



ข้อเสนอเชิงนโยบาย:

การจัดทำมาตรฐานการสร้างไฟล์ดิจิทัลเอกสารตัวเขียน เพื่อการอนุรักษ์มรดกภูมิปัญญาและการเชื่อมโยงคลังข้อมูล

1. บทสรุปผู้บริหาร

ประเทศไทยมีเอกสารตัวเขียนอันเป็นมรดกภูมิปัญญาจำนวนมาก กระจายอยู่ในหน่วยงานและพื้นที่หลากหลาย ทั้งหอสมุด หอจดหมายเหตุ พิพิธภัณฑ์ วัด แหล่งเรียนรู้ท้องถิ่น สถาบันอุดมศึกษา แหล่งครอบครัวของชุมชน และเอกชน เอกสารเหล่านี้ประกอบด้วยคัมภีร์โบราณ หนังสือสมุดไทย สมุดข่อย พับสาหนังสือบุต สมุดฝรั่ง จารึก และเอกสารจดหมายเหตุ ซึ่งล้วนมีคุณค่าต่อการศึกษาประวัติศาสตร์ ศาสนา วรรณกรรม กฎหมาย เวชศาสตร์พื้นบ้าน พิธีกรรม ภูมิปัญญาท้องถิ่น ประวัติศาสตร์ชุมชน และการศึกษาวิจัยด้านมนุษยศาสตร์

ปัจจุบัน หน่วยงานต่าง ๆ มีความพยายามในการจัดทำสำเนาดิจิทัลเอกสารตัวเขียน เพื่อป้องกันความเสียหายของต้นฉบับ ลดการหยิบจับเอกสารจริง และขยายโอกาสในการเข้าถึงข้อมูลให้กว้างขวางขึ้น อย่างไรก็ตาม การดำเนินงานจำนวนมากยังมีลักษณะกระจัดกระจาย ต่างคนต่างทำ และขาดมาตรฐานกลางร่วมกัน ทั้งในด้านการประเมินสภาพเอกสารก่อนถ่ายภาพ วิธีจับต้องเอกสาร การเลือกใช้อุปกรณ์ คุณภาพไฟล์ภาพ การจัดการสี การตั้งชื่อไฟล์ โครงสร้างไฟล์เดออร์ การบันทึกข้อมูลกำกับ การตรวจคุณภาพไฟล์ การสำรองข้อมูล การส่งมอบไฟล์ และการสงวนรักษาระยะยาว

ปัญหาดังกล่าวส่งผลกระทบต่ออย่างน้อย 4 ประการ ได้แก่ ประการแรก เอกสารต้นฉบับอาจได้รับความเสียหายจากการใช้เครื่องมือหรือวิธีการถ่ายภาพที่ไม่เหมาะสม เช่น การใช้เครื่องสแกนแบบแท่นเรียบกับสมุดไทยหรือสมุดฝรั่งที่สันเปราะ ประการที่สอง ไฟล์ภาพดิจิทัลจำนวนมากหนึ่งยังมีคุณภาพไม่เพียงพอต่อการใช้เป็นหลักฐานทางวิชาการ เช่น ความละเอียดต่ำ สีผิดเพี้ยน ไม่มีมาตราส่วน ไม่มีเป้าเทียบสี หรือถูกบีบอัดจนสูญเสียรายละเอียด ประการที่สาม ข้อมูลกำกับไฟล์และทะเบียนเอกสารยังไม่เป็นมาตรฐานเดียวกัน ทำให้ฐานข้อมูลของแต่ละหน่วยงานไม่สามารถเชื่อมโยงกันได้ และประการที่สี่ ไฟล์ดิจิทัลจำนวนมากขาดระบบสำรอง ตรวจสอบความสมบูรณ์ และสงวนรักษาระยะยาว จึงเสี่ยงต่อการสูญหาย เสื่อมสภาพ หรือไม่สามารถเปิดใช้งานได้ในอนาคต

ข้อเสนอเชิงนโยบายฉบับนี้จึงเสนอให้จัดทำ “มาตรฐานการสร้างไฟล์ดิจิทัลและการบริหารจัดการไฟล์สำเนาดิจิทัลเอกสารตัวเขียน” เพื่อเป็นแนวทางกลางสำหรับหน่วยงานภาครัฐ สถาบันการศึกษา วัด ชุมชน พิพิธภัณฑ์ท้องถิ่น และภาคีเครือข่ายที่ดำเนินงานด้านเอกสารตัวเขียน โดยมาตรฐานดังกล่าวควรครอบคลุมกระบวนการตั้งแต่การสำรวจและประเมินสภาพเอกสาร การเตรียมเอกสาร การถ่ายภาพอย่างปลอดภัย การ

กำหนดคุณภาพไฟล์ต้นฉบับเพื่อการสงวนรักษา การสร้างไฟล์ใช้งาน การตั้งชื่อไฟล์ การจัดโครงสร้างไฟล์เตอร์ การบันทึก metadata ขั้นต่ำ การตรวจคุณภาพ การสำรองข้อมูล และการเตรียมความพร้อมเพื่อเชื่อมโยงเข้าสู่คลังข้อมูลดิจิทัลในระยะยาว

