต้น ดังนั้นเมื่อต้องการตรวจสอบถึงที่มาของข้อมูลก็สามารถดูได้จากหมายเลขอ้างอิงที่ตรงกันดังกล่าว

2. หลังข้อมูลบางอย่างจะมีเลขอ้างอิงกำกับไว้ เพื่อให้สามารถอ้างอิงกลับไปถึงแหล่งที่มาได้

ศัพท์ภาษาอังกฤษ จะมีรหัสอ้างอิงแสดงที่มาของข้อมูลว่ามาจากไฟล์ใดในคลังข้อมูลภาษา ในกรณีที่ปรากฏอยู่หลายไฟล์จะใส่เฉพาะไฟล์แรกที่พบเท่านั้น

ศัพท์ภาษาไทย มีรหัสอ้างอิงเช่นเดียวกับภาษาอังกฤษ แตกต่างไปตรงที่ในกรณีที่นำมาจากศัพท์เดิมที่มีใช้อยู่แล้วในคลังข้อมูลภาษาก็จะวงเล็บรหัสอ้างอิงของไฟล์ไว้ เช่น 'น้ำบาดาล (CTB01.txt)' เป็นต้น ศัพท์เดิมที่รับมาใช้โดยผ่านเกณฑ์การยอมรับและความคุ้นเคย และเกณฑ์ความถูกต้องเหมาะสมจะไม่มีเชิงอรรถกำกับหลังแหล่งอ้างอิง แต่ศัพท์เดิมที่ไม่ผ่านเกณฑ์การยอมรับจะมีสัญลักษณ์ (+) ปรากฏแหล่งอ้างอิงของศัพท์นั้นๆ อีกที เช่น 'ปริมาณสมดุลย์ (CTP01.txt)+' เป็นต้น นอกจากนั้น ถ้าเป็นศัพท์ที่ปรับแก้จากของเดิมจะมีสัญลักษณ์ (#) กำกับอยู่หลังศัพท์นั้นๆ แต่ถ้าเป็นศัพท์ที่บัญญัติขึ้นมาใหม่จะมีไม่มีสัญลักษณ์ใดๆ ต่อท้ายแต่จะกำกับไว้ด้วยเชิงอรรถเพื่อบอกถึงวิธีการและรายละเอียดในการบัญญัติ

นิยาม จะแสดงรหัสหนังสืออ้างอิงและรหัสของไฟล์ข้อมูลที่เป็นที่มาของนิยาม ซึ่งในการกำหนดนิยามนี้ เป็นการผสมผสานนิยามที่ได้จากคลังข้อมูลภาษา นำมาตรวจสอบและเสริมข้อมูลให้ชัดเจนยิ่งขึ้นด้วยข้อมูลจากหนังสืออ้างอิง โดยรหัสของหนังสืออ้างอิงที่ใช้ในบันทึกข้อมูลศัพท์นี้ มีดังนี้

R01 = Fetter, C.W. Applied Hydrogeology Glossary. Third Edition. New Jersey: Prentice Hall, 1994.

R02 = High Plains Aquifer Evaluation Project. Glossary[Online]. 2000. Available from: http://www.kgs.ukans.edu/Highplains/atlas/glossary.html[2000, July 17].

R03 = U.S.Geological Survey. Water Science Glossary of Terms[online]. 2000. Available from: http://www.ga.usgs.gov/edu/dictionary.html[2000, July 17].

