

doi: 10.14456/jiskku.2021.11

บทเรียนจาก “ปลายทาง”: กรณีศึกษาการจัดการฐานข้อมูลงาน

ศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร

Lessons Learned from the End: The Case Study of SAC's
Research Database at the Princess Maha Chakri Sirindhorn
Anthropology Centre

สิทธิศักดิ์ รุ่งเจริญสุขศรี¹, Wachiraporn Klungthanaboon^{2*}

Sittisak Rungcharoensuksri¹, Wachiraporn Klungthanaboon^{2*}

*Corresponding author email: Wachiraporn.k@chula.ac.th

Received: November 5, 2020

Revised: January 18, 2021

Accepted: February 19, 2021

บทคัดย่อ

บทความนี้ เป็นการถอดบทเรียนการออกแบบและพัฒนาฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร (The SAC's Research Database) โดยแบ่งการนำเสนอออกเป็น 4 ส่วน คือ 1) นำเสนอที่มาและความสำคัญของการจัดการงานศึกษาวิจัยศูนย์ฯ และการพัฒนาฐานข้อมูล 2) ชี้ให้เห็นปัญหาที่พบในระหว่างกระบวนการสำรวจ และจัดทำทะเบียนข้อมูลงานศึกษาวิจัยสำหรับฐานข้อมูล 3) อธิบายเงื่อนไขและแนวคิดการออกแบบฐานข้อมูล และการจัดทำระบบอ้างอิงอัตโนมัติ และ 4) วิเคราะห์และอภิปรายปัญหาที่ประสบ พร้อมทั้งเสนอแนวทางสำหรับการแก้ไขปัญหาสำหรับศูนย์มานุษยวิทยาสิรินธร

¹ สำนักวิชาการและสารสนเทศ ศูนย์มานุษยวิทยาสิรินธร (องค์การมหาชน); Office of Academic Affairs and Information, The Princess Maha Chakri Sirindhorn Anthropology Centre (SAC).

² ภาควิชาบรรณารักษศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, และหน่วยปฏิบัติการวิจัย The Arc of Memory คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย; The Arc of Memory Research Unit and Department of Library Science, Faculty of Arts, Chulalongkorn University.

คำสำคัญ: การดูแลรักษาดิจิทัล, เมทาดาตา, ฐานข้อมูล, คลังสารสนเทศสถาบัน

Abstract

This article describes awareness and experiences of lessons learned from the design and development of the database of Princess Maha Chakri Sirindhorn (SAC) Anthropology Research Centre. The paper is divided into 4 parts. Part one provides the historical background and management of the SAC and its database development. Part 2 points out the problems encountered while conducting survey and registering research documents for its database development. Part 3 describes requirements and concepts for designing the database. Part 4 analyzes and discusses the problems and provides recommendations for solving them.

Keyword: Digital curation, Metadata, Databases, Institutional repositories

บทนำ

ในสังคมวิชาการ การเผยแพร่ข้อมูลและองค์ความรู้ที่ได้จากการค้นคว้าและวิจัย ถือเป็นเป้าหมายที่สถาบันวิชาการและนักวิจัยส่วนใหญ่ให้ความสำคัญ เพราะมุ่งหวังให้เกิดการนำไปใช้ค้นคว้าและอ้างอิง เพื่อพัฒนาต่อยอดผลงานวิจัยขึ้นมาใหม่ ศูนย์มานุษยวิทยาสิรินธร (องค์การมหาชน) ก็มีพันธกิจดังกล่าวด้วยเช่นกัน องค์การแห่งนี้ได้ผลิตผลงานวิชาการเผยแพร่สู่สาธารณชนเป็นเวลาต่อเนื่องเกือบ 30 ปี โดยมีตัวอย่างผลงานที่มีผู้คนนำไปใช้ในการเรียนการสอนทางด้านสังคมศาสตร์และมนุษยศาสตร์ เช่น พิพิธภัณฑสถานแห่งชาติ: ประสบการณ์จากคนลองทำ (2551) มรดกภูมิปัญญาทางวัฒนธรรมของชาติ: บทเรียนจากเพื่อนบ้าน (2560) การจัดการทรัพยากรทางวัฒนธรรม (2554, 2559) เป็นต้น (Princess Maha Chakri Sirindhorn Anthropology, 2019)

นอกเหนือจากผลงานวิชาการที่เผยแพร่สู่สาธารณชนในรูปแบบ หนังสือเล่มแล้ว ยังมีรายงานการศึกษาและเอกสารจากโครงการต่าง ๆ เช่น รายงานวิจัยฉบับสมบูรณ์ “โครงการวิจัยและพัฒนาพิพิธภัณฑสถาน” (2547-2551) ประเพณี 12 เดือน: พิธีกรรมที่เปลี่ยนไป (2546) เอกสารการสัมมนาทางวิชาการ “พิพิธภัณฑสถานในประเทศไทย” 23 - 24 พฤศจิกายน 2539 (2540) เป็นต้น ผลงานเหล่านี้เผยแพร่ในวงจำกัด ทั้ง ๆ ที่มีข้อมูลและองค์ความรู้ที่มีคุณค่า เป็นประโยชน์สำหรับการค้นคว้าและอ้างอิงได้เช่นกัน

ด้วยเหตุนี้ ศูนย์ฯ จึงได้เริ่มโครงการพัฒนาฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร (The SAC's Research Database) <http://www.sac.or.th/databases/sac-research/index.php>) เมื่อ พ.ศ. 2561 โดยมีวัตถุประสงค์เพื่อรวบรวมและจัดเก็บผลงานการศึกษาค้นคว้าวิจัยของศูนย์ฯ ในฐานะทรัพยากรสารสนเทศ

ปัญญาขององค์กร พร้อมทั้งเผยแพร่ผลงานเหล่านี้ในรูปแบบฐานข้อมูลออนไลน์ที่อนุญาตให้ผู้ใช้เข้าถึงและดาวน์โหลดเอกสารฉบับเต็ม (Full text) เพื่อนำไปใช้อ้างอิงพัฒนาต่อยอดงานวิจัยได้

อย่างไรก็ดีในระหว่างการทำเนิงานโครงการนี้ กลับประสบอุปสรรคที่ทำทาย เช่น ความซ้ำซ้อนในกระบวนการจัดเก็บ การรวบรวมงานศึกษาวิจัยทั้งที่เป็นสิ่งพิมพ์และไฟล์ดิจิทัลจากผู้รับผิดชอบ การออกแบบและดัดแปลงชุดเมตาดาตา (Metadata) ให้เหมาะสมกับทรัพยากรและตอบสนองความต้องการของกลุ่มผู้ใช้งาน เป็นต้น รวมถึงคำถามสำคัญที่ชวนให้ขบคิดว่ โครงการนี้อาจเป็นเพียงการแก้ปัญหาที่ “ปลายทาง” เท่านั้นหรือแล้ว “ต้นทาง” ของปัญหาอยู่ที่ใด หากต้นทางของปัญหายังไม่ได้รับการแก้ไข ปัญหาเหล่านี้จะย้อนกลับมาเกิดขึ้นอีกครั้งในอนาคตหรือไม่

บทความนี้จึงเป็นการแบ่งปันประสบการณ์การทำงานจากการศึกษาและพัฒนาฐานข้อมูลงานศึกษาวิจัยศูนย์ฯ ตลอดจนปัญหาอุปสรรค เพื่อเป็นกรณีศึกษาสำหรับการออกแบบวิธีบริหารจัดการและจัดเก็บผลงานวิชาการงานวิจัยที่มีประสิทธิภาพสำหรับประเทศต่อไปในอนาคต

