

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

แหล่งรวมสื่อและข้อมูลดิจิทัลบนเว็บ
DIGITAL COLLECTIONS ON THE WEB



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2545

b.....
i.....

เลขหมู่เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
เลขทะเบียน 49923 ห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
วัน,เดือน,ปี 2 เม.ย. 2547

แหล่งรวมสื่อและข้อมูลดิจิทัลบนเว็บ
DIGITAL COLLECTIONS ON THE WEB



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2545

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

๒๕๔๕

ปริญญานิพนธ์ปีการศึกษา 2545

ภาควิชา วิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง แหล่งรวมสื่อและข้อมูลดิจิทัลบนเว็บ

DIGITAL COLLECTIONS ON THE WEB

ผู้จัดทำ

1. นาย สุรพันธ์ อัสสรรัตน์ รหัสประจำตัว 42010401

2. นาย อาทิตย์ สิริอารีย์ รหัสประจำตัว 42010441



อาจารย์ที่ปรึกษา

(ดร. วิศิษฐ์ หิรัญกิตติ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แหล่งรวมสื่อและข้อมูลดิจิทัลบนเว็บ

นาย สุรพันธ์ อัสสรณ์ 42010401
 นาย อาทิตย์ สิริอารีย์ 42010441
 ดร. วิศิษฐ์ หิรัญกิตติ อาจารย์ที่ปรึกษา
 ปีการศึกษา 2545

บทคัดย่อ

ในปัจจุบันอินเทอร์เน็ตเป็นที่นิยมและแพร่หลายอย่างมาก ผู้ใช้บริการอินเทอร์เน็ตสามารถค้นหาและได้รับข่าวสารข้อมูลต่าง ๆ มากมายจากอินเทอร์เน็ต อินเทอร์เน็ตจึงเป็นแหล่งรวบรวมข้อมูลและสื่อดิจิทัลขนาดใหญ่ที่สุดของโลก อย่างไรก็ตามข้อมูลและสื่อดิจิทัลต่างๆ ในอินเทอร์เน็ตที่มีอยู่มากมายยังคงอยู่กันอย่างกระจัดกระจาย

เพื่อแก้ไขปัญหาดังกล่าว โครงการนี้จึงได้พัฒนาเว็บไซต์เพื่อเป็นแหล่งรวบรวมข้อมูลและสื่อดิจิทัลทุกรูปแบบที่เป็นที่ต้องการของผู้ใช้บริการไว้ในแหล่งเดียวกัน เพื่อให้ผู้ใช้สามารถจัดเก็บ ค้นหา และคัดลอกสื่อและข้อมูลได้อย่างหลากหลายมากมายที่อยู่บนอินเทอร์เน็ต โดยเชื่อมผ่านทางเว็บเบราว์เซอร์

ระบบทั้งหมดเป็นเว็บแอปพลิเคชันที่พัฒนาโดยใช้ภาษา เซิร์ฟเล็ท ที่ทำหน้าที่จัดการบริหารจัดการเก็บสื่อข้อมูลและสื่อดิจิทัลในรูปแบบของไฟล์ โดยที่ข้อมูลและสื่อดิจิทัลเหล่านี้จะได้รับการอธิบายความหมาย ซึ่งจัดเก็บไว้ในรูปแบบของเอกสารอิเล็กทรอนิกส์ ระบบมีการบริหารฐานข้อมูลด้วยคีย์ โดย มี เอกซ์เอ็มแอล เอกซ์เทนเซอร์ช่วยในการจัดเก็บกลุ่มของเอกสารอิเล็กทรอนิกส์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Digital Collections on the web

Surapun Assaratt

Arthit Siriaree

Dr. Visit Hirankitti Adviser

ABSTRACT

At present the Internet is very popular and widely used. People can find enormous information from the Internet. Undoubtedly the Internet is the largest source of information and digital contents in the world. However, the information is scattered and stored in different places on the Internet, this brings some difficulty for the user to find his/her intended information easily. It would be better if there is one place where useful information are organized and stored together.

The attempt in this thesis is to develop a web site which gathers all forms of digital information and contents, e.g. images, sound, videos, etc. in one place. The user can access to the site in order to store, copy, and search this information by themselves using a web browser.

The web site is developed as a web application using Java Servlet. The application allows the user to store and manage information and digital content in the form of files. Each content is described by an XML(Extensible Mark-up Language) document, in order to provide meta information about the content, e.g. information about the author, type and size of the content. These XML documents are stored and managed by DB2 using XML Extender so that we can search, store, copy, a digital contents form our web application.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จได้ด้วยดีสืบเนื่องมาจากการได้รับความช่วยเหลือและความร่วมที่ดีจากบุคคลต่างๆหลายฝ่าย บุคคลแรกที่มีส่วนอย่างมากในการทำให้ปริญญานิพนธ์ฉบับนี้สำเร็จได้ก็คือ ดร. วิศิษฐ์ หิรัญกิตติ ซึ่งเป็นอาจารย์ที่ปรึกษาปริญญานิพนธ์ของผู้เขียน อาจารย์ให้ความเอาใจใส่ คำแนะนำ และความช่วยเหลือเสมอมาตั้งแต่เริ่มต้นจนกระทั่งปริญญานิพนธ์สำเร็จได้ด้วยดี จึงต้องขอขอบพระคุณเป็นอย่างยิ่ง

บุคคลที่สองที่ต้องขอขอบคุณเป็นอย่างมากก็คือ พี่อู๋ที่ได้ให้คำแนะนำมาโดยตลอด และเพื่อนๆ ทุกคนที่คอยช่วยเหลือ ให้คำแนะนำ และให้กำลังใจ ช่วยแก้ไขปัญหาร่วมกันตลอดมา รวมทั้งต้องขอบคุณ ห้อง Intelligence Communication Lab (ICL) ที่เอื้อเฟื้อสถานที่ในการทำปริญญานิพนธ์นี้

สุดท้ายขอขอบพระคุณบุคคลสำคัญที่สุดที่ทำให้ข้าพเจ้ามีวันนี้ก็คือ บิดา มารดา อันเป็นที่เคารพรักยิ่ง ซึ่งได้เลี้ยงดูผู้เขียนมาเป็นอย่างดี พร้อมทั้งให้โอกาสในการศึกษาอย่างเต็มที่และยังให้กำลังใจเอาใจใส่เสมอมาในทุกด้านอันหาที่เปรียบมิได้ ข้าพเจ้าขอระลึกในพระคุณอันสุดประมาณและขอกราบขอบพระคุณมา ณ ที่นี้

สุรพันธ์ อัสสะรัตน์
อาทิตย์ สิริอารีย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VII
สารบัญรูปภาพ	VIII
บทที่ 1 บทนำ	1
1.1 ความสำคัญ และที่มา	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตของงานวิจัย	2
1.4 วิธีการดำเนินงาน	2
บทที่ 2 เซิร์ฟเล็ต (Servlet)	3
2.1 จาวาเซิร์ฟเล็ต	3
2.2 ข้อแตกต่างของเซิร์ฟเล็ตเมื่อเทียบกับซีจีไอ (CGI)	3
2.3 การอ่านข้อมูลจากเว็บ	4
2.4 สร้างเซิร์ฟเล็ตอ่านข้อมูลจากฟอร์ม	6
2.5 การใช้งานคุกกี้ด้วยเซิร์ฟเล็ต	7
2.6 เจดีบีซี (JDBC) ช่องทางติดต่อฐานข้อมูลด้วยจาวา	9
บทที่ 3 เอ็กซ์เอ็มแอล และฐานข้อมูลที่สามารถทำงานกับเอ็กซ์เอ็มแอลได้	14
3.1 ภาษาเอ็กซ์เอ็มแอล	14
3.2 ลักษณะ โครงสร้างของเอกสารเอ็กซ์เอ็มแอล	14
3.3 ดีทีดีข้อกำหนดโครงสร้างเอ็กซ์เอ็มแอล	17
3.3.1 ดีทีดี คืออะไร	17
3.3.2 ประโยชน์ของเอกสารดีทีดี	18
3.4 เอกสารเอ็กซ์เอ็มแอลที่ถูกต้อง	18
3.5 เอกสารเอ็กซ์เอ็มแอลดีและถูกต้อง	20
3.6 ประโยชน์ของเอ็กซ์เอ็มแอล	21
3.7 ฐานข้อมูลที่ทำงานร่วมกับเอ็กซ์เอ็มแอล	21
3.7.1 เหตุผลที่ใช้เอ็กซ์เอ็มแอลและ ดีบีทู	21
3.7.2 การทำงานร่วมกันของเอ็กซ์เอ็มแอลและ ดีบีทู	22
3.7.3 เอ็กซ์เอ็มแอลคอลัมน์	23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

3.7.4 ไซด์เทเบิล	24
3.7.5 ไฟล์ดีเอดี	24
3.7.6 ขั้นตอนการทำให้ฐานข้อมูลสามารถทำงานกับเอ็กซ์เอ็มแอล	24
3.7.7 สร้างตารางเอ็กซ์เอ็มแอล	26
3.7.8 จัดเก็บไฟล์ดีทีดีลงฐานข้อมูล	26
3.7.9 ขั้นตอนการทำให้คอมลัมน์สามารถทำงานกับเอ็กซ์เอ็มแอล	27
บทที่ 4 การสร้าง และการออกแบบ	29
4.1 บทนำ	29
4.2 เอกสารเอ็กซ์เอ็มแอล	30
4.2.1 วิเคราะห์และออกแบบโครงสร้างเอกสารเอ็กซ์เอ็มแอล	30
4.2.2 โครงสร้างดีทีดี	33
4.3 การจัดหมวดหมู่ในการจัดเก็บไฟล์ข้อมูล	34
4.4 การสร้างฐานข้อมูลด้วยคีย์	50
4.4.1 การสร้างตารางและจัดการกับเอกสารเอ็กซ์เอ็มแอลด้วยคีย์	50
4.4.2 ไฟล์ดีเอดี	50
4.4.3 ตารางต่างๆบนฐานข้อมูล	53
4.4.4 การจัดเก็บเอกสารเอ็กซ์เอ็มแอลบนฐานข้อมูล	53
4.5 การทำงานของระบบ	54
4.5.1 การสมัครสมาชิก	54
4.5.2 การเข้าสู่ระบบ	54
4.5.3 การจัดเก็บสื่อและข้อมูลดิจิทัล	54
4.5.4 การเรียกดูสื่อและข้อมูลดิจิทัล	55
4.5.5 การแสดงผลสื่อและข้อมูลดิจิทัล	55
4.5.6 การดาวน์โหลดสื่อและข้อมูลดิจิทัล	56
4.5.7 การค้นหาสื่อและข้อมูลดิจิทัล	56
4.5.8 การลบข้อมูลสื่อและข้อมูลดิจิทัล	56
4.5.9 การลบหมวดหมู่ของสื่อและข้อมูลดิจิทัล	57
4.5.10การสร้างหมวดหมู่ของสื่อและข้อมูลดิจิทัล	57
4.5.11การออกจากระบบ	57
4.6 ระบบค้นหา	57
4.6.1 การแปลความหมายของคีย์เวิร์ด	58
4.6.2 การจัดอันดับของเอกสาร	59

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

4.6.3 การคำนวณในการจัดลำดับในเอกสาร	61
บทที่ 5 ผลการทำงานของระบบ	64
5.1 หน้าจอหลักของเว็บไซต์	64
5.2 การสมัครสมาชิกใหม่	65
5.3 การจัดเก็บข้อมูลสื่อ และข้อมูลดิจิทัล	66
5.4 การเรียกดูไฟล์ข้อมูลบนเซิร์ฟเวอร์	71
5.5 การแสดงผลข้อมูล	72
บทที่ 6 บทสรุป และวิจารณ์	73
6.1 สรุปและวิจารณ์	73
6.2 ปัญหาที่เกิดขึ้น	73
6.3 แนวทางการพัฒนา	74
ภาคผนวก ก ไฟล์ข้อมูลที่สามารถแสดงผลทางหน้าเว็บ	75
ก.1 ไฟล์ที่เล่นได้บนควิกไทม์ 6	75
ก.2 ไฟล์ที่เล่นได้บนวินโดวส์มีเดียเพลย์เยอร์	77
ภาคผนวก ข เวกเตอร์โมเดลสำหรับหาความคล้ายคลึงกันของเอกสาร	78
ข.1 บทนำ	78
ข.2 สัญลักษณ์	78
ข.3 ทฤษฎีและการคำนวณ	78
บรรณานุกรม	81

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

รูปที่ 3-1 เอกสารอิเล็กทรอนิกส์ที่มีการสร้างตารางไฮค์เทเบิล	24
รูปที่ 4-1 แสดงโครงสร้างทั้งหมดของโครงการ	29
รูปที่ 4-2 แสดงหมวดหมู่ Upload	34
รูปที่ 4-3 แสดงหมวดหมู่ Animation	35
รูปที่ 4.4 แสดงหมวดหมู่ Document	36
รูปที่ 4.5 แสดงหมวดหมู่ Game	37
รูปที่ 4-6 แสดงหมวดหมู่ Image	39
รูปที่ 4-7 แสดงหมวดหมู่ Music	41
รูปที่ 4-8 แสดงหมวดหมู่ News	43
รูปที่ 4-9 แสดงหมวดหมู่ Program	46
รูปที่ 4-10 แสดงหมวดหมู่ Sound	47
รูปที่ 4-11 แสดงหมวดหมู่ Video	48
รูปที่ 5-1 หน้าจากหลักของเว็บไซต์	64
รูปที่ 5-2 แสดงหน้าเว็บการสมัครสมาชิกใหม่	65
รูปที่ 5-3 แสดงหน้าเว็บการเลือกหมวดหมู่ในอันดับแรก	66
รูปที่ 5-4 แสดงหน้าเว็บการเลือกหมวดหมู่ในขั้นที่สอง	67
รูปที่ 5-5 แสดงหน้าเว็บการเลือกหมวดหมู่ในขั้นที่สอง(ต่อ)	68
รูปที่ 5-6 แสดงหน้าเว็บการเลือกหมวดหมู่ในขั้นที่สอง(ต่อ)	69
รูปที่ 5-7 แสดงฐานข้อมูลที่ทำกรจัดเก็บเอกสารอิเล็กทรอนิกส์ โดยเครื่องมือของดีบีทู(1)	69
รูปที่ 5-8 แสดงฐานข้อมูลที่ทำกรจัดเก็บเอกสารอิเล็กทรอนิกส์ โดยเครื่องมือของดีบีทู(2)	70
รูปที่ 5-9 แสดงการเรียกดูไฟล์ข้อมูลที่อยู่บนเซิร์ฟเวอร์โดยผู้ให้บริการทั่วไป	70
รูปที่ 5-10 แสดงการเรียกดูไฟล์ที่อยู่บนเซิร์ฟเวอร์โดยผู้ดูแลระบบ	71
รูปที่ 5-11 แสดงหน้าแสดงผลในการเล่นไฟล์ข้อมูลด้วยวินโดวมีเดียเพลเยอร์	72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ 2-1 แสดงรายละเอียดที่เกี่ยวข้องกับการใช้งานคุกกี้	9
ตารางที่ 3-1 แสดงชนิดของข้อมูลในคอลัมน์ต่างในดีทีดีเรฟเฟอเรนซ์	27
ตารางที่ 4-1 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Upload	35
ตารางที่ 4-2 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Animation	37
ตารางที่ 4-3 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Document	38
ตารางที่ 4-4 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Game	39
ตารางที่ 4-5 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Image	40
ตารางที่ 4-6 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Art	40
ตารางที่ 4-7 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ People	40
ตารางที่ 4-8 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Places	41
ตารางที่ 4-9 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Music	42
ตารางที่ 4-10 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ News	44
ตารางที่ 4-11 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Entertainment	44
ตารางที่ 4-12 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ SCI-TECH	44
ตารางที่ 4-13 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Sports	45
ตารางที่ 4-14 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ World	45
ตารางที่ 4-15 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Program	46
ตารางที่ 4-16 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Driver	47
ตารางที่ 4-17 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Mobile	47
ตารางที่ 4-18 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Sound	48
ตารางที่ 4-19 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Video	49
ตารางที่ 4-20 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Movie	49
ตารางที่ 4-21 แสดงการคำนวณค่าความคล้ายคลึงด้วยค่าน้ำหนัก	62
ตารางที่ 4-22 แสดงการคำนวณค่าความคล้ายคลึงด้วยค่า <i>idf</i>	63
ตาราง ก-1 ชนิดของไฟล์ที่ วินโดวส์ มีเดียเพลย์เยอร์เล่นได้	77

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มา

สืบเนื่องมาจากเทคโนโลยีทางด้านอินเทอร์เน็ตในปัจจุบันได้เข้ามามีบทบาทต่อการดำรงชีวิตประจำวันของผู้คนจำนวนมาก ทั้งในแง่ของการติดต่อสื่อสารส่วนตัว และการติดต่อทางด้านธุรกิจ เป็นแหล่งเก็บรวบรวมข้อมูลและข่าวสารต่างมากมาย การให้บริการข้อมูลบนเครือข่ายอินเทอร์เน็ตมีการนำเสนอในรูปแบบต่าง ๆ กันออกไป เช่น บริการให้ข้อมูลการศึกษา บริการส่งจดหมายอิเล็กทรอนิกส์ เกิดการแลกเปลี่ยนข้อมูลสื่อและข้อมูลดิจิทัลกันมากมาย มีหลายเว็บไซต์ที่ได้ให้บริการด้านการจัดเก็บไฟล์ข้อมูล อาทิเช่น ไฟล์รูปภาพ ไฟล์เสียง ไฟล์เอกสารเวิร์ด และโปรแกรมต่างๆ อินเทอร์เน็ตจึงเป็นที่นิยมอย่างมาก

อย่างไรก็ตามข้อมูลที่เกิดขึ้นต่างมากมายกระจัดกระจายไปตามเว็บไซต์มากมายโดยที่ไม่อาจที่จะทราบได้ว่าไฟล์เหล่านั้นไปเก็บไว้ที่ใด โดยใครเมื่อไหร่ เนื้อหาข้อมูลของไฟล์นั้นเป็นอย่างไร เช่น ไฟล์เพลงของวงดนตรีชื่ออะไร ไฟล์หนังเรื่องนี้เป็นเรื่องอะไร เนื้อเรื่องเป็นอย่างไรเราไม่สามารถที่จะทราบถึงข้อมูลต่างๆเหล่านี้ได้เลย ทำให้ไม่สามารถค้นหาได้อีกหรือค้นหาได้แต่ไม่ได้ข้อมูลที่ต้องการจริงๆ จากการที่ข้อมูลต่างๆกระจัดกระจายไปตามเว็บไซต์แบบ ไม่มีความหมายหากมีการจัดเก็บข้อมูลเหล่านี้ให้เป็นที่เดียวกันและมีความหมาย จะทำให้ค้นหาทำได้ง่ายขึ้น

วิทยานิพนธ์นี้จึงเป็นอีกแนวทางหนึ่งที่จะทำการพัฒนาเว็บเซิร์ฟเวอร์ โดยมีแนวความคิดสร้างแหล่งรวบรวมข้อมูลและสื่อดิจิทัลทุกประเภทในรูปแบบของไฟล์ผ่านทางหน้าเว็บไซต์ ให้ความสามารถในการจัดเก็บไฟล์จากผู้เข้ามาใช้บริการ เพื่อเป็นการรวบรวมไฟล์หลากหลายประเภทมาไว้บนเซิร์ฟเวอร์ ไฟล์ข้อมูลและสื่อที่ถูกจัดเก็บไว้นั้นจะถูกทำให้มีความหมายในตัวเองด้วย นอกจากจัดเก็บแล้วเซิร์ฟเวอร์ยังมีความสามารถที่จะค้นหา เรียกดู หรือเล่นไฟล์ที่เป็นภาพและเสียงได้ การจัดเก็บข้อมูลผ่านเว็บนั้นใช้เทคโนโลยีของจาวาเซิร์ฟเล็ต (Java Servlet) เป็นเทคโนโลยี ที่มีรูปแบบการทำงานเหมือนกับเทคโนโลยีซีจีไอที่มีการใช้งานอยู่ เป็นตัวกลางในการจัดเก็บข้อมูล ค้นหา และเรียกดูข้อมูลต่างๆ โดยที่ผู้ใช้จะต้องทำการกรอกข้อมูลต่างๆเพื่อใช้สื่อความหมายของข้อมูลนั้นที่จะทำการจัดเก็บ โดยจะจัดเก็บไปตามหมวดหมู่ของชนิดข้อมูลนั้นๆ เช่น ข้อมูลที่เป็นเพลง แนวดนตรีเป็นแนวป๊อปก็จะอยู่ในหมวดหมู่เพลง/ ป๊อป ซึ่งผู้ใช้ต้องทำการเลือกให้ถูกหมวดหมู่ โดยในขั้นตอนการจัดเก็บข้อมูลจะใช้มาตรฐานที่เรียกว่าเอ็กซ์เอ็มแอล (XML หรือ Extension markup language) เป็นมาตรฐานที่ใช้ในการส่งข้อมูลผ่านเว็บทำการจัดเก็บลงในฐานข้อมูลที่สามารถทำงานกับเอกสารเอ็กซ์เอ็มแอลได้ ข้อมูลที่อยู่ในแท็กเอกสารเอ็กซ์เอ็มแอลนั้นจะถูกเก็บลงในตาราง ด้วยเทคโนโลยีของฐานข้อมูลดีบีทู เอ็กซ์เทนเคอร์ ที่มีให้เลือกใช้ในโอบีเอ็มดีบีทู ทำให้สามารถที่จะเรียกดู แก้ไขเปลี่ยนแปลง ความหมายของข้อมูลในเอกสารเอ็กซ์เอ็มแอลได้ง่าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของงานวิจัย

