

รายงานการวิจัย

โครงการห้องสมุดวิทยานิพนธ์อิเล็กทรอนิกส์
Electronic Thesis Library

จัดทำโดย

ดร.วิศิษฎ์ หิรัญกิตติ

นาย เกียรติณรงค์ ทองประเสริฐ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

บทคัดย่อ

โครงการนี้ศึกษาระบบห้องสมุดดิจิทัลเพื่อนำมาใช้จัดเก็บบทคัดย่อวิทยานิพนธ์ของนักศึกษาระดับบัณฑิตศึกษา สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยนำเสนอการจัดเก็บเมตาดาต้าของบทคัดย่อโดยใช้ ภาษา XML กำหนดไวยากรณ์เอกสาร และออกแบบแท็กที่ใช้ในการนำเสนอวิทยานิพนธ์เพื่อให้ข้อมูลที่จัดเก็บมีความหมายในตัวเอง ทำให้การค้นคืนข้อมูลสามารถทำได้โดยมีประสิทธิภาพและถูกต้องตรงตามความต้องการของผู้ใช้ ในส่วนของการจัดเก็บจะจัดเก็บไฟล์ XML ลงในฐานข้อมูลที่สนับสนุนเอกสาร XML ทำให้เข้าถึงเอกสารได้โดยง่ายและเป็นระบบ ส่วนการสืบค้นจะสืบค้นเข้าไปในเอกสาร XML แล้วทำการดึงเอกสาร XML ที่ต้องการนั้นออกมา ซึ่งทำการติดต่อผู้ใช้ผ่านทางระบบอินเทอร์เน็ต



RCH

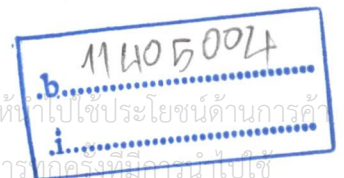
Z
678.9

0762 ค

เลขหมู่..... ๑๖ |

เลขทะเบียน... 54635

วัน,เดือน,ปี... 24 ส.ค. 2548



ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ผู้ดำเนินการวิจัยขอขอบคุณ บัณฑิตวิทยาลัย สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ในการสนับสนุนทุนอุดหนุนการวิจัยสำหรับโครงการนี้ รวมทั้งยังเอื้อเฟื้อสถานที่และอุปกรณ์เครื่องมือที่จำเป็นสำหรับงานวิจัยตลอดระยะเวลาของโครงการ

ท้ายที่สุด ผู้ดำเนินการวิจัยขอขอบพระคุณ บิดา มารดา อาจารย์ที่ปรึกษา (ดร.วิศิษฐ์ หิรัญกิตติ) คุณธีระพงศ์ เหมปรัชญกุล คุณชัยวัฒน์ ตรีวิรัชกร คุณวโรดม วีระพันธ์ และคุณจิตติมา อินทภรณ์ ที่มีส่วนช่วยให้งานวิจัยนี้สำเร็จลงได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อ	i
กิตติกรรมประกาศ	ii
สารบัญ	iii
บทที่ 1 บทนำ	1
1.1 ความสำคัญและที่มา	1
1.2 วัตถุประสงค์ของโครงการวิจัย	2
บทที่ 2 XML และภาษาที่ใช้กำหนดไวยากรณ์	3
2.1 XML	3
2.1.1 DTD	3
2.1.2 XML Schema	9
บทที่ 3 เมตาดาต้า และ โครงสร้างเมตาดาต้าที่ใช้แสดงวิทยานิพนธ์	13
3.1 Descriptive Metadata	13
3.2 โครงสร้างเมตาดาต้าที่ใช้แสดงวิทยานิพนธ์	13
บทที่ 4 การทดลอง	23
4.1 อุปกรณ์	23
4.2 วิธีการ	23
4.3 ขั้นตอนวิธี	28
บทที่ 5 ผลการทดลองและการวิเคราะห์	29
บทที่ 6 สรุปงานวิจัยและการวิจัยในอนาคต	32
ภาคผนวก ก ความหมายของแท็กที่ใช้เก็บวิทยานิพนธ์	33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1 บทนำ

1.1 ความสำคัญและที่มา

มหาวิทยาลัยเป็นแหล่งองค์ความรู้อันสำคัญของประเทศ มีองค์ความรู้จำนวนมากนับไม่ถ้วนที่เก็บอยู่ในวิทยานิพนธ์ของนักศึกษาในระดับบัณฑิตศึกษา และวิทยานิพนธ์นี้ถูกเก็บไว้ในห้องสมุดของมหาวิทยาลัยนั้นๆ ทำให้ไม่สะดวกในการค้นหาโดยเฉพาะสำหรับบุคคลภายนอกมหาวิทยาลัย ความรู้ที่มีค่าเหล่านี้สมควรที่จะนำมาทำประโยชน์ให้เกิดกับประชาชนในชาติ โดยเฉพาะอย่างยิ่ง จะช่วยในการแข่งขันของประเทศในด้านต่างๆ โดยเฉพาะในด้านการผลิตและด้านธุรกิจในสภาพโลกาภิวัตน์ที่มีการแข่งขันเสรีในทุกๆ เรื่อง

จะมีวิธีการใดที่จะนำเอาองค์ความรู้ดังกล่าว จากมหาวิทยาลัยออกเผยแพร่สู่โลกภายนอก โดยที่อาศัยทรัพยากรอย่างมีประสิทธิภาพมากที่สุด เพื่อให้เป็นประโยชน์ในการแลกเปลี่ยนความรู้เพื่อการค้นคว้าวิจัยที่จะเป็นประโยชน์ต่อภาคการผลิตและภาคบริการในประเทศเพื่อการแข่งขัน และเป็นประโยชน์ต่อประชาชนเพื่อเพิ่มพูนความรู้ประสบการณ์ตลอดจนทักษะอาชีพ

การสร้างห้องสมุดวิทยานิพนธ์ระดับบัณฑิตศึกษานบนเครือข่ายอินเทอร์เน็ต หรือที่เรียกว่า “ห้องสมุดวิทยานิพนธ์อิเล็กทรอนิกส์” นี้ เป็น โครงการที่นำเอาวิทยานิพนธ์ระดับบัณฑิตศึกษาของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังตั้งแต่อดีตจนถึงปัจจุบัน มาจัดเก็บให้อยู่ในรูปแบบอิเล็กทรอนิกส์ที่คอมพิวเตอร์สามารถใช้สืบค้นเนื้อหาภายในได้อย่างละเอียดถี่ถ้วนและครอบคลุม ซึ่งวิทยานิพนธ์ในรูปแบบอิเล็กทรอนิกส์นี้จะนำมาจัดเก็บรวมกันไว้เป็นเสมือนห้องสมุดจำลอง แล้วบันทึกเก็บไว้ในเครื่องคอมพิวเตอร์ที่มีขนาดหน่วยความจำสูง (ระดับเครื่องเซิร์ฟเวอร์) ที่เชื่อมต่อเข้ากับระบบเครือข่ายอินเทอร์เน็ต เมื่อดำเนินการแล้วเสร็จ ผู้ใช้คอมพิวเตอร์ไม่ว่าจะอยู่ที่ใดๆ ก็จะสามารถสืบค้นความรู้ที่ต้องการแสวงหาจากวิทยานิพนธ์ในห้องสมุดอิเล็กทรอนิกส์นี้ได้ และสามารถดึงเอาบางส่วนหรือทั้งหมดของวิทยานิพนธ์ไปเก็บไว้ที่เครื่องคอมพิวเตอร์เพื่ออ่าน หรือพิมพ์ออกเครื่องพิมพ์เพื่อเก็บไว้ค้นคว้าในภายหลังได้ บริการของห้องสมุดอิเล็กทรอนิกส์ทั้งหมดจะสามารถเรียกใช้จากเว็บไซต์ของห้องสมุด ที่จะจัดทำขึ้นมาให้ง่ายต่อการใช้งานของผู้ใช้และประชาชนโดยทั่วไป

การมีเว็บไซต์ห้องสมุดวิทยานิพนธ์อิเล็กทรอนิกส์ จะเป็นช่องทางสำคัญอันหนึ่งทำให้มหาวิทยาลัยได้มีการถ่ายทอดความรู้ไปสู่สังคมและชุมชนได้ นักศึกษาเองก็จะได้ตระหนักถึงประโยชน์ที่เขาจะได้สร้างให้กับสังคม จากผลที่ได้นี้ยังสามารถขยายไปสู่การถ่ายทอดความรู้ทางวิชาการ ศิลปะวัฒนธรรม ในรูปของเอกสารอิเล็กทรอนิกส์อื่นๆ รวมทั้งในรูปของสื่ออื่นๆ ที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถแปลงให้อยู่รูปแบบอิเล็กทรอนิกส์ ไปสู่ชุมชนภายนอกมหาวิทยาลัยได้ด้วย เช่น ตำรา
บทความวิชาการ เอกสารเผยแพร่ต่างๆ ภาพ เสียง ภาพยนตร์ ที่ให้ความรู้ เป็นต้น

ปัญหาที่เกิดขึ้นกับการสร้างระบบห้องสมุดเช่นนี้ที่จำเป็นต้องพึงพาการวิจัย ได้แก่

- การวิจัยเพื่อหาวิธีการแทนวิทยานิพนธ์ฉบับจริงด้วยรูปแบบที่เป็นเอกสาร
อิเล็กทรอนิกส์โดยอาศัยภาษาที่ใช้บรรยายรูปแบบของเอกสาร ได้แก่ ภาษา XML
- การวิจัยการจัดหมวดหมู่ของวิทยานิพนธ์เพื่อให้การสืบค้นต่างๆ ทำได้อย่างรวดเร็ว
- การวิจัยการจัดเก็บเอกสารให้ประหยัดเนื้อที่ในฐานข้อมูลคอมพิวเตอร์
- การค้นคิดระบบสืบค้นที่สามารถค้นหาบทคัดย่อและวิทยานิพนธ์ได้อย่างรวดเร็ว

1.2 วัตถุประสงค์ของโครงการวิจัย

1. เพื่อเป็นช่องทางหนึ่งให้มหาวิทยาลัยสามารถถ่ายทอดความรู้ไปสู่สังคมได้ ซึ่งเป็นหน้าที่
สำคัญประการหนึ่งของมหาวิทยาลัย
2. เพื่อเป็นบริการหนึ่งของบัณฑิตวิทยาลัยให้กับนักศึกษาและอาจารย์ระดับบัณฑิตศึกษา
เพื่อสนับสนุนการวิจัยของนักศึกษาและอาจารย์ใน สจล. และ ทั่วประเทศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2 XML และภาษาที่ใช้กำหนดไวยากรณ์