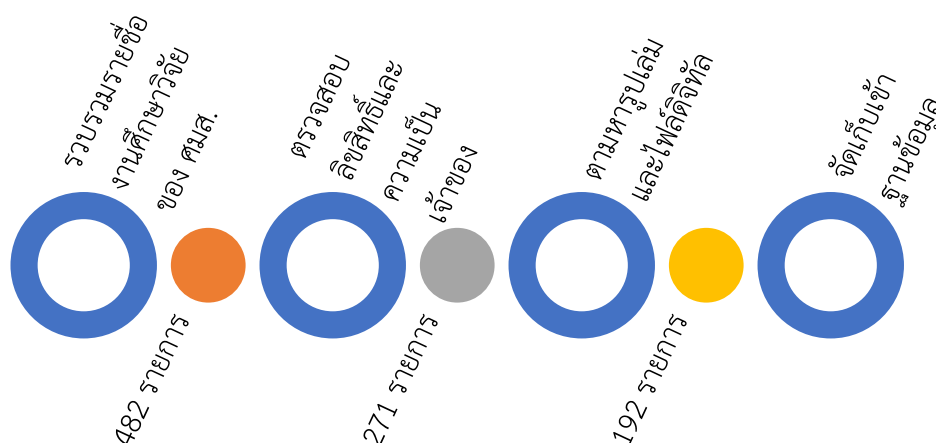
สำรวจและรวบรวม

คณะทำงานเริ่มงานด้วยการ “สำรวจรายชื่อ” งานศึกษาวิจัยที่ผ่านมาของศูนย์ฯ ในช่วงระหว่าง พ.ศ. 2534 – 2561 เหตุผลสำคัญที่ต้องสำรวจรายชื่องานศึกษาวิจัยทั้งหมดนั้น เนื่องจากศูนย์ฯ ไม่เคยจัดทำทะเบียนข้อมูลงานศึกษาวิจัยของตนเองอย่างเป็นระบบมาก่อน อีกทั้งผลงานที่เป็นไฟล์ดิจิทัลและรูปเล่มสิ่งพิมพ์ต่างถูกแยกกันเก็บกระจัดกระจายอยู่ที่นักวิชาการที่รับผิดชอบโครงการ เจ้าหน้าที่ประสานงาน และห้องสมุดศูนย์ฯ จากการสำรวจเบื้องต้นคณะทำงานรวบรวมรายชื่องานศึกษาวิจัยได้มากถึง 482 รายการ ภายใต้ชื่อเรื่อง (Title) ที่แตกต่างกันไป ซึ่งอาจแบ่งได้เป็น 2 กลุ่ม คือ 1) กลุ่มงานศึกษาวิจัยที่ตั้งชื่อเรื่องตามประเด็นการศึกษา เช่น เหย้าเรือนพื้นบ้านย่านตะวันตก (2545) เพศวิถีของกระเทยอีสานในพื้นที่หมอลำ (2555) พลวัตในจักรวาลทัศน์และโลกทัศน์ของลาวโซ่ง/ไทดำจากพิธีกรรมความตาย (2561) เป็นต้น และ 2) กลุ่มงานศึกษาวิจัยที่ตั้งชื่อเรื่องโดยใช้คำว่า “รายงานวิจัยฉบับสมบูรณ์” “รายงานวิจัย” “รายงานความก้าวหน้า” “รายงานสรุปผลการดำเนินงาน” หรือ “ชื่อโครงการ” แล้วตามด้วยประเด็นหรือพื้นที่การศึกษา เช่น รายงานโครงการวิจัยฉบับสมบูรณ์ เรื่อง “ความเป็นเมืองและความสัมพันธ์ระหว่างสินค้าและผู้คนย่านธนบุรี”: กรณีศึกษาชุมชนริมน้ำเจ้าพระยา ฝั่งธนบุรี (2561) รายงานสรุปผลการดำเนินงานโครงการวิจัยสืบสานภาษาและวัฒนธรรมชนเผ่าปกากะญอสู่การสร้างหลักสูตรท้องถิ่น (2555) โครงการสำรวจองค์ความรู้ชาติพันธุ์วิทยาในเวียดนาม (2546) เป็นต้น

อย่างไรก็ดี เมื่อทำการ “รวบรวมสิ่งพิมพ์และไฟล์ดิจิทัล” ของงานศึกษาวิจัยเหล่านี้ คณะทำงานกลับพบปัญหาที่คาดไม่ถึง 3 ประการ ประการแรก พบรายชื่อโครงการการให้ทุน รายชื่องานวิจัย รายชื่อหนังสือที่ซ้ำซ้อนกันเป็นจำนวนมาก อีกทั้งงานศึกษาวิจัยบางชิ้นที่เป็นการให้ทุนบุคคลภายนอกเพื่อดำเนินงาน กลับไม่สามารถตาม

หาหนังสือสัญญาการให้ทุน หรือหนังสือจัดจ้างเพื่อดำเนินงานทางวิชาการได้ ส่งผลให้ไม่สามารถตรวจสอบได้ว่า ผลงานการศึกษาดังกล่าวเป็นลิขสิทธิ์ของศูนย์ฯ หรือมีเงื่อนไขใดในการเผยแพร่ ด้วยเหตุนี้ คณะทำงานจึงต้องคัด รายชื่อผลงานในกลุ่มนี้ออกไปเป็นจำนวนมากถึง 211 รายการ ประการที่สอง เมื่อคณะทำงานตามหาสิ่งพิมพ์และ ไฟล์ดิจิทัลของผลงานการศึกษาดังกล่าวจากผู้ที่เกี่ยวข้อง กลับพบการสูญหายไปเป็นจำนวนมากเช่นกัน โดย ค้นพบเอกสารที่เป็นรูปเล่ม แต่ไม่ปรากฏไฟล์ดิจิทัล จำนวน 160 รายการ และผลงานศึกษาวิจัยที่มีเพียงรายชื่อ แต่ไม่สามารถตามหาไฟล์ดิจิทัลและเอกสารที่เป็นรูปเล่มได้ จำนวน 79 รายการ (Rungcharoensuksri, 2019b) และประการที่สาม สืบเนื่องจากปัญหาในข้อแรกได้ทำให้คณะทำงานตระหนักว่า นอกเหนือไปจากการเก็บรวบรวม สิ่งพิมพ์และไฟล์ดิจิทัล เพื่อดำเนินการจัดทำฐานข้อมูลแล้ว การให้ข้อมูลบริบทการสร้างผลงาน (Context of creation) ก็มีความสำคัญเช่นกัน เพราะข้อมูลเกี่ยวกับรายละเอียดการให้ทุน ข้อมูลเกี่ยวกับสิทธิ์การเผยแพร่และ การเป็นเจ้าของ หรือรายชื่อและบทบาทของผู้มีส่วนร่วมในการวิจัย ล้วนจำเป็นต่อผู้ปฏิบัติงานของศูนย์ฯ ใน การวางแผนโครงการ และบริหารจัดการผลงานวิชาการที่เป็นทรัพย์สินทางปัญญาขององค์กร

สุดท้ายแล้ว เนื่องจากปัญหาการบริหารจัดการผลงานวิจัยที่ไม่เป็นระบบ รายชื่อผลงานศึกษาวิจัยเกือบ 2 ใน 3 จำเป็นต้องถูกคัดออก และเหลือผลงานที่นำมาใช้เป็นข้อมูลในการจัดทำฐานข้อมูลเพียง 192 รายการ นอกจากนี้อีกหนึ่งประเด็นสำคัญที่คณะทำงานต้องขบคิดเพิ่ม ก่อนเริ่มลงมือออกแบบชุดเมทาดาทาสำหรับ ฐานข้อมูลงานศึกษาวิจัยศูนย์ฯ คือ นอกเหนือไปจากนักวิจัยและบุคคลทั่วไปที่จะเป็นผู้ใช้บริการหลักแล้ว ฐานข้อมูลนี้ควรตอบสนองการใช้งานของผู้ปฏิบัติงานภายในองค์กร ในฐานะผู้ใช้งานและผู้ดูแลระบบด้วยเช่นกัน