1. สร้างเว็บไซต์เวอร์ชันที่เป็นแหล่งรวบรวมสื่อและข้อมูลดิจิทัลให้บริการจัดเก็บ ค้นหา เล่น และดาวน์โหลดสื่อและข้อมูลดิจิทัลผ่านทางอินเทอร์เน็ตได้
2. เพื่อศึกษาการพัฒนาเว็บแอปพลิเคชัน โดยใช้ภาษาซีฟเลท
3. ศึกษากระบวนการจัดการฐานข้อมูลร่วมกับเอกสารอิเล็กทรอนิกส์

1.3 ขอบเขตของงานวิจัย

1. ศึกษาและพัฒนาการทำงานของซีฟเลท และภาษาจาวา ซึ่งเป็นภาษาหลักในการพัฒนาเพื่อนำมาประยุกต์ใช้ในการพัฒนาเว็บไซต์
2. ศึกษาและพัฒนาเว็บแอปพลิเคชันที่ให้บริการความสะดวกในการจัดเก็บ ค้นหา ดาวน์โหลดสื่อและข้อมูลดิจิทัลได้
3. ศึกษาและพัฒนาเกี่ยวกับการทำงานของเอกสารอิเล็กทรอนิกส์ เพื่อใช้แสดงและกำกับแฟ้มเอกสารข้อมูลดิจิทัล
4. ศึกษาและพัฒนาฐานข้อมูลที่สามารถจัดการจัดเก็บและใช้งานเอกสารอิเล็กทรอนิกส์โดยใช้คิวกับอิเล็กทรอนิกส์เทนเดอร์สำหรับจัดการกับเอกสารอิเล็กทรอนิกส์ได้ด้วยได้
5. ศึกษาและพัฒนาเสิร์จเอนจิน ที่ใช้ในการค้นหาข้อมูลบนฐานข้อมูลทำงานร่วมกับเอกสารอิเล็กทรอนิกส์ของเว็บไซต์ได้

1.4 วิธีการดำเนินงาน

1. ศึกษาการทำงานของซีฟเลท และความรู้พื้นฐานการรับส่งข้อมูลผ่านเว็บ
2. ศึกษาอิเล็กทรอนิกส์
3. ศึกษาการบริหารระบบฐานข้อมูลด้วยคิว เพื่อสร้างฐานข้อมูลให้ระบบ
4. ศึกษาและใช้งานเครื่องมืออิเล็กทรอนิกส์เทนเดอร์ เพื่อสร้างฐานข้อมูลเอกสารอิเล็กทรอนิกส์
5. ออกแบบระบบแหล่งรวมสื่อและข้อมูลดิจิทัล
6. พัฒนาระบบให้สามารถจัดเก็บ ค้นหา เรียกดู ตามที่ได้ออกแบบไว้
7. ทดสอบการทำงานของระบบ และแก้ไขปัญหที่เกิดขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

เซิร์ฟเล็ต

2.1 จาวาเซิร์ฟเล็ต

จาวาเซิร์ฟเล็ต คือเทคโนโลยี ที่เกิดขึ้นมาจากจาวา เพื่อใช้งานทางด้านการสร้างโปรแกรมในรูปแบบซีจีไอสำหรับทำงานบนเว็บเซิร์ฟเวอร์ เพื่อให้ผลการแสดงเว็บผู้ผู้อ่าน มีการเปลี่ยนแปลงตามการใช้งาน (On The Fly) ซึ่งเป็นเทคโนโลยีที่สืบเนื่องมาจากเหตุผลดังนี้

- ความต้องการให้หน้าเว็บเปลี่ยนตามข้อมูลที่ผู้อ่านส่งให้ เช่น เว็บที่ให้บริการระบบค้นหา (Search Engine) ทั้งหลาย ที่ได้รับข้อมูลจากผู้ใช้ คือข้อความที่ต้องการค้นหา และแสดงผลลัพธ์หน้าเว็บ ตามข้อมูลที่ค้นหาเจอแก่ผู้ใช้
- ความต้องการให้หน้าเว็บเปลี่ยนตามข้อมูลที่มีการเปลี่ยนแปลงบ่อยครั้ง เช่น เว็บที่เกี่ยวข้องกับการรายงานข่าว หรือสภาพอากาศ จะเห็นว่าข้อมูลที่รายงานสู่ผู้ใช้ ดังกล่าว มีการเปลี่ยนแปลงอยู่ตลอดเวลา อาจเป็นรายชั่วโมง หรือต่อนาทีก็ได้
- ความต้องการให้หน้าเว็บเปลี่ยนตามระบบฐานข้อมูล เช่น ระบบควบคุมคลังสินค้า ที่อาศัยการเก็บข้อมูลในระบบฐานข้อมูล สามารถแสดงผลจำนวนสินค้า ตามข้อมูลที่มีอยู่จริงในฐานข้อมูลออกสู่ผู้ใช้บนระบบเว็บ

2.2 ข้อแตกต่างของเซิร์ฟเล็ต เมื่อเทียบกับ ซีจีไอ

เซิร์ฟเล็ต เป็นเทคโนโลยี ที่มีรูปแบบการทำงานเหมือนกับ เทคโนโลยีซีจีไอ ที่มีการใช้งานอยู่ แต่เนื่องจาก เซิร์ฟเล็ต ถูกพัฒนามาจากเทคโนโลยีจาวาอีกที ดังนั้น ต่อไปนี้คือข้อเปรียบเทียบซึ่งแสดงให้เห็น ถึงความแตกต่างระหว่างเซิร์ฟเล็ตและซีจีไอ

- ประสิทธิภาพ การทำงานของซีจีไอจะมีการสร้างโปรเซสใหม่ทุกครั้งที่มีการเรียกใช้งานจากผู้ใช้ ดังนั้นหากมีการเรียกใช้งานจำนวนหลายครั้งในช่วงเวลาเดียวกัน ก็จะมีจำนวนโปรเซสเกิดขึ้นเท่ากับจำนวนครั้งที่เรียกใช้งาน โปรเซสจะมีขนาดและรูปแบบการทำงานที่เหมือนกันทำให้กินเวลาในการทำงานตามจำนวนโปรเซส แต่สำหรับเซิร์ฟเล็ตแล้วระบบเวอร์ชวลแมชีน จะถูกดึงขึ้นมาทำงานก่อน เพื่อทำให้เกิดโปรเซสตามลักษณะการเรียกใช้ของผู้ใช้เช่นกัน แต่การทำงานของเซิร์ฟเล็ตจะเป็นรูปแบบเธรด(Thread) แทนที่จะเรียกว่าโปรเซส ซึ่งแต่ละเธรดขนาดการทำงานเล็กกว่าเดิม และกินระยะเวลาการทำงานน้อยกว่า จึงทำให้รูปแบบของเซิร์ฟเล็ต มีประสิทธิภาพกว่านั่นเอง
- ความสะดวก หากคุณสามารถเขียนโปรแกรม ด้วยภาษาซีหรือ จาวาอยู่แล้ว จะเกิดความสะดวกและรวดเร็วในการสร้างเซิร์ฟเล็ตมากกว่า แทนที่จะเริ่มศึกษาภาษาคำสั่งใหม่ๆ และเซิร์ฟเล็ตก็มีความสามารถเพียงพอสำหรับการใช้งานบนเว็บอยู่แล้ว เช่น การรับข้อมูลจากฟอร์ม การอ่านค่า

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์หรือการสงวนสิทธิ์ในเนื้อหา ไม่อนุญาตให้นำไปใช้โดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอยู่ภายใต้เงื่อนไขของเอกสารนี้ทุกประการที่มีการนำไปใช้

- ความสามารถเซิร์ฟเล็ทมีข้อดีที่เห็นได้ชัดคือ สามารถติดต่อกับเว็บเซิร์ฟเวอร์ได้โดยตรง ในขณะที่ ซีจีไอ ปกติทำไม่ได้ (ยกเว้นรุ่นใหม่ๆ) และเซิร์ฟเล็ทยังสามารถแชร์ข้อมูลการทำงาน ร่วมกับโปรแกรมอื่นๆได้ด้วย
- การโยกย้าย เนื่องจากเซิร์ฟเล็ทถูกสร้างมาจากจาวา ซึ่งเป็นเทคโนโลยี ที่สามารถถูกใช้งาน ณ. ที่ใดก็ได้โดยไม่จำกัดรูปแบบ ดังนั้น หากพัฒนาเซิร์ฟเล็ทเพื่อใช้งานแล้วสามารถที่จะโยกย้าย ผลผลิตไปยัง ระบบอื่นๆ ได้ โดยไม่ต้องปรับปรุงหรือเปลี่ยนแปลง
- ไม่แพง ในปัจจุบันเว็บเซิร์ฟเวอร์ที่สนับสนุนการทำงานเซิร์ฟเล็ท มีอยู่มากมาย ทั้งที่เป็นแบบฟรี ความโหลด (Download) และซื้อ-ขายกัน นอกจากนี้ในอนาคตแผนการผลิตเว็บเซิร์ฟเวอร์อื่นๆ ยังมีการสนับสนุนการใช้งานเซิร์ฟเล็ทอีกด้วย

2.3 การอ่านข้อมูลจากเว็บ

สำหรับการใช้งานเว็บผ่านบราวเซอร์ โดยทั่วไปแล้ว ปกติในช่องที่ใช้กำหนดข้อมูลยูอาร์แอล (URL) จะประกอบด้วยชื่อเซิร์ฟเวอร์ ชื่อไดเรกทอรี และชื่อไฟล์เว็บเพจที่ต้องการดู เช่น หากข้อมูลยูอาร์แอล คือ <http://www.javacentrix.com/tutorials/api/index.htm> นั้นหมายถึงต้องการอ่านเว็บเพจที่ชื่อ index.htm จากไดเรกทอรี tutorials/api จากเว็บเซิร์ฟเวอร์ที่ชื่อ www.javacentrix.com

สำหรับยูอาร์แอล นั้นนอกจาก จะใช้ชี้ตำแหน่งไฟล์ที่ต้องการเรียกดูแล้ว ยังสามารถใช้สำหรับส่งข้อมูลให้กับระบบเซิร์ฟเวอร์ เพื่อใช้ในการทำงาน หรือดำเนินการอย่างใดอย่างหนึ่ง ลักษณะยูอาร์แอล ประเภทนี้หลังจากชื่อไฟล์แล้วจะตามด้วยเครื่องหมายคำถาม '?' และข้อมูลที่ต้องการส่งไปให้เซิร์ฟเวอร์ ซึ่งเรียกโดยทั่วไปว่าการส่งข้อมูลแบบ GET โดยมีรูปแบบคือ

URL?Parameter1=Data1&Parameter2=Data2

จากรูปแบบการส่งข้อมูลผ่านยูอาร์แอล โดย Parameter1 และ Parameter2 คือชื่อพารามิเตอร์ที่ถูกเซิร์ฟเวอร์อ้างอิงได้ ในขณะที่ Data1 และ Data2 คือข้อมูลที่อยู่ในพารามิเตอร์แต่ละตัว หากต้องการส่งข้อมูลมากกว่า 2 ตัว ก็จะเพิ่มต่อท้ายไปได้เรื่อยๆ เช่น

URL?Parameter1=Data1&Parameter2=Data2&Parameter3=Data3&Parameter4=Data4

ตัวอย่างเช่น

<http://www.javacentrix.com/index.htm?Name=JAVACENTRIX&Description=ThaiJavaWebSite>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าเครื่องหมาย “?” คือ ข้อมูลยูอาร์แอล ในขณะที่ส่วนที่เหลือคือ ข้อมูลที่ส่งให้กับเซิร์ฟเวอร์ อธิบายได้ดังนี้คือ เป็นเรียกใช้งานไฟล์ index.htm โดยที่มีการส่งข้อมูลไปให้กับเซิร์ฟเวอร์ 2 ตัวคือ Name ที่มีข้อมูล "JAVACENTRIX" และ Description ที่มีข้อมูล "Thai Java Web Site" เซิร์ฟเวอร์จะนำข้อมูลเหล่านี้ไปใช้ประกอบการทำงานกับไฟล์ index.htm ซึ่งในที่นี้หมายถึง เซิร์ฟเวอร์นั่นเอง

หากผู้อ่านเคยใช้เว็บที่มีระบบค้นหาและพิจารณาในช่องยูอาร์แอล ก็จะสังเกตเห็นข้อมูลที่เป็น เช่น อธิบายไว้หลักจากคลิกที่ปุ่มค้นหา แต่ข้อเสียของการส่งข้อมูลรูปแบบนี้ คือ ผู้ใช้ต้องพิมพ์ชื่อ พารามิเตอร์ ให้ถูก และพิมพ์ข้อมูลให้ถูกรูปแบบด้วย อีกทั้งการใช้งานยังไม่เป็นความลับ ซึ่งถูกมองเห็นได้จากคนทั่วไป

นอกจากผู้ใช้เว็บจะต้องพิมพ์ข้อมูลที่ยูอาร์แอล เองแล้ว การส่งข้อมูลแบบ GET นี้ยังสามารถใช้แท็ก FORM ในเอชทีเอ็มแอล เพื่อสร้างรูปแบบการกรอกข้อมูลแก่ผู้ใช้ ที่เรียกว่าฟอร์ม เมื่อผู้ใช้ข้อมูลกรอกข้อมูลผ่านฟอร์ม และกดปุ่ม Submit ข้อมูลเหล่านั้นจะถูกบราวเซอร์ ประกอบเป็นยูอาร์แอล ในรูปแบบ GET ให้อย่างอัตโนมัติ โดยที่ผู้อ่านไม่จำเป็นต้องพิมพ์ยูอาร์แอลเองให้เสียเวลาเลย

ในทำนองเดียวกัน การส่งข้อมูลจากผู้ใช้ นอกจากจะส่งในรูปแบบ GET แล้วยังมีอีกรูปแบบหนึ่งคือรูปแบบ POST ซึ่งแทนที่จะส่งไปที่ตัวยูอาร์แอลโดยตรง แต่เป็นการส่งด้วยโปรโตคอลเอชทีทีพี ที่ใช้ในเว็บเป็นตัวจัดการ การส่งแบบนี้มีข้อดีที่ผู้อื่น ไม่สามารถล่วงรู้ได้ว่าลักษณะการส่งข้อมูลไปใช้เซิร์ฟเวอร์ ส่งข้อมูลอะไรไปและส่งโดยการใช้พารามิเตอร์ชื่ออะไร การส่งข้อมูลในแบบ POST ทำได้โดยการใช้แท็ก FORM แต่กำหนดวิธีการส่งเป็น POST นอกนั้นการสร้างฟอร์มก็ใช้การสร้างยังส่งข้อมูลแบบ GET ได้ด้วย หากสร้างฟอร์มด้วยการใช้แท็ก FORM แต่กำหนดวิธีการส่งเป็น GET

การสร้างฟอร์มแบบ GET และ POST

```
<HTML><HEAD><TITLE>Test Form </TITLE></HEAD>
<BODY>
  <FORM method="get" action="myParameter">
    ชื่อ <input type="text" name="name"> <BR>
    นามสกุล <input type="text" name="surname"> <BR>
    ชาย<input type="radio" name="sex" value="M">
    หญิง<input type="radio" name="sex" value="F"> <BR>
    ที่อยู่ <textarea name="add"></textarea><br>
    E-mail <input type="text" name="mail"> <BR>
    <input type="submit" name="Submit" value="Submit">
  </FORM>
</BODY>
</HTML>
```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากคำสั่งเซ็ทที่เอนแวลที่เห็นจากตัวอย่าง มีช่องสำหรับให้ผู้ใช้กรอก 3 ช่องคือเท็กชบ์ล็อกที่ชื่อ **name**, **surname** และ **mail** (สังเกตจากคำสั่งในเท็ก FORM ของ HTML) และมีการเรียกใช้รูปแบบการส่งข้อมูลแบบ GET (สังเกตจากในเท็กฟอร์ม แอททริบิวท์ `method="get"`) ผู้อ่านสามารถทดลองเปลี่ยนคำว่า `get` เป็นคำว่า `post` ได้หากต้องการส่งข้อมูลรูปแบบ POST จำไว้ว่าการส่งแบบ GET ข้อมูลจะถูกพร้อมกับยูอาร์แอล ในขณะที่ส่งแบบ POST ข้อมูลถูกส่งด้วยรูปแบบคำสั่งโปรโตคอลเอชทีทีพีแทน

2.4 สร้างเซิร์ฟเล็ตอ่านข้อมูลจากฟอร์ม

คำสั่งที่ใช้สำหรับอ่านข้อมูลจากผู้ใช้คือ `getParameter` ที่อยู่ในคลาส `HttpServletRequest`

```
import java.io.*;
import javax.servlet.*;
import javax.servlet.http.*;
import java.util.*;

public class ShowFormData extends HttpServlet
{ public void doGet(HttpServletRequest request, HttpServletResponse response)
  throws ServletException, IOException {
    response.setContentType("text/html");
    String title = "Retrieving Information from Web-Form";
    PrintWriter out = response.getWriter();
    out.println("<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0
Transitional//EN"> " +
        "<HTML> " +
        "<HEAD><TITLE>Hello WWW</TITLE></HEAD> " +
        "<BODY> " +
        "<H1>" + title + "</H1> ");

    out.print("<TABLE BGCOLOR=\"#EEEEEE\"?> ");
    out.print("<TR><TD align=\"center\">Parameter Name</TD> ");
    out.print("      <TD align=\"center\">Parameter Value(s)</TD></TR> ");
    Enumeration paramNames = request.getParameterNames();
    while(paramNames.hasMoreElements()) {
        String paramName = (String) paramNames.nextElement();
        out.print("<TR><TD valign=\"top\">" + paramName + "</TD>");
```



```

String[] paramValues = request.getParameterValues(paramName);
if (paramValues.length == 1) {
    String paramValue = paramValues[0];
    if (paramValue.length() == 0)
        out.print("<TD>No Value</TD>");
    else
        out.print("<TD>" + paramValue + "</TD>");
} else {
    out.println("<TD>");
    for(int i=0; i<paramValues.length; i++)
        out.println(paramValues[i] + "<BR>");
    out.println("</TD>");
}
out.print("</TR> ");
}
out.println("</TABLE> " +
"</BODY></HTML>");
}

public void doPost(HttpServletRequest request, HttpServletResponse response)
throws ServletException, IOException {
    doGet(request, response);
}
}

```

2.5 การใช้งานคุกกี้ด้วยเซิร์ฟเล็ต

ในการใช้งานเว็บปกติเบราว์เซอร์จะได้รับเนื้อข้อมูลซึ่งอาจจะอยู่ในรูปของข้อความ รูป, เสียง หรือภาพเคลื่อนไหวและข้อมูลอื่นๆ โดยที่เบราว์เซอร์สามารถจัดเก็บไว้ที่พื้นที่ที่เรียกว่าแคช เพื่อนำมาใช้ งานในภายหลังได้และข้อมูลดังกล่าวที่ใช้ในการจัดเก็บถือได้ว่าเป็นเนื้อข้อมูลที่เว็บเซิร์ฟเวอร์ต้องการที่จะ นำเสนอต่อเบราว์เซอร์เท่านั้น แต่หากเว็บเซิร์ฟเวอร์ต้องการส่งข้อมูลนอกเหนือจากนั้นมาจัดเก็บไว้ใน พื้นที่ใดๆของเบราว์เซอร์ก่อน เพื่อในภายหลังหากเบราว์เซอร์เรียกใช้งานเว็บเซิร์ฟเวอร์ตัวเดิม เว็บ เซิร์ฟเวอร์ก็จะหยิบเอาข้อมูลนี้มาใช้งานในภายหลัง ซึ่งข้อมูลนี้เรียกกันโดยทั่วไปว่าคุกกี้ การใช้ประโยชน์ จากคุกกี้ก็สามารถนำมาประยุกต์การใช้งานได้ เช่น หากเว็บเซิร์ฟเวอร์ต้องการชื่อและนามสกุลผู้ใช้งาน ครั้งแรก และเบราว์เซอร์ได้บอกข้อมูลเหล่านี้ผ่านทางฟอร์มไปแล้ว เว็บเซิร์ฟเวอร์ก็จะส่งชื่อและนามสกุล ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผู้ใช้งานเก็บไว้ที่เบราว์เซอร์ และหลังจากนั้นหากเบราว์เซอร์มีการเรียกใช้เว็บเซิร์ฟเวอร์ตัวเดียวกันนี้อีก เว็บเซิร์ฟเวอร์ก็จะอ่านชื่อและนามสกุลที่ฝากไว้ที่เบราว์เซอร์มาใช้ได้เลย โดยที่ผู้ใช้งานเบราว์เซอร์ไม่จำเป็นต้องกรอกข้อมูลชื่อและนามสกุลเพื่อบอกให้เว็บเซิร์ฟเวอร์ทราบอีก ซึ่งเป็นการอำนวยความสะดวกในการใช้งานเกิดขึ้นนั่นเอง

เซิร์ฟเล็ตสามารถควบคุมการใช้งานคุกกี้นั้นผ่านทางข้อมูลส่วนหัวของโปรโตคอลเอชทีทีพี โดยหากต้องการจัดเก็บข้อมูลคุกกี้นั้นไว้ที่เบราว์เซอร์ก็กำหนดผ่านข้อมูลส่วนหัวได้ตอบ (ส่วน Set-Cookie) ไปยังเบราว์เซอร์ก่อน หลังจากนั้นหากเว็บเซิร์ฟเวอร์มีการอ้างจากเบราว์เซอร์ในเส้นทางเดิมอีกเบราว์เซอร์ก็จะส่งข้อมูลคุกกี้นั้นมายังข้อมูลส่วนหัวร้องขอ (ส่วนคุกกี้นั้น) มาให้เว็บเซิร์ฟเวอร์ด้วย ดังนั้นเซิร์ฟเล็ตก็สามารถอ่านข้อมูลดังกล่าวมาใช้งานได้ แต่สำหรับการใช้งานคุกกี้นั้นแล้วคลาส HttpServletRequest และ HttpServletResponse มีเมธอดเฉพาะสำหรับการดำเนินงานในส่วนคุกกี้นี้ให้ใช้งานได้อย่างสะดวก โดยไม่จำเป็นต้องจัดการข้อมูลส่วนหัวให้ยุ่งยากเลย ดังมีรายละเอียดคือ