3. ข้อมูล Subject Field แบ่งตามมิติที่แสดงไว้ในบันทึกมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น โดยเป็น 5 หมวดหลักๆ ดังนี้

A1 = Relationship with relevant sciences

A2 = Basic Elements

A3 = Reasons to artificial recharge

A4 = Evolution of artificial recharge

A5 = Methods of artificial recharge

A6 = Type of recharge water

แต่ละหมวดก็แบ่งเป็นสายย่อยๆ ลงไปได้อีกดังนี้

A1 = Relationship with relevant sciences

A2 = Basic Elements

A3 = Reasons to artificial recharge

- Causes of artificial recharge
  - Overdraft
    - Land Subsidence
    - Saltwater Intrusion
    - Declining Water Level

A4 = Evolution of artificial recharge

- Water Filtration
- Deep-well Injection
- ASR

A5 = Methods of artificial recharge

- Surface Recharge
- Well Recharge
- Induced Recharge

A6 = Type of recharge water

ในการนำเสนอจะเขียนในรูปแบบของหมวดเรื่องไล่จากกลุ่มใหญ่ไปหากลุ่มย่อยที่มโนทัศน์นั้นๆ อยู่ เช่น Infiltration rate อยู่ในหมวด Surface recharge ซึ่งอยู่ในกลุ่มความสัมพันธ์ Recharge method การนำเสนอก็จะออกมาในรูปของ Methods of artificial recharge > Surface recharge เป็นต้น

4. ในส่วนของ Linguistic Specification มีตัวย่อดังนี้

Syn. = Synonym ของศัพท์หลัก

Ant. = Antonym ของศัพท์หลัก

5. ภาพประกอบที่ใช้ในบันทึกข้อมูลศัพท์นี้ ส่วนใหญ่คือภาพประกอบที่ใช้ในบันทึก มโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้นมาแล้วนั่นเอง เลขอ้างอิงจึงใช้ของเดิม

6. การนำเสนอของบันทึกข้อมูลศัพท์นี้ จะอยู่ในรูปของการจัดเป็นชุดศัพท์ตามความสัมพันธ์ โดยแบ่งเป็น ห้ากลุ่ม ตามที่แจกแจงไว้ในบันทึกมโนทัศน์สัมพันธ์ โดยในหน้าก่อนจะเริ่มบันทึกข้อมูล

ศัพท์แต่ละชุด จะมีการสรุปชุดศัพท์อยู่ในชุดนั้นๆ และเพื่อง่ายต่อการค้นหา ได้มีการจัดทำดัชนีค้นศัพท์ทั้งหมดไว้ท้ายภาคผนวก ค แล้ว

(ดูบันทึกข้อมูลศัพท์ได้ที่ ภาคผนวก ค)

สรุปผลบันทึกข้อมูลศัพท์และบทวิเคราะห์

สรุปผลจากบันทึกข้อมูลศัพท์

เมื่อจัดทำบันทึกข้อมูลศัพท์เสร็จสิ้น ได้ข้อสรุปเกี่ยวกับข้อมูลศัพท์ในประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาลดังนี้

- จำนวนบันทึกข้อมูลศัพท์ตรงกับบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น และเลขบันทึกทั้งคู่ก็ตรงกัน

- จำนวนมโนทัศน์ที่มีการนำเสนอด้วยศัพท์ภาษาไทยมากกว่าหนึ่งศัพท์ มีทั้งหมด 21 มโนทัศน์ ส่วนมโนทัศน์ที่มีศัพท์ภาษาอังกฤษสื่อถึงมากกว่าหนึ่งศัพท์มีทั้งหมด 26 มโนทัศน์

- มโนทัศน์ที่ไม่มีศัพท์ภาษาไทยสื่อถึงเลย มีทั้งหมด 15 มโนทัศน์ (มีศัพท์ภาษาอังกฤษสื่อถึงทุกมโนทัศน์)

- ศัพท์ทั้งหมดมี 96 ศัพท์ เป็นศัพท์เดี่ยว (Single Term) 23 ศัพท์ เป็นศัพท์ที่มีโครงสร้างซับซ้อน (Complex Term) 73 ศัพท์ ซึ่งประกอบด้วย 2-word term จำนวน 57 ศัพท์ และ multiword term จำนวน 16 ศัพท์