ภาพที่ 1 กระบวนการสำรวจและรวบรวมผลงานศึกษาวิจัย และจำนวนเอกสารที่จัดเก็บได้

สร้างและเสริม

เพื่อให้ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร เป็นแหล่งจัดเก็บและเผยแพร่ข้อมูลงานวิจัยในรูปแบบออนไลน์ที่เป็นไปตามมาตรฐานสากล รวมถึงสามารถตอบสนองการใช้งานของผู้ใช้บริการภายในและภายนอกองค์กรได้อย่างเหมาะสม คณะทำงานจึงมีข้อสรุปเบื้องต้นร่วมกัน 3 ข้อ เพื่อเป็นเงื่อนไขในการออกแบบและพัฒนาฐานข้อมูล คือ ข้อที่หนึ่ง ฐานข้อมูลจะต้องออกแบบโดยอ้างอิงกับระบบโครงสร้างข้อมูลที่เป็นมาตรฐานสากล และสืบค้นผ่านระบบสืบค้นรวม (Single search) ของระบบคลังข้อมูลดิจิทัลศูนย์มานุษยวิทยาสิรินธรได้ (<https://www.sac.or.th/main/th/database/index>) ข้อที่สอง ชุดเมทาดาทาที่ใช้อธิบายรายละเอียดทรัพยากร ต้องตอบสนองผู้ใช้บริการได้ทั้ง 3 กลุ่ม คือ กลุ่มแรก ผู้ใช้บริการที่เป็นนักวิจัยและบุคคลทั่วไป กลุ่มที่สอง เจ้าหน้าที่ของศูนย์ฯ ได้แก่ นักวิชาการ ผู้ช่วยนักวิจัย และเจ้าหน้าที่ประสานงาน กลุ่มที่สาม ผู้ดูแลระบบ ได้แก่ นักวิชาการที่รับผิดชอบโครงการ และเจ้าหน้าที่ฐานข้อมูล เนื่องจากผู้ใช้งานทั้งสามกลุ่มต่างมีเงื่อนไขในการเข้าถึง และใช้งานที่แตกต่างกันออกไป ข้อที่สาม เพื่อเพิ่มคุณค่าให้กับผลงานการศึกษาที่เผยแพร่ในรูปแบบออนไลน์ และอำนวยความสะดวกให้กับผู้ใช้บริการในการอ้างอิงงานวิจัยของศูนย์ฯ คณะทำงานมีแผนที่จะสร้างระบบอ้างอิงอัตโนมัติ ด้วยเหตุนี้ ชุดเมทาดาทาที่จะนำมาใช้อธิบายงานศึกษาวิจัย ต้องสามารถให้รายละเอียดเพียงพอต่อการนำไปเรียบเรียงและสร้างบรรณานุกรมที่เป็นตัวแทนของทรัพยากรได้ (Rungcharoensuksri, 2018)

ภายหลังการศึกษาและประชุมร่วมกันของคณะทำงาน ที่ประชุมตัดสินใจเลือกใช้ชุดดัตต์ลินคอร์เมทาดาทา (DCMI, 2005) ในการออกแบบโครงสร้างและให้รายละเอียดฐานข้อมูลด้วยเหตุผลสองประการ ดังนี้

ประการแรก ตั้งแต่ พ.ศ. 2559 เป็นต้นมา ฐานข้อมูลทางวิชาการของศูนย์ฯ ได้รับการพัฒนาและออกแบบขึ้นบนโครงสร้างข้อมูลที่เป็นมาตรฐานสากล รวมถึงมีความสามารถในการเชื่อมโยง และสืบค้นข้อมูลผ่านระบบสืบค้นรวมคลังข้อมูลดิจิทัลของศูนย์ฯ ได้ โดยนักวิชาการที่รับผิดชอบในการออกแบบฐานข้อมูล มีอิสระในการคัดเลือกและออกแบบชุดเมทาดาทาที่จะใช้อธิบายทรัพยากรได้เอง ตามความเหมาะสมของเนื้อหาและวัตถุประสงค์โครงการ (Rungcharoensuksri, 2016) อย่างไรก็ตาม เมื่อพิจารณาจากความสามารถในการเชื่อมโยงกับฐานข้อมูลอื่น ๆ และการดูแลรักษาฐานข้อมูลในระยะยาวแล้ว ประกอบกับการพิจารณาจากชุดเมทาดาทาที่ฐานข้อมูลส่วนใหญ่ของศูนย์ฯ เลือกใช้จะพบว่า จากจำนวนฐานข้อมูล 26 ฐาน ที่ให้บริการอยู่ในขณะนี้ มีฐานข้อมูล que เลือกใช้มาตรฐานดัตต์ลินคอร์เมทาดาทาเป็นชุดโครงสร้างมาตรฐานในการให้รายละเอียดทรัพยากรมากถึง

19 ฐานข้อมูล³ การเลือกใช้ชุดดัชนีคอร์เมทาตาทาเพื่ออธิบายงานศึกษาวิจัยของศูนย์ฯ จึงดูจะเป็นทางเลือกที่เหมาะสมในความเห็นของคณะทำงาน

ประการที่สอง เมื่อพิจารณาจากเนื้อหา รูปแบบที่ปรากฏให้ใช้งาน และทรัพย์สินทางปัญญาของงานศึกษาวิจัยของศูนย์ฯ แล้ว แม้บางส่วนจะมีความใกล้เคียงกับหนังสือ ซึ่งสามารถใช้มาตรฐานการลงรายการตามรูปแบบ MARC21 ได้ ซึ่งนิยมใช้กันสำหรับฐานข้อมูลทรัพยากรสารสนเทศในห้องสมุด แต่เมื่อพิจารณารายละเอียดมาตรฐานดังกล่าวกลับไม่สามารถตอบสนองความต้องการของกลุ่มผู้ใช้งานฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร ตามที่ได้อธิบายไปแล้วเบื้องต้น กลุ่มผู้ใช้งานฐานข้อมูลนี้มีสามกลุ่มหลัก คือ กลุ่มแรก ได้แก่ นักวิจัยและผู้ให้บริการทั่วไป ซึ่งเน้นการเข้าถึงเอกสารฉบับเต็มและใช้งานเมทาตาทาเชิงพรรณนา (Descriptive metadata) ที่อธิบายเนื้อหา (Content) ของงานวิจัย เช่น ชื่อเรื่อง (Title) สารบัญ (Table of Contents) สถานที่ (Spatial) เป็นต้น กลุ่มที่สอง ได้แก่ นักวิชาการ ผู้ช่วยนักวิจัย และเจ้าหน้าที่ประสานงานที่เป็นผู้ปฏิบัติงานของศูนย์ฯ ซึ่งใช้งานทั้งเนื้อหาและข้อมูลที่เกี่ยวข้องกับทรัพย์สินทางปัญญาของทรัพยากร เช่น ประวัติการสร้างและการเก็บรักษา (Provenance) หนังสือสัญญาเลขที่ (Consent Form No.) ชื่อนักวิจัย (Creator) เป็นต้น เพื่อใช้ในการวางแผนและบริหารจัดการทรัพยากรของสถาบัน และกลุ่มสุดท้าย ได้แก่ นักวิชาการที่รับผิดชอบโครงการและเจ้าหน้าที่ฐานข้อมูล ซึ่งรับหน้าที่ในการจัดทำคำอธิบายทรัพยากร นำเข้าและดูแลฐานข้อมูล

เมื่อพิจารณาเงื่อนไขและความต้องการของผู้ใช้งานทั้งสามกลุ่มแล้ว คณะทำงานมีความเห็นร่วมกันว่า ผู้ใช้งานกลุ่มสุดท้ายควรมีความสำคัญเป็นอันดับแรก เพราะกลุ่มคนเหล่านี้ต้องเป็นผู้ให้รายละเอียดทรัพยากร และดูแลฐานข้อมูลโดยตรง (Rungcharoensuksri, 2018a) ด้วยเหตุนี้ ชุดเมทาตาทาที่จะใช้อธิบายรายละเอียดทรัพยากร จึงควรเป็นชุดเมทาตาทาทันักวิชาการที่รับผิดชอบโครงการและเจ้าหน้าที่ฐานข้อมูลของศูนย์ฯ มีความคุ้นเคย และใช้งานอยู่เป็นประจำ อีกทั้งเมื่อพิจารณาพื้นฐานการศึกษาของกลุ่มคนเหล่านี้พบว่าส่วนใหญ่มีพื้นฐานความรู้มาจากสายสังคมศาสตร์และมนุษยศาสตร์เป็นหลัก ซึ่งมีความคุ้นเคยน้อยมากกับมาตรฐานการลงรายการตามรูปแบบ MARC21 ที่นิยมใช้กันในงานสายบรรณารักษศาสตร์ ดังนั้น คณะทำงานจึงตัดสินใจเลือกใช้ชุดดัชนีคอร์เมทาตาทาสวนขยายเพื่อออกแบบโครงสร้างฐานข้อมูล

³ โครงสร้างฐานข้อมูลที่ถูกนำมาประยุกต์ใช้กับฐานข้อมูลดิจิทัลศูนย์มานุษยวิทยาสิรินธรอยู่ 6 รูปแบบ ได้แก่ 1) ออกแบบขึ้นเองตามประเด็นวิชาการ (In-house metadata set) 2) ดัชนีคอร์เมทาตาทา 3) International Standard Archival Description (ISAD(G)) 4) Encoded Archival Context - Corporate bodies, Persons, and Families (EAC-CPF) 5) Categories for the Description of Works of Art (CDWA) และ 6) MACHine-Readable Cataloging (MARC21)

