

บทที่ 4

การสร้างมโนทัศน์สัมพันธ์ (Conceptual Network)

หลังจากที่ได้คลังข้อมูลภาษา (Corpus) แล้ว ขั้นตอนต่อมาในการจัดทำประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาลคือ การสร้างมโนทัศน์สัมพันธ์ ซึ่งนับเป็นขั้นตอนที่สำคัญที่สุดอีกขั้นตอนหนึ่ง ขั้นตอนนี้ต้องใช้เวลาและความรู้เกี่ยวกับศาสตร์เฉพาะด้านอย่างสูง มโนทัศน์สัมพันธ์ที่ได้จากขั้นตอนนี้จะเป็นส่วนสำคัญยิ่งในบันทึกข้อมูลศัพท์ (Terminological Records) ซึ่งเป็นขั้นตอนต่อไป

ว่าด้วยเรื่องมโนทัศน์สัมพันธ์

มโนทัศน์สัมพันธ์ (Conceptual Network) เป็นสิ่งที่นักศัพทวิทยาให้ความสำคัญมาโดยตลอด ทั้งในฐานะที่เป็นการแสดงออกถึงขอบเขตของศาสตร์เฉพาะด้าน และเป็นตัวชี้วัดมโนทัศน์ (Concept) ของศัพท์เฉพาะด้านหรือความสัมพันธ์ที่ศัพท์เหล่านั้นมีต่อกัน ความสัมพันธ์ของมโนทัศน์นี้เองที่จะเป็นการพิสูจน์ให้เห็นว่าชุดศัพท์ในประมวลศัพท์เป็นชุดศัพท์ชุดเดียวกันและมีความเกี่ยวข้องสัมพันธ์กันอยู่ในศาสตร์เดียวกัน

มโนทัศน์สัมพันธ์ประกอบไปด้วยมโนทัศน์ (Concept) ที่มีศัพท์ (Terms) เป็นสัญลักษณ์แทนมโนทัศน์เหล่านั้น อย่างไรก็ตาม นักศัพทวิทยา เช่น Cabré (1998) มองว่า มโนทัศน์มีความเป็นอิสระจากศัพท์และเกิดขึ้นมาก่อนศัพท์ การสร้างมโนทัศน์สัมพันธ์คือการดึงเอามโนทัศน์ทั้งหมดของศาสตร์เฉพาะด้านออกมาแล้วใช้ศัพท์เป็นสัญลักษณ์ในการสื่อมโนทัศน์นั้นๆ ออกมา โดยมีการอธิบายอย่างชัดเจนถึงองค์ประกอบของมโนทัศน์นั้นๆ และความสัมพันธ์ที่มีกับมโนทัศน์อื่นๆ การอธิบายนี้อาจทำได้ด้วยการเขียนอธิบายเป็นตัวอักษร (Verbal Forms) หรือด้วยสื่ออื่นๆ (Nonverbal Visual Forms) เช่น แผนภาพ หรือรูปภาพ เป็นต้น แต่โดยมากจะใช้ประกอบกัน

นอกจากมโนทัศน์สัมพันธ์จะทำหน้าที่แสดงความสัมพันธ์ระหว่างมโนทัศน์ในศาสตร์เฉพาะด้านแล้ว ในขั้นตอนการทำมโนทัศน์สัมพันธ์ยังเป็นการช่วยตรวจสอบด้วยว่ามีศัพท์ใดที่หลุดลอยหายตาของผู้ทำประมวลศัพท์ไปได้บ้าง เพราะในกรณีที่มีมโนทัศน์ใดที่ขาดหายไปจากระบบความไม่ต่อเนื่องหรือช่องโหว่ของระบบมโนทัศน์สัมพันธ์จะพ้องถึงความผิดพลาดที่เกิดขึ้นทันที เป็นการเตือนว่ามีมโนทัศน์ที่ขาดไปจากระบบ ตัวอย่างเช่น เมื่อทำมโนทัศน์สัมพันธ์ของกลุ่มวิธีการเติมน้ำผ่านบ่อ (Well Recharge) แต่ไม่ได้พูดถึงบ่อเติมน้ำ (Recharge Well) การแสดงมโนทัศน์สัมพันธ์

ของชุดศัพท์จะสะดุดและไม่ต่อเนื่องไปทันที เป็นต้น มโนทัศน์แต่ละมโนทัศน์ในศาสตร์เรื่องเดียวกัน จะมีความสัมพันธ์ที่สอดคล้องประสานซึ่งเกี่ยวกันไปเหมือนตาข่าย ซึ่งเมื่อแสดงมโนทัศน์หนึ่งออกมา ก็จะเชื่อมโยงให้เห็นถึงอีกมโนทัศน์หรืออีกหลายมโนทัศน์ตามมา ทำให้เห็นความเกี่ยวเนื่องกันของมโนทัศน์ในศาสตร์นั้นๆ และเป็นการบังคับให้ประมวลศัพท์ที่ใช้วิธีการนี้มีชุดศัพท์ที่ครบถ้วนและมีความสัมพันธ์เกี่ยวข้องกันทั้งหมดอยู่ในเครือข่ายของเรื่องศาสตร์เฉพาะด้านนั้นๆ

นอกจากนั้น การสร้างมโนทัศน์สัมพันธ์ยังช่วยให้ผู้จัดทำเกิดความเข้าใจในศาสตร์เฉพาะด้านมากขึ้น โดยเฉพาะผู้จัดทำที่ไม่ใช่ผู้เชี่ยวชาญในศาสตร์เฉพาะด้าน การจัดระบบความสัมพันธ์ระหว่างมโนทัศน์จะช่วยเสริมความรู้เกี่ยวกับศาสตร์เฉพาะด้านนั้นๆ ซึ่งก็จะเป็นการเพิ่มประสิทธิภาพในการทำงาน และยังช่วยให้สื่อสารกับผู้เชี่ยวชาญศาสตร์เฉพาะด้านได้เข้าใจมากขึ้นอีกด้วย นับเป็นประโยชน์โดยตรงต่อคุณภาพของงานวิจัย

โดยรวมแล้ว การสร้างมโนทัศน์สัมพันธ์จะช่วยให้เราสามารถควบคุมคุณภาพของประมวลศัพท์ที่ออกมาได้ อีกทั้งชุดศัพท์ในประมวลศัพท์ก็จะมีความสัมพันธ์ที่เกี่ยวข้องเชื่อมโยงกันชัดเจนและครอบคลุมศัพท์ในส่วนที่เกี่ยวข้องได้ทั้งหมด ทั้งนี้ด้วยวิธีการจัดระบบการเข้าถึงศาสตร์เฉพาะด้านโดยอาศัยความเชื่อมโยงกันของมโนทัศน์ที่เป็นระบบ นอกจากนั้น ด้วยการสร้างมโนทัศน์สัมพันธ์ เรายังอาจได้นิยามของศัพท์รวมถึงรายละเอียดอื่นๆ มาบันทึกไว้ในบันทึกข้อมูลศัพท์เบื้องต้น อีกทั้งระบบศัพท์ที่เชื่อมโยงกันเป็นเครือข่ายยังช่วยให้การค้นข้อมูลทำได้สะดวกและง่ายขึ้น ท้ายสุด การมีมโนทัศน์สัมพันธ์ยังจะช่วยให้การบัญญัติศัพท์ไทยใหม่ให้สอดคล้องกับศัพท์เดิมที่มีอยู่แล้วในศาสตร์เดียวกันได้ ซึ่งประโยชน์ข้อนี้จะได้นำเสนอรายละเอียดต่อไปในบทที่ 6

วิวิธวิทยาในการสร้างมโนทัศน์สัมพันธ์

แม้ว่าในทางทฤษฎีแล้ว นักศัพท์วิทยาจะถูกคาดหวังให้สร้างระบบความสัมพันธ์ของมโนทัศน์ขึ้นมาก่อน แล้วจึงจัดรูปศัพท์ให้กับมโนทัศน์ในระบบนั้นๆ และในขั้นตอนต่อมาจึงจะดึงข้อมูลเกี่ยวกับศัพท์ออกมาจากข้อมูลภาษาแม่เก็บไว้ในบันทึกข้อมูลศัพท์เบื้องต้น (Extraction Records) (Sager, 1990: 21) แต่ในทางปฏิบัติแล้ว กระบวนการนี้เป็นกระบวนการที่เชื่อมโยงเป็นวงจร เราไม่สามารถแบ่งขั้นตอนทั้งหมดออกจากกันโดยเด็ดขาดได้ มโนทัศน์นั้นแสดงออกอยู่แล้วในรูปศัพท์ในคลังข้อมูลภาษา ดังนั้นการจะแสดงความสัมพันธ์ระหว่างมโนทัศน์ได้ เราจำเป็นต้องดึงเอาสัญลักษณ์แทนมโนทัศน์หรือศัพท์ (Terms) ออกมาจากคลังข้อมูลภาษาให้ได้ก่อน การเลือกชุดศัพท์จากคลังข้อมูลภาษาในขั้นตอนนี้ จะนำมาเป็นชุดศัพท์พื้นฐานในการสร้างมโนทัศน์สัมพันธ์ อย่างไรก็ตาม