นโยบายนี้เสนอแนะให้กำหนดมาตรฐานกลางที่ทุกหน่วยงานและชุมชนสามารถนำไปใช้ได้ โดยให้เจ้าของเอกสารยังคงมีสิทธิในการดูแล กำกับ อธิบาย และกำหนดระดับการเข้าถึงเอกสารของตนเอง ขณะเดียวกันไฟล์ดิจิทัลและข้อมูลกำกับควรมีรูปแบบที่สามารถเชื่อมโยงกับระบบทะเบียนกลาง คลังข้อมูล และเครือข่ายมาตรฐานสากล

การดำเนินนโยบายนี้จะช่วยยกระดับการดิจิทัลเอกสารตัวเขียนของประเทศไทยจากการดำเนินงานแบบรายโครงการไปสู่ระบบมาตรฐานระดับชาติที่โปร่งใส ตรวจสอบได้ เคารพสิทธิชุมชน และยั่งยืนในระยะยาว อีกทั้งยังเป็นการวางโครงสร้างพื้นฐานข้อมูลสำหรับการถ่ายทอด ปรึวรรต ศึกษาวิจัย พัฒนาฐานข้อมูล และต่อยอดสู่เทคโนโลยีเพื่อการรู้จำอักษรโบราณในอนาคต

2. บทนำ

เอกสารตัวเขียนเป็นหลักฐานทางวัฒนธรรมที่บันทึกองค์ความรู้ของสังคมไทยและภูมิภาคเอเชียตะวันออกเฉียงใต้ไว้เป็นจำนวนมาก เอกสารเหล่านี้มีได้เป็นเพียงวัตถุโบราณ แต่เป็น “คลังข้อมูลทางภูมิปัญญา” ที่สะท้อนโลกทัศน์ ระบบความรู้ ความเชื่อ ประวัติศาสตร์ ภาษา วรรณกรรม กฎหมาย การแพทย์ พื้นบ้าน พิธีกรรม ศาสนา และความทรงจำของชุมชนในแต่ละภูมิภาค

ลักษณะสำคัญของเอกสารตัวเขียนคือความหลากหลายทั้งด้านวัสดุ รูปแบบ อักษร ภาษา และบริบทการใช้งาน เช่น ใบลานมักใช้บันทึกคัมภีร์ศาสนา ตำรา พงศาวดาร หรือวรรณกรรม โดยมีรอยจารที่อาจฉีก ลบเลือน หรือแตกหักตามอายุของใบลาน สมุดไทยและสมุดข่อยมีลักษณะพับทบและเปิดอ่านขึ้นด้านบน พับสาของภาคเหนือทำจากกระดาษสาซึ่งมีความฟูและเปื่อยง่าย หนังสือบุคของภาคใต้มีวัสดุและการเคลือบผิวที่เสี่ยงต่อเชื้อรา ส่วนสมุดฝรั่งหรือเอกสารกระดาษสมัยใหม่บางประเภทอาจมีปัญหากรดในกระดาษ ทำให้กระดาษกรอบ เปราะ และเสื่อมสภาพจากภายใน

การจัดทำสำเนาดิจิทัลจึงเป็นกระบวนการสำคัญในการลดการใช้ต้นฉบับจริง และช่วยให้ภาพของเอกสารสามารถนำไปอ่าน ศึกษา ตรวจสอบ ถ่ายทอด ปรึวรรต วิเคราะห์ และเผยแพร่ได้โดยไม่จำเป็นต้องหยิบจับวัตถุต้นฉบับซ้ำ ๆ อย่างไรก็ตาม การจัดทำสำเนาดิจิทัลไม่ควรถูกเข้าใจว่าเป็นเพียงการถ่ายภาพหรือสแกนเอกสารทั่วไป หากแต่เป็นกระบวนการอนุรักษ์เชิงป้องกันและการจัดการข้อมูลดิจิทัลที่ต้องอาศัยมาตรฐานทางเทคนิค ความรู้ด้านเอกสารตัวเขียน และความเข้าใจในข้อจำกัดของวัสดุต้นฉบับ

ในทางปฏิบัติ การดิจิทัลเอกสารตัวเขียนจำนวนมากประสบปัญหาความไม่เป็นเอกภาพ เช่น ใช้เครื่องมือถ่ายภาพต่างกัน ใช้ความละเอียดของไฟล์ไม่เท่ากัน ตั้งชื่อไฟล์ไม่เป็นระบบ ขาดการจัดลำดับภาพ ขาดการบันทึกข้อมูลกำกับไฟล์ (metadata) ไม่มีไฟล์ต้นฉบับ (master) ที่เหมาะสมต่อการสงวนรักษาในระยะ

ยาว และไม่มีแนวทางชัดเจนในการจัดชุดไฟล์ส่งมอบ สิ่งเหล่านี้ทำให้ข้อมูลดิจิทัลที่สร้างขึ้นมีความเสี่ยงทั้งด้านคุณภาพ ความน่าเชื่อถือ และความยั่งยืน

นอกจากนี้ การกระจายตัวของเอกสารในหลายพื้นที่ยังทำให้การบริหารข้อมูลไม่สามารถใช้แนวทางการรวมศูนย์เพียงอย่างเดียวได้ เนื่องจากเอกสารจำนวนมากเป็นสมบัติของวัด ชุมชน หรือบุคคล ซึ่งมีบริบทความเชื่อ และเงื่อนไขการเข้าถึงแตกต่างกัน นโยบายที่เหมาะสมจึงต้องสร้างมาตรฐานกลางที่ทุกฝ่ายใช้ร่วมกันได้ โดยไม่พรากอำนาจการดูแลและการตีความเอกสารไปจากเจ้าของเอกสารและชุมชน