ทั้งนี้ ชุดเมทาดาทาฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร สามารถแบ่งหน้าที่การใช้งานออกเป็น 3 กลุ่ม คือ เนื้อหา รูปแบบที่ปรากฏให้ใช้งาน และทรัพยากรทางปัญญา โดยแต่ละกลุ่มประกอบไปด้วยองค์ประกอบหลัก (Core elements) และองค์ประกอบข้อมูลรายละเอียด (Refinement) ดังนี้ (Rungcharoensuksri, 2019a)⁴

1. เนื้อหา

1.1 ชื่อเรื่อง (Title)	เพื่อระบุ “ชื่อเรื่อง” งานศึกษาวิจัยที่กำหนดโดยนักวิจัย หรือสำนักพิมพ์ ได้แก่ ชื่อเรื่องหลัก (Title) และชื่อรอง (Alternative Title)
1.2 ลักษณะ/รายละเอียด (Description)	เพื่อระบุ “ลักษณะ/รายละเอียด” ของงานศึกษาวิจัย ได้แก่ คำอธิบาย/บทคัดย่อ (Abstract) สารบัญ (Table of Contents) ประวัติการสร้างและการเก็บรักษา (Provenance) สถาบัน/หน่วยงาน (Statement of Responsibility) และงบประมาณ (Budget)
1.3 หัวเรื่อง/คำสำคัญ (Subject)	เพื่อระบุ “หัวเรื่อง/คำสำคัญ” ของงานศึกษาวิจัย ซึ่งใช้เป็นช่องทางในการค้นคืนข้อมูล
1.4 เรื่องที่เกี่ยวข้อง (Relation)	เพื่อระบุ “เรื่องที่เกี่ยวข้อง” และความสัมพันธ์อื่น ๆ ที่เกี่ยวข้องกับงานศึกษาวิจัย
1.5 ขอบเขต (Coverage)	เพื่อระบุ “ขอบเขต/สถานที่และเวลา” ที่ปรากฏ หรือเกี่ยวข้องกับงานศึกษาวิจัย ได้แก่ สถานที่ (Spatial) พิกัด DMS (DMS Location) พิกัด DD (DD Location) และเวลา (Temporal)

2. รูปแบบที่ปรากฏให้ใช้งาน

2.1 รหัส (Identifier)	เพื่อระบุ “รหัส” ที่ใช้เรียกแทน และค้นคืนผลงานศึกษาวิจัย ได้แก่ รหัส LC (LC Call No.) รหัส (Identifier)
2.2 ประเภท (Type)	เพื่อระบุ “ประเภท” ของผลงานศึกษาวิจัย ได้แก่ หนังสือ (Book) วารสาร (Journal) บทความ (Article) วิทยานิพนธ์ (Dissertation) รายงานวิจัย (Research) รายงาน (Report) และอื่น ๆ (Other)
2.3 การจัดพิมพ์ (Publisher)	เพื่อระบุข้อมูลต่าง ๆ ที่เกี่ยวข้องกับ “การจัดพิมพ์” หรือเผยแพร่ผลงานศึกษาวิจัย ได้แก่ สถานที่พิมพ์ (Publication Place) สำนักพิมพ์ (Publisher) ปีที่พิมพ์

⁴ ผู้ที่สนใจสามารถดูโครงสร้างและขอบเขตการให้รายละเอียดข้อมูลฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธรฉบับเต็มได้ทาง <http://www.sac.or.th/databases/sac-research/structure.php>

	(Publication Year) เดือนที่พิมพ์ (Publication Month) ครั้งที่พิมพ์ (Edition Statement) ปี (Volume) ฉบับ (Issue) และหน้า (Page)
2.4 รูปแบบ (Format)	เพื่อระบุ “รูปแบบ” ที่ปรากฏของผลงานศึกษาวิจัย ได้แก่ จำนวนหน้า (Total Pages) สื่อ (Medium) และความยาว (Extent)
2.5 แหล่งที่มา (Source)	เพื่อระบุ “แหล่งที่มา” ของงานศึกษาวิจัย และรายละเอียดต่าง ๆ ที่เกี่ยวข้องกับต้นฉบับ ได้แก่ ไฟล์ต้นฉบับ (Digital File) ISBN DOI วันที่เข้าถึง (Day Accessed) เดือนที่เข้าถึง (Month Accessed) ปีที่เข้าถึง (Year Accessed) และ URL
2.6 ภาษา (Language)	เพื่อระบุ “ภาษา” ที่ใช้ในการอธิบาย และปรากฏในงานศึกษาวิจัย ได้แก่ รูปแบบภาษา (Language) ภาษาต้นฉบับ (Original Language)
2.7 วันที่ (Date)	เพื่อระบุ “วันที่” ที่เกี่ยวข้องกับการสร้างและการเปลี่ยนแปลงข้อมูล รวมถึงคำอธิบายต่าง ๆ ที่เกี่ยวข้อง ได้แก่ วันที่สร้างชุดคำอธิบาย (Description Created Date) วันที่ส่งเข้าระบบ (Submitted Date) วันที่เผยแพร่ (Issued Date) วันที่ปรับปรุงข้อมูล (Updated Date) รายละเอียดการแก้ไข (Maintenance Notes)

3. ทรัพย์สินทางปัญญา

3.1 นักวิจัย ⁵ (Creator)	เพื่อระบุ “นักวิจัย” ที่สร้างสรรค์ทรัพยากร ได้แก่ ชื่อนักวิจัย (First Name of Creator) ชื่อกลางนักวิจัย (Middle Name of Creator) นามสกุลนักวิจัย (Last Name of Creator)
3.2 ผู้มีส่วนร่วม (Contributor)	เพื่อระบุ “ผู้มีส่วนร่วม” ในการสร้างสรรค์ผลงานศึกษาวิจัย ได้แก่ ชื่อผู้มีส่วนร่วม (First Name of Contributor) ชื่อกลางผู้มีส่วนร่วม (Middle Name of Contributor) นามสกุลผู้มีส่วนร่วม (Last Name of Contributor) บทบาทของผู้มีส่วนร่วม (Role of Contributor)
3.3 สิทธิ์ (Rights)	เพื่อระบุ “สิทธิ์” ในผลงานศึกษาวิจัย และข้อมูลต่าง ๆ ที่จำเป็นต่อการนำไปใช้งาน ได้แก่ สิทธิ์ (Rights) รูปแบบลิขสิทธิ์ (Type of License) หนังสือสัญญาเลขที่ (Consent Form No.) ไฟล์หนังสือสัญญา (Digital File of Consent Form)
3.4 ผู้บริหารจัดการ (Administrator)	เพื่อระบุรายนามและหน้าที่ของผู้ที่เกี่ยวข้องกับการ “บริหารจัดการ” ผลงานการศึกษาวิจัย ได้แก่ ผู้จัดทำ (Cataloger Name) ผู้ตรวจทาน (Digital Editor) ผู้ดูแลและนำเข้าข้อมูล (Database Administrator)

⁵ ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธรเลือกที่จะแปลคำว่า “Creator” เป็นภาษาไทยว่า “นักวิจัย” (Researcher) แทนที่จะใช้คำว่า “ผู้สร้างสรรค์” เนื่องจากคำว่า “นักวิจัย” สามารถอธิบายบริบทในการผลิตผลงานได้เหมาะสมกว่า อีกทั้งยังสอดคล้องกับชื่อฐานข้อมูลที่เป็นแหล่งจัดเก็บผลงานของนักวิจัย

อย่างไรก็ดี เนื่องจากฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร จำเป็นต้องให้รายละเอียดข้อมูลเฉพาะบางอย่างเพิ่มเติม เพื่อประโยชน์ในการบริหารจัดการภายในองค์กร ซึ่งองค์ประกอบข้อมูลเหล่านี้ไม่ปรากฏอยู่ในชุดอธิบายข้อมูลส่วนขยายของดublincoreเมทาตาhta อีกทั้งการให้รายละเอียดในบางองค์ประกอบข้อมูลมีความซับซ้อนเกินไปสำหรับผู้ดูแลฐานข้อมูล ด้วยเหตุนี้ คณะทำงานจึงดัดแปลงและประยุกต์ใช้ชุดคำอธิบายดublincoreเมทาตาhtaในสามประเด็น ดังนี้ (Rungcharoensuksri, 2018b)