2.1 XML

XML ย่อมาจาก Extensible Markup Language คือ ภาษาที่ใช้บนเว็บรุ่นใหม่ และถูกพิจารณาให้เป็นมาตรฐานใหม่ แทนที่ Hypertext Markup Language (HTML) โดยตัวมันเป็นภาษาที่ใช้อธิบายข้อความในเอกสารเพื่อทำให้ข้อมูลนั้นมีโครงสร้าง โดยผู้ใช้สามารถกำหนดแท็กขึ้นมาใช้ตัวเอง ในการกำหนดแท็กขึ้นมาใช้เองนั้นจะต้องมีการประกาศโครงสร้างของแท็กขึ้นมาก่อน โดยใช้ DTD หรือ XML Schema

2.1.1 DTD

DTD ย่อมาจาก Document Type Definition ซึ่งทำหน้าที่เป็นเหมือนไวยากรณ์ของเอกสาร XML ที่กำหนดขึ้น แต่ในปัจจุบันมีการใช้น้อยลงเนื่องจากมีข้อจำกัดมาก แต่ก็ยังมีใช้อยู่บ้างในภาษาเช่น WML และ แอปพลิเคชันบางตัว เช่น DB2 XML Extender ที่ใช้ DTD ในการกำหนดโครงสร้างของ เอกสาร XML ที่จะใช้งาน

2.1.1.1 ลักษณะของเอกสาร DTD

ตัวอย่างด้านล่างแสดง DTD อย่างง่าย ๆ ของเอกสาร XML ที่ใช้เก็บข้อมูลแคตตาล็อกหนังสือ

```
<!ELEMENT BookCatalogue (Book)*>
<!ELEMENT Book (Title, Author, Date, ISBN, Publisher)>
<!ELEMENT Title (#PCDATA)>
<!ELEMENT Author (#PCDATA)>
<!ELEMENT Date (#PCDATA)>
<!ELEMENT ISBN (#PCDATA)>
<!ELEMENT Publisher (#PCDATA)>
```

จาก DTD ข้างต้นสามารถสร้างเอกสาร XML ที่เป็นไปตามข้อกำหนดข้างต้นได้ดังนี้

```
<?xml version="1.0"?>
<BookCatalogue>
  <Book>
    <Title>Modern Information Retrieval</Title>
    <Author>Ricardo Baeza-Yates</Author>
    <Date>1999</Date>
    <ISBN>0-201-39829-9</ISBN>
    <Publisher>ACM Press</Publisher>
  </Book>
</BookCatalogue>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.2 การสร้าง DTD

การสร้าง DTD แบ่งออกเป็น 3 หัวข้อหลัก ๆ คือ การประกาศค่าอิลิเมนต์, การประกาศค่าแอตทริบิวต์ และการประกาศค่าเอ็นทิตี

2.1.1.2.1 การประกาศค่าอิลิเมนต์

รูปแบบการประกาศค่าอิลิเมนต์เป็นดังนี้

```
<!ELEMENT ชื่ออิลิเมนต์ (เนื้อหาภายในอิลิเมนต์)>
```

เนื้อหาภายในอิลิเมนต์หนึ่ง ๆ อาจเป็นได้ 3 อย่างคือ อาจเป็นอิลิเมนต์อื่น ๆ, ข้อความปกติ หรืออาจจะไม่มีอะไรอยู่เลย (เป็นแท็กว่าง) หรือบางครั้งอาจจะมีอิลิเมนต์ปนอยู่กับข้อความปกติได้

ตัวอย่างแบบต่าง ๆ

- แบบมีอิลิเมนต์อื่น ๆ อยู่ภายใน เช่น

```
<Book>
  <Title>Modern Information Retrieval</Title>
  <Author>Ricardo Baeza-Yates</Author>
</Book>
```

ในตัวอย่างนี้อิลิเมนต์ Book มีอิลิเมนต์ Title และ Author อยู่ภายใน

- แบบที่มีข้อความอยู่ภายใน เช่น

```
<Publisher>ACM Press</Publisher>
```

- แบบไม่มีอะไรอยู่ หรือ แท็กว่าง เช่น

```
<Book isbn="975-85421-9-6"/>
```

- แบบที่มีอิลิเมนต์ปนอยู่กับข้อความ เรียกว่า “Mixed content” เช่น

```
<Message> Hello my XML
  <Note>See again</Note>
</Message>
```

ต่อไปจะกล่าวถึงวิธีประกาศ DTD

- กรณีที่อยู่ในอิลิเมนต์นั้นเป็นอะไรก็ได้ให้ระบุด้วย ANY

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรณีนี้จะใช้เมื่อไม่บังคับว่าข้างในอิลิเมนต์เป็นอะไร นั่นคือจะมีอะไรก็ได้

```
<!ELEMENT ชื่ออิลิเมนต์ ANY>
```

➤ ข้อความธรรมดาให้ระบุด้วย PCDATA

กรณีที่ต้องการบังคับให้เนื้อหาภายในอิลิเมนต์เป็นข้อความธรรมดาเท่านั้น

```
<!ELEMENT ชื่ออิลิเมนต์ (#PCDATA)>
```

โดยที่ PCDATA ย่อมาจาก Parsed Character Data หมายความว่าเนื้อหาส่วนที่เป็นข้อความนี้ เอาไว้ให้โปรแกรม XML Parser อ่านเพื่อประมวลผล

➤ อิลิเมนต์ว่าง ๆ ระบุด้วย EMPTY

กรณีของอิลิเมนต์ว่าง ให้ประกาศตามรูปแบบดังนี้

```
<!ELEMENT ชื่ออิลิเมนต์ EMPTY>
```

➤ การใช้ลิสต์ลำดับและตัวเลือก

การประกาศแบบเรียงลำดับจะใช้ (,) (คอมมา) เพื่อแสดงลำดับของอิลิเมนต์ และมีอย่างละ 1 อิลิเมนต์เท่านั้น และไม่มีไม่ได้ รูปแบบการประกาศเป็นดังนี้

```
<!ELEMENT ชื่ออิลิเมนต์ (อิลิเมนต์ 1, อิลิเมนต์ 2, อิลิเมนต์ 3)>
```

การประกาศแบบตัวเลือกจะใช้ (|) (ไปป์) คั่นระหว่างอิลิเมนต์ หมายความว่า ให้เลือกเพียงอิลิเมนต์ใดอิลิเมนต์หนึ่งเท่านั้น รูปแบบการประกาศเป็นดังนี้

```
<!ELEMENT ชื่ออิลิเมนต์ (อิลิเมนต์ 1 | อิลิเมนต์ 2 | อิลิเมนต์ 3)>
```

การประกาศเมื่อภายในอิลิเมนต์นั้นมีทั้งเนื้อหาและข้อความ รูปแบบการประกาศเป็นดังนี้

```
<!ELEMENT ชื่ออิลิเมนต์ (#PCDATA, อิลิเมนต์ 1, อิลิเมนต์ 2)>
```

➤ การประกาศอิลิเมนต์ โดยระบุจำนวน

ใน XML มีเครื่องหมายพิเศษ 3 แบบ คือ +, * และ ? ใช้เพื่อแสดงว่าอิลิเมนต์นั้นจะปรากฏได้กี่ครั้ง โดยแต่ละอันมีความหมายดังนี้

○ + หมายถึง จะต้องใช้อิลิเมนต์นั้นตั้งแต่ 1 อิลิเมนต์ขึ้นไป (คือ 1, 2, 3)

○ * หมายถึง จะมีอิลิเมนต์กี่อิลิเมนต์ก็ได้ หรือไม่มีก็ได้ (คือ 0, 1, 2)

เอกสารนี้เป็นเอกสารทสวงนวิธานหรับการเขางานเพื่อการศึกษาเท่านั้น เมื่อนูญาดเห็นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ? หมายถึง จะต้องมียุติเมนต์นั้นเพียง 1 ยุติเมนต์เท่านั้น หรือไม่มีก็ได้ (คือ 0, 1)

2.1.1.2.2 การประกาศค่าแอตทริบิวต์

การประกาศแอตทริบิวต์มีรูปแบบมาตรฐานดังนี้

```
<!ATTLIST ชื่ออิลิเมนต์ ชื่อแอตทริบิวต์ ชนิดข้อมูลของแอตทริบิวต์ (#REQUIRED | #IMPLIED | #FIXED) ค่าปกติของแอตทริบิวต์>
```

ตัวอย่าง

```
<!ATTLIST Document lang CDATA #FIXED "TH" >
```

Document คือ ชื่ออิลิเมนต์

lang คือ ชื่อแอตทริบิวต์

Document lang คือ การประกาศแอตทริบิวต์ชื่อ lang ซึ่งอยู่ภายในอิลิเมนต์

Document

CDATA คือการประกาศชนิดของค่าแอตทริบิวต์ ว่าเป็นข้อมูลประเภทตัวอักษร

#FIXED "TH" บอกให้ทราบว่า กำหนดค่าแอตทริบิวต์ไว้ตายตัวเป็นคำว่า "TH"

ชนิดของแอตทริบิวต์

ชนิดของแอตทริบิวต์	คำอธิบาย
CDATA	ข้อมูลตัวอักษร
enumerate คือ (ค่าที่ 1 ค่าที่ 2 ...)	enumerate หมายถึงชุดข้อมูลที่มีให้เลือกหลายค่า แต่ต้องใช้เพียงค่าเดียว โดยมีเครื่องหมายไปป์คั่นระหว่างค่าแต่ละค่าที่ให้เลือก ข้อมูลชนิดนี้แตกต่างจากชนิดอื่น ๆ คือ ไม่ได้ระบุชื่อชนิดว่า enumerate ลงไปโดยตรง แต่ระบุรายชื่อข้อมูลที่มีให้เลือก ลงไปเลย ในขณะที่ชนิดอื่น ๆ จะต้องระบุชื่อชนิดลงไป
ID	เป็นข้อมูลชนิดข้อความ ซึ่งใช้สำหรับกำหนดรหัสหรือหมายเลขประจำตัวของอิลิเมนต์ ดังนั้นค่า ID ของแต่ละอิลิเมนต์จะต้องมีค่าเดียวและไม่ซ้ำกัน (unique)
IDREF	เป็นข้อมูล ID ของอิลิเมนต์อื่น ซึ่งเอาไว้ระบุเพื่ออ้างถึงอิลิเมนต์อื่น เพียงหนึ่งอิลิเมนต์เท่านั้น
IDREFS	เป็นข้อมูล ID ของอิลิเมนต์อื่น ซึ่งเอาไว้ระบุเพื่ออ้างถึงอิลิเมนต์อื่น ได้ตั้งแต่หนึ่งอิลิเมนต์ขึ้นไป (IDREF เป็นสับเซตของ IDREFS)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	แล้วแต่จะระบุ แต่ถ้าไม่ระบุค่าไว้ XML Parser จะกำหนดให้มีค่าเท่ากับค่าปกติของแอตทริบิวต์เลย
--	---

2.1.1.2.3 การประกาศค่าเอ็นทิตี

เอ็นทิตี จะใช้สำหรับอ้างอิง “ทรัพยากร” หรือข้อมูลที่นิยามไว้ก่อน ทำให้สามารถนำข้อมูลนั้นกลับมาใช้ใหม่โดยไม่ต้องเขียนบ่อย ๆ การประกาศเอ็นทิตีมีรูปแบบดังนี้