- โครงสร้างของศัพท์ ประเภท 2-word term มีลักษณะดังนี้

noun + noun เช่น Storm Runoff, Groundwater Mound, Surface Water เป็นต้น

Adj. + noun เช่น Hydrologic Cycle, Natural Recharge, Potable Water เป็นต้น

- โครงสร้างศัพท์ประเภท Multiword term มีลักษณะดังนี้

noun + noun + noun เช่น Water Treatment Plant เป็นต้น

noun + of + noun เช่น Cone of Recharge, Cone of Discharge เป็นต้น

noun + of + noun + noun เช่น Swelling of Soil Particle เป็นต้น

บทวิเคราะห์

จากผลการจัดทำบันทึกข้อมูลศัพท์ที่ปรากฏข้างต้น พบข้อสังเกตที่น่าสนใจดังนี้

- มโนทัศน์ที่มีศัพท์ภาษาไทยสื่อถึงมากกว่าหนึ่งศัพท์นั้น มีเป็นจำนวนมากถึง 21 มโนทัศน์ เมื่อเทียบกับศัพท์ภาษาอังกฤษซึ่งมีถึง 26 มโนทัศน์ที่มีศัพท์สื่อถึงมากกว่าหนึ่งศัพท์แล้ว แม้ศัพท์ไทยจะมีจำนวนศัพท์ที่ซ้ำซ้อนน้อยกว่า แต่ข้อมูลภาษาอังกฤษเป็นข้อมูลที่มีที่มาหลากหลายกว่าข้อมูลภาษาไทย คือมาจากหลายประเทศที่มีการศึกษาในเรื่องศาสตร์การเติมน้ำลงชั้นน้ำบาดาล ซึ่งเป็นธรรมดาสำหรับศาสตร์ที่ยังค่อนข้างใหม่เช่นนี้ ที่จะมีชุดศัพท์ที่ใช้ในแต่ละประเทศต่างกันไป แต่สำหรับภาษาไทยแล้ว เป็นการรวบรวมข้อมูลที่มาจากแหล่งเดียว คือกลุ่มนักวิชาการและนักเทคนิคที่ทำงานเกี่ยวข้องกับศาสตร์การเติมน้ำลงชั้นน้ำบาดาลในประเทศไทย และคนกลุ่มนี้ก็ยังมีจำนวนไม่มากนักและทำงานเกี่ยวข้องกันเองเป็นส่วนใหญ่ ศัพท์ที่ใช้จึงน่าจะมีความเป็นอันหนึ่งอันเดียวกันมากกว่าศัพท์ภาษาอังกฤษที่ใช้สื่อสารกันข้ามประเทศ แต่ผลที่ออกมากลับเป็นตรงข้าม ซึ่งก็ชี้ให้เห็นว่าแต่ละกลุ่มทำงานที่เกี่ยวข้องกับศาสตร์การเติมน้ำลงชั้นน้ำบาดาลในประเทศไทยยังไม่มีความเป็นอันหนึ่งอันเดียวกันในการสื่อสาร ทั้งที่เป็นการสื่อสารในวงแคบแต่กลับมีปัญหาการสื่อมโนทัศน์ด้วยศัพท์ที่ต่างกันไปหลายรูปแบบเช่นนี้ ชี้ให้เห็นถึงปัญหาสำคัญในอนาคตของการพัฒนาศาสตร์ด้านนี้ในประเทศไทย และนัยที่แฝงอยู่ก็คือลักษณะนิสัยการใช้ภาษาของคนไทย ที่กำหนดศัพท์ขึ้นใช้เพื่อสนองความจำเป็นในการสื่อสารของตนเองเป็นหลัก โดยไม่สนใจว่าได้มีการใช้ศัพท์อื่นที่สื่อถึงมโนทัศน์ดังกล่าวแล้วหรือยัง ส่วนใหญ่มักจะบัญญัติขึ้นใหม่ตามความสะดวกของตน จึงทำให้เกิดความสับสนในการสื่อสารเช่นนี้ ซึ่งถ้าปล่อยให้ปัญหาลักษณะนี้ดำเนินต่อไป อาจเป็นอุปสรรคสำคัญอย่างหนึ่งต่อพัฒนาการของศาสตร์การเติมน้ำลงชั้นน้ำบาดาลในประเทศไทยอย่างแน่นอน

