

การจัดการระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามและการเชื่อมโยงแบบเปิด

Managing the Knowledge System of Ethnic Groups in Vietnam and Linked Open Data

Thai Trinh Ngoc^{1,*} และ วิระพงศ์ จันทร์สนาม²

Thai Trinh Ngoc^{1,*} และ Wirapong Chansanam²

¹ สาขาวิชากลุ่มน้ำโขงศึกษา คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น ประเทศไทย; Mekong River Basin Studies Faculty of Humanities and Social Sciences Khon Kaen University, Thailand

² สาขาวิชาสารสนเทศศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น ประเทศไทย; Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Thailand

* Corresponding author email: chiwu2020@gmail.com

บทคัดย่อ

วัตถุประสงค์: เพื่อจัดการระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามและการเชื่อมโยงแบบเปิด

วิธีการศึกษา: ทำการค้นคว้าข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามบนอินเทอร์เน็ต จัดระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนาม และทำการเชื่อมโยงข้อมูลกับแหล่งข้อมูลอื่น ๆ บนอินเทอร์เน็ต โดยใช้ออนโทโลยี และระบบ Reconciliation linking data by OpenRefine and RDF extension

ข้อค้นพบ: 1) การสร้าง SPARQL และภาษาแบบสอบถาม RDF เกี่ยวกับกลุ่มชาติพันธุ์เวียดนามบนอินเทอร์เน็ต อนุญาตให้ผู้ใช้สืบค้นข้อมูลจากฐานข้อมูลหรือแหล่งข้อมูลใด ๆ ที่สามารถแมปกับ RDF มาตรฐาน SPARQL ได้รับการออกแบบและรับรองโดย W3C และช่วยให้ผู้ใช้และนักพัฒนามุ่งเน้นไปที่สิ่งที่พวกเขาต้องการทราบแทนการจัดระเบียบฐานข้อมูล 2) การเลือก URI และคำศัพท์เพื่อจัดประเภทและกำหนดทรัพยากร การระบุดังกล่าวช่วยให้สามารถโต้ตอบกับอินสแตนซ์ของทรัพยากรบนเครือข่ายโดยใช้โปรโตคอลเฉพาะ สคีมาที่ระบุไวยากรณ์มาตรฐานและโปรโตคอลที่เกี่ยวข้องสำหรับการกำหนด URI แต่ละรายการ

การประยุกต์ใช้จากการศึกษา: งานวิจัยนี้เป็นแนวทางแก่หน่วยงานของรัฐ สถาบันต่าง ๆ และผู้ที่สนใจไปใช้ในการเรียนรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนาม เนื่องจากเวียดนามกำลังอยู่ในขั้นตอนของการเปลี่ยนแปลงทางดิจิทัลสำหรับหน่วยงานภาครัฐ ข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามส่วนใหญ่เป็นเอกสารทางราชการที่ยังไม่ได้รับการเผยแพร่เท่าที่ควร ดังนั้น งานวิจัยชิ้นนี้ถือว่าเป็นงานตัวอย่างเพื่อเปลี่ยนแปลงทางดิจิทัลในด้านอื่น ๆ ของหน่วยงานภาครัฐ

คำสำคัญ: การจัดการระบบองค์ความรู้ กลุ่มชาติพันธุ์ เวียดนาม การเชื่อมโยงแบบเปิด

Abstract

Purpose: The research aims to manage knowledge on ethnic groups and Linked Open Data (LOD).

Methodology: Information on ethnic groups in Vietnam is gathered from online sources. The Vietnamese ethnic group's knowledge system is then managed to provide an overview and facilitate further study. Ontology and Reconciliation linking data using Open Refine and RDF extension are used to connect the knowledge system to other online resources.

Findings: The results demonstrate that 1) the creation of SPARQL and RDF query language enables users to retrieve information from any database or data source mapped to RDF. The standardized SPARQL allows users to focus on their queries instead of organizing databases; 2) the selection of URIs and terminology for resource classification and definition enables interaction with resource instances on the network using specific protocols and standardized syntax.

Applications of this study: The research findings serve as a reference for government agencies, institutions, and individuals interested in learning about Vietnamese ethnic groups. Given Vietnam's ongoing digital transformation, where official information on ethnic groups is often limited, this research serves as a model for digital transformation in government agencies.

Keywords: Knowledge management system, Ethnic groups, Vietnam, An open link

1. บทนำ

กลุ่มชาติพันธุ์ในเวียดนาม หรือ คนเวียดนาม เป็นคำนามสามัญ หมายถึง ประชาชนที่อาศัยอยู่ในดินแดนเวียดนาม ปัจจุบัน เวียดนามมีกลุ่มชาติพันธุ์ 54 กลุ่ม เช่น กิ่งห์ (Kinh) ไท (Thái) ม้ง (Mông) ฯลฯ อยู่ในรายชื่อชาติพันธุ์ในเวียดนาม สำนักงานสถิติทั่วไปของเวียดนาม (เป็นหน่วยงานโดยตรงในสังกัดกระทรวงการวางแผนและการลงทุน ทำหน้าที่ให้คำปรึกษาและช่วยเหลือรัฐมนตรีว่าการกระทรวงการวางแผนและการลงทุนในการจัดการสถิติของรัฐ ประสานงานกิจกรรมทางสถิติและจัดกิจกรรมทางสถิติ และให้ข้อมูลสถิติทางเศรษฐกิจและสังคมแก่หน่วยงาน องค์กรและบุคคลตามที่กฎหมายกำหนด) เป็นองค์กรตัดสินใจจำนวนกลุ่มชาติพันธุ์อยู่ในเวียดนามตามหมายเลข 121-TCTK/PPCĐ 2 มีนาคม พ.ศ. 2522 (Ngoc Khanh, 2019) และได้รับการยอมรับจากคณะกรรมการกิจการชาติพันธุ์และรัฐบาลเวียดนาม ลักษณะเด่นของกลุ่มชาติพันธุ์เวียดนามคืออาศัยอยู่ด้วยกัน ติดกันตลอดกระบวนการสร้างและป้องกันประเทศ แต่ละกลุ่มชาติพันธุ์มีลักษณะทางวัฒนธรรมเป็นของตัวเอง ซึ่งแสดงออกด้วยภาษา การแต่งกาย ประชากร ขนบธรรมเนียม การปฏิบัติ ฯลฯ ทำให้วัฒนธรรมเวียดนามมีความสมบูรณ์ยิ่งขึ้นในอัตลักษณ์ ในบรรดากลุ่มชาติพันธุ์ต่าง ๆ ในเวียดนาม กลุ่ม