รายละเอียดที่เกี่ยวข้องกับการใช้งานคุกกี้นั้น	
Cookie()	คอนสตรักเตอร์ของคลาส Cookie
Cookie(String name, String value)	คอนสตรักเตอร์ของคลาส Cookie สำหรับกำหนดชื่อ name และข้อมูลใน value โดยการตั้งชื่อคุกกี้นั้นให้ใช้ภาษาอังกฤษสามารถมีช่องว่าง และใช้สัญลักษณ์ [] () = . " / ? @ : ; ได้
เมธอดที่เกี่ยวข้องกับการกำหนดค่าและการอ่านค่าของคลาส Cookie	
public void setComment(String comment)	ใช้กำหนดข้อมูลเสริมของคุกกี้นั้นเหมือนการใส่หมายเหตุเพื่ออธิบายรายละเอียดของการใช้งานคุกกี้นั้น โดยแต่ละคุกกี้นั้นสามารถมีข้อมูลหมายเหตุแยกแต่ละตัวได้
public String getComment()	สำหรับอ่านข้อมูลหมายเหตุของคุกกี้นั้น
public void setDomain(String domain)	สำหรับกำหนดชื่อโดเมน เพื่อให้ใช้งานคุกกี้นั้นได้ โดยรูปแบบโดเมนต้องมีสัญลักษณ์ . อยู่มากกว่า 2 ตำแหน่ง เช่น .javacentrix.com หรือ .kmitl.ac.th เป็นต้น
public String getDomain()	สำหรับอ่านค่าโดเมนเพื่อใช้งาน
public void setMaxAge(int lifetime)	สำหรับกำหนดอายุคุกกี้นั้นที่จะเก็บไว้ที่เบราว์เซอร์ หากเกินกำหนดเบราว์เซอร์จะลบออกจากเครื่อง มีหน่วยเป็นวินาที หากกำหนดเป็น 0 หรือค่าติดลบหมายถึงให้ลบทันทีที่เลิกใช้งานเบราว์เซอร์
public int getMaxAge()	สำหรับอ่านข้อมูลอายุของคุกกี้นั้น

public void setName(String name)	สำหรับกำหนดชื่อให้กับคุกกี้
public String getName()	สำหรับอ่านชื่อของคุกกี้
public void setPath(String path)	สำหรับกำหนดเส้นทางที่จะใช้งานคุกกี้ได้ โดยระบุเป็นข้อมูลยูอาร์แอล โดยที่ไม่ต้องระบุไฟล์เอกสาร เช่น http://online.javacentrix.com/order/ แต่ไม่ใช่ http://online.javacentrix.com/order/new.html
public String getPath()	สำหรับอ่านข้อมูลเส้นทางที่คุกกี้จัดเก็บไว้สำหรับใช้งาน
public void setSecure(boolean flag)	สำหรับค่าสถานการณ์รักษาความปลอดภัยในการเข้ารหัสข้อมูลแบบ SSL คือเริ่มต้นคือ false
public boolean getSecure()	สำหรับค่าสถานการณ์รักษาความปลอดภัย
public void setValue(String value)	สำหรับกำหนดข้อมูลให้กับคุกกี้
public String getValue()	สำหรับอ่านชื่อคุกกี้
public void setVersion(int version)	สำหรับกำหนดเวอร์ชันให้กับคุกกี้
public int getVersion()	สำหรับอ่านเวอร์ชันของคุกกี้

ตารางที่ 2-1 แสดงรายละเอียดที่เกี่ยวข้องกับการใช้งานคุกกี้

การส่งคุกกี้ไปเก็บที่เบราว์เซอร์ ในการส่งข้อมูลไปเก็บไว้เป็นคุกกี้ในเบราว์เซอร์จำเป็นต้องเริ่มต้นจากเซิร์ฟเล็ตส่งไปให้ในครั้งแรกก่อน ซึ่งการจะเก็บได้หรือไม่ได้ขึ้นอยู่กับความยินยอมที่กำหนดจากโปรแกรมเบราว์เซอร์เอง ซึ่งการจัดเก็บคุกกี้นี้หากเบราว์เซอร์ยินยอมแล้วผู้ใช้จะไม่สามารถสังเกตเห็นเลย ว่ามีการจัดเก็บข้อมูลคุกกี้หรือไม่จากการใช้งานปกติ (ยกเว้นผู้ใช้เข้าไปดูคุกกี้ภายในเครื่องใช้งานเอง) โดยต้องอาศัยเมธอด addCookie ของคลาส HttpServletResponse จัดเก็บดังนี้

การอ่านคุกกี้จากเบราว์เซอร์ สำหรับการอ่านข้อมูลคุกกี้ที่เซิร์ฟเล็ตที่ทำหน้าที่อ่านข้อมูลจากคุกกี้ได้นั้น สามารถเรียกใช้เมธอด getCookie จากคลาส HttpServletRequest ได้เลย หากเคยส่งคุกกี้ไปเก็บไว้ที่เบราว์เซอร์แล้ว แต่หากไม่เคยเมธอดดังกล่าวจะคือค่า null กลับมา และเมื่อเซิร์ฟเล็ตได้รับคุกกี้กลับมา ก็สามารถเรียกใช้เมธอดคุกกี้ในกลุ่ม getXxx ที่มีรายละเอียดอยู่ก่อนหน้านี้อ่านข้อมูลมาใช้งานได้ต่อไป

2.6 เจดีบีซี (JDBC) ช่องทางติดต่อฐานข้อมูลด้วยจาวา

จาวาด้านเบส คอนเนกทิวิตี (Java Database Connectivity) หรือ เจดีบีซี (JDBC) คือมิดเดิลแวร์ที่ถูกใช้บนเทคโนโลยีจาวา ในการเข้าถึงระบบฐานข้อมูล โดยสามารถเรียกใช้งานข้อมูล หรือจัดการระบบฐานข้อมูลผ่าน คำสั่ง SQL-92 หรือภาษาเอสคิวแอลมาตรฐาน

การใช้งานเจดีบีซี

อย่างที่ทราบกันว่าการเข้าใช้งานฐานข้อมูลด้วยภาษาจาวา มีรูปแบบอยู่ด้วยกันถึง 4 ชนิด การที่โปรแกรมเมอร์จะเลือกใช้วิธีใดก็ขึ้นอยู่กับความเหมาะสมของระบบว่าระบบถูกออกแบบมาให้มีลักษณะการทำงานแบบใด แต่ถึงแม้ว่าจะได้ข้อสรุปอย่างไรก็ตาม วิธีการลงมือเขียนโปรแกรมก็มีขั้นตอนที่คล้ายกัน ดังนี้

- ขั้นตอนเริ่มติดต่อกับฐานข้อมูล (Establish Database)
- ขั้นตอนการส่งคำสั่งเอสคิวแอลให้กับฐานข้อมูล (Send Query)
- ขั้นตอนการแสดงผลลัพธ์ผู้ใช้ (Display Result)
- ขั้นตอนการปิดการติดต่อกับฐานข้อมูล (Close Connection)

ขั้นตอนที่ 1 เริ่มติดต่อกับฐานข้อมูล

แบ่งออกเป็น 2 ขั้นตอนย่อยๆดังนี้ คือ

A) ขั้นตอนการโหลดไดรเวอร์เพื่อใช้งาน

รูปแบบ

```
Class.forName ("Driver_Name")
```

คลาสคักจับข้อผิดพลาด

```
ClassNotFoundException
```

เมื่อ

Driver_Name คือชื่อไดรเวอร์สำหรับใช้งานระบบฐานข้อมูล ที่ต้องการโหลดขึ้นมาใช้งาน ซึ่งจะต้องระบุชื่อให้ตรงกับระบบฐานข้อมูลที่ต้องการติดต่อ

```
try {
    Class.forName ("sun.jdbc.odbc.JdbcOdbcDriver");
} catch (ClassNotFoundException e) {
    System.err.print (e.getMessage ());
}
```

B) ขั้นตอนการติดต่อไปยังระบบฐานข้อมูลปลายทาง

รูปแบบ

```
Connection Instance_Name = DriverManager.getConnection (URL, Login, Password)
```

คลาสคักจับข้อผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SQLException

รูปแบบ URL

jdbc:<subprotocol>:<subname>

เมื่อ

<subprotocol> คือ โพรโทคอลการสื่อสารข้อมูล เพื่อแจ้งไปยังเครื่องหรือระบบปลายทาง
ซึ่งมีข้อมูลแตกต่างกันไปตามรูปแบบการสื่อสาร เช่น

odbc สำหรับการสื่อสารผ่านระบบ ODBC

dbnet สำหรับการสื่อสารผ่านโปรโตคอล TCP/IP

<subname> คือฐานข้อมูลภายในระบบที่อ้างผ่าน subprotocol หรือแหล่งข้อมูล

ขั้นตอนที่ 2 การส่งคำสั่งเอสคิวแอล (SQL)

เมธอดที่ถูกกำหนดไว้สำหรับส่งคำสั่งเอสคิวแอลไปให้กับฐานข้อมูล ถูกแบ่งตามรูปแบบคำสั่งเอสคิวแอล ได้ 2 แบบ โดยผู้ใช้ต้องเลือกเมธอดที่เหมาะสมกับคำสั่งเอสคิวแอลที่ใช้งานคือ

หากคำสั่งเอสคิวแอลเป็นประเภทดำเนินการ ซึ่งสร้างด้วยคีย์เวิร์ดจำพวก ตรีเอท (CREATE), อัลเทอ (ALTER), ดรอป (DROP), อินเซท (INSERT), อัปเดต (UPDATE), ดีลิต (DELETE), ? การทำงานของคำสั่งลักษณะนี้ เป็นการส่งคำสั่งไปยังฐานข้อมูล เพื่อให้เกิดการเปลี่ยนแปลงใดๆในระบบฐานข้อมูล ดังนั้นผู้ใช้คำสั่งเพียงต้องการรับทราบแค่ว่าทำงานสำเร็จหรือไม่เท่านั้น การเรียกใช้คำสั่งเหล่านี้ให้ทำงานจึงต้องใช้งานผ่านเมธอด เอกสคิวอัปเดต (ExecuteUpdate)

หากคำสั่งเอสคิวแอล เป็นประเภทสืบค้นข้อมูล ซึ่งสร้างด้วยคีย์เวิร์ด ซีเลกท (SELECT) เนื่องจากคำสั่งประเภทสืบค้นข้อมูลมีลักษณะพิเศษที่เมื่อระบบฐานข้อมูลได้รับคำสั่งเอสคิวแอล แล้วจะต้องส่งข้อมูล ผลลัพธ์ที่ต้องการกลับมาในรูปแบบของกลุ่มข้อมูลเรคคอร์ด (Record Set) การเรียกใช้งานคำสั่งเอสคิวแอล ประเภทนี้จึงต้องเรียกใช้งานผ่านเมธอดเอกสคิวคิวรี่ (ExecuteQuery)และต้องมีคลาสมารับผลลัพธ์ที่ได้รับกลับมาในรูปแบบของกลุ่มเรคคอร์ดคือคลาส รีซอลทเซท (ResultSet) เมธอดทั้งสองรูปแบบดังกล่าว สามารถเรียกใช้งานได้จากคลาส Statement และคลาส PreparedStatement อย่างใดอย่างหนึ่ง แต่ทั้งสองคลาสดังกล่าวใช้สำหรับการสร้างคำสั่งเอสคิวแอล ที่แตกต่างกัน คือ

Statement ใช้สำหรับสร้างคำสั่งเอสคิวแอล ที่เป็นข้อมูลคงที่ (ข้อมูลในคำสั่งเอสคิวแอลไม่เปลี่ยนแปลง)

PreparedStatement ใช้สำหรับสร้างคำสั่งเอสคิวแอล ที่สามารถเปลี่ยนแปลงข้อมูลบางส่วนภายในตามพารามิเตอร์ที่ต้องการ

รูปแบบคลาส Statement

int ExecuteUpdate()

สำหรับส่งงานคำสั่ง SQL ประเภทจัดการข้อมูล

ResultSet executeQuery()

สำหรับส่งงานคำสั่ง SQL ประเภทสืบค้นข้อมูล

รูปแบบคลาส PreparedStatement

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่หรือใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามนำเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<code>void setBoolean(int index, boolean x)</code>	กำหนดข้อมูล boolean ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setByte(int index, byte x)</code>	กำหนดข้อมูล byte5 ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setShort(int index, short x)</code>	กำหนดข้อมูล short ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setInt(int index, int x)</code>	กำหนดข้อมูล int ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setLong(int index, long x)</code>	กำหนดข้อมูล long ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setFloat(int index, float x)</code>	กำหนดข้อมูล float ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setDouble(int index, double x)</code>	กำหนดข้อมูล double ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setString(int index, String x)</code>	กำหนดข้อมูล String ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setDate(int index, Date x)</code>	กำหนดข้อมูล Date ให้กับพารามิเตอร์ตำแหน่ง index
<code>void setTime(int index, Time x)</code>	กำหนดข้อมูล Time ให้กับพารามิเตอร์ตำแหน่ง index
<code>void clearParameter()</code>	ลบข้อมูลออกจากพารามิเตอร์ทุกตัว
<code>int executeUpdate()</code>	สำหรับสั่งงานคำสั่ง SQL ประเภทจัดการข้อมูล
<code>ResultSet executeQuery()</code>	สำหรับสั่งงานคำสั่ง SQL ประเภทสืบค้นข้อมูล
<code>void close()</code>	ปิดช่องทางการสื่อสาร

การขยับตัวชี้เพื่ออ้างอิงไปยังแถวที่ต้องการสามารถขยับได้โดยใช้เมธอดของ **ResultSet** ดังนี้

ResultSet ดังนี้

- หากต้องการย้ายไปที่ข้อมูลแถวแรกใช้เมธอด **first()**
- หากต้องการย้ายไปที่ข้อมูลแถวสุดท้ายใช้เมธอด **last()**
- หากต้องการย้ายไปที่ข้อมูลแถวถัดไปจากตำแหน่งปัจจุบันใช้เมธอด **next()**
- หากต้องการย้ายไปที่ข้อมูลแถวก่อนหน้าจากตำแหน่งปัจจุบันใช้เมธอด **previous()**
- หากต้องการย้ายไปที่ข้อมูลตามแถวตัวเลขที่ต้องการใช้เมธอด **absolute(แถวที่ต้องการ)**

ในกลุ่มเมธอดที่ใช้ในการตรวจสอบตำแหน่งแถวว่าอยู่แถวใด มีเมธอดในกลุ่มที่เกี่ยวข้องดังนี้

- หากต้องการทราบว่าขณะนี้ตัวชี้อยู่ที่ตำแหน่งแถวใดให้ใช้เมธอด **getRow()**
- หากต้องการทราบว่าขณะนี้ตัวชี้อยู่ตำแหน่ง Before First หรือไม่ให้ใช้เมธอด **isBeforeFirst()**
- หากต้องการทราบว่าขณะนี้ตัวชี้อยู่ตำแหน่ง First หรือไม่ให้ใช้เมธอด **isFirst()**
- หากต้องการทราบว่าขณะนี้ตัวชี้อยู่ตำแหน่ง After Last หรือไม่ให้ใช้เมธอด **isAfterLast()**
- หากต้องการทราบว่าขณะนี้ตัวชี้อยู่ตำแหน่ง Last หรือไม่ให้ใช้เมธอด **isLast()**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 3 การแสดงผลลัพธ์สู่ผู้ใช้

ผลลัพธ์ในชุดรีซอลท์เซต (ResultSet) มีโครงสร้างคล้ายกับข้อมูลในเทเบิล คือแบบออกเป็นแนวแถวเรียกว่าโรล Row หรือ เรคอร์ด และแนวคอลัมน์เรียกว่า Column หรือ ฟิลด์ ดังนั้นหลักการดึงข้อมูลคือต้องระบุโรลให้ชัดเจนก่อนแล้วจึงค่อยระบุคอลัมน์ที่ต้องการ

รูปแบบคลาส ResultSet

boolean getBoolean(String ColumnName)	อ่านข้อมูล boolean ในฟิลด์ ColumnName
byte getByte(String ColumnName)	อ่านข้อมูล byte ในฟิลด์ ColumnName
short getShort(String ColumnName)	อ่านข้อมูล short ในฟิลด์ ColumnName
int getInt(String ColumnName)	อ่านข้อมูล int ในฟิลด์ ColumnName
long getLong(String ColumnName)	อ่านข้อมูล long ในฟิลด์ ColumnName
float getFloat(String ColumnName)	อ่านข้อมูล float ในฟิลด์ ColumnName
double getDouble(String ColumnName)	อ่านข้อมูล double ในฟิลด์ ColumnName
String getString(String ColumnName)	อ่านข้อมูล String ในฟิลด์ ColumnName
Date getDate(String ColumnName)	อ่านข้อมูล Date ในฟิลด์ ColumnName
Time getTime(String ColumnName)	อ่านข้อมูล Time ในฟิลด์ ColumnName

ขั้นตอนที่ 4 การยกเลิกการติดต่อ

สำหรับในขั้นตอนยกเลิกการติดต่อนี้เป็นขั้นตอนที่ง่ายและไม่ยุ่งยากในการทำงาน หากเมื่อพิจารณาจากการดำเนินการติดต่อข้อมูล มีอยู่ 2 ส่วนด้วยกัน ส่วนแรกคือช่องทางการติดต่อที่สร้างมาจากเมธอด getConnection และส่วนที่สองคือชุดคำสั่ง ResultSet ที่ได้จากเมธอด executeQuery

```
Connection con = DriverManager.getConnection ( url, "myLogin", "myPassword");
Statement stmt = con.createStatement();
ResultSet rs = stmt.executeQuery("SELECT * FROM Products")
?
?
?
rs.close();
con.close();
```


บทที่ 3

เอ็กซ์เอ็มแอลและฐานข้อมูลที่สามารถทำงานกับเอ็กซ์เอ็มแอลได้

3.1 ภาษาเอ็กซ์เอ็มแอล

เนื่องจากทุกวันนี้โลกของเรามีการนำเทคโนโลยีสารสนเทศเข้ามาใช้ในชีวิตประจำวันมากขึ้น และได้เข้ามาเป็นส่วนหนึ่งของการทำงาน ที่เห็นได้ชัดคือ เทคโนโลยีบนอินเทอร์เน็ต ไม่ว่าจะเป็นการเข้าไปเยี่ยมชมเว็บไซต์ต่าง ๆ หรือการส่งจดหมายอิเล็กทรอนิกส์ที่เรียกว่า อีเมล ซึ่งกำลังเป็นมาตรฐานของการติดต่อสื่อสารสำหรับอนาคตทำให้ต้องมีการคิดเพื่อพัฒนาให้มีความก้าวหน้ามากขึ้น การเขียนเว็บไซต์ในปัจจุบันนี้ โดยปกติแล้วสิ่งที่เราจะสร้างเว็บได้คือ ภาษาเอชทีทีที ที่เรารู้จักกันดีเท่านั้น แต่ว่าวันนี้โลกได้มีการพัฒนาเทคโนโลยีใหม่ที่ใช้สำหรับการเขียนเว็บ นั่นคือเอ็กซ์เอ็มแอลซึ่งเป็นสิ่งที่หลายผลิตภัณฑ์ให้การสนับสนุน

เอ็กซ์เอ็มแอลเป็นภาษาที่ให้ความชัดเจนในการให้รายละเอียดเกี่ยวกับข้อมูลและการเปลี่ยนแปลงข้อมูลโดยแอปพลิเคชันบนเว็บและใช้ฟอร์มที่ยืดหยุ่นได้ตามมาตรฐานเอชทีเอ็มแอลได้เปิดโลกแห่งการแสดงผลต่าง ๆ มากมาย ส่วนเอ็กซ์เอ็มแอลจะทำให้การทำงานกับข้อมูลโดยตรงที่เสริมกับการทำงานของ เอชทีเอ็มแอล

เอ็กซ์เอ็มแอลเป็นฟอร์มที่อธิบายถึงรายละเอียดของโครงสร้างและแบบของข้อมูลเป็นภาษาหรือชุดคำสั่งเกี่ยวกับข้อมูลบนเว็บ ที่ให้การพัฒนาและมีศักยภาพในส่วนของการสร้างข้อมูลจากหลากหลายแอปพลิเคชันมานำเสนอบนเครื่องเดสก์ทอป ด้วยเอ็กซ์เอ็มแอลจะทำให้การจัดการข้อมูลหรือเรียกใช้ข้อมูลจากแอปพลิเคชันต่างๆ จะเข้าสู่มาตรฐานเดียวกัน

