CS 2559/51

โครงงานคอมพิวเตอร์

ระบบจัดการโครงงานและงานวิจัยคอมพิวเตอร์

Computer Researches and Projects Management System

โดย

573020361-9 นายคมเคียว ตั้งประเสริฐ

573021086-0 นางสาวคุณัญญา ยุปาระมี

อาจารย์ที่ปรึกษา: ดร.นันท์นภัส เบญจมาศ

รายงานเล่มนี้เป็นส่วนหนึ่งของการศึกษาวิชา 322 391 ระเบียบวิธีวิจัย

ภาคเรียนที่ 2 ปีการศึกษา 2559

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์

มหาวิทยาลัยขอนแก่น

เดือน เมษายน พ.ศ. 2560

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

เนื่องจากในแต่ละปีนักศึกษาภาควิชาวิทยาการคอมพิวเตอร์ ซึ่งประกอบไปด้วย 3 สาขาวิชา ได้แก่ สาขาวิทยาการคอมพิวเตอร์ สาขาเทคโนโลยีสารสนเทศและการสื่อสาร สาขาภูมิสารสนเทศศาสตร์ ต้องจัดทำโครงงานคอมพิวเตอร์เพื่อวัดประสิทธิภาพก่อนการจบการศึกษา ภายใต้การควบคุมดูแลของอาจารย์ที่ปรึกษา และคณะกรรมการ เพื่อให้โครงงานเป็นไปตามเป้าหมายวัตถุประสงค์ของรายวิชา

ในปัจจุบัน มีเว็บไซต์การจัดการโครงงานและงานวิจัยคอมพิวเตอร์ ที่ใช้บันทึกจัดเก็บไฟล์เอกสารต่าง ๆ ของแต่ละโครงงาน แต่ระบบเดิมนี้ไม่สามารถรองรับความต้องการของผู้ใช้ และนโยบายของภาควิชาที่มุ่งเน้นการนำสำนักงานอิเล็กทรอนิกส์ (E-Office) ซึ่งใช้ระบบดิจิทัลเข้ามาทดแทนระบบเอกสารแบบเดิม เพื่อลดระยะเวลาในการปฏิบัติงาน และลดการใช้งานกระดาษ (Paperless)

ทางผู้จัดทำจึงเห็นสมควรว่าจะพัฒนาต่อยอดเว็บไซต์รวบรวมโครงงานคอมพิวเตอร์ ให้สามารถใช้แท็ก (Tag) ในการแยกหมวดหมู่ของโครงงานและงานวิจัย และสามารถค้นหาข้อมูลจากเนื้อหาของโครงงานได้สะดวกและแม่นยำมากขึ้น นอกจากนี้จะมีการพัฒนาระบบจัดการที่ปรึกษาโปรเจค การจัดกลุ่มสอบโปรเจค การสอบโปรเจค และเปลี่ยนการลงนามเอกสารไปเป็นการลงนามเอกสารแบบดิจิทัล (Digital Signature) ซึ่งมีคุณสมบัติทางด้านความปลอดภัย มีความน่าเชื่อถือและเป็นที่ยอมรับ อีกทั้งในปัจจุบันกฎหมายได้รับรองการลงนามเอกสารแบบดิจิทัลจึงสามารถใช้เป็นหลักฐานในทางกฎหมายได้ โดยเว็บไซต์จัดการโครงงานและงานวิจัยคอมพิวเตอร์นี้ พัฒนาขึ้นโดยใช้การค้นคืนสารสนเทศ และการลงนามเอกสารแบบดิจิทัลสำหรับเอกสารทั้งหมด เพื่อรองรับความต้องการของผู้ใช้และนโยบายของภาควิชา

1.2 วัตถุประสงคการวิจัย

1.2.1 เพื่อพัฒนาเว็บไซต์จัดการโครงงานและงานวิจัยคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ ที่สามารถรองรับการลงนามเอกสารแบบดิจิทัล

1.2.2 เพื่อพัฒนาเว็บไซต์จัดการโครงงานและงานวิจัยคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ ที่สามารถค้นหาข้อมูลจากเนื้อหาของโครงงานได้

1.3 ขอบเขตและขอจำกัด

1.3.1 เว็บไซต์รวบรวมเฉพาะโครงงานของนักศึกษาปี 4 ภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่นเท่านั้น

1.3.2 ระบบถูกใช้งานผ่านเว็บบราวเซอร์บนคอมพิวเตอร์และอุปกรณ์เคลื่อนที่

1.3.3 โครงงานที่ทำการรวบรวมเป็นโครงงานของนักศึกษา ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น เท่านั้น

1.3.5 เข้าใช้งานได้เฉพาะภาควิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยขอนแก่นเท่านั้น

1.3.6 หากไม่ได้ใช้บริการอินเตอร์เน็ตของมหาวิทยาลัยต้องเชื่อมต่อ VPN เพื่อใช้งาน

1.4 ประโยชนที่คาดวาจะไดรับ

1.4.1 ผู้ใช้สามารถลงนามเอกสาร ได้สะดวก รวดเร็วมากขึ้น โดยใช้การลงนามเอกสารแบบดิจิทัล

1.4.2 ผู้ใช้สามารถสืบค้นโครงงานออนไลน์ได้แม่นยำมากขึ้น โดยใช้การค้นหาเชิงความหมาย

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ลายเซ็นดิจิตอล (Digital Signature)

การลงนามดิจิตอล (Digital Signature) เป็นสิ่งที่ใช้ยืนยันตัวบุคคล (Identification) ในโลกดิจิตอล เปรียบเสมือนกับการเซ็นลายเซ็นลงนามเอกสารในความเป็นจริง ทำให้สามารถมั่นใจได้ว่าข้อมูลที่นั้นเป็นของใคร โดยข้อมูลจะถูกรับรองว่าเป็นของคนๆนั้นจริงๆ มีความน่าเชื่อถือ ไม่ถูกแก้ไขเพิ่มเติมภายหลังโดยบุคคลที่ไม่มีสิทธิ์ หากมีการแก้ไขจะสามารถตรวจจับได้ มีการแจ้งเตือนไปยังผู้ใช้หรือไม่สามารถเปิดเข้าไปดูข้อมูลได้ ซึ่งในปัจจุบันลายเซ็นดิจิตอลได้รับการยอมรับตามกฎหมายแล้ว จึงสามารถที่จะนำมาใช้เป็นหลักฐานในทางกฎหมายได้

โดยวิธีการในการลงนามแบบดิจิตอลที่ศึกษามีอยู่ 2 วิธี คือ Electronic Signature และ Digital Signature [1] ซึ่งมีความหมายและข้อแตกต่างกันดังนี้

2.1.1.1 Electronic Signature

เป็นการลงนามเอกสารโดยใช้สัญลักษณ์ หรือลายเซ็นที่อยู่ในรูปแบบอิเล็กทรอนิกส์ โดยเป็นการลงนามโดยบุคคลที่เป็นเจ้าของเอกสารหรือต้องการรับรองเอกสารนั้น สัญลักษณ์ที่นิยมใช้ได้แก่ รูปภาพลายเซ็นที่เซ็นด้วยหมึกปากกาลงในกระดาษแล้วสแกนเข้าสู่ระบบคอมพิวเตอร์ การใช้เมาส์ ลายนิ้วมือ stylus วาดรูปลายเซ็นบนหน้าจอคอมพิวเตอร์ ลายเซ็นที่แนบท้ายอีเมลล์ การพิมพ์ชื่อด้วยคีย์บอร์ด รูปภาพลายนิ้วมือ การคลิก “I Agree” ใน ข้อตกลง ต่างๆ เป็นต้น [2]

2.1.1.2 Digital Signature

Digital Signature นั้นมีจุดประสงค์เดียวกันกับ Electronic Signature เนื่องจากเป็นลายเซ็นที่อยู่ในรูปแบบของอิเล็กทรอนิกส์เหมือนกัน การลงนามก็จะใช้วิธีแบบเดียวกัน แต่มีการเพิ่มเติมคุณสมบัติทางด้านความปลอดภัยเข้าไป เพื่อให้มีความน่าเชื่อถือมากยิ่งขึ้น โดยคุณสมบัติดังกล่าวประกอบด้วย

1) Signer Authentication

เป็นคุณสมบัติที่ใช้ในการพิสูจน์ว่าใครเป็นคนลงนามเอกสารนั้น ตัวลายเซ็นจะมีเอกลักษณ์เฉพาะตัวที่สามารถใช้ในการเชื่อมโยงไปยังบุคคลที่ลงนามเอกสารได้

2) Data Integrity

เป็นคุณสมบัติที่ใช้ในการตรวจสอบ หรือพิสูจน์ว่ามีการแก้ไขเปลี่ยนแปลงเนื้อหาของเอกสารหลังจากที่ได้มีการลงนามไปแล้วหรือไม่ หากมีการแก้ไขก็จะทำให้เอกสารนั้นตกเป็นโมฆะ โดยจะมีการแจ้งเตือนไปยังผู้อ่าน หรือทำให้เอกสารนั้นไม่สามารถที่จะเปิดอ่านได้อีกต่อไป

3) Non-repudiation

การไม่สามารถปฏิเสธความรับผิดชอบได้ เนื่องจากลายเซ็นที่สร้างขึ้นมีเอกลักษณ์สามารถพิสูจน์ในชั้นศาลได้ว่าใครเป็นผู้เซ็นเอกสาร และในปัจจุบันลายเซ็นดิจิตอลได้รับการยอมรับตามกฎหมายแล้ว จึงสามารถที่จะนำมาใช้เป็นหลักฐานในทางกฎหมายได้

ตารางที่ 1 ข้อดีข้อเสียของ Electronic Signatures และ Digital Signature [1]

วิธีในการลงนามแบบดิจิตอล	ข้อดี	ข้อเสีย
Electronic Signature	- เป็นการใช้ สัญลักษณ์ หรือลายเซ็นที่อยู่ในรูปแบบอิเล็กทรอนิกส์ ที่มีเอกลักษณ์เฉพาะตัวที่ผู้ใช้สามารถเห็นได้ง่าย ทำให้ทราบได้ว่าใครเป็นเจ้าของลายเซ็น	– ถูกคัดลอกและปลอมแปลงไปใส่ในเอกสารอื่นได้ง่าย – ไม่สามารถตรวจสอบ หรือพิสูจน์ได้ว่ามีการแก้ไขเปลี่ยนแปลงเนื้อหาของเอกสารหลังจากที่มีการลงนามไปแล้วหรือไม่ –ไม่สามารถพิสูจน์ได้ว่าผู้ที่ลงนามในเอกสารนั้นเป็นตัวจริงหรือไม่ จึงโดนปฏิเสธความรับผิดชอบได้ และไม่ได้รับการรับรองตามกฎหมาย
Digital Signature	– เอกสารที่ผ่านการลงนามแล้วจะไม่สามารถแก้ไขได้ หากมีการแก้ไขก็จะสามารถตรวจสอบได้ – ผู้เซ็นเอกสารจะไม่สามารถปฏิเสธความรับผิดชอบได้ – สามารถใช้เป็นหลักฐานได้ตามกฎหมายได้ เทียบเท่ากับการเซ็นเอกสารในกระดาษด้วยหมึกปากกา	– ความปลอดภัยขึ้นอยู่กับการเข้ารหัสลับ - มีความยุ่งยากในการเชื่อมโยงการลงนามแบบดิจิตอลกับลายเซ็นที่แนบไปในเอกสาร