ชาติพันธุ์เวียดนาม (ชาว Kinh) มีประชากรมากที่สุด คิดเป็นประมาณ 86% ของประชากรในประเทศ ซึ่งเป็นกลุ่มชาติพันธุ์ที่มีประสบการณ์กว้างขวางในการปลูกข้าวนาดำ (Ngoc Khanh, 2019) งานฝีมือที่ล้ำสมัย และแรงงานจำนวนมากในอุตสาหกรรม บริการ วิทยาศาสตร์และเทคโนโลยี นอกจากนี้ ชาติพันธุ์ยังมีประชากรและระดับการพัฒนาทางเศรษฐกิจที่แตกต่างกัน แต่ละกลุ่มชาติพันธุ์มีประสบการณ์ของตนเองในด้านต่าง ๆ เช่น พืชผลทางอุตสาหกรรม ไม้ผล ปศุสัตว์และหัตถกรรม กิจกรรมทางอุตสาหกรรม การบริการ วัฒนธรรม วิทยาศาสตร์และเทคนิคของเวียดนามล้วนมีส่วนร่วมจากชาติพันธุ์ อย่างไรก็ตาม ในปัจจุบันบนอินเทอร์เน็ตยังไม่มีกรอบการบรรยายสำหรับเทคโนโลยีเว็บ อินเทอร์เน็ต และเครื่องมือการจัดการในการเชื่อมโยงและนำเสนอข้อมูลเกี่ยวกับชาวเวียดนามในด้านต่าง ๆ มากมาย เช่น ที่มา ชื่อ ภาษา และการเขียน ฯลฯ รวมถึงการไม่มีเครื่องมือในการสร้างการเชื่อมโยงข้อเท็จจริงไปยังแหล่งความรู้ที่เกี่ยวข้องออนไลน์เป็นวิธีการจัดระบบความรู้และการแยกข้อมูลเชิงลึกเพื่อช่วยรวบรวมความเข้าใจแลกเปลี่ยนความเข้าใจ (Noy and McGuinness, 2001) บทความนี้นำเสนอผลลัพธ์ของการสร้างชุดข้อมูลการแบ่งปันทางอินเทอร์เน็ตของกลุ่มชาติพันธุ์ในประเทศเวียดนามโดยการประยุกต์ใช้เทคนิคลิงก์เดต้าแบบออนไลน์

ข้อมูลเปิดอาจเป็นชุดข้อมูลที่รัฐบาลรวบรวม (เช่น data.gov, data.gov.uk เป็นต้น) ข้อมูลจากการศึกษาและวารสารทางวิทยาศาสตร์ (Open access journals) ข้อมูลข้อความหรือข้อมูลมัลติมีเดียจากชุมชน (เช่น วิกีพีเดีย วิกีฮาว ฯลฯ) ลักษณะทั่วไปของแหล่งข้อมูลทั้งหมดคือมีการเชื่อมโยงกันทั้งหมดและสร้างขึ้นจากการมีส่วนร่วมของชุมชนผู้ใช้งานใบอนุญาตที่เหมือนกัน ดังนั้นข้อมูลเปิดจึงเป็นข้อมูลอิเล็กทรอนิกส์หรือดิจิทัลที่ต้องเปิดให้ทุกคนใช้และเผยแพร่ใหม่ได้ตามต้องการโดยไม่มีการจำกัดลิขสิทธิ์ สิทธิบัตร หรือกลไกอื่น ๆ การควบคุมอื่น ๆ (Chansanam, Tuamsuk, Chaikhambung & Sugimoto, 2020)

ลิงก์เดต้า (Linked Data) เป็นหนึ่งในแนวคิดหลักของเว็บเชิงความหมาย (Semantic Web) หรือเป็นเว็บของข้อมูล (Web of Data) เว็บเชิงความหมาย เป็นเว็บเกี่ยวกับการเชื่อมโยงระหว่างชุดข้อมูลที่เข้าใจได้ไม่เฉพาะกับมนุษย์เท่านั้น แต่ยังรวมถึงคอมพิวเตอร์ด้วย และลิงก์เดตตานำเสนอแนวทางปฏิบัติที่ดีที่สุดสำหรับการสร้างการเชื่อมโยงข้อมูลเหล่านั้น ลิงก์เดต้าคือชุดของหลักการออกแบบสำหรับการแบ่งปันลิงก์เดต้าผ่านเว็บที่สามารถอ่านและทำความเข้าใจด้วยคอมพิวเตอร์ได้ Tim Berners-Lee เป็นผู้ประดิษฐ์ World Wide Web และผู้สร้าง เว็บเชิงความหมาย และ Linked Data ขึ้นมาในปี 2006 เขาได้แนะนำหลักการออกแบบ 4 ประการ ได้แก่ 1) ใช้ URI เป็นชื่อสำหรับทุกสิ่ง 2) ใช้ HTTP URI เพื่อให้ผู้คนสามารถค้นหาได้ เมื่อมีคนค้นหา URI 3) ให้ข้อมูลที่เป็นประโยชน์โดยใช้มาตรฐาน (RDF, SPARQL) และ 4) ทำการรวมลิงก์ไปยัง URI อื่น ๆ เพื่อให้ค้นพบได้มากขึ้น การใช้ลิงก์เดต้าที่เพิ่มขึ้นอย่างรวดเร็วทำให้เว็บเติบโตซึ่งมีตัวเลขทางสถิติระดับโลกที่ได้รับการยืนยันว่ามีเอกสารและข้อความที่เชื่อมโยงกันเป็นพันล้านครั้ง ความก้าวหน้าของเว็บช่วยให้สามารถตรวจสอบความสัมพันธ์ใหม่ระหว่างบันทึกและการปรับปรุงที่ตามมาสำหรับแอปพลิเคชันใหม่ (Chansanam, Tuamsuk & Supnithi, 2020)

ข้อมูล Linked Open Data – LOD เป็นการผสมผสานระหว่าง data ที่เชื่อมโยงเข้ากับข้อมูลเปิด ดังนั้นจึงอาจกล่าวได้ว่า LOD เป็นชุดลิงก์เดต้าที่สร้างขึ้นจากแหล่งข้อมูลเปิด เพื่อให้ได้แหล่ง LOD ที่มีคุณภาพจะต้องมีการลงทุนอย่างเป็นระบบและความสามารถเข้าถึงเทคโนโลยีได้อย่างมีประสิทธิภาพ โดยเฉพาะอย่างยิ่ง

ยิ่งการเข้าถึงแบบเปิด การพัฒนาอย่างรวดเร็วของเทคโนโลยีเว็บ โดยเฉพาะเทคโนโลยีเว็บเชิงความหมาย ได้เปลี่ยนวิธีการจัดเก็บและเผยแพร่ข้อมูล โดยในกระบวนการนั้นมี LOD ซึ่งจะสร้างการเชื่อมต่อข้อมูลที่มีโครงสร้างบนเว็บเชิงความหมายซึ่งจะช่วยส่งเสริมกระบวนการเผยแพร่ข้อมูลอย่างรวดเร็วและมีประสิทธิภาพ LOD มีประโยชน์มากมายและนำไปใช้ได้จริงในทางปฏิบัติ ตัวอย่างทั่วไปของ LOD คือฐานข้อมูล Dbpedia นี่เป็นโครงการที่มีผู้คนรวมตัวกันอย่างหนาแน่นเพื่อดึงข้อมูลที่มีโครงสร้างจาก Wikipedia และทำให้ข้อมูลนี้ปรากฏบนเว็บเพื่อให้พร้อมใช้งานสำหรับแอปพลิเคชันอื่น ๆ LOD ใน Dbpedia มีความหลากหลายมากและมีโอกาสพัฒนามากขึ้นในอนาคต

