# วัตถุประสงค์และเป้าหมาย

### วัตถุประสงค์

- 1. รองรับการปรับปรุงข้อมูลสมุนไพรไทยจากหลายแหล่งข้อมูลค้วยวิธีการอัตโนมัติ เพื่อเพิ่มความ สะควกและรวคเร็วในการจัดการฐานข้อมูลสมุนไพรไทย
- 2. รองรับการปรับปรุงข้อมูลสมุนไพรไทยจากหลากหลายแหล่งข้อมูลอย่างสะควกและรวคเร็ว
- 3. ใช้ออนโทโลจีในการจัดลำดับชั้นโครงสร้างข้อมูลสมุนไพรไทย เช่น ชื่อสมุนไพร, ชื่อวงศ์, ลักษณะ ของสมุนไพร, ข้อมูลของสมุนไพรไทยในด้านการแพทย์ เป็นต้น
- 4. ใช้ออนโทโลจีในการรวบรวมข้อมูล ซึ่งส่วนใหญ่จะถูกจัดแบ่งออกเป็นหัวข้อต่างๆ
- 5. พัฒนา GUI เพื่อเพิ่มความสะควกแก่ผู้ใช้งาน
- 6. รองรับการสืบค้นข้อมูลสมุนไพรไทยได้อย่างต่อเนื่อง

#### เป้าหมาย

- 1. พัฒนาขึ้นเป็นเว็บไซต์เพื่อเผยแพร่ ต่อบุคคลทั่วไป และผู้ซึ่งต้องการควบรู้ด้านสมุนไพรไทย
- 2. เพื่อเป็นการรวบรวมข้อมูล และความรู้ที่เกี่ยวกับสมุนไพรไทยให้อยู่ในรูปแบบเดียวกัน
- 3. เพื่อพัฒนาระบบสืบค้น และจัดหมวดหมู่สมุนไพรไทยโดยใช้วิธีการออนโทโลจี ซึ่งมีลักษณะ ดังต่อไปนี้
  - 3.1 รองรับการปรับปรุงข้อมูลสมุนไพรไทยจากหลากหลายแหล่งข้อมูลอย่างสะควกและ รวคเร็ว
  - 3.2 รองรับการนำเสนอข้อมูลสมุนไพรไทยโดยใช้วิธีการออนโทโลจี
  - 3.3 รองรับการสืบค้นข้อมูลสมุนไพรไทยได้อย่างต่อเนื่อง

# 1 ทฤษฎีและความรู้ที่เกี่ยวข้อง

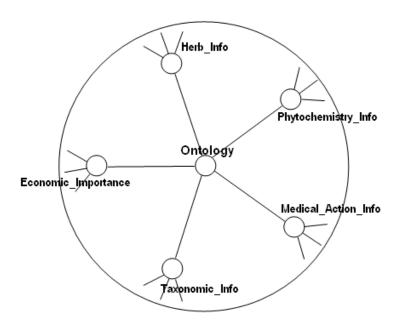
### 2.1 องค์ความรู้ (Ontology)

องค์ความรู้ (Ontology) คือ การอธิบายความสัมพันธ์โครงสร้างความรู้ให้อยู่ในรูปแบบลำคับชั้นเชิง วัตถุ เมื่อมีการเปลี่ยนแปลงฐานข้อมูล แล้วองค์ความรู้จะ ไม่มีการเปลี่ยนแปลง การเพิ่มหรือเปลี่ยนแปลง ฐานข้อมูลจะมีผลต่อส่วนของการเชื่อมต่อกับฐานข้อมูล เช่นเคียวกับโครงสร้างข้อมูล องค์ความรู้มีการ เปลี่ยนแปลง เมื่อมีการเปลี่ยนแปลงโครงสร้างความรู้ ตัวอย่างงานที่ใช้องค์ความรู้ เช่น BACIIS

ในระบบรวบรวมข้อมูล (Data Integration System) องค์ความรู้จะแสดงกลุ่มของคอนเซปต์ที่อยู่ใน โคเมน ทำให้การร้องขอข้อมูลถูกกำหนดด้วยคำที่เป็นมาตรฐาน สามารถอ้างถึงได้ในทุกๆ แหล่งข้อมูลที่ เกี่ยวข้องกัน การทำงานขององค์ความรู้ในระบบรวบรวมข้อมูลมีดังนี้

- ลดความซ้ำซ้อนในการเข้าถึงและการดึงข้อมูลจากแหล่งข้อมูลต่างๆ
- อธิบายข้อมูลในแต่ละแหล่งข้อมูลตามรูปแบบของคำที่กำหนดในองค์ความรู้
- กำหนดโครงสร้างของข้อมูลในแหล่งข้อมูลตามคอนเซปต์ขององค์ความรู้
- สร้างคอนเซปต์ใหม่ๆ ได้โดยมีพื้นฐานจากคอนเซปต์ที่ได้กำหนดไว้แล้ว
- จัดเตรียมกลไกที่สามารถเลือกคอนเซปต์ในแต่ละองค์ความรู้ที่ต้องการร้องขอได้
- แสดงผลลัพธ์ของการร้องขอในรูปแบบมีโครงสร้าง

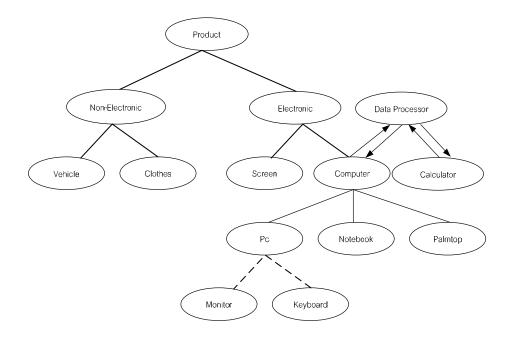
จากรายละเอียดของ Ontology ดังที่ได้กล่าวมาแล้วนั้น เมื่อนำมาประยุกต์เข้ากับข้อมูลสมุนไพรไทย จะสามารถแบ่งหัวข้อออกได้ ดังตัวอย่างในรูปที่ 2.1



รูปที่ 2.1 ตัวอย่างองก์ความรู้ (Ontology) ทางสมุนไพรไทย

### 2.1.2 การประยุกต์ใช้ Ontology

ออนโทโลขี เป็นการจัดแบ่งหมวดหมู่ความรู้ในขอบเขตความสนใจหนึ่ง ๆ (Domain) ซึ่งออนโทโล ยีนั้น ถูกอ้างถึงว่าเป็นการจัดความรู้แนวความคิดหรือคำต่าง ๆ ให้เป็นระเบียบ (Formalization) อยู่ใน ขอบเขตโดเมนที่สนใจฐานความรู้ที่มีจึงสามารถนำมาแลกเปลี่ยน หรือนำมาใช้ซ้ำ (Reuse) และเผยแพร่ ให้แก่ผู้อื่นได้นำไปใช้งานต่อ ๆ ไปอย่างกว้างขวางได้โดย ออนโทโลขีนั้นประกอบด้วยออนโทโลขีคอน เซพ ที่แสดงกลุ่มคำ และความสัมพันธ์เชิงความหมายระหว่างคำในโดเมน ความสัมพันธ์ของคำในโดเมนที่ มีหลากหลายรูปแบบ เช่น is-a , part-of ซึ่งความสัมพันธ์เหล่านี้จะเชื่อมโยงกลุ่มคำในโดเมนเป็นลำดับชั้น รูปแบบโครงสร้างอออนโทโลขีจึงมีลักษณะเป็นลำดับชั้น เช่น โครงสร้างต้นไม้ (Tree), กราฟ (Graph) ตัวอย่างการจัดหมวดหมู่ความรู้ในโดเมนที่สนใจเป็นลำดับชั้น เช่นการจำแนกพืช ชนิดต่าง ๆ ออกเป็น Genus และ Species เป็นต้น แสดงตัวอย่างออนโทโลขีกลุ่มสินค้าได้ดังรูปที่ 3.2 ซึ่งอธิบายการจัดกลุ่มสินค้า ที่ประกอบไปด้วยความสัมพันธ์เชิงความหมาย คือ is-a เช่น PC is-a Computer , part-of เช่น Keyboard part-of PC และ syn-of เช่น Data Processor syn-of Computer เป็นดัน



รูปที่ 2.2 ตัวอย่างออนโทโลจีกลุ่มสินค้า

ส่วนต่อไปจะอธิบายถึงความรู้พื้นฐานที่เกี่ยวข้องกับออนโทโลยี ได้แก่ องค์ประกอบของออน โทโลยี และภาษาที่ใช้ในการอธิบายออนโทโลยี รวมถึงการนำออนโทโลยีไปประยุกต์ใช้

### องค์ประกอบของออนโทโลยี (Components of ontology)

ออนโทโลยีประกอบไปด้วยคอนเซพ ซึ่งเป็นพื้นฐานสำคัญในการสร้างฐานความรู้ โดยคอนเซพ เหล่านี้จัดเรียงอยู่ในลำดับชั้นการถ่ายทอดความสัมพันธ์ และมีคุณสมบัติเฉพาะในแต่ละคอนเซพ โดยสรุป แล้วองค์ประ กอบของออนโทโลยีประกอบด้วยส่วนต่าง ๆ ดังต่อไปนี้

- แนวคิด/คอนเซพ (Concept) : ความคิดทั่วไปหรือนามธรรมในโคเมนที่เราสนใจ เช่น Computer, PC, Notebook, Keyboard เป็นต้น
- คุณลักษณะ (Property) : คุณสมบัติของคอนเซพ เช่น สี น้ำหนัก เป็นต้น
- ความสัมพันธ์ (Relationship) : ความสัมพันธ์เชิงความหมายระหว่างคอนเซพออนโทโลยีโดย ส่วนใหญ่จะประกอบไปด้วยความสัมพันธ์ ได้แก่