เอ็กซ์เอ็มแอลจะเป็นส่วนหนึ่งของเอชทีเอ็มแอล ซึ่งเอ็กซ์เอ็มแอลจะให้รายละเอียดเกี่ยวกับข้อมูล เช่น ชื่อเมือง อุณหภูมิ ความกดอากาศ ส่วนเอชทีเอ็มแอลเป็นการกำหนดแท็กต่างๆ ที่จะทำให้ข้อมูลแสดงออกมาในรูปแบบไหน ซึ่งข้อมูลสามารถแสดงออกมาได้หลายรูปแบบไม่ว่าจะเป็นตารางหรือข้อความธรรมดา ขึ้นอยู่กับการกำหนดของเอชทีเอ็มแอลและในปัจจุบันนี้ ด้วยเอ็กซ์เอ็มแอลจะมีการให้รายละเอียดของเนื้อหาเอกสารที่เรียกว่าดีทีดี (DTD หรือ Document Type Definition) ที่ให้รายละเอียดเกี่ยวกับตัวเอกสารว่าจะแสดงหรือซ่อนส่วนไหนของเอกสารบ้าง ซึ่งดีทีดีจะเป็นส่วนที่เพิ่มเติมสำหรับ เอ็กซ์เอ็มแอล ถ้าหากมีการส่งข้อมูลในรูปแบบดีทีดีก็รู้กันว่าเป็นเอ็กซ์เอ็มแอล

3.2 ลักษณะโครงสร้างของเอกสารเอ็กซ์เอ็มแอล

เอ็กซ์เอ็มแอลเป็นการใช้ข้อความเพื่อบ่งบอกโครงสร้างของเอกสาร พิจารณาตัวอย่างรูปแบบโครงสร้างของหนังสือ เมื่อหนังสือประกอบด้วยจำนวนบท 2 บทในแต่ละบทประกอบด้วยเนื้อความ

Begin Book

Begin Chapter 1

Text for Chapter 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

End Chapter 1
Begin Chapter 2
Text for Chapter 2
End Chapter 2
End Book

หนังสือที่มีอยู่ในปัจจุบัน จะมีโครงสร้างที่มีรายละเอียดที่ซับซ้อนมากกว่านี้ เช่น บทนำ, สารบัญ, เป็นต้น เช่นเดียวกัน ภายในส่วนเนื้อหา (Text) ยังประกอบด้วย โครงสร้างย่อย คือ ย่อหน้า (Paragraph) แต่ละย่อหน้ายังประกอบขึ้นจากประโยคและตัวอักษรด้วย ลักษณะของเอกสารอิเล็กทรอนิกส์นั้น สามารถอธิบายโดยใช้ตัวอย่างที่ 1 ได้ ดังนี้

ตัวอย่างที่ 1

```
<?xmlversion="1.0"encoding="windows-874"?>
<mali>
<malisorn>ขึ้นต้นด้วยมะลิซ้อน</malisorn>
<malila>พอแตกใบอ่อนเป็นมะลิลา</malila>
</mali>
<?xmlstylesheet type="text/xsl" href=""?>
```

บรรทัดที่ 1 นั้นหมายความว่าเราประกาศเอกสารนี้เป็นเอกสาร เอกซ์เอ็มแอล และมีการเข้ารหัสอักขระเป็น windows-874 เพื่อให้ใช้ภาษาไทยได้ จากนั้นเราจะเห็นสิ่งที่เราคุ้นเคยที่เราเรียกว่า แท็ก แท้จริงแล้วในภาษาเอกซ์เอ็มแอลจะแบ่ง โครงสร้างเป็น 2 ส่วนใหญ่ๆ คือแท็ก & อิลิเมนต์ สามารถอธิบายเพิ่มเติมได้จากตัวอย่างที่ 2

ตัวอย่างที่ 2

```
<root>
  <element>
    <tag></tag>
  </element>
</root>
```

ความหมายของแท็ก สำหรับในเอกซ์เอ็มแอล แล้วแท็ก มีความหมายในลักษณะเดียวกับที่ใช้ในเอกสารเป็นเอกสารที่สิ่งหนึ่งหรือสองสิ่งนั้นคือหนึ่งคือหนึ่งและหนึ่งนั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

-แท็ก เปิด (Start tag)

<book>

จากตัวอย่างที่แสดง ด้านบนถูกเรียกว่าแท็ก เปิด ดังนั้นแท็ก เปิดจึงมีสัญลักษณ์คือ <...>

-แท็ก ปิด (End Tag)

</book>

แท็กที่ถูกเรียกว่าแท็ก ปิด ต่อเมื่อในแท็ก มีเครื่องหมาย / อยู่หลังสัญลักษณ์ < ดังนั้นลักษณะของแท็ก ปิดจึง มีรูปแบบคือ </...> หากพิจารณาระหว่างแท็ก เปิดกับแท็ก ปิดแล้ว ข้อแตกต่างอีกข้อหนึ่งคือแท็ก เปิด เป็นแท็ก ที่สามารถใส่ข้อมูลแอททริบิวต์ ลงไปภายในแท็ก ได้ แต่แท็ก ปิดจะไม่ทำกัน

อิลิเมนต์ ในที่นี้คือ โครงสร้างหลักของ เอกซ์เอ็มแอล ซึ่งอยู่ในรูปของแท็ก เช่นเดียวกัน ตามตัวอย่างข้างบน อิลิเมนต์ ก็จะมีลักษณะซ้อนกันเป็นชั้นๆ

```
<root>
  <element>
    </element>
  </root>
```

- Root element เป็น อิลิเมนต์ แรกสุดของเอกสาร เอกซ์เอ็มแอล

อิลิเมนต์เป็นส่วนประกอบของ คอนเท้นท์และแอททริบิวต์เพื่อให้ผู้อ่าน ได้เข้าใจความหมายของศัพท์หลายๆ คำศัพท์ ก่อนที่จะนำไปใช้ในการสร้างเอกซ์เอ็มแอล ซึ่งจำเป็นอย่างมาก ลองมาพิจารณาคำจำกัดความของ อิลิเมนต์ กันก่อน

<chap number="1">Text for Chapter 1</chap>

สังเกตจากลักษณะ รูปทางด้านบน ทั้งหมดตั้งแต่ <chap จนถึง </chap> ถูกเรียกว่า อิลิเมนต์ หรือถ้ามองง่ายก็คือ อิลิเมนต์ เริ่มต้นที่แท็ก เปิด และสิ้นสุดที่แท็ก ปิดในแท็กคำสั่งเดียวกัน

คอนเท้นท์หรือเนื้อหา ความ คอนเท้นท์ถือได้ว่าเป็นข้อมูลเพื่อใช้ในการแสดง ให้ผู้อ่านเอกสารได้ เห็นหรือกล่าวอีกนัยหนึ่งคือคอนเท้นท์อยู่หลังแท็กเปิดและจบที่ก่อนถึงแท็กปิดนั่นเอง

แอททริบิวต์จะถูกบรรจุอยู่ภายในแท็ก เปิด และแอททริบิวต์ นี้ในบางครั้ง อาจจะมีหรือไม่มีก็ได้ รวมถึงถ้ามีอาจมีได้มากกว่าหนึ่งตัวในแท็ก เปิดแอททริบิวต์คือข้อมูลความหมายเพิ่มเติม แน่นอนถ้ากล่าวถึงสิ่งของ หากเราไม่ได้ให้ความหมายเพิ่มเติมเราก็จะไม่รู้ และแยกแยะสิ่งที่เหมือนกันออกจากกันได้ เช่น ถ้าพูดถึง บท ในหนังสือ แค่นี้เป็นความหมายโดยรวม แต่ถ้าบอกว่า บทที่ 1 ในหนังสือ เลข 1 ในที่นี้คือความหมายเพิ่มเติมให้กับบท หรืออีกซีกตัวอย่าง ถ้ากล่าวถึง แอปเปิ้ล อย่างเดียว คนทั่วไปจะนึกถึงผลไม้ที่ทานได้ แต่ถ้าบอกแอปเปิ้ล สีแดง รสเปรี้ยว นั่นคือการบอกถึง แอททริบิวต์ 2 ตัวคือ สีแดงและรสเปรี้ยว ทำให้ทราบว่าแอปเปิ้ลที่ไม่มีใครอยากได้มาทาน (ยกเว้นคนที่ชอบของเปรี้ยวๆ)

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ดีทีดี ข้อกำหนดโครงสร้างสำหรับ เอกซ์เอ็มแอล

ส่วนนี้ จะมาทำความรู้จักกับเอกสารในรูปแบบการจัดสร้างที่ถูกต้อง และส่วนควบคุม โครงสร้างเอกสารที่เรียกว่า ดีทีดี

3.3.1 ดีทีดี คืออะไร

ดีทีดี คือแฟ้มข้อมูลหรือหลายแฟ้มข้อมูลที่ใช้งานร่วมกัน ซึ่งบรรจุข้อกำหนด และกฎเกณฑ์ของเอกสาร ชุดข้อกำหนดเหล่านี้ สำหรับการกำหนดรูปแบบ อิลิเมนต์ ตัวอย่างเช่น หากต้องการเอกสารที่มีอิลิเมนต์ <LIST> ที่มี อิลิเมนต์ <ITEM> บรรจุอยู่ภายใน ข้อกำหนดในแฟ้มข้อมูล ดีทีดี จะมีรูปแบบดังนี้

<!ELEMENT item (#pcdata)>

<!ELEMENT list (item)+>

ซึ่งอธิบายความหมาย คืออิลิเมนต์ items บรรจุข้อความใดๆ และอิลิเมนต์ลิสต์ บรรจุอิลิเมนต์ใดเพิ่มอีกที ดังนั้นดีทีดีเป็นรูปแบบภาษา ซึ่งทำให้สามารถตรวจสอบเอกสารที่นำเอาข้อกำหนดดีทีดีไปใช้ว่าถูกจัดสร้างตามความต้องการหรือไม่ ทำให้ระบบการแสดงผล สามารถเข้าใจตัวเอกสารได้ดี และดึงไปใช้งานได้อย่างถูกต้อง

ดีทีดี มีความซับซ้อน จากคำอธิบายในส่วนต้นดีทีดีมีความซับซ้อน หมายถึงการสร้างดีทีดีมีรายละเอียดและขั้นตอนที่ยุ่งยากพอสมควร ถึงแม้ ดีทีดี จะทำให้การสร้างเอกสารมีความยุ่งยากก็ตาม แต่มีเหตุผลที่ควรรู้ก่อน ก็คือเอกซ์เอ็มแอลไม่จำเป็นต้องใช้ดีทีดีเสมอไป ถ้าแม้จะต้องใช้ดีทีดีแต่ก็จะมีโครงสร้างดีทีดีที่เป็นมาตรฐานให้ใช้อยู่แล้ว ดีทีดีที่เป็นมาตรฐาน ถูกพัฒนาโดยผู้ที่สนใจในเอกสารเฉพาะด้าน และมีให้พวกเรา ได้ดึงมาใช้งาน เพื่อให้เกิดรูปแบบมาตรฐานที่ชัดเจน และเข้าใจข้อมูลกัน ในเอกสารประเภทเดียวกัน 3 ส่วนสำคัญ เอกสารเอกซ์เอ็มแอลมีส่วนประกอบที่สำคัญอยู่สามส่วน สองส่วนเป็นสิ่งที่จำเป็นในขณะที่อีกส่วนเป็นทางเลือกที่จะมีหรือไม่ก็ได้

1. ส่วนแรก คือ ส่วนของเนื้อหาหรือคอนเท้นต์นั่นเอง ซึ่งทำให้เอกสารมีข้อมูลสำหรับดูโดยผู้อ่าน คอนเท้นต์ นี้เป็นได้ทั้งข้อความ รูปภาพ ส่วนนี้ถูกสร้างขึ้นมาจากอิลิเมนต์ ในที่นี้คือไฟล์เอกซ์เอ็มแอล

2. ส่วนที่สอง คือ กฎเกณฑ์และข้อกำหนดโครงสร้างของเอกสาร ในที่นี้คือไฟล์ดีทีดีส่วนนี้ถือเป็นทางเลือกซึ่งจะเลือกใช้หรือไม่ก็ได้

3. ส่วนสุดท้าย คือ สไตลชีท คือ ลักษณะข้อกำหนดสำหรับการแสดงผลฟร้นเอง ในที่นี้คือไฟล์เอกซ์เอสแอล

ดีทีดีคืออุปสรรคในขณะที่สไตลชีทไม่ใช่เอกสารเอกซ์เอ็มแอล สามารถถูกบังคับโครงสร้างด้วยดีทีดี ซึ่งจะใช้หรือไม่ใช้ก็ได้แต่สไตลชีท คือสิ่งที่จำเป็นที่จะสามารถนำข้อมูลในเอกสาร มาแสดงด้วยการเรนเดอร์ สไตลชีทในบางครั้งไม่ต้องการ หากเรามีกลไกแสดงผล (rendering) ที่จัดสร้างไว้พร้อมแล้วสำหรับเอกสารเอกซ์เอ็มแอลที่ต้องผ่านขบวนการแปลงด้วยเอกซ์เอสแอล เพื่อให้เป็น เอกซ์เอ็มแอลในการแสดงผ่านบราวเซอร์ อย่างนี้ต้องจำเป็นในการใช้งานสไตลชีท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 ประโยชน์ของเอกสารดีที่ดี

ในการจัดสร้างดีที่ดีทำให้เอกสารมีความเป็นรูปแบบที่ดีและถูกต้อง เรียกว่ากันว่าเอกสารที่ถูกต้อง ในบางครั้งหากเราต้องการสร้างเอกสารอิเล็กทรอนิกส์ที่สามารถแลกเปลี่ยนกับคนอื่น ๆ หรือส่วนอื่นๆ และให้สามารถเข้าใจโครงสร้างเดียวกัน นั่นก็คือเอกสารอิเล็กทรอนิกส์ของเราและในส่วนอื่นๆ จำเป็นต้องมีรูปแบบโครงสร้างเดียวกัน นั่นก็คือควรที่จะเลือกใช้ดีที่ดีเดียวกัน เพื่อบังคับโครงสร้างให้เหมือนกัน

3.4 เอกสาร อิเล็กทรอนิกส์ ที่ถูกต้อง

เอกสารที่อยู่ในรูปแบบที่ดี (well-formed) เป็นเอกสารที่เรียกว่าถูกต้องตามข้อกำหนดของอิเล็กทรอนิกส์ แต่เอกสารที่ไม่ถูกต้องตามรูปแบบอิเล็กทรอนิกส์ก็สามารถเป็นเอกสารที่ดีได้เช่นกัน เนื่องจากในปัจจุบันข้อมูล บนระบบอินเทอร์เน็ตมีอยู่มากมายที่ไม่ได้สร้างเป็นเอกสารที่ถูกต้องตามแบบอิเล็กทรอนิกส์ นั่นคือความต้องการส่วนหนึ่ง ที่ดีที่ดีจะสามารถบังคับเอกสารอิเล็กทรอนิกส์ที่สร้างมาจากดีที่ดี เดียวกัน ให้เป็นเอกสารที่ถูกต้องตามแบบอิเล็กทรอนิกส์ (ต้องเป็นดีที่ดีที่ถูกต้องด้วย)

เอกสารรูปแบบที่ถูกต้อง แนวคิดของเอกสารแบบที่ดี ถูกนำมาใช้สำหรับอิเล็กทรอนิกส์ ในกรณีที่ไม่สามารถสร้างเอกสารอิเล็กทรอนิกส์ด้วยข้อกำหนดดีที่ดี "จากตัวอย่างของแท็ก ใน เอกซ์เอ็มแอล เป็นแท็กที่เรียกว่า อิลิเมนต์เปล่า เนื่องจากว่าแท็ก ดังกล่าว ไม่จำเป็นต้องอาศัยแท็กปิด เช่นเดียวกัน หาก อิเล็กทรอนิกส์ มีอิลิเมนต์ที่มีแท็กปิดจึงไม่เป็นเอกสารในรูปแบบที่ดี

หากต้องการสร้างเอกสารอิเล็กทรอนิกส์แล้วจำไว้เสมอว่าจะต้องสร้างเอกสารที่เป็นรูปแบบที่ดี เอกสารที่ดีต้องเป็นอย่างไร

1. อิลิเมนต์ที่ถูกต้องคือ อิลิเมนต์ที่มีแท็กเปิด และต้องปิดด้วยด้วยแท็กปิดเสมอ เช่น <...>.....</...> ยกเว้นอิลิเมนต์ว่าง
2. ข้อมูลแอททริบิวต์ต้องอยู่ภายในเครื่องหมาย '...' หรือ '...' หาก แอททริบิวต์ไม่ได้อยู่ในสัญลักษณ์ดังกล่าวเอกสารจะไม่ถือว่าเป็นแบบที่ดี
3. อิลิเมนต์ว่างหากต้องการใช้งานอิลิเมนต์ว่าง ต้องสร้างอิลิเมนต์ว่างให้ถูกต้องเอกสารอิเล็กทรอนิกส์ ไม่จำเป็นต้องมีอิลิเมนต์ที่มีแท็กเปิดและแท็กปิดเสมอ เรายังสามารถสร้างอิลิเมนต์ว่างบรรจุลงในเอกสารได้ด้วยอิลิเมนต์ว่างจะต้องไม่บรรจุตัวอักษรใดๆลงในอิลิเมนต์ดังกล่าว สามารถเขียนได้ในสองรูปแบบอิลิเมนต์ที่มีแท็กเปิดและแท็กปิด โดยที่ไม่มีข้อความหรือคอนเท้นท์อยู่ระหว่างแท็ก (ยกเว้นตัวอักษรขึ้นบรรทัดใหม่)

<book></book>

อิลิเมนต์เฉพาะที่เรียกว่าอิลิเมนต์ว่างตามรูปแบบ <.../>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<book/>

อิลิเมนต์ว่างที่มีแอททริบิวต์แน่นอนอิลิเมนต์ว่าง สามารถบรรจุข้อมูลแอททริบิวต์ได้ด้วย ดังตัวอย่างต่อไปนี้

<book author="baldwin" price="\$9.95" />

4. ต้องไม่มีตัวอักษรประเภทมาร์กอัปอยู่ในส่วนคอนเท้นท์ คือข้อมูล < หรือ & เนื่องจากตัวอักษร < ถูกใช้สำหรับการสร้างแท็ก และ & ถูกใช้สำหรับการแสดงสัญลักษณ์พิเศษ หากต้องการแสดงข้อมูลตัวอักษรดังกล่าว ในส่วน คอนเท้นท์ จริงๆ ให้ใช้สัญลักษณ์ ใช้งาน < สำหรับแสดงตัวอักษร < และ & สำหรับแสดงตัวอักษร &

5. อิลิเมนต์ที่ต้องซ้อนกันต้องมีรูปแบบ หากมีการสร้างเอกสาร เอ็กซ์เอ็มแอล ที่มีการวางซ้อนกันของ อิลิเมนต์ ในหลายๆชั้น การวางซ้อนกันต้องมีรูปแบบ เป็นชั้นๆห้ามวางอิลิเมนต์ในลักษณะที่ซ้อนไขว้กัน

```
<book>
  <chapter number="1">
    <paragraph>
    </paragraph>
  </chapter>
</book>
```

ตัวอย่างด้านบนเป็นเอกสารที่อิลิเมนต์ว่างซ้อนกันอย่างถูกต้อง

```
<book>
  <chapter number="1">
    <paragraph>
    </chapter>
    < paragraph>
  </book>
```

ตัวอย่างด้านบนเป็นเอกสารที่วางซ้อนอิลิเมนต์อย่างไม่ถูกต้อง

กฎและคำศัพท์ที่เกี่ยวข้อง อันดับแรกการที่เราจะใช้งานเอ็กซ์เอ็มแอล นั้น เราต้องอาศัยส่วนของเอกสารที่อธิบายถึงสิ่งที่เรียกว่าดีทีดี (Document Type Definition) ในการกำหนดลักษณะและส่วนประกอบต่างๆของอิลิเมนต์ เรียกว่าไวยากรณ์ (grammar) ของภาษาเอ็กซ์เอ็มแอลนั้นแหละ ซึ่งดีไปใช้

ทีนี้เราสามารถเขียนขึ้นมาใช้งานได้ง่ายๆหรือถ้ามีความซับซ้อนมากๆของเอกสารเอ็กซ์เอ็มแอลเราอาจจะใช้คิตีที่มีผู้เขียนไว้แล้วสำหรับเอกสารเฉพาะอย่าง ใช้ก็ได้

3.5 เอกสารเอ็กซ์เอ็มแอลที่ดีและถูกต้อง

เอกสารที่ดีและถูกต้องนั้นศัพท์ 2 คำนี้จะใช้ตรวจสอบว่าเอกสารเอ็กซ์เอ็มแอล ที่เราเขียนขึ้นนั้น ถูกต้องหรือไม่ ซึ่งประกอบด้วย 2 ประการคือ

1. เอกสารเอ็กซ์เอ็มแอลนั้นแท็กต่างๆต้องมีรูปแบบที่ถูกต้องเช่น ขึ้นต้นแท็กด้วยตัวเล็กต้องจบแท็ก ด้วยตัวเล็กเป็นต้นและหากมีแท็กเปิดต้องมีแท็กปิดเสมอ เช่น

```
<boy></Boy> ---- ผิด , <BOY></boy> -- ผิด
<boy> ---- ผิด เพราะไม่มีแท็กปิด
<girl></girl> ---- ถูก เพราะมีแท็กปิด
```

แบบนี้เราจะเรียกว่าเอกสารที่ดี คือ โครงสร้างถูกต้องทั้งหมด

2. เอกสารเอ็กซ์เอ็มแอลนั้นต้องเป็นไปตามกฎของคิตีที่ดีของเอกสารนั้นๆ ด้วยจึงจะเรียกได้ว่า ถูกต้องเราลองเพิ่มโค้ดเพื่ออ้างอิงคิตีที่ดีกันจะได้เป็น

```
<?xml version="1.0" encoding="windows-874"?>
<!DOCTYPE mali [
<!ELEMENT mali (malisorn,malila)>
<!ELEMENT malisorn (#PCDATA)>
<!ELEMENT malila (#PCDATA)>
]>
<mali>
<malisorn>ขึ้นต้นด้วยมะลิซ้อน</malisorn>
<malila>พอแตกใบอ่อนเป็นมะลิลา</malila>
</mali>
<?xmlstylesheet type="text/xsl" href=""?>
```