ในเวียดนาม งานวิจัยเกี่ยวกับลิงค์เดต้าและลิงค์เดต้าแบบเปิดยังไม่ได้ได้รับความนิยมอย่างเท่าที่ควร แต่ก็ยังพอมองงานวิจัยที่เกี่ยวข้องกับประเด็นเหล่านี้อยู่บ้าง เช่น วิจัยค้นหาแบบ Semantic โดยใช้ ออนโทโลยี และแอปพลิเคชันเพื่อสร้างระบบ look up และค้นหาตัวอย่างข้อความของผู้ป่วย (ในด้านการแพทย์) (Nguyen Hong Son, 2020) เชื่อมโยงแอปพลิเคชันข้อมูลแบบเปิดเพื่อปรับปรุงคุณภาพของทรัพยากรการศึกษาแบบเปิด (สาขาการศึกษา) (Nguyen Danh Minh Tri, 2020) ฯลฯ ซึ่งในงานวิจัยเหล่านั้นมีงานวิจัยเกี่ยวกับเว็บเชิงความหมายเพื่อสร้างระบบการดึงข้อมูลเกี่ยวกับวัฒนธรรมเขมรได้ โดยผู้แต่ง Nguyen Thi Kim Cuong, (2013) เป็นการศึกษาเดียวที่ใช้ออนโทโลยีเพื่อสร้างระบบค้นหาชาติพันธุ์อยู่ในเวียดนาม และมีวัตถุประสงค์เพื่อสร้างมาตรฐานฐานข้อมูลวัฒนธรรมเขมรในเวียดนาม และประยุกต์ใช้เทคโนโลยีเว็บเชิงความหมาย โดยการสร้างระบบสืบค้นข้อมูลเกี่ยวกับวัฒนธรรมเขมรในเวียดนามเพื่อสนับสนุนการค้นหา ตอบสนองความต้องการของการค้นหาที่ถูกต้อง ให้ความรู้ที่จำเป็นเกี่ยวกับวัฒนธรรมเขมรในเวียดนาม มีส่วนช่วยในการดำเนินการอนุรักษ์และส่งเสริมเอกลักษณ์ทางวัฒนธรรมของชาติได้เป็นอย่างดี โดยใช้ซอฟต์แวร์ Protégé เพื่อสร้างออนโทโลยีวัฒนธรรมเขมรภาคใต้ของเวียดนาม และระบบค้นหาวัฒนธรรมเขมรอยู่ในเวียดนาม การใช้งานแอปพลิเคชันได้ผลลัพธ์เชิงบวกในขั้นต้น ระบบสามารถแก้ไขปัญหาพื้นฐาน ตอบสนองความต้องการในการค้นหาอย่างรวดเร็วและแม่นยำ ให้ข้อมูลที่จำเป็นเกี่ยวกับวัฒนธรรมเขมรในเวียดนามอย่างมีประสิทธิภาพ อย่างไรก็ตาม งานวิจัยข้างต้นไม่ได้แสดงภาพรวมของการจัดการข้อมูลของกลุ่มชาติพันธุ์เวียดนามบนอินเทอร์เน็ต แต่เน้นที่การใช้ประโยชน์จากข้อมูลที่เกี่ยวข้องกับวัฒนธรรมเขมรในเวียดนามเท่านั้น ในขณะเดียวกัน จะเห็นได้ว่าเทคนิคการเชื่อมโยงข้อมูลแบบเปิดยังคงเป็นเทคนิคใหม่และยังไม่ได้ดำเนินการในเวียดนาม ดังนั้นจนถึงขณะนี้ยังไม่มีงานใดที่ใช้เทคนิคการเชื่อมโยงข้อมูลแบบเปิดเพื่อสร้างลิงค์เดต้าโดยทั่วไป และข้อมูลเกี่ยวกับชาติพันธุ์เวียดนามโดยเฉพาะ บทความนี้จะนำเสนอการจัดการระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามและการเชื่อมโยงแบบเปิดโดยใช้เทคนิคการเชื่อมโยงข้อมูลแบบเปิด

2. วัตถุประสงค์

- 1) เพื่อเป็นการจัดชุดข้อมูลระบบองค์ความรู้ของกลุ่มชาติพันธุ์ในประเทศเวียดนามทั้ง 54 กลุ่มชาติพันธุ์
- 2) เพื่อทำการเชื่อมโยงข้อมูลกลุ่มชาติพันธุ์ในประเทศเวียดนามกับแหล่งข้อมูลมาตรฐานสากลภายนอกอื่น ๆ

3. วิธีการศึกษา

ดำเนินการโดยใช้เทคนิคการเชื่อมโยงข้อมูลแบบเปิด (Ontology และ ระบบ Reconciliation linking data by OpenRefine and RDF extension) เป็นเครื่องมือเพื่อจัดระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามและการเชื่อมโยงแบบเปิด โดยใช้แนวคิดของ Chansanam, Tuamsuk, Chaikham-bung & Sugimoto, (2020) และมีทั้งหมด 5 ขั้นตอน ได้แก่ 1) การสร้างออนโทโลยี (Ontology) 2) การจัดการข้อมูล (Unify and Cleaning) 3) การเชื่อมโยงกับข้อมูลแหล่งอื่น ๆ (Linking to other data) 4) การจัดการข้อมูล RDF และ 5) การนำเสนอการเชื่อมโยงข้อมูลแบบเปิดผ่านเว็บเบราว์เซอร์

3.1 การสร้างออนโทโลยี (Ontology)

วิธีการจะดำเนินการสองขั้นตอนพื้นฐาน เช่น การสร้างโครงสร้างคลาสแบบลำดับชั้นและการกำหนดคุณสมบัติสำหรับคลาส อันที่จริงการพัฒนา ออนโทโลยี เพื่ออธิบายประเด็นที่น่าสนใจไม่ใช่เรื่องง่าย ขึ้นอยู่กับเครื่องมือที่ใช้ ธรรมชาติ ขนาด ความถี่ของการเปลี่ยนแปลงโดเมน ตลอดจนความสัมพันธ์ที่ซับซ้อน ในนั้น ปัญหาเหล่านี้ต้องการให้การสร้าง ออนโทโลยี เป็นกระบวนการวนซ้ำ การทำซ้ำแต่ละครั้งจะปรับปรุง ปรับแต่ง และพัฒนาผลิตภัณฑ์มากกว่ากระบวนการกรอบงานที่มีขั้นตอนแยกกัน การสร้าง ออนโทโลยี ยังต้องคำนึงถึงความเป็นไปได้ของการขยายสาขาที่น่าสนใจในอนาคต ความสามารถในการสืบทอดระบบ ออนโทโลยี ที่มีอยู่ รวมไปถึงความยืดหยุ่นในการทำให้ ออนโทโลยี สามารถอธิบายความสัมพันธ์ที่ซับซ้อนในโลกแห่งความเป็นจริงได้ดีที่สุด ดังนั้น ในขั้นตอนนี้ ผู้เขียนจึงใช้วิธีสร้างออนโทโลยีชาติพันธุ์ในประเทศเวียดนาม (Ontology Building) และ ส่งออกข้อมูล (Export) เป็นไฟล์เว็บ_ontology (Ontology Web Language: owl) เพื่อสร้างออนโทโลยีเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนาม

3.2 การจัดการข้อมูล (Unify and cleaning)

ในขั้นตอนนี้ ผู้ศึกษาจัดเตรียมชุดข้อมูลจากไฟล์ตารางทำงาน (ethnic_vietnam.xlsx) โดยใช้วิธีค้นคว้าข้อมูลจากแหล่งข้อมูลต่าง ๆ และจัดการข้อมูลตามกลุ่มภาษาชาติพันธุ์อยู่ในเวียดนาม การรวบรวมข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามเป็นกิจกรรมที่สำคัญและมีจุดมุ่งหมาย กระบวนการเก็บรวบรวมข้อมูลต้องตอบคำถามต่อไปนี้โดยเฉพาะ ข้อมูลนี้เก็บรวบรวมเพื่ออะไร ทำหน้าที่อะไร และประเด็นใดที่เกี่ยวข้องกับปัญหา ฯลฯ

ethnic_vietnam - Excel										
F11										
Cul, Cur, Cu, Thổ, Việt gốc Miên, Khơ Me, Króm										
Languages/Family										
id	Language	id	self name	meaning	ethnonyms	official name english	official name thai	ical name viet	Origin	speaking
1	Vietnamese/Muong Proto-Vietic	Kinh	Kinh	người Việt cổ	Việt (Kinh)	Kinh	Kinh	Vietname Việt Nam	Việt	1
2	Tay Thai	Tay	Tay	sinh sống chủ yếu ở vùng miền núi thấp	Thổ, Ngai, Phien, Thù Lao, Tày	Tay	Tay	China Trung Quốc	Tiếng Việt	2
3	Hmong Dao	Thai	Thai	Tay Khao hoặc Đon (Thai Tr) Thai	Tay Khao hoặc Đon (Thai Tr) Thai	Thai	Thai	China Trung Quốc	Tiếng Việt	3
4	Ka Dai	Muong	Muong	là dân tộc sống ở khu vực trung du và	Mai (Mau), Mon, Moan,	Muong	Muong	Vietname Việt Nam	Tiếng Việt	4
5	Tibetan-Burmese	Khmer	Khmer	là dân tộc sống ở khu vực trung du và	Mai (Mau), Mon, Moan,	Khmer	Khmer	Vietname Việt Nam	Tiếng Việt	5
6	Mon Khmer	Hoa	Hoa	là người Việt gốc Hoa	Triều Châu, Phúc Kiến, Quai chiness	Hoa	Hoa	China Trung Quốc	Tiếng Việt	6
7	Austroasiatic	Nung	Nung	là người Việt gốc Hoa	Nung, Nung, Nung, Nung	Nung	Nung	China Trung Quốc	Tiếng Việt	7
8	Sinitic	Hmong	Hmong	là người Việt gốc Hoa	Hmong, Hmong, Hmong, Hmong	Hmong	Hmong	China Trung Quốc	Tiếng Việt	8
9		Dao	Dao	là người Việt gốc Hoa	Mai, Dao, Tai, Xi, Diu, M, Dao	Dao	Dao	China Trung Quốc	Tiếng Việt	9
10		Gia Rai	Gia Rai	là người Việt gốc Hoa	Jorai, Jorai, có các nhóm p) Gia Rai	Gia Rai	Gia Rai	Vietname Việt Nam	Tiếng Việt	10
11		Ê đê	Ê đê	là một dân tộc có vùng cư trú truyền thống	Ê đê, Ê đê, Ê đê, Ê đê, Ê đê	Ê đê	Ê đê	Hải đảo Thái Bình Dương	Tiếng Việt	11
12		Ba Na	Ba Na	là một dân tộc có vùng cư trú truyền thống	Ba Na, Ba Na, Ba Na, Ba Na	Ba Na	Ba Na	Indonesia	Tiếng Việt	12
13		San Chay	San Chay	là một dân tộc có vùng cư trú truyền thống	San Chay, San Chay, San Chay, San Chay	San Chay	San Chay	China Trung Quốc	Tiếng Việt	13
14		Cham	Cham	là một dân tộc có vùng cư trú truyền thống	Cham, Cham, Cham, Cham	Cham	Cham	Đông Nam Á	Tiếng Việt	14
15		Cơ Ho	Cơ Ho	là một dân tộc có vùng cư trú truyền thống	Xơ, Nổ (Tu Lốp), Cơ Ho, Cơ Ho	Cơ Ho	Cơ Ho	Vietname Việt Nam	Tiếng Việt	15
16		Xơ Đăng	Xơ Đăng	là một dân tộc có vùng cư trú truyền thống	Xơ Đăng, Xơ Đăng, Xơ Đăng, Xơ Đăng	Xơ Đăng	Xơ Đăng	Vietname Việt Nam	Tiếng Việt	16
17		Sán Dư	Sán Dư	là một dân tộc có vùng cư trú truyền thống	Sán Dư, Sán Dư, Sán Dư, Sán Dư	Sán Dư	Sán Dư	China Trung Quốc	Tiếng Việt	17
18		Hre	Hre	là một dân tộc có vùng cư trú truyền thống	Cham, Rê, Mui, Chom, Kie, L Hre	Hre	Hre	Vietname Việt Nam	Tiếng Việt	18
19		Rai	Rai	là một dân tộc có vùng cư trú truyền thống	Rai, Rai, Rai, Rai, Rai, Rai	Rai	Rai	China Trung Quốc	Tiếng Việt	19
20		Mnong	Mnong	là một dân tộc có vùng cư trú truyền thống	Phong, Mông, Mông, Mông	Mnong	Mnong	Vietname Việt Nam	Tiếng Việt	20
21		Thổ	Thổ	là một dân tộc có vùng cư trú truyền thống	Thổ, Thổ, Thổ, Thổ, Thổ, Thổ	Thổ	Thổ	Vietname Việt Nam	Tiếng Việt	21
22		Stieng	Stieng	là một dân tộc có vùng cư trú truyền thống	Sa, Stieng, Sa, Stieng, Sa, Stieng	Stieng	Stieng	Vietname Việt Nam	Tiếng Việt	22
23		Khmu	Khmu	là một dân tộc có vùng cư trú truyền thống	Xa, Khmu, Khmu, Khmu, Khmu, Khmu	Khmu	Khmu	Lào	Tiếng Việt	23
24		Bru	Bru	là một dân tộc có vùng cư trú truyền thống	Mang, Coong, Tri, Bru, Bru, Bru	Bru	Bru	Lào	Tiếng Việt	24
25		Co Tu	Co Tu	là một dân tộc có vùng cư trú truyền thống	Ca, Tu, Cao, Ha, Phom, Ca, Co Tu	Co Tu	Co Tu	Vietname Việt Nam	Tiếng Việt	25
26		Giay	Giay	là một dân tộc có vùng cư trú truyền thống	Nhông, Dông, Pù, Giay, Pù, Giay	Giay	Giay	China Trung Quốc	Tiếng Việt	26
27		Ta Oi	Ta Oi	là một dân tộc có vùng cư trú truyền thống	Ta Oi, Pa Oi, Pa Oi, Pa Oi, Pa Oi	Ta Oi	Ta Oi	Vietname Việt Nam	Tiếng Việt	27
28		Ma	Ma	là một dân tộc có vùng cư trú truyền thống	Chau Ma, Chô Ma, Chô Ma, Chô Ma	Ma	Ma	Vietname Việt Nam	Tiếng Việt	28