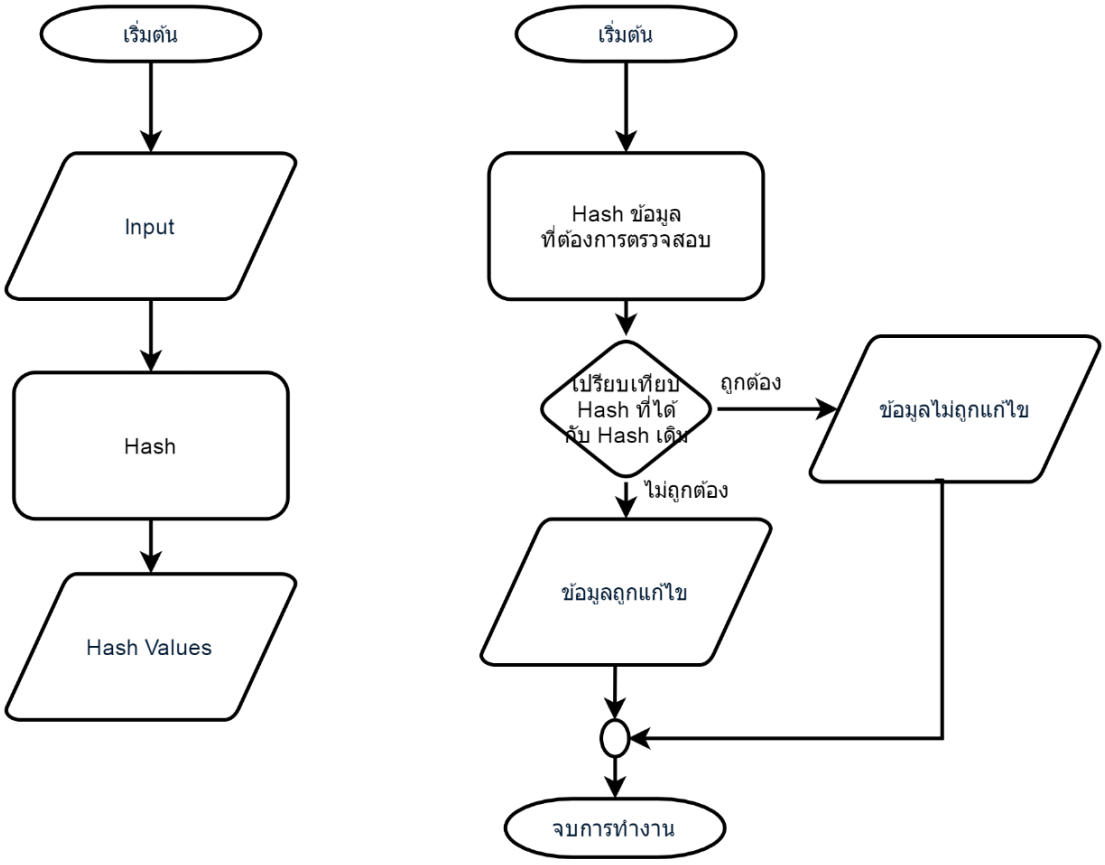
จากตารางที่ 1 จะเห็นว่าการลงนามเอกสารแบบอิเล็กทรอนิกส์ทั้ง 2 รูปแบบนั้น มีข้อดี และข้อเสียที่แตกต่างกันไป การเลือกรูปแบบที่จะมาใช้ในการพัฒนานั้นขึ้นอยู่กับว่า ความต้องการของระบบนั้นมีอะไรบ้าง เมื่อพิจารณาความต้องการของระบบจัดการโครงงานและงานวิจัยคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น พบว่าต้องการความปลอดภัยและน่าเชื่อถือ และจะปฏิเสธความรับผิดชอบไม่ได้จึงควรเลือกใช้รูปแบบ Digital Signature

2.1.2 การเข้ารหัสข้อมูล (Cryptography)

การเข้ารหัสข้อมูลโดยพื้นฐานแล้วจะเกี่ยวข้องกับวิธีการทางคณิตศาสตร์เพื่อใช้ในการปกป้องข้อมูลหรือข้อความตั้งต้นที่ต้องการส่งไปถึงผู้รับ ข้อมูลตั้งต้นจะถูกแปรเปลี่ยนไปสู่ข้อมูลหรือข้อความอีกรูปแบบหนึ่งที่ไม่สามารถอ่านเข้าใจได้โดยใครก็ตามที่ไม่มีกุญแจสำหรับเปิดดูข้อมูลนั้น เราเรียกกระบวนการในการแปรรูปของข้อมูลตั้งต้นว่า การเข้ารหัสข้อมูล (Encryption) และกระบวนการในการแปลงข้อความที่ไม่สามารถอ่านและทำความเข้าใจได้ให้กลับไปสู่ข้อความดั้งเดิม ว่าการถอดรหัสข้อมูล (Decryption) [3]

2.1.3 การตรวจสอบความถูกต้องของข้อมูล

เป็นการใช้ขั้นตอนวิธี (Algorithm) บางอย่างในการสร้างค่าสำหรับตรวจสอบความถูกต้อง (Integrity) ของข้อมูล ซึ่งถ้าหากข้อมูลถูกแก้ไขก็จะสามารถตรวจสอบได้ โดยการทำขั้นตอนวิธีเดิมซ้ำอีกครั้งจากนั้นนำค่าที่ได้มาเปรียบเทียบกัน หากไม่มีการแก้ไขข้อมูลค่าที่ได้ก็จะเป็นค่าเดียวกัน ดังภาพที่ 1 โดยขั้นตอนวิธีในการสร้างค่า Hash นั้นมีอยู่หลายวิธี ซึ่งแต่ระวิธีก็จะมีข้อดีและข้อเสียแตกต่างกันไป ยกตัวอย่างเช่น Hash, Message Digest (MD), SHA เป็นต้น [4]



ภาพที่ 1 แสดงขั้นตอนวิธีการตรวจสอบความถูกต้องของข้อมูล

2.1.4 ชุดรหัสผ่านแบบใช้งานครั้งเดียว (One Time Password)

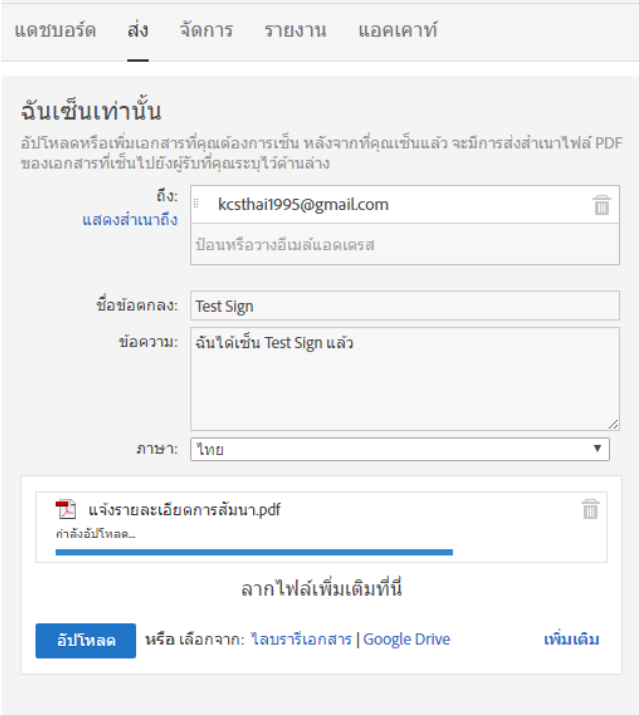
One Time Password (OTP) คือ รหัสผ่าน (Password) ในการเข้าสู่ระบบต่างๆ ที่ต้องการความปลอดภัยมากกว่าปกติ เช่นระบบชำระเงินออนไลน์ ระบบที่เกี่ยวข้องกับธนาคาร เป็นต้น โดยรหัสผ่านจะสามารถใช้ได้ครั้งเดียว และเป็นรหัสผ่านที่ถูกสร้างขึ้นให้ไม่ซ้ำกันโดยใช้ขั้นตอนวิธี หรือการสุ่ม (Random) ขึ้นมา ในปัจจุบันนิยมใช้ควบคู่กับการเข้าสู่ระบบด้วยรหัสผ่านแบบเดิมที่เป็นรหัสผ่านแบบตายตัว (Fixed Password) ซึ่งจะทำให้ระบบดังกล่าวมีระดับความปลอดภัยที่สูงขึ้น

ระบบ One-Time Password (OTP) มีหลายรูปแบบ เช่น อุปกรณ์แบบพกพา, Email และ SMS แต่ที่ได้รับความนิยมจะเป็นการใช้อุปกรณ์พกพาหรือ Token ในการสร้างรหัสผ่าน และการส่งรหัสผ่านไปยังโทรศัพท์มือถือที่ได้ลงทะเบียนไว้ [5]

2.1.5 ระบบ Adobe Sign

เป็นระบบการลงนามเอกสารออนไลน์ผ่านเว็บไซต์ที่ถูกพัฒนาโดย Adobe Systems Incorporated เป็นการลงนามที่สามารถนำไปใช้อ้างอิงได้ตามกฎหมาย โดยการใช้งานนั้น ผู้ใช้งานจะต้องทำการสมัครสมาชิกกับ Adobe ก่อนโดยข้อมูลที่ใช้สมัครนั้นควรจะเป็นข้อมูลจริง เพราะข้อมูลในส่วนนี้จะถูกนำไปเป็นส่วนประกอบของลายเซ็น หลังจากที่สมัครสมาชิกเรียบร้อยแล้ว จะมีให้ทดลองใช้ได้ 30 วัน หากครบกำหนดการทดลองใช้แล้วจะต้องเสียค่าใช้จ่ายในการใช้งานเป็นรายเดือน โดยจะมีขั้นตอนในการลงนามเอกสารดังต่อไปนี้