สรุปคือ สำหรับเอกสารที่มีรูปแบบที่ดีคือ เอกสารเอ็กซ์เอ็มแอลที่สร้างถูกต้อง แต่หากต้องการให้ เกิดบังคับโครงสร้างแล้ว มักนิยมสร้างเอกสารเอ็กซ์เอ็มแอล โดยการอาศัยคิตีที่ติมาช่วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มีความสามารถในการค้นหาข้อมูลภายในเอกสารได้ตาม อิลิเมนต์และแอททริบิวต์รวมทั้งค้นหาแบบ ตัวอักษร (text search)

เอ็กซ์เอ็มแอล เอ็กซ์เทนเดอร์ สามารถทำอะไรกับโปรแกรมได้บ้าง

- จัดการเก็บไฟล์ข้อมูลเอ็กซ์เอ็มแอลในรูปของข้อมูลแบบคอลัมน์ หรือเป็นแบบไฟล์ภายนอก โดย การจัดเก็บข้อมูลต่างๆแต่ละอิลิเมนต์ หรือ แอททริบิวต์ลงในตารางมีการสร้างอินเด็กซ์ เพื่อที่จะ ค้นหาได้อย่างรวดเร็ว ในการจัดเก็บข้อมูลแบบคอลัมน์ มีความสามารถที่จะทำงานได้ดังนี้
 - สามารถค้นหาข้อมูลได้อย่างรวดเร็วตามอิลิเมนต์หรือ แอททริบิวต์ของข้อมูลที่จัดเก็บใน ตารางโดยการใช้อีสคิวแอล และการทำอินเด็กซ์
 - เปลี่ยนแปลง (Update) ข้อมูลภายในไฟล์เอ็กซ์เอ็มแอล
 - แยกเนื้อหาเอกสารเอ็กซ์เอ็มแอล (Extract XML) โดยใช้อีสคิวแอล
 - ตรวจสอบความถูกต้องของเอกสารเอ็กซ์เอ็มแอล (Validate XML) ขณะแทรกข้อมูล หรือ ขณะแก้ไขข้อมูล
 - สามารถค้นหาแบบตัวอักษรได้ (text search) โดยใช้ เท็กซ์เอ็กซ์เทนเดอร์ (Text Extender)
- แต่งเอกสาร หรือแยกเอกสารเอ็กซ์เอ็มแอลตามความสัมพันธ์ของตารางโดยวิธีการที่เรียกว่า เอ็กซ์เอ็มแอล คอลเล็กชัน

3.7.2 การทำงานร่วมกันของดีบีทูและเอ็กซ์เอ็มแอล

เอ็กซ์เอ็มแอล เอ็กซ์เทนเดอร์ มีคุณสมบัติต่างๆที่ช่วยในการจัดการระหว่างข้อมูลของ เอ็กซ์เอ็มแอล กับ ดีบีทู ดังต่อไปนี้

- เครื่องมือของผู้ดูแลระบบ (Administration tools) ใช้ช่วยในการรวบรวมข้อมูลเอ็กซ์เอ็มแอลใน ตารางความสัมพันธ์
- การจัดเก็บและการเข้าถึงข้อมูลเอ็กซ์เอ็มแอลภายในฐานข้อมูล
- การตรวจสอบความถูกต้องของข้อมูลเอ็กซ์เอ็มแอลด้วยการกำหนดความหมายของข้อมูลที่ดี (Data Type Definition (DTD))
- การจัดทำเอกสารเอ็กซ์เอ็มแอลให้เป็นข้อมูลที่มีความสัมพันธ์กัน โดยใช้ไฟล์สำหรับจัดทำ เอกสาร (mapping file) ที่เรียกว่าไฟล์ดีเอดี

เครื่องมือของผู้ดูแลระบบ (Administration tools) :

ใช้ช่วยในการทำให้ฐานข้อมูลและคอลัมน์ของตารางนั้นๆสามารถที่จะใช้งานกับเอกสารเอ็กซ์เอ็มแอลได้รวมทั้งสามารถที่จะจัดการข้อมูลของเอ็กซ์เอ็มแอลไปเป็น โครงสร้างความสัมพันธ์ของข้อมูลในดีบีทู

เอกสารนี้เป็นเอกสารต้นฉบับที่จัดทำขึ้นเพื่อใช้ในการศึกษาเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การจัดเก็บ และการเข้าถึงข้อมูล :

เอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์มีวิธีการจัดเก็บสองวิธี และการเข้าถึงของข้อมูลสำหรับการรวบรวมเอกสารเอ็กซ์เอ็มแอลเข้ากับโครงสร้างข้อมูลของดีบีทู วิธีการจัดเก็บข้อมูลมีสองวิธีทั้งสองมีวิธีการที่แตกต่างกัน แต่สามารถที่จะใช้ในโปรแกรมเดียวกันได้ คือ

- **วิธีเอ็กซ์เอ็มแอลคอลัมน์ (XML column method)**

วิธีนี้ช่วยในการจัดการเก็บเอกสารเอ็กซ์เอ็มแอลลงในฐานข้อมูล วิธีเอ็กซ์เอ็มแอล คอลัมน์ใช้ได้สำหรับเอกสารที่เป็นเอกสารสำคัญ เอกสารต่างๆนี้ที่ถูกแทรกจัดเก็บลงไปในคอลัมน์ที่สามารถใช้งานกับเอ็กซ์เอ็มแอล จะสามารถที่จะปรับปรุงเอกสาร นำเอกสารขึ้นมา ค้นหาเอกสารได้ ข้อมูลอิลิเมนต์ และแอททริบิวต์ จะถูกจัดการเก็บลงในตารางไซค์เทเบิล ตารางที่สร้างขึ้นมานี้สามารถทำอินเด็คซ์ เพื่อให้การค้นหาได้รวดเร็วขึ้น

- **วิธีเอ็กซ์เอ็มแอลคอลเล็คชั่น (XML collection method)**

วิธีนี้ช่วยในการจัดการเก็บ โครงสร้างของเอกสารเอ็กซ์เอ็มแอลลงในตาราง ดังนั้นคุณจึงสามารถที่จะดึงเอกสารเอ็กซ์เอ็มแอลจากข้อมูลที่มีอยู่ หรือแกะเอกสารเอ็กซ์เอ็มแอล แล้วทำการจัดเก็บข้อมูลแบบที่ไม่มีเท็กในตาราง วิธีนี้ดีสำหรับที่จะมีการแลกเปลี่ยนข้อมูลระหว่างโปรแกรม โดยเฉพาะข้อมูลที่บรรจุภายในเอกสารเอ็กซ์เอ็มแอลที่มีการเปลี่ยนแปลงแก้ไขบ่อยๆ

การตรวจสอบความถูกต้องของข้อมูลเอ็กซ์เอ็มแอล :

เอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์อนุญาตให้คุณที่จะจัดเก็บไฟล์ดีทีดี เมื่อฐานข้อมูลถูกทำให้สามารถที่จะทำงานกับเอ็กซ์เอ็มแอล แล้วตาราง DTD_REF จะถูกสร้างขึ้นในแต่ละแถวของตารางนี้จะแสดงดีทีดี และข้อมูลต่างๆผู้ใช้สามารถที่จะแทรกข้อมูลดีทีดีของตนลงไปได้ ดีทีดีใช้สำหรับตรวจสอบความถูกต้องของเอกสารเอ็กซ์เอ็มแอล

การจัดทำเอกสารเอ็กซ์เอ็มแอล :

ในการสร้างต้องระบุไปว่าภายในเอกสารเอ็กซ์เอ็มแอล นั้นมีโครงสร้างอย่างไร ข้อมูลเหล่านี้จะถูกดำเนินการโดยเอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์ โดยใช้ข้อมูลจากไฟล์ดีเอ็ดซึ่งตัวเอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์จะทำการจัดการ โครงสร้างของเอกสารเข้ากับตารางคุณต้องใช้ไฟล์ดีเอ็ดนี้ในการจัดเก็บเอกสารเอ็กซ์เอ็มแอลในเอ็กซ์เอ็มแอลคอลัมน์ หรือเมื่อต้องการที่จะแต่งหรือแกะข้อมูลเอ็กซ์เอ็มแอลในเอ็กซ์เอ็มแอลคอลเล็คชั่น

3.7.3 เอ็กซ์เอ็มแอลคอลัมน์

การจัดการข้อมูลในเอ็กซ์เอ็มแอลคอลัมน์ เมื่อใช้เอ็กซ์เอ็มแอลคอลัมน์ในการเก็บข้อมูล เอกสารเอ็กซ์เอ็มแอลทั้งหมดจะถูกเก็บลงในฐานข้อมูล ในรูปแบบคอลัมน์ทำให้สามารถที่จะทำอินเด็คซ์ และค้นหาเอกสาร นำข้อมูลออกมาจากเอกสารและเปลี่ยนแปลงเอกสาร หลังจากที่ได้ทำคำสั่งที่ทำให้ฐานข้อมูลสามารถทำงานกับเอ็กซ์เอ็มแอล แล้วก็จะมีการมี ยูดีที (user-defined type หรือ UDTs) ให้ใช้ดังนี้

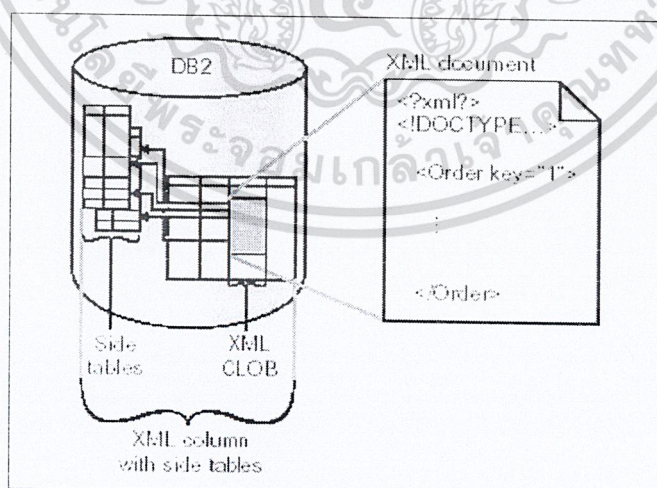
XMLCLOB	ใช้สำหรับเอกสารที่ต้องการจัดเก็บในรูปของอักขระจำนวนมาก character large object (CLOB) เหมาะสำหรับเอกสารที่มีขนาดใหญ่
XMLVARCHAR	ใช้สำหรับเอกสารที่ต้องการจัดเก็บในรูปของ VARCHAR เหมาะสำหรับเอกสารที่มีขนาดเล็ก
XMLFile	ใช้สำหรับเอกสารที่ต้องการจัดเก็บในรูปของไฟล์ที่อยู่บนเครื่องจัดเก็บ นอกฐานข้อมูล

หลังจากนี้ต้องสร้างตารางที่มีคอลัมน์ที่ทำการเก็บข้อมูลแบบยูติที่ UDT ตารางนี้เรียกว่าตาราง (XML Table) เมื่อสร้างตารางเสร็จแล้วต้องทำให้คอลัมน์นั้นสามารถจัดการกับข้อมูลเอ็กซ์เอ็มแอลได้ หลังจากนั้นจะสามารถทำการ

- จัดเก็บเอกสารเอ็กซ์เอ็มแอล
- อ่านค่าข้อมูลจากเอกสารเอ็กซ์เอ็มแอล
- เปลี่ยนแปลงข้อมูลเอกสารเอ็กซ์เอ็มแอล
- ลบข้อมูลหรือเอกสารเอ็กซ์เอ็มแอล

3.7.4 ไซด์เทเบิล (Side Table)

ไซด์เทเบิลเป็นตารางย่อยที่ใช้ในการเก็บข้อมูลที่ถอดออกมาจากเอกสารเอ็กซ์เอ็มแอล ที่ถูกใช้ในการค้นหาบ่อย โดยทำการจัดการค่าของแต่ละอิลิเมนต์หรือแอททริบิวต์เข้ากับตารางและคอลัมน์ เมื่อตารางเอกสารเอ็กซ์เอ็มแอลมีการเปลี่ยนแปลงข้อมูลภายในตาราง (Side table) ก็จะเปลี่ยนแปลงไปด้วย ก่อนที่จะทำการสร้างไซด์เทเบิลที่ใช้ในการค้นหาควรที่จะพิจารณาว่าตารางที่จะสร้างจะมีกี่ตารางมีอิลิเมนต์หรือแอททริบิวต์ไหนที่เกิดขึ้นได้หลายครั้ง



รูปที่ 3-1 เอกสารเอ็กซ์เอ็มแอลที่มีการสร้างตารางไซด์เทเบิล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.7.5 ไฟล์ดีเอดี (DAD)

สำหรับเอ็กซ์เอ็มแอลคอลัมน์ไฟล์ดีเอดีจะทำการระบุว่าเอกสารที่จะจัดเก็บภายในคอลัมน์ที่สามารถทำงานกับเอ็กซ์เอ็มแอลได้นั้นจะทำอินเดียช้อย่างไร ไฟล์ดีเอดีนั้นระบุตำแหน่งของไฟล์ดีทีดีไว้สำหรับตรวจสอบเอกสารที่จะนำมาเก็บในคอลัมน์ไว้ในแท็ก <dtid> ถ้าต้องการให้ตรวจสอบเอกสารก่อนที่จะทำการจัดเก็บลงตารางให้ใส่ค่า “YES” ภายในแท็ก <validation> หากไม่ต้องการให้ใส่ “NO” ในการระบุคอลัมน์ที่ใช้ในการดึงข้อมูลออกและจัดเก็บสามารถระบุไว้ในแท็ก <Xcolumn>

```
<?xml version = "1.0" encoding="UTF-8"?>
<!DOCTYPE DAD SYSTEM "E:\Project\Upload\webapp\save\Multimedia.dtd">
<DAD>
    <dtid>E:\Project\Upload\webapp\save\Multimedia.dtd</dtid>
    <validation>NO</validation>
    <Xcolumn>
        <table name = "TXML">
            <columnname = "Title"
                type = "varchar(50)"
                path = "/Multimedia/Title"
                multi_occurrence = "NO"
            />
        </table>
    </Xcolumn>
</DAD>
```

3.7.6 ขั้นตอนการทำให้ฐานข้อมูลสามารถทำงานกับเอ็กซ์เอ็มแอล

การจัดเก็บเอกสารและการดึงค่าเอกสารจากฐานข้อมูลด้วยเอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์ ต้องทำคำสั่งให้ฐานข้อมูลสามารถทำงานกับเอ็กซ์เอ็มแอลได้ก่อน สามารถทำได้โดยใช้คำสั่ง

การใช้คำสั่งให้ฐานข้อมูลสามารถทำงานกับเอ็กซ์เอ็มแอลจากฐานข้อมูลที่สร้างไว้แล้ว

```
dxadm enable_db dbName
dbName ชื่อของฐานที่ต้องการให้ทำงานกับเอ็กซ์เอ็มแอล
```

ตัวอย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ภายในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.7.7 สร้างตารางเอ็กซ์เอ็มแอล

ในขั้นตอนนี้เป็นส่วนที่มีขั้นตอนซับซ้อนในการสร้างตารางและการทำให้คอลัมน์สามารถทำงานกับเอ็กซ์เอ็มแอลได้

ตารางเอ็กซ์เอ็มแอลใช้สำหรับเก็บเอกสารเอ็กซ์เอ็มแอล ในการเก็บเอกสารลงฐานข้อมูลนั้น เอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์ต้องทำการ สร้างตารางที่มีคอลัมน์ที่มีชนิดเป็นเอ็กซ์เอ็มแอล ยูติที (User-defined type หรือ UDT) เอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์นั้นมียูติทีให้ทำการเก็บเอกสารเอ็กซ์เอ็มแอล สามอย่างด้วยกัน คือ XMLVARCHAR, XMLCLOB และ XMLFILE สามารถที่จะสร้างตารางที่มีคอลัมน์เป็นชนิดของเอ็กซ์เอ็มแอล ได้โดย

การสร้างตารางที่มีคอลัมน์เป็นชนิดของเอ็กซ์เอ็มแอล โดยใช้คำสั่ง

ตัวอย่าง

```
CREATE TABLE TF (
    FILENAME XMLFILE);
```

หลังจากที่สร้างตารางแล้วขั้นตอนต่อไปคือการทำให้คอลัมน์ **FILENAME** สามารถทำงานกับข้อมูลเอ็กซ์เอ็มแอลได้

3.7.8 จัดเก็บไฟล์ดีทีดีลงฐานข้อมูล

คุณสามารถที่จะใช้ ดีทีดี ในการตรวจสอบข้อมูลเอ็กซ์เอ็มแอลใน เอ็กซ์เอ็มแอลคอลัมน์ หรือ เอ็กซ์เอ็มแอลคอลเล็กชัน (XML collection) แต่ละดีทีดีที่อยู่ในตาราง (DTD_REF) มีไอดี (ID) ของแต่ละดีทีดี เอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์จะสร้างตาราง (DTD_REF) ให้หลังจากที่คำสั่งที่ทำให้ฐานข้อมูลสามารถทำงานกับข้อมูลเอ็กซ์เอ็มแอล ได้ คุณสามารถที่จะแทรกข้อมูลดีทีดีลงตารางได้โดย

การแทรกแทรกข้อมูลดีทีดีลงตารางใช้คำสั่งแทรกของเอสคิวแอลได้

ตัวอย่าง

```
db2 => INSERT INTO DB2XML.DTD_REF (
    DTDID, CONTENT, USAGE_COUNT, AUTHOR, CREATOR, UPDATOR)
VALUES ('E:\Project\Upload\webapp\save\Multimedia.dtd',
'E:\Project\Upload\webapp\save\Multimedia.dtd', 0, 'user1', 'user1', 'user1')
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อคอลัมน์	ชนิดของข้อมูล	คำอธิบาย
DTDID	VARCHAR(128)	ID ของดีทีดี
CONTENT	XMLCLOB	ตำแหน่งของไฟล์ดีทีดี
USAGE_COUNT	INTEGER	จำนวนเอ็กซ์เอ็มแอลคอลัมน์ ในฐานข้อมูลที่ใช้ดีทีดี ในการอ้างอิงของไฟล์ดีเอดี
AUTHOR	VARCHAR(128)	ชื่อผู้แต่งไฟล์ดีทีดี
CREATOR	VARCHAR(128)	ID ของผู้ใช้ที่แทรกข้อมูลคนแรก
UPDATOR	VARCHAR(128)	ID ของผู้ใช้ที่เปลี่ยนแปลงคนสุดท้าย

ตารางที่ 3-1 แสดงชนิดของข้อมูลในคอลัมน์ต่างในดีทีดีรีเฟรฟเวอร์

3.7.9 ขั้นตอนการทำให้คอลัมน์สามารถทำงานกับเอกสารเอ็กซ์เอ็มแอล

ในการจัดเก็บเอกสารเอ็กซ์เอ็มแอลในฐานข้อมูลต้องทำคำสั่งให้คอลัมน์นั้นสามารถทำงานกับเอ็กซ์เอ็มแอลได้เสียก่อน การที่ทำคำสั่งนั้นเป็นการสร้างตารางขึ้นมาให้โดยมีการทำอินเด็กซ์ ทำให้สามารถค้นหาได้รวดเร็วขึ้น โดยที่คอลัมน์นั้นต้องเป็นชนิดเอ็กซ์เอ็มแอลยูคิที

ในการทำคำสั่งให้คอลัมน์นั้นสามารถทำงานกับเอ็กซ์เอ็มแอลนั้น เอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์มีขั้นตอนการทำงานดังนี้

- อ่านไฟล์ดีเอดีเพื่อที่จะ
 - ทำการตรวจสอบการมีอยู่ของดีทีดีในตาราง (DTD_REF) ที่ได้ทำการระบุค่าไว้ในคอลัมน์ (DTDID)
 - สร้างตาราง (Side table) บนเอ็กซ์เอ็มแอลคอลัมน์เพื่อที่จะทำอินเด็กซ์
 - สร้างคอลัมน์เพื่อรองรับข้อมูลเอ็กซ์เอ็มแอล
- สร้างคอลัมน์ ROOT ID ถ้ายังไม่ได้ระบุไว้ตอนสร้างตาราง

การทำให้คอลัมน์สามารถทำงานกับเอ็กซ์เอ็มแอลโดยใช้คำสั่ง

```
dxxadm enable_column dbName tbName colName DAD_file
dbName      ชื่อของฐานข้อมูล
tbName      ชื่อของตารางที่มีคอลัมน์ที่ต้องการทำงานกับเอ็กซ์เอ็มแอล
colName      ชื่อของ XML column ที่ทำงานกับเอ็กซ์เอ็มแอล
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DAD_file ชื่อตำแหน่งของไฟล์ดีเอดี document access definition (DAD)