- is-a คือ ความสัมพันธ์ที่มีคุณสมบัติการถ่ายทอด คุณสมบัติของคอนเซพแม่
- ไปยังคอนเซพลูก เช่น PC is-a Computer ซึ่งอธิบายได้ว่า PC มีคุณสมบัติเป็น
   Computer
- part-of คือ ความสัมพันธ์ที่หมายถึงการเป็นส่วนประกอบ เช่น Keyboard
- part-of Computer ซึ่งอธิบายได้ว่า Computer จะต้องประกอบไปด้วย Keyboard
- syn-of คือ ความสัมพันธ์ที่แสดงถึงคอนเซพที่มีความเหมือนเชิงความหมายต่อกัน เช่น
  Data Processor syn-of Computer ซึ่งอธิบายได้ว่า Computer มีความหมายเคียวกันกับ
  Data Processor สามารถใช้แทนกันได้
- instance-of คือ ความสัมพันธ์ที่แสดงถึงการเป็นตัวแทน หรือสมาชิกของ คอน เซพ เช่น HP instance-of Computer ซึ่งอธิบายได้ว่า HP เป็นคอมพิวเตอร์ประเภทหนึ่ง นอกจากนี้ออนโทโลยียังประกอบไปด้วยความสัมพันธ์เชิงความหมายอื่น ๆ ที่สอดคล้องกับโดเมน ซึ่งกำหนดโดยผู้เชี่ยวชาญ
  - ข้อความอธิบายคอนเซพ (Axiom) เงื่อนไขหรือ ข้อกำหนดเฉพาะในแต่ละคอนเซพเป็นกลไก สำคัญสำหรับใช้ในการอนุมานความรู้ เพื่อการสร้างความรู้ใหม่จากออนโทโลยี ตัวอย่างเช่น คอมพิวเตอร์จะต้องประกอบด้วยซีพียู เป็นต้น

### ภาษาที่ใช้อธิบายออนโทโลยี (Ontology languages)

ภาษาที่ใช้พัฒนาออนโทโลยี แบ่งตามรูปแบบภาษาได้คังต่อไปนี้ ลอจิคเบส (First order logic based), เฟรมเบส (Frame logic based) และเว็บเบส (Web based)

• ลอจิคเบส รูปแบบภาษาที่อธิบายออนโทโลยี โดยการใช้ตรรกะเพื่อการอนุมานความรู้ โดยมีส่วนประกอบ คือ ภาคแสดง (Predicate) อาร์กิวเมนต์ (Argument) และตัวบ่งปริมาณ (Quantifier) ตัวอย่าง เช่น  $\forall x$ :  $PC(X) \longrightarrow haspart(keyboard(x)) หมายถึง PC ทุกเครื่องจะ ประกอบด้วย keyboard$ 

กุณสมบัติของภาษาออนโทโลยีที่พัฒนาแบบลอจิกเบส คือ วากยสัมพันธ์ (Syntax) มีความชัดเจน และมีรูปแบบเป็นทางการ จึงสามารถสร้างกฎอนุมานใค้ (Reference rule) ตัวอย่างภาษา เช่น CYCL, CLASSIC, LOOM เป็นต้น

• เฟรมเบส รูปแบบภาษาที่อธิบายออนโทโลยี โดยใช้หลักการของการคิดเชิงปฏิสัมพันธ์ของ มนุษย์ เมื่อนึกถึงวัตถุ หรือสิ่งใดจะเชื่อมโยงกับคุณลักษณะเด่นของวัตถุหรือสิ่งนั้นด้วย ดังนั้น เฟรมจึงประกอบไปด้วย เซตของกุณสมบัติ (Attribute) หรือ สล๊อท(slot) และแฟ็กเซ็ต (Facets) ข้อความอธิบายคอนเซพหรือคำอธิบายสล๊อท คุณสมบัติของภาษาออน โทโลยีที่พัฒนาแบบ เฟรมเบส คือ เข้าใจได้ง่าย ตัวอย่างภาษาที่พัฒนาแบบเฟรมเบส เช่น OKBC, F-logic แสดง ตัวอย่างการใช้เฟรมแทนออน โทโลยีคอนเซพ "PC" ได้ดังนี้

จากตัวอย่างเฟรมข้างต้นเป็นเฟรมที่อธิบายองค์ประกอบการแทนคอนเซพ "PC" ซึ่งประกอบด้วย ค่าคุณสมบัติ 3 ค่า ได้แก่ ชื่อรุ่น (Model), สี (Color) และราคา (Price) และการแทนค่าคุณสมบัติของ ตัวอย่างสมาชิกของคอนเซพ (Instance) คือ "HP"

• เว็บเบส รูปแบบภาษาที่อธิบายออนโทโลยีซึ่งพัฒนาจากภาษาที่ใช้สำหรับอธิบายทรัพยกรบน เว็บ ได้แก่ XML (Extensible Markup Language), RDF (Resource Description Framework) ซึ่งอธิบายออนโทโลยีโดยใช้พื้นฐานของลอจิกเบส และเฟรมเบส ภาษาที่พัฒนาขึ้นจึงมี หลักการและแบบแผน สามารถอ้างอิงได้ และอยู่ในรูปแบบที่มนุษย์สามารถเข้าใจได้ง่าย ตัวอย่างภาษาเช่น DAML + OIL (DARPA Agent Markup Language + Ontology Interface Language), OWL (Web Ontology Language) เป็นต้น

### การประยุกต์ใช้ออนโทโลยี (Ontology application)

ออนโทโลยีถูกนำไปประยุกต์ใช้ในหลาย ๆ งาน สามารถแบ่งกลุ่มโปรแกรมที่ประยุกต์ใช้ออน โทโลยี ได้ดังนี้

- การนำออนโทโลยีไปใช้เพื่อแปลงข้อมูลให้อยู่ในรูปแบบภาษาต่าง ๆ (Neutral authoring) นำ
  ออนโทโลยีไปใช้เพื่อแปลงข้อมูลให้อยู่ในรูปแบบภาษาต่าง ๆเพื่อให้โปรแกรมอื่น ๆ สามารถ
  ใช้งานได้ประโยชน์ที่ได้จาการประยุกต์ใช้ออนโทโลยี คือ การนำความรู้มาใช้ได้อีก
  (Knowledge reuse)
- การนำออนโทโลยีมาใช้เพื่อกำหนดรายละเอียดของซอฟต์แวร์ (Ontology as specification)
   ประยุกต์ใช้ออนโทโลยีเพื่อออกแบบซอฟต์แวร์ในโดเมน และรวบรวมคำศัพท์สำหรับกำหนด
   ความต้องการในการพัฒนาซอฟต์แวร์ ประโยชน์ที่ได้จาการประยุกต์ใช้ออนโทโลยี คือ การทำ
   คู่มือโปรแกรม การบำรุงรักษาซอฟต์แวร์ และ การนำกลับมาใช้ใหม่
- การประยุกต์ใช้ออนโทโลยี เพื่อการเข้าถึงข้อมูลที่มีโครงสร้าง หรือรูปแบบต่างกัน (Common Access to Information) ออนโทโลยีจัดเตรียมคำที่สามารถเข้าใจได้ตรงกัน หรือจัดกลุ่มคำที่มี ความหมายเดียวกัน ประโยชน์ที่ได้ คือ การทำงานร่วมกัน (Inter-operability) และการนำ กลับมาใช้ใหม่
- การประยุกต์ใช้ออนโทโลยีเพื่อการสืบค้นข้อมูลจากแหล่งข้อมูลต่าง ๆ (Ontology-based search) ประยุกต์ใช้ออนโทโลยีเพื่อการสืบค้นข้อมูลจากแหล่งข้อมูลต่าง ๆ เช่น เอกสาร เว็บ เพจ หรือฐานข้อมูล แนวทางนี้ประยุกต์ใช้ออนโทโลยีในการกำหนดคอนเซพที่สอดคล้องกับ คำสืบค้นของผู้ใช้ และใช้คอนเซพนั้นในการสืบค้นข้อมูล ทำให้ผลการสืบค้นมีความถูกต้อง มากยิ่งขึ้น และเวลาที่ใช้ในการสืบค้นลดลง ซึ่งจะอธิบายรายละเอียดเพิ่มเติมในหัวข้อถัดไป

### การสืบค้นข้อมูลโดยประยุกต์ใช้ออนโทโลยี

ในปัจจุบันออนโทโลยี มีความสำคัญต่อการพัฒนาแอพพลิเคชันในด้านต่างๆ ให้มีประสิทธิภาพ มากขึ้น โดยเฉพาะการพัฒนาแอพพลิเคชันสำหรับสืบค้นข้อมูลที่ประยุกต์ใช้ออนโทโลยี เพื่อสืบค้นข้อมูล จากแหล่งข้อมูลต่าง ๆ เช่น เอกสาร เว็บเพจ และการสืบค้นข้อมูลในโคเมนที่เกี่ยวข้องกับวิทยาศาสตร์เพื่อ ชีวิต เช่น Gene Ontology Project [GO] (1999), Plant Ontology Consortium [POC] (2004)

การประยุกต์ใช้ออนโทโลยีในการสืบค้นข้อมูล สามารถอธิบายได้ดังนี้ การประยุกต์ใช้ออนโทโลยี เป็นโครงสร้างพื้นฐานในการจัดการข้อมูล เช่น ประยุกต์ใช้ออนโทโลยีเป็นโครงสร้างของคอนเซพ เพื่อให้ ผู้ใช้ทราบถึงขอบเขตของแหล่งข้อมูล ตัวอย่างการใช้เช่น Yahoo taxonomy ซึ่งจะแบ่งกลุ่มข้อมูลเป็น 14 กลุ่มใหญ่ที่ประกอบด้วยกลุ่มย่อยที่สอดคล้องกัน, การประยุกต์ใช้ออนโทโลยีเป็นคำศัพท์สำหรับกำหนด เมตาดาต้า (Metadata language) สำหรับสร้างดัชนี (Index) หรือสร้างแท็ก (Tag) ให้กับข้อมูล อีกแนวทาง หนึ่ง คือ การประยุกต์ใช้ออนโทโลยี เพื่อแปลงคำสั่งสืบค้น (Query transformation) เช่น การประยุกต์ใช้ ออนโทโลยีเพื่อถ้างอิงถึงคอนเซพที่สอดคล้องกับคำสืบค้น

ตัวอย่างระบบการสืบค้นข้อมูลโดยใช้ออนโทโลยี ได้แก่ Knowledge-Based Discovery Tool เมื่อ ผู้ใช้กำหนดคำสืบค้นที่ต้องการ ระบบจะพิจารณาคำสืบค้นกับออนโทโลยี WordNet เพื่อพิจารณาหาคอน เซพที่สอดคล้องกับคำสืบค้น และให้ผู้ใช้กำหนดคอนเซพที่ถูกต้อง หลังจากนั้นคอนเซพที่มีความหมายไม่ ตรงกับคำสืบค้นก็จะถูกคัดออกไป เพื่อลดความกำกวมของผลการสืบค้น