ดังนั้น ประเทศไทยจึงควรจัดทำมาตรฐานการสร้างไฟล์ดิจิทัลและบริหารจัดการไฟล์สำเนาดิจิทัลเอกสารตัวเขียนอย่างเป็นระบบ เพื่อให้ไฟล์ที่สร้างขึ้นมีคุณภาพเพียงพอสำหรับการสงวนรักษา ใช้งานทางวิชาการ และเชื่อมโยงข้อมูลระหว่างหน่วยงาน

3. ปัญหาและความท้าทาย

3.1 ความเสี่ยงต่อวัตถุดิบจากกระบวนการดิจิทัลที่ไม่เหมาะสม

เอกสารตัวเขียนจำนวนมากมีสภาพเปราะบางจากอายุ วัสดุ ความชื้น เชื้อรา แมลง ฝุ่น กรดในกระดาษ หรือการซ่อมแซมเดิม หากนำเข้าสู่กระบวนการทำสำเนาดิจิทัลโดยไม่มีการประเมินสภาพก่อน อาจเกิดความเสียหายต่อวัตถุดิบ เช่น ไบลานแตกหัก สมุดไทยฉีกขาด สันสมุดฝรั่งแตก เป็นต้น

การใช้เครื่องสแกนแบบแท่นเรียบกับเอกสารเปราะบางเป็นความเสี่ยงสำคัญ เนื่องจากต้องคว่ำและกดเอกสารกับกระจก ซึ่งอาจทำให้ไบลานที่แห้งกรอบแตกหักได้ ขณะเดียวกันการใช้แสงที่ให้ความร้อนสูงหรือมีรังสีที่เป็นอันตรายอาจเร่งการเสื่อมสภาพของหมึก สี และวัสดุรองรับ

3.2 คุณภาพไฟล์ไม่เพียงพอต่อการสงวนรักษาและการใช้เป็นหลักฐานทางวิชาการ

ไฟล์ภาพจำนวนมากถูกสร้างขึ้นเพื่อการใช้งานทั่วไป เช่น JPEG ความละเอียดต่ำ หรือ PDF ที่ผ่านการบีบอัดสูง ทำให้สูญเสียรายละเอียดของตัวอักษร เส้นจาร ร่องรอยหมึก ลักษณะวัสดุ และความแตกต่างของสี ทำให้ไม่สามารถตรวจสอบสี ขนาด และสัดส่วนของเอกสารได้อย่างน่าเชื่อถือ

รายละเอียดเพียงเล็กน้อยของเอกสารตัวเขียน เช่น รอยจาร เส้นขีด รูปร่างตัวอักษร รอยลบ รอยซ่อม การเปลี่ยนสีของหมึก หรือร่องรอยการชำรุด มีความสำคัญต่อการอ่าน การถ่ายทอด การปริวรรต และการวิเคราะห์เชิงอักษรวิทยา หากไฟล์ภาพไม่มีคุณภาพเพียงพออาจทำให้ต้องกลับไปใช้ต้นฉบับจริงซ้ำอีก

3.3 การตั้งชื่อไฟล์และโครงสร้างโฟลเดอร์ไม่เป็นระบบ

การตั้งชื่อไฟล์ตามลำดับอัตโนมัติจากกล้องหรือสแกนเนอร์ เช่น IMG_0001 หรือ DSC_0001 โดยไม่เชื่อมโยงกับรหัสเอกสาร ประเภทเอกสาร ผูก เล่ม ใบ หน้า หรือแหล่งที่มา ทำให้เมื่อเวลาผ่านไปไม่สามารถตรวจสอบย้อนกลับได้ว่าไฟล์ใดสัมพันธ์กับต้นฉบับใด ลำดับใด หรือชุดเอกสารใด ปัญหานี้ส่งผลกระทบต่อการจัดเก็บ การส่งมอบ การตรวจรับ การนำเข้าฐานข้อมูล การเผยแพร่ และการสำรองข้อมูลในระยะยาว โดยเฉพาะเมื่อมีไฟล์จำนวนมากจากหลายพื้นที่และหลายหน่วยงาน

3.4 ข้อมูลกำกับไฟล์ (metadata) ไม่สอดคล้องกัน

ข้อมูลกำกับไฟล์ (metadata) จะช่วยอธิบายว่าเอกสารคืออะไร อยู่ที่ใด สร้างจากวัสดุใด ใช้อักษรและภาษาอะไร มีจำนวนเท่าใด ถ่ายภาพเมื่อใด ใช้อุปกรณ์ใด ใครเป็นผู้ถ่าย ใครเป็นเจ้าของ และเปิดให้เข้าถึงได้ในระดับใด หากขาดข้อมูลกำกับไฟล์ (metadata) ไฟล์ภาพจะกลายเป็นข้อมูลที่ค้นหา เชื่อมโยง และตรวจสอบได้ยาก

3.5 การขาดระบบตรวจสอบความสมบูรณ์และการสงวนรักษาระยะยาว

ไฟล์ดิจิทัลมีความเสี่ยงต่อการสูญหายจากอุปกรณ์เสียหาย การคัดลอกผิดพลาด ไฟล์เสียหาย การเปลี่ยนแปลงรูปแบบเทคโนโลยี หรือการขาดระบบสำรองข้อมูล และแผนการอพยพข้อมูล ไฟล์สำเนาดิจิทัลที่สร้างขึ้นด้วยงบประมาณและแรงงานจำนวนมากอาจสูญหายหรือเปิดใช้งานไม่ได้ในอนาคต