ตัวอย่าง

`dxadm enable_column TF TXML FILENAME`

`E:\Project\Upload\webapp\save\Multimedia.dad`

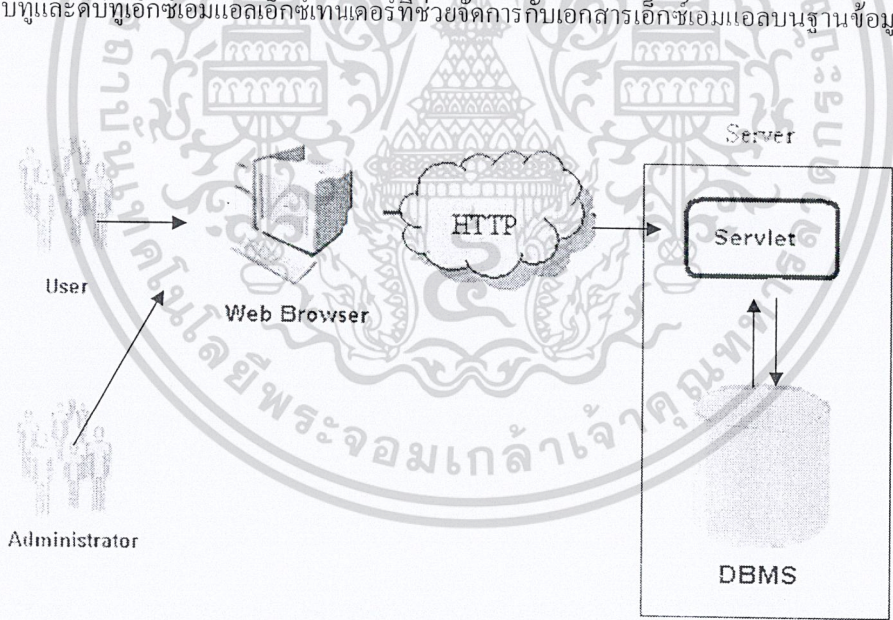


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
โครงสร้างและการออกแบบ

4.1 บทนำ

การออกแบบและสร้างเว็บแอปพลิเคชันนี้มีแนวคิดเริ่มมาจากการที่อินเทอร์เน็ตมีข้อมูลและสื่อดิจิทัลปริมาณมากมายมหาศาลอยู่ตามแหล่งต่างๆ และมีผู้ใช้บริการอินเทอร์เน็ตจำนวนมากที่ต้องการไฟล์ข้อมูลเหล่านี้ แต่อย่างไรก็ตามข้อมูลเหล่านั้นกระจัดกระจายไปทั่วโลก ดังนั้นจุดประสงค์ของโครงการคือ ต้องการสร้างแหล่งรวบรวมข้อมูลและสื่อดิจิทัลต่างๆ ไว้ในที่เดียวกัน เพื่อให้ผู้ใช้สามารถจัดเก็บ ค้นหา และดาวน์โหลดข้อมูลได้อย่างหลากหลาย โดยผู้ใช้สามารถใช้บริการแหล่งรวมข้อมูลและสื่อดิจิทัลผ่านทางหน้าเว็บไซต์ จุดเด่นของเว็บเซิร์ฟเวอร์นี้คือ การที่มีผู้เข้าใช้บริการต่างๆ เข้ามาจัดเก็บไฟล์ผู้เข้าใช้บริการจึงเหมือนเป็นผู้ที่ช่วยกันรวบรวมไฟล์หลากหลายประเภทไว้บนเซิร์ฟเวอร์เองด้วย โดยเว็บแอปพลิเคชันถูกพัฒนาด้วยเทคโนโลยีจาวาเซิร์ฟเล็ตเป็นตัวกลางสื่อสารกับผู้ใช้ทางหน้าเว็บไซต์ในการจัดการกับไฟล์ เว็บเซิร์ฟเวอร์มีหน้าที่ให้บริการต่างๆ กับผู้ใช้ นอกจากนี้ได้นำเทคโนโลยีภาษาเอ็กซ์เอ็มแอลเข้ามาช่วยในการสร้างเอกสารที่ช่วยสื่อถึงความหมายของข้อมูลเพื่อให้ไฟล์ข้อมูลเหล่านั้นมีความหมายในตัวเอง เพิ่มความสะดวกในการทำระบบค้นหาอีกด้วย เซิร์ฟเวอร์จัดการบริหารฐานข้อมูลด้วยคิวบีทูและคิวบีทูเอ็กซ์เอ็มแอลเอ็กซ์เทนเดอร์ที่ช่วยจัดการกับเอกสารเอ็กซ์เอ็มแอลบนฐานข้อมูล



รูปที่ 4-1 แสดงโครงสร้างทั้งหมดของโครงการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 เอกสารอิเล็กทรอนิกส์

ในโครงการนี้เราใช้เอกสารอิเล็กทรอนิกส์มาช่วยในการเก็บค่าคอนเทนต์ต่างๆที่รับมาจากผู้ที่ทำการจัดเก็บไฟล์ข้อมูลบนเซิร์ฟเวอร์ผ่านทางเว็บไซต์ โดยเก็บค่าคอนเทนต์ที่สื่อถึงความหมายของไฟล์นั้นๆ โดยค่าคอนเทนต์ต่างๆผู้ที่ทำการจัดเก็บจะต้องสร้างและป้อนให้กับเซิร์ฟเวอร์เอง ค่าคอนเทนต์เหล่านี้จะถูกจัดเก็บไว้ในเอกสารอิเล็กทรอนิกส์บนฐานข้อมูล

4.2.1 วิเคราะห์และออกแบบโครงสร้างเอกสารอิเล็กทรอนิกส์

ในโครงการนำเอกสารอิเล็กทรอนิกส์มาใช้สำหรับเก็บค่าคอนเทนต์ต่างๆจากผู้ใช้ได้ทำการจัดเก็บไฟล์นั้นๆไว้เพื่อทำให้ตัวไฟล์เอกสารที่ทำการจัดเก็บในฐานข้อมูลมีความหมาย เมื่อไฟล์มีคำอธิบายที่มีความหมายแล้ว เราสามารถนำค่าต่างๆในคอนเทนต์ไปแสดงต่อผู้ใช้งานว่าไฟล์เอกสารนั้นๆคือไฟล์ประเภทไหน มีชื่อว่าอะไร อยู่ในหมวดหมู่อะไรได้

จุดประสงค์หลักอีกอันหนึ่งของการกำกับความหมายให้กับไฟล์ข้อมูลในรูปแบบของเอกสารอิเล็กทรอนิกส์ ก็เพื่อให้สามารถค้นหาไฟล์ข้อมูลจากความหมายได้ ดังนั้นการกำกับความหมายจึงต้องคำนึงถึงจุดนี้ เมื่อต้องการจะค้นหาไฟล์ข้อมูลอะไรบางอย่างผู้ต้องการค้นหาจะนึกถึงคีย์เวิร์ดอะไรบางอย่างที่จะช่วยในการค้นหาเพื่อจะได้ไฟล์ข้อมูลตรงตามต้องการ เช่น ถ้าเป็นไฟล์เพลง สิ่งแรกที่นึกถึงคือชื่อเพลง อันดับต่อไปก็อาจเป็นชื่อนักร้องหรือผู้ประพันธ์เพลง จากความคิดนี้เราได้คีย์เวิร์ดสามตัว คือ ชื่อของเพลง ชื่อนักร้องหรือนักประพันธ์และสุดท้ายชนิดของไฟล์คือเป็นไฟล์ประเภทเสียง ดังนั้นหากจะสร้างเอกสารที่กำกับความหมายให้กับไฟล์ข้อมูลหนึ่งๆอย่างน้อยจะต้องมีทั้งสามคอนเทนต์นี้อยู่ คือ หัวเรื่องหรือชื่อที่เป็นที่รู้จักกันทั่วไป คำอธิบายอื่นๆนอกจากนั้นและประเภทของไฟล์ จากแนวความคิดนี้เราสามารถนำไปสร้างเอกสารอิเล็กทรอนิกส์ที่มีแท็กอธิบายถึงคอนเทนต์ดังกล่าวได้

เอกสารอิเล็กทรอนิกส์ในโครงการนี้ มีโครงสร้าง ดังนี้

```
<Multimedia>
  <Title>...</Title>
  <FileName>...</FileName>
  <FileType>...</FileType>
  <Category>...</Category>
  <Size>...</Size>
  <Date>...</Time>
  <Description>...</Description>
  <Author>...</Author>
  <Email>...</Email>
</Multimedia>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากโครงสร้างเอกสารอิเล็กทรอนิกส์ด้านบน ค่าของคอนเท้นท์ในแต่ละแท็กมีความหมายและความสำคัญดังต่อไปนี้

<Multimedia>	เป็นรูปอิลลิเมนต์ของเอกสารที่จัดเก็บแท็กที่มีค่าคอนเท้นท์ต่างๆอยู่
<Title></Title>	ใช้เก็บค่าคอนเท้นท์ที่แสดงหัวข้อหรือชื่อที่เป็นที่รู้จักกันของเอกสาร เช่น ชื่อเพลง ชื่อ หนังสือ ชื่อภาพยนตร์ เป็นต้น เป็นคอนเท้นท์สำคัญมากที่ใช้ในการค้นหา เพราะเป็นคอนเท้นท์ที่ระบุได้ตรงประเด็นมากที่สุด
<FileName></FileName>	ใช้เก็บค่าคอนเท้นท์ที่แสดงชื่อไฟล์ของเอกสาร ใช้ในการอ้างอิงถึงไฟล์ เช่น “Java Programming.pdf” “More than word.mp3” เป็นต้น
<FileType></FileType>	ใช้เก็บค่าคอนเท้นท์ที่แสดงชนิดของไฟล์ของเอกสารมีประโยชน์ในขั้นตอนใช้ตรวจสอบว่าเป็นไฟล์ประเภทไหน เพื่อเลือกโปรแกรมที่จะใช้ในการแสดงผลได้อย่างถูกต้อง
<Category></Category>	ใช้เก็บค่าคอนเท้นท์ที่แสดงหมวดหมู่หรือประเภทของไฟล์ของเอกสาร นำไปใช้ในการจัดเก็บเอกสารให้เป็นหมวดหมู่ที่เหมาะสมเพื่อให้สะดวกในการค้นหายิ่งขึ้น
<Size></Size>	ใช้เก็บค่าคอนเท้นท์ที่แสดงขนาดของเอกสาร มีหน่วยเป็นไบต์ ใช้แสดงแก่ผู้ค้นหา เพื่อเป็นทางเลือกในการเลือกไฟล์ เช่น ในกรณีไฟล์พีดีเอฟ มีหัวข้อเรื่องเหมือนกัน แต่ขนาดไฟล์ไม่เท่ากัน ผู้ค้นหาอาจต้องการไฟล์ที่มีรายละเอียดมากกว่าก็จะเลือกไฟล์ที่มีขนาดมากกว่า
<Date></Date>	ใช้เก็บค่าคอนเท้นท์ที่แสดงวันที่ทำการเก็บไฟล์เอกสารไว้บนฐานข้อมูลของเซิร์ฟเวอร์ ใช้ในการช่วยพิจารณาเช่นเดียวกับ Size
<Time></Time>	ใช้เก็บค่าคอนเท้นท์ที่แสดงเวลาที่ทำการเก็บไฟล์เอกสารไว้บนฐานข้อมูลของเซิร์ฟเวอร์ ใช้ในการช่วยพิจารณาเช่นเดียวกับ Size
<Description></Description>	ใช้เก็บค่าคอนเท้นท์ที่แสดงคำอธิบายอื่นๆที่รับมาจากผู้ทำการจัดเก็บไฟล์ของเอกสาร เป็นอีกคอนเท้นท์ที่มีความสำคัญมาก เพราะจะเก็บรายละเอียดต่างๆที่นอกเหนือไปจาก หัวเรื่อง
<Author></Author>	ใช้เก็บค่าคอนเท้นท์ที่แสดงยูสเซอร์เนมของผู้ทำการอัปโหลดไฟล์เอกสาร ช่วยเพิ่มความสะดวกในการค้นหา เช่น ต้องการไฟล์ที่ถูกจัดเก็บโดยยูสเซอร์คนนั้นๆ
<Email></Email>	ใช้เก็บค่าคอนเท้นท์ที่แสดงอีเมลแอดเดรสของผู้ทำการอัปโหลดไฟล์เอกสาร เช่นเดียวกับ Author

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างของไฟล์เอกสารเอ็กซ์เอ็มแอลที่ใช้ในโครงงานนี้

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE Multimedia SYSTEM "E:\Project\save\multimedia.dtd">
<Multimedia>
  <Title>More than word</Title>
  <FileName>More than word.mp3</FileName>
  <FileType>audio/mpeg</FileType>
  <Category>Music/Pop</Category>
  <Size>2846</Size>
  <Date>December 13, 2002</Date>
  <Time>2:14:05 AM GMT+07:00</Time>
  <Description>Extreme</Description>
  <Author>Por</Author>
  <Email>arthieg7@hotmail.com</Email>
</Multimedia>
```

จากตัวอย่างเอกสารเอ็กซ์เอ็มแอลนี้ เราสร้างเอกสารเอ็กซ์เอ็มแอลเพื่ออธิบายความหมายของไฟล์เพลงเอ็มพีสามไฟล์ ต่างๆดังตัวอย่าง เมื่อเราอ่านไฟล์เอกสารเอ็กซ์เอ็มแอลนี้จะได้ความหมายของไฟล์ที่ถูกจัดเก็บ ดังนี้

จากแท็ก **Title** จะจัดเก็บชื่อเรื่องของไฟล์เอกสารคือในกรณีนี้เป็นชื่อเพลงว่า “more than word”

จากแท็ก **FileName** มีชื่อไฟล์ว่า “more than word.mp3”

จากแท็ก **FileType** บ่งบอกไฟล์นี้เป็นไฟล์ประเภท “audio/mpeg”

จากแท็ก **Category** บ่งบอกหมวดหมู่ที่ไฟล์ถูกจัดเก็บไว้ “Music/Pop”

จากแท็ก **Size** บ่งบอกว่าขนาดของไฟล์ มีขนาด “2486” ไบต์

จากแท็ก **Date** บ่งบอกว่าไฟล์นี้ถูกจัดเก็บไว้บนเซิร์ฟเวอร์ ณ วันที่ 13 ธันวาคม ปี 2002

จากแท็ก **Time** บ่งบอกว่าไฟล์นี้ถูกจัดเก็บไว้บนเซิร์ฟเวอร์ ณ เวลา 2:14:05 am เทียบกับเวลา GMT+ 7:00

จากแท็ก **Description** บ่งบอกคำอธิบายอื่นที่เกี่ยวข้องกับไฟล์นี้ ซึ่ง “Extreme” เป็นชื่อของกลุ่มศิลปินที่ร้องและเป็นเจ้าของเพลงนี้

จากแท็ก **Author** บ่งบอกว่าผู้ที่ทำการจัดเก็บไฟล์นี้ มีชื่อยูสเซอร์เนมว่า “Por”

จากแท็ก **Email** บ่งบอกว่า “arthieg7@hotmail” เป็นอีเมลแอดเดรสของยูสเซอร์เนม “Por”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การจัดเก็บค่าคอนเทนท์ต่างๆ มีที่มา ดังนี้

จากแท็ก Title ได้ค่าคอนเทนท์มาจากการอ่านค่าพารามิเตอร์ที่ป้อนเข้ามาของผู้ใช้บริการ

จากแท็ก FileName ได้ค่าคอนเทนท์มาจากการอ่านค่าพารามิเตอร์ที่ป้อนเข้ามาของผู้ใช้บริการ

จากแท็ก FileType ได้ค่าคอนเทนท์มาจากการแปลงค่าพารามิเตอร์ที่ป้อนเข้ามาของผู้ใช้บริการ

จากแท็ก Category ได้ค่าคอนเทนท์มาจากการอ่านค่าพารามิเตอร์ที่ผู้ใช้บริการเลือกไว้

จากแท็ก Size ได้ค่าคอนเทนท์มาจากการอ่านขนาดจากไฟล์ที่ทำการจัดเก็บ

จากแท็ก Date ได้ค่าคอนเทนท์มาจากการวันที่ทำการจัดเก็บไฟล์

จากแท็ก Time ได้ค่าคอนเทนท์มาจากการเวลาที่ทำการจัดเก็บไฟล์

จากแท็ก UserName ได้ค่าคอนเทนท์มาจากการชื่อของยูสเซอร์เนมที่ใช้ในการล็อกอินอยู่และทำการจัดเก็บไฟล์

จากแท็ก Email ได้ค่าคอนเทนท์มาจากการอ่านค่าพารามิเตอร์ที่ป้อนเข้ามาของผู้ใช้บริการ โดยถ้าไม่มีการเปลี่ยนแปลงใช้ค่าอีเมลแอดเดรสของผู้ที่ทำการล็อกอินอยู่และทำการจัดเก็บไฟล์

4.2.2 โครงสร้างดีทีดี

ดีทีดี คือ ไฟล์เอกสารที่ใช้กำหนดโครงสร้างของเอกสารเอกซเอ็มแอลให้อยู่ในรูปแบบที่ถูกต้อง และตรงกันทุกๆเอกสารเอกซเอ็มแอลในฐานข้อมูล