ตาม ระบบมโนทัศน์สัมพันธ์และชุดศัพท์อาจมีการปรับเปลี่ยนได้อยู่ตลอดเวลาในช่วงที่มีการสร้างมโนทัศน์สัมพันธ์ เนื่องจากเมื่อมีการกำหนดความสัมพันธ์ของมโนทัศน์ไปเรื่อยๆ และพบว่ายังมีมโนทัศน์บางส่วนที่ขาดหายไป ก็ทำให้ต้องกลับไปค้นคลังข้อมูลภาษาเพื่อหาศัพท์ที่แทนมโนทัศน์นั้นๆ ซึ่งเป็นการตรวจสอบความบกพร่องในขั้นตอนการเลือกศัพท์ครั้งแรกไปในตัว และในระหว่างขั้นตอนการรวบรวมชุดศัพท์และจัดระบบความสัมพันธ์ของมโนทัศน์นี้ ต้องอาศัยการอ้างอิงข้อมูลจากคลังข้อมูลภาษา (Corpus) ซึ่งข้อมูลเหล่านั้นนอกจากจะแสดงความสัมพันธ์ระหว่างมโนทัศน์แล้ว ยังแสดงรายละเอียดเกี่ยวกับศัพท์ เช่น นิยาม หรือตัวอย่างการใช้งาน เป็นต้น นับเป็นข้อมูลที่จะบันทึกไว้ในบันทึกข้อมูลศัพท์เบื้องต้น (Extraction Records) จะเห็นได้ว่าขั้นตอนต่างๆ วนเวียนเป็นวงจรที่เชื่อมโยงกันไปมา ซึ่งผู้ทำประมวลศัพท์ต้องทำเกี่ยวพันกันไป ไม่สามารถทำขั้นตอนใดขั้นตอนหนึ่งให้สำเร็จเสร็จสิ้นเด็ดขาดไปเลยได้ ลักษณะการทำงานในส่วนของการสร้างมโนทัศน์สัมพันธ์นี้จึงค่อนข้างแตกต่างจากขั้นตอนอื่น อย่างไรก็ตามเพื่อความเข้าใจที่ง่าย ในรายงานการวิจัยครั้งนี้จึงจะขอแสดงวิธีวิทยาการสร้างมโนทัศน์สัมพันธ์โดยแบ่งเป็นขั้นตอนใหญ่ๆ สามขั้นตอน และจะได้พูดในรายละเอียดของแต่ละขั้นตอนต่อไป แต่สิ่งที่ต้องคำนึงถึงก็คือขั้นตอนแต่ละขั้นตอนไม่ได้จบลงในตัว แต่อาจต้องกลับมาปฏิบัติซ้ำอีกหมุนเวียนไปมาจนกว่าการสร้างมโนทัศน์สัมพันธ์จะเสร็จสมบูรณ์

ขั้นตอน การเลือกชุดศัพท์

ในการเลือกชุดศัพท์จากคลังข้อมูลภาษา สิ่งที่ต้องคำนึงถึงก็คือ ศัพท์ (Terms) ต่างจากคำ (Words) และเราต้องแยกแยะความแตกต่างให้ได้ เพื่อให้สามารถคัดเลือกชุดศัพท์เฉพาะด้านที่แท้จริงออกมาได้ แม้ว่าในขั้นตอนการสร้างคลังข้อมูลภาษา จะได้มีการเลือกข้อมูลที่เป็นศาสตร์เฉพาะด้านมารวบรวมไว้อยู่แล้ว แต่ศัพท์ที่ปรากฏในคลังข้อมูลภาษาก็ไม่ใช่ศัพท์เฉพาะด้านทั้งหมด ดังนั้นนอกจากการแยกแยะระหว่างศัพท์และคำแล้ว ผู้วิจัยยังต้องสามารถรู้ได้ด้วยว่าศัพท์ใดเป็นศัพท์เฉพาะด้านที่ต้องการบ้าง

นิยามคร่าวๆ ของ ศัพท์ และความแตกต่างของศัพท์กับคำ ได้รับการสรุปไว้ในงานของ Sager (1990) ว่า

The items which are characterised by special reference within a discipline are the 'term' of that discipline, and collectively they form its 'terminology'; those which function in general reference over a variety of sublanguages are simply called 'words', and their totality the 'vocabulary'.
(Sager, 1990: 19)

ซึ่งเกณฑ์ที่ Sager ใช้ก็คือความเกี่ยวพันของศัพท์กับศาสตร์เฉพาะด้าน แต่ในการจะชี้ชัดว่าอะไรคือสิ่งที่ตัดสินความสัมพันธ์ดังกล่าวได้ ก็ยังเป็นสิ่งที่ไม่ชัดเจน อย่างไรก็ตามนิยามของ Sager นับว่าได้ช่วยให้เห็นความแตกต่างระหว่างศัพท์กับคำได้ในระดับหนึ่ง

ต่อมา Pearson (1998) ได้พยายามประมวลนิยามของ 'ศัพท์' และเกณฑ์วัดความแตกต่างระหว่างศัพท์กับคำ ที่นักศัพทวิทยาและนักภาษาศาสตร์หลายๆ ท่านได้เคยรวบรวมไว้ และได้รายงานผลสรุปไว้ในหนังสือ Terms in Context ว่า ยังไม่มีเกณฑ์ใดที่สามารถใช้เป็นบรรทัดฐานโดยไม่มีข้อโต้แย้งได้ Pearson (1998) เสนอวิธีการใหม่ขึ้นมา ซึ่งประกอบด้วยเกณฑ์สองอย่าง คือ Generic Reference Criterion และ Co-occurrence Criterion โดยเกณฑ์แรกจะเป็นการวัดความเป็นสากลของศัพท์ โดยมีสมมุติฐานว่าศัพท์เป็นคำที่อ้างถึงกลุ่ม (class) ของสิ่งทั่วไปไม่ใช่สิ่งเฉพาะเจาะจง จึงปรากฏในรูปที่มี indefinite article นำหน้า หรือไม่มีอะไรนำหน้าเลย อีกเกณฑ์หนึ่งคือ Co-occurrence Criterion จะถือว่าในกรณีที่คำปรากฏในบริบทที่มีลักษณะเฉพาะที่กำหนดไว้ มากกว่าสองครั้ง คำดังกล่าวก็มีแนวโน้มที่จะเป็นศัพท์ เช่น คำที่ปรากฏอยู่ท้ายประโยค แล้วประโยคต่อมาขึ้นต้นด้วยวลีอย่าง This process, This method หรือ The Device ก็มีความเป็นไปได้ที่จะเป็นศัพท์ได้เป็นต้น

เมื่อลองนำวิธีการดังกล่าวไปทดลองใช้งานพบว่า เกณฑ์แรกคือ Generic Reference Criterion สามารถแยกแยะศัพท์ออกจากคำได้ เช่น Artificial Recharge มีการใช้โดดๆ ไม่มีอะไรนำหน้า ก็เป็นศัพท์เฉพาะด้าน หรือ คำว่า Aquifer มี indefinite article นำหน้าพบเห็นทั่วไปในบริบท ก็เป็นศัพท์เฉพาะด้าน เป็นต้น แต่การจะพิจารณาคำที่ละคำเพื่อดูว่าเป็นคำที่ลักษณะดังกล่าวหรือเปล่าก็เป็นวิธีที่ใช้เวลาและอาจเกิดความผิดพลาดได้ง่าย และถ้าเป็นกรณีที่คลังข้อมูลภาษามีขนาดใหญ่มากๆ การใช้เกณฑ์ดังกล่าวจะทำให้เสียเวลามาก ดังนั้น วิธีการนี้ยังไม่น่าจะเพียงพอในการคัดเลือกศัพท์ ส่วนเกณฑ์ที่สอง คือ Co-occurrence Criterion เมื่อปฏิบัติจริงพบว่าสามารถใช้คัดเลือกศัพท์ได้จริงเช่นเดียวกับเกณฑ์แรก เช่น ตัวอย่างสองตัวอย่างที่ยกมาจากคลังข้อมูลภาษา

- (1) Table 11.2 summarizes representative rates obtained by Basin Recharge. In this method water is distributed... (CEB04.txt)
- (2) In this circumstance, basin recharge is the most effective one. This method involves surface spreading...(CEP04.txt)

ตัวอย่างทั้งสองเป็นไปตามเงื่อนไขที่กำหนดโดยเกณฑ์ดังกล่าวที่ว่า ในกรณีที่มีคำที่ถูกอ้างถึงอีกครั้งด้วยคำอย่าง 'this method' หรือ 'the process' ก็มีความเป็นไปได้สูงที่จะเป็นศัพท์ ซึ่งตามตัวอย่างก็แสดงให้เห็นว่า Basin Recharge เป็นศัพท์ แต่เกณฑ์นี้ก็มีข้อเสีย โดยเฉพาะเมื่อนำมาใช้กับข้อมูลในศาสตร์การเติมน้ำลงชั้นน้ำบาดาล เพราะศาสตร์นี้มีความเกี่ยวข้องกับศาสตร์อื่นๆ ที่ใกล้เคียงกันสูง เช่น วิศวกรรมน้ำ ธรณีวิทยา น้ำบาดาล เป็นต้น ซึ่งศัพท์เฉพาะด้านในศาสตร์เหล่านี้ก็มีคุณสมบัติเป็นศัพท์ตามเกณฑ์นี้ได้ทั้งหมด ทำให้เกิดปัญหาที่ไม่สามารถแยกแยะได้ว่าศัพท์ใดเป็นศัพท์เฉพาะของศาสตร์เรื่องการเติมน้ำลงชั้นน้ำบาดาล นอกจากนั้นเกณฑ์นี้ก็ยังปัญหาเช่นเดียวกับเกณฑ์แรกคือ ต้องใช้การพิจารณาเป็นรายประโยคไป ทำให้เสียเวลามาก ดังนั้นการใช้เพียงเกณฑ์ทั้งสองของ Pearson (1998) จึงยังไม่เพียงพอ

ทัศนะของนักศัพทวิทยาท่านต่อมา เกี่ยวกับการกำหนดเกณฑ์การคัดเลือกศัพท์ ก็คือวิธีการของ Alain Rey (1995) ที่ว่าการที่คำๆ ใดจะถูกเรียกว่า 'ศัพท์' ได้ จะต้องมีความแตกต่างจากคำอื่นๆ ในแง่ที่เป็นองค์ประกอบของศาสตร์เฉพาะด้านเรื่องใดเรื่องหนึ่ง ซึ่งการแสดงถึงความแตกต่างนี้ทำได้โดยกระบวนการให้คำนิยามของศัพท์นั้นๆ (Rey, 1995: 40-41)

เมื่อลองนำแนวคิดดังกล่าวมาประยุกต์ใช้ในทางปฏิบัติ โดยการอาศัยรูปภาษาที่แสดงให้เห็นถึงการให้คำนิยามหรือความหมายของศัพท์ โดยตั้งสมมติฐานว่า ศัพท์ที่ผู้เขียนใช้ในแง่ที่ให้คำนิยาม หรือพยายามให้ความหมายเพื่อให้ผู้อ่านเข้าใจถึงมโนทัศน์ของศัพท์นั้นๆ มีความเป็นไปได้ที่จะเป็นศัพท์เฉพาะด้าน ซึ่งเมื่อทดลองใช้ดูก็ได้ผลดี เช่น *Artificial recharge may be defined as the practice of increasing, by artificial means, the amount of water that enters a groundwater reservoir.* (CEB03.txt) การพบรูปประโยคที่ผู้เขียนพยายามให้คำอธิบายถึงศัพท์บางศัพท์ ทำให้สันนิษฐานได้ว่าศัพท์ดังกล่าวเป็นศัพท์เฉพาะด้านในศาสตร์ที่ข้อมูลนั้นอ้างอิงถึง