ภาพที่ 1 การเตรียมชุดข้อมูลกลุ่มภาษาชาติพันธุ์ในเวียดนาม

จากภาพที่ 1 เราจะเห็นได้ว่า ผู้ศึกษาเตรียมชุดข้อมูลอย่างละเอียดของ 54 กลุ่มชาติพันธุ์ในเวียดนาม และจำแนกเป็นแต่ละประเด็น เช่น กลุ่มภาษา ชื่อทางการ ชื่อคนอื่นเรียก ฯลฯ ข้อมูลนี้ถือว่าเป็นข้อมูลเบื้องต้นและพื้นฐานของขั้นตอนต่อไป

หลังจากการเตรียมชุดข้อมูลดังกล่าว ผู้ศึกษาได้นำข้อมูลนี้เข้า Google Open Refine เพื่อดำเนินต่อไป ในการจัดการทำความสะอาดข้อมูลและจัดระบบข้อมูลดังกล่าวโดยเชื่อมโยงกันกับข้อมูลแหล่งอื่น ๆ ทั้งภายในและภายนอก กำหนดการยกเว้นจำนวนแถวที่ไม่ต้องการ ทำให้บรรทัดที่ไม่มีข้อมูลหรือว่างให้ตัดออก ลบคอลัมน์ที่ไม่จำเป็นและไม่ได้ใช้งานออกไป ฯลฯ งานนี้เรียกว่าการล้างข้อมูลโดยไม่จำเป็น นี่เป็นขั้นตอนบังคับเบื้องต้นในการประมวลผลหรือลบข้อมูลที่ซ้ำกัน ผิดรูปแบบ ไม่ถูกต้องหรือไม่สมบูรณ์ จุดประสงค์ของการล้างข้อมูลไม่ใช่เพียงเพื่อเพิ่มพื้นที่จัดเก็บเท่านั้น แต่ยังเพิ่มความแม่นยำของชุดข้อมูลก่อนทำการวิเคราะห์และวิจัยอีกด้วย

3.3 การเชื่อมโยงกับข้อมูลแหล่งอื่น ๆ (Linking to other data)

การจัดการระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามและการเชื่อมโยงแบบเปิดผู้ศึกษาจึงใช้วิธีการกระหายอดเชื่อมโยงข้อมูลโดย OpenRefine และส่วนขยาย RDF

OpenRefine ช่วยให้ผู้ใช้ศึกษาสำรวจชุดข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามได้อย่างง่ายดาย ใช้เพื่อเชื่อมโยงและขยายชุดข้อมูลของผู้ศึกษาด้วยบริการเว็บต่าง ๆ ดาวนโหนดข้อมูลที่ทำความสะอาดของกลุ่มชาติพันธุ์ในเวียดนามไปยังฐานข้อมูลกลาง เช่น Wikidata OpenRefine ทำงานร่วมกับไฟล์เครื่องหรือ

ข้อมูลจากที่อยู่เว็บในรูปแบบไฟล์ต่าง ๆ รวมถึง CSV TSV XLS XML และอื่น ๆ ความสามารถในการค้นหาข้อมูลที่ซ้ำกัน ข้อมูลที่ไม่สอดคล้องกัน และรูปแบบข้อผิดพลาดสำหรับการแก้ไขและล้างข้อมูลจำนวนมาก OpenRefine มีความสามารถในการกรองหรือค้นหาข้อมูลบางอย่างเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามที่จำเป็นต้องเปลี่ยนแปลง ซึ่งจะจำกัดมุมมองเฉพาะเซลล์ แถว หรือคอลัมน์ที่เกี่ยวข้องที่มีข้อมูลเท่านั้น

The screenshot shows the OpenRefine web interface. At the top, there's a browser address bar with the URL '127.0.0.1:3333/project?project=1730609993618'. Below it, a toolbar contains icons for various applications like Facebook, Twitter, and others. The main interface has a sidebar on the left with 'Using facets and filters' and a main table area. The table has columns: 'Name', 'Language', 'Autonym', 'Meaning', 'Exonym', 'Ethnonym_english', 'Ethnonym_Thai', 'Ethnonym_vietnam', 'Origin', 'Spreading', 'Ethnicity', and 'Lang'. The table lists 14 records of ethnic groups in Vietnam, including Kinh, Tay, Hmong, and others, with their respective languages and meanings.

ภาพที่ 2 การกระหายยอดเชื่อมโยงข้อมูลโดย OpenRefine และส่วนขยาย RDF (Reconciliation linking data by OpenRefine and RDF extension)

DBpedia และ WikiData เป็นฐานข้อมูลที่รวมเนื้อหาที่ปรับมาจากข้อมูล Wikipedia ในรูปแบบ RDF ทำให้การเชื่อมโยงกับ WikiData และ DBpedia มีบทบาทที่สำคัญอย่างยิ่ง การสร้างลิงก์เป็นเรื่องง่าย ทำให้ผู้ใช้สามารถเข้าไปยังข้อมูลที่เกี่ยวข้องกับบทความที่กำลังอ่านได้ทันที ช่วยเพิ่มประโยชน์ของ Wikipedia DBpedia ในกระบวนการนี้ ผู้ศึกษาต้องทำผ่านหลายขั้นตอน อย่างเช่น ทำการ Reconciling ข้อมูลด้วย RDF Extension การตรวจสอบความถูกต้องของข้อมูลและปรับแต่งข้อมูลที่มีการเชื่อมโยงให้ถูกต้อง ยืนยันข้อมูลการเชื่อมโยงที่ถูกต้องและคะแนนค่าความคล้ายคลึงสูง ฯลฯ ผลการเชื่อมโยงข้อมูลเกี่ยวกับชาติพันธุ์เวียดนามกับข้อมูลแหล่งอื่นได้แสดงดังภาพด้านล่าง

3.4 การจัดการข้อมูล RDF

Converting to Resource Description Framework (RDF) เป็นเครื่องมือกำหนดมาตรฐานสำหรับภาษาที่ใช้สืบค้น SPARQL เป็นมาตรฐานสำหรับการแลกเปลี่ยนข้อมูลบนเว็บที่เผยแพร่โดย World Wide Web (WWW) RDF เป็นชุดของหลักการสำหรับภาษามาร์กอัพที่จัดเตรียมแบบจำลองข้อมูลและไวยากรณ์อย่างง่ายที่สามารถแลกเปลี่ยนและใช้งานโดยระบบอิสระ ในขณะเดียวกันก็ได้รับการออกแบบเพื่อให้ระบบคอมพิวเตอร์สามารถเข้าใจและอ่านข้อมูลได้ ไม่ใช่แค่การนำเสนอข้อมูลให้ผู้ใช้ทราบไวยากรณ์

ของ RDF ที่สร้างขึ้นบนตัวแบบข้อมูล และแบบจำลองนี้ส่งผลต่อวิธีการอธิบายคุณสมบัติ และยังทำให้โครงสร้างของคำอธิบายชัดเจนอีกด้วย