ในหัวข้อถัดไปจะอธิบายถึงการสืบค้นข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี ซึ่งเป็น แนวทางที่ประยุกต์ใช้ออนโทโลยี เพื่อเพิ่มประสิทธิภาพในการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ ซึ่ง สืบค้นโดยพิจารณาคำสืบค้นโดยตรงให้สามารถสืบค้นโดยพิจารณาคำสืบค้นเชิงความหมายได้ (Semantic search)

## การสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี

การสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์โดยประยุกต์ใช้ออนโทโลยี เป็นแนวทางเพิ่ม
ประสิทธิภาพการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ให้สามารถสืบค้นข้อมูลเชิงความหมายและข้อมูลที่
มีลักษณะเป็นลำดับชั้นใค้ (Concept Hierarchy) เช่น การสืบค้นข้อฐานสมุนไพรในระดับตระกูล (Order)
และวงศ์ (Family) โดยประยุกต์ใช้ออนโทโลยีเพื่อแปลงคำสั่งสืบค้นของผู้ใช้ให้อยู่ในรูปแบบคำสั่งสืบค้น
เชิงความหมาย เช่น การพิจารณาคอนเซพในออนโทโลยีที่เฉพาะเจาะจง (Specialization) คือ การหาโหนดที่
เป็นโหนดลูก (Child node) ของโหนดที่เราสนใจ, การพิจารณาคอนเซพในออนโทโลยีในระดับทั่วไป
(Generalization) คือ การหาโหนดบรรพบุรุษ (Parent Node) ของโหนดที่เราพิจารณา และการพิจารณาคอน
เซพข้างเคียงในออนโทโลยี (Neighborhood) ทำให้ผลการสืบค้นที่ได้มีความครบถ้วนมากยิ่งขึ้น เมื่อเทียบ

กับการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ทั่ว ๆ ไปที่สืบค้นโดยใช้คำสืบค้นจากผู้ใช้เพียงเท่านั้นอีกทั้งยัง ลดเวลาที่ใช้ในการสืบค้นอีกด้วย

พิจารณาตัวอย่างการสืบค้นข้อมูลจากตารางที่ 2.2 ถ้าผู้ใช้ต้องการสืบค้นคำว่า "Computer"

Select \* from Item where Name='Computer'

ผลการสืบค้นที่ได้จะมีเพียงเรคคอรค์เคียวเท่านั้น คือ RecID ที่ 123 แต่เมื่อพิจารณาออนโทโลยีกลุ่ม สินค้าคังรูปที่ 2.2 จะเห็นได้ว่า Computer มีความหมายเหมือนกันกับ "Calculator" และ "Data Processor" ซึ่งอธิบายค้วย ความสัมพันธ์แบบ syn-of ในออนโทโลยี นอกจากนี้ เมื่อพิจารณา ความสัมพันธ์แบบ is-a จะพบว่า "PC", "Notebook" และ "Palmtop" สอคคล้องกับคำสืบค้น คังนั้นผลการสืบค้น คือ RecID ที่ 123, 125,127, 141

ตาราง 2.1 ตาราง Item

RecID	Name	Model	Price
123	Computer	IBM	3000\$
124	IntelPC	TOSHIBA	5000\$
125	Notebook	DELL	4000\$
127	PC	COMPAQ	2500\$
128	Product	НР	3000\$

RecID	Name	Model	Price
129	Monitor	ELSA	1000\$
135	Keyboard	ITT	80\$
136	Desktop	IBM	1000\$
140	MacPC	MAC	2000\$
141	Calculator	SIEMEN	1500\$

การพัฒนาออนโทโลยีเพื่อเพิ่มประสิทธิภาพการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ ต้อง พิจารณาถึงแนวทางการจัดเก็บและสืบค้นออนโทโลยี รวมถึงวิธีการสืบค้นข้อมูลในฐานข้อมูลเชิงสัมพันธ์ที่ สอคคล้องกับออนโทโลยี

#### 2.2 Search Engine

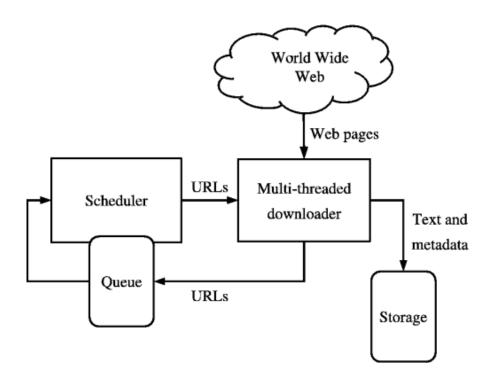
Search Engine คือ เครื่องมือการค้นหาข้อมูลผ่านอินเตอร์เน็ต ที่ทุกคนสามารถเข้าไปค้นหาข้อมูล ในอินเตอร์เน็ต ได้อย่างง่ายดาย โดยกรอกข้อมูลที่ต้องการค้นหา หรือ Keyword เข้าไปที่ช่อง Search Box เท่านั้น โดยข้อมูลที่เราต้องการค้นหาจะถูกแสดงออกมาเป็นจำนวนมากเพื่อให้เราเลือกข้อมูลที่เราสนใจมาก ที่สุดเอามาใช้งาน ลักษณะของการแสดงผลของ Search Engine นั้นจะทำการแสดงผลแบบเรียงอันดับ Search Result ผ่านหน้าจอคอมพิวเตอร์ โดย Search Engine สามารถแบ่งการหลักการทำงานที่แตกต่างกัน และการจัดอันดับข้อมูลออกได้เป็น 3 ประเภท คือ

#### 1) Crawler Based Search Engines

เป็นเครื่องมือการค้นหาบนอินเตอร์เน็ตแบบอาศัยการบันทึกข้อมูล และจัดเก็บข้อมูลเป็นหลัก ซึ่งจะ เป็นจำพวก Search Engine ที่ได้รับความนิยมสูงสุด เนื่องจากให้ผลการค้นหาที่แม่นยำที่สุด และการ ประมวลผลการค้นหาสามารถทำได้อย่างรวดเร็ว จึงทำให้มีบทบาทในการค้นหาข้อมูลมากที่สุดในปัจจุบัน มีการทำงานดังในรูปที่ 2.3

### โดยมืองค์ประกอบหลักอยู่ 2 ส่วน คือ

- ฐานข้อมูล โดยส่วนใหญ่แล้ว Crawler Based Search Engine จะมีฐานข้อมูลเป็นของตัวเอง ที่ มีระบบการประมวลผล และจัดอันดับที่เฉพาะเป็นเอกลักษณ์ของตนเอง
- Software คือ เครื่องมือหลักสำคัญที่สุดอีกส่วนหนึ่งสำหรับ Search Engine ประเภทนี้ เนื่องจากต้องอาศัยโปรแกรมเล็กๆ ทำหน้าที่ในการตรวจหา และทำการจัดเก็บข้อมูล หน้าเพจ หรือเว็บไซค์ต่างๆ ในรูปแบบของการทำสำเนาข้อมูล เหมือนกับต้นฉบับทุกอย่าง ซึ่งเราจะ รู้จักกันในนาม Spider หรือ Web Crawler หรือ Search Engine Robots



รูปที่ 2.3 ลักษณะการทำงานของ Web Crawler

การทำ Web Crawler แบบเฉพาะเจาะจงนั้น มักจะไม่ได้เน้นในส่วนของการทำ Multi-Thread เพื่อ แบ่งย่อยหลายๆ งาน แล้วแบ่งๆ แยกย้ายกันไปอ่านตามแต่ละลิงค์ แต่อยู่ตรงที่ตัว Web Crawler เองนั้น ต้อง รู้จักการ Login เข้าไปสู่ระบบคังกล่าว แล้วจึงเข้าไปยังหน้าเว็บไซค์ลึกๆ เพื่ออ่านข้อมูล จากนั้นจึงเจาะจง เลือกเฉพาะบางลิงค์ บางตำแหน่งเท่านั้นเพื่อกระโคคต่อไป

นอกจากนี้การออกแบบโครงสร้างข้อมูลใน Virtual Memory ให้สอดรับกับ Tag HTML ทั้งปวง ของเว็บไซค์นั้นๆ อีกทั้งต้องสามารถกำหนดจุดเริ่มต้น หรือสิ้นสุดในแต่ละบล็อคที่ต้องการอ่านลิงค์อีกด้วย และหากการเขียน Software ดังกล่าวไม่เขียนเป็นระบบเปิด ก็จะใช้ได้เฉพาะกับเว็บไซค์ใดเว็บไซค์หนึ่ง เท่านั้น ดังนั้นการทำ Web Crawler แบบเฉพาะเจาะจง จึงหวังที่ผลลัพธ์ที่ได้จากเว็บไซค์นั้นเป็นสำคัญ

ตัวอย่างเว็บไซค์ที่เป็นระบบปิค เมื่อต้องการค้นหาข้อมูล จำเป็นจะต้องล็อกอินเข้าสู่ระบบก่อนทุก ครั้ง คังรูปที่ 2.4



รูปที่ 2.4 ตัวอย่างเว็ปใชด์ที่ต้อง Login ก่อนเข้าไปอ่านข้อมูลได้

ตัวอย่างเว็บที่ใช้เทคนิคนี้ เช่น Google, Yahoo, Live, Search, Technorati เป็นต้น ซึ่งลักษณะการ ทำงานและการเก็บข้อมูลของ Web Crawler แต่ละแห่งจะมีวิธีการเก็บและการจัดอันดับข้อมูลที่แตกต่างกัน ดังรูปที่ 2.5



Advertising Programs - Business Solutions - About Google - Go to Google Thailand

Make Google Your Homepage!