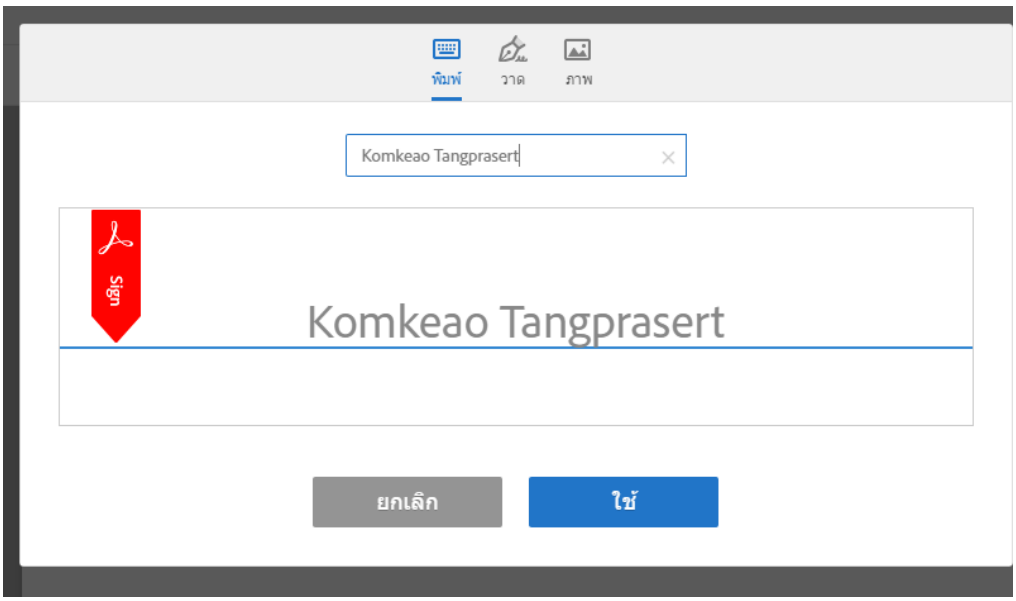
2.1.5.1 ผู้ใช้งานต้องทำการอัพโหลดเอกสารเข้าไปในระบบ ดังภาพที่ 2 โดยระบบจะให้ระบุ Email ของผู้ที่ต้องการให้ลงนามในเอกสารนั้น โดยจะสามารถระบุได้มากกว่า 1 คน ผู้ที่จะลงนามได้นั้น จะเป็นต้องสมาชิกกับทาง Adobe ด้วยเช่นกัน



ภาพที่ 2 ขั้นตอนการอัพโหลดเอกสารเข้าสู่ระบบ

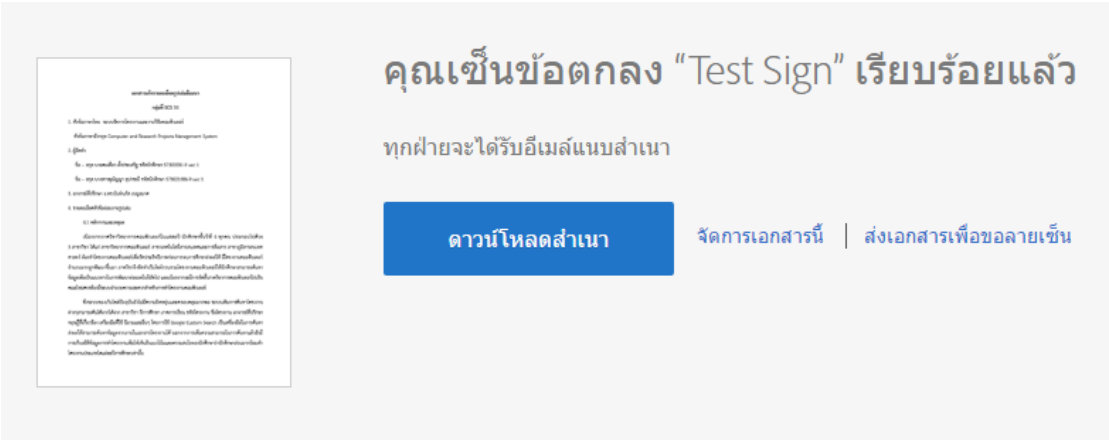
2.1.5.2 ระบบจะแสดงเนื้อหาของเอกสารที่อัพโหลดเสร็จแล้ว เพื่อให้ผู้ใช้งานเลือกตำแหน่งที่จะทำการเซ็นบนเอกสาร

2.1.5.3 เมื่อเลือกลงนามเอกสาร จะสามารถเลือกวิธีในการลงนามได้ 3 วิธี คือ การพิมพ์ข้อความเป็นตัวอักษร การใช้ภาพลายเซ็น และการวาดลายเซ็นเอง ดังภาพที่ 3



ภาพที่ 3 ขั้นตอนการสร้างรายเซ็นดิจิตอล

2.1.5.4 หากทำการลงนามเอกสารและตรวจสอบข้อมูลแล้ว ก็สามารถเลือกดาวโหลดเอกสารได้ทันที หรือในกรณีที่มีผู้ลงนามเอกสารหลายคน ก็จะต้องรอให้ทุกคนทำการลงนามก่อน ดังภาพที่ 4



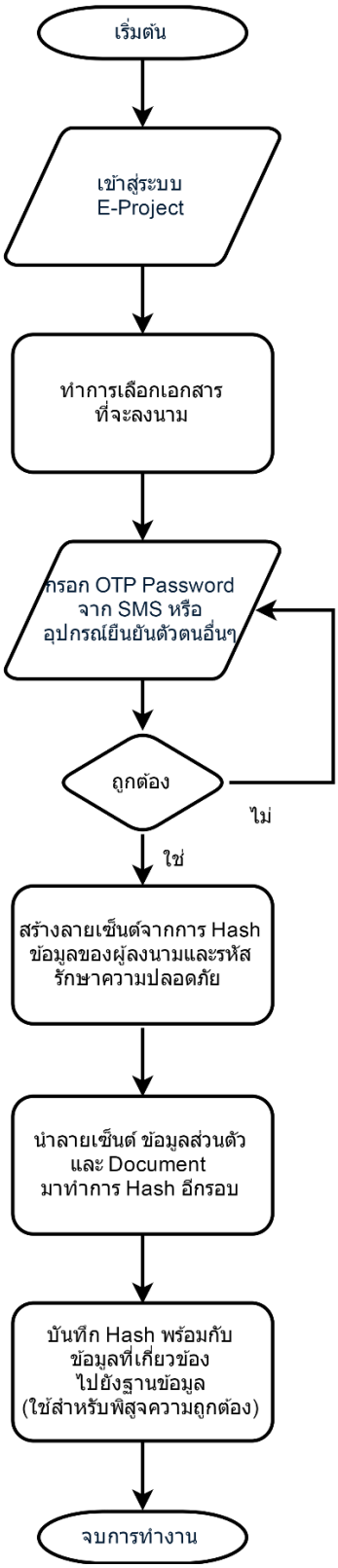
ภาพที่ 4 ขั้นตอนการดาวโหลดเอกสารที่ลงนามแล้ว

จะเห็นได้ว่าการลงนามเอกสารแบบดิจิตอลในระบบ Adobe Sign ของ Adobe Systems Incorporated นั้นมีความสะดวกสบายใช้งานง่าย เป็นมิตรต่อผู้ใช้ และยังคงคุณสมบัติทั้งสามข้อคือ Signer Authentication, Data Integrity และ Non-repudiation ไว้ได้อย่างครบถ้วน เนื่องจากสามารถเชื่อมโยงไปยังผู้ลงนาม เมื่อทำการลงนามไปแล้วก็จะไม่สามารถแก้ไขข้อความในเอกสารได้ รวมทั้งได้รับการยอมรับตามกฎหมายอีกด้วย

การลงนามดิจิตอล (Digital Signature) เป็นสิ่งที่ใช้ยืนยันตัวบุคคล (Identification) ในโลกดิจิตอล เปรียบเสมือนกับการเซ็นลายเซ็นลงนามเอกสารในความเป็นจริง ทำให้สามารถมั่นใจได้ว่าข้อมูลที่นั้นเป็นของใคร โดยข้อมูลจะถูกรับรองว่าเป็นของคนๆนั้นจริงๆ มีความน่าเชื่อถือ ไม่ถูกแก้ไขเพิ่มเติมภายหลังโดยบุคคลที่ไม่มีสิทธิ์ หากมีการแก้ไขจะสามารถตรวจจับได้ มีการแจ้งเตือนไปยังผู้ใช้หรือไม่สามารถเปิดเข้าไปดูข้อมูลได้ ซึ่งในปัจจุบันลายเซ็นดิจิตอลได้รับการยอมรับตามกฎหมายแล้ว จึงสามารถที่จะนำมาใช้เป็นหลักฐานในทางกฎหมายได้

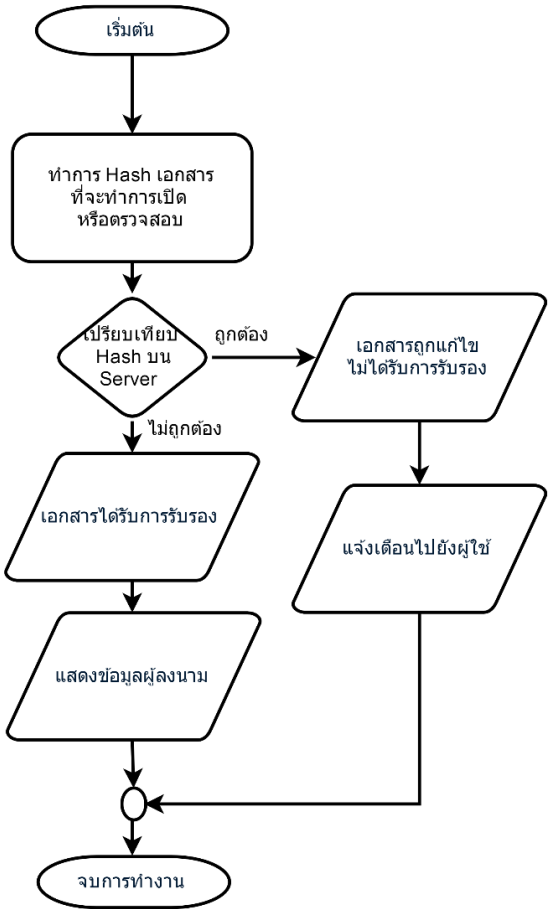
เนื่องจากการลงนามดิจิตอล (Digital Signature) เป็นขั้นตอนวิธีที่มีความน่าเชื่อถือ และมีความปลอดภัยในระดับที่สูง จึงเหมาะแก่การนำมาพัฒนาระบบจัดการโครงงานและงานวิจัยคอมพิวเตอร์ เพื่อที่จะเปลี่ยนการลงนามเอกสารทั้งหมดให้เป็นแบบดิจิตอล โดยได้มีการศึกษาขั้นตอนวิธีในการลงนามเอกสารและการตรวจสอบความถูกต้องของเอกสาร ดังภาพที่ 2 และ 3 เพื่อนำมาเป็นต้นแบบในการพัฒนาระบบต่อไป

ขั้นตอนการทำงานของการลงนามดิจิตอล จะใช้การยืนยันตัวตนของผู้ลงนามแบบสองชั้น โดยชั้นแรกคือการเข้าสู่ระบบด้วยชื่อผู้ใช้และรหัสผ่าน จากนั้นเมื่อต้องการลงนามเอกสารจะมีการใช้ OTP (One Time Password) เพื่อเพิ่มระดับความปลอดภัยอีกชั้นหนึ่ง ให้มั่นใจได้ว่าผู้ที่ลงนามนั้นเป็นตัวจริง เมื่อทำการยืนยันตัวตนเรียบร้อยแล้วก็จะเข้าสู่กระบวนการสร้างลายเซ็นต์ โดยจะใช้ข้อมูลส่วนตัวของผู้ลงนามและรหัสความปลอดภัยในการสร้าง จากนั้นจะแนบลายเซ็นไปกับเอกสารที่ลงนาม จากนั้นจึงสร้างค่า Hash ขึ้นมาเพื่อใช้ในการตรวจสอบความถูกต้องของข้อมูล และเก็บข้อมูลที่เกี่ยวข้องทั้งหมดไปยังฐานข้อมูลที่มีความปลอดภัยสูง ดังภาพที่ 5



