

ระบบเชื่อมโยงข้อมูลทางบรรณานุกรม
ของผลงานวิชาการจากฐานข้อมูลสากล (International Databases)
สู่ฐานข้อมูลซึ่งใช้โปรแกรม DSpace ได้แบบอัตโนมัติ
Automatic Synchronization System for Bibliographic Data of
Academic Publications from International Databases to
Databases under DSpace Program

ปราชญ์ สวางศักดิ์, ดวงฤทัย ฦ ทุงฝ่าย, ภิญญาพัชญ์ ญาณะคำ
สำนักหอสมุด มหาวิทยาลัยเชียงใหม่ e-mail: pratchaya.s@cmu.ac.th

บทคัดย่อ

ระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล (International databases) สู่ฐานข้อมูลซึ่งใช้โปรแกรม DSpace ได้แบบอัตโนมัติ ถูกพัฒนาขึ้นโดยมีวัตถุประสงค์ เพื่อระบบเชื่อมโยงที่สามารถดึงข้อมูลทางบรรณานุกรมของผลงานวิชาการของมหาวิทยาลัยเชียงใหม่ที่เผยแพร่ในฐานข้อมูลสากล ให้บันทึกลงในคลังปัญญามหาวิทยาลัยเชียงใหม่ (CMU IR) แบบอัตโนมัติ ระบบนี้พัฒนาขึ้นจากเฟรมเวิร์ค Angular โดยเชื่อมโยงข้อมูลกับ International Database API และ DSpace API นอกจากนี้ ยังสามารถเชื่อมโยงข้อมูลกับคลังข้อมูลของหน่วยงานอื่นที่พัฒนาค้างข้อมูลโดยใช้ DSpace จากการทดสอบระบบด้วยการเชื่อมโยงข้อมูลระหว่างคลังปัญญามหาวิทยาลัยเชียงใหม่ (CMU IR) กับฐานข้อมูลสากล พบว่า ข้อมูลมีความถูกต้องในอัตราร้อยละ 100 ประโยชน์ที่ได้รับจากการพัฒนาระบบดังกล่าว สามารถลดภาระงาน และลดขั้นตอนการสร้างระเบียบบรรณานุกรมในคลังปัญญา CMU IR ได้เป็นอย่างดี

คำสำคัญ:

การเชื่อมโยงข้อมูล, คลังปัญญามหาวิทยาลัยเชียงใหม่, DSpace, API, ฐานข้อมูลสากล, การพัฒนาโปรแกรม

Abstract

The automatic synchronization system for bibliographic data from the international databases to the one under DSpace program has been created with the aim to develop the synchronization system that could extract bibliographic databases of Chiang Mai University's academic publications published in international databases and automatically enter the data in Chiang Mai University Intellectual Repository (CMU IR). This system has been developed from Angular framework by synchronizing the

databases with International Database API and DSpace API. Besides, the system could also synchronize the databases with repository of other organizations that developed the databases by applying DSpace. After testing the system by synchronizing the databases between CMU IR and international databases, it was found that the accuracy of the data was 100%. Thus, the benefit from developing this system could help decreasing the amount of work load and also perfectly help reducing the number of steps in creating bibliographic records in CMU IR.

Keyword:

Data Synchronization, Chiang Mai University Intellectual Repository (CMU IR), DSpace, API, International Databases, Program Development

บทนำ

มหาวิทยาลัยเชียงใหม่ได้พัฒนาค้นปัญญามหาวิทยาลัยเชียงใหม่ (CMU Intellectual Repository: CMU IR) เมื่อปี พ.ศ. 2557 โดยมีวัตถุประสงค์เพื่อพัฒนาค้นผลงานวิชาการของมหาวิทยาลัยเชียงใหม่ ซึ่งการดำเนินงานตามโครงการดังกล่าว เป็นการทำงานในรูปคณะกรรมการ ซึ่งมีคณะกรรมการมาจาก 4 ส่วนงานที่เกี่ยวข้อง ได้แก่ สำนักหอสมุด สำนักบริการเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย และศูนย์บริหารการวิจัย