อย่างไรก็ตาม การแยกแยะระหว่างศัพท์และคำด้วยการพิจารณาจากโครงสร้างการให้นิยาม จะทำให้ขั้นตอนการคัดเลือกชุดศัพท์ต้องใช้ความรู้เกี่ยวกับศาสตร์เฉพาะด้านเข้ามาช่วยอย่างมาก เพื่อให้ทราบว่าคำใดมีความหมายในศาสตร์เฉพาะด้านและมีคุณสมบัติเป็นศัพท์เฉพาะด้านของเรื่องการเติมน้ำลงชั้นน้ำบาดาล ซึ่งก็เป็นปัญหาที่ต้องหาทางแก้ไขในขั้นตอนการลงมือปฏิบัติจริงต่อไป และนอกเหนือจากการเลือกศัพท์แล้ว อีกสิ่งที่ต้องคำนึงก็คือการแยกแยะศัพท์เฉพาะด้านเรื่องการเติมน้ำลงชั้นน้ำบาดาลออกจากศาสตร์เฉพาะด้านอื่นๆ ในศาสตร์ที่กว้างกว่าหรือใกล้เคียงกัน อย่างเช่น วิศวกรรมน้ำ (Water Engineering) น้ำบาดาลศาสตร์ (Groundwater Study) หรือ

เทคโนโลยีการบำบัดน้ำ (Water Treatment) เป็นต้น ซึ่งมีความเป็นไปได้สูงที่คลังข้อมูลภาษาที่รวบรวมข้อมูลเกี่ยวกับเรื่องการเติมน้ำลงชั้นน้ำบาดาลอาจจะรวมเอาศัพท์จากศาสตร์ใกล้เคียงดังกล่าวเอาไว้ด้วย เพราะเป็นศาสตร์ที่มีความเกี่ยวข้องกันมากและมักจะถูกกล่าวหาตีพิมพ์รวมๆ กันไปด้วยเสมอ อีกทั้งยังมีการใช้ศัพท์เฉพาะด้านบางส่วนที่ซ้ำซ้อนกันอยู่ด้วย ซึ่งการแยกแยะศัพท์ดังกล่าวออกมาได้ก็หนีไม่พ้นการใช้ความรู้เกี่ยวกับศาสตร์เฉพาะด้านที่มากพอ ซึ่งในขั้นตอนนี้ผู้ทำวิจัยจำเป็นต้องอาศัยความช่วยเหลือจากผู้เชี่ยวชาญ (Specialist) อย่างมาก ทั้งในแง่ของการตรวจสอบความถูกต้องและการให้คำแนะนำ

มาถึงขั้นตอนนี้ การผสมผสานเกณฑ์ทั้งสองของ Pearson (1998) กับเกณฑ์การดูรูปภาษาที่ให้นิยาม ทำให้ได้ชุดหลักเกณฑ์ที่มีประสิทธิภาพและเหมาะกับการทำวิจัยครั้งนี้ แต่เกณฑ์ทั้งสามก็ยังไม่สามารถแก้ปัญหาด้านหนึ่งได้ นั่นก็คือ ปัญหาเรื่องความล่าช้าของการไล่ดูศัพท์ไปที่ละคำที่คลังข้อมูลภาษา ซึ่งก็เป็นการใช้เวลามากและอาจเกิดความผิดพลาดได้ง่าย เช่น ซ้ำคำศัพท์บางคำศัพท์ไปเพราะความหลงลืม เป็นต้น ดังนั้น ในขั้นแรกจึงต้องมีการคัดศัพท์ที่สันนิษฐานว่าเป็นศัพท์เฉพาะด้านออกมาก่อน แล้วจึงนำมาตรวจสอบเพื่อคัดเลือกด้วยนิยามต่อไป การตั้งสมมติฐานขั้นต้นนี้ ผู้ทำวิจัยได้ใช้ความถี่ที่ศัพท์ปรากฏในคลังข้อมูลภาษาเป็นหลัก สมมติฐานก็คือ ศัพท์ที่พบบ่อยในคลังข้อมูลภาษาเฉพาะด้าน มีความเป็นไปได้สูงที่จะเป็นศัพท์เฉพาะด้านในเรื่องนั้นๆ

การค้นหาศัพท์ที่มีความถี่สูง 200 อันดับแรกจากคลังข้อมูลทำได้โดยการใช้โปรแกรม Win Concordance ให้แสดงศัพท์ที่พบในคลังข้อมูลภาษาออกมา และให้เรียงลำดับตามความถี่ โดยศัพท์ชุดนี้ได้ผ่านการตัดส่วนที่เป็น Article และ Preposition ออกแล้ว ให้เหลือเฉพาะที่เป็นศัพท์ล้วนๆ จากนั้นจึงให้โปรแกรม Win Concordance แสดงบริบทที่ปรากฏการใช้ศัพท์ในคลังข้อมูลภาษาไล่ไปที่ละศัพท์ เพื่อทำการตรวจสอบด้วยเกณฑ์ที่ประยุกต์มาจากเกณฑ์ทั้งสองของ Pearson (1998) และเกณฑ์การให้นิยามของ Rey (1995) โดยมีรายละเอียดดังนี้

(1) มีรูปภาษาที่ชี้ว่าเป็นศัพท์เฉพาะด้าน เช่น มีคำว่า This process, This method หรือ The Device ใช้แทนที่ศัพท์ดังกล่าว เป็นต้น ตัวอย่าง The process is often referred to as evaporation, but this leads to confusion as to whether the meteorological processes in the atmosphere are included. (CEB02.txt) ในกรณีนี้เรารู้ได้ว่า evaporation เป็นศัพท์เฉพาะด้าน เป็นต้น

(2) รูปภาษาแสดงถึงการให้คำนิยามหรือความหมายของศัพท์นั้นๆ เช่น The term groundwater is usually reserved for the subsurface water that occurs beneath the water table in soils and geologic formations that are fully saturated. (CEB02.txt) เป็นต้น

(3) ในกรณีที่มีคำนิยามให้มาด้วย ก็ใช้ความรู้ด้านศาสตร์เฉพาะด้าน พิจารณานิยามนั้นเลยว่าสื่อถึงมโนทัศน์ที่เกี่ยวข้องกับศาสตร์การเติมน้ำลงชั้นน้ำบาดาลหรือไม่ โดยอาจสอบถามกับผู้เชี่ยวชาญหรือตรวจสอบจากหนังสืออ้างอิง ตัวอย่าง Any withdrawal in excess of safe yield is an overdraft . (CEB02.txt) เมื่อพิจารณาดูแล้วเป็นมโนทัศน์ที่เกี่ยวข้องในเรื่องการเติมน้ำลงชั้นน้ำบาดาล ดังนั้นจึงตัดคำว่า overdraft ไว้เป็นศัพท์เฉพาะด้านได้ เป็นต้น

(4) เกี่ยวข้องในแง่ไวยากรณ์กับศัพท์เฉพาะด้าน เช่น เป็นรูปกริยาของศัพท์ที่เป็นศาสตร์เฉพาะด้าน ตัวอย่าง Too much discharge overdrafted this confined aquifer. (CEB02.txt) จะเห็นได้ว่า overdraft นี้เป็นรูปกริยา (verb) ของ overdraft ที่เป็นนาม และได้รับการพิจารณาให้เป็นศัพท์เฉพาะด้านไปแล้วด้วยเกณฑ์อื่น ดังนั้นรูปกริยาของศัพท์เฉพาะด้าน ก็น่าจะเป็นศัพท์เฉพาะด้านด้วย เป็นต้น

(5) มีลักษณะที่ใช้ได้ทั่วไป เป็นอิสระจากบริบท เช่น มี Indefinite article นำหน้า หรือไม่มี article นำหน้าเลย เช่น Obtain suspended-solids removal by filtration through the ground. (CEB03.txt) คำว่า filtration ไม่มี article นำหน้า สันนิษฐานได้ว่าน่าจะเป็นศัพท์ เป็นต้น

ในหลายๆ กรณีการใช้เกณฑ์การตรวจสอบเพียงเกณฑ์เดียวอาจไม่เพียงพอ เช่น ใช้เกณฑ์ที่ (5) เพียงเกณฑ์เดียวอาจไม่เพียงพอในการตัดสินว่าศัพท์ดังกล่าวเป็นศัพท์เฉพาะด้านหรือไม่ ต้องใช้เกณฑ์อื่นๆ มาช่วยเช่นเกณฑ์การตรวจสอบนิยาม หรือเกณฑ์รูปศัพท์ เป็นต้น

ผลจากการใช้เกณฑ์ทั้งหมดในกระบวนการตรวจสอบชุดศัพท์ 200 ศัพท์ที่เลือกตามความถี่ พบว่ามีศัพท์ที่ผ่านเกณฑ์เป็นศัพท์เฉพาะด้าน ทั้งหมด 28 ศัพท์ คิดเป็นร้อยละ 14 จากทั้งหมด 200 ศัพท์ และทั้งหมดอยู่ในรูปของศัพท์เดี่ยว (Single term) สาเหตุที่เป็นเช่นนี้ ไม่ใช่เพราะวิธีการคัดเลือกด้วยความถี่ไม่มีคุณภาพ เพราะเมื่อพิจารณาจากชุดศัพท์จริงที่ใช้ในประมวลศัพท์ที่เสร็จสมบูรณ์แล้ว มีศัพท์ที่อยู่ในรูปศัพท์เดี่ยวทั้งหมด 32 ศัพท์เท่านั้น ซึ่งเมื่อเทียบกับจำนวน 28 ศัพท์ที่ได้จากการคัดเลือกด้วยความถี่และเกณฑ์เบื้องต้นแล้ว ถือว่าได้ผลเกือบร้อยละ 80 จึงถือว่ากระบวนการคัดเลือกชุดศัพท์ที่ใช้นี้มีประสิทธิภาพในระดับที่น่าพอใจ ปัญหาอีกด้านก็คือ ศัพท์ที่ได้จากการใช้โปรแกรม Win Concordance ดึงออกมาแสดงตามความถี่ ทำได้เฉพาะศัพท์ที่เป็นศัพท์เดี่ยวเท่านั้น ไม่สามารถใช้กับศัพท์ที่ประกอบด้วยมากกว่าหนึ่งคำได้ (Complex terms) ได้ ซึ่งก็แก้ปัญหาดังกล่าว