@2006 Google

# รูปที่ 2.5 ตัวอย่างเว็ปไซด์ที่ทำงานแบบ Web Crawler

### 2) Web Directory หรือ Blog Directory

เป็นสารบัญเว็บไซด์ที่สามารถค้นหาข่าวสารข้อมูล ด้วยหมวดหมู่ข่าวสารข้อมูลที่เกี่ยวข้องกัน ใน ปริมาณมากๆ คล้ายๆ กับสมุดหน้าเหลือง ซึ่งจะมีการสร้างครรชนี มีการระบุหมวดหมู่อย่างชัดเจน ซึ่งจะ ช่วยให้การค้นหาข้อมูลต่างๆ ตามหมวดหมู่นั้นๆ ได้รับการเปรียบเทียบอ้างอิง เพื่อหาข้อเท็จจริงได้ ในขณะ ที่เราค้นหาข้อมูล เพราะจะมีเว็บไซด์มากมาย หรือ Blog มากมายที่มีเนื้อหาคล้ายๆ กัน ในหมวดหมู่เดียวกัน ให้เราเลือกที่จะหาข้อมูลได้อย่างตรงประเด็นที่สุด ซึ่งจะลดระยะเวลาในการค้นหาได้มาก ดังรูปที่ 2.6



over 4 million sites - 74,455 editors - over 590,000 categories

## รูปที่ 2.6 ตัวอย่าง Web Directory ที่มี Search Engine มากมายใช้เป็นฐานข้อมูล

- OPD หรือ Dmoz ที่เป็นที่นิยมเป็น Web Directory ที่ใหญ่ที่สุดในโลก Search Engine หลายๆ แห่ง มักใช้ข้อมูลจากแห่งนี้เกือบทั้งสิ้น เช่น Google, AOL, Yahoo, Netscape และอื่นๆ อีก มากมาย ODP มีการบันทึกข้อมูลประมาณ 80 ภาษาทั่วโลก (URL: <a href="http://www.dmoz.org">http://www.dmoz.org</a>)
- สารบัญเว็ปไทย Sanook ก็เป็น Web Directory ที่มีชื่อเสียงอีกแห่งหนึ่ง และเป็นที่รู้จักกันมาก ในเมืองไทย (URL : <a href="http://webindex.sanook.com">http://webindex.sanook.com</a>)
- Blog Directory อย่าง BlogFlux Directory ที่มีการเก็บข้อมูลเกี่ยวกับบลีอกมากมายตาม หมวดหมู่ต่างๆ

#### 3) Meta Search Engine

เป็น Search Engine ที่ใช้หลักการในการค้นหาโดยอาศัย Meta Tag ในภาษา HTML ซึ่งจะมีการ ประกาศชุดคำสั่งต่างๆ เป็นรูปแบบของ Text Editor ด้วยภาษา HTML คือ ชื่อผู้พัฒนา คำค้นหา เจ้าของเว็บ หรือบล็อก คำอธิบายเว็บ หรือบล็อกอย่างย่อ

ผลการค้นหาของ Meta Search Engine นี้มักไม่แม่นยำอย่างที่กิด เนื่องจากบางครั้งผู้ให้บริการหรือ ผู้ออกแบบเว็บสามารถใส่อะไรเข้าไปก็ได้มากมายเพื่อให้เกิดการค้นหาและพบเว็บ หรือ บล็อกของตนเอง และ อีกประการหนึ่งก็คือ มีการอาศัย Search Engine Index Server หลายๆ แห่งมาประมวลผลรวมกัน จึงทำ ให้ผลการค้นหาข้อมูลต่าง ๆ ไม่เที่ยงตรงเท่าที่ควร

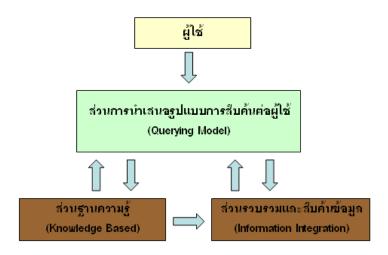
ตัวอย่างคำสั่งของ Tag META จะมีการประกาศรายละเอียดของเว็บไซค์นั้นๆ ดังรูปที่ 2.7

```
<HEAD>
<TITLE>Stamp Collecting World</TITLE>
<META name="description" content="Everything you wanted to know about stamps, from prices to history.">
<META name="keywords" content="stamps, stamp collecting, stamp history, prices, stamps for sale">
</HEAD>
```

รูปที่ 2.7 ตัวอย่างคำสั่งใน Tag META เพื่อประกาศข้อมูลของเว็บ

### 2.3 ระบบสืบค้นข้อมูล

จากการศึกษาการสืบค้นข้อมูลผ่านทางระบบเครือข่าย พบว่า ข้อมูลมีการกระจัดกระจายกันอยู่ใน ฐานข้อมูลต่างๆ ซึ่งยุ่งยากและเสียเวลาในการสืบค้น ดังนั้นจึงได้จัดการแบ่งระบบสืบค้นออกเป็น 3 ส่วน ได้แก่ ฐานความรู้ (Knowledge Based) การรวบรวมและสืบค้นข้อมูล (Information Integration) และการ นำเสนอรูปแบบการสืบค้นต่อผู้ใช้ (Querying Model)



รูปที่ 2.8 สถาปัตยกรรมทั่วไปในการสืบค้นข้อมูล

จากรูปที่ 2.8 จะแสดงภาพรวมของระบบ เริ่มจากผู้ใช้จะทำการสร้างคำถาม และสิ่งที่ต้องการ คำตอบผ่านทางรูปแบบการสืบค้นต่อผู้ใช้ หลังจากนั้นระบบจะติดต่อระหว่างส่วนการรวบรวมและสืบค้น ข้อมูล และฐานข้อมูลเพื่อประมวลคำตอบ หลังจากนั้นคำตอบจะถูกส่งไปยังรูปแบบการสืบค้น เพื่อแสดงผล ของการสืบค้นต่อผู้ใช้

# 2.4 ความรู้พื้นฐานเกี่ยวกับสมุนไพร

### 2.4.1 การจำแนกสมุนไพร

การจำแนกข้อมูลทางสมุนไพรนี้ ได้แนวคิดมาจากนางสาวบุญรัตน์ เผดิมรอด ซึ่งได้ให้คำแนะนำใน การทำโครงงานระบบสำหรับรวบรวมและสืบค้นข้อมูลสมุนไพร สามารถจำแนกได้ ดังนี้

## <u>ข้อมูลทั่วไป</u>

- ชื่อสามัญ (Common Name)
- ชื่ออังกฤษ
- ชื่อวิทยาศาสตร์ (Science Name)
- ชื่อวงศ์ (Family Name)
- ชื่ออื่นๆ
- ชื่อฟ้อง (Synonym Name)

### ชีววิทยา - นิเวศวิทยา

- ลักษณะพืช
  - ประเภทพืช เช่น ไม้ล้มลุก ไม้ยืนต้น ไม้พุ่ม
  - ความสูง
  - ลักษณะเปลือกไม้
  - ลักษณะใบ ประกอบด้วย รูปใบ ความกว้าง/ยาว
  - ลักษณะดอก ประกอบด้วย สี จำนวนกลีบดอก รูปทรงช่อดอก
- ลักษณะทางภูมิศาสตร์
  - ลักษณะป่า
  - ลักษณะอากาศ
  - แสงแคค
  - ลักษณะดิน
  - ปริมาณน้ำฝน
- แหล่งที่พบ
  - จังหวัดที่พบ
  - ลักษณะทางภูมิศาสตร์ที่พบ
- ประโยชน์ทางยา
  - ส่วนที่ใช้

- รสทางยา
- สรรพคุณ (บอกชื่อโรคที่รักษา แยกตามส่วนที่ใช้ / บอกสรรพคุณ โดยรวม)
- ภูมิปัญญาหมอพื้นบ้าน
- การขยายพันธุ์ (การเพาะปลูก)
  - การขยายพันธุ์แบบไม่อาศัยเพศ
  - การขยายพันธุ์แบบอาศัยเพศ

### ตัวอย่างของสมุนไพร

สมุนไพรที่จะยกตัวอย่างนี้ คือ สะเคา ซึ่งเป็นสมุนไพรชนิคหนึ่งที่ได้รับความสนใจมากขึ้น และมี รายละเอียคตามแนวคิดการจำแนกสมุนไพรดังที่ได้กล่าวมาแล้ว คือ

# <u>ข้อมูลทั่วไป</u>

- ชื่อสามัญ : Neem Tree
- ชื่ออังกฤษ : Nim, Margosa
- ชื่อวิทยาศาสตร์ : Azadirachta indica A. Juss. var. siamesis Valeton
- ชื่อวงศ์ : Meliaceae
- ชื่ออื่นๆ: กะเดา(ภาคใต้), จะตัง(ส่วย), สะเดา(ภาคกลาง), สะเลียม(ภาคเหนือ), สะเดา บ้าน(ทั่วไป)
- ชื่อพ้อง : Azadirachta indica A. Juss. var. siamensis Backer,

Azadirachta indica A. Juss. var. Minor Veleton,

Azadirachta indica A. Juss. var. Minor Backer

### ชีววิทยา - นิเวศวิทยา

ข้อมูลทั่วไปเกี่ยวกับพืช (Botanical Characteristic)

ลักษณะของพืช: เป็นต้นไม้ขนาดกลาง สูง 12 – 20 เมตร ทุกส่วนมีรสขม เส้นรอบวงของลำต้น สะเคายาวประมาณ 30 – 60 ซม. เรือนยอดเป็นพุ่มกลมหรือคล้ายเจดีย์ต่ำ เปลือกของต้นสีน้ำตาลเทาหรือเทา ปนดำ แตกระแหงเป็นร่องเล็กๆ หรือเป็นสะเก็ดยาวตามต้น แต่เปลือกของกิ่งอ่อนเรียบ ลักษณะทางภูมิศาสตร์: ไม้สะเคาสามารถเจริญงอกงามในท้องถิ่นที่มีอากาศร้อนและชื้น ซึ่งมี อุณหภูมิสูงถึง 44 องศาเซลเซียส ที่ระดับความสูง 50-1500 เมตร จากระดับน้ำทะเล ปริมาณน้ำฝนเฉลี่ย ประมาณ 450-1150 มิลลิเมตรต่อปี สามารถขึ้นได้ในสภาพคินที่มีความแห้งแล้ง คินหิน คินเหนียวและตื้น แต่คินที่เหมาะสมต่อการเจริญเติบโตของไม้สะเคาควรมีความเป็นกรคค่อนข้างสูง คือ มี pH ประมาณ 5 ไม้ สะเคาก็สามารถปรับตัวขึ้นอยู่ได้ แต่พื้นที่ซึ่งถูกน้ำท่วมขังอยู่เสมอและพื้นคินเค็มจัดจะมีการเจริญเติบโตช้า หรือตายได้

#### ประ โยชน์ทางยา

- ใบอ่อน แก้โรคผิวหนัง น้ำเหลืองเสีย พุพอง
- ใบแก่ บำรุงธาตุ ช่วยย่อยอาหาร ฆ่าแมลงศัตรูพืช
- ก้าน แก้ไข บำรุงน้ำดี แก้ร้อนในกระหายน้ำ
- คอก แก้พิษ โลหิต พิษกำเดา แก้ริดสีดวงเป็นเม็ดยอด คันในลำคอ บำรุงชาตุ
- ลูก บำรุงหัวใจ ให้เต้นเป็นปกติ ฆ่าแมลงศัตรูพืช
- ลูกอ่อน แก้ลมหทัยวาตะ เจริญอาหาร ฆ่าพยาธิ แก้ริคสีควง แก้ปัสสาวะพิการ
- เปลือกต้น แก้ท้องเดิน แก้บิด มูกเลือด แก้ไข แก้กษัย แก้ในกองเสมหะ
- แก่น แก้คลื่นเหียนอาเจียน แก้ไขจับสั่น ไข้ตัวร้อน บำรุงโลหิต บำรุงไฟธาตุ
- ราก แก้เสมหะจุกคอ แก้เสมหะที่เกาะแน่นในทรวงอก
- ยางใช้ดับพิษร้อน

การขยายพันธุ์ : เป็นการขยายพันธุ์แบบไม่อาศัยเพศ คือ แตกหน่อ และการขยายพันธุ์แบบอาศัยเพศ คือ เพาะเมล็ด

# 3 อุปกรณ์และเครื่องมือที่ใช้

#### Hardware

- Notebook Pentium M Processor 1.73 GHz
- RAM 1.00 GB
- Hard disk 60 GBytes

#### Operation System

Microsoft Windows XP Professional

#### Database

• MySQL Version 2005

#### Tools

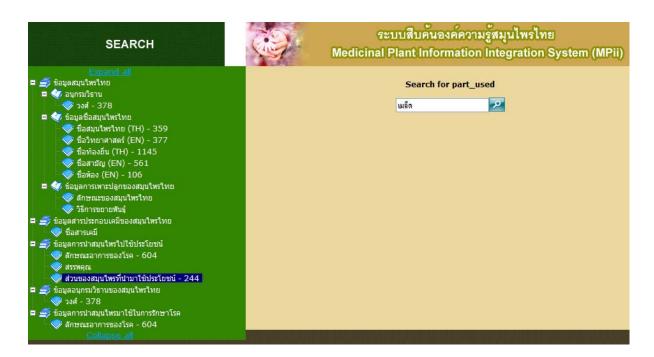
- EditPlus Version 3.01
- Apache Web Server Version 2.5
- PHP Version 5
- Microsoft Office Version 2007 (Microsoft Word, PowerPoint)

# 4 รายละเอียดโปรแกรม

โครงงานนี้จัดทำขึ้นในรูปแบบของเว็บเพจ โดยใช้ภาษา PHP เป็นหลัก

#### • Input/Output Specification

Input: ผู้ใช้สามารถอินพุค ชื่อวงศ์, ชื่อสมุนไพร, ชื่อวิทยาศาสตร์, ชื่อท้องถิ่น, ชื่อสามัญ, ชื่อพ้อง, ลักษณะอาการของโรค หรือ ส่วนของสมุนไพรที่นำมาใช้ประโยชน์



รูปที่ 4.1 ตัวอย่างการอินพุดข้อมูล

Output: ชื่อสมุน ไพรทั้งหมดที่ ได้จากผลการค้นหา ซึ่งประกอบด้วยข้อมูลที่เกี่ยวข้องทั้งหมดที่มีอยู่ ในฐานข้อมูล เช่น ข้อมูลชื่อสมุน ไพร ไทย แหล่งที่อยู่ การขยายพันธุ์ ฤดูการออกดอก ลักษณะพืช ชื่อสารเคมี ส่วนที่นำมาใช้ประโยชน์เป็นต้น



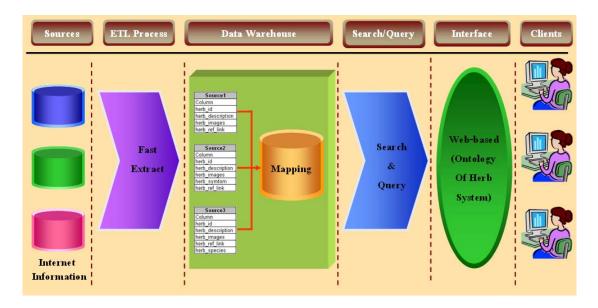
รูปที่ 4.2 ตัวอย่างการแสดงผลข้อมูล

#### • Functional Specification

- มี Dynamic dropdown list ที่มีชื่อทุกๆ ฐานข้อมูล และสามารถเพิ่มฐานข้อมูลเข้าไปได้ในที่ เดียวกัน
- สามารถตัดคำได้อย่างถูกต้อง
- สามารถค้นหาคำผ่านทางรูปสมุนไพรได้
- การค้นหาแบบต่อเนื่อง (Partial searching)
- ออกแบบออนโทโลจีให้มีหลายชั้น
- มีการเก็บสถิติจากการค้นหาคำ
- แสดงฐานข้อมูลของผลลัพธ์ที่ได้

- สามารถรายงานข้อมูลออกมาได้ ในรูปแบบของ Excel โดยใช้ฟังก์ชั่นการรวม (Merging function)

# • โครงสร้างของซอฟต์แวร์ (Design)



รูปที่4.3 กระบวนการทำงานของการค้นหาข้อมูลสมุนไพรไทย

จากรูปที่ 4.3 กระบวนการทำงานของระบบ แบ่งออกเป็น 2 ส่วน ได้แก่ ส่วน Data and User Management และ ส่วน Ontology-Based User Interface

### ส่วนที่ 1 Data and User Management Module

- 1. ใช้ crawler เก็บข้อมูลจากเว็บไซต์ ที่เราจะทำการศึกษา
- 2. นำข้อมูลที่เราทำการ crawler มาตัดคำ (Parser)
- 3. บรรจุลงในฐานข้อมูล

# ส่วนที่ 2 Ontology-Based User Interface Module

- 1. การจัดลำดับชั้นโครงสร้างข้อมูลสมุนไพรไทยในการออกแบบออนโทโลจี
- 2. พัฒนา GUI เพื่อเพิ่มความสะควกแก่ผู้ใช้งาน
- 3. เมื่อผู้ใช้ต้องการค้นหาคำ ระบบจะใช้เทคนิคการ query ในฐานข้อมูล และแสดงผลลัพธ์ที่ได้ในหน้า เว็บเพจ
- 4. มีการเก็บสถิติจากชื่อสมุนไพรไทยที่ผู้ใช้ค้นหา
- 5. ผู้ใช้สามารถค้นหาข้อมูลสมุนไพรไทยจากรูปสมุนไพรได้

# 5 ขอบเขตและข้อจำกัดของโปรแกรม

การศึกษาเพื่อพัฒนาระบบสืบค้นและจัดหมวดหมู่สมุนไพรไทย โคยใช้วิธีการออนโทโลจีครั้งนี้ได้ จำกัดขอบเขตของงานไว้ โดยจะใช้แหล่งข้อมูลสมุนไพรไทย ดังต่อไปนี้

- 1. ข้อมูลจากเวปไซด์ <a href="http://www.rspg.org/palace/chitralada/">http://www.rspg.org/palace/chitralada/</a> ซึ่งจัดทำโดยสำนักงานโครงการอนุรักษ์ พันธุกรรมพืชอันเนื่องมาจากพระราชดำริ สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี
- 2. ข้อมูลจากเวปไซค์ <a href="http://www.medplant.mahidol.ac.th/pubhealth/index.html">http://www.medplant.mahidol.ac.th/pubhealth/index.html</a> ซึ่งจัดทำโดย สำนักงานข้อมูลสมุนไพร คณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล

# 6 กลุ่มผู้ใช้โปรแกรม

- นักวิจัย
- อาจารย์
- นักเรียน นักศึกษา
- ผู้ที่สนใจทั่วไป

# 7 การทดลองและและการประเมินผล

# ในการทดสอบระบบและประเมินผลการดำเนินงานโครงการ แสดงรายละเอียดได้ดังต่อไปนี้

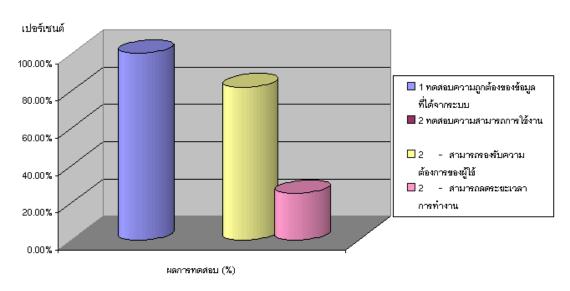
#### 7.1 การทดสอบระบบ

ในการทดสอบการทำงานของระบบใหม่ที่พัฒนาขึ้นมานี้ ได้ผลการทดสอบการทำงานดังตารางที่ 7.1 โดยแสดงเป็นแผนภูมิรูปภาพ ดังรูปที่ 7.1

ตาราง 7.1 ตารางแสดงการทดสอบระบบ

ลำดับ	ดัชนีชี้วัด	ผลการทดสอบ (%)	
1	ทดสอบความถูกต้องของข้อมูลที่ได้จากระบบ	100.00%	
2	ทคสอบความสามารถการใช้งาน		
	- สามารถรองรับความต้องการของผู้ใช้	82.00%	
	- สามารถลดระยะเวลาการทำงาน	25.00%	

#### การทดสอบระบบ



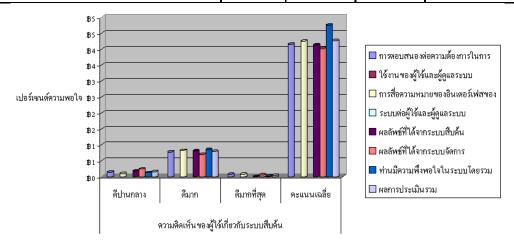
รูปที่ 7.1 ภาพแผนภูมิการทดสอบระบบ

#### 7.2 การประเมินระบบ

ในการประเมินการทำงานของระบบใหม่ที่พัฒนาขึ้นมานี้ ได้ผลการประเมิน ตามตารางที่ 7.2 โดย แสดงเป็นแผนภูมิรูปภาพ ดังรูปที่ 7.2

ตาราง 7.2 ตารางแสดงผลการประเมินระบบ

	ความคิดเห็นของผู้ใช้เกี่ยวกับระบบสืบค้น			
หัวข้อการประเมิน	ดีปาน กลาง	ดีมาก	ดีมากที่สุด	คะแนนเฉลี่ย
การตอบสนองต่อความต้องการในการ				
ใช้งานของผู้ใช้และผู้ดูแลระบบ	14.00%	78.00%	8.00%	4.15
การสื่อความหมายของอินเตอร์เฟสของ				
ระบบต่อผู้ใช้และผู้ดูแลระบบ	10.00%	82.50%	7.50%	4.25
ผลลัพธ์ที่ได้จากระบบสืบค้น	18.00 %	81.00%	1.00%	4.125
ผลลัพธ์ที่ได้จากระบบจัดการ	24.50 %	70.00%	5.50%	4.02
ท่านมีความพึ่งพอใจในระบบโดยรวม	13.00 %	85.00%	2.00%	4.75
ผลการประเมินรวม	15.90 %	79.30%	4.80%	4.259



รูปที่ 7.2 ภาพแผนภูมิแสดงระดับความพึงพอใจการใช้งานระบบ

# 8 ข้อเสนอแนะในการพัฒนาต่อไป

ระบบที่พัฒนาขึ้นนี้ มีการระบุขอบเขตการนำข้อมูลสมุนไพรไทยเข้าสู่ฐานข้อมูล ทำให้ข้อมูล สมุนไพรไทยยังไม่หลากหลายมากเพียงพอ อีกทั้งการจัดการไฟล์และฐานข้อมูลยังไม่สมบูรณ์มากพอ เนื่องจากยังต้องมีการสร้างฐานข้อมูลผ่านระบบ PHPMyAdmin ก่อน จึงจะสามารถปรับปรุงฐานข้อมูลผ่าน ระบบการจัดการไฟล์ได้

หากมีการพัฒนาระบบนี้ต่อไปในอนาคต ควรปรับปรุงในส่วนต่อไปนี้

- 1. ระบบการจัดการไฟล์
  - รองรับการสร้างฐานข้อมูลใหม่

เมื่อต้องการเพิ่มข้อมูลจากเว็บไซด์ใดๆ ที่ยังไม่เคยปรับปรุงผ่านระบบการจัดการไฟล์มาก่อน ให้ สามารถสร้างฐานข้อมูลใหม่ได้ ภายใต้การให้บริการการจัดการระบบ หรือรองรับการจัดการฐานข้อมูลผ่าน หน้าการให้บริการระบบ โดยอาจเชื่อมโยงไปยัง PHPMyAdmin ด้วยหลักการ Single-SignOn (SSO) [19]

• รองรับเว็บไซค์ต้นทางที่หลากหลาย

ในการเพิ่มเว็บไซค์ต้นทางเพื่อรองรับการสืบค้นที่มีประสิทธิภาพยิ่งขึ้น จำเป็นจะต้องอาศัย การศึกษาโครงสร้างของเว็บไซค์นั้นๆ และทำการตัดคำในแต่ละหน้าของข้อมูลสมุนไพรไทย ทั้งนี้ควร ศึกษาวิธีการคึงข้อมูลจากเว็บไซค์แบบต่างๆ ทั้งแบบเปิดให้บุคคลทั่วไปเข้าชมโคยไม่ต้องล็อกอิน และแบบ ปิดที่จำเป็นต้องล็อกอิน เพื่อให้เว็บไซค์ต้นทางตรวจสอบสถิติผู้เข้าเยี่ยมชมได้

- การตัดคำแบบอิงตามธรรมชาติภาษา (Natural Language)[20] เพื่อสนับสนุนการสืบค้น จากการพัฒนาระบบใหม่ พบว่า การตัดคำจากเว็บไซด์ต้นทางที่ได้ทำการคึงข้อมูลสมุนไพรไทยมา นั้น ยังไม่ฉลาดเพียงพอ เนื่องจากจำเป็นจะต้องกำหนดหัวข้อสมุนไพรไทย ให้อย่างชัดเจน และต้องตรงกัน ทุกตัวอักษร ในการพัฒนาต่อไปในอนาคต การตัดคำควรอิงจากธรรมชาติภาษา คือ แต่ละคำที่ตัดได้จาก ข้อมูลสมุนไพรไทยนั้น จะต้องมีความหมาย มีแนวโน้ม สถิติ ว่าควรจะรวบรวมกันไว้ หรือตัดต่อกันจนเกิด เป็นคำใด และเข้ากับหมวดหมู่สมุนไพรไทยที่ได้จัดแบ่งไว้ตามหลักการออนโทโลจีส่วนไหน
  - 2. ระบบสืบค้นสมุนไพรไทย
    - ปรังปรุงผลการรายงานทั้งในรูปแบบหน้าเว็บและการแสดงผลเป็นตาราง เป็นต้น
    - รองรับการค้นคำแบบลำดับชั้นของแต่ละหมวดหมู่สมุนไพรไทย

ในการพัฒนาระบบสมุนไพรไทย โดยใช้หลักการออนโทโลจี ทำให้มีการจัดแบ่งหมวดหมู่ของข้อมูล สมุนไพรไทย และแบ่งลำดับชั้นข้อมูลที่แยกย่อยลงไปอีก เช่น

หมวดหมู่การใช้ประโยชน์ แบ่งออกเป็นข้อย่อย คือ สรรพคุณ, ลักษณะอาการโรค และ ประโยชน์อื่นๆ เป็นต้น

หากต้องการก้นหาว่า สมุนไพรไทยชนิดไหนที่นิยมใช้รักษาอาการใช้ สามารถเลือกได้จาก หมวดหมู่การใช้ประโยชน์ ที่เป็นหัวข้อหลัก หรือเลือกจากลักษณะอาการของโรคซึ่งเป็นหัวข้อย่อยได้ เช่นกัน

# 9 เอกสารอ้างอิง

- ธรารัตน์ น้อยพรหม, ระบบสำหรับรวบรวมและสืบค้นข้อมูลทางสมุนไพร, วิทยานิพนธ์ปริญญาตรี สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ, 2549.
- ธิติวรรณ ทิมาภากร,ระบบสืบค้นและจัดหมวดหมู่สมุนไพรไทยโดยใช้วิธีการออนโทโลจี, วิทยานิพนธ์ปริญญาโท คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ, 2550
- รศ.คร. กฤษณะ ไวยมัย, Data Analysis Knowledge Discovery Laboratory (DAKDL),
   <a href="http://kdl.cpe.ku.ac.th/">http://kdl.cpe.ku.ac.th/</a>, 2008
- หน่วยบริการฐานข้อมูลสมุนไพรไทย, **สำนักงานข้อมูลสมุนไพรไทย**, คณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล (<a href="http://www.medplant.mahidol.ac.th/index.asp">http://www.medplant.mahidol.ac.th/index.asp</a>), 2000
- สำนักงานโครงการอนุรักษ์พันธุกรรมพืช อันเนื่องมาจากพระราชคำริ, สวนจิตรลดาพฤกษาพรรณ (http://www.rspg.org/palace/chitralada/cld6-2.htm), 2544
- สำนักบริการคอมพิวเตอร์ มหาวิทยาลัยเกษตรศาสตร์, การเกษตรเพื่ออาชีพ (<a href="http://web.ku.ac.th/agri">http://web.ku.ac.th/agri</a>), 2541
- ธรารัตน์ น้อยพรหม, ระบบสำหรับรวบรวมและสืบค้นข้อมูลทางสมุนไพร, วิทยานิพนธ์ปริญญาตรี สาขาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ, 2549.
- สิริวรรณ แต้วิจิตร, รายงานความก้าวหน้าการดำเนินงานโครงการวิจัย (Project) โครงการวิจัย ทุนอุดหนุนวิจัย มก. ปีงบประมาณ 2549, ห้องปฏิบัติการ DAKDL สาขาวิชาวิสวกรรมคอมพิวเตอร์ คณะวิสวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ, 2550.
- Ben-Miled, Z., N. Li, G.M. Kellett, B.Sipes and O. Bukhres, Proceedings of the IEEE, Complex Life Science Multi-Database Queries, Vol.90, No.11, 2002.
- Bomber, Ontology Language คืออะไร, NARISA.COM Forum (<a href="http://www.narisa.com/forums/index.php?act=idx">http://www.narisa.com/forums/index.php?act=idx</a>), 2005.
- สิทธิสักดิ์ บุญมาก, Search Engine คืออะไร, Blog Marketing (http://www.makemany.com), 2006
- PeeTai, Web Crawler, Mr.PeeTai (http://www.peetai.com), 2007
- Boonlert Aroonpiboon, <u>เทคนิคการพัฒนาเว็ปเพจ ตอน เตรียมเอกสารเว็ปสำหรับ Search Engine</u>,
   NECTEC's Web Based Laerning (http://www.nectec.or.th), 2006
- Deborah L., Frank van Harmelen, <u>OWL Web Ontology Language Overview</u>, W3C Org (http://www.w3.org), 2004.
- Macrakis, Ontology (Computer Science), Wikipedia (http://en.wikipedia.org), 2007.

- วิลาศ วูวงศ์, **การบูรณาการสารสนเทศด้วยเซแมนติกเวปเซอร์วิส**, ศูนย์สารสนเทศ กรมส่งเสริม การเกษตร, สถาบันเทคโนโลยีแห่งเอเชีย, 2007
- IT Guides, <u>Search Engine เครื่องมือในการค้นหาข้อมูลทางอินเตอร์เนต</u>, IT-Guides.COM (http://www.it-guides.com), 2004
- ชญานี ใพนุสิน, <u>Search Engine คือ</u>, Goto Know.Org (http://gotoknow.org/), 2005
- Keng, <u>Blog ช่วยให้อันดับใน Search Engine สูงขึ้นจริงหรือ?</u> , Keng.COM Blog & New Media Specialist (http://www.keng.com), 2005
- เพ็ญพรรณ อัศวนพเกียรติ, ระบบสำหรับรวบรวมและสืบค้นข้อมูลทางชีววิทยาโดยใช้ชีวออนโทโล จี, วิทยานิพนธ์ปริญญาโท คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ, 2547
- SineBot, <u>Single Sign-On (SSO)</u>, Wikipedia: the free encyclopedia (<a href="http://en.wikipedia.org/wiki/Single\_sign-on">http://en.wikipedia.org/wiki/Single\_sign-on</a>), 2008
- Domokato, <u>Natural Language Processing</u>, Wikipedia: the free encyclopedia (<u>http://en.wikipedia.org/wiki/Natural\_language\_processing</u>), 2008
- Boonyarat Phadermrod, A Novel Multi-Relation Querying method on RDBMS-Based
   Ontology, DAKDL, 2008
- <a href="http://www.rspg.org/palace/chitralada/">http://www.rspg.org/palace/chitralada/</a>
- <a href="http://www.medplant.mahidol.ac.th">http://www.medplant.mahidol.ac.th</a>