- โครงสร้างของรูปศัพท์ที่ปรากฏในประมวลศัพท์ มีลักษณะที่ชี้ให้เห็นถึงการสร้างศัพท์ในศาสตร์เฉพาะด้านนี้ โครงสร้างศัพท์ที่ปรากฏมีลักษณะเป็นศัพท์ที่มีความซับซ้อนประกอบไปด้วยคำมากกว่าหนึ่งคำมารวมกัน (73ศัพท์) มากกว่าจะเป็นศัพท์เดี่ยว (23 ศัพท์) สาเหตุที่เป็นเช่นนี้

น่าจะเป็นเพราะ ศาสตร์เรื่องการเติมน้ำลงชั้นน้ำบาดาลเป็นศาสตร์ที่พัฒนามาจากศาสตร์เรื่องน้ำบาดาลและวิศวกรรมน้ำ การกำหนดศัพท์จึงนำศัพท์เดิมที่มีใช้อยู่แล้วในศาสตร์ทั้งสองมารวมกันเพื่อสร้างศัพท์ใหม่ ตัวอย่าง Groundwater Mound มาจากการนำศัพท์ 'Groundwater'ที่มีใช้อยู่แล้วในศาสตร์น้ำบาดาล มาผนวกกับศัพท์ 'Mound' ที่เป็นลักษณะของตะกอนที่พบอยู่ก้นบ่อบาดาลในเรื่องวิศวกรรมน้ำ มารวมกันเป็นศัพท์ใหม่เพื่อใช้สื่อถึงภาวะที่ตะกอนดินที่สะสมตัวกันอยู่ก้นแอ่งน้ำที่ใช้ในการเติมน้ำทางผิวดินทำให้อัตราการเติมน้ำลดลงเนื่องจากการอุดตันของตะกอนดินดังกล่าว เป็นต้น จะเห็นได้ว่าลักษณะของศัพท์ที่นำมารวมกันนี้ มักจะได้เป็นศัพท์ใหม่ที่มีความหมายตามศัพท์เดิมที่นำมารวมเข้าด้วยกัน เช่น Well Clogging มาจาก 'Well' ที่มีความหมายถึงบ่อบาดาล บวกกับ 'Clogging' ที่หมายถึงการติดขัดเนื่องจากมีสิ่งอุดตัน เมื่อมารวมกันก็มีความหมายถึงภาวะที่บ่อเติมน้ำเกิดการอุดตัน เป็นต้น แต่กระนั้นก็ยังมีศัพท์ที่เกิดจากคำมารวมกันแต่มีความหมายผิดไปจากรูปศัพท์เดิม เช่น Water Table มาจากคำว่า 'water' ที่แปลว่าน้ำ กับ'table'ที่แปลว่าโต๊ะ แต่เมื่อนำมารวมกันแปลได้ว่าระดับน้ำในชั้นน้ำเปิด เป็นต้น นอกจากนั้น ศัพท์บางส่วนก็ยังเกิดจากการนำกริยามาเติม –ing เพื่อทำให้เป็นนาม แล้วนำมารวมกับคำนามเป็นศัพท์ใหม่ ลักษณะนี้ปรากฏอยู่มากในศาสตร์เรื่องนี้ เช่น Pumping Well, Collecting Ditch, Monitoring Well เป็นต้น หรือนำกริยามาเปลี่ยนให้อยู่ในรูปของคำขยาย เช่น Confined Aquifer, Induced Recharge เป็นต้น ลักษณะเช่นนี้ชี้ให้เห็นว่าศัพท์ในชุดนี้ผู้บัญญัติต้องการบัญญัติให้ง่ายต่อการใช้งาน เมื่อมองดูศัพท์แล้วสามารถเข้าใจได้ทันทีว่าน่าจะสื่อถึงมโนทัศน์อะไร ดังนั้นการบัญญัติศัพท์จึงเน้นความสะดวกและประสิทธิภาพในการใช้งาน มากกว่าความประณีตสวยงามของรูปภาษา