3.6 ความเหลื่อมล้ำด้านเครื่องมือ บุคลากร และองค์ความรู้

แหล่งเอกสารระดับท้องถิ่น เช่น วัด พิพิธภัณฑ์ชุมชน หอวัฒนธรรม หรือสถาบันขนาดเล็ก มักมีข้อจำกัดด้านงบประมาณ เครื่องมือ บุคลากร และความรู้ทางเทคนิค หากไม่มีมาตรฐานที่เหมาะสมและระบบสนับสนุนจากส่วนกลาง เอกสารสำคัญจำนวนมากอาจถูกละเลย หรือการทำสำเนาดิจิทัลด้วยวิธีที่ไม่สามารถประโยชน์ได้ในระยะยาว

4. คำนิยาม

เอกสารตัวเขียน หมายถึง เอกสารที่บันทึกข้อความด้วยการจาร เขียน หรือคัดลอกด้วยมือบนวัสดุต่าง ๆ เช่น ใบลาน สมุดไทย สมุดข่อย พับสา หนังสือบุต สมุดฝรั่ง กระดาษเพลลา จารึก หรือวัสดุพื้นถิ่นอื่น ๆ

การจัดทำสำเนาดิจิทัล หมายถึง กระบวนการแปลงสภาพเอกสารตัวเขียนทางกายภาพให้เป็นไฟล์ภาพดิจิทัล โดยคำนึงถึงความปลอดภัยของวัตถุต้นฉบับ คุณภาพของไฟล์ การจัดการข้อมูลกำกับ และการสงวนรักษาในระยะยาว

ไฟล์ต้นฉบับเพื่อการสงวนรักษา (master) หมายถึง ไฟล์ภาพดิจิทัลคุณภาพสูงสุดที่สร้างจากต้นฉบับ ใช้สำหรับเก็บรักษาระยะยาวและเป็นหลักฐานอ้างอิง ไม่ควรถูกแก้ไขหรือเผยแพร่แทนไฟล์ใช้งานทั่วไป

ไฟล์ใช้งาน หมายถึง ไฟล์ที่สร้างจากไฟล์ต้นฉบับเพื่อใช้ในการอ่าน ค้นดู เผยแพร่ หรือให้บริการ โดยมีขนาดและรูปแบบเหมาะสมต่อการใช้งาน

ข้อมูลกำกับ (metadata) หมายถึง ข้อมูลที่ใช้อธิบายเอกสารและไฟล์ดิจิทัล เช่น รหัสเอกสาร ชื่อเรื่อง ประเภทเอกสาร วัสดุ อักษร ภาษา สถานที่เก็บ วันที่ถ่ายภาพ ผู้ถ่ายภาพ อุปกรณ์ รูปแบบไฟล์ สิทธิ และระดับการเข้าถึง

5. วัตถุประสงค์

1. เพื่อจัดทำมาตรฐานกลางในการสร้างไฟล์ดิจิทัลเอกสารตัวเขียนที่มีคุณภาพ สอดคล้องกับหลักการอนุรักษ์ และเหมาะสมต่อการใช้งานทางวิชาการ

6. กลุ่มเป้าหมายและผู้มีส่วนได้ส่วนเสีย

กลุ่ม	บทบาท
หน่วยงานกำหนดนโยบาย	กระทรวงวัฒนธรรม กรมศิลปากร สำนักหอสมุดแห่งชาติ หอจดหมายเหตุแห่งชาติ และหน่วยงานด้านดิจิทัลภาครัฐ
หน่วยงานวิชาการและวิจัย	สถาบันอุดมศึกษา ศูนย์มานุษยวิทยาสิรินธร และหน่วยบริหารทุนวิจัย
หน่วยปฏิบัติการ	ห้องสมุด พิพิธภัณฑ์ท้องถิ่น หอวัฒนธรรม อปท. และหน่วยงานภาคสนาม
เจ้าของหรือผู้ครอบครองเอกสาร	วัด พระภิกษุ ชุมชน ครอบครัวเก่าแก่ นักสะสม และเอกชน
ผู้เชี่ยวชาญเฉพาะด้าน	นักอนุรักษ์ นักจดหมายเหตุ นักสารสนเทศ ช่างภาพดิจิทัล ผู้เชี่ยวชาญอักษรโบราณ และผู้เชี่ยวชาญเทคโนโลยี
ผู้ใช้ประโยชน์	นักวิชาการ นักศึกษา ประชาชน ผู้พัฒนาฐานข้อมูล และนักพัฒนา AI

7. ขั้นตอนการดำเนินงาน

การขับเคลื่อนนโยบายการจัดทำมาตรฐานไฟล์สำเนาดิจิทัลเอกสารตัวเขียน แบ่งการดำเนินงานออกเป็น 5 ระยะ ดังนี้

ระยะที่ 1: การเตรียมการและการประเมินสถานภาพ

ก่อนเริ่มกระบวนการสแกนหรือถ่ายภาพ ต้องมีการประเมินและบริหารจัดการตัวเอกสารทางกายภาพอย่างรัดกุม