ประเด็นแรก คณะทำงานได้ออกแบบ “องค์ประกอบข้อมูลส่วนขยาย สมส.” (SAC's qualifiers) ชุดคำอธิบายนี้มีพื้นฐานจากการสร้างองค์ประกอบส่วนขยาย เพิ่มจากแกนหลักของชุดคำอธิบายดublincoreเมทาตาhta เช่น องค์ประกอบข้อมูล “รหัส LC” <dc:identifier.lc> ถูกออกแบบขึ้นเพื่อให้รายละเอียดเลขทะเบียนหนังสือที่จัดเก็บอยู่ในห้องสมุดศูนย์ฯ ซึ่งให้รหัสตามระบบหอสมุดรัฐสภาอเมริกัน (Library of Congress Classification) เพื่อให้ผู้ใช้บริการสามารถขอดูเอกสารที่เป็นรูปเล่มได้ และองค์ประกอบข้อมูล “ประวัติการสร้างและการเก็บรักษา” <dc:description.provenance> ถูกออกแบบขึ้นเพื่อให้รายละเอียดเกี่ยวกับการสร้างและการเก็บรักษาผลงานการศึกษาที่ผ่านมา เช่น ชื่อโครงการวิจัยที่ผลิตงานการศึกษาชิ้นนี้ ชื่อนักวิชาการ หรือเจ้าหน้าที่ประสานงานที่ อนุมัติ และเกี่ยวข้องกับโครงการ หรือองค์ประกอบข้อมูล “ผู้จัดทำ” <dc:contributor.catalogerName> ถูกออกแบบขึ้นเพื่อระบุชื่อนักวิชาการ หรือเจ้าหน้าที่ฐานข้อมูลที่เป็นผู้ให้รายละเอียดเมทาตาhtaทางงานศึกษาวิจัยรายการนี้

ประเด็นที่สอง เพื่ออำนวยความสะดวกให้ผู้ดูแลฐานข้อมูลในการให้รายละเอียดข้อมูล องค์ประกอบข้อมูลบางรายการถูกแตกย่อย เช่น องค์ประกอบข้อมูล “นักวิจัย” (Creator) ถูกแตกย่อยออกเป็น 3 องค์ประกอบคือ “ชื่อนักวิจัย” <dc:creator.firstName> “ชื่อกลางนักวิจัย” <dc:creator.middleName> และ “นามสกุลนักวิจัย” <dc:creator.lastName> หรือหากเป็นองค์ประกอบข้อมูลที่จะต้องระบุเนื้อหาเดิมซ้ำ ๆ ก็จะถูกเปลี่ยนลักษณะ (attribute) จากข้อความ (text) ให้เป็นการเลือกแทน (choice type) เพื่อประหยัดเวลาและลดความผิดพลาดในการกรอกข้อมูล เช่น “ประเภท” <dc:type> และ “รูปแบบภาษา” <dc:language> เป็นต้น

ส่วนการจัดทำระบบอ้างอิงอัตโนมิติ ซึ่งเป็นเงื่อนไขที่สามในการออกแบบและพัฒนาระบบนั้น เป็นอีกหนึ่งความท้าทายที่คณะทำงานต้องการทดลองดำเนินการ โดยมีวัตถุประสงค์เพื่อเสริมคุณค่าให้กับระเบียบข้อมูลงานศึกษาวิจัยของศูนย์ฯ ที่เผยแพร่ผ่านระบบฐานข้อมูล เนื่องจากที่ผ่านมาแวดวงวิชาการมักมีความเชื่อว่า ผลงานที่ได้รับการ “ตีพิมพ์” เป็นรูปเล่มเท่านั้นที่มีความน่าเชื่อถือ สามารถนำไปใช้อ้างอิงทางวิชาการได้ แต่กลับละเลยที่จะอ้างอิงถึงข้อมูล หรือเอกสารอิเล็กทรอนิกส์ต่าง ๆ ที่เผยแพร่ในระบบออนไลน์ เพราะมองว่าข้อมูลเหล่านี้ขาดความน่าเชื่อถือ ไม่มีหลักแหล่งถาวรในการจัดเก็บ และบริบทในการสร้างที่ชัดเจน แต่ในมุมมองของศูนย์ฯ ผลงานวิชาการเหล่านี้ก็มีคุณค่าไม่ยิ่งหย่อนไปกว่าผลงานที่ได้รับการตีพิมพ์ เช่นกัน เพราะเป็นผลงานที่สร้างสรรค์

ขึ้นโดยผู้เชี่ยวชาญในประเด็นการศึกษานั้น ๆ อีกทั้งยังได้รับการตรวจสอบจากผู้ทรงคุณวุฒิก่อนส่งมอบงานด้วยเหตุนี้ คณะทำงานจึงมีแผนที่จะยกระดับ และเสริมคุณค่างานการศึกษาวิจัยของศูนย์ฯ ผ่านระบบอ้างอิงอัตโนมัติที่ผู้ใช้บริการสามารถคัดลอก (Copy) และนำรูปแบบการอ้างอิงที่ต้องการไปวาง (Paste) ประกอบรายงานการศึกษาได้ถึง 3 แบบ คือ APA MLA และ HARVARD

ส่วนขั้นตอนการออกแบบนั้น เริ่มจากการ “เทียบเคียง” โครงสร้างฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร ซึ่งได้รับการออกแบบไว้แล้วตามชุดมาตรฐานการให้รายละเอียดข้อมูลดับลินคอร์⁶ กับระบบอ้างอิง APA MLA และ HARVARD (University of Reading, 2019) เพื่อพิจารณาว่าองค์ประกอบใดที่เหมือนกัน เช่น Creator-Author, Publication Year-Year, Publication Place-City เป็นต้น แล้วจึงนำองค์ประกอบข้อมูลทั้งหมดมา “เรียบเรียง” ใหม่อีกครั้ง ตามเงื่อนไขการจัดเรียงของแต่ละระบบอ้างอิง พร้อมกับระบุตำแหน่งการใส่เครื่องหมาย การเว้นวรรค คำเชื่อม หรือลักษณะพิเศษต่าง ๆ เช่น จุลภาค (,) อัญประกาศ (“...”) จาก (from) ค้นคืนเมื่อ (Retrieved) ตัวเอน (Italic) เป็นต้น⁷ นอกจากนี้ เพื่อให้เกิดความเข้าใจที่ชัดเจนในการทำงานร่วมกันระหว่างทีมออกแบบ (นักวิชาการและเจ้าหน้าที่ฐานข้อมูล) และโปรแกรมเมอร์ที่รับผิดชอบในการสร้างฐานข้อมูล คณะทำงานตัดสินใจจัดทำเอกสารประกอบการออกแบบโครงสร้างฐานข้อมูล และเอกสารอธิบายวิธีการจัดทำระบบอ้างอิงอัตโนมัติ เพื่อใช้เป็นเครื่องมือสื่อสารการทำงานระหว่างคณะทำงาน และใช้เป็นคู่มือประกอบการทำงานสำหรับเจ้าหน้าที่ฐานข้อมูล หรือเจ้าหน้าที่อื่น ๆ ที่เกี่ยวข้อง ซึ่งจะต้องเป็นผู้ดูแลและนำเข้าข้อมูล

ตารางที่ 1 การออกแบบและจัดเรียงบรรณานุกรมระบบอ้างอิงอัตโนมัติ ตามระบบ APA MLA และ HARVARD

อ้างอิงในเนื้อเรื่อง	
APA	(Creator, Publication Year, Page)
MLA	(Creator Page)
HARVARD	(Creator, Page)