```
<!ENTITY ชื่อเอ็นทิตี ทรัพยากร >
```

ส่วนใหญ่แล้วทรัพยากรหรือข้อมูล ก็คือ ข้อความ (สตริง) หรือไฟล์

เอ็นทิตีในภาษา XML มี 2 ประเภทใหญ่ ๆ คือ

- General Entity
- Parameter Entity เป็น เอ็นทิตีประเภทที่ค่อนข้างซับซ้อนพอสมควรและแทบจะไม่ได้ใช้ จึงไม่ขอกล่าวในที่นี้

General Entity แบ่งออกได้เป็น 2 ประเภทย่อย คือ Parsed Entity และ Unparsed Entity

โดย Parsed Entity เป็นเอ็นทิตีที่ XML Parser อ่านหรือแปล ได้ ซึ่งก็คือข้อความประเภทตัวอักษร

Unparsed Entity คือข้อมูลที่ XML Parser อ่านไม่ได้หรือแปลไม่ได้ เช่น ข้อมูลประเภทรูปภาพ ซึ่งเป็นไฟล์จำพวกไบนารี

Parsed Entity ใช้กับข้อมูลประเภทข้อความ

ข้อมูลที่ใช้อ้างอิงใน Parsed Entity จำแนกตามที่อยู่ของข้อมูลได้ 2 แบบ ได้แก่

แบบ Internal และแบบ External แบบ internal คือประกาศเอ็นทิตีนั้นไว้ใน DTD เลย

ส่วนแบบ external คือข้อมูลแยกไปอยู่อีกไฟล์หนึ่งต่างหาก ซึ่งมีรูปแบบการประกาศดังนี้

```
<!ENTITY ชื่อเอ็นทิตี SYSTEM “ชื่อไฟล์และพาร” >
```

Unparsed Entity ใช้กับข้อมูลที่ไม่ใช่ข้อความ

ข้อมูลที่เป็น Parsed Entity มี 2 แบบ คือ internal และ external แต่สำหรับ Unparsed Entity มีเฉพาะ external อย่างเดียว เพราะข้อมูลที่อ้างถึงนี้อยู่คนละไฟล์กับเอกสาร XML

โดยการประกาศใน DTD จะต้องมี 2 ขั้นตอนหลัก ๆ คือ

1. ประกาศโนเทชัน (NOTATION) เพื่อบอกชนิดข้อมูลก่อน
2. ประกาศเอ็นทิตี

การประกาศโนเทชัน มีรูปแบบดังนี้

```
<!NOTATION ชื่อโนเทชัน SYSTEM "External_ID" >
```

External_ID คือ โปรแกรมภายนอกที่จะมาจัดการกับข้อมูลนั้น ซึ่งส่วนใหญ่จะระบุเป็น MIME type แทน เช่น ไฟล์รูปภาพ ก็ระบุเป็น image/jpg หรือ image/gif ขึ้นอยู่กับว่ารูปภาพนั้นเป็นไฟล์ประเภทไหน ตัวอย่างเช่น ประกาศว่า <!NOTATION jpg SYSTEM "image/jpg" >

ต่อไปเป็นการประกาศเอ็นทิตี ซึ่งมีรูปแบบพิเศษออกไป ซึ่งประกาศดังนี้

```
<!ENTITY ชื่อเอ็นทิตี SYSTEM "ชื่อไฟล์และพาท" NDATA ชื่อโนเทชัน >
```

ส่วนที่เพิ่มขึ้นมาคือคีย์เวิร์ด NDATA และชื่อ โนเทชัน เพื่อเชื่อมโยงไปยังโนเทชันที่ประกาศไว้ก่อนแล้ว

2.1.2 XML Schema

XML Schema เป็นวิธีที่ใช้ในการกำหนดโครงสร้างของเอกสาร XML และเป็นวิธีที่ใหม่กว่า DTD มาก เพราะ DTD มีใช้มาตั้งแต่สมัยภาษา SGML และ XML Schema ได้กลายมาเป็นมาตรฐานของ W3C เอกสาร XML จะถูกตรวจดูความถูกต้องโดย XML parser โดยใช้ XML Schema ของมัน ตัวมันทำงานเหมือนไวยากรณ์ของเอกสารที่ใช้เพื่อให้ความหมายกับเอกสารที่มีโครงสร้าง

ข้อดีของ XML Schema ที่ดีกว่า DTD

- XML Schema มีไวยากรณ์ที่เหมือนกับ XML
- XML Schema รู้จักชนิดข้อมูลมากกว่า DTD และสามารถระบุเจาะจงชนิดข้อมูลได้
- ชนิดข้อมูลที่ XML Schema รองรับเป็นชนิดข้อมูลที่รองรับในโปรแกรมฐานข้อมูลที่มีอยู่ในปัจจุบัน ทำให้สามารถนำข้อมูลที่อยู่ในรูปเอกสาร XML ไปเก็บในฐานข้อมูลได้ง่ายโดยพิจารณาจากชนิดข้อมูลใน XML Schema
- สามารถกำหนดรูปแบบข้อมูลเองได้ เช่น กำหนดว่าข้อมูลหมายเลขโทรศัพท์ต้องมีรูปแบบเป็น "d-dddd-dddd" คือ ตัวเลข 1 ตัวตามด้วยขีด (-) และตัวเลข 4 ตัว จากนั้นจึงตามด้วยขีด (-) อีกครั้งและตัวเลขอีก 4 ตัว เป็นต้น
- สามารถสร้างชนิดข้อมูลใหม่โดยอิงจากชนิดของข้อมูลที่มีอยู่เดิม
- สามารถกำหนดจำนวนและลำดับของอิลิเมนต์ลูกได้ (อิลิเมนต์ลูกคืออิลิเมนต์ที่อยู่ในอิลิเมนต์ปัจจุบัน)
- XML Schema รองรับการใช้เนมสเปซ (namespace)

2.1.2.1 ลักษณะของ XML Schema

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2.2 Simple Type และ Complex Type

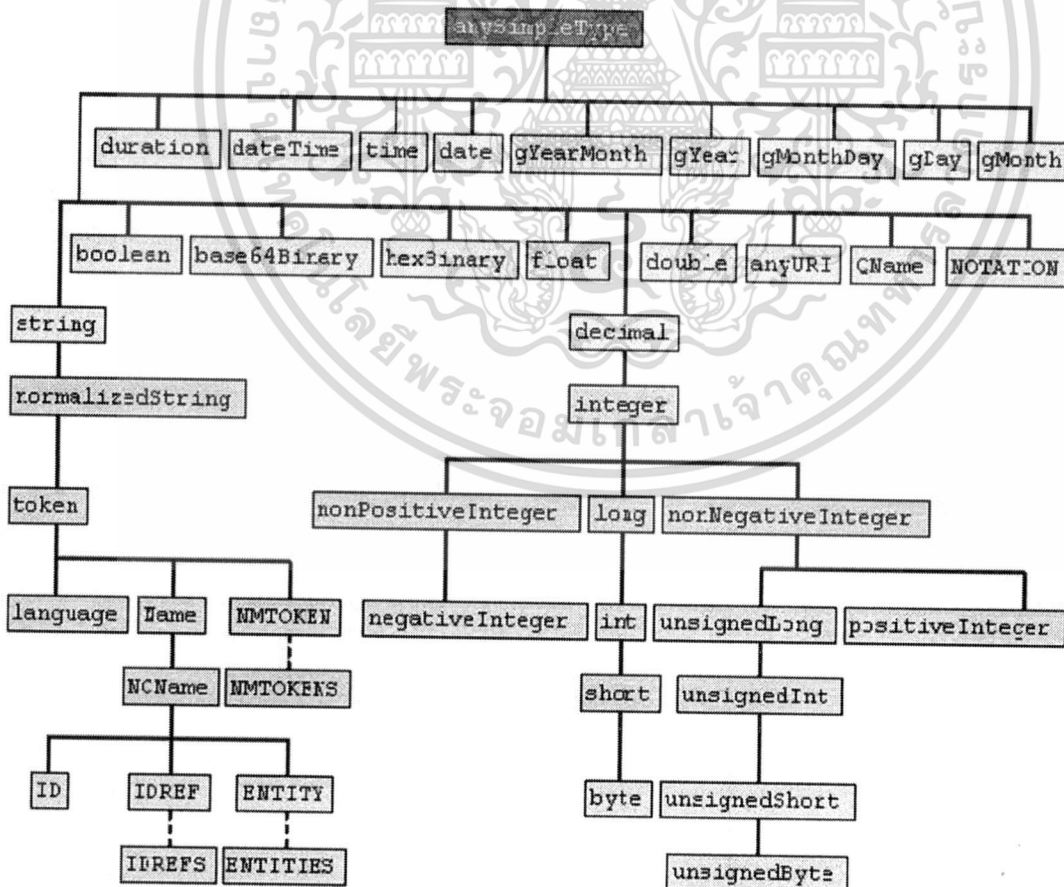
ใน XML Schema แบ่งอิลิเมนต์ออกเป็น 2 แบบคือ แบบ Simple Type และแบบ Complex Type โดย Simple Type คือ อิลิเมนต์ที่มีข้อมูลภายในเป็นข้อมูลพื้นฐาน อย่างเช่น สตริง, ตัวเลข เป็นต้น ส่วน Complex Type คือ อิลิเมนต์ที่มีแอตทริบิวต์หรือมีข้อมูลภายในเป็นอิลิเมนต์อื่น ๆ

❖ ชนิดข้อมูลแบบ Simple Type

Simple Type คือ อิลิเมนต์ที่มีข้อมูลภายในเป็นสตริง หรือตัวเลข หรือข้อมูลพื้นฐานอื่น ๆ นอกจากนี้ยังประกอบด้วยข้อมูลพื้นฐานอื่น ๆ อีก

ชนิดข้อมูลของ Simple Type แบ่งออกเป็น 2 กลุ่มคือ

1. Primitive datatype คือ ชนิดข้อมูลพื้นฐาน เช่น string, boolean, float, double เป็นต้น
2. Derived datatype คือ ชนิดข้อมูลที่ขยายมาจาก primitive datatype อีกที่ดังแผนภาพที่ 2_1



รูปที่ 2_1 แสดงการแจกแจงชนิดข้อมูลแบบ Simple Type ของ XML Schema

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การประกาศ XML Schema สำหรับ อีลิเมนต์ที่เป็น Simple Type มีรูปแบบพื้นฐานดังนี้

```
<element name="ชื่ออีลิเมนต์" type="ชนิดของข้อมูล">
```

❖ ชนิดข้อมูลแบบ Complex Type

Complex Type เป็นอีลิเมนต์ที่บรรจุอีลิเมนต์อื่น ๆ ไว้ภายใน หรือเป็นอีลิเมนต์ที่มีแอตทริบิวต์ด้วย

ต่อไปเป็นตัวอย่างการประกาศอีลิเมนต์แบบ Complex Type ใน XML Schema