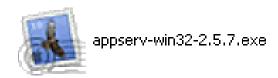
10 ภาคผนวก

# คู่มือการใช้งาน และติดตั้งโปรแกรมอัพเดตฐานข้อมูลอัตโนมัติ (Automatic crawler and parser)

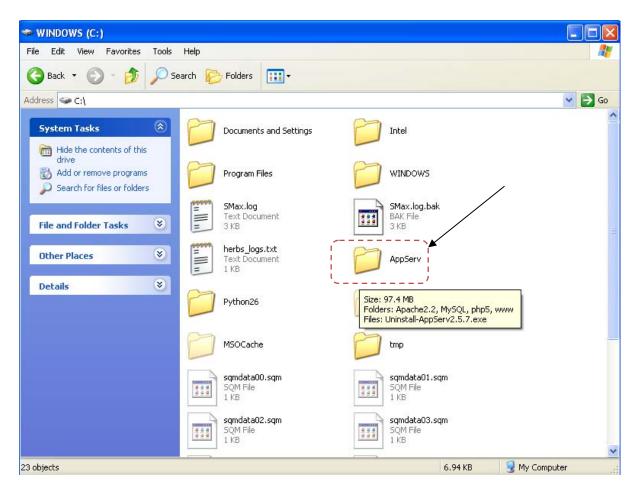
1.เริ่มต้นด้วยการลงโปแกรม JDKเพื่อรองรับการใช้ภาษาJAVA



2. โปรแกรม appserv-win32-2.5.7.exe เพื่อลองรับการใช้งานdatabase และ web-application

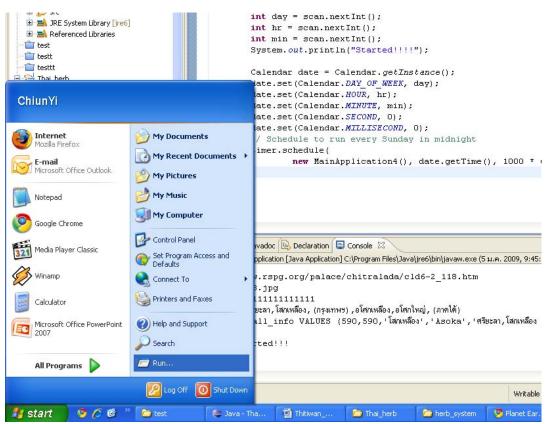


3.copy folder name "AppServ", "test" and paste in path "C:\"

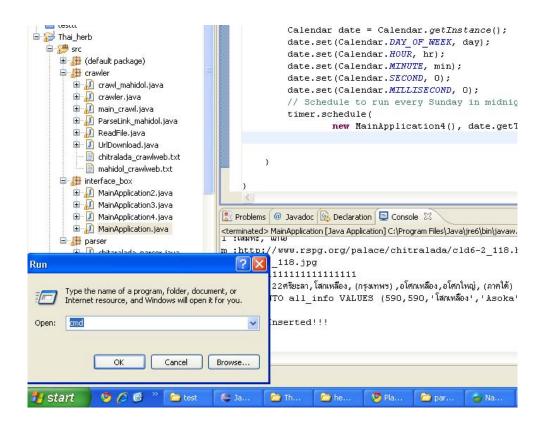


4. เริ่มต้นการรันระบบ อับเคตฐานข้อมูลอัตโนมัติ ด้วยการสั่ง cmd คั่งนี้

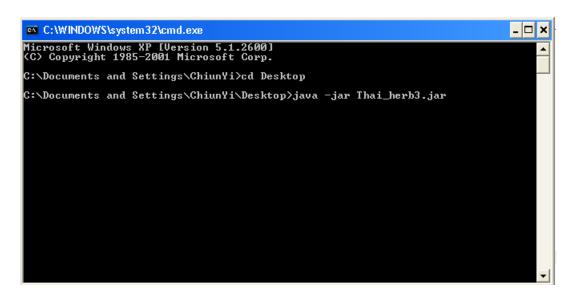
4.1 ไปที่ start แล้วคลิ้ก run



4.2 เมื่อขึ้นหน้าต่าง run ให้พิมพ์ cmd ลงในช่อง แล้วคลิ้ก ok



4.3 เปลี่ยน directory ใปยัง path ที่ที่วางไฟล์ htmlparser.jar, Thai\_herb3.jar และUtil.jar ไว้ พร้อม พิมพ์คำสั่ง ดังนี้ java –jar Thai\_herb3.jar แล้วกด Enter



4.4 เลือก วัน เวลา เป็นตัวเลขเพื่อใช้ในการ อัพเดท ข้อมูล ทุกรอบวันในหนึ่งสัปดาห์ โดยเริ่มต้นที่
 วันอาทิตย์ = 1 ไปจนถึงวันสุดท้ายคือวันเสาร์ = 7 ส่วนชั่วโมงคือแบ่งเป็น 12 ชั่วโมง เช่น เราต้องการตั้งเวลา
 ในการอัพเดตข้อมูลทุกวันจันทร์ เวลา 10.50น. จะต้องพิมพ์เลขดังรูปนี้ แล้วกด enter

```
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\ChiunYi\cdot Desktop

C:\Documents and Settings\ChiunYi\Desktop\java -jar Thai_herb3.jar
Input sechedule time with day, hr, min:
Sunday = 1 - Saturday = 7 totalhour = 12

2

10

50
```

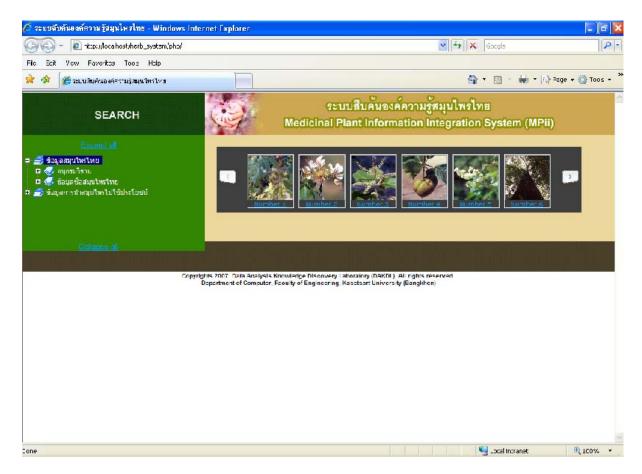
4.5 เมื่อถึงเวลาระบบจะเริ่มทำการรันเองอัตโนมัติ ดังรูป

```
processing http://www.rspg.org/palace/chitralada/cld6-2_098.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_099.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_199.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_199.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_190.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_192.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_193.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_195.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_196.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_196.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_198.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_199.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_199.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_119.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_111.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_111.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_111.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_113.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_115.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_115.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_115.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_115.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_115.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_116.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_116.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_118.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_118.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_118.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_118.htm
processing http://www.rspg.org/palace/chitralada/cld6-2_118.htm
processing http://www.rspg.org/palace/chitralada/cld6-3_htm
```

4.6 หากต้องการหยุดหรือออกจากระบบให้กด ctrl+c

#### คู่มือการใช้งาน

# ส่วนติดต่อกับผู้ใช้ (Graphic User Interface)



รูปที่ 1 หน้าเว็บของระบบสืบค้นสมุนไพรไทย

โดยระบบสืบค้นสมุนไพรไทยนี้ ผู้ใช้สามารถสืบค้นได้ทั้ง 2 วิธี คือ

- 1. พิมพ์คีย์เวิดของสมุนไพรไทย
- 2. เลือกรูปภาพสมุนไพรไทย

## วิธีที่ 1 พิมพ์คีย์เวิด

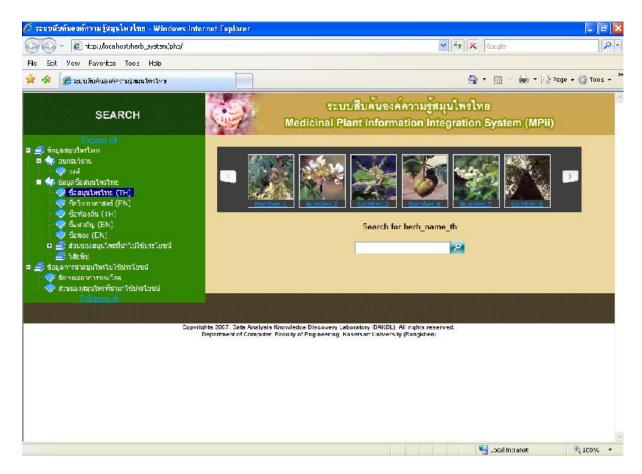
ผู้ใช้สามารถกดที่ ๒ เพื่อเปิดข้อมูลที่เป็นส่วยย่อยทีละขั้น หรือกด "Expand all" เพื่อเปิดข้อมูลที่เป็น ส่วนย่อยทั้งหมด หรือ กด ๒ เพื่อปิดข้อมูลที่เป็นส่วยย่อยทีละขั้น หรือ กด "Collapse all" เพื่อต้องการปิด ข้อมูลส่วนย่อยทั้งหมด

ข้อมูลส่วนย่อยประกอบด้วย :

- ข้อมูลสมุนไพรไทย
  - o อนุกรมวิชาน
    - วงศ์
  - O ข้อมูลชื่อสมุนไพรไทย
    - ชื่อสมุนไพรไทย
    - ชื่อวิทยาศาสตร์
    - ชื่อท้องถิ่น
    - ชื่อสามัญ
    - ชื่อพ้อง
  - ส่วนของสมุนไพรที่นำไปใช้ประโยชน์
    - เมล็ด
      - สรรพคุณ
        - 0 รักษาโรค
    - เนื้อ
      - สรรพคุณ
        - 0 รักษาโรค
    - คอก
    - ราก
    - เปลือก
    - ต้น
    - กลีบ
    - **■** ผถ
    - ใบ
    - ห้ว
    - แก่น
    - กิ่ง
    - อื่นๆ
  - 0 วิสัยพืช