ภาพที่ 5 ต้นแบบขั้นตอนวิธีในการลงนามเอกสารแบบดิจิตอล

ขั้นการตรวจสอบการลงนามและความถูกต้องของเอกสาร จะทำได้โดยการนำเอกสารที่ต้องการตรวจสอบมาทำการ Hash และนำค่าที่ได้ไปเปรียบเทียบกับค่า Hash ที่เก็บไว้ในฐานข้อมูล หากตรงกัน ก็จะสามารถสรุปได้ว่าข้อมูลนั้นได้รับการลงนามอย่างถูกต้องและไม่ถูกแก้ไข โดยจะใช้ขั้นตอนวิธีที่แสดงในภาพที่ 6



ภาพที่ 6 ต้นแบบขั้นตอนวิธีในการตรวจสอบความถูกต้องของเอกสารที่ได้รับการลงนาม

2.1.6 การค้นคืนสารสนเทศ (Information Retrieval)

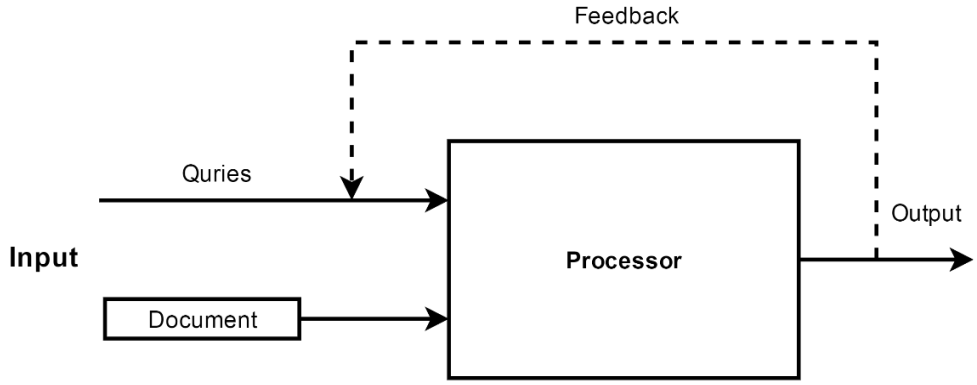
ระบบคนคืนสารสนเทศ (Information Retrieval System หรือ IR) เปนระบบที่จัดการประมวลผลสารสนเทศประเภทเอกสาร (Document) ในรูปแบบตางๆ เชน หนังสือ, วารสาร, บทความ เป็นต้น โดยเกี่ยวข้องในเรื่องการแสดงรูปแบบ, การเก็บบันทึก, การดึงเอกสาร โดยระบบค้นคืนสารสนเทศจะทำการสร้างตัวแทนขึ้นมาแทนข้อความในเอกสารทั้งหมดและเก็บไว้ในรูปแบบ index เพื่อให้นำข้อคำถามที่ผู้ใช้ต้องการค้นมาเปรียบเทียบถ้าเจอก็จะนำข้อมูลมาจัดลำดับ เพื่อให้ตอบสนองกับความต้องการของผู้ใช้มากที่สุด โดยแยกแยะความเกี่ยวข้อง (Relevant) และความไม่เกี่ยวข้อง (Non-relevant) [6]

2.1.6.1 ระบบค้นคืนสารสนเทศสามารถแบ่งได้ 3 ประเภทคือ

1) ระบบค้นคืนที่ให้คำถาม คำตอบ เป็นการบริการค้นคำตอบสำหรับคำถามที่ต้องการคำตอบที่เป็นข้อเท็จจริง

2) ระบบค้นคืนที่ให้ข้อมูลเป็นตัวเลขหรือสัญลักษณ์ เป็นระบบจัดเก็บข้อมูลทางฟิสิกส์ เคมี สำมะโนประชากร เป็นต้น

3) ระบบค้นค้นข้อความจากวารสาร เป็นระบบที่จัดเก็บตัวเนื้อหา เอกสารและสามารถเรียกข้อความส่วนใดส่วนหนึ่งของเอกสารได้ เช่น ฐานข้อมูลทางกฎหมายเป็นต้น [9]



ภาพที่ 7 แสดงส่วนประกอบของระบบค้นคืนสารสนเทศ [6]

2.1.6.2 ส่วนประกอบของระบบค้นคืนสารสนเทศดังภาพที่ 7 แบ่งได้เป็น 3 ส่วนได้แก่

1) ส่วนนำเข้าข้อมูล (Input) เป็นส่วนของการป้อนข้อคำถาม (query) จากผู้ใช้ซึ่งเป็นภาษาธรรมชาติหรืออาจเป็นการนำเข้า Metadata ซึ่งเป็นสารสนเทศเกี่ยวกับเอกสารหรืออาจไม่เป็นส่วนหนึ่งของเอกสารก็ได้แต่เป็นข้อมูลเกี่ยวกับข้อมูล (data about data) อาทิเช่น

ก) Descriptive Metadata เป็นการนำเขาสารสนเทศที่เป็นความหมายของเอกสารที่อยู่ภายนอก เช่น ผู้แต่ง(Author), ชื่อเรื่อง(Title), แหล่งที่มา (Source), วันที่ (Date), ISBN, สำนักพิมพ (Publisher), ความยาว (Length)

ข) Semantic Metadata Concerns The Content เป็นการนำเขาเนื้อความ ที่มีความหมายเชน บทคัดย่อ (Abstract), คำสำคัญ (Keywords), รหัสของหัวเรื่อง (SubjectCodes) ซึ่งอาจเปน Library of Congress หรือ Dewey Decimal หรือ UMLS (Unified Medical Language System) ก็ได

ค) เทอมของหัวเรื่อง (Subject terms) ซึ่งอาจมาจาก ontologies พิเศษเปนลำดับขั้นของเทอมมาตรฐาน (hierarchical taxonomies of standardized semantic terms)

ง) อาจเปนสารสนเทศของเว็บ (Web Metadata) ก็ได เชน META tag in HTML ระบบค้นคืนสารสนเทศจะนำสารสนเทศเหลานี้ผานการประมวลผลแบบเชื่อมโดยตรงกับระบบคอมพิวเตอร ซึ่งผูใชจะมีการโตตอบหรือปฏิสัมพันธกับระบบโดยตรง

2) โปรเซสเซอร (Processor) เปนส่วนของการประมวลผล ไดแก การจัดโครงสรางของสารสนเทศในรูปแบบที่เหมาะสม ประกอบดวย การสรางตัวแทนเอกสาร, การแบงแยกกลุมของเอกสาร, การจัดเก็บสารสนเทศ, การดึงขอมูลตามที่ผูใชตองการ การทำงานนั้นจะนำขอคำถามไปเปรียบเทียบกับตัวแทนเอกสารที่มีอยู เพื่อดึงเอกสารที่ใกล้เคียงนำออกมาให แกผูสอบถาม

3) สวนของผลลัพท (Output) ผลลัพธที่ไดจากระบบเป็นข้อความสั้นๆ เชน ชื่อหนังสือ, หมายเลขเอกสาร, ชื่อผูแต่ง, สำนักพิมพ เปนตน ผูใชสามารถพิจารณาจากขอมูลตางๆ ที่ไดจากระบบถาเอกสารที่ได มีจำนวนมากเกินไปหรือไมใกลเคียงกับสิ่งที่ตองการ ผูใชสามารถปรับปรุงขอคำถามใหมเพื่อใหขอคำถามนั้นสืบคนสารสนเทศไดตรงกับความตองการมากที่สุด เป็นระบบตอบกลับ (feedback) ดังนั้นผลลัพธ์ที่ไดจึงขึ้นอยูกับ ขอคำถาม (Query) ของผูใช้ [6]

2.1.6.3 หลักการทำงานของการค้นคืนสารสนเทศ

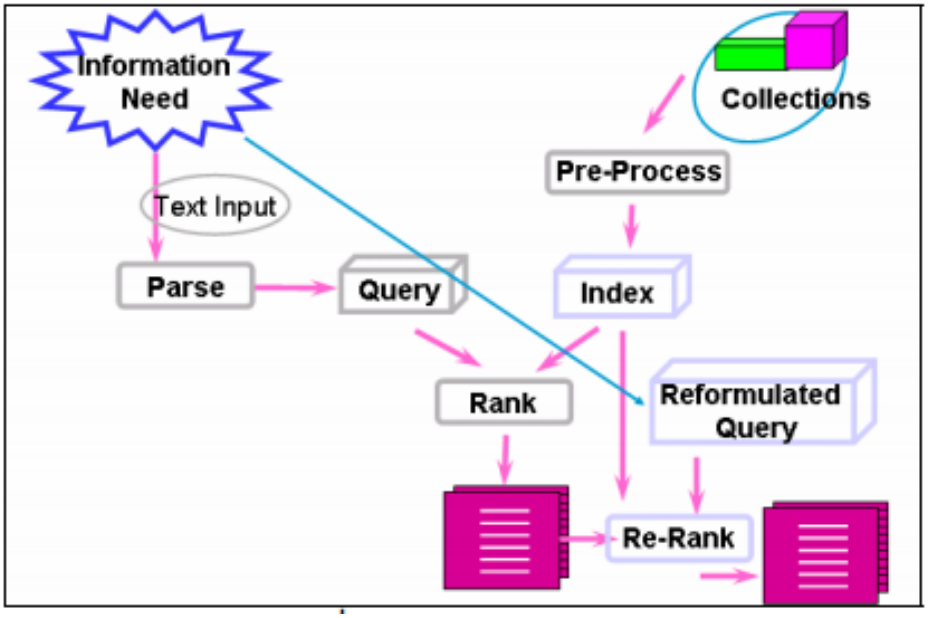
1) การคัดเลือก เป็นการการรวบรวมเอกสารตามเกณฑ์ และนโยบายที่กำหนดไว้ซึ่งสอดคล้องกับความต้องการของผู้ใช้งาน

2) การวิเคราะห์เอกสาร ได้แก่ การจัดหมวดหมู่ การจัดทำ รายการการทำดรรชนี และการทำสาระสังเขป

3) การจัดระเบียบแฟ้มข้อมูล

4) การค้นคืน วิเคราะห์แนวคิดและกำหนดศัพท์ หลังจากนั้นก็นำคำศัพท์ไปดำเนินการค้น ถ้าคำศัพท์ตรงกับดรรชนี คำค้นของเอกสารนั้นจะได้รับ เอกสารจำนวนหนึ่ง หรือผู้ค้นจะทำการปรับปรุงเอกสารให้เป็นที่พอใจของผู้ใช้บริการ

5) การแจกจ่าย เป็นการนำส่งผลการค้นคืนให้แกผู้ใช้ ที่มีความต้องการเอกสาร เอกสารในเรื่องนั้นๆ [9]