1.1 การสำรวจและจัดทำทะเบียนเบื้องต้น สำรวจแหล่งจัดเก็บทั่วประเทศ จัดทำบัญชีรายชื่อ เพื่อป้องกันเอกสารสูญหายหรือข้อมูลสลับเปลี่ยนระหว่างดำเนินการ

1.2 การประเมินสภาพและคัดกรอง ประเมินสภาพเอกสาร ทำความสะอาด ทำทะเบียน และจัดเตรียมเอกสารสำเนาดิจิทัล

- หากพบเอกสารต้นฉบับอยู่ในสภาพไม่สมบูรณ์ อาทิ สมุดไทย หนังสือบุค หรือพับสา มีเชื้อราเกาะแน่นหรือเปื่อยยุ่ยจนหยิบจับไม่ได้ ต้องส่งเข้ากระบวนการ “อนุรักษ์เชิงฟื้นฟู” เช่น การทำความสะอาดแบบแห้ง การซ่อมแซมก่อนนำไปทำสำเนาดิจิทัล เป็นต้น
- เอกสารที่สภาพสมบูรณ์สามารถเข้าสู่กระบวนการเตรียมจัดทำสำเนาดิจิทัลได้ทันที โดยตรวจสอบการเรียงหน้า เช่น ไบลานจัดเรียงหน้าตามลำดับangkaให้เรียบร้อย สมุดข่อย หนังสือบุค หรือพับสา ให้ตรวจหาหน้าต้นหน้าปลาย เป็นต้น

ระยะที่ 2: การกำหนดมาตรฐานทางเทคนิคและข้อมูลกำกับ

2.1 มาตรฐานไฟล์ภาพ

ระดับไฟล์	รูปแบบไฟล์	ความละเอียด ขั้นต่ำ	สี/Bit depth	การใช้งาน	เงื่อนไข
Preservation master	T I F F uncompressed หรือ JPEG2000 lossless	400–600 PPI	24-bit หรือ 48-bit color	เก็บถาวรและ ใช้ เป็น หลักฐาน	ห้ามปรับแต่ง เพื่อ ความ สวยงาม
Production master	TIFF หรือ JP2 คุณภาพสูง	อ้างอิงจาก master	24-bit color	สร้างไฟล์ บริการ	ต้องสร้างจาก master
Access file	J P E G , J P 2 , P D F / A หรือ WebP	150–300 PPI	sRGB	ให้บริการและ อ่านทั่วไป	ใช้ตามระดับ สิทธิ
Thumbnail	JPEG หรือ WebP	72–150 PPI	sRGB	แสดงตัวอย่าง	ไม่ใช่ แทน หลักฐาน วิชาการ
Metadata/Log file	CSV, XML, JSON, TXT	ไม่เกี่ยวข้อง	ไม่เกี่ยวข้อง	ข้อมูลกำกับ ไฟล์	ต้องแนบกับ ชุดส่งมอบ

2.2 การตั้งชื่อไฟล์

การกำหนดโครงสร้างการตั้งชื่อไฟล์ให้เป็นระบบและอ่านเข้าใจได้ง่าย เช่น [รหัสจังหวัด][ลำดับแหล่งข้อมูล]-[ลำดับเอกสาร].[ชื่อเรื่อง]-[ลำดับฉบับ, ผูก]-[ลำดับไฟล์เอกสาร] ตัวอย่าง: BKK001-001.PhraMalai-001-001.jpg หมายถึง BKK-กรุงเทพฯ, 001-วัดหนังสือราชวรวิหาร, 001-เอกสารลำดับที่ 1, PhraMalai-พระมาลัย, 001-ฉบับที่ 1, 001-ไฟล์ดิจิทัลที่ 1 ตัวอย่างนี้เป็นการตั้งชื่อไฟล์ของศูนย์มานุษยวิทยาสิรินธร (องค์การมหาชน) ทั้งนี้การตั้งชื่อไฟล์ควรเป็นตัวเลขอารบิกและตัวอักษรภาษาอังกฤษ ไม่ควรใช้เลขไทยและอักษรไทยเพราะอาจทำให้การความผิดพลาดในการเรียกลำดับและการแปลงผลได้ ทั้งนี้ควรจัดทำคู่มือการตั้งชื่อไฟล์ไว้สำหรับตรวจสอบด้วย

2.3 โครงสร้างข้อมูลกำกับ (Metadata Schema): พัฒนาตามมาตรฐาน Dublin Core โดยให้รายละเอียดข้อมูล ดังนี้

หมวด	รายการข้อมูล
รหัส	Identifier, persistent ID, รหัสเดิมของหน่วยงาน
ชื่อเรื่อง	ชื่อเรื่องหลัก ชื่อรอง
ประเภทเอกสาร	ใบลาน สมุดไทย พับสา หนังสือбуд สมุดฝรั่ง จารึก