ได้ด้วยการใช้โปรแกรม N-gram Frequency List ซึ่งสามารถดึงศัพท์ที่ประกอบด้วยคำมากกว่าหนึ่งคำได้ตามที่ต้องการ แล้วนำชุดศัพท์ได้มาใหม่นี้มาให้โปรแกรม Win Concordance แสดงบริบทที่พบศัพท์ดังกล่าวเพื่อตรวจสอบด้วยเกณฑ์ทั้งห้าข้ออีกครั้ง และได้ชุดศัพท์เพิ่มขึ้นมาอีก 57 ศัพท์ รวมเป็นทั้งหมด 85 ศัพท์ นับเป็นร้อยละ 82 ของชุดศัพท์จริง (ทั้งหมด 96 ศัพท์) ซึ่งแสดงให้เห็นว่าเกณฑ์การคัดเลือกศัพท์ด้วยความถี่ถ้ามีการควบคุมด้วยเกณฑ์การคัดเลือกอื่นๆ ที่เหมาะสมด้วย ก็อาจใช้เป็นเกณฑ์พื้นฐานในการคัดเลือกชุดศัพท์ได้

ชุดศัพท์ที่ได้มาทั้งหมด 85 ศัพท์นี้ แม้จะได้ผ่านการตรวจสอบด้วยเกณฑ์ต่างๆ ดังที่กล่าวมาแล้ว ก็ถือเป็นการตรวจสอบโดยคร่าวๆ เท่านั้น เพราะยังขาดการตรวจสอบในแง่ของความครบถ้วน ว่าศัพท์ที่ได้มามีความเหมาะสมในทัศนทั้งหมดในเรื่องศาสตร์การเติมน้ำลงชั้นน้ำบาดาลแล้วหรือยัง ซึ่งจะได้มีการตรวจสอบในเรื่องนี้ต่อไปในขั้นตอนการสร้างมโนทัศน์สัมพันธ์ ซึ่งอาจทำให้ต้องมีการเพิ่มศัพท์เพื่อใช้สื่อถึงมโนทัศน์ที่หายไป หรืออาจต้องมีการตัดศัพท์บางศัพท์ออกไป เพราะสื่อถึงมโนทัศน์ที่ไม่เกี่ยวข้องกับศาสตร์การเติมน้ำลงชั้นน้ำบาดาลอย่างแท้จริง

ขั้นตอน การสร้างมโนทัศน์สัมพันธ์

ขั้นตอนนี้เป็นการนำชุดศัพท์เบื้องต้นที่ได้มาจากคลังข้อมูลภาษาและมีการตรวจสอบจากบริบทโดยคร่าวๆ แล้วว่าเป็นศัพท์ที่อยู่ในเรื่องการเติมน้ำลงชั้นน้ำบาดาล มาเป็นตัวแทนมโนทัศน์ แล้วจัดระบบความสัมพันธ์ของมโนทัศน์เหล่านั้นให้เป็นระบบเพื่อสื่อให้เห็นถึงขอบเขตและแง่มุมทุกอย่างของศาสตร์เรื่องการเติมน้ำลงชั้นน้ำบาดาล

ในการกำหนดความสัมพันธ์ของมโนทัศน์ต้องอาศัยความรู้ความเข้าใจในศาสตร์เฉพาะด้าน ซึ่งการนำเสนอระบบความสัมพันธ์ของมโนทัศน์เหล่านี้ อาจทำได้หลายรูปแบบ ขึ้นอยู่กับความเหมาะสมตามแต่กรณีไป เช่น อาจนำเสนอด้วยคำอธิบาย (Description) แผนภาพ (Network) รูปภาพ (Illustration) เป็นต้น และรูปแบบการนำเสนอแต่ละอย่างก็ยังมีรายละเอียดปลีกย่อยอีก เช่น แผนผังที่แสดงความเชื่อมโยงระหว่างมโนทัศน์อาจเป็นในรูปแผนภูมิต้นไม้ (Tree Diagram) หรืออาจจะอยู่ในรูปแบบหลายมิติ (Multidimensionality) ซึ่งแสดงให้เห็นมโนทัศน์ในมิติต่างๆ ของศาสตร์เฉพาะด้านนั้นๆ

ในการทำวิจัยครั้งนี้ ผู้ทำวิจัยได้เลือกใช้วิธีการนำเสนอระบบความสัมพันธ์ของมโนทัศน์ต่างๆ ในเรื่องการเติมน้ำลงชั้นน้ำบาดาลด้วยวิธีการหลายรูปแบบ แต่ที่เลือกใช้เป็นรูปแบบหลัก

คือ แผนภาพแสดงความเชื่อมโยงทั้งแบบธรรมดาและหลายมิติประกอบกับคำบรรยายเป็นลายลักษณ์อักษร ซึ่งจากการค้นคว้าในเรื่องการเติมน้ำลงชั้นน้ำบาดาล พบว่าศาสตร์เรื่องนี้มีมิติที่เหลื่อมซ้อนกันอยู่หลายมิติด้วยกัน เช่น มิติของวิธีการเติมน้ำที่มีรูปแบบการปฏิบัติหลายรูปแบบด้วยกัน มิติที่มาของการเติมน้ำซึ่งแสดงถึงสาเหตุปัญหาที่ทำให้ต้องเติมน้ำ หรือมิติพัฒนาการของการเติมน้ำลงชั้นน้ำบาดาลว่ามีการวิวัฒนาการมาอย่างไร เป็นต้น การที่จะแสดงออกถึงมิติทั้งหลายได้อย่างครบถ้วนจึงต้องอาศัยการแสดงออกด้วยแผนภาพแบบหลายมิติ แต่ในส่วนของความสัมพันธ์ย่อยๆ ในแต่ละมิติที่ไม่มีความซับซ้อนมากนักก็จะให้แสดงด้วยแผนภาพธรรมดาหรือรูปภาพประกอบคำบรรยาย แม้ว่าในการจัดทำบันทึกมโนทัศน์สัมพันธ์จะใช้แผนภาพเป็นหลักในการแสดงความสัมพันธ์ แต่บางครั้งรูปภาพก็มีความจำเป็นในการบรรยายมโนทัศน์ในเรื่องการเติมน้ำลงชั้นน้ำบาดาล โดยเฉพาะในส่วนของกระบวนการเติมน้ำ ซึ่งการใช้คำบรรยายอย่างเดียวอาจทำให้เกิดความสับสนได้ แต่อย่างไรก็ตาม ก็มีข้อควรระวังในการใช้รูปภาพ โดยควรใช้รูปภาพเป็นส่วนประกอบช่วยในการอธิบายเสริมความเข้าใจเท่านั้น ไม่ใช่นำมาใช้แทนที่คำนิยามของศัพท์นั้นๆ ไปเลย ซึ่งในเรื่องนี้ Catherine Bowman, Diane Michaud, และ Heidi Suonuuti (1997) ได้เคยให้ข้อคิดไว้ในบทความชื่อ Do's and Don'ts of Terminology Management ว่า รูปภาพใช้ช่วยในการอธิบายความหมายได้ แต่ไม่อาจใช้แทนที่การให้คำนิยามที่ถูกต้องได้ (Cited in Wright & Budin, 1997: 216)

เพื่อให้การจัดสร้างมโนทัศน์มีระบบเพื่อการอ้างอิงในขั้นตอนต่อไป จึงได้มีการวางรูปแบบของการบันทึกมโนทัศน์สัมพันธ์ ไว้ในรูปของแบบฟอร์ม หนึ่งในแบบฟอร์มบันทึกหนึ่งมโนทัศน์ องค์ประกอบของแบบฟอร์มจะเป็นการแจกแจงรายละเอียดของศัพท์และการแสดงความสัมพันธ์ของมโนทัศน์นั้นๆ กับมโนทัศน์อื่นๆ ในเรื่องการเติมน้ำลงชั้นน้ำบาดาล โดยมีรายละเอียดดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 1 แบบฟอร์มบันทึกมโนทัศน์สัมพันธ์

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 5 | | | |
| 6 | | | |

1 = ลำดับ (no.)
 2 = มโนทัศน์ (Concept)
 3 = ศัพท์ภาษาอังกฤษ (English Terms)
 4 = ศัพท์ภาษาไทย (Thai Terms)
 5 = รายละเอียดศัพท์ (Terms Features)
 6 = ความสัมพันธ์กับมโนทัศน์อื่น (Conceptual Relation)

เรื่องสำคัญอีกอย่างหนึ่งในขั้นตอนการสร้างมโนทัศน์สัมพันธ์นี้ ก็คือ การกำหนดรูปแบบความสัมพันธ์ของมโนทัศน์ (Conceptual Relations) มโนทัศน์ทั้งหลายในศาสตร์เฉพาะด้านมีความสัมพันธ์กันในหลายๆ รูปแบบ Sager (1995) เคยให้ความเห็นไว้ในเรื่องนี้ว่า ในทางปฏิบัติแล้วประเภทของความสัมพันธ์นี้จะมีกี่แบบก็ได้ ขึ้นอยู่กับความต้องการใช้งานและความเหมาะสมตามแต่กรณีไป (Sager, 1990: 35) ซึ่งจากการประมวลรูปแบบความสัมพันธ์ที่ปรากฏอยู่ตามตำราเกี่ยวกับศัพท์วิทยาต่างๆ สามารถแบ่งรูปแบบความสัมพันธ์ออกเป็นสองประเภทใหญ่ๆ คือ

(1) รูปแบบความสัมพันธ์มาตรฐาน (Traditional Relationship) ความสัมพันธ์ในกลุ่มนี้คือความสัมพันธ์ที่มีการวางรูปแบบความสัมพันธ์ไว้แล้วตายตัว และมักจะมีผู้นำไปใช้ในการทำประมวลศัพท์อยู่เป็นประจำ ลักษณะเด่นคือ เป็นรูปแบบความสัมพันธ์ที่พบบ่อยในประมวลศัพท์ทั่วไป ซึ่งรวมทั้งประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาลด้วย ตัวอย่างความสัมพันธ์แบบนี้ เช่น