```
1: <xsd:complexType name="USAddress"
2: <xsd:sequence>
3: <xsd:element name="name" type="xsd:string"/>
4: <xsd:element name="street" type="xsd:string"/>
5: <xsd:element name="city" type="xsd:string"/>
6: <xsd:element name="state" type="xsd:string"/>
7: <xsd:element name="zip" type="xsd:decimal"/>
8: </xsd:sequence>
9: <xsd:attribute name="country" type="xsd:NMTOKEN"
   fixed="US"/>
10: </xsd:complexType>
```

XML Schema ในตัวอย่างข้างต้นกำหนดอีลิเมนต์ประเภท USAddress ขึ้นมาเป็นแบบ Complex Type (บรรทัดที่ 1) ซึ่งประกอบไปด้วย อีลิเมนต์ name, street, city, state, zip (บรรทัดที่ 3 ถึงบรรทัดที่ 7) และมีแอตทริบิวต์ country ด้วย (บรรทัดที่ 9) นอกจากนี้ยังระบุชนิดข้อมูลของแต่ละอีลิเมนต์ ซึ่งส่วนใหญ่เป็นข้อมูลชนิดสตริง ยกเว้น zip ที่เป็นตัวเลขทศนิยม (decimal) และ country เป็นชนิด NMTOKEN ซึ่งก็คือสตริงประเภทที่ไม่มีช่องว่าง

2.1.2.3 เนมสเปซ

ในภาษา XML จะอาศัยข้อกำหนดที่สำคัญอย่างหนึ่ง ซึ่งเรียกว่า เนมสเปซ (namespace) เพื่อใช้สำหรับป้องกันความสับสนในการระบุชื่อองค์ประกอบใด ๆ ภายในเอกสาร XML คนละไฟล์ แต่ใช้ชื่อเดียวกัน เนื่องจากอาจเป็นไปได้ว่าเอกสาร XML ที่ใช้ในงานที่ต่างกัน มีชื่ออีลิเมนต์เหมือนกัน ชื่อชนิดข้อมูลเหมือนกัน แล้วอาจจะถูกเรียกใช้ภายในเอกสาร XML ไฟล์เดียวกัน

การประกาศ prefix และ เนมสเปซ จะอาศัยแอตทริบิวต์ xmlns (ย่อมาจาก XML namespace) ตามรูปแบบดังนี้

```
<xmlns:prefix="ชื่อเนมสเปซ">
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3 เมตาดาต้า และ โครงสร้างเมตาดาต้าที่ใช้แสดงวิทยานิพนธ์

เมตาดาต้า คือ ข้อมูลเกี่ยวกับข้อมูลอื่น ๆ โดยทั่ว ๆ ครอบคลุมเป็นเมตาดาต้าที่ใช้บรรยายเช่น ข้อมูลบรรณานุกรม เมตาดาต้าโครงสร้างเป็นข้อมูลที่บอกรูปแบบและโครงสร้าง และ เมตาดาต้าที่ใช้ในการจัดการ ซึ่งใช้ในการจัดการกับข้อมูล

3.1 Descriptive Metadata

มีวิธีการมากมายในการค้นหาข้อมูลโดยไม่ต้องเข้าไปหาในเนื้อข้อมูลนั้น โดยตรง ซึ่งหนึ่งในนั้นก็คือการค้นหาจาก การบรรยายด้วยเมตาดาต้า(Descriptive Metadata) เมตาดาต้าของอ็อบเจกต์หนึ่ง ๆ อาจจะเป็น ข้อมูลของ แคตตาล็อก, ข้อมูลอินเด็กซ์ หรือบทคัดย่อ ซึ่งโดยปกติแล้ว การบรรยายด้วยเมตาดาต้าจะเก็บแยกจากอ็อบเจกต์ที่มันกล่าวถึง แต่มีบางครั้งที่มันถูกเก็บรวมกันไว้เลย

การบรรยายด้วยเมตาดาต้าโดยปกติแล้วจะแสดงในรูปของตัวอักษร ซึ่งนอกจากจะใช้กับอ็อบเจกต์ที่เป็นข้อความแล้วยังนำไปใช้กับ อ็อบเจกต์ที่เป็นรูปภาพ, เสียง, แผนที่ หรือโปรแกรมคอมพิวเตอร์ได้ โดยในการค้นหาผู้ใช้สามารถค้นหาข้อมูลเหล่านี้ได้โดยการค้นหาผ่านข้อความที่บรรยายอ็อบเจกต์นั้น

โดยปกติแล้วการบรรยายด้วยเมตาดาต้าจะทำโดยผู้เชี่ยวชาญในสาขานั้น ๆ ในการทำแคตตาล็อกหนังสือและดัชนีทางวิทยาศาสตร์เป็นงานที่ต้องใช้ทักษะของคนและเวลา ดังนั้นจึงจำเป็นที่จะต้องหาหนทางที่ถูกและรวดเร็วในการสร้างเมตาดาต้าโดยใช้คอมพิวเตอร์เพื่อขยายความสามารถของคน

3.2 โครงสร้างเมตาดาต้าที่ใช้แสดงวิทยานิพนธ์

ในการเก็บข้อมูลวิทยานิพนธ์จะทำการเก็บข้อมูลทั้งหมดของวิทยานิพนธ์ โดยเพิ่มแท็กที่แสดงถึงความหมายในแต่ละส่วนของวิทยานิพนธ์เพื่อทำให้ข้อมูลที่จัดเก็บอยู่มีความหมายง่ายต่อการค้นหาและการทำดัชนีเอกสาร ต่อ ไปนี้แสดงตัวอย่างของวิทยานิพนธ์ที่จัดเก็บ

หัวข้อวิทยานิพนธ์ การบีบอัดข้อมูลภาพ โดยการเข้ารหัสแบบบล็อกหลายขนาด
 นักศึกษา นายอำนาจ ขาวเน
 รหัสประจำตัว 42061037
 ปริญญา วิศวกรรมศาสตรมหาบัณฑิต
 สาขาวิชา วิศวกรรมไฟฟ้า
 พ.ศ. 2544
 อาจารย์ผู้ควบคุมวิทยานิพนธ์ ผศ.ดร.สุทธิชัย นพนาสีพงษ์

บทคัดย่อ

วิทยานิพนธ์เล่มนี้ได้นำเสนอวิธีการบีบอัดข้อมูลภาพ โดยการเข้ารหัสแบบบล็อกหลายขนาด ซึ่งใช้วิธีการแยกองค์ประกอบควอดทรีแบบใหม่ โดยอาศัยหลักการของลากรากิเยนมัดดิฟลายเออร์ การแยกองค์ประกอบควอดทรีเป็นวิธีการที่ใช้แสดงภาพในลักษณะที่เป็นลำดับชั้น แต่อย่างไรก็ตามการที่จะแบ่งระดับชั้นของภาพได้นั้นจำเป็นต้องกำหนดค่าเทรชโฮล ซึ่งเป็นเรื่องยากที่จะหาค่าเทรชโฮลที่เหมาะสมในการแบ่งภาพทั้งภาพ ในที่นี้ได้นำเสนอวิธีการตัดสินใจในการแบ่งบล็อกโดยใช้ค่าของลา กรากิเยนมาเป็นตัวกำหนดการเติบโตของโครงสร้างต้นไม้ ซึ่งในวิธีการที่ได้นำเสนอนั้นจะให้โครงสร้างต้นไม้เติบโตที่ละโหนดแตกต่างกับการเติบโตของโครงสร้างต้นไม้แบบเดิม ซึ่งเติบโตทีละชั้น และผลที่ได้ของวิธีการที่ได้นำเสนอมีข้อดีกว่าวิธีการแบบเดิม

Thesis Title	Image Compression Based on Variable Block Truncation Code
Student	Mr.Amnach Khawne
Student ID.	42061037
Degree	Master of Engineering
Programme	Electrical Engineering
Year	2001
Thesis Advisor	Asst. Dr. Suthichai Noppanakeepong

ABSTRACT

In this thesis, we proposes image coding by using variable block truncation code, which uses a new quadtree growing decided by the Lagrangian function. Quadtree decomposition (QT) is a technique for representing an image data as hierarchical data structure. However, QT technique requires the threshold for the test-block criterion, It is difficult to define the optimum threshold value which is covered the whole image data. In our algorithm, we use Lagrangian cost instead of the threshold for decision the tree grows. In our algorithm, the tree grows one node at a time that differ from the method of the quadtree growing which grows tree one layer at the time, and our results shown that the algorithm obtained a good quality of image reconstruction.