หมวด	รายการข้อมูล
วัสดุ	ไบลาน กระดาษข่อย กระดาษสา กระดาษฝรั่ง ผ้า ไม้ โลหะ
อักษร	ขอมไทย ธรรมล้านนา ธรรมอีสาน ไทยน้อย มอญ พม่า ยาวี ฯลฯ
ภาษา	ไทย บาลี สันสกฤต ล้านนา ลาว มอญ เขมร ฯลฯ
ลักษณะกายภาพ	จำนวนใบ หน้า ผูก เล่ม ขนาด สภาพ
แหล่งที่มา	เจ้าของ สถานที่เก็บเดิม สถานที่เก็บปัจจุบัน
การดิจิทัล	วันที่ถ่ายภาพ ผู้ถ่ายภาพ อุปกรณ์ ความละเอียด รูปแบบไฟล์
สีและสเกล	Color target, scale, color profile
สิทธิ์	เจ้าของสิทธิ์ เงื่อนไขการใช้ ระดับการเข้าถึง
การสงวนรักษา	checksum, storage location, backup, migration history
ความสัมพันธ์	ชุดเอกสาร ผูก เล่ม ภาพ access file , รหัสเดิมของหน่วยงาน, ข้อมูลปริวรรตถ้ามี
คำสำคัญ	วรรณกรรม เวชศาสตร์ โสยศาสตร์ โหราศาสตร์ กฎหมาย ประวัติศาสตร์ อักษรศาสตร์ ธรรมคดี อักษร ภาษา แหล่งข้อมูล เจ้าของข้อมูล ฯลฯ

ระยะที่ 3: กระบวนการถ่ายภาพและควบคุมคุณภาพ

3.1 มาตรฐานอุปกรณ์การถ่ายทำ

การทำสำเนาดิจิทัลสามารถใช้กล้อง DSLR, Mirrorless หรือ Medium Format ที่มีความละเอียดสูง (24 Megapixels ขึ้นไป) และใช้ขาตั้งกล้องที่ปรับเป็นแนวนอนได้ และจัดแสงสว่างโดยใช้ไฟ LED เกรตสตูดิโอที่ไม่แผ่รังสี UV และ IR และจัดแสงแบบเฉียง 45 องศาเพื่อหลีกเลี่ยงแสงสะท้อน นอกจากนี้ยังสามารถใช้เครื่องสแกนทรง V-Shape สำหรับสแกนเอกสารได้ แต่เครื่องสแกนลักษณะนี้จะมีขนาดใหญ่และขนย้ายลำบาก ไม่เหมาะกับการนำไปใช้นอกสถานที่

หมายเหตุ การใช้เครื่องสแกนแบบแท่นเรียบ ที่ต้องคว่ำเอกสารกดกับกระจก อาจจะทำให้ไบลานที่แห้งกรอบแตกหักได้

3.2 กระบวนการถ่ายทำ

การถ่ายภาพเพื่อทำสำเนาดิจิทัลควรมีเป้าเทียบสี (Color Checker) และไม้บรรทัดมาตราส่วน (Scale) ปรากฏอยู่ในขอบภาพของทุกเฟรม เพื่อให้สามารถปรับตั้งสีให้ตรงกับกระดาษข่อย หรือไบลานของจริง และควรถ่ายภาพให้เห็น "ขอบของเอกสารครบทั้ง 4 ด้าน" เพื่อเป็นหลักฐานว่าไม่มีข้อความใดถูกตัดทอน

3.3 การควบคุมคุณภาพ

หลังจากถ่ายภาพหรือสแกนเอกสารเรียบร้อยแล้วต้องตรวจเช็คความคมชัด สี สัน ตรวจสอบว่าไม่มีหน้าใดตกหล่นโดยเทียบกับทะเบียน

3.4 ชุดไฟล์ส่งมอบมาตรฐาน

รายการส่งมอบ	รายละเอียด
Master files	ไฟล์ TIFF/JP2 lossless
Access files	ไฟล์ JPEG, JP2, PDF/A หรือ WebP
Thumbnail files	ไฟล์ตัวอย่างขนาดเล็ก
Metadata files	CSV, XML, JSON หรือรูปแบบที่กำหนด
Checksum manifest	ค่า SHA-256 หรือค่าตรวจสอบที่กำหนด
QC report	รายงานตรวจคุณภาพไฟล์
Capture log	บันทึกวันถ่ายภาพ อุปกรณ์ ผู้ปฏิบัติงาน และค่ากล้อง
Condition report	รายงานสภาพเอกสารก่อนถ่ายภาพ
Rights statement	ข้อมูลสิทธิและระดับการเข้าถึง
README file	คำอธิบายโครงสร้างไฟล์ วิธีอ่านไฟล์ และข้อจำกัด

3.5 เกณฑ์ตรวจรับไฟล์สำเนาดิจิทัล

เกณฑ์ตรวจรับ	เงื่อนไขผ่าน
ความครบถ้วน	ภาพครบทุกหน้า ใบ ผูก หรือเล่ม ตามทะเบียน
ลำดับภาพ	ไม่สลับ ไม่ขาด ไม่ซ้ำโดยไม่ระบุเหตุ
ความคมชัด	ตัวอักษร เส้นจาร หมึก และขอบวัสดุอ่านได้ชัด
สี	มี color target และสีไม่ผิดเพี้ยนเกินเกณฑ์ที่กำหนด
ขนาด	เห็นขอบเอกสารครบทั้ง 4 ด้าน
แสง	ไม่มีเงา แสงสะท้อน หรือ overexposure ที่บังข้อมูล
ชื่อไฟล์	ตรงตาม naming convention
รูปแบบไฟล์	ตรงตามระดับไฟล์ที่กำหนด
Checksum	มี checksum ครบทุก master file
Metadata	มี metadata ขั้นต่ำครบถ้วน
สิทธิ	ระบุระดับการเข้าถึงครบทุกชุดเอกสาร
รายงานส่งมอบ	มี QC report, capture log และ README