⁶ ดับลินคอร์เมทาตาได้กำหนดไว้ว่า ในตาราง Creator หากเป็นภาษาอังกฤษให้ระบุ นามสกุล-ชื่อ-ชื่อกลาง เช่น Fairbank, John K. แต่หากเป็นภาษาไทยให้ระบุ ชื่อ-ชื่อกลาง-นามสกุล เช่น ปรีดา เฉลิมเผ่า กอนันทกุล อย่างไรก็ตาม เจ็อนไซดังกล่าวได้สร้างความสับสนให้กับผู้ดูแลฐานข้อมูลของศูนย์ฯ ที่มีได้มีความรู้ด้านบรรณารักษศาสตร์ ด้วยเหตุนี้ เพื่ออำนวยความสะดวกในการทำงานให้กับผู้ดูแลและนำเข้าข้อมูลองค์ประกอบข้อมูล Creator จึงถูกแยกออกเป็น 3 ตาราง เพื่อระบุชื่อ-ชื่อกลาง-นามสกุล นอกจากนี้ การแตกองค์ประกอบข้อมูลดังกล่าว ยังจำเป็นต้องการเรียกและจัดเรียงชื่อ-ชื่อกลาง-นามสกุล ตามเงื่อนไขของการสร้างระบบอ้างอิงอัตโนมัติอีกด้วย

⁷ ผู้ที่สนใจสามารถดูขั้นตอนการจัดทำระบบอ้างอิงอัตโนมัติ ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาฉบับเต็มได้ทาง <http://www.sac.or.th/databases/sac-research/citation.php>

หนังสือ	
APA	Creator. (Publication Year). <i>Title</i> . Publication Place: Publisher. ค้นคืนเมื่อ Date of Accession, จาก ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร: URL.
MLA	Creator. <i>Title</i> . Publisher, Publication Year. ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร [Online]. ค้นคืนเมื่อ Date of Accession <URL>
HARVARD	Creator, Publication Year. <i>Title</i> . [Online]. Publication Place: Publisher. จาก ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร: URL [ค้นคืนเมื่อ Date of Accession]
บทความในวารสาร	
APA	Creator. (Publication Year). <i>Title</i> . ใน Contributor (Role of Contributor), Alternative Title (Page), Publication Place: Publisher. ค้นคืนเมื่อ Date of Accession, จาก ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร: URL.
MLA	Creator. "Title." Alternative Title. บรรณาธิการโดยContributor, Publisher, Publication Year, Page. ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร[Online], ค้นคืนเมื่อ Date of Accession <URL>
HARVARD	Creator, Publication Year. "Title", ใน Contributor (Role of Contributor), Alternative Title. [Online]. Publication Place: Publisher. Page. จาก ฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร: URL [ค้นคืนเมื่อ Date of Accession]

ภาพที่ 2 หน้าระบบบริหารจัดการฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร
ซึ่งนักวิชาการและเจ้าหน้าที่ฐานข้อมูลจะต้องเข้ามากรอกรายละเอียดทรัพยากร

ภาพที่ 3 หน้าแสดงผลการเข้าถึงข้อมูลของฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร

ผู้ให้บริการสามารถเข้าถึงคำอธิบายรายละเอียดข้อมูล และดาวน์โหลดไฟล์งานศึกษาวิจัยได้ทั้งฉบับและแบบแยกบท นอกจากนี้ยังได้ข้อมูลสำหรับการอ้างอิงงานศึกษาวิจัยรายการนี้ ผ่านระบบอ้างอิงอัตโนมัติที่มีให้เลือก 3 รูปแบบ คือ APA MLA และ HARVARD

สะท้อนบทเรียนเพื่อการปรับปรุง

ปัจจุบันสังคมวิชาการสากลได้ร่วมมือกันผลักดันและส่งเสริมให้มีการเผยแพร่ผลการศึกษาวิจัยในทุก รูปแบบ เพื่อให้ผู้คนในสังคมได้มีสิทธิ์เข้าถึงและนำไปใช้ประโยชน์ได้ โดยปราศจากข้อจำกัดทางด้านค่าใช้จ่าย เวลา และสิทธิ์การนำไปใช้ เนื่องด้วยแนวคิดที่มองว่า ทุนสนับสนุนการศึกษาวิจัยมาจากภาษีรายได้ของประชาชน ดังนั้นประชาชนทั่วไปจึงควรได้รับโอกาสที่จะเข้าถึง และใช้ประโยชน์จากผลการค้นคว้าวิจัยนั้น เพื่อเสริมความรู้และพัฒนาต่อยอดองค์ความรู้ (Suber, 2013)

การพัฒนาฐานข้อมูลผลงานศึกษาวิจัยของศูนย์มานุษยวิทยาสิรินธร ถือเป็นแนวทางหนึ่งที่ส่งเสริมแนวคิดวิทยาการแบบเปิด (Open science) ช่วยให้ผู้คนทั่วไปสามารถเข้าถึงเอกสารฉบับเต็มของผลงานศึกษาวิจัยที่ศูนย์ฯ ได้ผลิตเผยแพร่และสนับสนุนทุนวิจัย ได้ทุกที่ตลอดเวลา ซึ่งสอดคล้องกับพฤติกรรมของผู้ใช้งานในปัจจุบัน และเอื้อประโยชน์ต่อการพัฒนาองค์ความรู้ด้านสังคมศาสตร์และมนุษยศาสตร์ ในขณะเดียวกันศูนย์ฯ เอง ก็ได้รับประโยชน์เช่นเดียวกัน กล่าวคือ ช่วยส่งเสริมพันธกิจขององค์กรที่มุ่งหวังเป็นแหล่งสารสนเทศด้านมานุษยวิทยาและสาขาที่เกี่ยวข้องในระดับชาติและนานาชาติ นอกจากนี้ยังช่วยให้ข้อมูลป้อนกลับแก่ศูนย์ฯ สำหรับการวางแผนการส่งเสริมการศึกษาวิจัยได้อีกด้วย

จากแนวคิด “เอกสารที่มีการตีพิมพ์เผยแพร่ในวงจำกัด” (Grey literature) สู่ “เอกสารตามหลักการ FAIR” (FAIR Principles)

ข้อมูลที่รวบรวมไว้ในฐานข้อมูลงานศึกษาวิจัยศูนย์ฯ นั้น ประกอบด้วยเอกสารที่มีการตีพิมพ์เผยแพร่อย่างแพร่หลาย และเอกสารที่มีการตีพิมพ์เผยแพร่ในวงจำกัด (Grey literature)

Grey literature stands for manifold document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories, but not controlled by commercial publishers i.e., where publishing is not the primary activity of the producing body. (Schöpfel, 2010)

ตามนิยามที่ได้ทบทวนใหม่ของคำว่า “เอกสารที่มีการตีพิมพ์เผยแพร่ในวงจำกัด” ทำให้เห็นว่าลักษณะของเอกสารที่มีการตีพิมพ์เผยแพร่ในวงจำกัดนั้นไม่ได้จำกัดอยู่ในลักษณะรูปเล่ม (Printed documents) อย่าง

เดียว แต่ครอบคลุมถึงรูปแบบอิเล็กทรอนิกส์ ซึ่งสอดคล้องกับการเปลี่ยนแปลงของการสื่อสารทางวิชาการในปัจจุบัน นักวิจัยส่วนใหญ่ผลิตผลงานศึกษาวิจัยในรูปดิจิทัลตั้งแต่ต้น (Born-digital documents) ตลอดจนแหล่งให้ทุนวิจัยก็กำหนดให้นักวิจัยส่งรายงานการศึกษาวิจัยทั้งตัวเล่มและไฟล์ข้อมูลตั้งนั้นจึงเอื้อให้ศูนย์สารสนเทศและแหล่งทุนสามารถจัดเก็บและเผยแพร่เอกสารที่มีคุณค่าเหล่านี้สู่สังคมทางเครือข่ายออนไลน์ได้อย่างกว้างขวางและสะดวกมากขึ้น