ในการจัดการข้อมูล RDF ของกลุ่มชาติพันธุ์อยู่ที่เวียดนาม ผู้ศึกษาได้กำหนดโครงร่างของ RDF (RDF skeleton) และแปลงไฟล์เป็น TTL (Turtle, the Terse RDF Triple Language, Turtle is an extension of NTriples ([N-TRIPLES]))

3.5 การนำเสนอ Linked Open Data ผ่านเว็บเบราว์เซอร์

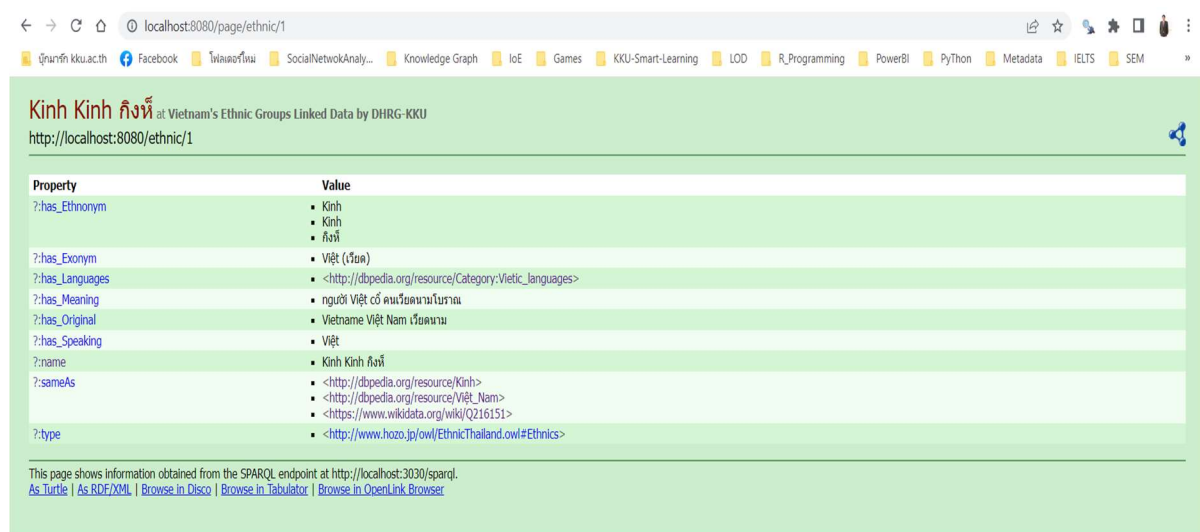
Linked Open Data (LOD) เป็นการผสมผสานระหว่าง data ที่เชื่อมโยงกัน และเปิดข้อมูล ดังนั้นจึงอาจกล่าวได้ว่า LOD เป็นชุดข้อมูลที่เชื่อมโยงเข้าด้วยกันและสร้างจากแหล่งข้อมูลแบบเปิด เพื่อให้ได้แหล่ง LOD ที่มีคุณภาพจะต้องลงทุนการเข้าถึงอย่างเป็นระบบและมีประสิทธิภาพในด้านเทคโนโลยี โดยเฉพาะอย่างยิ่งการเข้าถึงแบบเปิด การพัฒนานโยบายการเข้าถึงแบบเปิดเป็นเงื่อนไขที่จำเป็นสำหรับการพัฒนาทรัพยากร OER ที่มีคุณภาพ การพัฒนาอย่างรวดเร็วของเทคโนโลยีเว็บ โดยเฉพาะเทคโนโลยี Semantic Web (SW) ได้เปลี่ยนแปลงวิธีการจัดเก็บและเผยแพร่ข้อมูล รวมถึง LOD และ LOD จะสร้างการเชื่อมต่อข้อมูลที่มีโครงสร้างบน SW ซึ่งจะช่วยสนับสนุนกระบวนการเผยแพร่ข้อมูลอย่างรวดเร็วและมีประสิทธิภาพ

ในขั้นตอนนี้ ผู้ศึกษาทำการเก็บหรือคัดลอกไฟล์ Ethnic_vietnam.ttl จากขั้นตอนที่ 4 ไปยังโฟลเดอร์ DE:\LOD_GMS\jena-fuseki-0.2.6-SNAPSHOT\pages ไปที่ command prompt ของ windows ทดสอบการ Query ข้อมูลด้วย SPARQL Endpoint ไปที่โฟลเดอร์ D:\LOD_GMS\jetty-distribution-9.0.0.M3\webapps\ROOT\WEB-INF ทดสอบการทำงานบนเว็บเบราว์เซอร์ของ LOD กลุ่มชาติพันธุ์ในประเทศไทย โดยไปที่ <http://localhost:8080/page/ethnic/1>

หลังจากใช้ 5 ขั้นตอนข้างต้นของ LOD สำหรับองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์อยู่ในเวียดนาม รวมกับเว็บ LOD ที่จัดตั้งขึ้น อนุญาตให้ชุดข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามอ้างอิงถึงกันผ่านเงื่อนไขที่เข้าร่วมกันที่กำหนดไว้ในรูปแบบของพจนานุกรมข้อมูล นี่เป็นคุณลักษณะที่สำคัญที่สุดเนื่องจากช่วยให้ข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์อยู่ในเวียดนามที่สร้างขึ้นในองค์กรหนึ่ง สามารถอ้างอิงถึงข้อมูลที่สร้างโดยองค์กรอื่นได้ (หมายความว่า พื้นที่ของข้อมูลจะไม่จำกัดภายในองค์กร) และถือว่าเป็นรากฐานที่มั่นคงในการปรับปรุงคุณภาพของข้อมูลกลุ่มชาติพันธุ์อยู่ในเวียดนาม ปัญหาที่เหลืออยู่ในกลยุทธ์ในการสร้างและรักษาชุมชนของผู้ร่วมให้ข้อมูลเพื่อเพิ่มประสิทธิภาพวิธีการรวบรวมและจัดการข้อมูลที่เปลี่ยนแปลงหรือเกิดขึ้นใหม่ กระบวนการนี้ใช้เวลานานและถูกต้องตามเป้าหมายการพัฒนาของแหล่งข้อมูลแต่ละแหล่งเกี่ยวกับกลุ่มชาติพันธุ์ต่าง ๆ ในเวียดนาม กระแสการใช้ LOD จะเป็นที่ยินยมนมากขึ้นเรื่อย ๆ ไม่เพียงแต่ในประเทศที่พัฒนาแล้วเท่านั้น แต่จะแพร่กระจายไปยังประเทศอื่น ๆ ทั่วโลกอีกด้วย แหล่งที่มาของ LOD จะมีส่วนสนับสนุนที่สำคัญในการสร้างแหล่งข้อมูลที่มีคุณภาพมากมายเกี่ยวกับกลุ่มชาติพันธุ์ของเวียดนาม เพื่อให้บริการตามวัตถุประสงค์ในการค้นหาข้อมูลของบุคคลและองค์กรที่สนใจในกลุ่มชาติพันธุ์ในเวียดนาม