สำนักหอสมุด มหาวิทยาลัยเชียงใหม่ ในฐานะส่วนงานที่รับผิดชอบการจัดหาและจัดระบบทรัพยากรสารสนเทศเพื่อสนับสนุนการเรียนการสอนและการวิจัยของมหาวิทยาลัย จึงได้รับมอบหมายให้รับผิดชอบในการรวบรวมข้อมูลผลงานวิชาการของมหาวิทยาลัยเชียงใหม่และนำเข้าสู่คลังปัญญา CMU IR ซึ่งพัฒนาขึ้นโดยใช้โปรแกรม DSpace ซึ่งเป็นโปรแกรมที่มีมาตรฐานการจัดเก็บในระดับสากล และมีมหาวิทยาลัยในประเทศไทยหลายแห่งนำไปใช้พัฒนาฐานข้อมูล เช่น คลังปัญญาจุฬาลงกรณ์มหาวิทยาลัย (Chulalongkorn University Intellectual Repository-CUIR), ฐานข้อมูลงานวิจัยและผลงานวิชาการ บุคลากรมหาวิทยาลัยมหิดล (Mahidol University Institutional Repository Database-Mahidol-IR) และคลังปัญญามหาวิทยาลัยสงขลานครินทร์ (PSU Knowledge Bank) โดยขอบเขตของข้อมูลที่สำนักหอสมุดได้รับมอบหมายให้ทำการรวบรวม มีดังนี้

1. งานวิจัย (Technical reports)
2. วิทยานิพนธ์ (Theses) และการค้นคว้าแบบอิสระ (Independent study)
3. บทความวิชาการ (Journal articles)
4. Conference papers
5. Working papers
6. Official report or Project
7. วารสารอิเล็กทรอนิกส์ (e-Journals)
8. สิทธิบัตร (Patents)
9. Book or book chapter

จากการรวบรวมข้อมูลพบว่า สามารถแยกข้อมูลได้เป็น 2 ประเภท คือ ข้อมูลที่มีการจัดเก็บอยู่แล้วในสำนักหอสมุด เช่น วิทยานิพนธ์ การค้นคว้าแบบอิสระ และข้อมูลผลงานวิชาการของมหาวิทยาลัยเชียงใหม่ที่เผยแพร่

ในฐานข้อมูลสากล ซึ่งข้อมูลที่สำนักหอสมุดไม่ได้จัดเก็บนี้สำนักหอสมุดไม่สามารถคัดลอกข้อมูลบรรณานุกรมไปยังคลังปัญญาแบบอัตโนมัติได้ ดังนั้นเพื่อเป็นการลดภาระงานและลดขั้นตอนการปฏิบัติงานของบุคลากรสำนักหอสมุด จึงจำเป็นต้องพัฒนาระบบเชื่อมโยงข้อมูลทางบรรณานุกรมของผลงานวิชาการมหาวิทยาลัยเชียงใหม่ที่เผยแพร่ในฐานข้อมูลสากล สู่ระบบคลังปัญญามหาวิทยาลัยเชียงใหม่