- Inclusion Relationship หรือ Generic Relationship (Generic-Specific) เป็นความสัมพันธ์ระหว่างสองมโนทัศน์ที่มโนทัศน์หนึ่งมีขอบเขตที่กว้างกว่า และอีกมโนทัศน์หนึ่งมีขอบเขตที่แคบกว่าและเป็นส่วนหนึ่งของมโนทัศน์แรก โดยที่มโนทัศน์หลังจะมีคุณสมบัติเช่นเดียวกับมโนทัศน์แรกทุกประการ แต่ก็จะมีคุณสมบัติเฉพาะตัวที่ทำให้แตกต่างจากมโนทัศน์แรก เช่น Flower-rose ซึ่ง rose มีลักษณะเป็น flower ทุกประการ แต่จะมีคุณสมบัติพิเศษที่ทำให้มีความแตกต่างจาก Flower ทั่วไป คือ rose มีสีแดง มีหนาม มีกลิ่นหอม เป็นต้น

- Coordination Relationship หรือ Partitive Relationship (Part-Whole) เป็นรูปแบบความสัมพันธ์ที่มีโน้ตศน์หนึ่งมีส่วนประกอบหลายส่วน และอีกมโน้ตศน์หนึ่งก็เป็นหนึ่งในส่วนประกอบนั้นๆ เช่น Car-Wheel

- Chain Relationship (Cause-Effect) ความสัมพันธ์นี้จะมีมโน้ตศน์หนึ่งที่เป็นผลมาจากอีกมโน้ตศน์หนึ่ง เช่น Explosion-Fallout

ซึ่งในการจัดทำประมวลศัพท์ครั้งนี้ ได้ใช้รูปแบบความสัมพันธ์มาตรฐานทั้งหมด 3 รูปแบบความสัมพันธ์ด้วยกัน และมีรายละเอียดดังนี้

Generic - Specific แสดงความสัมพันธ์ในแบบมาตรฐาน ระหว่างมโน้ตศน์หนึ่งที่มีขอบเขตกว้างครอบคลุมอีกมโน้ตศน์หนึ่ง โดยมีมโน้ตศน์ที่เล็กกว่าจะมีคุณสมบัติเดียวกับมโน้ตศน์ที่ใหญ่กว่าทุกประการ แต่ในขณะที่เดียวกันก็มีคุณสมบัติเฉพาะตัวที่ต่างออกไป ตัวอย่างเช่น Artificial recharge – Surface recharge ซึ่ง Artificial recharge สื่อถึงมโน้ตศน์รวมของการเติมน้ำลงชั้นน้ำบาดาล มีขอบเขตครอบคลุมมโน้ตศน์ของ Surface recharge ซึ่งเป็นวิธีการเติมน้ำลงชั้นน้ำบาดาลวิธีหนึ่ง อย่างไรก็ตาม Surface recharge ก็มีลักษณะทั่วไปของ Artificial recharge ทุกประการ คือ เป็นการนำน้ำลงไปเติมในชั้นน้ำบาดาล เพียงแต่จะมีลักษณะเฉพาะตัวที่ต่างไปคือ เป็นวิธีการเติมน้ำที่ทำจากผิวดินเท่านั้น เป็นต้น

Whole – Part แสดงความสัมพันธ์ที่มีมโน้ตศน์หนึ่งเป็นส่วนหนึ่งที่ประกอบเป็นอีกมโน้ตศน์หนึ่ง ตัวอย่าง Recharge Well – Well Casing โดยที่ Well Casing เป็นส่วนประกอบหนึ่งของ Recharge Well เป็นต้น

Cause – Effect แสดงความสัมพันธ์ในรูปที่มีมโน้ตศน์หนึ่งเป็นสาเหตุที่ทำให้เกิดอีกมโน้ตศน์หนึ่ง หรือมองได้ว่ามโน้ตศน์หนึ่งทำให้เกิดอีกมโน้ตศน์หนึ่งนั่นเอง ตัวอย่าง Overdraft - Saltwater intrusion โดยที่ Overdraft เป็นสาเหตุที่ทำให้เกิดวิกฤตการณ์ Saltwater intrusion เป็นต้น

(2) **รูปแบบความสัมพันธ์ซับซ้อน (Complex Relationship)** ได้แก่รูปแบบความสัมพันธ์ที่กำหนดขึ้นเพื่อตอบสนองต่อปัญหาและความต้องการในแต่ละกรณี ความสัมพันธ์ของมโน้ตศน์อาจมีได้หลายแบบตามแต่สถานการณ์ใช้งาน Sager (1995) ได้ยกตัวอย่างความสัมพันธ์รูปแบบนี้หลายรูปแบบ เช่น (Sager, 1995: 35)

Material – Product เช่น Steel - Girder

Material – Property เช่น Glass - Brittle

| | |
|-------------------------|---------------------------------------|
| Material – State | เช่น Iron - Corrosion |
| Process – Product | เช่น Weaving - Cloth |
| Process – Instrument | เช่น Incision - Scalpel |
| Process – Method | เช่น Water Recharge – Water Injection |
| Object – Material | เช่น Bridge - Iron |
| Object – Quality | เช่น Petrol – High Octane |
| Object – Operation | เช่น Drill bit - Drilling |
| Object – Characteristic | เช่น Fuel - Smokeless |
| Object – Form | เช่น Book – Paperback |
| Activity – Place | เช่น Coalmining – Coalmine |

ซึ่งความสัมพันธ์เหล่านี้เป็นเพียงตัวอย่างเท่านั้น สามารถเพิ่มเติมได้ในระหว่างกระบวนการ ซึ่งการจัดความสัมพันธ์เช่นนี้ขึ้นอยู่กับมุมมองของผู้ใช้งาน (Facet) ซึ่งอาจมีได้หลายมุมมอง ไม่มีรูปแบบตายตัวแต่อย่างใด ซึ่งในกระบวนการจัดทำประมวลศัพท์เรื่องการเติมน้ำลงชั้นน้ำบาดาล ได้กำหนดรูปแบบความสัมพันธ์ซับซ้อนขึ้นเพื่อใช้ในการแสดงรูปแบบความสัมพันธ์ของมโนทัศน์ต่างๆ ในเรื่องนี้ไว้ทั้งหมด 9 ความสัมพันธ์ ได้แก่

Problem – Solution แสดงความสัมพันธ์ที่มโนทัศน์หนึ่งเป็นปัญหาในขณะที่อีกมโนทัศน์หนึ่งเป็นวิธีแก้ปัญหานั้น ๆ ตัวอย่าง Saltwater intrusion – Artificial recharge โดยที่ Artificial recharge เป็นวิธีการหนึ่งที่ใช้แก้ปัญหามโนทัศน์ Saltwater intrusion ได้

Process – Target แสดงความสัมพันธ์ระหว่างมโนทัศน์ที่สื่อถึงกระบวนการกับอีกมโนทัศน์ที่เป็นเป้าหมายของกระบวนการดังกล่าว ตัวอย่าง Recharge – Aquifer โดยที่ Recharge สื่อถึงกระบวนการเติมน้ำลงชั้นน้ำบาดาล ส่วน Aquifer เป็นเป้าหมายของการเติมน้ำลงชั้นน้ำบาดาล เป็นต้น

Process – Material แสดงความสัมพันธ์ที่มโนทัศน์หนึ่งสื่อถึงกระบวนการ ส่วนอีกมโนทัศน์หนึ่งคือวัตถุดิบหรือสิ่งที่ใช้ในกระบวนการนั้นๆ ตัวอย่าง Infiltration – Surface Water โดยที่ Infiltration เป็นกระบวนการที่น้ำผิวดินไหลซึมลงไปในพื้นที่น้ำเปิด ซึ่งวัตถุดิบที่ใช้ในกระบวนการดังกล่าวก็คือ Surface Water เป็นต้น

Process – Method แสดงความสัมพันธ์ที่มีโน้ตคนหนึ่งสื่อถึงกระบวนการ อีกรโน้ตคนสื่อถึงวิธีการที่ใช้ในการดำเนินการกระบวนการนั้น ตัวอย่าง Natural Recharge - Infiltration โดยที่ Infiltration เป็นวิธีการที่ใช้ในกระบวนการ Natural Recharge เป็นต้น

Process – Operator แสดงความสัมพันธ์ระหว่างมโนทัศน์หนึ่งสื่อถึง กระบวนการกับอีกรโน้ตคนหนึ่งที่สื่อถึงผู้ดำเนินการกระบวนการดังกล่าว ตัวอย่าง Artificial recharge - Human โดยที่ Artificial recharge เป็นกระบวนการที่ต้องดำเนินการโดยมนุษย์เท่านั้น เป็นต้น

Process – Problem แสดงความสัมพันธ์ระหว่างมโนทัศน์ที่เป็นกระบวนการกับมโนทัศน์ที่เป็นปัญหาของกระบวนการนั้น ๆ เช่น Well recharge – Well Clogging เนื่องจาก Well Clogging เป็นปัญหาหนึ่งที่เกิดกับกระบวนการ Well recharge เป็นต้น

Object – Qualification แสดงความสัมพันธ์ระหว่างมโนทัศน์หนึ่งสื่อถึง คุณสมบัติของอีกรโน้ตคนหนึ่ง ตัวอย่าง Unsaturated Zone - Peameability โดยบริเวณที่เรียกว่า Unsaturated Zone มีคุณสมบัติที่โดดเด่นคือ Peameability สูง เป็นต้น

Evolution เป็นการแสดงความสัมพันธ์ในรูปแบบที่แตกต่างจากความสัมพันธ์มาตรฐานที่ใช้ในการจัดทำประมวลศัพท์ทั่วไป ทั้งนี้เพราะในศาสตร์เรื่องการเติมน้ำลงชั้นน้ำบาดาลมีรูปแบบความสัมพันธ์อีกรูปแบบหนึ่งที่มีความสำคัญ นั่นคือ กระบวนการที่เกิดการพัฒนาการ มาเป็นลำดับ ก่อนที่จะมาเป็นการเติมน้ำลงชั้นน้ำบาดาลในปัจจุบัน ดังนั้นเพื่อแสดงความสัมพันธ์ ในรูปแบบนี้ จึงต้องมีการกำหนดรูปแบบความสัมพันธ์นี้ขึ้นมา