ไฟล์ดีทีดีที่ใช้ในโครงงานมีรูปแบบดังนี้

```
<?xml version="1.0" encoding="UTF-8" ?>
<!ELEMENT Multimedia ( Title, FileName, FileType, Category, Size, Date, Time, Description, Author,
Email ) >
<!ELEMENT Author ( #PCDATA ) >
<!ELEMENT Category ( #PCDATA ) >
<!ELEMENT Date ( #PCDATA ) >
<!ELEMENT Description ( #PCDATA ) >
<!ELEMENT Email ( #PCDATA ) >
<!ELEMENT FileName ( #PCDATA ) >
<!ELEMENT FileType ( #PCDATA ) >
<!ELEMENT Size ( #PCDATA ) >
<!ELEMENT Title ( #PCDATA ) >
<!ELEMENT Time ( #PCDATA ) >
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากไฟล์ที่ดี จะแสดงโครงสร้างของเอกสารอิเล็กทรอนิกส์ไว้ดังนี้

เวอร์ชันและการเข้ารหัสที่ใช้ในเอกสารอิเล็กทรอนิกส์ จะต้องตรงกับ เวอร์ชัน 1.0 กับ “UTF-8

ตามลำดับ อิลิเมนต์แรกมีเพียงตัวเดียว คือ Multimedia โดยมี อิลิเมนต์ลูกต่างๆ คือ

Title,FileName,FileType,Category,Size,Date,Time,Description,Author และ Email ตามลำดับ

โดยค่าคอนเท้นท์ในแต่ละอิลิเมนต์ลูกทั้งหมดจะต้องเป็นแบบ PCDATA(Parse Character Data) ซึ่งเป็น

ข้อความที่จะถูกแปรภาษาก่อนแล้วค่อยแสดงข้อความออกมาได้

4.3 การจัดหมวดหมู่ในการจัดเก็บไฟล์ข้อมูล

เนื่องจากเว็บเซิร์ฟเวอร์เป็นแหล่งรวบรวมไฟล์ข้อมูลจำนวนมาก จำเป็นจะต้องจัดเก็บไฟล์เป็น

หมวดหมู่เพื่อง่ายต่อการบริหารต่างๆเกี่ยวกับไฟล์ โดยเฉพาะในเรื่องการค้นหา ถ้าไม่จัดที่อยู่ของไฟล์ไว้

ให้เป็นระเบียบจะใช้เวลามากกว่าในการหาไฟล์หนึ่งๆ

หมวดหมู่ที่ใช้ในการจัดเก็บไฟล์ต่างๆมีดังนี้



รูปที่ 4-2 แสดงหมวดหมู่ใน Upload

Upload เป็นหมวดหมู่รากของทุกหมวดหมู่ ดังนั้นไฟล์ข้อมูลและสื่อดิจิทัลที่อยู่ในระบบ

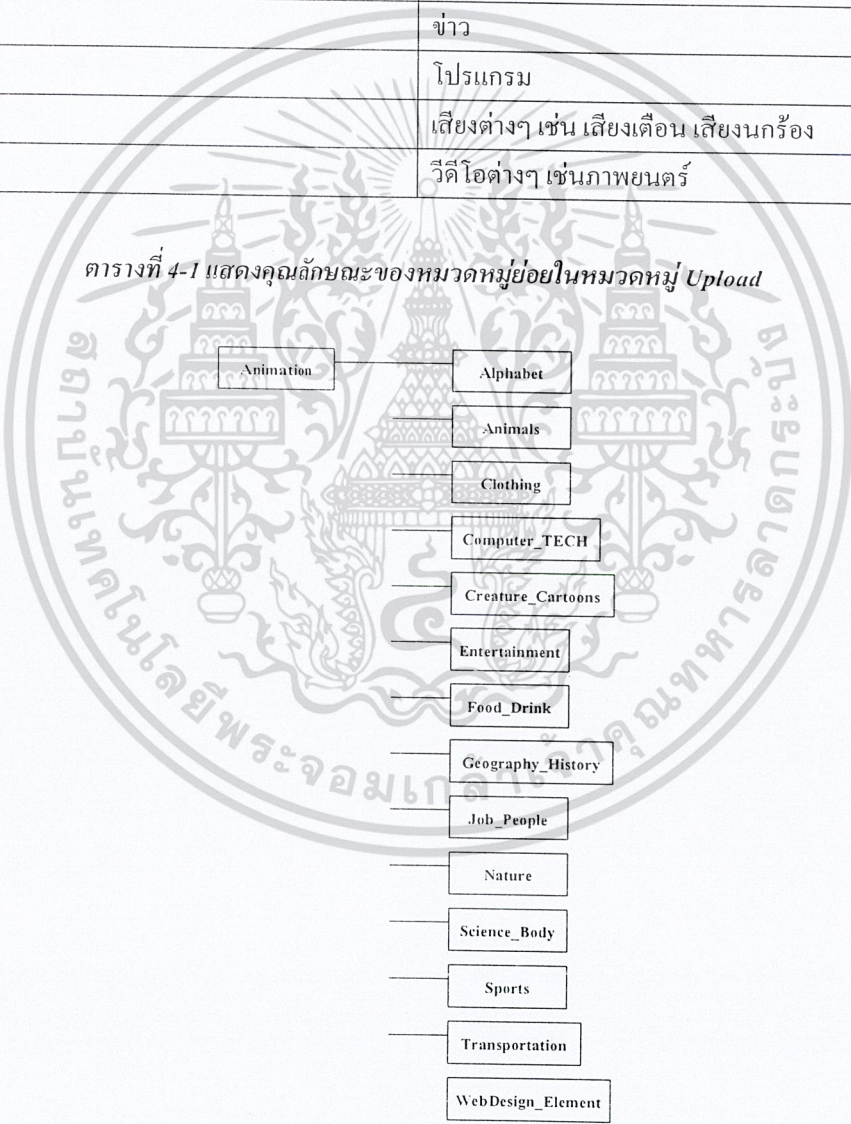
ทั้งหมดจะอยู่ในหมวดหมู่ Upload เป็น โดยในหมวดหมู่นี้ถูกแบ่งเป็นหมวดหมู่ย่อยออกไป ในแต่ละ

หมวดหมู่ย่อยบางหมวดหมู่มียังมีหมวดหมู่แบ่งแยกลงไปอีก บางหมวดหมู่ย่อยก็ไม่มี

รายละเอียดหมวดหมู่ย่อยใน Upload

หมวดหมู่	คุณลักษณะ
Animation	ภาพเคลื่อนไหว
Document	เอกสารต่างๆ เช่น เอกสารเวิร์ด เอกสารพีดีเอฟ
Game	เกมต่างๆ
Image	รูปภาพ
Music	เพลงต่างๆ เช่น ไฟล์เอ็มพีสาม
News	ข่าว
Program	โปรแกรม
Sound	เสียงต่างๆ เช่น เสียงเตือน เสียงนกร้อง
Video	วิดีโอต่างๆ เช่น ภาพยนตร์

ตารางที่ 4-1 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Upload



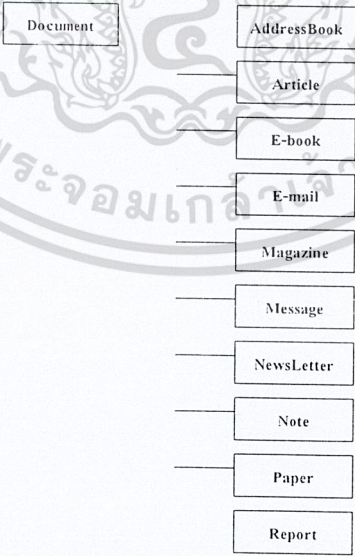
รูปที่ 4-3 แสดงหมวดหมู่ Animation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Animation

หมวดหมู่	คุณลักษณะ
Alphabet	ตัวอักษร ตัวเลขต่างๆ
Animals	จำพวกสัตว์
Clothing	เสื้อผ้า เครื่องนุ่งห่ม
Computer_TECH	เทคโนโลยีทางคอมพิวเตอร์
Creature_Cartoons	คน สัตว์ สิ่งมีชีวิต และการ์ตูน
Entertainment	สื่อบันเทิง
Food_Drink	อาหาร เครื่องดื่ม
Geography_History	ภูมิศาสตร์ ประวัติศาสตร์
Job_People	งาน คน
Nature	ธรรมชาติ
Science_Body	วิทยาศาสตร์เกี่ยวกับร่างกายสิ่งมีชีวิต
Sports	กีฬา
Transportation	การขนส่งต่างๆ
WebDesign_Element	ที่ใช้ตกแต่งหน้าเว็บเพจ

ตารางที่ 4-2 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่ Animation



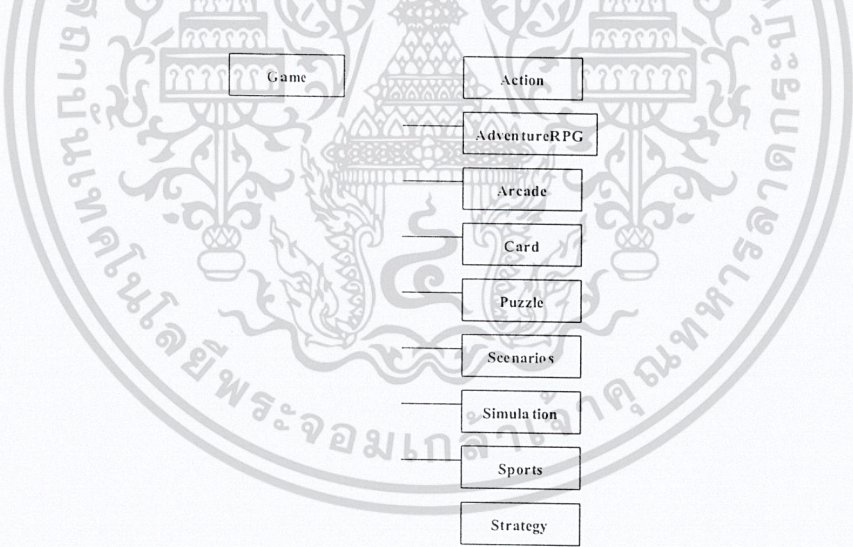
รูปที่ 4-4 แสดงหมวดหมู่Document

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Document

หมวดหมู่	คุณลักษณะ
AddressBook	สมุดบันทึกที่อยู่
Article	บทความทั่วไป
E-Book	หนังสือไฟฟ้า
E-Mail	จดหมายไฟฟ้า
Magazine	นิตยสาร
Message	ข้อความ
NewsLetter	จดหมาย
Note	ข้อความสั้นๆ
Paper	บทความทางวิชาการ
Report	รายงาน

ตารางที่ 4-3 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่Document



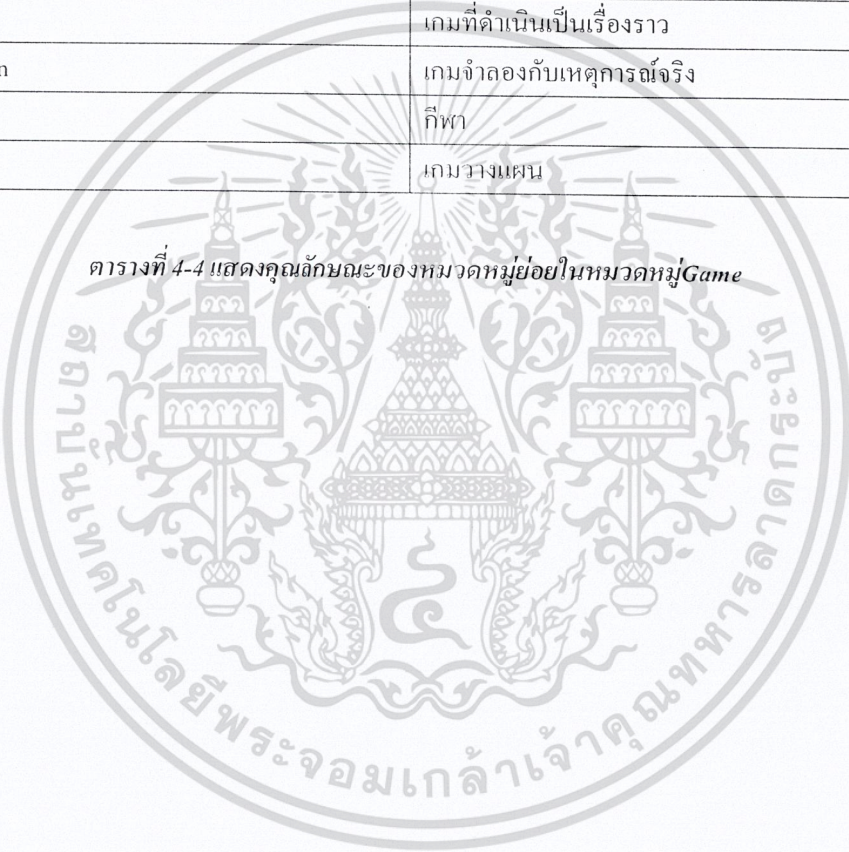
รูปที่ 4-5 แสดงหมวดหมู่Game

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

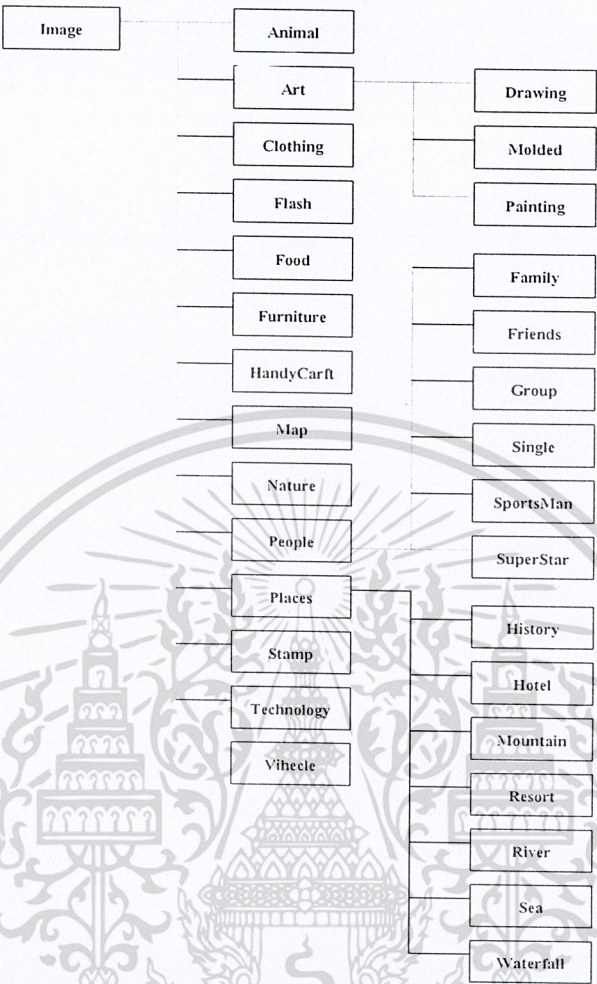
รายละเอียดหมวดหมู่ย่อยใน Game

หมวดหมู่	คุณลักษณะ
Action	เกมแนวต่อสู้
AdventureRPG	เกมผจญภัย โดยสวมบทบาทเป็นตัวละคร
Arcade	เกมตะลุยไปตามฉาก
Card	เกมไพ่
Puzzle	เกมตัวต่อ
Scenarios	เกมที่ดำเนินเป็นเรื่องราว
Simulation	เกมจำลองกับเหตุการณ์จริง
Sports	กีฬา
Strategy	เกมวางแผน

ตารางที่ 4-4 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่Game



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-6 แสดงหมวดหมู่ Image

รายละเอียดหมวดหมู่ย่อยใน Image

หมวดหมู่	คุณลักษณะ
Animal	รูปสัตว์ต่างๆ
Art	รูปศิลปะ
Clothing	เสื้อผ้า
Flash	ไฟล์ประเภทแฟลช
Food	อาหาร
Furniture	เครื่องตกแต่งเรือน
HandyCraft	ผลิตภัณฑ์ทางหัตถกรรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Map	แผนที่
Nature	ธรรมชาติ
People	รูปคนทั่วไป
Places	สถานที่ต่างๆ
Stamp	แสตมป์
Technology	เทคโนโลยี
Vehicle	ยานพาหนะ

ตารางที่ 4-5 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่Image

รายละเอียดหมวดหมู่ย่อยใน Art

หมวดหมู่	คุณลักษณะ
Drawing	ภาพวาด
Modeled	รูปปั้น
Painting	ภาพวาดพู่กัน

ตารางที่ 4-6 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่Art

รายละเอียดหมวดหมู่ย่อยใน People

หมวดหมู่	คุณลักษณะ
Family	ครอบครัว
Friends	เพื่อนๆ
Group	หมู่คณะ
Single	ภาพเดี่ยว
SportsMan	นักกีฬา
SuperStar	ดารา

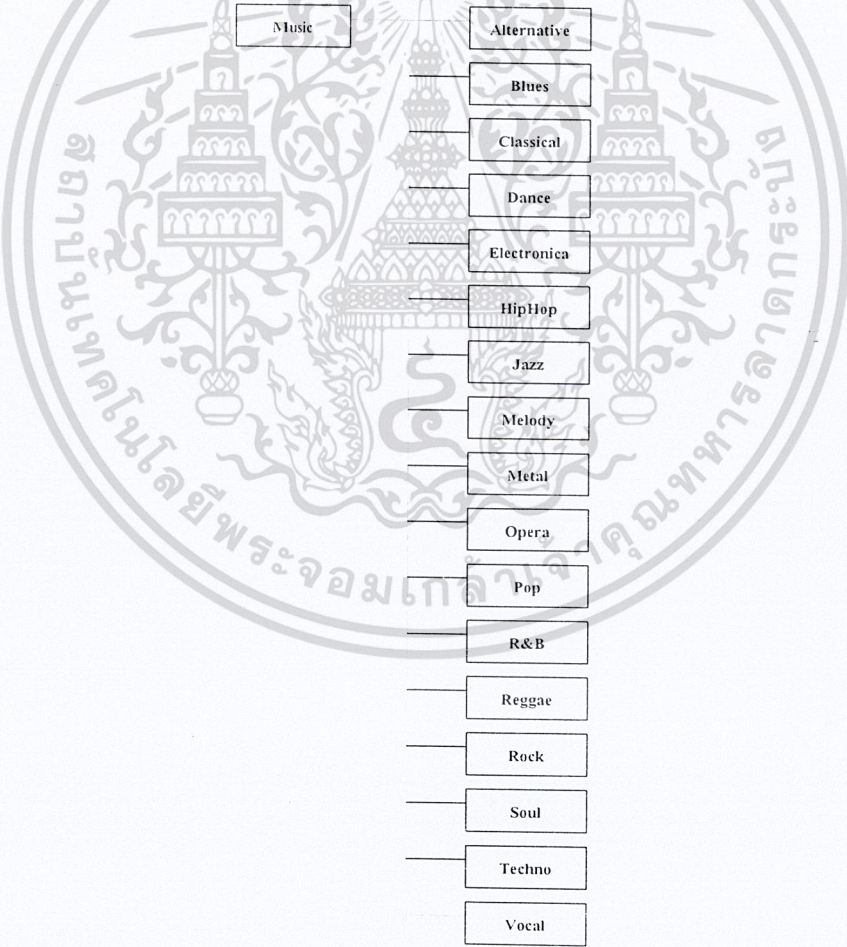
ตารางที่ 4-7 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่People

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แสดงหมวดหมู่ Places

หมวดหมู่	คุณลักษณะ
History	สถานที่ทางประวัติศาสตร์
Hotel	โรงแรม
Mountain	ภูเขา
Resort	บ้านพักตากส
River	แม่น้ำ
Sea	ทะเล
Waterfall	น้ำตก

ตารางที่ 4-8 แสดงคุณลักษณะของหมวดหมู่ย่อยในหมวดหมู่Places



รูปที่ 4-7 แสดงหมวดหมู่Music

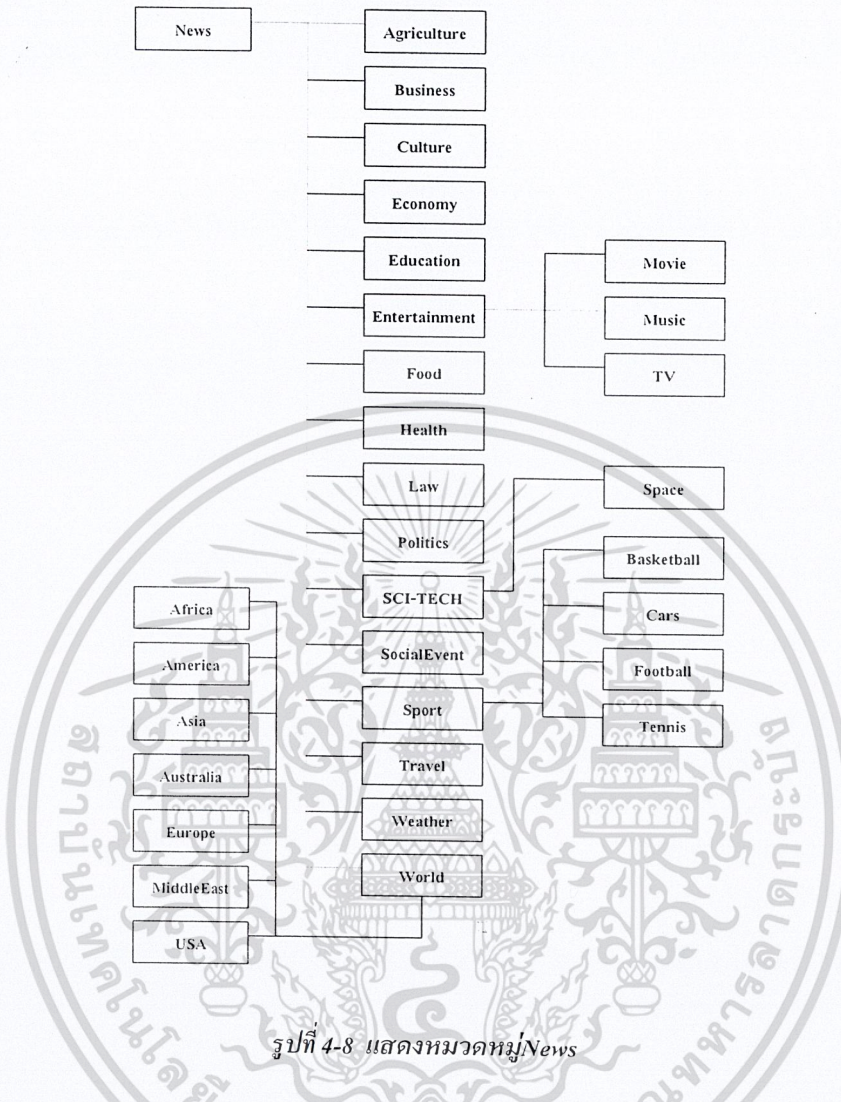
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Music

หมวดหมู่	คุณลักษณะ
Alternative	แนวลัทธิร็อก
Blues	แนวลู
Classical	เพลงคลาสสิก
Dance	เพลงเต้น
Electronica	แนวลัทธิอิเล็กทรอนิกส์
HipHop	ฮิปฮอป
Jazz	แจ๊ซ
Melody	ทำนองเพลง
Metal	แนวมทัล
Opera	โอเปร่า
Pop	ป๊อป
R&B	อาร์แอนด์บี
Reggae	เรกเก้
Rock	ร็อก
Soul	แนวลู
Techno	เทคโน
Vocal	เพลงประสานเสียง

ตารางที่ 4-9 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่ Music

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-8 แสดงหมวดหมู่News

รายละเอียดหมวดหมู่ย่อยใน News

หมวดหมู่	คุณลักษณะ
Agriculture	ข่าวเกษตรกรรม
Business	ข่าวธุรกิจ
Culture	ข่าววัฒนธรรม
Economy	ข่าวเศรษฐศาสตร์
Education	ข่าวการศึกษา
Entertainment	ข่าวบันเทิง
Food	ข่าวเกี่ยวกับการโภชนา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Health	ข่าวสุขภาพ
Law	ข่าวเกี่ยวกับกฎหมาย
Politics	ข่าวการเมือง
SCI_TECH	ข่าวเทคโนโลยีทางวิทยาศาสตร์
SocialEvent	ข่าวสังคม
Sport	ข่าวกีฬา
Travel	ข่าวการท่องเที่ยว
Weather	ข่าวพยากรณ์อากาศ
World	ข่าวรอบโลก

ตารางที่ 4-10 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่News

รายละเอียดหมวดหมู่ย่อยใน Entertainment

หมวดหมู่	คุณลักษณะ
Movie	ข่าวเกี่ยวกับภาพยนตร์
Music	ข่าวทางดนตรี
TV	ข่าวโทรทัศน์

ตารางที่ 4-11 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Entertainment

รายละเอียดหมวดหมู่ย่อยใน SCI-TECH

หมวดหมู่	คุณลักษณะ
SPACE	ข่าวเกี่ยวกับอวกาศ

ตารางที่ 4-12 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่SCI-TECH

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Sports

หมวดหมู่	คุณลักษณะ
Basketball	ข่าวบาสเก็ตบอล
Cars	ข่าวการแข่งขันรถ
Football	ข่าวฟุตบอล
Tennis	ข่าวเทนนิส

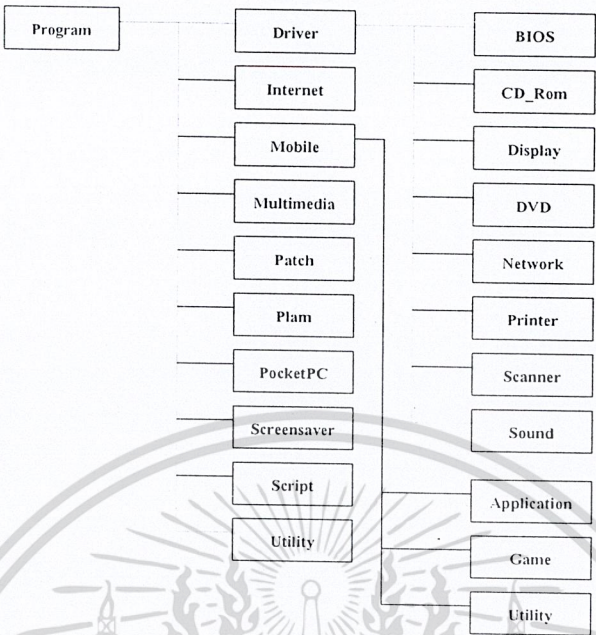
ตารางที่ 4-13 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Sports

รายละเอียดหมวดหมู่ย่อยใน World

หมวดหมู่	คุณลักษณะ
Africa	ข่าวในทวีปแอฟริกา
America	ข่าวในทวีปอเมริกา
Asia	ข่าวในทวีปเอเชีย
Australia	ข่าวในทวีปออสเตรเลีย
Europe	ข่าวในทวีปยุโรป
MiddleEast	ข่าวในแถบตะวันออกกลาง
USA	ข่าวในทวีปอเมริกา

ตารางที่ 4-14 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่World

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4-9 แสดงหมวดหมู่Program

รายละเอียดหมวดหมู่ย่อยใน Program

หมวดหมู่	คุณลักษณะ
Driver	โปรแกรมไดรฟ์เวอร์
Internet	เกี่ยวกับอินเทอร์เน็ต
Mobile	เกี่ยวกับมือถือ
Multimedia	เกี่ยวกับมัลติมีเดีย
Patch	แพทช์ที่ใช้กับโปรแกรมต่างๆ
Palm	โปรแกรมบนปาล์ม
PocketPc	โปรแกรมบนคอมพิวเตอร์พกพา
ScreenSaver	สกรีนเซฟเวอร์
Script	สคริปต่างๆ
Utility	ยูทิลิตี้ต่างๆ

ตารางที่ 4-15 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Program

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Driver

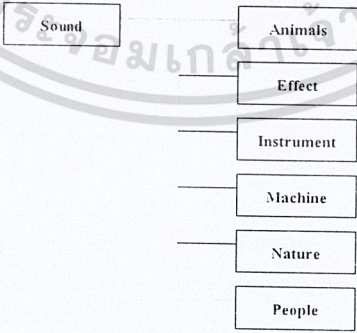
หมวดหมู่	คุณลักษณะ
BIOS	ใช้งานกับไบออส
CD_ROM	สำหรับซีดีรอมไดรฟ์
Display	สำหรับการ์ดแสดงผล
DVD	สำหรับดีวีดีไดรฟ์
Network	สำหรับอุปกรณ์ทางเน็ตเวิร์ค
Printer	สำหรับเครื่องพิมพ์
Scanner	สำหรับเครื่องสแกนเนอร์
Sound	สำหรับการ์ดเสียง

ตารางที่ 4-16 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Driver

รายละเอียดหมวดหมู่ย่อยใน Mobile

หมวดหมู่	คุณลักษณะ
Application	แอปพลิเคชันต่างๆที่ทำงานบนโมบายได้
Game	เกมบนมือถือ
Utility	ยูทิลิตีบนมือถือ

ตารางที่ 4-17 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Mobile



รูปที่ 4-10 แสดงหมวดหมู่Sound

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Sound

หมวดหมู่	คุณลักษณะ
Animals	เสียงสัตว์
Effect	เสียงประกอบต่างๆ
Instrument	เสียงเครื่องมือต่างๆ
Machine	เสียงเครื่องจักร
Nature	เสียงธรรมชาติ
People	เสียงคน

ตารางที่ 4-18 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Sound



รูปที่ 4-11 แสดงหมวดหมู่ Video

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดหมวดหมู่ย่อยใน Video

หมวดหมู่	คุณลักษณะ
Advertisement	ภาพยนตร์โฆษณา
Clip	ภาพเคลื่อนไหวสั้นๆ
FasionShow	การแสดงแฟชั่น
Feature	สารคดี
Movie	ภาพยนตร์
MusicVideo	ภาพยนตร์ประกอบเพลง

ตารางที่ 4-19 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Video

รายละเอียดหมวดหมู่ย่อยใน Movie

หมวดหมู่	คุณลักษณะ
Action	แนวบู๊ ตื่นเต้น
Adventure	แนวผจญภัย
Animation	ภาพยนตร์การ์ตูน
Comedy	แนวตลก
Crime	แนวฆาตกรรม
Drama	แนวชีวิต
Family	แนวชีวิตครอบครัว สนุกสนาน
Fantasy	แนวนิยายมหัศจรรย์
AnimalLife	ชีวิตสัตว์
Education	เพื่อการศึกษา
History	ประวัติศาสตร์
Horor	น่ากลัว สยองขวัญ
Mystery	ลึกลับ ปริศนา
Romance	หนังรัก
SCI-FI	นิยายวิทยาศาสตร์
Thriller	ตื่นเต้น
War	สงคราม

ตารางที่ 4-20 แสดงคุณลักษณะหมวดหมู่ย่อยในหมวดหมู่Movie

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 การสร้างฐานข้อมูลด้วยดีบีทู

ทำการติดตั้งดีบีทู ยูทีบี เวอร์ชัน 8.1

4.4.1 การสร้างตารางและจัดการกับเอกสารเอ็กซ์เอ็มแอลด้วยดีบีทู

เริ่มจากเรียกหน้าต่างคำสั่งขึ้นมาใช้งาน ทำการสร้างฐานข้อมูลชื่อ DC โดยใส่คำสั่งเอสคิวแอล