4. ผลการศึกษา

ผู้ศึกษาได้สร้างโปรโตคอล SPARQL และภาษาแบบสอบถาม RDF เกี่ยวกับกลุ่มชาติพันธุ์เวียดนามบนอินเทอร์เน็ต SPARQL ถูกสร้างขึ้นเป็นมาตรฐานสำหรับการเข้าถึงข้อมูล RDF ที่ทำงานในกลุ่ม (DAWG) ของ World Wide Web Consortium และถือเป็นหนึ่งในเทคโนโลยีหลักของความหมายของเว็บ โปรโตคอล SPARQL HTTP อนุญาตการเข้าถึงแบบสาธารณะ การเข้าถึงแบบเปิดนี้อนุญาตให้ทุกคนสอบถามที่เก็บข้อมูลโดยใช้ภาษา SPARQL แต่ควรสังเกตว่าการเข้าถึงนี้สามารถจำกัดได้โดยใช้กลไกและการกำหนดค่าความปลอดภัย HTTP มาตรฐาน รูปแบบเฉพาะของพื้นที่เก็บข้อมูล ไฟล์ที่มีกลุ่มชาติพันธุ์ในข้อมูล RDF ของเวียดนามก็มีให้ออนไลน์เช่นกัน ตัวอย่าง SPARQL HTTP Protocol ของกลุ่มชาติพันธุ์ในเวียดนามแสดงอยู่ในภาพที่ 3



ภาพที่ 3 ตัวอย่างของหน้าข้อมูลเปิดที่เชื่อมโยงแบบกำหนดเองของเวียดนาม

RDF แบบกำหนดเองของเวียดนามทั้งหมดที่เขียนด้วย Turtle ต้องใช้ภายในส่วนภาษาแบบสอบถามของโปรโตคอล SPARQL และ RDF Query Language ซึ่งใช้ไวยากรณ์แฟชัน Turtle/N3 สำหรับรูปแบบ Triple และสำหรับ RDF triples ในอนุประโยค CONSTRUCT สิ่งนี้อำนวยความสะดวกในการใช้ RDF ที่เขียนด้วย Turtle เพื่อเปิดใช้งานการสร้าง "แบบสอบถามโดยใช้ตัวอย่าง" โดยใช้ข้อมูลเพื่อสร้างคำถามเบื้องต้น ซึ่งสามารถแก้ไขได้เพื่อใช้ตัวแปรที่ต้องการการเชื่อมโยง ดูรูปภาพที่ 4



[dbpedia.org/page/Vietnamese_people](#)

[qinrnf kku.ac.th](#)
[Facebook](#)
[Tinauifnu](#)
[SocialNetworkAnaly...](#)
[Knowledge Graph](#)
[IOE](#)
[Games](#)
[KKU-Smart-Learning](#)
[LOD](#)
[R_Programming](#)
[PowerBI](#)
[PyThon](#)
[Metadata](#)
[IELTS](#)
[SEM](#)

[Browse using ▾](#)
[Formats ▾](#)

[Faceted Browser](#)
[Sparql Endpoint](#)

About: [Vietnamese people](#)

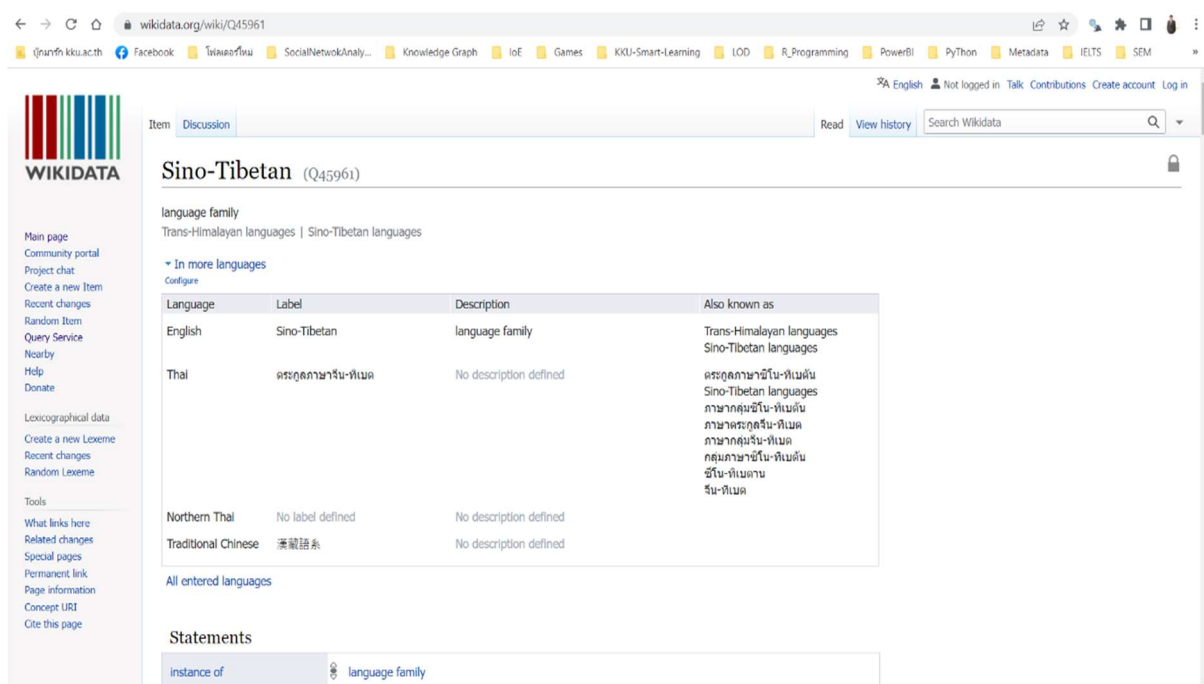
An Entity of Type: [ethnic group](#), from Named Graph: [http://dbpedia.org](#), within Data Space: [dbpedia.org](#)

The Vietnamese people (Vietnamese: người Việt) or Kinh people (Vietnamese: người Kinh) are a Southeast Asian ethnic group originally native to modern-day Northern Vietnam and Southern China. The native language is Vietnamese, the most widely spoken Austroasiatic language. Its vocabulary was influenced by Chinese early on. During the French colonial era, French was an official language in Vietnam. Afterwards, the Vietnamese language codified in the Latin alphabet emerged.



Property	Value
dbpedia:abstract	The Vietnamese people (Vietnamese: người Việt) or Kinh people (Vietnamese: người Kinh) are a Southeast Asian ethnic group originally native to modern-day Northern Vietnam and Southern China. The native language is Vietnamese, the most widely spoken Austroasiatic language. Its vocabulary was influenced by Chinese early on. During the French colonial era, French was an official language in Vietnam. Afterwards, the Vietnamese language codified in the Latin alphabet emerged. Vietnamese Kinh people account for just over 85.32% of the population of Vietnam in the 2019 census, and are officially known as Kinh people (người Kinh) to distinguish them from the other minority groups residing in the country such as the Hmong, Cham or Muong. The earliest recorded name for the ancient Kinh people in Vietnamese history books is Lạc or Lạc Việt. The Vietnamese are one of the four main groups of Vietic speakers in Vietnam, the others being the Muong, Thô and Chứt people. They are related to the Gin or the Jing people, a Vietnamese ethnic group in China. (en)
dbpedia:language	dbp:Vietnamese language

ภาพที่ 5 DBpedia เชื่อมโยงข้อมูล (DBpedia linked data)