ปัญหาในเรื่องการจัดทำบันทึกข้อมูลศัพท์

ขั้นตอนการบันทึกข้อมูลศัพท์ เป็นส่วนของการวิเคราะห์ข้อมูลดิบและข้อมูลเบื้องต้นเพื่อให้ได้ผลลัพธ์นำมาบันทึกไว้เป็นข้อมูลของประมวลศัพท์ ปัญหาในส่วนของข้อมูลจะไม่พบมากเท่าไรนัก ส่วนที่ต้องใช้ข้อมูลดิบในการวิเคราะห์มีน้อยกว่าส่วนที่อาศัยข้อมูลพื้นฐานเบื้องต้นที่ผ่านการวิเคราะห์วิจัยมาบ้างแล้ว ดังนั้นปัญหาที่ประสบในขั้นตอนนี้ส่วนใหญ่จึงอยู่ที่การจัดรูปแบบของประมวลศัพท์มากกว่าปัญหาในการวิเคราะห์ข้อมูล

- *ปัญหาการจัดชุดศัพท์แบบอิงมโนทัศน์ (Concept-Orientation)* การจัดทำประมวลศัพท์ครั้งนี้ใช้วิธีการเรียงชุดศัพท์แบบที่ Sager (1990) เรียกว่า การจัดชุดศัพท์แบบอิงมโนทัศน์

(Concept-Orientation) ซึ่งเป็นการจัดแบบอาศัยมโนทัศน์เป็นหลักสำคัญ ศัพท์ใดๆ ที่แสดงมโนทัศน์เดียวกันก็จะจัดมาอยู่ในบันทึกเดียวกัน วิธีการนี้มีประโยชน์ในแง่ของการใช้งานสำหรับนักวิชาการและช่วยให้ข้อมูลที่ได้เป็นระเบียบ แต่อาจเป็นอุปสรรคต่อผู้ใช้งานทั่วไปที่อาจพบศัพท์ได้หลากหลาย บางศัพท์อาจไม่ได้ปรากฏอยู่ในรูปศัพท์หลัก แต่เป็นคำเหมือน (synonym) ซึ่งกลุ่มเป้าหมายที่เป็นบุคคลทั่วไปอาจไม่มีความรู้พื้นฐานว่าศัพท์ที่พบเป็น synonym ของศัพท์ใด จึงทำให้หาศัพท์ดังกล่าวไม่พบ วิธีการแก้ปัญหาทำได้โดยการจัดให้มีดัชนีค้นคำไว้ท้ายประมวลศัพท์ โดยในดัชนีนี้จะแสดงศัพท์ทุกศัพท์ที่ปรากฏในบันทึกมโนทัศน์ ไม่ว่าจะเป็นศัพท์หลักหรือรูปศัพท์อื่น เช่น synonym, variant หรือ abbreviation ก็ตาม

- *การคัดเลือกองค์ประกอบของบันทึกประมวลศัพท์* องค์ประกอบของบันทึกประมวลศัพท์ที่หนังสือตำราต่างๆ แนะนำไว้นั้น มีอยู่มากมายหลายรูปแบบ การจะเลือกว่าองค์ประกอบใดจะเหมาะกับการทำงานในแต่ละกรณีก็ไม่ใช่เรื่องง่าย วิธีแก้ปัญหาก็คือ พิจารณาคุณสมบัติของข้อมูลที่ต้องการก่อน ข้อมูลใดบ้างจึงจะเป็นประโยชน์ต่อกลุ่มเป้าหมายผู้ใช้งาน โดยคงองค์ประกอบหลักๆ ที่จำเป็นในประมวลศัพท์ไว้ เช่น ศัพท์หลัก นิยาม เป็นต้น จากนั้น จึงพิจารณาเลือกองค์ประกอบที่เข้าข่ายคุณสมบัติที่ต้องการต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย