

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อ 1) ออกแบบและพัฒนาคลังคำด้วยการจัดทำดัชนีเชิงความหมายสำหรับผลไม้เศรษฐกิจไทย และ 2) เพื่อออกแบบ พัฒนา และประเมินตัวแบบค้นคืนเชิงความหมายสำหรับผลไม้เศรษฐกิจไทย เพื่อช่วยให้ผู้ใช้ได้เข้าถึงความรู้เกี่ยวกับผลไม้เศรษฐกิจไทยได้อย่างสะดวกและรวดเร็ว ผลการวิจัยได้พัฒนาคลังคำโดยใช้แนวทางการจัดทำดัชนีเชิงความหมายโดยนำภาษาเอสเคโอเอสมาประยุกต์ใช้ ซึ่งคลังคำที่พัฒนาขึ้นจำแนกตามหมวดหมู่ได้ 10 คอลเล็กชัน ประกอบด้วยคอนเซ็ปต์จำนวน 101 คอนเซ็ปต์ การออกแบบออนโทโลยีตามแนวทางกระบวนการวิศวกรรมออนโทโลยี ประกอบด้วยคลาสหลักจำนวน 4 คลาส และคลาสย่อยจำนวน 12 คลาส การออกแบบออนโทโลยีโดยสร้างคุณสมบัติประเภท Datatype properties จำนวน 1 คุณสมบัติ คุณสมบัติประเภท Object properties จำนวน 12 คุณสมบัติสำหรับจัดเก็บรายการความรู้ เชื่อมโยงรายการความรู้ภายในฐานความรู้ และเชื่อมโยงรายการความรู้กับคำสำคัญในคลังคำ ใช้ภาษา SWRL สำหรับการสร้างกฎเพื่อเพิ่มเติมความรู้ในฐานความรู้ ตัวแบบการค้นคืนเชิงความหมายประกอบด้วย 3 ขั้นตอน 1) การประมวลคำสำคัญ โดยมีการประยุกต์ใช้ค่าสัมประสิทธิ์ความคล้ายคลึงของแจ็กการ์ด เพื่อระบุคำค้น 2) การประมวลผลคำค้น เพื่อสร้างคำสั่งสำหรับการค้นหาด้วยภาษา SPARQL สำหรับการค้นหาแบบ Concept-based และ Keyword-based และ 3) การแสดงผลลัพธ์ที่มีการคำนวณ ค่าความคล้ายคลึงเชิงมุมโคไซน์เพื่อหาคล้ายคลึงของแต่ละรายการความรู้ ผลการประเมินตัวแบบโดยการทดลองการค้นหา 30 ชุดคำค้นที่แตกต่างกัน แล้วนำมาวิเคราะห์ค่าความแม่นยำได้ร้อยละ 91.39 ค่าความระลึกร้อยละ 83.44 และค่าอัตราการรู้จำร้อยละ 87.23 ผลการประเมินระบบการค้นคืนเชิงความหมายที่พัฒนาขึ้นมาโดยทดสอบการค้นหาทั้งหมด 400 รายการ ผลการประเมินระบบในภาพรวมได้ค่าความแม่นยำร้อยละ 63.31 ค่าความระลึกร้อยละ 95.74 และค่าอัตราการรู้จำร้อยละ 76.22

คำสำคัญ: ตัวแบบการค้นคืนเชิงความหมาย, ดัชนีเชิงความหมาย, การวัดค่าความคล้ายคลึง, ผลไม้เศรษฐกิจไทย

Abstract

This research aimed to 1) design and develop a word repository by using semantic indexation for Thai economic fruits and 2) design, develop, and evaluate a semantic search model to support users to quickly and conveniently access Thai economic fruits knowledge. The research results were to develop word repository using a semantic indexing approach by applying the SKOS language. The developed word repository can be classified into 10 collections, consisting of 101 concepts. Designing an ontology based on an ontology engineering approach. It consists of 4 main classes and 12 subclasses. The ontology was designed by 1 datatype property and 12 object properties for describing knowledge and associating knowledge within their knowledge base and between the word repository. The SWRL language was used for defining rules to build additional knowledge in the knowledge base. The semantic retrieval model consists of 3 steps: 1) keyword processing 2) query processing and 3) display of results. The Jaccard similarity coefficient was applied to identify keywords to search knowledge in the knowledge base. The SPARQL language was used for querying based on concept-based and keyword-based search. The Cosine Similarity Measurement was applied to identify the similarity of knowledge for displaying the results to users. The evaluation of semantic retrieval model by searching experiments on 30 different search terms showed a Precision of 91.39%, Recall of 83.44% and F-measure of 87.23%. The evaluation of semantic retrieval system by using experiments on 400 different search terms showed a Precision of 63.31%, Recall of 95.74% and F-measure of 76.22%.

Keywords: Semantic retrieval model, Semantic indexation, Similarity measure, Thai economic fruits