ส่วนด้านล่างนี้เป็นตัวอย่างของวิทยานิพนธ์ที่จัดเก็บไว้ในรูปแบบของไฟล์ XML ที่กำหนดแท็กที่แสดงถึงความหมายในแต่ละส่วนของวิทยานิพนธ์ ในส่วนของความหมายของแต่ละแท็กแสดงไว้ในภาคผนวก ก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	<pre><?xml version="1.0" encoding="UTF-8"?> <!--Sample XML file generated by XML Spy v4.3 U (http://www.xmlspy.com)--> <Thesis xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="D:\Study\MasterProjectXMLThesis\ThesisDB.xsd"> <Abstract> <Title leng="th">การบีบอัดข้อมูลภาพโดยการเข้ารหัสแบบบล็อกหลายขนาด</Title> <Title leng="en">Image Compression Based on Variable Block Truncation Code </Title> <Creator leng="th"> <Initial>นาย</Initial> <Firstname>อำนาจ</Firstname> <Lastname>ชวาน</Lastname> </Creator> <Creator leng="en"> <Initial>Mr.</Initial> <Firstname>Amnach</Firstname> <Lastname>Khawne</Lastname> </Creator> <StudentID>42061037</StudentID> <Degree leng="th">วิศวกรรมศาสตรมหาบัณฑิต</Degree> <Degree leng="en">Master of Engineering</Degree> <Programme leng="th">วิศวกรรมไฟฟ้า</Programme> <Programme leng="en">Electrical Engineering</Programme> <Faculty leng="th">คณะวิศวกรรมศาสตร์</Faculty> <Faculty leng="en">Faculty of Engineering</Faculty> <Year type="AD">2544</Year> <Year type="BC">2001</Year> <Advisor leng="th"> <Initial>ผศ.ดร.</Initial> <Firstname>สุทธิชัย</Firstname> <Middlename/> <Lastname>นพนาศิพงษ์</Lastname> </Advisor> <Advisor leng="en"> <Initial>Asst. Dr.</Initial> <Firstname>Suthichai</Firstname> <Middlename/> <Lastname>Noppanakepong</Lastname> </Advisor> <Abstract.content leng="th">วิทยานิพนธ์เล่มนี้ได้นำเสนอวิธีการบีบอัดข้อมูลภาพ โดยการเข้ารหัสแบบบล็อกหลายขนาด ซึ่งใช้วิธีการแยกองค์ประกอบควอดทรีแบบใหม่ โดยอาศัยหลักการของลากรางจ์เกี่ยวกับผลดีฟายเออร์ การแยกองค์ประกอบควอดทรีเป็นวิธีการที่ใช้แสดงภาพในลักษณะที่เป็นลำดับขั้น แต่อย่างไรก็ตามการที่จะแบ่งระดับขั้นของภาพได้นั้นจำเป็นต้องกำหนดค่าเทรชโวล ซึ่งเป็นเรื่องยากที่จะหาค่าเทรชโวลที่เหมาะสมในการแบ่งภาพทั้งภาพ ในที่นี้ได้นำเสนอวิธีการตัดสินใจในการแบ่งบล็อกโดยใช้ค่าของลากรางจ์มาเป็นตัวกำหนดการเติบโตของโครงสร้างต้นไม้ ซึ่งในวิธีการที่ได้นำเสนอนั้นจะให้โครงสร้างต้นไม้เติบโตที่ละเอียดแตกต่างกับการเติบโตของโครงสร้างต้นไม้แบบเดิม ซึ่งเติบโตที่ละเอียดและผลที่ได้ของวิธีการที่ได้นำเสนอมีข้อดีกว่าวิธีการแบบเดิม</Abstract.content> <Abstract.content leng="th">In this thesis, we proposes image coding by using variable block truncation code, which uses a new quadtree growing decided by the Lagrangian function. Quadtree decomposition (QT) is a technique for representing an image data as hierarchical data structure. However, QT technique requires the threshold for the test-block criterion, It is difficult to define the optimum threshold value which is covered the whole image data. In our algorithm, we use Lagrangian cost instead of the threshold for decision the tree grows. In our algorithm, the tree grows one node at a time that differ from the method of the quadtree growing which grows tree one layer at the time, and our results shown that the algorithm obtained a good quality of image reconstruction. </Abstract.content> <Keywords> <Keyword>ควอดทรี</Keyword> <Keyword>Variable Block Truncation Code</Keyword> </Keywords> </Abstract></pre>
ส่วนบทคัดย่อ	
หมายเลข ISBN	<ISBN>974-648-152-8</ISBN>
ส่วนกิตติกรรมประกาศ	<p><Acknowledgements>การจัดทำวิทยานิพนธ์ฉบับนี้ได้รับความสำเร็จลุล่วงไปด้วยดีก็เพราะการช่วยเหลือและให้ความอนุเคราะห์จากบุคคลหลายท่าน ขอขอบพระคุณแม่ ขอขอบคุณพี่ๆและน้องๆห้องวิจัย การสื่อสารดาวเทียมที่ให้ความช่วยเหลือและให้กำลังใจเป็นอย่างมาก ขอขอบพระคุณ ผศ.ดร.สุทธิชัย นพนาศิพงษ์ ในฐานะที่เป็นอาจารย์ที่ปรึกษาที่ให้ความดูแลและเอาใจใส่อย่างสม่ำเสมอ ขอขอบพระคุณ ผศ.กฤษณ์ วงศ์จรูญและผศ.ดร. ปิฎญา ฐิติมีดมขีมา ที่ให้คำปรึกษา ขอขอบพระคุณ รศ.ดร. ชูโตมุ กาวาบะตะ (Assoc. Prof. Tsutomu KAWABATA) ที่ให้คำแนะนำและให้คำปรึกษา พร้อมทั้งเอาใจใส่ดูแลเสมอมาในการทำวิจัย ขอขอบพระคุณห้องวิจัยของ ศาสตราจารย์ ทาคะเดะ มิทซึโอะ (Prof. Takeda MITSUO) ที่ได้เชื้อเพลิงสถานที่ในการทำวิจัย</Acknowledgements></p>

เอกสารนี้เป็นเอกสาร

เผยแพร่เป็นการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนสารบัญ	<pre> <TableofContent> <RowofTable> <Content>บทคัดย่อภาษาไทย</Content> <PageNumber>I</PageNumber> </RowofTable> <RowofTable> <Content>บทคัดย่อภาษาอังกฤษ</Content> <PageNumber>II</PageNumber> </RowofTable> <RowofTable> <Content>กิตติกรรมประกาศ</Content> <PageNumber>III</PageNumber> </RowofTable> <RowofTable> <Content>สารบัญ</Content> <PageNumber>IV</PageNumber> </RowofTable> <Content>บรรณานุกรม</Content> <PageNumber>81</PageNumber> </RowofTable> <RowofTable> <Content>ภาคผนวก </Content> <PageNumber>84</PageNumber> </RowofTable> <RowofTable> <Row.Number>ภาคผนวก ก</Row.Number> <Content>วิธีการของตัวคูณลากรางจ์</Content> <PageNumber>85</PageNumber> </RowofTable> <RowofTable> <Row.Number>ภาคผนวก ข</Row.Number> <Content>โปรแกรม MATLAB ที่ใช้ในการทดลอง</Content> <PageNumber>89</PageNumber> </RowofTable> <RowofTable> <Row.Number>ภาคผนวก ค</Row.Number> <Content>ผลงานที่ได้รับการตีพิมพ์</Content> <PageNumber>108</PageNumber> </RowofTable> <RowofTable> <Content>ประวัติผู้เขียน</Content> <PageNumber>109</PageNumber> </RowofTable> </TableofContent> </pre>
ส่วนสารบัญ ตาราง	<pre> <ListofTable> <RowofTableT> <Table.Number>2.1</Table.Number> <Content>ความน่าจะเป็นของสัญลักษณ์ในการเข้ารหัสด้วยวิธีซีฟแมน</Content> <PageNumber>30</PageNumber> </RowofTableT> <RowofTableT> <Table.Number>2.2</Table.Number> <Content>แสดงการแบ่งขอบเขตตามความน่าจะเป็นสะสมในการเข้ารหัสเลขคณิต </Content> <PageNumber>32</PageNumber> </RowofTableT> <RowofTableT> <Table.Number>2.3</Table.Number> <Content>ปทานุกรมสำหรับวิธีการของเลมเปิล-ชิป</Content> <PageNumber>35</PageNumber> </RowofTableT> </ListofTable> </pre>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<p>ส่วนเนื้อหา วิทยานิพนธ์</p>	<pre> <Page Page.number="5"> <Paragraph order="1" position="left"> <paragraph.text> ถูกเรียกว่าการแซมปลิง (sampling) ภาพ และถ้าเป็น การทำดิจิทัลทางขนาดจะเรียกว่าการควอนไทซ์ในระดับเทา (gray-level quantization) [20] โดย ส่วนมากภาพสองมิติที่มีขนาด จะแสดงดังสมการที่ 2.1 </paragraph.text> </Paragraph> <Paragraph order="2" position="center"> <math>สมการ 2.1</math> </Paragraph> <Paragraph order="3" position="left"> <paragraph.text> ในทางด้านขวามือของสมการ 2.1 ปกติจะเรียกว่าภาพ ดิจิตอล (digital image) ความต้องการเบื้องต้นสำหรับการประมวลผลภาพด้วยคอมพิวเตอร์แสดงดังรูปที่ 2.1 นั่นคือภาพควรจะถูกจัดอยู่ในรูปแบบดิจิตอล การประมวลผลภาพเริ่มจากภาพถูกแซมเปิลหรือการสุ่มสัญญาณให้อยู่ ในรูปแบบจุดภาพในแต่ละแซมเปิลหรือจุดภาพจะถูกควอนไทซ์โดยใช้จำนวนบิตที่จำกัด ภาพที่ถูกควอนไทซ์ แล้วสามารถประมวลผลได้ด้วยคอมพิวเตอร์ เพื่อที่จะแสดงผลภาพในรูปแบบที่เป็นดิจิตอลนั้น อันดับแรกภาพ จะถูกแปลงไปเป็นสัญญาณอนาลอก ซึ่งภาพที่แสดงบนจอภาพนั้น ได้จากการสแกนของหลอดภาพ </paragraph.text> </Paragraph> <Paragraph order="3" position="center"> <image gotofig="2.1"/> </Paragraph> <Paragraph order="4" position="center"> <paragraph.head> <paragraph.headcontent>รูปที่ 2.1 การประมวลผลภาพด้วย คอมพิวเตอร์</paragraph.headcontent> </paragraph.head> </Paragraph> <Paragraph order="4" position="left"> <paragraph.text>การควอนไทซ์มีอยู่สองแบบด้วยกันคือ การควอนไทซ์ แบบสเกลาร์ ซึ่งทำการควอนไทซ์ข้อมูลทำการสุ่มโดยการทดลองตัวอย่าง ส่วนการควอนไทซ์แบบเวกเตอร์ จะทำการควอนไทซ์บล็อกของข้อมูลโดยใช้บล็อก (ที่มีความยาวบล็อกเท่าๆกัน) การควอนไทซ์แบบสเกลาร์มี ข้อดีตรงที่มีความง่าย แต่ความผิดเพี้ยนมากกว่าการควอนไทซ์แบบเวกเตอร์ที่มีความยุ่งยากมากกว่า </paragraph.text> </Paragraph> <Paragraph order="5" position="left"> <paragraph.head> <paragraph.headcontent>2.2 การควอนไทซ์แบบสเกลาร์ที่ใช้ระดับ คงที่และไม่คงที่ (Uniform and Non-Uniform Scalar Quantization)</paragraph.headcontent> </paragraph.head> </Paragraph> <Paragraph order="6" position="left"> <paragraph.text>เวกเตอร์ข้อมูลที่จะถูกลดคือ สมมติว่าข้อมูลตัวอย่าง เป็น ค่าจริงซึ่งในทางปฏิบัติมีความเป็นไปได้หลายค่าสำหรับเวกเตอร์ข้อมูล วัตถุประสงค์ของการควอนไทซ์คือ </paragraph.text> </Paragraph> </Page> <Page Page.number="2"> <Paragraph order="1" position="center"> <math>สมการ 2.70.1</math> </Paragraph> <Paragraph order="2" position="center"> <math>สมการ 2.70.2</math> </Paragraph> <Paragraph order="3" position="left"> <paragraph.head> <paragraph.headcontent>ตัวอย่างการเข้ารหัสแบบฮัฟแมน </paragraph.headcontent> </paragraph.head> </Paragraph> <Paragraph order="4" position="left"> <paragraph.text>สมมติให้องค์ประกอบข่าวสารมี 7 องค์ประกอบคือ โดยที่ มีความน่าจะเป็นของการเกิดเป็น 0.35, 0.30, 0.20, 0.10, 0.04, 0.005, 0.005 ตามลำดับ เมื่อเข้ารหัส ด้วยวิธีการเข้ารหัสฮัฟแมนจะได้ดังนี้</paragraph.text> </Paragraph> <Paragraph order="5" position="center"> <image gotofig="2.11"/> </Paragraph> <Paragraph order="6" position="center"> <paragraph.head> <paragraph.headcontent>รูปที่ 2.11 การสร้างรหัสฐานสองด้วยวิธีฮัฟ แมน</paragraph.headcontent> </paragraph.head> </Paragraph> </pre>
------------------------------------	--