ในการจัดทำมโนทัศน์สัมพันธ์ของเรื่องการเติมน้ำลงชั้นน้ำบาดาลนี้ ผู้ทำวิจัยได้เลือกใช้รูปแบบความสัมพันธ์ที่กล่าวมาโดยรวมๆ กันไป ทั้งรูปแบบความสัมพันธ์มาตรฐานและรูปแบบ ความสัมพันธ์ซับซ้อน และใช้อักษรย่อแสดงรูปแบบความสัมพันธ์ในแผนภาพ มีรายละเอียดดังนี้

B = Process – Problem

C = Cause – Effect

E = Process – Advanced process (Evaluation)

G = Generic – Specific

M = Process – Material

O = Object – Qualification

P = Problem – Solution

PM = Process – Method

PO = Process – Operator

PT = Process – Target

W = Whole – Part

เมื่อบันทึกมโนทัศน์สัมพันธ์ได้รับการบันทึกข้อมูลเสร็จสิ้นแล้ว ในระหว่างขั้นตอนนี้อาจมีการค้นพบว่ายังมีบางมโนทัศน์ที่ยังไม่ได้รวบรวมมา สิ่งที่จะชี้ให้เห็นข้อบกพร่องนี้ได้ก็คือระบบความสัมพันธ์ที่สร้างขึ้นนี้เอง โดยมีข้อสันนิษฐานว่า ถ้าการจัดระบบมโนทัศน์เป็นไปโดยครบถ้วนแล้วมโนทัศน์ทั้งหมดจะเชื่อมโยงถึงกันโดยไม่มีการขาดตอน ซึ่งถ้ามีก็หมายความว่าอาจมีมโนทัศน์ที่หลุดรอดสายตาไปในการรวบรวมศัพท์ชุดแรก และอาจทำให้ต้องกลับไปดำเนินการขั้นตอนการเลือกชุดศัพท์ใหม่อีกครั้ง เพื่อหาศัพท์ที่แสดงมโนทัศน์ที่ขาดไปให้ครบถ้วน หรือในระหว่างกระบวนการจัดสร้างมโนทัศน์สัมพันธ์ อาจพบเห็นมิติบางมิติของเรื่องที่เคยมองข้ามไปในครั้งแรก ด้วยเหตุนี้ การดำเนินการในขั้นตอนแรกและขั้นตอนที่สองจึงอาจจะมีการดำเนินการที่เชื่อมโยงกันอยู่ อาจจำเป็นต้องดำเนินการกลับไปกลับมาจนกว่าจะได้มโนทัศน์สัมพันธ์ที่สมบูรณ์ตามวัตถุประสงค์การสร้างประมวลศัพท์

ขั้นตอน การบันทึกข้อมูลศัพท์เบื้องต้น (Extraction Records)

ตามทฤษฎีการทำประมวลศัพท์ที่ Cabré (1998) ได้เสนอไว้ ขั้นตอนต่อจากการสร้างมโนทัศน์สัมพันธ์ก็คือ การบันทึกข้อมูลศัพท์เบื้องต้นไว้ในบันทึกที่เรียกว่า บันทึกข้อมูลศัพท์เบื้องต้น (Extraction Records) โดยการดำเนินการในขั้นตอนนี้จะเป็นการค้นข้อมูลศัพท์จากคลังข้อมูลภาษา โดยอาศัยข้อมูลบริบทที่ศัพท์นั้นๆ ปรากฏเป็นหลัก แล้วนำข้อมูลดังกล่าว เช่น นิยามศัพท์ (Definition) หรือคำเหมือน (Synonym) เป็นต้น มาบันทึกไว้ ซึ่งบันทึกนี้จะเป็นเสมือนการรวบรวมข้อมูลศัพท์แต่ละศัพท์จากคลังข้อมูลภาษามาแยกไว้ต่างหาก แล้วใช้ประกอบกับข้อมูลมโนทัศน์สัมพันธ์ในการจัดทำบันทึกข้อมูลศัพท์ในขั้นตอนต่อไป ข้อดีของการแยกเป็นบันทึกย่อยๆ เช่นนี้ก็คือ ช่วยให้ข้อมูลมีความเป็นระเบียบเป็นสัดส่วน และสามารถหยิบยกมาใช้งานตามความต้องการเป็นเรื่องๆ ไป เช่น เมื่อต้องการข้อมูลเกี่ยวกับศัพท์ในแง่ของหน่วยภาษา ก็นำบันทึกข้อมูลศัพท์เบื้องต้นมาใช้ หรือเมื่อต้องการดูความสัมพันธ์ระหว่างมโนทัศน์ก็อาจนำบันทึกมโนทัศน์สัมพันธ์มาพิจารณาได้ เป็นต้น

อย่างไรก็ตาม การปฏิบัติเช่นนั้นก็มีข้อเสียที่โดดเด่น นั่นคือ การทำงานซ้ำซ้อน ทำให้เสียเวลา และในบางครั้งข้อมูลที่มีมากชุดอาจทำให้ใช้งานได้ไม่สะดวก ขั้นตอนการทำบันทึกข้อมูลศัพท์เบื้องต้นและการทำบันทึกมโนทัศน์สัมพันธ์มีส่วนที่ซ้อนทับกันอยู่ คือในส่วนของ การค้นข้อมูลศัพท์จากคลังข้อมูลภาษา ซึ่งต่างก็อาศัยการวิเคราะห์หาข้อมูลจากบริบทที่ศัพท์ปรากฏเป็นหลัก ผู้ทำวิจัยจึงเห็นว่าถ้าในขั้นตอนการทำมโนทัศน์สัมพันธ์ ได้มีการคัดข้อมูลจากคลังข้อมูลภาษามาทบันทึกได้เลย ก็จะช่วยประหยัดเวลาทำงานได้ อีกทั้งการรวมบันทึกทั้งสองเป็นบันทึกเดียวจะช่วยให้มีความคล่องตัวในการทำงานมากขึ้น เมื่อพิจารณาองค์ประกอบต่างๆ ที่ควรจะมีอยู่ในบันทึกข้อมูลศัพท์เบื้องต้นแล้ว ก็พบว่ามีการซ้ำซ้อนอยู่ในหลายประการซึ่งอาจรวมเข้าด้วยกันได้ ในรูปแบบต่อไปนี้



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 2 แบบฟอร์มบันทึกกรรมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 5 | | | |
| 6 | | | |
| 7 | | | |
| 8 | | | |

1 = ลำดับ (no.)
 2 = มโนทัศน์ (Concept)
 3 = ศัพท์ภาษาอังกฤษ (English Terms)
 4 = ศัพท์ภาษาไทย (Thai Terms)
 5 = รายละเอียดศัพท์ (Terms Features)
 6 = ความสัมพันธ์กับมโนทัศน์อื่น (Conceptual Relation)
 7 = ตัวอย่างการใช้งาน (Examples)
 8 = รูปไวยากรณ์ของศัพท์ (Grammatical Category)

ส่วนที่เพิ่มเติมจากรูปแบบเดิมขึ้นมาคือส่วนที่ 7 และ 8 ซึ่งในส่วนที่ 7 จะเป็นการยกตัวอย่างข้อมูล (Examples) ที่ได้จากคลังข้อมูลภาษา และส่วนที่ 8 เป็นการแสดงรูปทางไวยากรณ์ (Grammatical Category) ของศัพท์นั้นๆ ข้อมูลดังกล่าวได้มาจากคลังข้อมูลภาษา ข้อมูลที่จะนำมาบันทึกไว้ก็คือเป็นตัวอย่างข้อมูลที่จะเป็นตัวแทนของข้อมูลที่เกี่ยวข้องกับศัพท์นั้นๆ ซึ่งจะนำไปใช้ในขั้นตอนต่อไป ดังนั้นการเลือกข้อมูลที่เหมาะสมจึงเป็นสิ่งที่ควรให้ความสำคัญ

การพิจารณาเลือกข้อมูลมาบันทึกไว้ในส่วนนี้ นักศัพทวิทยาจะพิจารณาจากบริบทของข้อมูลเป็นสำคัญ โดยบริบทที่อาจนำมาบันทึกไว้ในส่วนนี้ได้ แบ่งออกเป็น สามประเภทด้วยกัน (Cabré, 1998: 138) คือ

- Defining Context - หมายถึงบริบทที่อาจให้นิยามของศัพท์นั้นๆ ทั้งหมดหรือเพียงบางส่วน บริบทประเภทนี้นับว่ามีคุณค่ามากที่สุด และเป็นบริบทประเภทที่ควรเลือกมาเก็บไว้ในบันทึกข้อมูลศัพท์มากที่สุด คำสำคัญ (Key words) ที่ชี้ว่าเป็นบริบทประเภท Defining Context เช่น is/are, was/were, means, means that, can be defined as เป็นต้น ตัวอย่างบริบทประเภทนี้

Groundwater discharge *can be defined as* the removal of water from the saturated zone across the water-table surface, together with the associated flow toward the water table within the saturated zone.

- **Metalinguistic Context** หรือ Implicative Context – หมายถึงบริบทที่แสดงข้อมูลอื่นๆ เกี่ยวกับศัพท์ในลักษณะที่เป็นหน่วยภาษา เช่น คำเหมือน (Synonym) หรือคำตรงข้าม (Antonym) เป็นต้น คำสำคัญที่อาจพบในบริบทประเภทนี้ เช่น also known as, called, may be named as, or, opposite to เป็นต้น บริบทประเภทนี้มีความสำคัญเป็นอันดับสองในการเลือกมาบันทึก ตัวอย่างบริบท:

A pumping cone, *or* cone of depression, will form in the aquifer around a pumping well as the water level declines.