วัตถุประสงค์

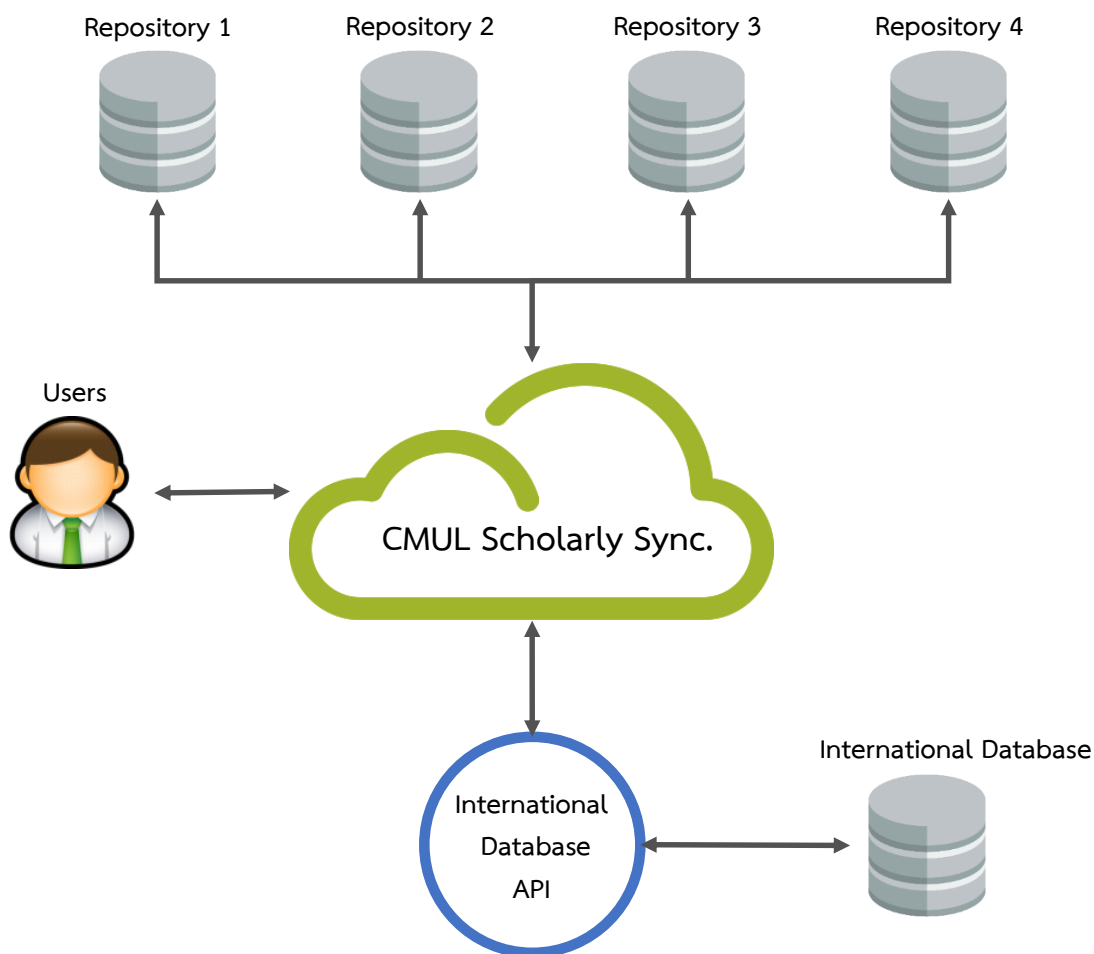
1. เพื่อพัฒนาระบบเชื่อมโยงฐานข้อมูลสากล ให้สามารถนำเข้าข้อมูลทางบรรณานุกรมของผลงานวิชาการสู่คลังปัญญา (Intellectual repository) ซึ่งพัฒนาโดยใช้โปรแกรม DSpace ได้แบบอัตโนมัติ
2. เพื่อพัฒนาระบบเชื่อมโยงที่สามารถตรวจสอบความถูกต้องและความซ้ำซ้อนของข้อมูลที่มีอยู่ในโปรแกรม DSpace
3. เพื่อจัดทำ Cloud Platform ให้หน่วยงานอื่นที่ใช้โปรแกรม DSpace ในการพัฒนาฐานข้อมูล ให้สามารถเชื่อมโยงข้อมูลจากฐานข้อมูลสากลกับฐานข้อมูลของหน่วยงานได้

ขั้นตอนและวิธีการดำเนินงาน

การดำเนินงานพัฒนาระบบเชื่อมโยงบรรณานุกรมจากฐานข้อมูลสากลนั้น ผู้พัฒนาได้เลือกใช้วิธีการพัฒนาแบบ Agile ซึ่งมีการกำหนดรอบอายุของระบบในรูปแบบ Sprint ที่ทำให้ระบบถูกปรับปรุงให้สนับสนุนความต้องการของผู้ใช้งานอยู่เสมอ โดยในแต่ละรอบอายุของระบบสามารถแบ่งออกเป็นกระบวนการย่อย ดังนี้

1. วิเคราะห์และออกแบบ (System analysis and design)

ในขั้นตอนการวิเคราะห์และออกแบบระบบมีวัตถุประสงค์ในการวางแผนการพัฒนาระบบโดยดูจากกลุ่มผู้ใช้งาน (Stakeholder) ระบบที่เกี่ยวข้อง เครื่องมือที่ใช้ในการพัฒนา และนำมาร้อยเรียงเป็นขั้นตอนและกระบวนการทำงานของระบบที่พัฒนา ซึ่งระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล สามารถออกแบบการทำงานออกมาได้เป็นกระบวนการ ดังนี้