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการ $$ เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<p>ส่วนเนื้อหา วิทยานิพนธ์</p>	<pre> </Paragraph> <Paragraph order="7" position="left"> <paragraph.head> <paragraph.headcontent>ตารางที่ 2.1 ความน่าจะเป็นของสัญลักษณ์ ในการเข้ารหัสตัวบิตวีซีเอฟแมน</paragraph.headcontent> </paragraph.head> </Paragraph> <Paragraph order="8" position="center"> <table> <tr> <td>สัญลักษณ์</td> <td>ความน่าจะเป็น</td> <td>ค่ารหัส</td> <td>ความยาวรหัส</td> </tr> <tr> <td>x1</td> <td>0.35</td> <td>00</td> <td>2</td> </tr> </table> </Paragraph> </Page> </Pages> </Chapter> </Body> </pre>
<p>ส่วนบรรณานุกรม</p>	<pre> <Bibliography> <Citation> <Number.Citation>1</Number.Citation> <Name.Citation>Three levels block truncation coding</Name.Citation> <Name.Creator.FirstName>Alsaka</Name.Creator.FirstName> <Name.Creator.LastName>Y. A.</Name.Creator.LastName> <Name.Creator.Other>Lee D. A.</Name.Creator.Other> <Publish.Name>Proceeding IEEE</Publish.Name> <Publish.Year>1990</Publish.Year> </Citation> <Citation> <Number.Citation>2</Number.Citation> <Name.Citation>Image Compression Using Block Truncation Coding</Name.Citation> <Name.Creator.FirstName>Delp</Name.Creator.FirstName> <Name.Creator.LastName>E. J.</Name.Creator.LastName> <Name.Creator.Other>Mitchell O. R.</Name.Creator.Other> <Publish.Name>IEEE Transactions on Communications</Publish.Name> <Publish.Place>String</Publish.Place> <Publish.Year>1979</Publish.Year> <Publish.Volume>27</Publish.Volume> <Publish.Date>Sep.</Publish.Date> <Citation.page>1335-1342</Citation.page> </Citation> </Bibliography> </pre>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<p>ส่วนภาคผนวก</p>	<pre> <Appendix> <Chapter> <Chapter.number>๗</Chapter.number> <Chapter.name>โปรแกรม MATLAB ที่ใช้ในการทดลอง</Chapter.name> <Pages> <Page Page.number="89"> <Paragraph order="1" position="center"> <paragraph.head> <paragraph.headcontent>ภาคผนวก ๗ </paragraph.headcontent> </paragraph.headcontent> </paragraph.head> </Paragraph> <Paragraph order="2" position="center"> <paragraph.head> <paragraph.headcontent>โปรแกรม MATLAB ที่ใช้ในการทดลอง </paragraph.headcontent> </paragraph.headcontent> </paragraph.head> <Paragraph order="3" position="left"> <paragraph.text>ฟังก์ชันหาค่าลากรางจ์ของบล็อกภาพ และการเข้ารหัส แบบการเข้ารหัสโดยเก็บรักษาค่าสีมบรณโมเมนต์ clc; clear; I = imread('image file'); [I1 I2 I3] = size(I); if I3 == 1 II = I; else II = rgb2gray(I); end </paragraph.text> </Paragraph> </Pages> </Chapter> </Appendix> </pre>
<p>ส่วนประวัติผู้เขียน</p>	<pre> <Biography> <Creator leng="th"> <Initial>นาย</Initial> <Firstname>อำนาจ</Firstname> <Lastname>ชานน</Lastname> </Creator> <Creator leng="en"> <Initial>Mr.</Initial> <Firstname>Amnach</Firstname> <Lastname>Khawne</Lastname> </Creator> <Date.Birth> <Day>7</Day> <Month>สิงหาคม</Month> <Year type="AD">2517</Year> </Date.Birth> <Place.Birth>สุพรรณบุรี</Place.Birth> <Study_History>ปีการศึกษา 2538-2540 ระดับปริญญาตรีวิศวกรรมศาสตรบัณฑิต สาขาวิศวกรรมโทรคมนาคม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง </Study_History> </Biography> </pre>
	<pre> </Thesis> </pre>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในแต่ละส่วนจะเก็บข้อมูลทั้งหมดของวิทยานิพนธ์ไว้ในไฟล์เอกสาร XML หลังจากนั้นจะนำมาจัดเก็บและแสดงผลต่อไป



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4 การทดลอง

4.1 อุปกรณ์

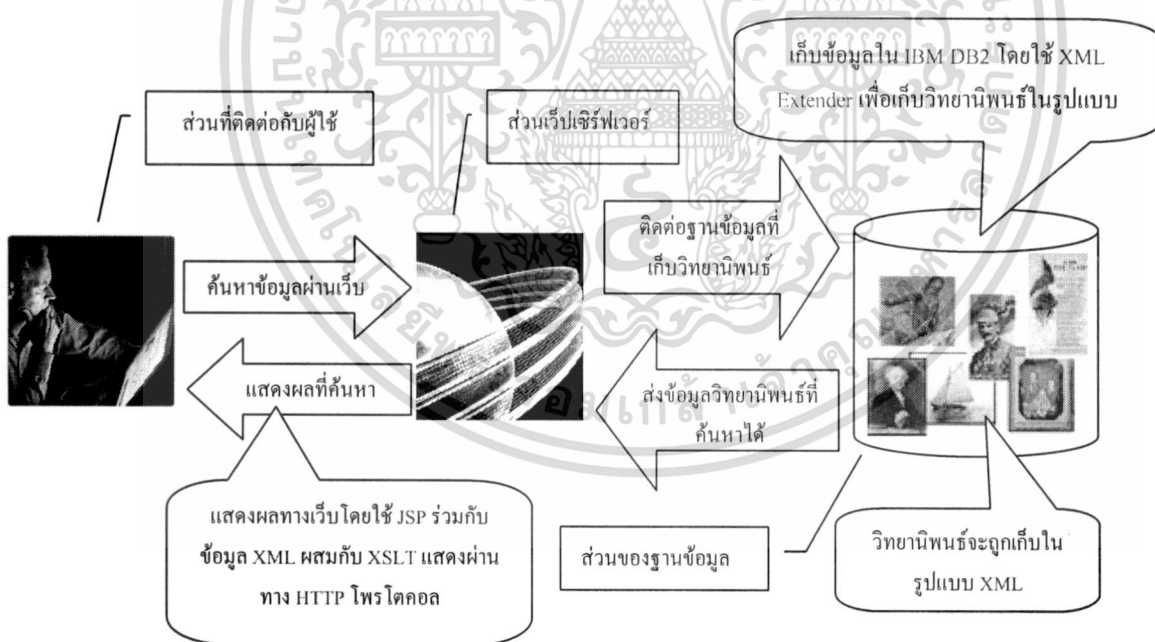
1. เครื่องคอมพิวเตอร์ 1 เครื่อง เพนเทียมทรี ความเร็วสัญญาณนาฬิกา 1 GH, หน่วยความจำ 256 MB เนื้อที่ในการเก็บข้อมูลประมาณ 20 GB เพื่อใช้เป็นเซิร์ฟเวอร์
2. Java SDK version 1.4.1
3. Apache web server version 4.0.6
4. IBM DB2 UDB V 8.1
5. Internet Explorer

4.2 วิธีการ

ประกอบไปด้วยภาพรวมของระบบต้นแบบ, ซอฟต์แวร์เครื่องมือ และฐานข้อมูลที่เกี่ยวข้อง และกระบวนการ

4.2.1 ภาพรวมของระบบต้นแบบ

ระบบห้องสมุดวิทยานิพนธ์นี้ประกอบด้วย 3 ส่วนคือ ส่วนที่ติดต่อกับผู้ใช้ ส่วนเว็บเซิร์ฟเวอร์ และส่วนฐานข้อมูล ดังแสดงดังรูปที่ 4_1



รูปที่ 4_1 แสดงโครงสร้างของระบบห้องสมุดวิทยานิพนธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.1 ส่วนติดต่อผู้ใช้

ส่วนการติดต่อผู้ใช้ ทำการติดต่อผ่านทาง WWW และแสดงผลของวิทยานิพนธ์ที่สืบค้นผ่านทางเว็บไซต์

1.2 ส่วนเว็บเซิร์ฟเวอร์

เป็นส่วนที่ใช้ในการติดต่อระหว่างผู้ใช้กับฐานข้อมูล เป็นตัวจัดการการทำงานของเว็บกับผู้ใช้แต่ละคน

1.3 ส่วนฐานข้อมูล

ในการเก็บข้อมูล XML นี้จะเก็บในฐานข้อมูลเพื่อให้สามารถสืบค้นได้อย่างรวดเร็ว และทำให้การจัดการกับข้อมูลเป็นไปอย่างเป็นระบบ ในส่วนของฐานข้อมูลใช้ DB2 UDB เป็นฐานข้อมูล

4.2.2 การเตรียมความพร้อมของข้อมูลและซอฟต์แวร์เครื่องมือ

4.2.2.1 เอกสาร XML

เมื่อได้วิทยานิพนธ์ของนักศึกษาที่สำเร็จการศึกษามาแล้วจะทำการเปลี่ยนข้อมูลเหล่านั้นให้อยู่ในรูปแบบของเอกสาร XML ที่ประกอบด้วยแท็กที่ได้กำหนดไว้ โดยในส่วนแรกนี้จะทำการเก็บข้อมูลเฉพาะส่วนบทคัดย่อเท่านั้นเนื่องจากข้อมูลที่ทางบัณฑิตวิทยาลัยจัดเก็บไว้มีเฉพาะบทคัดย่อของนักศึกษาเท่านั้น

4.2.2.2 ซอฟต์แวร์เครื่องมือ

เนื่องจากการเก็บข้อมูล XML ลงในฐานข้อมูลเลยทั้งทำได้แต่ให้ผลไม่ดีเท่าที่ควร คือไม่มีการจัดการเอกสาร XML และไม่ช่วยให้การค้นหาทำได้อย่างรวดเร็วจึงมีซอฟต์แวร์เพื่อช่วยในการจัดการกับเอกสาร XML โดยใช้ DB2 XML Extender

2.2.1 DB2 XML Extender

DB2 XML Extender เป็นโปรแกรมที่ช่วยในการเก็บเอกสาร XML ประกอบเอกสาร XML จากตารางที่มีอยู่ หรือ แยกส่วนเอกสาร XML เข้าไปในตาราง ซึ่งซอฟต์แวร์นี้ได้จัดเตรียมชนิดข้อมูล ฟังก์ชัน และ Stored procedure ที่จะใช้จัดการกับเอกสาร XML ขึ้นมาใหม่ นอกจากนี้ยังสามารถเข้าถึงเอกสาร XML ได้โดยใช้ SQL ซึ่งทำให้่ง่ายในการใช้งาน