ภาพที่ 8 แสดงหลักการทำงานของการค้นคืนสารสนเทศ [9]

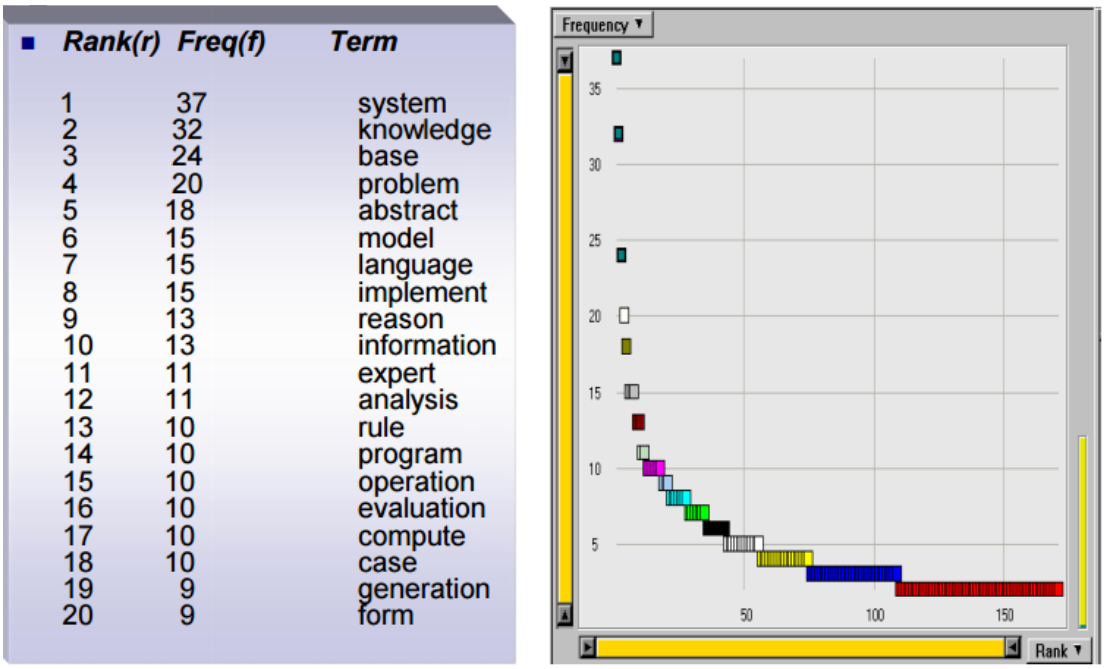
2.1.6.4 การสรางระบบคนคืนสารสนเทศ แบงออกเปน 4 ขั้นตอนคือ

1) การวิเคราะหขอความ (Text Analysis) เป็นวิธี การวิเคราะห์ข้อความให้ได้มาซึ่งตัวแทนเอกสาร แบ่งออกเป็น 2 วิธีคือ

ก) วิเคราะห์ทางด้านภาษาศาสตร์ วิธีการนี้ มีความยุ่งยากและเสียค่าใช้จ่ายสูงมาก

ข) วิเคราะห์ทางด้าน สถิติศาสตร์เป็นวิธีการที่นิยม

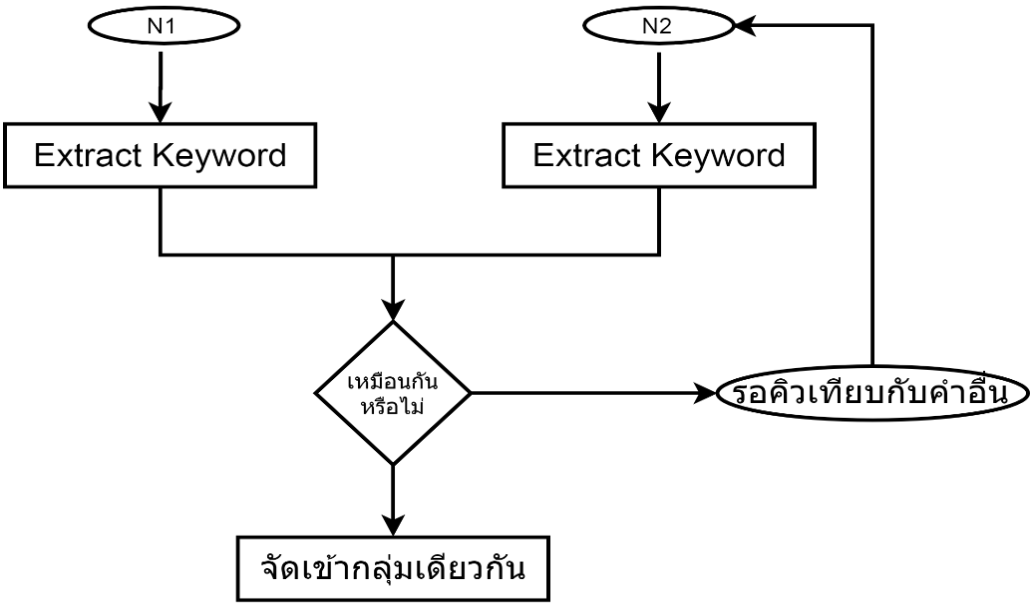
ซึ่งวิธีการนี้ได้มีการ ตรวจสอบและทดลองเพื่อให้สามารถนา มาประยุกต์ใช้กับ ระบบค้นคืนสารสนเทศ (IR) ได้ดีซึ้งได้ศึกษาหลัก การพื้นฐานโดยใช้ทฤษฎี ของลุนซ์ (Luhn) เข้ามาช่วยในการในการสร้างตัวแทนเอกสารโดยสมมุติให้ f เป็นความถี่ (Frequency) ของการเกิดขึ้นของคำใหม่ ในตำแหน่ง ต่างๆของข้อความ และให้ r เป็น Rank Order หรือลำดับของระดับ ของการเกิดขึ้นของคำๆ นั้นในเอกสาร และความสัมพันธ์ของ f และค่าของ f สูงจะส่งผลให้ r มีค่าต่ำ ดังภาพที่ 9 ดังนั้นความถี่ของข้อมูล (Frequency Data) สามารถถูกใช้เพื่อนำมาใช้วัด ความสำคัญของคำและประโยคที่ใช้แทนเอกสารหนึ่งได้ [10]



Rank(r)	Freq(f)	Term
1	37	system
2	32	knowledge
3	24	base
4	20	problem
5	18	abstract
6	15	model
7	15	language
8	15	implement
9	13	reason
10	13	information
11	11	expert
12	11	analysis
13	10	rule
14	10	program
15	10	operation
16	10	evaluation
17	10	compute
18	10	case
19	9	generation
20	9	form

ภาพที่ 9 แสดงการวิเคราะหขอความโดยใช้ทฤษฎีของลุนซ์ (Luhn) [10]

2) การจัดแบงกลุมขอมูล (Classification)

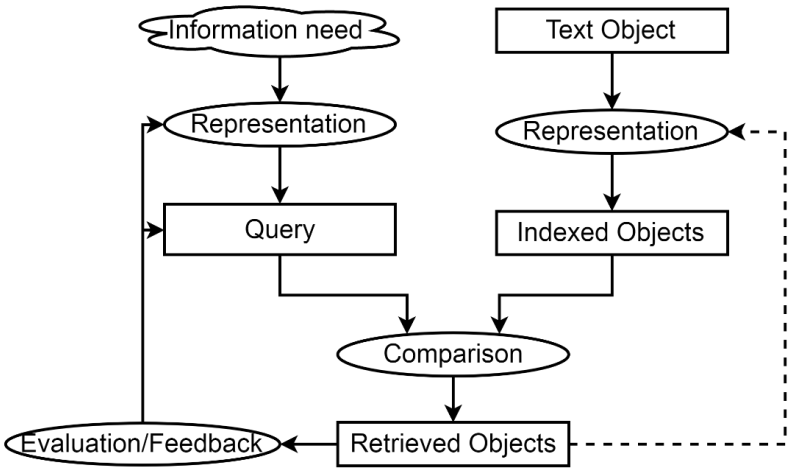


ภาพที่ 10 แสดงการจัดแบงกลุมขอมูล (Classification) [11]

เมื่อได้ตัวแทนเอกสารแล้ว จากนั้นเราจะนำตัวแทนเอกสารที่ได้ มาเปรียบเทียบกัน ว่าเหมือนกันหรือไม่ ถ้าเหมือนก็จะนำมาจัดเข้ากลุ่มเดียวกัน แต่ถ้าไม่เหมือนกันจะไปรอคิวเพื่อเทียบกับคำอื่นต่อไปดังภาพที่ 10

3) การเก็บบันทกข้อมูลลงในแฟมขอมูล เราก็จะนำตัวแทนเอกสารที่ได้มาจัดลงแฟ้มแทนข้อความของเอกสารทั้งหมดให้สมบูรณ์ให้อยู่ในรูปแบบดรรชนี (Index)

4) การคนคืนสารสนเทศ



ภาพที่ 11 แสดงการคนคืนสารสนเทศ

การค้นคืน เราจะนำข้อคำถาม (Query) ไปดำเนินการค้นในแฟ้ม (Indexed Objects) ว่ามีอะไรที่ตรงหรือใกล้เคียงกับตัวแทนเอกสารหรือดรรชนี (Indexed Objects) ไหม ถ้าข้อคำถาม (Query) กับดรรชนี (Indexed Objects) ตรงกัน คำระบบจะได้รับ เอกสารจำนวนหนึ่งนำม
าจัดลำดับและส่งไปแสดงให้ผู้ใช้ หรือผู้ใช้สามารถจะทำการปรับปรุง (Feedback) โดยสงความต้องการมาใหม่ ระบบจะทำการจัดลำดับใหม่ เพื่อให้เอกสารให้เป็นที่พอใจของ ผู้ใช้บริการตามภาพที่ 11

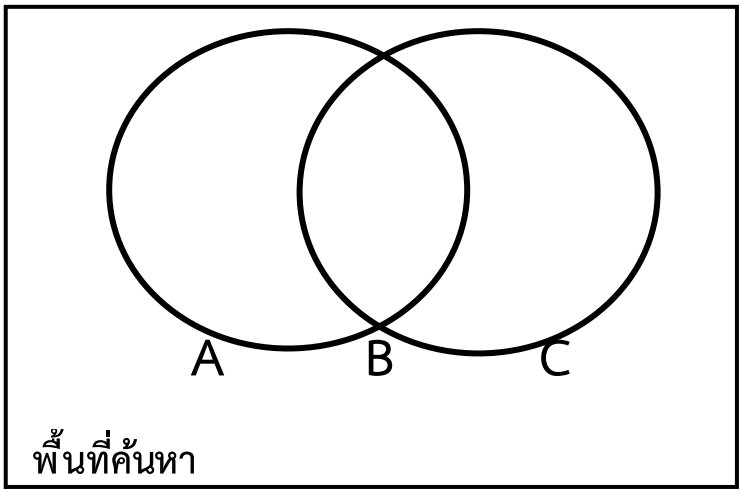
2.1.6.5 การประเมินผลระบบคนคืนสารสนเทศ (Evaluation of IR System)