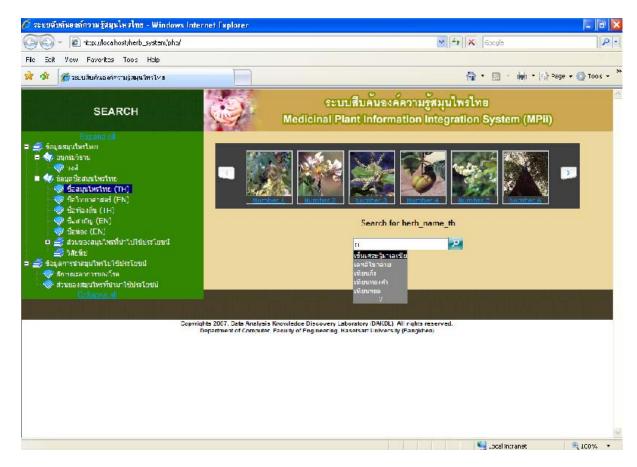
- ข้อมูลการนำสมุนไพรไปใช้ประโยชน์
  - ลักษณะอาการของโรค
  - ส่วนของสมุนไพรที่นำไปใช้ประโยชน์

เมื่อผู้ใช้เลือกหัวข้อของสมุนไพรที่ต้องการค้นหาแล้ว ระบบจะแสดงกล่องข้อความ เพื่อให้ผู้ใช้พิมพ์คำที่ ต้องการค้นหา ดังรูปที่ 2



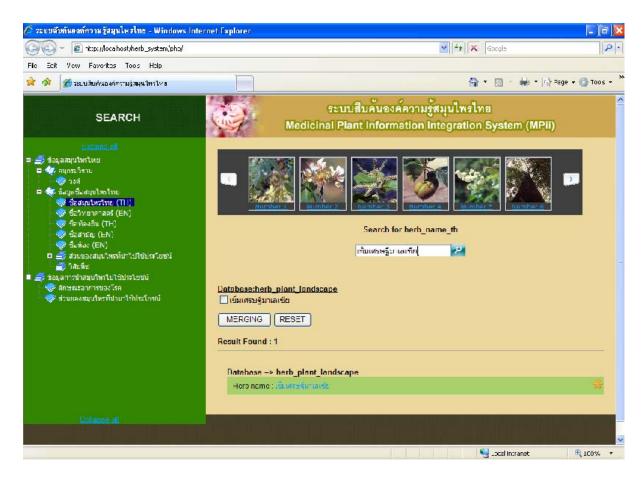
รูปที่ 2 ระบบแสดงกล่องข้อความ

เมื่อผู้ใช้พิมพ์คำที่ต้องการค้นหา ระบบจะแสดงคำที่เกี่ยวข้องขึ้นมาทันที เพื่อช่วยให้ผู้ใช้ใช้งานได้ง่ายยิ่งขึ้น ดังรูป 3



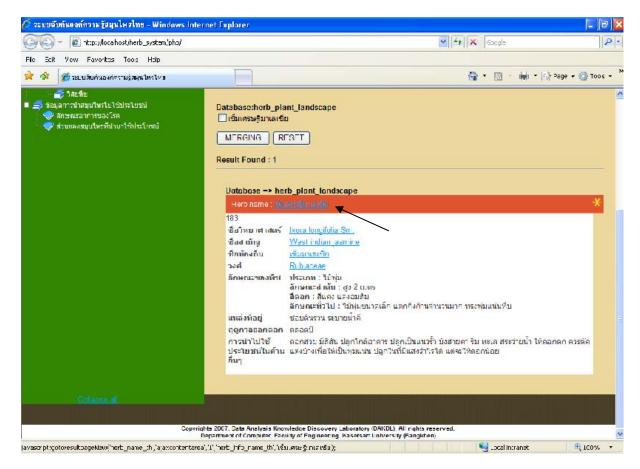
รูปที่ 3 ระบบแสดงคำที่เกี่ยวข้องเพื่อช่วยในการค้นหา

เมื่อเลือกคำที่ต้องการค้นหาแล้ว ให้กด 🌌 จากนั้นระบบจะทำการแสดงผลลัพธ์ ดังรูป 4



รูปที่ 4 แสคงผลลัพธ์

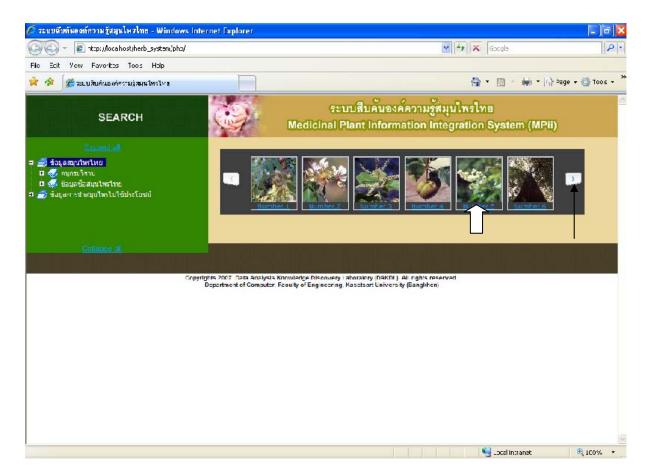
กคที่ชื่อสมุนไพรไทย เพื่อแสดงข้อมูลทั้งหมด ดังรูปที่ 5



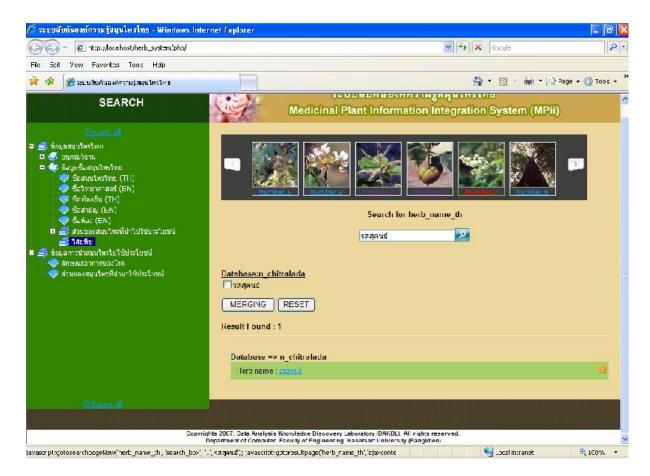
รูปที่ 5 แสคงข้อมูลทั้งหมค

และถ้าผู้ใช้ต้องการค้นหาคีย์เวิดอื่นๆ ก็ทำดังตามขั้นตอนเช่นกัน

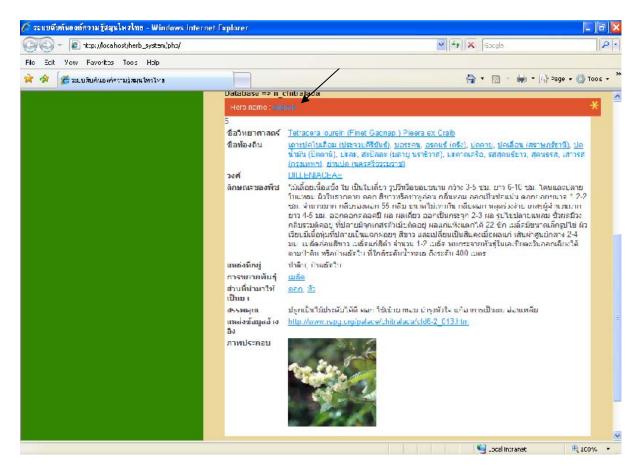
# วิธีที่ 2 เลือกรูปภาพสมุนไพรไทย



รูปที่ 6 สไลค์รูปสมุนไพร



รูปที่ 7 แสคงผลลัพธ์สมุนใพร



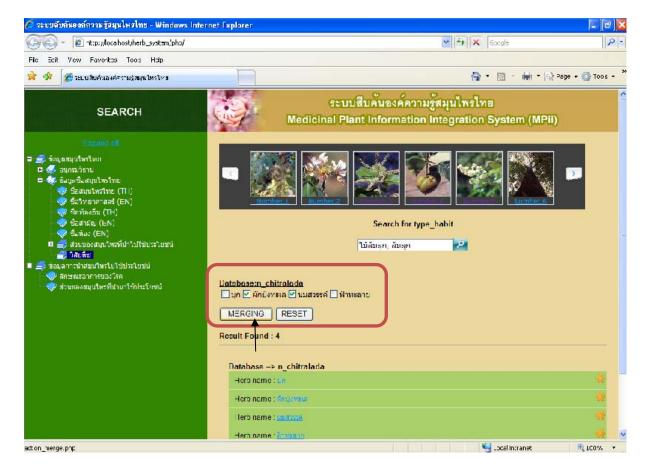
รูปที่ 8 เมื่อกคชื่อสมุนไพรแล้ว ระบบจะแสดงข้อมูลทั้งหมด

และเมื่อผู้ใช้ต้องการค้นหาจากรูปสมุนไพร ก็ทำเช่นขั้นตอนที่ได้กล่าวมา

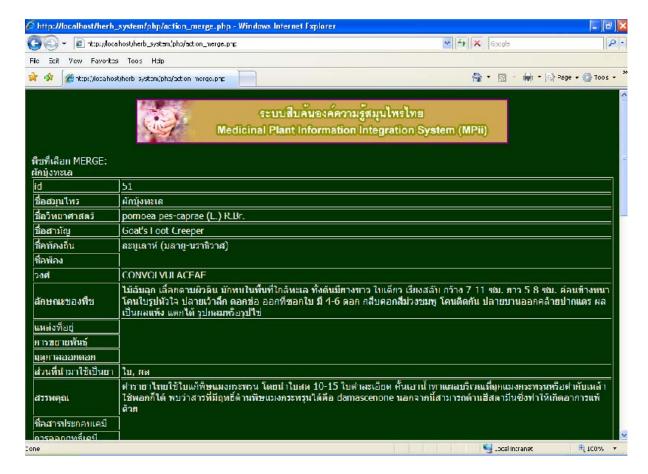
# ฟังก์ชั่นเพิ่มเติม

ฟังก์ชั่นเพิ่มเติมของระบบนี้คือ ฟังก์ชั่นการรวม (Merging) เมื่อระบบแสดงผลลัพธ์จากการค้นหา ผู้ใช้ สามารถเลือกสมุนไพรที่ต้องการ เพื่อแสดงข้อมูลสมุนไพรในรูปแบบของตารางโดยการกดที่ checkbox จากนั้นกดปุ่ม MERGING

หรือกดปุ่ม RESET เพื่อยกเลิกสมุนไพรที่ได้เลือกไป



รูปที่ 9 ฟังก์ชั่นการรวม (Merging function)



รูปที่ 10 ระบบแสดงผลลัพธ์จากการรวม





และเพื่อกลับไปยังหน้าหลักอีกครั้ง