ภาพที่ 1 กระบวนการทำงานของระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล

จากการออกแบบกระบวนการทำงานของระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล จะเห็นว่าระบบถูกออกแบบให้สามารถเชื่อมโยงข้อมูลกับคลังปัญญาได้หลายแห่งขึ้นอยู่กับสิทธิที่ได้รับในการแก้ไขข้อมูลของผู้ใช้งาน โดยไม่จำกัดว่าจะต้องสามารถเชื่อมโยงกับคลังปัญญาของมหาวิทยาลัยเชียงใหม่เท่านั้น โดยเมื่อผู้ใช้งานตั้งค่าการเชื่อมโยงข้อมูลกับส่วนติดต่อกับผู้ใช้งาน (User interface) ระบบจะทำการเชื่อมโยงข้อมูลตามที่ตั้งค่าไว้กับฐานข้อมูลปลายทางทันที

2. พัฒนาระบบ (System development)

จากการวิเคราะห์และออกแบบระบบทำให้ผู้พัฒนาสามารถกำหนดขอบเขตของการพัฒนาระบบ ผังกระบวนการทำงานของระบบ จะพบว่าระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากลนั้น มีส่วนเชื่อมต่อ (Interface) อยู่ทั้งหมด 3 จุด คือ

2.1 ส่วนเชื่อมต่อกับผู้ใช้งาน

2.2 ส่วนเชื่อมต่อกับฐานข้อมูลสากล

2.3 ส่วนเชื่อมต่อกับคลังปัญญา

โดยส่วนเชื่อมต่อทั้ง 3 จุดจะทำงานตามลำดับคือ ส่วนเชื่อมต่อกับผู้ใช้งานจะรับข้อมูลการตั้งค่าการเชื่อมโยงข้อมูลกับผู้ใช้งาน จากนั้นระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล จะทำการร้องขอ (Request) ข้อมูลจากฐานข้อมูลสากลตามผู้ใช้งานได้ตั้งค่าไว้ และทำการบันทึกข้อมูลที่รวบรวมได้ลงฐานข้อมูลคลังปัญญาผ่านส่วนเชื่อมต่อกับคลังปัญญา ดังนั้น ในการพัฒนาระบบ ผู้พัฒนาจึงเลือกใช้เครื่องมือในการพัฒนาจากการเชื่อมต่อทั้ง 3 ของระบบดังนี้

1. Angular สำหรับพัฒนาระบบในส่วนเชื่อมต่อกับผู้ใช้งานในการสร้างระบบในส่วนของ Front-end ผ่านเว็บไซต์เพื่อให้ผู้ใช้งานสามารถใช้งานได้ง่ายในทุกที่ ทุกเวลา และทุกอุปกรณ์สื่อสาร

2. International database API สำหรับพัฒนาระบบในส่วนเชื่อมต่อกับฐานข้อมูลสากล ในการรวบรวมข้อมูล (Data harvesting) เมทาเดตาของระเบียบบรรณานุกรมจากฐานข้อมูลตามมาตรฐาน PRISM Dublin Core Metadata Specification (PRISMDCMS) ในรูปแบบโครงสร้างข้อมูลแบบ JSON (JavaScript Object Notation)

3. DSpace API สำหรับพัฒนาในส่วนเชื่อมต่อกับคลังปัญญา เพื่อเป็นช่องทางในการตรวจสอบความซ้ำซ้อนของข้อมูลที่รวบรวมมา และบันทึกข้อมูลลงในคลังปัญญา การพัฒนาระบบโดยแบ่งส่วนเชื่อมต่อออกเป็น 3 จุดนี้ ยังมีข้อดีในการบูรณาการเชื่อมโยงกับฐานข้อมูลสากลได้ทุกฐานที่มีการเปิดให้เชื่อมโยงข้อมูลด้วย API โดยไม่จำเป็นต้องพัฒนาโปรแกรมใหม่ทั้งหมด