ค่าความครบถ้วน (Recall) คือเป็นการวัดความสามารถของระบบในการดึงเอกสารที่เกี่ยวข้องออกมา โดยให้ ค่าของจำนวนเอกสารที่ถูกเรียกคืนออกมาได้และตรงความต้องการ หารด้วยจำนวนเอกสารที่ตรงความต้องการทั้งที่ถูกเรียกคืนและไม่ถูกเรียกคืน ซึ่งสามารถเขียนเป็นสูตรคณิตศาสตร์ [8] ได้ดังนี้

$$\text{Recall} = \frac{(|\{relevant\ document\} \cap \{document\ retrieved\}|)}{(|\{relevant\ document\}|)}$$

ส่วนค่าความแม่นยำ (Precision) คือเป็นการวัดความสามารถของระบบในการขจัดเอกสารที่ไม่เกี่ยวข้องออกไป โดยใช้ค่าของจำนวนเอกสารที่ถูกเรียกคืนและตรงความต้องการ หารด้วยจำนวนเอกสารที่ถูกเรียกคืนออกมาได้ทั้งหมดไม่ว่าจะตรงหรือไม่ตรงความต้องการ ซึ่งสามารถเขียนเป็นสูตรคณิตศาสตร์ได้ดังนี้

$$\text{Precision} = \frac{(|\{relevant\ documents\} \cap \{document\ retrieved\}|)}{(|\{relevant\ document\}|)}$$



ภาพที่ 12 ภาพอธิบายวิธีการคำนวณค่าความครบถ้วน และค่าความแม่นยำ [8]

จากภาพที่ 12 จะอธิบายวิธีการคำนวณค่าความครบถ้วน และค่าความแม่นยำดังนี้ วงกลม A+B คือ เอกสารที่ตรงกับความต้องการ วงกลม C+B คือ เอกสารที่ระบบเรียกคืนออกมาได้ ส่วน B คือ เอกสารที่ระบบเรียกคืนออกมาได้และตรงกับความต้องการ ค่าความครบถ้วน มีวิธีคิดโดย นำค่าจำนวนเอกสารที่ระบบเรียกคืนออกมาได้และ ตรงกับความต้องการ (B) หารด้วยผลรวมของจำนวนเอกสารที่ตรงกบความต้องการและจำนวนเอกสารที่ระบบเรียกคืนออกมาได้และตรงกับความต้องการ (A+B) ส่วนคาความแม่นยำ มีวิธีคิด โดยนำค่าจำนวนเอกสารที่ระบบเรียกคืนออกมาได้และตรงกบความต้องการ (B) หารด้วยผลรวมของเอกสารที่ระบบเรียกคืนออกมาได้และตรงกับความต้องการและเอกสารที่ระบบเรียกคืนออกมา (B+C) ซึ่งสามารถเขียนเป็นสูตรคณิตศาสตร์ได้ดังนี้

$$\text{ค่าความครบถ้วน} = \frac{B}{A+B}$$

$$\text{ค่าความแม่นยำ} = \frac{B}{B+C}$$

ค่าความครบถ้วน และค่าความแม่นยำจะเป็นค่าที่มีจำนวนเต็มเท่ากับหนึ่ง แต่เพื่อให้ทำความเข้าใจได้ง่ายขึ้น ผู้วิจัยจึงคูณเลขหนึ่งร้อยเพิ่มเข้าไปในการคำนวณค่าดังกล่าวเพื่อเลื่อนจุดทศนิยมขึ้น 2 ตำแหน่ง ทำให้มีจำนวนเต็มเท่ากับหนึ่งร้อย และเนื่องจากค่าความครบถ้วนและค่าความแม่นยำเป็นค่าที่มองจากมิติที่ต่างกัน หากต้องการผลที่สะท้อนค่าทั้งสองนี้จำเป็นจะต้องนำค่าทั้งสองไปคำนวณหาคาเฉลี่ยที่ เรียกว่า ค่าเอฟ-เมเซอร์ (F-measure) ซึ่งจะให้น้ำหนักของค่าความครบถ้วนและค่าความแม่นย่ำเท่าๆกัน ค่าเอฟ-เมเซอร์ จึงเป็นค่าที่นิยมใช้บอกประสิทธิภาพของระบบ ค่าเอฟ-เมเซอร์สามารถคำนวณได้ โดยนำค่าความครบถ้วนคูณด้วยค่าความแมนยำคูณด้วยสอง แล้วหารด้วยผลรวมของคาความครบถ้วนและค่าความแม่นยำ ซึ่งเขียนเป็นสูตรคณิตศาสตร์ได้ [8] ดังนี้

$$F = \frac{2 * (precision * recall)}{(\ precision + recall\ )}$$

ตารางที่ 2 ความแตกต่างระหว่าง Database และการค้นคืนสารสนเทศ [7]

	Database	IR
การค้นข้อมูล	ข้อมูลมีโครงสร้าง มีความหมายที่ชัดเจน	จะไม่ใช้โครงสร้างข้อมูล ส่วนใหญ่จะเป็น Metadata
การดึงข้อมูลข้อมูล	ตรงไปตรงมา	คลุมเครือ ไม่แน่ชัด (ตามธรรมชาติของภาษา)
ผลลัพท์ที่ได้	แน่นอนและถูกต้อง	เกี่ยวข้องเป็นบางครั้ง
การติดต่อกับระบบ	เป็นการ Query แบบ ครั้งเดียวจบ	มีการโต้ตอบและปรับปรุงผลลัพท์ ให้มีความถูกต้องมากขึ้น

จากตารางที่ 2 จะเห็นว่าการค้นคืนสารสนเทศ (IR) มีความยืดหยุ่นในการค้นหาข้อมูลมากกว่า Database เพราะเป็นการค้นหาข้อมูลจากลักษณะเด่นภายในเอกสาร และ Metadata นอกจากนี้ยังสามารถที่จะปรับปรุงการค้นหา ให้มีความถูกต้องมากยิ่งขึ้น

2.2 งานวิจัยเกี่ยวข้อง

การศึกษาค้นคว้าข้อมูลเพื่อใช้พัฒนาระบบจัดการโครงงานและงานวิจัยคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ได้เลือกศึกษาบทความทั้งหมดสามบทความ ได้แก่ Self-Contained Digitally Signed Documents: Approaching ‘ What You See Is What You Sign โดย H. Soderstrom ซึ่งจะพูดถึงการทำให้ลายเซ็นแบบดิจิทัลเชื่อมโยงกับลายเซ็นเสมือนที่สามารถมองเห็น และระบุตัวตนของผู้ที่ลงนามได้ และงานวิจัยที่สองได้แก่ Electronic signature based on digital signature and digital watermarking โดย L. Zhu and L. Zhu ซึ่งได้ประยุกต์ใช้ Time Stamp เข้ามาช่วยในการตรวจสอบความถูกต้องของเอกสาร และป้องกันการคัดลอกลายเซ็นไปใช้ในเอกสารที่ไม่ได้ลงนาม งานวิจัยที่สาม A PKI based timestamped secure signing tool for e-documents โดย S. Goswami, S. Misra และ M. Mukesh ซึ่งวิธีนี้จะนำ Time Stamp มาใช้ และการที่จะลงนามเอกสารได้จำเป็นต้องทำการติดตั้งใบรับรอง (Certificate) ในเครื่องก่อน โดยรายละเอียดจะนำเสนอต่อไป

2.2.1 Self-Contained Digitally Signed Documents

ในบทความนี้เสนอเกี่ยวกับการใช้งาน Digital Signature ซึ่งคือการลงนามเอกสารแบบดิจิตอล “สิ่งที่คุณเห็นคือลายเซ็นที่คุณเซ็นไป” [8] เป็นความท้าทายที่เป็นส่วนหนึ่งของการสร้างลายเซ็นดิจิตอลมาตั้งแต่แรกเริ่ม ลายเซ็น Digital ถูกนำไปใช้ในระดับบิท ทั้งนี้ ผู้ใช้งานจะเห็นในระดับที่สูงกว่านั้น ซึ่งเขาจะรู้ได้อย่างไรว่า เขาเซ็นอะไรไป การสุ่มตัวอย่างการนำไปใช้ในชีวิตจริงชี้ให้เห็นว่า ประเด็นดังกล่าวยังคงเป็นปัญหา ซึ่งบทความนี้จะนำเสนอวิธีที่จะทำให้มั่นใจว่า “สิ่งที่คุณเห็นคือลายเซ็นที่คุณเซ็นไป” ซึ่งได้ถูกพบจากหลักง่ายๆ คือ

1) เอกสารที่ได้รับการลงนาม คือ เอกสารที่มีการประทับลายเซ็นลงไป ทั้งเอกสารก็จะถือว่าได้รับการลงนามแล้ว

2) หลังจากลงนาม คู่กรณีก็ได้รับสำเนาเอกสารที่มีลายเซ็น พวกเขามีสิทธิที่จะจัดการเอกสารของเขา

3) โดยทั้งหมดนี้จะใช้ขั้นตอนวิธีของ Public Key - Cryptographic ซึ่งไฟล์ที่ใช้การลงนามจะเป็นไฟล์ PDF/A เนื่องจากว่าสามารถใช้ได้หลาย Platform อีกทั้งยังสามารถเก็บข้อมูลต่างๆ ไว้ในไฟล์เดียวกัน ไม่ว่าจะเป็น Font metadata และรายละเอียดอื่นๆ ซึ่งทำให้เราสามารถเก็บลายเซ็นเข้าไปในเอกสารนั้นได้เช่นกัน การตรวจสอบความถูกต้องของเอกสารว่าถูกแก้ไขไปแล้วหรือไม่นั้น จะใช้ค่า Checksum ในการตรวจสอบ โดยอัลกอริธึมที่ใช้จะเป็น SHA-256 โดยการเชคผ่านเว็บ นอกจากนี้ยังมีการเก็บ Log การลงนามต่างๆ ไว้แยกต่างหากจากไฟล์เอกสารด้วย

ระบบนี้จะทำงานบน Web Server ซึ่งมีหลายระบบเช่น การทำรายการธุรกรรมในธนาคาร การสมัครสวัสดิการต่างๆ การจ่ายภาษี เป็นต้น โดยจะมีการใช้รหัสผ่านอีกชุด (PIN) ในการยืนยันก่อนลงนามเอกสาร ซึ่งเราจะได้มาจาก card reader (smart card) โดยจะยกตัวอย่างเป็นการสมัครสวัสดิการจะมีขั้นตอนดังนี้

1) ใส่ข้อมูลพื้นฐาน เช่น วันเดือนปีเกิด บริการที่ร้องขอ (ใช้งานครั้งแรก)

