



## แบบรายงานการประชุม/ฝึกอบรม/สัมมนา/ศึกษาดูงาน

หน่วยงาน ศูนย์วิทยบริการ สำนักวิทยบริการและเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏสุรินทร์

### ๑. ข้อมูลส่วนบุคคล

ชื่อ-สกุล นายกฤษฏีกา แก้วกรอง ตำแหน่ง นักวิชาการคอมพิวเตอร์

กลุ่มบุคลากร  สายวิชาการ  สายสนับสนุนวิชาการ

### ๒. หลักฐานหรือเรื่องที่เกี่ยวข้องประชุม/ฝึกอบรม/สัมมนา/ศึกษาดูงาน

“สารสนเทศดิจิทัล: ความท้าทายในงานห้องสมุด จดหมายเหตุและพิพิธภัณฑ์”

### ๓. วิทยากรในการสัมมนา

๑. ผศ.ดร. วชิราภรณ์ คลังธนบูรณ์

๒. นายมณฑล กาญจนไพศาล

๓. นางจิราภรณ์ ศิริธร

### ๔. สถาบันหรือหน่วยงานที่จัดสัมมนา

ศูนย์สารสนเทศสิทธิมนุษยชน สำนักงานคณะกรรมการสิทธิมนุษยชนแห่งชาติ

### ๕. ระยะเวลาที่เข้ารับการศึกษา

ระหว่างวันที่ ๓ กุมภาพันธ์ ๒๕๖๖ เวลา ๐๘.๐๐ – ๑๖.๓๐ น. อบรมออนไลน์ผ่านโปรแกรม Zoom

### ๖. งบประมาณที่ใช้ในการสัมมนา

ไม่มีค่าใช้จ่าย

### ๗. วัตถุประสงค์ของการสัมมนา

๑. ทราบถึงแนวคิดในการพัฒนาห้องสมุด เพื่อตอบสนองความต้องการของผู้ใช้บริการ

๒. สามารถจัดการข้อมูลที่ได้รับไปใช้ประโยชน์สูงสุด

### ๘. สรุปเนื้อหาสาระของการสัมมนา

#### การบริหารจัดการสารสนเทศดิจิทัล: การเก็บรักษาและอนุรักษ์

การสงวนรักษา (Preservation) หมายถึง การป้องกัน เก็บรักษาให้ปลอดภัย และรอดพ้นจากสิ่งที่จะมีผลต่อการเสื่อมสลายของทรัพยากรสารสนเทศ เช่น การกัดกินจากแมลง ภัยจากอัคคีภัย รวมถึงการใช้งานทรัพยากรสารสนเทศ อย่างระมัดระวัง ถูกวิธี เพื่อมิให้ถูกทำลาย สูญหาย หรือเสื่อมสภาพ

การอนุรักษ์ (Conservation) หมายถึง การซ่อมแซม การบำรุงรักษา การแก้ไข การเปลี่ยนแปลงโครงสร้างทางเคมี หรือโครงสร้างทางกายภาพของทรัพยากรสารสนเทศที่เกิดการชำรุด เสียหาย หรือเสื่อมสภาพ ให้อยู่ในสภาพที่ดีหรือ แข็งแรงขึ้น

การสงวนรักษาทรัพยากรสารสนเทศในรูปแบบดิจิทัล (Digital Preservation) คือ การประยุกต์ใช้เทคโนโลยีดิจิทัลเพื่อการอนุรักษ์ทรัพยากรสารสนเทศ หรือองค์ความรู้ที่สามารถ จับต้องได้ (Tangible) และองค์ความรู้ที่ไม่สามารถจับต้องได้(Intangible) โดยประกอบไปด้วย ๓ กระบวนการ คือ

- การแปลงข้อมูล (Digitization)
- การจัดเก็บและการจัดการข้อมูล (Information Management)
- การเข้าถึงข้อมูล (Information Access)

ประโยชน์ของการสงวนรักษาทรัพยากรสารสนเทศในรูปแบบดิจิทัล

### อนุรักษ์

ช่วยในการอนุรักษ์ข้อมูลสำคัญ โดยลดการใช้งานข้อมูลต้นฉบับที่ อ่อนไหวต่อการถูกทำลาย เสี่ยงต่อการเสียหายหรือสูญหายจาก การเข้าถึงได้โดยตรงจากผู้ใช้

### จัดการ

มีเครื่องมือช่วยในการบริหารจัดการข้อมูลที่มีจำนวนมากอย่างเป็นระบบระเบียบ และแลกเปลี่ยนหรือเผยแพร่ ข้อมูลผ่านระบบคอมพิวเตอร์และเครือข่ายอินเทอร์เน็ต

### ห้องสมุดดิจิทัล

ใช้ระบบการจัดเก็บในรูปแบบเทคโนโลยี OCR

### เทคโนโลยี OCR (Optical Character Recognition)

คือเทคโนโลยีการรู้จำอักขระ ที่สามารถนำมาใช้เพื่อสแกนข้อมูลต่าง ๆ บนรูปภาพหรือเอกสาร จากนั้นจะประมวลผลออกมาในรูปแบบของข้อความหรือตัวอักษร ซึ่งจะช่วยให้ผู้ใช้งานสามารถจัดการกับข้อมูลจากรูปภาพหรือเอกสารได้สะดวกมากยิ่งขึ้น ไม่ว่าจะในด้านการจัดเก็บข้อมูล หรือการนำข้อมูลไปวิเคราะห์ต่อก็ตาม ซึ่งเทคโนโลยี OCR นี้ เมื่อได้มีการนำ AI เข้ามาทำงานอยู่เบื้องหลัง จะยิ่งช่วยเพิ่มประสิทธิภาพในการทำงาน และเพิ่มความแม่นยำให้กับการอ่านและรู้จำอักขระต่าง ๆ