- **Testimonial Context** หรือ Associative Context – หมายถึงบริบทที่ไม่ได้แสดงข้อมูลพิเศษใดๆ เพียงแต่แสดงให้เห็นถึงการปรากฏของศัพท์ในเรื่องเฉพาะด้านนั้นๆ เท่านั้น เพื่อเป็นการยืนยันว่าศัพท์นั้นๆ เป็นศัพท์เฉพาะด้านจริง และอาจใช้บริบทประเภทนี้เป็นตัวอย่างการใช้งานและรูปไวยากรณ์ของศัพท์ได้ บริบทประเภทนี้จะพบได้ทั่วไปในคลังข้อมูลภาษา ข้อสังเกตคือบริบททั้งหมดที่ไม่จัดอยู่ในประเภทที่ (1) และ (2) ก็จะจัดอยู่ในประเภทนี้ นักศัพทวิทยาจะเลือกบันทึกบริบทประเภทนี้ในกรณีที่ไม่มีพบบริบทสองประเภทแรก

ตามประเภทบริบทข้างต้น ผู้ทำวิจัยได้พิจารณาเลือกข้อมูลบริบทมาบันทึกไว้ โดยให้ความสำคัญกับประเภทของบริบทประเภท Defining Context และ Metalinguistic Context เป็นหลัก แต่ในกรณีที่ไม่มีสองประเภทดังกล่าว การบันทึก Testimonial Context ไว้อย่างน้อยสองตัวอย่างก็น่าจะเป็นการเพียงพอที่จะใช้พิสูจน์ยืนยันความเป็นศัพท์เฉพาะด้านของศัพท์นั้นๆ ได้ ดังนั้นข้อมูลที่เลือกมาบันทึกในส่วนนี้จึงจะมีอย่างน้อยสองตัวอย่างสำหรับศัพท์แต่ละศัพท์

ในส่วนของรูปไวยากรณ์ (ช่องที่ 8) ก็บันทึกตามที่ปรากฏจริงในคลังข้อมูลภาษา ในกรณีที่ศัพท์หนึ่งมีรูปไวยากรณ์หลายแบบ เช่น Recharge ที่เป็นทั้ง verb และ noun ก็จะมีบันทึกข้อมูลทั้งสองไว้ด้วยกัน เพราะมีการสื่อมโนทัศน์ที่เหมือนกัน

สรุปขั้นตอนวิธีวิทยาการสร้างมโนทัศน์สัมพันธ์ เมื่อสิ้นสุดขั้นตอนนี้ ก็จะได้ชุดมโนทัศน์สัมพันธ์ของศาสตร์เฉพาะด้านทั้งหมด รวมถึงชุดศัพท์ และรายละเอียดเบื้องต้นของศัพท์ด้วย ซึ่ง

ข้อมูลที่ได้จะมีทั้งข้อมูลแสดงมโนทัศน์สัมพันธ์ของชุดศัพท์ และรายละเอียดเบื้องต้นของศัพท์นั้นๆ ซึ่งรวมกันอยู่ในบันทึกมโนทัศน์สัมพันธ์ การรวมบันทึกข้อมูลศัพท์เบื้องต้นไว้ด้วยนี้ จะช่วยให้การทำวิจัยครั้งนี้ข้ามผ่านขั้นตอนของการบันทึกข้อมูลศัพท์เบื้องต้นไป ขั้นตอนต่อมาซึ่งเป็นขั้นตอนที่สำคัญยิ่งก็คือขั้นตอนการบันทึกข้อมูลศัพท์ (Terminological Records) ซึ่งจะได้กล่าวถึงในบทต่อไป ตัวอย่างของบันทึกรวมมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้นดูหน้าต่อไป

ข้อสังเกตในการศึกษาบันทึกมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น

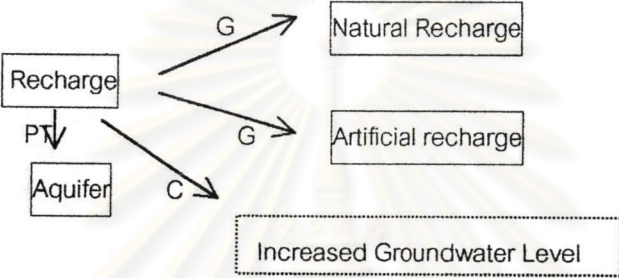
ในการจัดทำบันทึกมโนทัศน์สัมพันธ์และข้อมูลศัพท์นี้ ผู้จัดทำได้ประยุกต์ใช้รูปแบบใหม่บางประการในการบันทึก เพื่อความเข้าใจของผู้ใช้งานจึงขออธิบายวิธีการศึกษาบันทึกนี้โดยคร่าวๆ ดังนี้

- ในส่วนของหมายเลขบันทึก จะอยู่ในรูป CNxxx ซึ่ง CN เป็นตัวย่อที่ใช้แทน Conceptual Network ส่วนตัวเลขตามตัวหลังคือหมายเลขของบันทึก หมายเลขบันทึกนี้จะใช้อ้างอิงอยู่บ่อยครั้ง ทั้งในบันทึกนี้เองและในบันทึกข้อมูลศัพท์ (Terminological Records)
- ในส่วนของศัพท์อังกฤษ (Eng Term) ศัพท์ที่นำมาบันทึกไว้ คือศัพท์ทุกศัพท์ที่พบในคลังข้อมูลภาษาว่าสื่อถึงมโนทัศน์นั้นๆ ซึ่งอาจมีได้หลายศัพท์ ตัวอย่างเช่น Piezometric Surface, Artesian Water Level และ Potentiometric Surface ซึ่งล้วนแล้วแต่ใช้สื่อถึงมโนทัศน์เดียวกัน ดังนั้นในช่องศัพท์ภาษาอังกฤษจึงปรากฏศัพท์ทั้งสามนี้ พร้อมระบุที่มาของศัพท์ด้วยรหัสของบทความในคลังข้อมูลภาษาที่ศัพท์นั้นๆ ปรากฏใช้ เช่น Piezometric Surface (CEB01.txt) เป็นต้น (ที่มาของรหัสดังกล่าวดูได้ในส่วนของคลังข้อมูลภาษา ในภาคผนวก ก) และผลจากการที่มีศัพท์ภาษาอังกฤษหลายศัพท์นี้เอง ทำให้ในส่วนที่เป็นตัวอย่างการใช้งานที่ยกมาจากคลังข้อมูลภาษา (Extraction) เป็นตัวอย่างการใช้ศัพท์ทุกศัพท์ที่บันทึกไว้ในส่วนของศัพท์ภาษาอังกฤษด้วย นอกจากนี้ในส่วนของบันทึกมโนทัศน์ (Concept) จะเลือกศัพท์ที่ปรากฏใช้มากที่สุดในคลังข้อมูลภาษา ในบรรดาศัพท์ภาษาอังกฤษที่สื่อถึงมโนทัศน์นี้ทั้งหมดมาเป็นตัวแทน และศัพท์นี้เองก็จะปรากฏเป็นศัพท์ที่จะใช้เป็นศัพท์หลักในการแสดงมโนทัศน์นั้นๆ ในบันทึกข้อมูลศัพท์ต่อไป เช่น ถ้าตามตัวอย่างก่อน พบว่ามีการใช้ Piezometric Surface ในคลังข้อมูลภาษาทั้งหมด 19 ครั้ง Artesian Water Level มีการใช้ทั้งหมด 11 ครั้ง และ Potentiometric Surface มีการใช้ทั้งหมด 9 ครั้ง ดังนั้นจึงใช้ Piezometric Surface เป็นตัวแทนศัพท์ที่ใช้แทนมโนทัศน์และใช้เป็นศัพท์หลักในบันทึกข้อมูลศัพท์ต่อไป เป็นต้น

- ในส่วนของมโนทัศน์สัมพันธ์ (Conceptual Relation) จะมีการแสดงความสัมพันธ์ดังกล่าวในสามรูปแบบคือ แผนผังแสดงความสัมพันธ์ ภาพประกอบ และคำอธิบาย โดยเน้นหนักอยู่ที่แผนผังและคำอธิบาย ส่วนภาพประกอบอาจมีเฉพาะบางมโนทัศน์ที่มีการมีภาพประกอบอาจช่วยให้เข้าใจความสัมพันธ์มากขึ้นเท่านั้น แต่กระนั้นแม้ว่าจะมีรูปประกอบหรือไม่ก็ตาม ในส่วนของคำอธิบายก็ต้องมีอยู่เสมอในทุกๆ มโนทัศน์ และการแสดงความสัมพันธ์ด้วยแผนผังจะมีสองลักษณะคือ แผนผังรวมในแต่ละมิติ และแผนผังย่อยในแต่ละมโนทัศน์ ซึ่งเป็นการแสดงถึงมโนทัศน์ที่เชื่อมโยงกัน สาเหตุที่แบ่งเป็นแผนผังย่อยเพื่อให้เข้าใจง่ายและเห็นภาพรวม
- ในส่วนของ Features จะเป็นข้อมูลของศัพท์ที่ได้จากบริบทของศัพท์ที่ปรากฏในคลังข้อมูลภาษาเท่านั้น บางครั้งก็อาจได้นิยามของศัพท์ หรือบางครั้งอาจได้เพียงลักษณะการใช้งานบางอย่างเท่านั้น ดังนั้นในบันทึกนี้จึงยังไม่ใช้บทสรุปสุดท้ายของประมวลศัพท์ เป็นเพียงการวิเคราะห์ข้อมูลที่ปรากฏในการใช้งานจริงจากคลังข้อมูลภาษาเท่านั้น ข้อมูลที่ได้จึงอาจยังไม่สมบูรณ์
- มโนทัศน์บางส่วนที่ปรากฏในบันทึกนี้อาจไม่ใช่มโนทัศน์ที่ปรากฏในศาสตร์เฉพาะด้านเรื่องการเติมน้ำลงชั้นน้ำบาดาลเพียงศาสตร์เดียว แต่อาจเป็นศาสตร์เฉพาะด้านที่มีการใช้มากกว่าหนึ่งศาสตร์เฉพาะด้านขึ้นไป แต่ก็ล้วนแล้วแต่เกี่ยวข้องอยู่ในศาสตร์ที่ใกล้เคียงกับศาสตร์การเติมน้ำลงชั้นน้ำบาดาลทั้งสิ้น ตัวอย่างเช่น Precipitation ซึ่งสื่อถึงที่มาของน้ำที่ใช้ใน Surface recharge โดยแท้จริงแล้ว Precipitation ไม่ใช่ศัพท์เฉพาะด้านในศาสตร์การเติมน้ำลงชั้นน้ำบาดาลเท่านั้น แต่มีใช้ในศาสตร์น้ำบาดาล ซึ่งเป็นศาสตร์ที่ครอบคลุมศาสตร์การเติมน้ำลงชั้นน้ำบาดาลอีกชั้นหนึ่งด้วย เป็นต้น สาเหตุที่ต้องนำมโนทัศน์ลักษณะดังกล่าวมารวมอยู่ด้วยกันก็เพื่อให้มโนทัศน์ทั้งหมดมีการเชื่อมโยงต่อเนื่องกันไปได้ ทั้งนี้เพราะมโนทัศน์บางมโนทัศน์แม้จะปรากฏอยู่ในศาสตร์อื่นด้วยแต่ก็มีความสำคัญต่อการทำความเข้าใจมโนทัศน์สัมพันธ์ในเรื่องการเติมน้ำลงชั้นน้ำบาดาลอย่างมาก จนถ้าตัดมโนทัศน์ดังกล่าวไป อาจทำให้มโนทัศน์สัมพันธ์ที่ได้ไม่ครบวงจรได้ นอกจากนั้น แม้ว่ามโนทัศน์เหล่านั้นจะปรากฏในศาสตร์อื่น แต่ก็ยังเป็นศาสตร์ที่เกี่ยวข้องกันอยู่ในที่ และก็เป็นมโนทัศน์ที่ผู้ใช้งานประมวลศัพท์มีโอกาสพอเจอได้ง่ายอีกด้วย