2) ระบบจะแสดงข้อมูลโดยสรุปให้ตรวสอบ

3) กด Sign your request เพื่อลงนามในเอกสาร

4) รับรหัส PIN จาก card reader และกรอกไปยังหน้าเว็บ

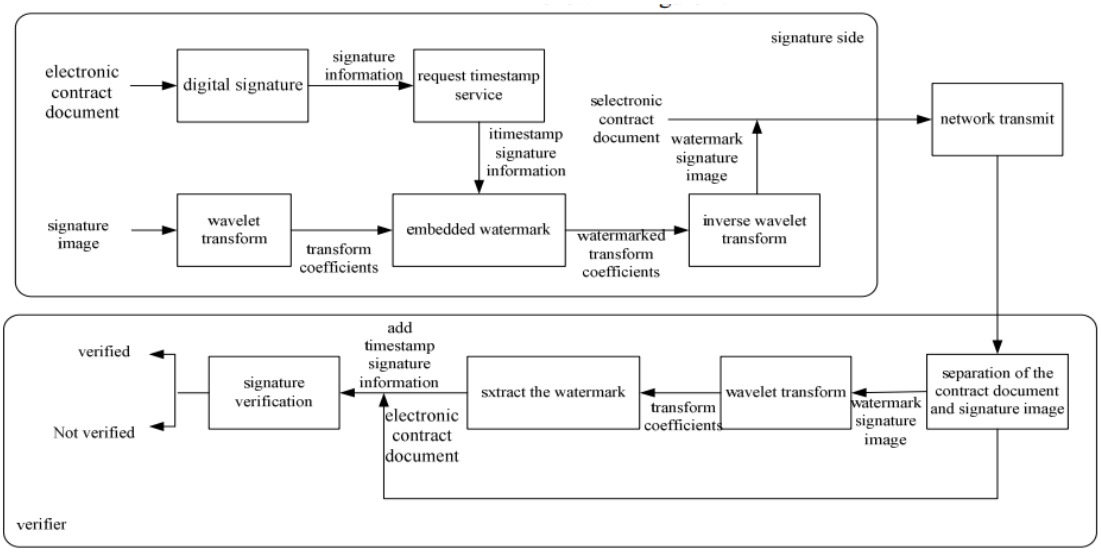
5) ระบบจะสร้างเอกสาร PDF ให้และทำการ Download และนำไปใช้งานต่อไป

ซึ่งวิธีการนี้สามารถใช้ได้จริงและถูกนำไปใช้ในระบบเว็บไซต์ของรัฐบาลที่เก็บข้อมูลในรูปแบบอิเล็กโทรนิค (e-government) ในเมืองหลวงที่สำคัญของสวีเดน

2.2.2 Electronic Signature Based on Digital Signature and Digital Watermarking

ในบทความนี้เสนอเกี่ยวกับการสร้างลายเซ็นดิจิทัล (Digital Signature) ด้วยเทคโนโลยี Digital Watermarking และ Timestamp โดยในการสร้างลายเซ็นดิจิทัล ขั้นตอนแรกจะทำการสร้างค่าที่ใช้สำหรับตรวจความถูกต้องจากขั้นตอนวิธี Hash จากนั้นค่าที่ได้จะถูกไปยัง Time Stamp Server เพื่อที่จะร้องขอการประทับเวลา (Time Stamp) สำหรับสร้างลายเซ็นดิจิทัล [9]

ขั้นตอนที่สองใช้ Wavelet Transformation ในการแปลงภาพลายเซ็นให้เป็นค่าสัมประสิทธิ์และนำ Time Stamp ที่ได้จาก Time Stamp Server แนบเข้าไป จากนั้นจะเข้าสู่กระบวนการ Inverse Wavelet Transform เพื่อแปลงกลับเป็นภาพลายเซ็นอีกครั้ง และนำไปแนบกับเอกสาร ซึ่งจะถูกเรียกว่า Signature Contract และถูกส่งไปยังคู่ค้า เพื่อพิสูจน์ตัวจริง (Authentication) ต่อไป ดังภาพที่ 13



ภาพที่ 13 แสดงกระบวนการสร้างและตรวจสอบลายเซ็นดิจิทัลจาก Electronic Signature Algorithm [9]

เมื่อคู่ค้าได้รับ Signature Contract ก็จะทำการแยกตัวเอกสารกับภาพลายเซ็นออกจากกัน จากนั้นจะใช้ Wavelet Transformation ในการแปลงภาพลายเซ็นให้เป็นค่าสัมประสิทธิ์ และทำการแยกข้อมูลออกจากค่าสัมประสิทธิ์

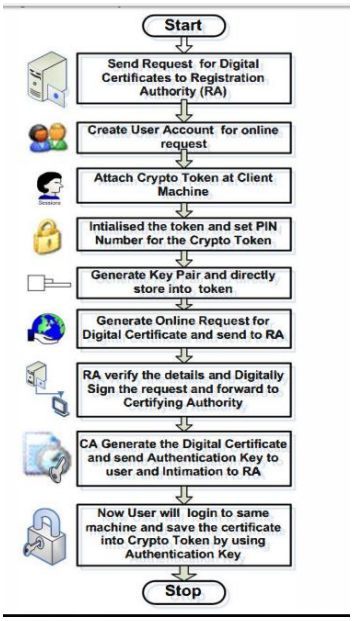
จากนั้นค่าสัมประสิทธิ์จะถูกแปลงให้เป็น Time Stamp Signature และใช้เทคโนโลยีพิสูจน์ตัวจริง (Authentication) หากลายเซ็นนี้ถูกตรวจสอบแล้วไม่พบความผิดปกติ จะสรุปได้ว่าไม่ถูกแก้ไขระหว่างทาง แต่หาก

ตรวจแล้วพบว่าไม่ถูกต้อง ก็จะสรุปได้ว่าเอกสารนี้ถูกแก้ไขระหว่างการส่งผ่านเครือข่าย และเอกสารนี้ก็จะไม่ได้รับการยอมรับ

จากการทดลองแสดงให้เห็นขั้นตอนวิธีนี้สามารถตอบสนองความต้องการของการลงนามในสัญญาธุรกรรมออนไลน์ได้เป็นอย่างดี โดยขั้นตอนวิธีนี้ไม่เพียงแค่ยืนยันตัวตน แต่ยังทำให้มั่นใจในความถูกต้องของสัญญาและไม่ยุ่งยาก

2.2.3 A PKI based Timestamped Secure Signing Tool for e-Documents

การแปลงเอกสารให้เป็นรูปแบบดิจิทัลทำให้เกิดการริเริ่มการวิจัยหลายอย่าง หนึ่งในนั่นคือกระบวนการของการรับรองและตรวจสอบความสมบูรณ์ของเอกสาร [10] ซึ่งลายเซ็นอิเล็กทรอนิกส์จะช่วยแก้ปัญหาได้ แต่ยังไม่สามารถพิสูจน์เอกลักษณ์ของผู้ลงนามได้ ใบรับรองดิจิทัลจึงถูกใช้หลังจากนั้นเป็นต้นมา มีลายเซ็นเพื่อพิสูจน์ตัวตนของผู้ลงนามในบทความนี้เขานำเสนอ Schema สำหรับการฝังลายเซ็นดิจิทัล รวมทั้งการรับรองและตรวจสอบเนื้อหาของเอกสารอิเล็กทรอนิกส์ในลักษณะที่ปลอดภัยและและป้องกันการปลอมตัว ลายเซ็นดิจิทัลถูกสร้างโดยอัลกอริทึมและการเข้ารหัส Private Key ของผู้ลงนามสุดท้าย เอกสารจะประทับตราเวลาโดยผเซิร์ฟเวอร์ Timestamp อีกขั้นตอนที่สำคัญในการลงนามแบบดิจิตอลคือการตรวจสอบความถูกต้องของลายเซ็นดิจิทัลเพื่อป้องกันจากเอกสารปลอมที่ใช้ลายเซ็นปลอมแปลงและเอกสารที่ดัดแปลง เนื้อหา การตรวจสอบใบรับรองสามารถทำได้โดยการตรวจสอบรายชื่อการเพิกถอนใบรับรอง (CRL) หรือสถานะใบรับรองออนไลน์ โปรโตคอล (OCSP) CRL ได้รับการอัปเดตและเมื่อผู้ใช้ดาวน์โหลดรายการและปรับปรุงข้อมูลในระบบของตนอย่างไรก็ตาม CRL อาจไม่สะท้อนถึงสถานะปัจจุบันของใบรับรอง OCSP ในเครื่องมืออื่น ๆ มักจะตรวจสอบความถูกต้อง ของใบรับรองออนไลน์ผ่านเซิร์ฟเวอร์ศูนย์กลางที่เชื่อถือได้ และจึงถือว่าน่าเชื่อถือมากขึ้น การยืนยันการโพสต์ค่าของเอกสารถูกส่งไปยัง Timestamp Authority สำหรับการสร้าง Timestamp ผู้มีเวลาประทับเก็บรักษาที่เก็บถาวรของ Timestamps ที่สร้างขึ้นโดยใช้มันสำหรับตรวจสอบ Timestamp ภายหลัง [10]



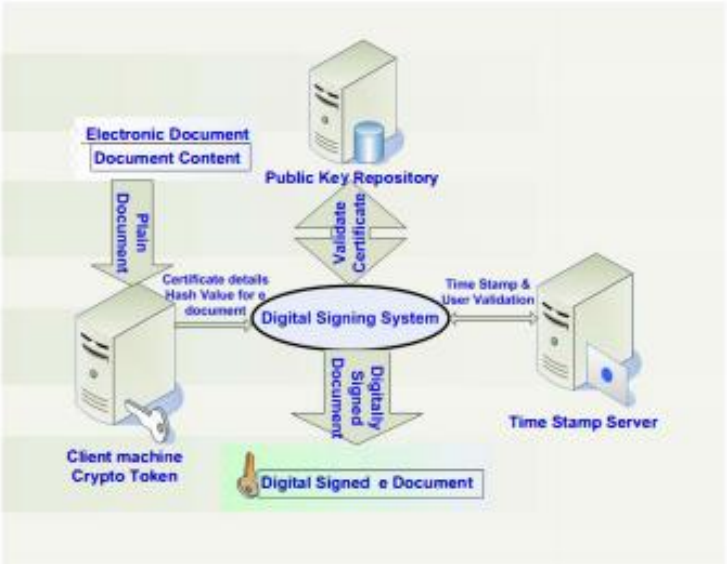
ภาพที่ 12 กระบวนการสร้างลายเซ็นดิจิทัล [9]

2.2.3.1 กระบวนการสร้างลายเซ็นดิจิทัล

1) ส่งคำขอใบรับรองดิจิทัลไปที่หน่วยลงทะเบียน
2) สร้างบัญชีผู้ใช้คำขอออนไลน์
3) แนบ Crypto Token ที่เครื่อง Client
4) เริ่มต้น Token และตั้งจำนวน PIN สำหรับ Token Crypto
5) สร้างคู่คีย์และส่งเก็บไว้ใน Token
6) สร้างคำขอออนไลน์สำหรับใบรับรองดิจิทัลและส่งถึง RA
7) RA ตรวจสอบรายละเอียดและแบบดิจิทัลลงชื่อเข้าใช้คำขอและส่งต่อไปที่เจ้าหน้าที่
8) CA สร้างใบรับรองดิจิทัลและส่งคีย์การรับรองความถูกต้องไปที่ผู้ใช้
9) ขณะที่ผู้ใช้จะเข้าสู่ระบบเครื่องเดียวกันและบันทึกใบรับรองลงใน Crypto