อย่างไรก็ตามแม้ว่าการเผยแพร่เอกสารที่มีการตีพิมพ์เผยแพร่ในวงจำกัดแก่สังคม เปิดโอกาสให้ผู้คนสามารถเข้าถึงออนไลน์ได้อย่างเสรี แต่การจะทำให้ทรัพยากรประเภทนี้ได้รับการยอมรับในฐานะเอกสารทางวิชาการที่สามารถนำไปอ้างอิงได้ ต้องอาศัยการปรับเปลี่ยนแนวคิดในการมองคุณค่าทรัพยากรสารสนเทศประเภทนี้ และการแสวงหากลยุทธ์การจัดการคอลเลกชันดิจิทัลให้เข้าถึงได้และใช้งานได้ จึงพบว่าองค์กรหลายแห่งได้พัฒนาลังสารสนเทศดิจิทัล (Digital repositories) หรือฐานข้อมูลเอกสารฉบับเต็ม เป็นแนวทางหรือกลยุทธ์ใหม่สำหรับเผยแพร่งานที่มีคุณค่าให้เกิดประโยชน์สูงสุด และเอื้อให้เกิดการเข้าถึงและต่อยอดผลการค้นคว้าในกลุ่มนักวิจัย นักวิชาการ และผู้ที่มีความสนใจ (Wichian, 2015) นอกจากนี้ศูนย์ฯ พยายามออกแบบระบบฐานข้อมูลและการจัดการคอลเลกชันดิจิทัลนี้ เพื่อสนับสนุนหลักการการจัดการข้อมูล FAIR ซึ่งเน้นการจัดการข้อมูลที่เอื้อให้ข้อมูลเหล่านั้น สามารถค้นหาได้เจอ (Findable) เข้าถึงได้ (Accessible) แลกเปลี่ยนถ่ายโอนข้ามระบบได้ (Interoperable) และนำกลับมาใช้ใหม่ได้ (Reuseable) (Wilkinson et al., 2016) ทั้งนี้ศูนย์ฯ ต้องปรับกระบวนการจัดการผลงานศึกษาวิจัยขององค์กรให้สอดคล้องกับการเปลี่ยนแปลงดังกล่าวด้วยเช่นกัน

แม้ว่าศูนย์ฯ มีประสบการณ์การจัดการสารสนเทศ และการพัฒนาฐานข้อมูลออนไลน์ด้านมานุษยวิทยา และสาขาที่เกี่ยวข้องมายาวนานกว่า 15 ปี แต่การวิเคราะห์และออกแบบระบบฐานข้อมูลแต่ละฐานก็แตกต่างกันไปตามลักษณะสารสนเทศ กลุ่มเป้าหมาย และรูปแบบการให้บริการ สำหรับการจัดทำฐานข้อมูลงานศึกษาวิจัยศูนย์ฯ นั้นก็มีลักษณะพิเศษซึ่งเป็นประเด็นท้าทาย คือ เป็นโครงการที่เพิ่งเริ่มดำเนินการใน พ.ศ. 2561 เพื่อรวบรวมผลงานศึกษาวิจัยย้อนหลังได้ตั้งแต่ พ.ศ. 2534–2561 ทั้งนี้การรวบรวมข้อมูลผลงานศึกษาวิจัยย้อนหลังไปนั้น อาจทำได้ไม่ถนัดนักตามที่ระบุไว้ในข้างต้น นอกจากนี้ ฐานข้อมูลจำเป็นต้องมีการนำเข้างานศึกษาวิจัยเพิ่มขึ้นอย่างต่อเนื่อง งานนี้จึงเป็นโครงการต่อเนื่องที่ต้องอาศัยความมุ่งมั่นของบุคลากรที่เกี่ยวข้อง เวลา และงบประมาณ ตลอดจนการดำเนินงานอย่างเป็นระบบในทุกขั้นตอน เพื่อมีข้อมูลครบถ้วนของงานศึกษาวิจัย ทั้งหมดที่เกี่ยวกับเนื้อหา รูปแบบ และทรัพย์สินทางปัญญา และไฟล์เอกสารฉบับเต็ม

การดูแลรักษาดิจิทัลกับผลงานศึกษาวิจัย (Digital curation and research outputs)

เมื่อพิจารณาประสบการณ์ที่เกิดขึ้นในช่วงสำรวจและเตรียมการการพัฒนาฐานข้อมูลนั้น อาจจำแนกปัญหาและความท้าทายที่ประสบได้ 4 กรณี คือ 1) รายชื่อเอกสารที่ได้มามีความซ้ำซ้อน 2) มีเพียงรายชื่อเอกสาร แต่ขาดเอกสารฉบับเต็ม 3) มีรายชื่อเอกสาร พบเพียงแต่รูปเล่มเอกสารฉบับเต็ม และ 4) มีรายชื่อเอกสาร พบแต่เพียงไฟล์เอกสารดิจิทัล นอกจากนี้บางรายการอาจไม่มีเอกสารที่แสดงสิทธิ์ประกอบจึงไม่อาจนำมาเผยแพร่ให้เข้าถึงออนไลน์ได้

ดังนั้น จึงอาจสรุปได้ว่า ความท้าทายทั้งหมดนี้ล้วนเกี่ยวกับการสูญเสียข้อมูล (Data loss) ซึ่งแบ่งได้เป็น 3 ประเภท ได้แก่ การสูญเสียเอกสารฉบับเต็ม (Loss of digital/printed full-text documents) การสูญเสียเมทาดาทาเพื่อการพรรณนาเอกสารงานวิจัย (Loss of associated descriptive metadata) และการสูญเสียเมทาดาทาเพื่อการบริหารจัดการ (Administrative metadata) เช่น ข้อมูลเกี่ยวกับโครงการวิจัย ข้อตกลงระหว่างศูนย์ฯ และผู้วิจัย ข้อมูลเรื่องสิทธิ์การเป็นเจ้าของ เป็นต้น หากข้อมูลเหล่านี้สูญหาย หรือไม่ครบถ้วนก็ส่งผลกระทบต่อความสมบูรณ์ของรายการเอกสารแต่ละรายการ การเข้าถึงและการแสดงผลของข้อมูล และความครบถ้วนของคอลเลกชัน ปัญหาดังกล่าวนี้นี้เกิดจากต้นทางของกระบวนการดูแลรักษาผลงานศึกษาวิจัยดิจิทัลไม่เป็นระบบ

การพัฒนาฐานข้อมูลเป็นเพียงการสร้างที่จัดเก็บ และช่องทางเผยแพร่ผลงานศึกษาวิจัย อย่างไรก็ตาม ฐานข้อมูลที่มีข้อมูลครบถ้วนและบริการที่มีคุณค่านั้น ต้องอาศัยการดูแลรักษาดิจิทัล (Digital curation) อย่างเป็นระบบตั้งแต่ต้นจนจบ และการได้รับความร่วมมือจากผู้มีส่วนได้ส่วนเสีย (Sabharwal, 2015) ศูนย์การดูแลรักษาดิจิทัล (Digital Curation Centre - DCC) ในสหราชอาณาจักร ได้พัฒนาโมเดลวงจรชีวิตการดูแลรักษาดิจิทัล (DCC curation lifecycle) เพื่อชี้ให้เห็นภาพรวมกระบวนการการดูแลรักษาดิจิทัล และปัจจัยที่จำเป็นสำหรับการจัดการโครงการการดูแลรักษาดิจิทัลด้วย นั่นคือ บุคลากร (People) เนื้อหา (Content) และเทคโนโลยี (Technology) บุคลากรในที่นี้ประกอบด้วยผู้มีส่วนได้ส่วนเสีย เช่น นักสารสนเทศ นักวิจัย นักเทคโนโลยี เจ้าหน้าที่โครงการ ผู้บริหาร เป็นต้น ส่วนเนื้อหานั้นหมายรวมถึงไฟล์ดิจิทัล ซึ่งอาจเป็นเอกสารหรือข้อมูลดิบ ประการสุดท้ายคือเทคโนโลยีที่ใช้สร้างคลังสารสนเทศดิจิทัล อุปกรณ์ฮาร์ดแวร์และซอฟต์แวร์ (Sabharwal, 2015)

หากทบทวนการดำเนินงานที่ผ่านมาโดยเทียบกับขั้นตอนต่าง ๆ ของโมเดลวงจรชีวิตการดูแลรักษาดิจิทัลข้างต้นพบว่า การจัดการคอลเลกชันผลงานศึกษาวิจัยดิจิทัลของศูนย์ฯ เริ่มขึ้นจากตรงกลางและปลายทางของกระบวนการ โดยการ “สำรวจและรวบรวม” อาจเทียบเคียงได้กับขั้นตอน “การประเมินและคัดเลือก” (Appraise & Select) และ “จำหน่ายออก” (Dispose) ส่วนการ “สร้างและเสริม” อาจเทียบเคียงได้กับ “การนำเข้า”