4.3 ขั้นตอนวิธี

4.3.1 การจัดเก็บเอกสาร

หลังจากได้เอกสาร XML ที่เป็นส่วนของบทคัดย่อวิทยานิพนธ์เรียบร้อยแล้วจะเตรียมนำมาจัดเก็บในฐานข้อมูล ซึ่งต้องมีการเตรียมการดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ทำการสร้างฐานข้อมูลที่จะเก็บข้อมูล และทำให้ฐานข้อมูลนั้นสามารถใช้กับเอกสาร XML ได้
2. สร้างคอลัมน์ที่จะเอาไว้เก็บเอกสาร XML
3. นำ DTD ที่เป็นโครงสร้างของเอกสารไปเก็บไว้ในฐานข้อมูล
4. จัดเตรียม DAD ไฟล์ ซึ่ง DAD ไฟล์นี้เป็นไฟล์ที่จะทำการสร้างตาราง โดยตารางนั้นเป็นตารางที่เก็บข้อมูลในเอกสาร XML เพื่อให้สามารถค้นหาข้อมูลในฐานข้อมูลได้เร็วขึ้น ต่อไปนี้เป็นตัวอย่าง DAD ไฟล์ที่ใช้สร้างตารางในฐานข้อมูล

```

<?xml version="1.0" encoding="UTF-8"?>
<DAD>
  <dtidid>D:\Study\MasterProject\XMLThesis\ThesisAbstract.dtd</dtidid>
  <validation>NO</validation>
  <Xcolumn>
    <table name="TitleSideTable">
      <column name="Title" type="varchar(900)" path="/Thesis/Abstract/Title"
multi_occurrence="YES"/>
    </table>
    <table name="StudentIDSideTable">
      <column name="Student" type="varchar(10)" path="/Thesis/Abstract/StudentID"
multi_occurrence="NO"/>
    </table>
    <table name="StudentFirstName">
      <column name="StudentFirstName" type="varchar(100)"
path="/Thesis/Abstract/Creator/FirstName" multi_occurrence="YES"/>
    </table>
    <table name="StudentLastName">
      <column name="StudentLastName" type="varchar(100)"
path="/Thesis/Abstract/Creator/LastName" multi_occurrence="YES"/>
    </table>
    <table name="AbstractContent">
      <column name="AbstractContent" type="clob(32K)"
path="/Thesis/Abstract/AbstractContent" multi_occurrence="YES"/>
    </table>
  </Xcolumn>
</DAD>

```

เอกสารนี้เป็นเอกสารที่สงวนเวลาสำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. หลังจากนั้นก็จัดเก็บเอกสาร XML เอกสารที่จัดเก็บแสดงได้ดังนี้