ไฟล์ที่ไม่ผ่านเกณฑ์ต้องถ่ายซ้ำหรือแก้ไขก่อนส่งมอบ เว้นแต่ต้นฉบับมีข้อจำกัดทางกายภาพที่ไม่สามารถถ่ายซ้ำได้ โดยต้องมีบันทึกเหตุผลประกอบการตรวจรับ

ระยะที่ 4: การจัดเก็บและการแสดงผล

การจัดการข้อมูลเพื่อไม่ให้ไฟล์ภาพสูญหายตามกาลเวลา ควรจัดเก็บไฟล์ไว้อย่างเป็นระบบ ดังนี้

4.1 การจัดเก็บและสำรองข้อมูล

1. การจัดการไฟล์ต้นฉบับ (Preservation Master) จัดเก็บไฟล์ภาพถ่ายต้นฉบับที่มีความละเอียดสูงและไม่ถูกบีบอัด เช่น รูปแบบไฟล์ TIFF หรือ RAW เพื่อใช้เป็นตัวแทนของเอกสารจริงในระยะยาว

2. การจัดเก็บสำเนาอย่างน้อย 3 ชุด จัดทำสำเนาข้อมูลทั้งหมดอย่างน้อย 3 ชุด ประกอบด้วย ข้อมูลหลักที่ใช้งาน 1 ชุด และข้อมูลสำรองอีก 2 ชุด

3. การจัดเก็บบนสื่ออย่างน้อย 2 ประเภท คือ บันทึกรหัสข้อมูลลงบนสื่อบันทึกข้อมูลที่แตกต่างกันอย่างน้อย 2 ชนิด เพื่อกระจายความเสี่ยง เช่น ฮาร์ดดิสก์ภายนอก (External HDD) และระบบคลาวด์ (Cloud Storage)

4. การสำรองข้อมูลอย่างน้อย 1 ชุด ไปจัดเก็บไว้ในสถานที่อื่นที่ต่างจากจุดปฏิบัติงานหลัก หรือเก็บบนระบบคลาวด์ที่มีความปลอดภัยสูง เพื่อป้องกันกรณีเกิดภัยพิบัติทางธรรมชาติหรือเหตุฉุกเฉิน

นอกจากนี้ ควรหมั่นตรวจสอบไฟล์ภาพอยู่เป็นระยะ และควรจัดทำแผนการอพยพข้อมูลเมื่อรูปแบบไฟล์หรือสื่อจัดเก็บเริ่มล้าสมัย

4.2 การจัดการเพื่อการแสดงผลและการเข้าถึง

1. การสร้างไฟล์สำหรับแสดงผล แปลงไฟล์ต้นฉบับให้เป็นไฟล์ที่มีขนาดเล็กลงและรองรับการเปิดใช้งานทั่วไป เช่น รูปแบบไฟล์ JPEG หรือ PDF เพื่อความรวดเร็วในการแสดงผลบนระบบเครือข่ายหรือเว็บไซต์

2. การจัดทำข้อมูลกำกับไฟล์ (Metadata) จัดทำระบบข้อมูลอธิบายรายละเอียดของเอกสารกำกับไว้กับไฟล์ภาพ เช่น ชื่อเอกสาร, หมวดหมู่, ปีที่บันทึก, สถานที่จัดเก็บต้นฉบับ และเนื้อหาสังเขป เพื่อให้ง่ายต่อการสืบค้น

3. การนำเข้าสู่ระบบฐานข้อมูลดิจิทัลหรือเว็บไซต์ของสถาบัน เพื่อให้คณาจารย์ นักศึกษา และนักวิจัยในเครือข่ายสามารถเข้าถึงและนำไปใช้ประโยชน์ทางวิชาการได้ตามสิทธิ์ที่กำหนด

ระยะที่ 5: การเสริมสร้างศักยภาพและการพัฒนาต่อยอด

5.1 การจัดตั้งหน่วยปฏิบัติการเคลื่อนที่

การสร้างทีมผู้เชี่ยวชาญพร้อมอุปกรณ์เคลื่อนที่ได้ ลงพื้นที่ไปยังวัดและชุมชนที่มีข้อจำกัดในการขนย้ายเอกสาร เช่น การลงพื้นที่ถ่ายภาพหนังสือบุคในจังหวัดนครศรีธรรมราช เป็นต้น

5.2 การอบรมและพัฒนาบุคลากร

การจัดทำหลักสูตรอบรมการอนุรักษ์ การทำทะเบียนเอกสาร การทำสำเนาดิจิทัล ให้กับ พระภิกษุ ปราชญ์ชาวบ้าน เจ้าหน้าที่หอสมุดท้องถิ่น ให้สามารถทำสำเนาดิจิทัลเอกสารของตนเองได้ตามมาตรฐานกลาง

5.3 การต่อยอดสู่มนุษยศาสตร์ดิจิทัล

การนำเทคโนโลยีปัญญาประดิษฐ์ (AI) ให้เรียนรู้ลายมืออักษรโบราณ เช่น ขอมไทย ธรรมล้านนา ธรรมอีสาน ไทยน้อย เพื่อถอดความจากภาพถ่ายเป็นข้อความภาษาไทยได้