(Ingest) “ขั้นตอนการสงวนรักษา” (Preservation action) และ “จัดเก็บ” (Store) อย่างไรก็ดี หากในอนาคต ศูนย์ฯ มีแผนที่จะจัดการคอลเลกชันงานศึกษาวิจัยของตนให้เป็นระบบ ศูนย์ฯ ควรนำวงจรชีวิตการดูแลรักษาข้อมูลดิจิทัล DCC Curation Lifecycle มาเป็นกรอบแนวทางปฏิบัติสำหรับการจัดการและดูแลรักษาข้อมูลดิจิทัลขององค์กรอย่างจริงจังและครบวงจร โดยมีผลบังคับใช้กับผู้ที่เกี่ยวข้องภายในองค์กร และผู้ที่ได้รับทุนสนับสนุนในการทำวิจัยจากศูนย์ฯ รวมถึงการมีนโยบายและแนวปฏิบัติเป็นลายลักษณ์อักษร เพื่อเผยแพร่แก่ผู้มีส่วนเกี่ยวข้องทั้งภายในและภายนอกศูนย์ฯ อาจช่วยสร้างความเข้าใจการดำเนินงานและนำไปสู่การจัดการคอลเลกชันดิจิทัลอย่างเป็นระบบและมีประสิทธิภาพ

นอกเหนือจากการจัดหารายการทรัพยากรเข้าจัดเก็บในคลังฐานข้อมูลผลงานศึกษาวิจัยศูนย์ฯ ให้ได้ครบถ้วนแล้ว ความก้าวหน้าทางเทคโนโลยีและพฤติกรรมการใช้งานของผู้คนในยุคดิจิทัล เป็นอีกสองเรื่องที่จะท้าทายกระบวนการทำงานขององค์กรในอนาคต ด้วยเหตุนี้ การตระหนักถึงความท้าทายของการเปลี่ยนแปลงทางเทคโนโลยีที่มีต่อการมีอยู่ และใช้งานได้ของข้อมูลดิจิทัลที่มีลักษณะเปราะบางและสูญเสียได้ง่าย รวมถึงการแสวงหาและเตรียมการสำหรับการสงวนรักษาดิจิทัล (Digital preservation) เพื่อให้สินทรัพย์ดิจิทัล (Digital assets) นั้นคงอยู่และสามารถเข้าถึงใช้งานได้ทั้งปัจจุบันและอนาคต จึงเป็นประเด็นที่ศูนย์ฯ ควรให้ความสำคัญและมีนโยบายในการรับมือ เพื่อให้คลังข้อมูลศูนย์ฯ มีมาตรฐานเป็นคลังที่ได้รับความเชื่อมั่น (Trusted repositories) จากนักวิจัยในสาขาทั้งในประเทศและต่างประเทศ

บทสรุป

แม้ว่าการจัดการคอลเลกชันดิจิทัลฐานข้อมูลผลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธร อาจเป็น การดำเนินงานที่เริ่มต้นจาก “ปลายทาง” ที่เต็มไปด้วยอุปสรรคและความท้าทาย แต่โดยแท้จริงแล้วกลับนำไปสู่ การทบทวน “ต้นทาง” ที่ควรจะต้องได้รับการแก้ไข ไม่ว่าจะเป็นการให้ความเข้าใจการจัดการข้อมูล การจัดการ คอลเลกชันดิจิทัล และการดูแลรักษาข้อมูลดิจิทัลให้เป็นระบบและเป็นองค์รวม ซึ่งจำเป็นต้องอาศัยการสนับสนุน เชิงนโยบายจากผู้มีส่วนได้ส่วนเสียในระดับการบริหาร และความมุ่งมั่นจากบุคลากรที่เกี่ยวข้อง รวมถึง ความร่วมมือจากนักวิจัยและหน่วยงานที่เป็นเครือข่ายศูนย์ฯ เพื่อปรับปรุงกระบวนการบริหารจัดการ การ สนับสนุนการศึกษาวิจัยของศูนย์ฯ ให้เป็นระบบ ตลอดจนการดูแลรักษาข้อมูลดิจิทัลให้สามารถเข้าถึงได้ในระยะ ยาว

นอกจากนี้ กระบวนการดังกล่าวจำเป็นต้องอาศัยมาตรฐานต่าง ๆ ที่เกี่ยวข้อง เพื่อให้ข้อมูลที่จะนำเข้า จัดเก็บฐานข้อมูลนั้นเป็นข้อมูลที่พร้อมใช้งาน ซึ่งควรเริ่มตั้งแต่การเริ่มโครงการศึกษาวิจัย การให้ทุน ศึกษาวิจัย การตรวจสอบความถูกต้องและคุณภาพของไฟล์เอกสารการศึกษาวิจัย และการจัดการไฟล์และเมทาดาตาที่ เกี่ยวข้องให้ครบถ้วนสมบูรณ์ ความครบถ้วนและความสมบูรณ์ของเมทาดาตานั้นจะเอื้อให้เกิดประโยชน์มากมาย แก่ผู้ใช้และศูนย์ฯ ตลอดจนนำไปสู่การสร้างสรรค์บริการใหม่บนพื้นฐานข้อมูลเมทาดาตาที่มี เช่น บริการข้อมูล

อ้างอิงอัตโนมัติ หรือ การนำเสนอสถิติผลการศึกษาวิจัยในรูปแบบที่ ซึ่งช่วยให้นักวิจัยที่สนใจพื้นที่ศึกษาสามารถเข้าถึงและจำกัดการสืบค้นได้ตรงประเด็น เป็นต้น

กิตติกรรมประกาศ

ผู้เขียนขอขอบคุณ คุณสุธาสิณี บุญเกิด (เจ้าหน้าที่ฐานข้อมูล) คุณธนาภา อึ้งธนาภา คุณรัตนาพร เจียงคำ และคุณวิสูตร สิงห์โส (ทีมเจ้าหน้าที่คอมพิวเตอร์) ที่ช่วยให้โครงการฐานข้อมูลงานศึกษาวิจัยศูนย์มานุษยวิทยาสิรินธรสามารถเสร็จลุล่วงได้

เอกสารอ้างอิง

- DCMI. (2005). **Using Dublin Core - Dublin Core qualifiers**. Retrieved 17 April 2018, from <http://www.dublincore.org/documents/usageguide/qualifiers/>
- Princess Maha Chakri Sirindhorn Anthropology. (2019). **Museums in Thailand**. Retrieved 17 March 2019, from <https://db.sac.or.th/museum/magazine/book>
- Rungcharoensuksri. S. (2016). **Guideline of data repository management**. (In Thai). Retrieved 22 May 2019, from <https://www.sac.or.th/main/th/contents/index/176>
- Rungcharoensuksri, S. (2018a). **1st meeting minute: Studies and development project of database of research of The Princess Maha Chakri Sirindhorn Anthropology Centre**. (In Thai). Bangkok: The Princess Maha Chakri Sirindhorn Anthropology Centre.
- Rungcharoensuksri, S. (2018b). **2nd meeting minute: Studies and development project of database of research of The Princess Maha Chakri Sirindhorn Anthropology Centre**. (In Thai). Bangkok: The Princess Maha Chakri Sirindhorn Anthropology Centre.
- Rungcharoensuksri. S. (2019a). **Database structure**. (In Thai). Retrieved 15 May 2019, from <http://www.sac.or.th/databases/sac-research/structure.php>
- Rungcharoensuksri, S. (2019b). **The procedures of developing research database**. (In Thai). Retrieved 15 May 2019, from <http://www.sac.or.th/databases/sac-research/process.php>
- Sabharwal, A. (2015). Digital curation lifecycle in practice. **Digital Curation in the Digital Humanities**, 95–108. Retrieved from <https://doi.org/10.1016/b978-0-08-100143-1.00005-2>

- Schöpfel, J. (2010). Towards a Prague definition of grey literature. In Twelfth International Conference on Grey Literature (Eds.), **Transparency in Grey Literature**. (pp. 11–26). Prague, Czech. Retrieved from https://archivesic.ccsd.cnrs.fr/sic_00581570/document
- Suber, P. (2013). Open access (The MIT Press Essential Knowledge Series). **The Australian Library Journal** (62). Retrieved from <https://doi.org/10.1080/00049670.2013.811775>
- University of Reading. (2019). **Citing references**. Retrieved 29 June 2019, from <https://libguides.reading.ac.uk/citing-references/referencingstyles#s-lg-box-9973347>
- Wichian, S. (2015). Grey literature in a digital age. (In Thai). **Journal of Information Science**, 33(1), 123-141. Retrieved from <https://www.tci-thaijo.org/index.php/jiskku/article/view/42503/35131>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. **The Nature**, 3(160018), 1–9. Retrieved from <https://doi.org/10.1038/sdata.2016.18>