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE Thesis SYSTEM "D:\Study\MasterProject\XMLThesis\ThesisAbstract.dtd">
<Thesis>
  <Abstract>
    <Title lang="th">การศึกษาการเลือกสร้างฟิล์มเพชรเฉพาะพื้นที่ ด้วยวิธี CVD แบบความร้อน</Title>
    <Title lang="en">A Study on Selective-Area Deposition of Diamond Films by Thermal CVD</Title>
    <Creator lang="th">
      <Initial>นาย</Initial>
      <FirstName>กอบศักดิ์</FirstName>
      <LastName>ศรีประภา</LastName>
    </Creator>
    <Creator lang="en">
      <Initial>Mr.</Initial>
      <FirstName>Kobsak</FirstName>
      <LastName>Sriprapha</LastName>
    </Creator>
    <StudentID>0</StudentID>
    <Degree lang="th">วิศวกรรมศาสตรมหาบัณฑิต</Degree>
    <Degree lang="en">Master of Engineering</Degree>
    <Program lang="th">วิศวกรรมไฟฟ้า</Program>
    <Program lang="en">Electrical Engineering</Program>
    <Faculty lang="th">วิศวกรรมศาสตร์</Faculty>
    <Faculty lang="en">Engineering</Faculty>
    <Year format="BE">2541</Year>
    <Year format="AD">1998</Year>
    <Advisor lang="th">
      <Initial>รศ.ดร.</Initial>
      <FirstName>วิสุทธิ์</FirstName>
      <LastName>ฐิติรุ่งเรือง</LastName>
    </Advisor>
  </Abstract>

```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.2 การสืบค้น

ในการสืบค้นเอกสารจะทำการสืบค้นโดยใช้ Boolean Model ในการสืบค้นข้อมูลภายในฐานข้อมูล โดยทำการหาคำที่ตรงกับคำค้นในฐานข้อมูล แล้วนำข้อมูลที่พบบนนั้นมาแสดงผล โดยแปลงข้อมูล XML ให้สามารถแสดงผลผ่านเว็บไซต์ โดยใช้ XSLT

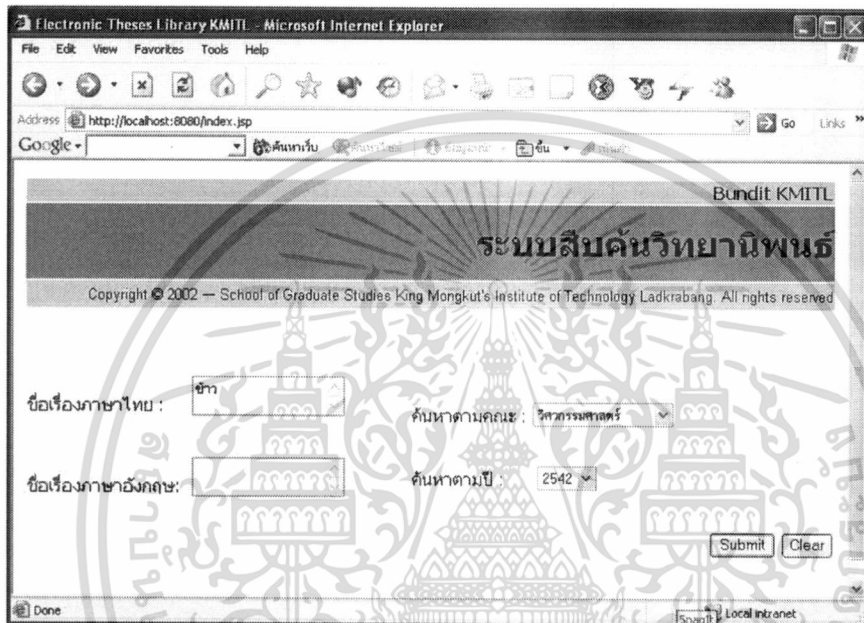


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5 ผลการทดลองและการวิเคราะห์

ในการทดลองนี้ทำการทดลองค้นหาผ่านทางเว็บไซต์ซึ่งสามารถค้นหาได้ทั้งคำภาษาไทยและภาษาอังกฤษ โดยสามารถสืบค้นได้จาก ชื่อเรื่องภาษาไทย ชื่อเรื่องภาษาอังกฤษ และ ค้นหาตามคณะ และ ตามปีการศึกษา โดยเว็บไซต์และผลการค้นหาแสดงดังรูปด้านล่าง

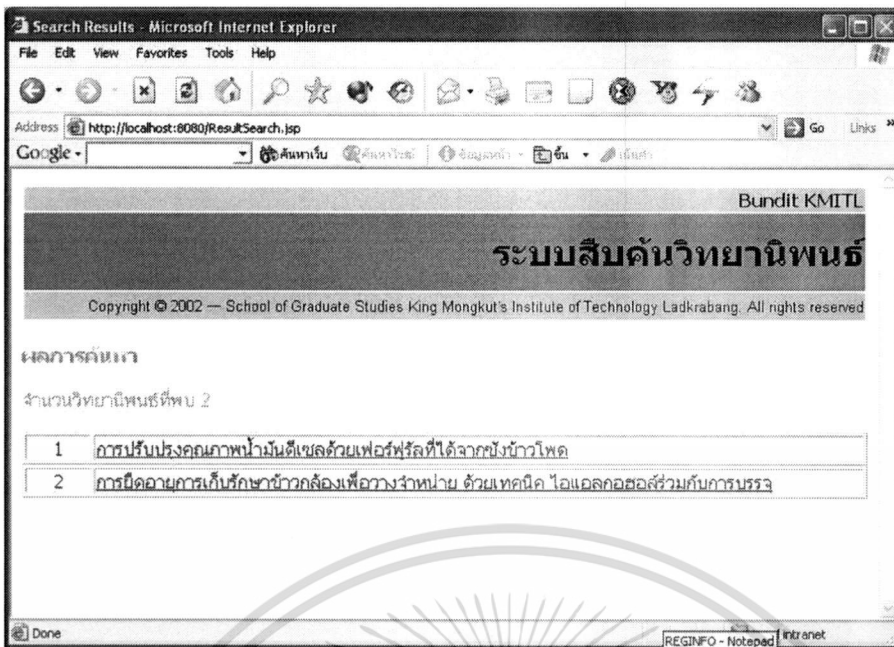
ผลของการทดลองสร้างต้นแบบห้องสมุดดิจิทัล ในที่นี้จะทำงานผ่านเว็บไซต์ซึ่งมีลักษณะดังรูป 5_1



รูปที่ 5_1 แสดงหน้าต่างของต้นแบบห้องสมุดดิจิทัล

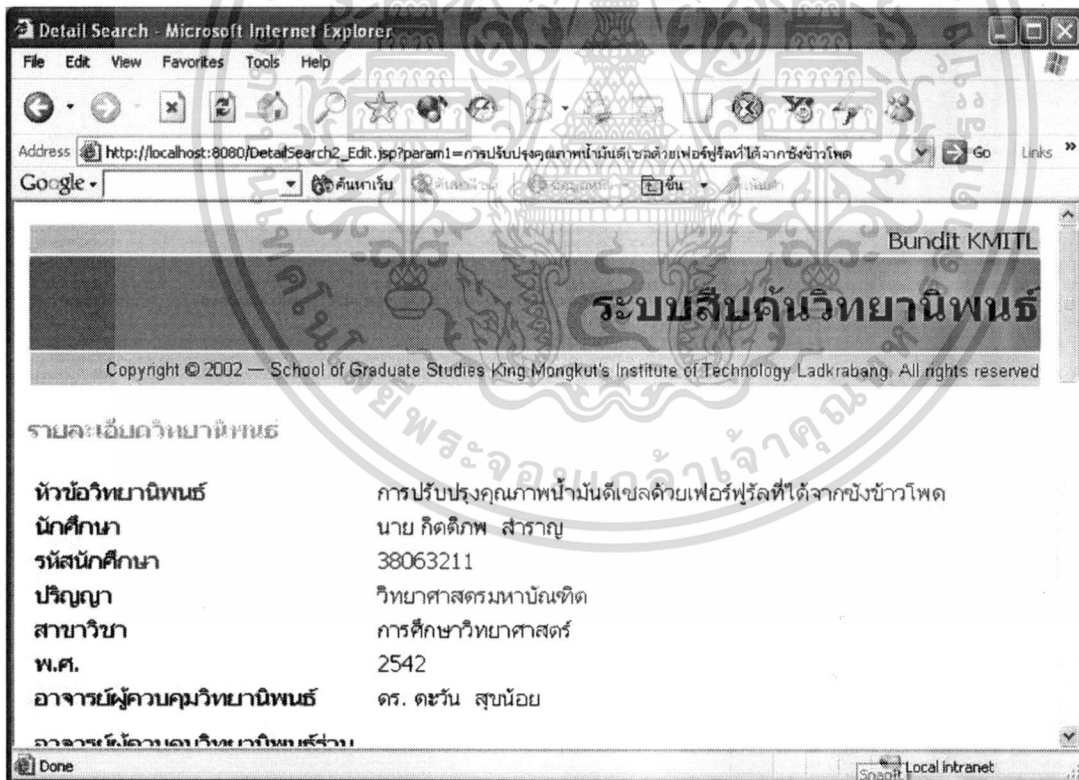
ในการสืบค้นสามารถสืบค้นข้อมูลของชื่อวิทยานิพนธ์ ชื่อผู้แต่ง คณะ ตามปี พ.ศ. ได้ โดยสามารถสืบค้นได้ทั้งเอกสาร โดยทำการค้นหาที่ต้องการค้นหาตลอดทั้งเอกสาร ผลของการค้นหาแสดงได้ดังรูปที่ 5_2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5_2 แสดงผลการค้นหาจากคำค้น

หลังจากนั้นสามารถดูข้อมูลทั้งหมดได้ซึ่งแสดงดังรูปที่ 5_3



รูปที่ 5_3 แสดงข้อมูลบทคัดย่อวิทยานิพนธ์ทั้งหมด

ในการค้นหาข้อมูลนี้จะทำการค้นหาคำค้นกับข้อมูลจริง ๆ ภายในฐานข้อมูลเลย จากการทดลองยังไม่เกิดความล่าช้ามากนักในการค้นหาเนื่องจากข้อมูลในฐานข้อมูลยังมีขนาดไม่มากนัก เอกสารนี้เป็นเอกสารที่ส่งมอบไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งยังมีให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต่อไปจะทำการเก็บข้อมูลที่เป็นบทความให้มากขึ้นเพื่อจะได้ทดสอบประสิทธิภาพการทำงานต่อไป

ในการพัฒนาขั้นต่อไปจะทำการเพิ่มข้อมูลในฐานข้อมูลและจัดทำโปรแกรมจัดการทำอินเด็กซ์ให้กับเอกสารที่ได้จัดเก็บไว้แล้ว



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6 สรุปงานวิจัยและการวิจัยในอนาคต

งานวิจัยนี้ทำการเก็บบทความของนักศึกษาที่สำเร็จการศึกษาไปแล้วตามฐานข้อมูลของบัณฑิตศึกษาที่ได้เก็บรวบรวมไว้ โดยเริ่มตั้งแต่ปี 2542 ถึง ปัจจุบัน จำนวนบทความที่จัดเก็บไว้แล้วมีประมาณ 230 บทความ โดยสามารถสืบค้นได้ตลอดทั้งเอกสารแบบซีเคิร์นซีเอสทำให้การทำงานกับฐานข้อมูลขนาดใหญ่มีปัญหาในความล่าช้า

งานวิจัยในอนาคตจะทำการจัดทำดัชนีแบบอัตโนมัติให้กับข้อมูลที่เก็บไว้ในฐานข้อมูลโดยสกัดส่วนที่เป็นคำสำคัญจากข้อมูล XML ที่ได้จัดเก็บไว้แล้วซึ่งจะทำให้ระบบสามารถตอบสนองต่อการค้นหาของผู้ใช้ได้ดียิ่งขึ้น มีความรวดเร็ว และความถูกต้องเพิ่มขึ้น



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ความหมายของแท็กที่ใช้เก็บวิทยานิพนธ์

วิทยานิพนธ์นี้จะเก็บในแท็ก Thesis ซึ่งภายในประกอบด้วยแท็กต่าง ๆ ดังนี้

1. ส่วนบทคัดย่อ

ประกอบด้วยแท็กต่าง ๆ ดังนี้

1.1 Abstract แสดงขอบเขตบทคัดย่อ

1.2 Title ใช้เก็บชื่อวิทยานิพนธ์ และใช้ แอตทริบิวต์ lang แสดงชนิดของภาษาที่เก็บ ซึ่งมีได้ 2 ค่า คือ en ใช้แทนภาษาอังกฤษ กับ th ใช้แทนภาษาไทย

1.3 Creator เป็นส่วนที่ใช้เก็บชื่อผู้แต่ง ประกอบด้วย

1.3.1 Initial คือ คำนำหน้าชื่อ

1.3.2 Firstname คือ ชื่อแรก

1.3.3 Middlename คือ ชื่อกลางถ้ามี

1.3.4 Lastname คือ นามสกุล

1.4 StudentID แสดงรหัสนักศึกษา

1.5 Degree แสดงชื่อปริญญา

1.6 Programme แสดงสาขาที่จบ

1.7 Faculty แสดงคณะที่จบการศึกษา

1.8 Year แสดงปีที่จบการศึกษา

1.9 Advisor แสดงชื่ออาจารย์ผู้ควบคุมวิทยานิพนธ์ ประกอบด้วย

1.9.1 1.9.1 Initial คือ คำนำหน้าชื่อ

1.9.2 Firstname คือ ชื่อแรก

1.9.3 Middlename คือ ชื่อกลางถ้ามี

1.9.4 Lastname คือ นามสกุล

1.10 Abstract.content แสดงส่วนที่เป็นบทคัดย่อ

1.11 Keywords แสดงขอบเขตของคำสำคัญ

1.12 Keyword แสดงคำสำคัญแต่ละคำในวิทยานิพนธ์

2. ส่วนของ ISBN

2.1 ISBN แสดงหมายเลข ISBN ของวิทยานิพนธ์

3. ส่วนกิตติกรรมประกาศ

3.1 Acknowledgements แสดงส่วนของกิตติกรรมประกาศ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ส่วนสารบัญ

4.1 TableofContent แสดงขอบเขตของสารบัญ ซึ่งประกอบด้วย

4.1.1 RowofTable แสดงขอบเขตของแถว ซึ่งในแต่ละแถวประกอบด้วย

4.1.1.1 Content แสดงหัวข้อ

4.1.1.2 PageNumber แสดงหมายเลขหน้าที่อ้างอิงถึง

5. ส่วนสารบัญตาราง

5.1 ListofTable แสดงขอบเขตของสารบัญตารางซึ่งประกอบด้วย

5.1.1 RowofTableT แสดงขอบเขตของแถว ซึ่งแต่ละแถวประกอบด้วย

5.1.1.1 Table.Number แสดงหมายเลขของตาราง

5.1.1.2 Content แสดงคำบรรยายขอตาราง

5.1.1.3 PageNumber แสดงหน้าที่ตารางนั้นอ้างอิงถึง

6. ส่วนสารบัญรูปภาพ

6.1 ListofFigure แสดงขอบเขตของสารบัญรูปภาพ ซึ่งประกอบด้วย

6.1.1 RowofTableF แสดงขอบเขตของแถว ซึ่งแต่ละแถวประกอบด้วย

6.1.1.1 Figure.Number แสดงหมายเลขของรูปภาพ

6.1.1.2 Content แสดงคำบรรยายรูปภาพ

6.1.1.3 PageNumber แสดงหน้าของรูปนั้น

7. ส่วนเนื้อหาวิทยานิพนธ์

7.1 Body แสดงขอบเขตของส่วนเนื้อหาวิทยานิพนธ์ ซึ่งประกอบด้วย

7.1.1 Chapter แสดงขอบเขตของบทในวิทยานิพนธ์ ซึ่งประกอบด้วย

7.1.1.1 Chapter.number แสดงเลขบท

7.1.1.2 Chapter.name แสดงชื่อบท

7.1.1.3 Pages แสดงขอบเขตของหน้า ซึ่งประกอบด้วย

7.1.1.3.1 Page แสดงขอบเขตของหน้า 1 หน้า ซึ่งประกอบด้วย

7.1.1.3.1.1 Page.number แสดงหมายเลขหน้า

7.1.1.3.1.2 Paragraph แสดงขอบเขตของย่อหน้า ซึ่ง

ประกอบด้วย

7.1.1.3.1.2.1 order แสดงลำดับของย่อหน้า

7.1.1.3.1.2.2 position แสดงตำแหน่งของย่อหน้า

7.1.1.3.1.2.3 paragraph.head แสดงขอบเขตของหัวข้อ

7.1.1.3.1.2.3.1 paragraph.headcontent แสดง

ส่วนเนื้อหาในส่วนหัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นเนื้อหา

7.1.1.3.1.2.4 paragraph.text แสดงส่วนของย่อหน้าที่

7.1.1.3.1.2.5 image แสดงรูปในย่อหน้า

7.1.1.3.1.2.6 table แสดงขอบเขตของตาราง ซึ่ง

ประกอบด้วย

7.1.1.3.1.2.6.1 tr แสดงขอบเขตของแถว

7.1.1.3.1.2.6.2 td แสดงข้อมูลในคอลัมน์

8. ส่วนบรรณานุกรม

8.1 Bibliography แสดงขอบเขตของบรรณานุกรม ซึ่งประกอบด้วย

8.1.1 Citation แสดงขอบเขตของบรรณานุกรม ซึ่งประกอบด้วย

8.1.1.1 Number.Citation แสดงหมายเลขของบรรณานุกรม

8.1.1.2 Name.Citation แสดงชื่อของเอกสารอ้างอิง

8.1.1.3 Name.Creator.FirstName แสดงชื่อแรกของผู้เขียน

8.1.1.4 Name.Creator.LastName แสดงนามสกุลของผู้เขียน

8.1.1.5 Name.Creator.Other แสดงชื่อและนามสกุลผู้เขียนคนอื่น ๆ

8.1.1.6 Public.Name แสดงสถานที่พิมพ์

8.1.1.7 Public.Year แสดงปีที่พิมพ์

8.1.1.8 Public.Date แสดงวันหรือเดือนที่พิมพ์

8.1.1.9 Citation.page แสดงหน้าของเอกสารอ้างอิง

9 ส่วนภาคผนวก

9.1 Appeindex แสดงขอบเขตของภาคผนวกซึ่งประกอบด้วย

9.1.1 Chapter แสดงขอบเขตของบท ซึ่งประกอบด้วย

9.1.1.1 Chapter.number แสดงหมายเลขของภาคผนวก

9.1.1.2 Chapter.name แสดงชื่อของภาคผนวก

9.1.1.3 Pages แสดงขอบเขตของหน้า ซึ่งประกอบด้วย

9.1.1.3.1 Page แสดงขอบเขตของหน้า 1 หน้า ซึ่งประกอบด้วย

9.1.1.3.1.1 Page.number แสดงหมายเลขหน้า

9.1.1.3.1.2 Paragraph แสดงขอบเขตของย่อหน้า ซึ่ง

ประกอบด้วย

9.1.1.3.1.2.1 order แสดงลำดับของย่อหน้า

9.1.1.3.1.2.2 position แสดงตำแหน่งของย่อหน้า

9.1.1.3.1.2.3 paragraph.head แสดงขอบเขตของหัวข้อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่สามารถเผยแพร่โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้