โปรแกรม OCR หรือซอฟต์แวร์ OCR ทำงานโดยมีขั้นตอนดังต่อไปนี้

### การรับภาพ

ตัวสแกนจะอ่านเอกสารและแปลงเป็นข้อมูลไบนารี จากนั้นซอฟต์แวร์ OCR จะวิเคราะห์ภาพที่สแกนและระบุส่วนที่สว่างเป็นพื้นหลัง และส่วนที่มีมืดเป็นข้อความ

### กระบวนการก่อนการประมวลผล

ซอฟต์แวร์ OCR จะทำความสะอาดรูปภาพก่อน และลบข้อผิดพลาดออกเพื่อเตรียมรูปภาพสำหรับการอ่าน โดยเทคนิคบางส่วนในการทำความสะอาดรูปภาพมีดังนี้:

- การปรับเอกสารให้ตรงหรือการเอียงเอกสารที่สแกนเล็กน้อยเพื่อแก้ไขปัญหาการจัดตำแหน่งระหว่างการสแกน
- การลบรอยขีดหรือรอยจุดในรูปภาพดิจิทัลออก หรือการปรับขอบของรูปภาพข้อความให้เรียบ
- การทำความสะอาดช่องและเส้นในรูปภาพ
- การรู้จำสคริปต์สำหรับเทคโนโลยี OCR แบบหลายภาษา

## การรู้จำข้อความ

อัลกอริทึมหรือกระบวนการด้านซอฟต์แวร์ OCR หลักสองประเภทที่ซอฟต์แวร์ OCR ใช้ในการรู้จำข้อความเรียกว่าการจับคู่รูปแบบและการแยกลักษณะ

### การจับคู่รูปแบบ

การจับคู่รูปแบบทำงานโดยการแยกภาพอักขระที่เรียกว่ารูปร่างอักษร และเปรียบเทียบกับรูปร่างอักษรที่จัดเก็บไว้ในลักษณะเดียวกัน การรู้จำรูปแบบจะทำงานได้ก็ต่อเมื่อรูปร่างอักษรที่จัดเก็บไว้มีแบบอักษรและมาตราส่วนใกล้เคียงกับรูปร่างอักษรที่ใช้ โดยวิธีการนี้ใช้ได้กับรูปภาพที่สแกนของเอกสารที่พิมพ์ด้วยแบบอักษรที่เป็นที่รู้จัก

### การแยกลักษณะ

การแยกลักษณะจะแบ่งหรือแยกย่อยรูปร่างอักษรออกเป็นคุณสมบัติต่างๆ เช่น เส้น วงปิด ทิศทางของเส้น และจุดตัดของเส้น จากนั้นจึงใช้คุณสมบัติเหล่านี้เพื่อค้นหาคู่ที่เหมาะสมที่สุดหรือตำแหน่งข้างเคียงที่ใกล้เคียงที่สุดในบรรดารูปร่างอักษรต่างๆ ที่จัดเก็บไว้

### กระบวนการหลังการประมวลผล

หลังจากการวิเคราะห์ ระบบจะแปลงข้อมูลตัวอักษรที่แยกออกมาเป็นไฟล์ที่ใช้ระบบคอมพิวเตอร์ ระบบ OCR บางส่วนสามารถสร้างไฟล์ PDF ที่มีคำอธิบายประกอบซึ่งมีทั้งเวอร์ชันก่อนและหลังของเอกสารที่สแกนได้

## ๙. ปัญหาอุปสรรคในการสัมมนา

### ๑๐. ประโยชน์ที่ได้รับจากการสัมมนา

- ต่อตนเอง ได้รู้ความรู้อีกเกี่ยวกับเทคโนโลยี OCR (Optical Character Recognition) ขั้นตอนและกระบวนการต่างๆที่ใช้จัดเก็บทรัพยากรให้อยู่ในรูปแบบข้อมูลดิจิทัล

- ต่อหน่วยงาน/มหาวิทยาลัย นำข้อมูลความรู้จากการสัมมนา มาวางแผนปรับปรุงขั้นตอนการทำงานของหน่วยงานให้มีประสิทธิภาพยิ่งขึ้น

### ๑๑. เอกสารหรืออื่น ๆ ที่เกี่ยวข้องที่ได้รับจากสัมมนา

### ๑๒. สำเนาประกาศนียบัตร/วุฒิบัตรฯ ที่ได้รับการประชุม/ฝึกอบรม/สัมมนา/ศึกษาดูงาน

### ๑๓. ความคิดเห็นและข้อเสนอแนะ อื่น ๆ

(ผู้รายงาน)..... กฤษฎีกา แก้วกรวง.....

(นายกฤษฎีกา แก้วกรวง)

วันที่ ๗ กุมภาพันธ์ ๒๕๖๖

ความคิดเห็นของผู้บังคับบัญชาชั้นต้น

.....  
.....  
.....  
.....  
.....

(ลงชื่อ). ..... *พช* .....

(อาจารย์ ดร.พิมพ์พลอย อีรสถิตย์ธรรม)

วันที่ ๗ กุมภาพันธ์ ๒๕๖๖