จุฬาลงกรณ์มหาวิทยาลัย

ภาพที่ 3 ตัวอย่างบัณฑิตกมโนทัศน์สัมพันธ์และข้อมูลศัพท์เบื้องต้น

| | | | |
|--|-------------------|---|--|
| CN004 | Concept: Recharge | Eng = Recharge -n. (CEB01.txt), Recharge -v. (CES02.txt) | Th = การเติมน้ำ -n. (CTB01.txt), การเพิ่มเติมน้ำ -n. (CTB01.txt), อด น้ำ -v. (CTB01.txt), เพิ่มเติมน้ำ - v. (CTB01.txt) |
| Features: การเพิ่มเติมน้ำลงสู่ Aquifer แบ่งออกเป็นสองประเภทย่อย คือ Natural Recharge และ Artificial recharge | | | |
| <p>Conceptual Relation:</p>  <p style="text-align: center;">N05: Recharge</p> <p>G = Generic - Specific PT = Process - Target C = Cause - Effect</p> <p>คำอธิบาย : Recharge คือการเติมน้ำลงสู่ Aquifer เพื่อจุดประสงค์ในการ Increased Groundwater Level นอก จากนั้น Recharge ยังมีสองประเภทคือ Natural Recharge กับ Artificial recharge</p> | | | |
| <p>Extraction: 1. (n.) Subsurface percolation of water, from basin-margin mountains, directly into aquifers in the valley basins is appreciable and should be considered as <u>recharge</u>. (CEB03.txt)</p> <p>2. (n.) <u>Recharge</u> can significantly increase the sustainable yield of an aquifer. (CEP04.txt)</p> <p>3. (v.) In the USA all water <u>to be recharged</u> must comply with EPA drinking water standards or similar criteria, and detailed suites of analyses are required. (CES02.txt)</p> <p>4. (v.) Recharge basins are frequently used <u>to recharge</u> unconfined aquifers, especially where land costs are high. (CEB01.txt)</p> | | | |
| Grammatical Category : noun, verb | | | |

ปัญหาและทางแก้ไขในขั้นตอนการสร้างมโนทัศน์สัมพันธ์

ปัญหาและวิธีแก้ไขแบ่งตามขั้นตอนทั้งสามได้ดังนี้

ปัญหาในขั้นตอนการคัดเลือกชุดศัพท์

- การใช้โปรแกรม Win Concordance ในการ list ศัพท์ โดยทั่วไปแล้ว รูปแบบของศัพท์ในประมวลศัพท์มีอยู่สามรูปแบบด้วยกัน ได้แก่ Single-Word Terms, Multiword Terms และ Collocations แต่ในการให้โปรแกรมแสดงชุดศัพท์ทั้งหมดในคลังข้อมูลภาษา สามารถทำได้เฉพาะศัพท์เดี่ยวเท่านั้น (Single-word Terms) แต่ไม่สามารถแสดงศัพท์ที่มีความซับซ้อนกว่าได้ (Multiword Terms และ Collocation) แม้ว่าศัพท์เดี่ยวจะเป็นองค์ประกอบของศัพท์ที่มีความซับซ้อน เช่น Water ที่อาจประกอบเป็นคำอื่นๆ อย่าง Reclaimed Water, Surface Water, Water Recharge ซึ่งก็อาจใช้การค้นคำว่า Water ออกมาให้หมดแล้วหาไปที่ละศัพท์ แต่ก็ เป็นวิธีการที่เสียเวลามาก ซึ่งนับเป็นจุดอ่อนของโปรแกรม Win Concordance วิธีการแก้ปัญหาคือใช้โปรแกรม N-gram Frequency List* ช่วยในการดึงศัพท์ที่ประกอบด้วยคำมากกว่าหนึ่งได้ ช่วยในการประหยัดเวลาในกระบวนการคัดเลือกชุดศัพท์ไปได้มาก

- ความรู้พื้นฐานของผู้ทำวิจัยไม่เพียงพอในการแยกแยะศัพท์เฉพาะด้าน ปัญหานี้เกิดขึ้นภายหลังขั้นตอนที่โปรแกรมคอมพิวเตอร์ดึงศัพท์ออกมาจากคลังข้อมูลภาษาแล้ว ผู้ทำวิจัยต้องคัดเลือกขั้นต้นว่าศัพท์ใดน่าจะเป็นศัพท์เฉพาะด้าน ซึ่งการอาศัยข้อมูลจากคลังข้อมูลภาษาอย่างเดียวบางครั้งไม่เพียงพอ วิธีแก้ไข คือหาความรู้เพิ่มเติมในศาสตร์เฉพาะด้าน และขอคำปรึกษาจากผู้เชี่ยวชาญศาสตร์เฉพาะด้าน

ปัญหาในขั้นตอนการสร้างมโนทัศน์สัมพันธ์

- การกำหนดรูปแบบความสัมพันธ์ของมโนทัศน์ ในส่วนของรูปแบบความสัมพันธ์ที่ต้องกำหนดเอง เนื่องจากไม่อยู่ในส่วนของรูปแบบมาตรฐาน มักจะเป็นไปอย่างไม่มีหลักเกณฑ์ทำให้รูปแบบความสัมพันธ์มีมากและบางครั้งซ้ำซ้อนกัน เช่น Process – Problem กับ Process – Obstacle ซึ่งจริงๆ แล้ว Problem กับ Obstacle ก็อยู่ในหมวดเดียวกันสามารถรวมไว้ด้วยกันได้ เป็นต้น วิธีแก้ปัญหาก็คือ เมื่อจะกำหนดรูปแบบความสัมพันธ์ใหม่ขึ้นมาในแต่ละสถานการณ์ ต้อง

* ดาวน์โหลดจาก <http://pioneer.chula.ac.th/~awirote/colloc/>

ตรวจสอบอย่างละเอียดกับรูปแบบที่มีอยู่เดิม ว่าความสัมพันธ์ที่เป็นปัญหานั้น สามารถปรับใช้กับความสัมพันธ์เดิมได้หรือไม่

- รายละเอียดศัพท์ในบันทึกมโนทัศน์สัมพันธ์ขาดความสมบูรณ์ บางครั้งข้อมูลศัพท์ที่ได้จากคลังข้อมูลภาษาไม่ได้ให้นิยามของศัพท์ไว้เลย มีเพียง Metalinguistic Context และ Testimonial Context ทำให้ไม่สามารถแสดงนิยามที่ชัดเจนของศัพท์ได้ แต่จริงๆ แล้วสิ่งที่บันทึกในบันทึกมโนทัศน์สัมพันธ์เป็นเพียงข้อมูลจริงที่ได้จากคลังข้อมูลภาษาเท่านั้นและจะใช้เป็นข้อมูลเบื้องต้นสำหรับการจัดทำประมวลศัพท์เท่านั้น จึงไม่จำเป็นว่าข้อมูลเหล่านี้จะปรากฏอยู่ในประมวลศัพท์ฉบับสมบูรณ์ ข้อมูลในส่วนที่ขาดหายไปนี้ ต้องค้นจากหนังสืออ้างอิงเพื่อมาเสริมข้อมูลศัพท์ให้มีความครบถ้วนสมบูรณ์

- ไม่พบศัพท์ภาษาไทยที่สื่อถึงมโนทัศน์ ปัญหานี้เกิดจากการที่ข้อมูลภาษาไทยในคลังข้อมูลภาษามีน้อย มโนทัศน์จำนวนหนึ่งจึงไม่มีศัพท์ภาษาไทยที่สื่อถึง ทำให้เกิดช่องโหว่ในการบันทึกศัพท์ภาษาไทยที่สื่อถึงบางมโนทัศน์ วิธีการแก้ปัญหาคือ การบัญญัติศัพท์ภาษาไทยเสริมเข้าไป ซึ่งเป็นขั้นตอนที่จะพูดถึงในบทที่ 6 การกำหนดศัพท์ภาษาไทยในประมวลศัพท์ ซึ่งในขั้นตอนนี้ให้บันทึกข้อมูลศัพท์ภาษาไทยเท่าที่มีการใช้จริงไปก่อน

ปัญหาในขั้นตอนการบันทึกข้อมูลศัพท์เบื้องต้น

- ค้นบริบทศัพท์ในคลังข้อมูลภาษาได้น้อย ส่วนใหญ่ปัญหานี้ไม่ได้เกิดจากปริมาณข้อมูลเกี่ยวกับศัพท์นั้นๆ ในคลังข้อมูลภาษามีน้อย แต่ปัญหานี้มักเกิดจากการเลือกใช้คำสำคัญ (Key words) ในการค้นข้อมูลด้วยโปรแกรม Win Concord เช่น เมื่อต้องการค้นบริบททั้งหมดที่มีการใช้คำว่า Recharge ที่เป็น verb การใช้คำสำคัญในการค้นข้อมูลว่า Recharge เฉยๆ อาจทำให้ได้ผลออกมาน้อยหรือได้ข้อมูลในฐานะที่เป็นนามมากกว่า วิธีแก้ปัญหาก็คือ ลองเปลี่ยนรูปแบบของคำสำคัญในการค้นดู อย่างเช่นกรณีตัวอย่างนี้ ถ้าลองเปลี่ยนมาใช้คำว่า Recharged ในการค้นจะพบข้อมูลบริบทมากกว่าเดิมถึง 10 เท่า เป็นต้น ดังนั้น บางครั้งการค้นข้อมูลจึงควรลองใช้คำสำคัญหลายๆ แบบ จะช่วยให้พบข้อมูลที่มากขึ้นกว่าเดิม