ภาพที่ 6 Wikidata เชื่อมโยงข้อมูล (Wikidata linked data)

จากการจัดระบบองค์ความรู้เกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนาม เราจะเห็นได้ว่า ในเวียดนามมีกลุ่มชาติพันธุ์ที่หลากหลาย กลุ่มชาติพันธุ์ในเวียดนามไม่ใช่มีจุดกำเนิดอยู่ในเวียดนามเท่านั้น แต่ยังมีแหล่งกำเนิดจากประเทศอื่น ๆ เช่น จีน ลาว ไทย เขมร ฯลฯ แต่ละกลุ่มชาติพันธุ์จะมีภาษาพูดที่แตกต่างกันแต่ยังใช้ภาษาเวียดนาม (ภาษา Kinh) เป็นภาษากลางในการสื่อสาร การสร้างลิงก์เดต้าที่เชื่อมต่อ Link อื่น ๆ มีส่วนทำให้เกิดความหลากหลายข้อมูลเกี่ยวกับชาติพันธุ์อยู่ในเวียดนาม และมีส่วนช่วยสำหรับคนอื่นในการค้นคว้าข้อมูลได้ง่ายและครบถ้วนมากยิ่งขึ้น

อย่างไรก็ตาม ข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์อยู่ในเวียดนามยังมีข้อจำกัดบางอย่าง เช่น จากลิงก์เดต้าเชื่อมต่อ Link อื่น ๆ ข้อมูลของแต่ละกลุ่มชาติพันธุ์ยังขาดเป็นจำนวนมาก รวมทั้งข้อมูลแบบพื้นฐาน (ชื่อ ภาษาอังกฤษ ภาษาพูด ฯลฯ) การแบ่งกลุ่มภาษาของเวียดนามยังไม่เป็นสากลจึงเป็นอุปสรรคในการค้นคว้าข้อมูล ดังนั้น ในอนาคตข้างหน้า คณะกรรมการกิจการชาติพันธุ์ของเวียดนามควรเป็นประธานและประสานงานกับหน่วยงานอื่น ๆ เช่น กระทรวงวิทยาศาสตร์และเทคโนโลยี กระทรวงสารสนเทศและการสื่อสาร ฯลฯ ในการเผยแพร่ข้อมูลที่เกี่ยวข้องกับกลุ่มชาติพันธุ์ของเวียดนามบนอินเทอร์เน็ตอย่างครบถ้วนและถูกต้องมากขึ้น ตอบสนองความต้องการของการเปลี่ยนแปลงด้านดิจิทัลของหน่วยงานภาครัฐ

5. สรุปผลการศึกษา

แอปพลิเคชัน LOD เป็นหนึ่งในวิธีการที่มีประสิทธิภาพในการปรับปรุงคุณภาพของแหล่งข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนาม LOD สร้างขึ้นบนแหล่งข้อมูลเปิดที่เชื่อมโยงกันในเครือข่ายของเว็บเชิงความหมาย ซึ่งมีการขยายและปรับปรุงอย่างต่อเนื่องเมื่อเวลาผ่านไป การนำชุดข้อมูลระดับ 5 ดาวไปใช้เพื่อสร้างการ

เชื่อมโยงข้อมูลแบบเปิดนั้นมีความสำคัญเป็นพิเศษโดยเฉพาะอย่างยิ่งในบริบทที่แหล่งเก็บข้อมูลแบบเปิดและแหล่งที่มายังอยู่ในสถานะที่กระจายกระจายอย่างมากและจำเป็นต้องปรับโฉมใหม่เพื่อให้มีความเกี่ยวข้องมากขึ้น ในการสร้างระบบของความรู้เกี่ยวกับกลุ่มชาติพันธุ์อยู่ในเวียดนาม กระบวนการสร้างและปรับปรุงคุณภาพของแหล่งข้อมูลเกี่ยวกับกลุ่มชาติพันธุ์ในเวียดนามไม่สามารถแยกออกจากกระบวนการสร้างแหล่งข้อมูล LOD ที่มีความหลากหลายมากขึ้นเรื่อย ๆ ดังนั้นเพื่อให้ได้แหล่ง LOD ที่มีคุณภาพจึงต้องใช้เงินลงทุนและความพยายามอย่างมาก ไม่เพียงแต่ในด้านด้านเทคนิค ใบอนุญาต แต่ยังรวมถึงวิสัยทัศน์เชิงกลยุทธ์ระยะยาว บวกกับการเชื่อมต่อจากหน่วยงาน องค์กร

อย่างไรก็ตาม จากการวิจัยนี้เป็นการเพิ่มข้อมูลที่ศึกษาจากงานวิชาการที่เกี่ยวข้องเป็นภาษาไทยและภาษาอังกฤษ และมีการศึกษาจากเอกสารออนไลน์ภาษาเวียดนาม ซึ่งทำให้ข้อมูลอาจจะยังไม่ครบถ้วน เห็นได้จากการวิจัยยังมีข้อมูลของกลุ่มชาติพันธุ์ที่ยังไม่มีในระบบ โดยหากมีการเพิ่มเติมการค้นคว้าเอกสารภาษาลาวเพิ่มเติม หรือศึกษาเพิ่มเติมเกี่ยวกับงานวิจัยหรือระบบฐานข้อมูลออนไลน์ของเวียดนามโดยตรง อาจจะช่วยเพิ่มข้อมูลที่เกี่ยวข้องกับชาติพันธุ์ของเวียดนามได้ หรือหากมีการเทียบเคียงข้อมูลที่เป็นภาษาอื่น ๆ ที่มีความใกล้เคียงกับกลุ่มชาติพันธุ์เวียดนาม เช่น ลาว ไทย หรือจีน อาจจะทำให้มีข้อมูลที่จะสามารถเพิ่มในระบบฐานข้อมูลออนไลน์ได้มากขึ้น

เอกสารอ้างอิง

- Chansanam, W., Tuamsuk, K., & Supnithi, T. (2020). Digital content management of Heet Sib Sorong custom for semantic search. *Periodicals of Engineering and Natural Sciences*, 8(3), 1935–1950.
- Chansanam, W., Tuamsuk, K., Chaikhambung, J., & Sugimoto, S. (2020). Linked open data framework for ethnic groups in Thailand learning. *International Journal of Emerging Technologies in Learning (IJET)*, 15(10), 140–156.
- Noy, N.F. and McGuinness, D. (2001). *Ontology development 101: A guide to creating your first ontology*. Knowledge Systems Laboratory Stanford University.
- Ngoc Khanh. (2019). **How many ethnic groups does Vietnam have today?**. Retrieved from <https://infonet.vietnamnet.vn/chuyen-la/viet-nam-co-bao-nhieu-dan-toc-hien-nay-1950.html>
- Nguyen Thi Kim Cuong. (2013). **The semantic web application to build an information retrieval system about Southern Khmer culture, Master's thesis in engineering**. University of Danang.
- Nguyen Danh Minh Tri. (2020). Linked open data application in advanced quality of open educational resources. *Proceedings of the Library Information and Communication Conference (LIC)*, University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City.
- Nguyen Hong Son. (2020). **Research on methods of semantic search using ontology and application of building a search system to search for patient text samples**, Doctoral thesis in Military Technical Academy, Vietnam.