8. ระดับการเปิดเผยข้อมูล

การจัดทำสำเนาดิจิทัลไม่ควรหมายถึงการเปิดเผยสาธารณะโดยอัตโนมัติ ควรกำหนดระดับการเข้าถึงดังนี้

ระดับ	รายละเอียด	ตัวอย่าง
ระดับ A เปิดสาธารณะ	เปิดภาพและ metadata ให้เข้าถึงได้ทั่วไป	เอกสารหมอดอายุลิขสิทธิ์ ไม่มีข้อมูลอ่อนไหว
ระดับ B เปิดเฉพาะ metadata	เปิดข้อมูลบรรณานุกรม แต่ไม่เปิดภาพเต็ม	เอกสารมีข้อจำกัดด้านลิขสิทธิ์
ระดับ C เปิดเฉพาะนักวิจัย	ต้องลงทะเบียนหรือขออนุญาต	เอกสารหายาก เปราะบาง หรือมีเงื่อนไขชุมชน
ระดับ D เปิดเฉพาะพื้นที่	เข้าถึงได้เฉพาะในห้องสมุดหรือศูนย์ข้อมูล	เอกสารที่เจ้าของไม่อนุญาตให้ดาวน์โหลด
ระดับ E ปิดหรือเลื่อนการเปิดเผย	ไม่เปิดเผยจนกว่าจะได้รับอนุญาต	เอกสารศักดิ์สิทธิ์ พิธีกรรม ข้อมูลส่วนบุคคล

9. จริยธรรม สิทธิชุมชน และข้อมูลอ่อนไหว

การจัดทำสำเนาดิจิทัลเอกสารตัวเขียนต้องเคารพสิทธิของเจ้าของเอกสาร ชุมชน และบริบททางวัฒนธรรม โดยยึดหลักสำคัญดังนี้

1. ต้องขออนุญาตก่อนถ่ายภาพและอธิบายวัตถุประสงค์อย่างชัดเจน
2. ต้องไม่ถือว่าการดิจิทัลเท่ากับการเปิดเผยสาธารณะโดยอัตโนมัติ
3. เจ้าของเอกสารควรมีสิทธิกำหนดระดับการเข้าถึง
4. ต้องระบุแหล่งที่มาและเงื่อนไขการใช้ข้อมูลทุกครั้ง
5. เอกสารที่เกี่ยวข้องกับพิธีกรรม ความเชื่อ ศาสนา หรือข้อมูลส่วนบุคคล ต้องได้รับการจัดการด้วยความระมัดระวัง
6. ต้องไม่พรางความเป็นเจ้าของทางวัฒนธรรมจากชุมชน
7. การเผยแพร่ภาพควรมีข้อมูลบริบทประกอบ เพื่อป้องกันการตีความผิดหรือการใช้ข้อมูลผิดบริบท

10. บทสรุป

ข้อเสนอเชิงนโยบายนี้มุ่งยกระดับการเจ้าหน้าที่ดิจิทัลเอกสารตัวเขียนของประเทศไทยจากการดำเนินงานแบบรายโครงการ ไปสู่ระบบมาตรฐานที่สามารถใช้ร่วมกันได้ในระดับประเทศ โดยให้ความสำคัญทั้งต่อความปลอดภัยของวัตถุดิบฉบับ คุณภาพไฟล์ดิจิทัล การจัดการไฟล์อย่างเป็นระบบ metadata ขั้นต่ำ การตรวจคุณภาพ การสำรองข้อมูล การสงวนรักษาระยะยาว และสิทธิของเจ้าของเอกสารหรือชุมชน

การมีมาตรฐานดังกล่าวจะช่วยลดความเสี่ยงจากการถ่ายภาพหรือสแกนที่ไม่เหมาะสม ทำให้ไฟล์ดิจิทัลมีคุณภาพเพียงพอต่อการใช้เป็นหลักฐานทางวิชาการ ลดปัญหาไฟล์สูญหายหรือไม่สามารถตรวจสอบย้อนกลับได้ และทำให้หน่วยงานต่าง ๆ สามารถเชื่อมโยงข้อมูลเข้าสู่ระบบกลางหรือมาตรฐานสากลได้ในอนาคต

มาตรฐานนี้จะเป็นรากฐานของการพัฒนามนุษยศาสตร์ดิจิทัลของประเทศ เพราะไฟล์ภาพคุณภาพสูง และชุดข้อมูลกำกับ (metadata) ที่เป็นระบบจะช่วยสนับสนุนการถ่ายทอด ปรึวรรต วิเคราะห์ข้อความ การพัฒนาคลังข้อมูล การพัฒนาระบบปัญญาประดิษฐ์ และการวิจัยเชิงลึกด้านประวัติศาสตร์ ภาษาวรรณกรรม ศาสนา และภูมิปัญญาท้องถิ่น

ดังนั้น การจัดทำมาตรฐานการสร้างไฟล์ดิจิทัลและการบริหารจัดการไฟล์สำเนาดิจิทัลเอกสารตัวเขียนจึงมิใช่เพียงการกำหนดวิธีถ่ายภาพเอกสาร แต่เป็นการวางโครงสร้างพื้นฐานทางวัฒนธรรมและข้อมูลของชาติ เพื่อให้มรดกเอกสารตัวเขียนของไทยได้รับการอนุรักษ์ เข้าถึง ใช้ประโยชน์ และส่งต่อสู่อนาคตได้อย่างยั่งยืน