2.2.3.2 ขั้นตอนการลงนามแบบดิจิทัล

ขั้นแรกระบบจะระบุพารามิเตอร์ความปลอดภัยของลายเซ็นดิจิตอลที่จำเป็นสำหรับการลงชื่อเข้าใช้ เอกสารอิเล็กทรอนิกส์ พารามิเตอร์มีที่มาจากสองตำแหน่งที่แตกต่างกันคือ Token การเข้ารหัสลับและระบบการตรวจสอบจากส่วนกลาง Token รหัสลับให้ค่าคงที่เช่นการเซ็นชื่อใบรับรอง,ข้อมูลส่วนตัว, Private Key และ Public Key ไม่ได้เปลี่ยนแปลงด้วยเอกสารอิเล็กทรอนิกส์ใด ๆ การตรวจสอบจากส่วนกลาง ค่าของระบบเป็นแบบไดนามิกเช่น Timestamp Authority (TSA) ใบรับรองเซิร์ฟเวอร์, รายการ CRL และสถานะใบรับรองออนไลน์พิธีสาร (OCSP) ค่าเหล่านี้เปลี่ยนแปลงในขณะที่เซ็นชื่อแตกต่างกัน เอกสารอิเล็กทรอนิกส์ พารามิเตอร์เหล่านี้จะถูกป้อนเข้าสู่ระบบที่ สร้าง "Message Digest" ขึ้นอยู่กับข้อมูลเหล่านี้ ข้อความนี้ ส่วนย่อยและคีย์ส่วนตัวใช้เพื่อสร้าง Digital ลายมือชื่อ คีย์สาธารณะที่เกี่ยวข้องมีอยู่ใน Root CA เพื่อให้ผู้ใช้สามารถทดสอบความน่าเชื่อถือและความถูกต้องของผู้ลงนาม ระบบจะสร้างระบบดิจิตอล ลายเซ็นและแนบไปกับเอกสารต้นฉบับ เอกสารอิเล็กทรอนิกส์ที่ลงลายมือชื่อแบบดิจิทัลสามารถใช้งานได้ง่าย ตรวจสอบความถูกต้องและความซื่อสัตย์



ภาพที่ 13 ขั้นตอนการลงนามแบบดิจิทัล [9]

จากงานวิจัยข้างต้นทั้งสามงานวิจัย จะเห็นว่าสิ่งที่จำเป็นสำหรับลายเซ็นแบบดิจิทัล คือต้องสามารถตรวจสอบได้ว่าเอกสารที่ถูกลงนามไปแล้วนั้นถูกแก้ไขหรือไม่ หากมีการแก้ไขเอกสารนั้นก็จะไม่ถูกรับรองทันที และอีกประเด็นหนึ่งที่สำคัญคือ “What you what you see” [8] คือลายเซ็นที่เซ็นลงไปนั้น จะต้องสามารถมองเห็นได้ด้วยตา และระบุตัวตนของผู้ที่ลงนามได้ชัดเจน นอกจากนี้ได้มีการนำ Time Stamp เข้ามาช่วยในการตรวจสอบการลงนามอีกชั้นหนึ่ง ว่าเอกสารนั้นได้ถูกลงนามเมื่อไหร่ และจะต้องมั่นใจว่าลายเซ็นจะไม่ถูกคัดลอกไปใช้กับเอกสารอื่น ซึ่งจะเปรียบเทียบในแต่ละงานวิจัยไว้ใน ตารางที่ 3

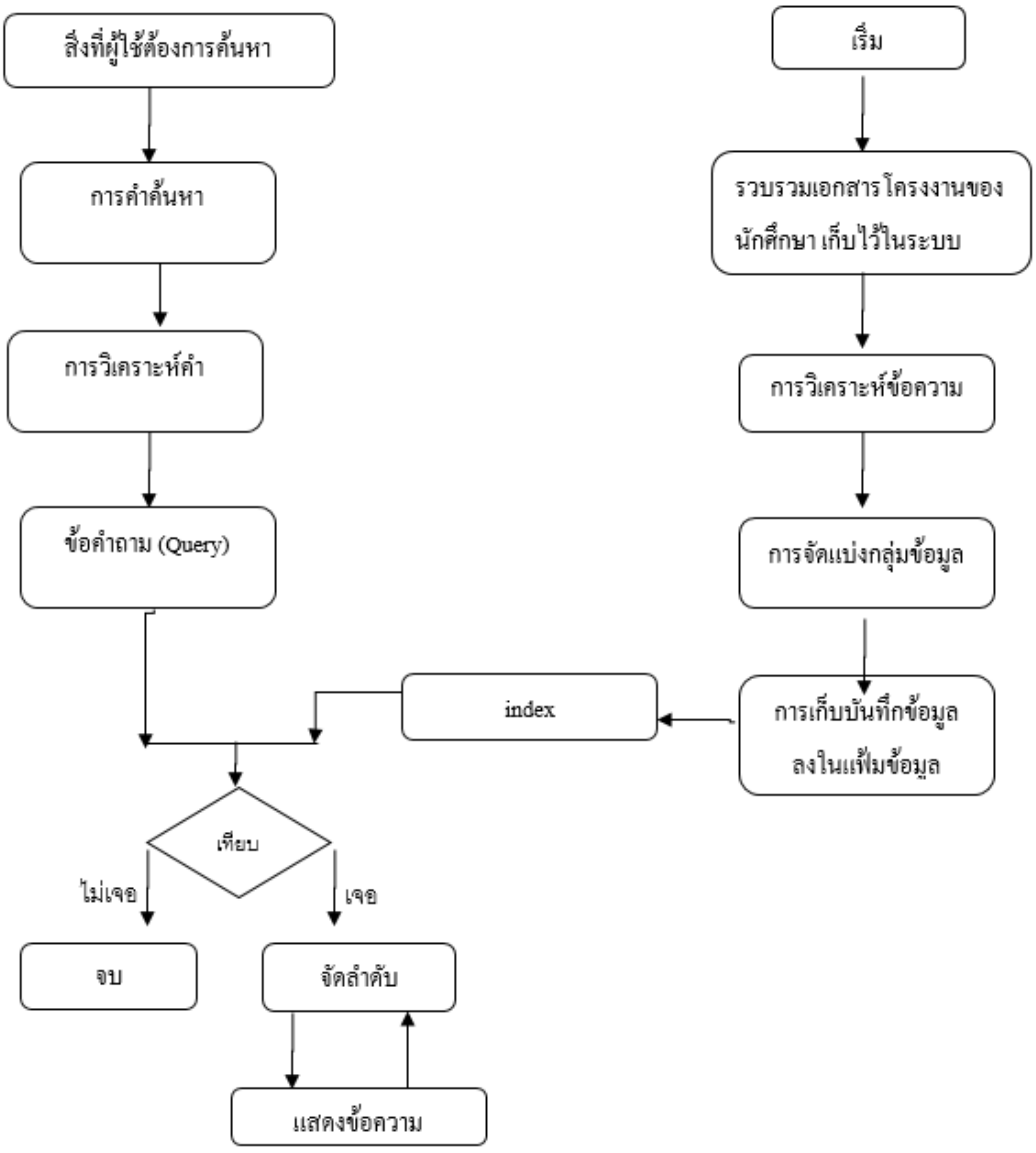
ตารางที่ 3 เปรียบเทียบงานวิจัยที่เกี่ยวของกับการลงนามเอกสารแบบดิจิทัล

ชื่องานวิจัย/วรรณกรรม	ใช้ TimeStamp เข้ามาช่วย	สามารถตรวจสอบการแก้ไขเอกสาร	มีความสะดวกในการใช้งาน	ลายเซ็นสามารถมองเห็นได้
Self-Contained Digitally Signed Documents		/	/	/
Electronic Signature Based on Digital Signature and Digital Watermarking	/	/	/	/
A PKI based Timestamped Secure Signing Tool for e-Documents	/	/		

จากการศึกษางานวิจัยเกี่ยวกับระบบค้นคืนสารสนเทศ (Information Retrieval System) หรือ IR สามารถนำไปเป็นแนวทางในการพัฒนาระบบจัดการโครงงาน และงานวิจัยคอมพิวเตอร์ในเรื่องของการค้นหาให้สะดวกมากขึ้น เนื่องจากมีจัดการโครงงาน และงานวิจัยคอมพิวเตอร์ที่ถูกเก็บไว้ในระบบมีจำนวนมาก โดยใช้การค้นคืนสารสนเทศ (Information Retrieval) เข้ามาช่วยดังนี้ 1. การวิเคราะห์ข้อความ (Text Analysis) เป็นการหาตัวแทนของเอกสารที่เหมาะสม เพื่อแทนการนำข้อความทั้งหมดในเอกสารไปเก็บในระบบ (ลดเวลาและค่าใช้จ่าย) 2. การจัดแบ่งกลุ่มข้อมูล (Classification) เป็นการจัดกลุ่มข้อมูลด้วยตัวแทนเอกสารที่ได้ 3. การเก็บบันทึกข้อมูลลงในแฟ้มข้อมูล เป็นการนำตัวแทนของเอกสารหรือดรรชนี (Index) ได้มาจัดเก็บแทนข้อความฉบับสมบูรณ์

การค้นคืนสารสนเทศ เป็นการเปรียบเทียบตัวแทนของเอกสารกับข้อคำถามของผู้ใช้ เพื่อวัดประสิทธิภาพ และประสิทธิผลของระบบเพื่อให้ผู้ใช้สามารถค้าหาได้ตรงความต้องการ โดยขั้นตอนนำมาปรับปรุงการทำงานของระบบค้นคืนโครงงานเป็นดังนี้ เริ่มจากเก็บรวบรวมเอกสารโครงงานของนักศึกษา เก็บไว้ในระบบ จากนั้นทำการวิเคราะห์ข้อคำถามเพื่อหาตัวแทนของเอกสาร นำตัวแทนเอกสารที่ได้มาจัดแบ่งกลุ่มข้อมูล นำไปเก็บ

ข้อมูลลงในแฟ้มเมื่อผู้ใช้ต้องการค้นหาเอกสารให้กรอกคำค้น นำคำที่ผู้ใช้กรอกเข้ามาวิเคราะห์คำ ได้ข้อคำถามนำไปค้นในดัชนี (index) ว่ามีคำที่เหมือนหรือใกล้เคียงหรือไม่ จากนั้นนำข้อมูลส่วนที่เหมือนหรือใกล้เคียงมาจำนวนหนึ่งจัดลำดับเพื่อให้ตรงกับความต้องการขอผู้ใช้ แสดงข้อมูลให้ผู้ใช้ และสุดท้ายผู้ใช้สามารถกรอกข้อคำถามเข้ามาใหม่ให้ตรงตามความต้องการ ระบบจะนำมาจัดลำดับใหม่เพื่อให้ตรงกับความต้องการของผู้ใช้งานในครั้งต่อไป



S

ภาพที่ 13 ต้นแบบขั้นตอนวิธีในการค้นคืนสารสนเทศ