เมื่อผู้พัฒนาได้พัฒนาระบบเสร็จสิ้นแล้ว จึงได้นำระบบอัปโหลดขึ้นไปยังเครื่องแม่ข่ายคอมพิวเตอร์เพื่อเตรียมความพร้อมในการวิพากษ์ระบบในขั้นตอนถัดไป

3. วิพากษ์ระบบ (System review)

หลังจากพัฒนาระบบเสร็จสิ้นผู้พัฒนาได้นำระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากลมาเสนอภายในงานวันนวัตกรรม สำนักหอสมุด 2561 (CMUL Innovation Day 2018) ให้ผู้เข้าร่วมงาน ซึ่งประกอบด้วยบุคลากรสำนักหอสมุด มหาวิทยาลัยเชียงใหม่ บุคลากร ภาควิชาบรรณารักษศาสตร์ มหาวิทยาลัยเชียงใหม่ บุคลากรจากสำนักหอสมุด มหาวิทยาลัยแม่โจ้ และบุคลากรจากสำนักหอสมุด มหาวิทยาลัยพายัพ เข้าร่วมการวิพากษ์ระบบ จำนวน 105 คน เพื่อแสดงความคิดเห็นในการพัฒนาระบบ และประเมินความพึงพอใจในระบบเป็นครั้งสุดท้ายก่อนนำระบบไปใช้งานจริง ซึ่งจากการวิพากษ์ระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล ได้รับความพึงพอใจในระดับมากที่สุด (4.30) และมีข้อเสนอแนะดังนี้

3.1 สำนักหอสมุด มหาวิทยาลัยเชียงใหม่ ควรจัดอบรมการใช้งานระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากลให้กับหน่วยงานภายนอก

3.2 สำนักหอสมุด มหาวิทยาลัยเชียงใหม่ ควรเปิดระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากลให้หน่วยงานภายนอกเข้าถึงได้

4. การใช้งานจริง (Launch)

เมื่อทำการแก้ไขระบบตามการวิพากษ์ผู้พัฒนาจึงได้อัปโหลดระบบขึ้นบนเครื่องแม่ข่ายคอมพิวเตอร์ที่มีคุณสมบัติดังนี้

4.1 ระบบปฏิบัติการ Windows 2016 Server

4.2 หน่วยประมวลผลกลาง 2 แกนประมวลผลแบบเสมือน (2 vCore)

4.3 หน่วยความจำหลัก 4 GB

4.4 ฮาร์ดดิสก์ 200 GB SSD

จากนั้น ทำการเชื่อมต่อกับคลังปัญญามหาวิทยาลัยเชียงใหม่ พร้อมกับแจ้งให้ผู้เกี่ยวข้องในการปรับปรุงคลังปัญญา เพื่อใช้งานจริงต่อไป

สรุปผล

ระบบเชื่อมโยงข้อมูลทางบรรณานุกรมของผลงานวิชาการจากฐานข้อมูลสากล (International databases) ฐานข้อมูล ซึ่งใช้โปรแกรม DSpace ได้แบบอัตโนมัติ ถูกพัฒนาขึ้นโดยมีวัตถุประสงค์ในการรวบรวมข้อมูลผลงานวิชาการของมหาวิทยาลัยเชียงใหม่ บันทึกคลังปัญญามหาวิทยาลัยเชียงใหม่ (CMU IR) ได้แบบอัตโนมัติ โดยระบบถูกออกแบบมาให้สามารถใช้งานร่วมกับฐานข้อมูลที่พัฒนาด้วยโปรแกรม DSpace ดังนั้น จึงสามารถใช้งานได้กับฐานข้อมูลของทุกหน่วยงาน ที่พัฒนาด้วยโปรแกรม DSpace นอกเหนือจากคลังปัญญามหาวิทยาลัยเชียงใหม่ซึ่งขึ้นอยู่กับ URL ปลายทางที่ผู้ใช้งานตั้งค่าที่หน้าโปรแกรม

The screenshot displays a three-step configuration process for a DSpace repository. Step 1, 'SETUP REPOSITORY', shows the 'Repository REST URL' field populated with 'http://cmuir.cmu.ac.th/rest/'. Step 2, 'SELECT TARGET COLLECTION', shows a dropdown menu for 'Target Collection' with 'CMUL: Journal Articles' selected. Step 3, 'DATA CLUSTERING', includes fields for 'Affiliation ID' (60000881), 'Article Year' (ปีที่เผยแพร่), 'Username' (ชื่อผู้ใช้), and 'Password' (รหัสผ่าน). A red 'IMPORT START' button is located at the bottom right of the interface.

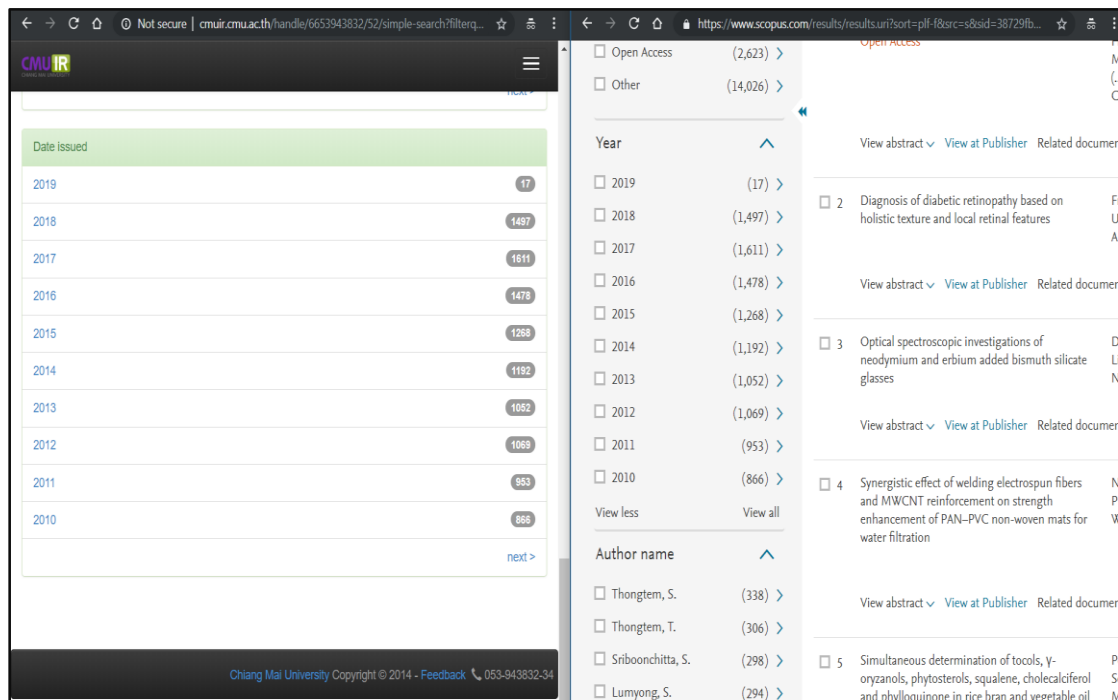
ภาพที่ 2 หน้าหลักระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล ฐานข้อมูลที่พัฒนาโดยใช้โปรแกรม DSpace

ระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล ถูกพัฒนาขึ้นให้ใช้งานง่าย (User friendly) โดยผู้ใช้งานตั้งค่าในการเชื่อมโยงเพียงแค่ 3 ขั้นตอน คือ

1. ตั้งค่า URL ของฐานข้อมูลปลายทาง
2. ระบุ Collection ปลายทางที่จะนำข้อมูลเข้า
3. กำหนดขอบเขตข้อมูลที่จะนำเข้าและการยืนยันตนเอง

จากการทดสอบระบบเชื่อมโยงข้อมูลบรรณานุกรมจากฐานข้อมูลสากล โดยการเชื่อมโยงข้อมูลระหว่างคลังปัญญามหาวิทยาลัยเชียงใหม่กับฐานข้อมูลสากล และรวบรวมระเบียบบรรณานุกรมตั้งแต่ปี พ.ศ. 2553-2562 (ค.ศ. 2010-2019) ณ วันที่ 1 ตุลาคม พ.ศ. 2561 พบว่า ระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล

สามารถรวบรวมระเบียบบรรณานุกรมลงคลังปัญญามหาวิทยาลัยเชียงใหม่ ได้จำนวน 11,003 ระเบียบ ซึ่งมีความคลาดเคลื่อนจากข้อมูลในฐานข้อมูลสากล ร้อยละ 0 ภายในระยะเวลา 1 ชั่วโมง



ภาพที่ 3 การเปรียบเทียบจำนวนข้อมูลจากฐานข้อมูลสากลกับจำนวนข้อมูลที่ระบบเชื่อมโยงข้อมูลทางบรรณานุกรมฯสามารถรวบรวมได้

ปัญหาและข้อเสนอแนะ

จากการทดสอบระบบโดยการเชื่อมโยงข้อมูลกับคลังปัญญามหาวิทยาลัยเชียงใหม่ พบว่าระบบระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากล สามารถใช้งานได้ดี มีความแม่นยำในระดับ 100% แต่เมื่อนำระบบมาทดสอบกับฐานข้อมูลอื่น ที่พัฒนาขึ้นจากโปรแกรม DSpaceกลับพบปัญหา ไม่สามารถบันทึกระเบียบบรรณานุกรมลงในเขตข้อมูลของฐานข้อมูลที่มีการแก้ไขเมตาตา (Metadata) เพิ่มเติมนอกเหนือจากที่โปรแกรม DSpace มีให้ เนื่องจากระบบเชื่อมโยงข้อมูลทางบรรณานุกรมฯ และฐานข้อมูลสากลที่ใช้ทดสอบการเชื่อมโยง ใช้มาตรฐาน PRISMDCMS ดังนั้น จึงไม่สามารถรองรับเขตข้อมูลที่อยู่นอกเหนือจากมาตรฐานได้

จากปัญหาที่พบผู้พัฒนาจึงมีข้อเสนอแนะในการพัฒนาระบบในรุ่นถัดไป ให้มีการตั้งค่าด้านการบริหารจัดการเขตข้อมูลในส่วนเชื่อมต่อกับผู้ใช้งาน เพื่อให้ผู้ใช้งานสามารถสร้างเขตขึ้นมาใหม่เพื่อให้รองรับกับเมตาตาที่เพิ่มขึ้นจากการเพิ่มเขตข้อมูลในโปรแกรม DSpace

การนำไปใช้ประโยชน์

1. ประโยชน์ที่มีต่อมหาวิทยาลัยเชียงใหม่

หลังจากพัฒนาระบบเชื่อมโยงข้อมูลทางบรรณานุกรมจากฐานข้อมูลสากลเสร็จสิ้นแล้ว สำนักหอสมุดได้เปิดระบบให้กับบุคคลภายนอกใช้ผ่านระบบอินเทอร์เน็ตและแจ้งให้กับผู้ดูแลระบบคลังปัญญามหาวิทยาลัยเชียงใหม่รับทราบ ซึ่งวันที่ 1 ตุลาคม พ.ศ. 2561 ระบบดังกล่าวได้รวบรวมผลงานวิชาการของมหาวิทยาลัยเชียงใหม่ บันทึกลงในคลังปัญญาแล้ว จำนวน 12,955 ระเบียบ การพัฒนาระบบดังกล่าว ช่วยให้สำนักหอสมุด มหาวิทยาลัยเชียงใหม่สามารถลดจำนวนบุคลากรในการสร้างระเบียบบรรณานุกรมผลงานวิชาการของมหาวิทยาลัยเชียงใหม่ที่ได้รับการเผยแพร่ในฐานข้อมูลสากล จำนวน 2 คน

2. ประโยชน์ที่มีต่อหน่วยงานอื่น

หน่วยงานอื่นที่ใช้โปรแกรม DSpace ในการพัฒนาระบบคลังปัญญาหรือฐานข้อมูลผลงานวิชาการ สามารถใช้ Cloud platform รวบรวมข้อมูล เพื่อบันทึกระเบียบบรรณานุกรมลงคลังปัญญาแบบอัตโนมัติได้

รายการอ้างอิง

- Highsmith, J. (2004). *Agile project management: Creating innovative products* (Agile software development series). Boston: Addison-Wesley.
- PRISM: *Publishing requirements for industry standard metadata*. (2012). Retrieved from. http://www.prismstandard.org/specifications/3.0/PRISM_Dublin_Core_Metadata_3.0.pdf