```
db2 "create database DC"
```

ติดต่อกับฐานข้อมูลด้วยคำสั่ง

```
db2 "connect to DC"
```

ใส่คำสั่งทำให้ฐานข้อมูลสามารถจัดเก็บเอกสารเอ็กซ์เอ็มแอลได้ ด้วยคำสั่ง

```
db2 "bind C:\Progra~1\IBM\SQLLIB\bnd\@dxxbind.lst"
```

```
dxxadm enable_db DC
```

ต่อไปทำการใส่ เอกสารดีทีดีเพื่อใช้เป็นแม่แบบให้กับเอกสารเอ็กซ์เอ็มแอลลงในตาราง DTD_REF ซึ่งดีบีทูจะสร้างให้อัตโนมัติเมื่อมีคำสั่งให้ฐานข้อมูลสามารถจัดการเอกสารเอ็กซ์เอ็มแอล

```
db2 "insert into db2xml.dtd_ref values('getstart.dtd',
```

```
db2xml.XMLClobFromFile('C:\Progra~1\IBM\SQLLIB\samples\db2xml\dtd\multimedia
.dtd'), 0, 'admin', 'admin', 'admin')"
```

ทำการสร้างตารางเพื่อจัดเก็บเอกสารเอ็กซ์เอ็มแอลด้วยคำสั่ง

```
db2 "create table MM_TAB(XML DB2XMLFILE)"
```

ก่อนที่จะใส่คำสั่งให้คอลัมน์ของเอกสารเอ็กซ์เอ็มแอลสามารถทำงานได้ ต้องสร้างเอกสารดีเอชไอให้เรียบร้อยก่อน โดยถูกจัดเก็บไว้ที่

```
C:\Progra~1\IBM\SQLLIB\samples\db2xml\dad\getstart_xcolumn.dad
```

แล้วใช้คำสั่ง

```
dxxadm enable_column sales_db sales_tab order
```

```
C:\Progra~1\IBM\SQLLIB\samples\db2xml\dad\getstart_xcolumn.dad -v MM_view
```

เมื่อเสร็จสิ้นขั้นตอนนี้ เราจะได้ตารางสำคัญ 2 ตารางที่นำไปใช้งาน คือ ตารางหลักชื่อ

MM_TAB และ TXML เป็นไชด์เทเบิลที่ถูกสร้างขึ้นตามที่ประกาศโครงสร้างไว้ในไฟล์ดีเอชไอ

4.4.2 เอกสารดีเอชไอ

เป็นไฟล์เอกสารที่ระบุว่าจะเอกสารที่จะจัดเก็บในคอลัมน์ที่สามารถทำงานร่วมกับ เอ็กซ์เอ็มแอล นั้นจะทำอินเด็กซ์ อย่างไร เอกสารดีเอชไอนั้นระบุไฟล์ ดีทีดี ไว้สำหรับตรวจสอบเอกสารที่จะนำมาเก็บในคอลัมน์ไว้ในแท็ก <didid> ถ้าต้องการให้ตรวจสอบเอกสารก่อนที่จะทำการจัดเก็บลงตารางให้ใส่ค่า “YES” ภายในแท็ก <validation> หากไม่ต้องการให้ใส่ “NO” ด้านล่างนี้เป็นไฟล์ดีเอชไอที่ใช้ในโครงการนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้


```
<?xml version = "1.0" encoding="UTF-8"?>
<!DOCTYPE DAD SYSTEM "E:\Project\db2\sample.dtd">
<!DOCTYPE DAD SYSTEM "E:\Project\db2\sample.dtd">
<DAD>
    <dtdid>sample.dtd</dtdid>
    <validation>NO</validation>
    <Xcolumn>
        <table name = "TXML">
            <columnname = "Title"
                type = "varchar(50)"
                path = "/Multimedia/Title"
                multi_occurrence = "NO"
            />
            <columnname = "FileName"
                type = "varchar(50)"
                path = "/Multimedia/FileName"
                multi_occurrence = "NO"
            />
            <columnname = "FileType"
                type = "varchar(50)"
                path = "/Multimedia/FileType"
                multi_occurrence = "NO"
            />
            <columnname = "Category"
                type = "varchar(50)"
                path = "/Multimedia/Category"
                multi_occurrence = "NO"
            />
            <columnname = "Size"
                type = "integer"
                path = "/Multimedia/Size"
                multi_occurrence = "NO"
            />
            <columnname = "Date"
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้


```

        type = "varchar(50)"
        path = "/Multimedia/Date"
        multi_occurrence = "NO"
    />

    <columnname = "Time"
        type = "varchar(50)"
        path = "/Multimedia/Time"
        multi_occurrence = "NO"
    />

    <columnname = "Description"
        type = "varchar(500)"
        path = "/Multimedia/Description"
        multi_occurrence = "NO"
    />

    <columnname = "Author"
        type = "varchar(50)"
        path = "/Multimedia/Author"
        multi_occurrence = "NO"
    />

    <columnname = "Email"
        type = "varchar(50)"
        path = "/Multimedia/Email"
        multi_occurrence = "NC"
    />

</table>

</Xcolumn>
</DAD>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.3 ตารางต่างๆบนฐานข้อมูล

MM_TAB

DXXROOT_ID	XML
------------	-----

โดย XML เป็นคอลัมน์ที่จัดเก็บเอกสารเอ็กซ์เอ็มแอล

TXML

DXXROOT_ID	TITLE	FILENAME	FILETYPE	
CATEGORY	SIZE	DATE	TIME	DESCRIPTION
AUTHOR	EMAIL			

TXML เป็นไซต์เทเบิลที่ถูกสร้างขึ้นตาม โครงสร้างในเอกสารดีเอดี

CAT_TAB

SUB1	SUB2	SUB3
------	------	------

CAT_TAB เป็นตารางเก็บข้อมูลหมวดหมู่ของประเภทไฟล์

USER_TAB

U_NAM	PASSWORD	EMAIL	GRANT
-------	----------	-------	-------

USER_TAB เป็นตารางเก็บข้อมูลของผู้ใช้

4.4.4 การจัดเก็บเอกสารเอ็กซ์เอ็มแอลบนฐานข้อมูล

เมื่อมีการรับไฟล์ข้อมูลเข้ามาเพื่อจัดเก็บ เอกสารเอ็กซ์เอ็มแอลจะถูกสร้างขึ้น และถูกจัดเก็บไว้บนตาราง MM_TAB ในตารางนี้จะมีสองคอลัมน์ คือ หนึ่งใช้จัดเก็บตำแหน่งของเอกสารเอ็กซ์เอ็มแอลในคอลัมน์ XML บนตาราง MM_TAB และสองคือ DXXROOT_ID ที่ดีบีทิวสร้างและกำหนดให้กับแต่ละเอกสารเอ็กซ์เอ็มแอล เมื่อตำแหน่งของเอกสารเอ็กซ์เอ็มแอลถูกจัดเก็บไว้บนตาราง MM_TAB แล้วดีบีทิวจะจัดการเก็บค่าคอนเท้นท์ต่างๆในแต่ละแท็ก ของเอกสารเอ็กซ์เอ็มแอลไปจัดเก็บใส่ไว้ในแต่ละคอลัมน์ของไซต์เทเบิล TXML โดยจะความสัมพันธ์จะตรงกันกับที่ประกาศไว้ในเอกสารดีเอดี โดยอัตโนมัติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 การทำงานของระบบ

4.5.1 การสมัครสมาชิก

ผู้ใช้ที่มีความประสงค์ในการใช้ระบบจะต้องทำการลงทะเบียนสมัครสมาชิกกับระบบก่อน ใน การสมัครสมาชิกผู้ใช้งานต้องกรอกรายละเอียดข้อมูลดังนี้

- ชื่อสมาชิก
- ชื่อจดหมายอิเล็กทรอนิกส์ (Email)
- รหัสผ่านที่ต้องการ
- รหัสผ่านที่ต้องการอีกครั้ง

เมื่อผู้ใช้ได้ทำการกรอกรายละเอียดข้อมูลข้างต้นแล้ว ระบบจะนำข้อมูลเหล่านี้มาทำการ ตรวจสอบข้อมูลที่ผู้ใช้งานได้ทำการกรอกมาหรือไม่ ขั้นต่อไปก็จะทำการตรวจสอบชื่อจดหมาย อิเล็กทรอนิกส์ที่กรอกมานั้นจะต้องมีเครื่องหมาย “@” อยู่ตามด้วย “.” ซึ่งตัวอักษร “@” และ “.” นั้นต้อง ไม่ติดกันอีกด้วย จากนั้นก็จะทำการตรวจสอบรหัสผ่านของผู้ใช้งานทำการกรอกมาเหมือนกันหรือไม่ ขั้น ต่อไปจะนำข้อมูลชื่อสมาชิกไปทำการตรวจสอบกับชื่อสมาชิกในฐานข้อมูลที่มีอยู่ ชื่อของสมาชิคนั้น จะต้องไม่ซ้ำกัน หากผ่านขั้นตอนเหล่านี้แล้วระบบก็จะทำการเก็บชื่อสมาชิก ชื่อจดหมายอิเล็กทรอนิกส์ และรหัสผ่านไว้ที่ฐานข้อมูล

4.5.2 การเข้าสู่ระบบ

ผู้ใช้ที่มีความประสงค์ที่จะใช้ระบบจะต้องทำการสมัครสมาชิกก่อนที่จะเข้าสู่ระบบ เมื่อสมัคร สมาชิกเรียบร้อยแล้ว ผู้ใช้ก็จะมาเริ่มเข้าสู่ระบบได้โดยการกรอกชื่อสมาชิกและรหัสผ่าน ระบบจะนำชื่อ สมาชิกและรหัสผ่านที่ผู้ใช้ทำการกรอกมาไปตรวจสอบกับฐานข้อมูลที่มีอยู่หากชื่อสมาชิกและรหัสผ่าน ถูกต้อง ระบบจะสร้างคุกกี้ (Cookie) แล้วส่งไปฝั่งตัวที่ทางฝั่งผู้ใช้ ระบบจะใช้คุกกี้เพื่อที่จะทำการ ตรวจสอบการเข้าสู่ระบบของผู้ใช้ต่อไป ในการสร้างคุกกี้ก็จะฝั่งข้อมูลสองอย่างดังนี้ ชื่อผู้ใช้ และ สิทธิ ของผู้ใช้ ซึ่งสิทธิของผู้ใช้นั้นขึ้นอยู่กับข้อมูลสิทธิที่อยู่ในฐานข้อมูล สิทธิของผู้ใช้แบ่งออกเป็นสองระดับ คือ ผู้ใช้ระดับธรรมดา กับ ผู้ใช้ที่เป็นผู้ดูแลระบบ

4.5.3 การจัดเก็บสื่อและข้อมูลดิจิทัล

การจัดเก็บสื่อและข้อมูลดิจิทัลผู้ใช้งานจะต้องทำการเข้าสู่ระบบก่อนที่จะทำการจัดเก็บได้ เมื่อผู้ใช้ เลือกที่จะทำการจัดเก็บระบบจะทำการตรวจสอบคุกกี้ที่ได้สร้างไว้ตอนที่ทำการเข้าสู่ระบบ หากทางฝั่ง ผู้ใช้มีคุกกี้ที่ระบบสร้างไว้ให้ก็จะสามารถทำการจัดเก็บสื่อและข้อมูลดิจิทัลได้โดย ผู้ใช้จะต้องทำการ เลือกหมวดหมู่หลักที่จะทำการจัดเก็บ เมื่อทำการเลือกแล้วขั้นตอนต่อไปผู้ใช้งานจะต้องทำการเลือกหมวดหมู่ ย่อยให้ตรงกับเนื้อหาของข้อมูลที่ผู้ใช้งานต้องการที่จะจัดเก็บ ผู้ใช้จะต้องทำการกรอกข้อมูลต่างๆดังนี้

- หัวเรื่อง (Title)
- ชื่อผู้ทำการจัดเก็บข้อมูล (Author)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ชื่อจดหมายอิเล็กทรอนิกส์ (Email)
- รายละเอียดอื่นของสื่อและข้อมูล (Description)
- สื่อและข้อมูลที่ประสงค์จะจัดเก็บ (FileName)

ผู้ใช้งานจำเป็นต้องทำการกรอกข้อมูลให้ครบและให้ตรงกับเนื้อหาของข้อมูลทำการจัดเก็บให้มากที่สุดเพื่อนำไปใช้ในการค้นหาข้อมูล เมื่อทำการกรอกข้อมูลต่างๆครบแล้ว ระบบจะทำการเก็บข้อมูลโดยการนำข้อมูลไปจัดเก็บภายในเซิร์ฟเวอร์ให้ตรงกับหมวดหมู่ของข้อมูลตามที่ใช้ได้ทำการเลือกมา และจะสร้างเอกสารเอ็กซ์เอ็มแอล (XML) การสร้างเอกสารเอ็กซ์เอ็มแอลนี้ข้อมูลที่บรรจุภายในเอกสารนี้ส่วนหนึ่งได้มาจากผู้ใช้และอีกส่วนหนึ่งได้มาจากข้อมูลทางฝั่งเซิร์ฟเวอร์ และข้อมูลของสื่อที่ผู้ใช้ประสงค์ที่จะจัดเก็บ เอกสารเอ็กซ์เอ็มแอลที่สร้างขึ้นมาจะถูกจัดเก็บลงภายในฐานข้อมูลดีบีทู (DB2) ซึ่งขั้นตอนการจัดเก็บข้อมูลนั้นได้กล่าวไว้ข้างต้นแล้ว

4.5.4 การเรียกดูสื่อและข้อมูลดิจิทัล

การเรียกดูสื่อและข้อมูลดิจิทัลผู้ใช้จะต้องทำการเข้าสู่ระบบก่อนเช่นเดียวกับการจัดเก็บสื่อและข้อมูลดิจิทัล เมื่อผู้ใช้เลือกที่จะทำการเรียกดูข้อมูล ผู้ใช้ต้องทำการเลือกหาข้อมูลตามหมวดหมู่ที่ผู้ใช้ต้องการโดยขั้นตอนการเรียกดูจะคล้ายคลึงกับการจัดเก็บข้อมูลกล่าวคือ ต้องทำการเลือกหมวดหมู่หลักตามด้วยหมวดหมู่ย่อยหากหมวดหมู่นั้นมีข้อมูลตามที่ใช้เรียกดูก็จะแสดงออกมา การแสดงผลการเรียกดูจะแสดงข้อมูล ชื่อของข้อมูลนั้นๆ วันที่ทำการจัดเก็บ ชื่อผู้ทำการจัดเก็บ หมวดหมู่ และ รายละเอียดเพิ่มเติมของข้อมูล หลังจากขั้นตอนการเรียกดูนี้ผู้ใช้สามารถที่จะค้นหาข้อมูลภายในหมวดหมู่ที่ผู้ใช้ประสงค์ที่จะเรียกดูต่อไป และยังสามารถที่จะเล่นข้อมูลประเภทภาพหรือเสียงได้ รวมทั้งสามารถดาวน์โหลดข้อมูลเหล่านั้นได้อีกด้วย

4.5.5 การแสดงผลสื่อและข้อมูลดิจิทัล

การแสดงผลสื่อและข้อมูลดิจิทัลผู้ใช้จะต้องทำการเข้าสู่ระบบก่อนเช่นเดียวกับการจัดเก็บสื่อและข้อมูลดิจิทัล ต่อจากเมื่อทำการเรียกดูสื่อและข้อมูลดิจิทัลแล้วนั้น หากผู้ใช้ประสงค์ที่จะให้แสดงผลสื่อและข้อมูลดิจิทัลสำหรับสื่อที่เป็นภาพหรือเสียงได้นั้น ระบบจะทำการตรวจสอบลูกก๊อของผู้ใช้เพื่อให้แน่ใจว่าได้ทำการเข้าสู่ระบบมาก่อนแล้ว โดยการแสดงผลสื่อและข้อมูลดิจิทัลนั้นจะแสดงผ่านแท็ก ตัวอย่างนี้

```
<object width="378" height="292">
```

```
<embed src="../../../MediaTest/trippygaia1.mid" width="378" height="292"></embed>
```

```
</object>
```

ภายในแท็กอ็อบเจกต์ <object> จะทำการกำหนดความกว้างและสูงของส่วนแสดงผล

แท็กเอ็มเบด <embed> เอสอาซี src เป็นการอ้างถึงข้อมูลสื่อที่ต้องการแสดงผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.6 การดาวน์โหลดสื่อและข้อมูลดิจิทัล

การดาวน์โหลดสื่อและข้อมูลดิจิทัลผู้ใช้จะต้องทำการเข้าสู่ระบบก่อนเช่นเดียวกันกับการแสดงผลสื่อและข้อมูลดิจิทัล ต่อจากเมื่อทำการเรียกดูสื่อและข้อมูลดิจิทัลแล้วนั้น หากผู้ใช้ประสงค์ที่จะทำการดาวน์โหลดสื่อและข้อมูลดิจิทัล ก็สามารถทำได้โดยเลือกไปที่การดาวน์โหลดตรงส่วนนี้เป็นการสร้างแท็กที่อ้างอิงไฟล์ธรรมดา

4.5.7 การค้นหาสื่อและข้อมูลดิจิทัล

การค้นหาสื่อและข้อมูลดิจิทัล เมื่อผู้ใช้เข้าสู่ระบบแล้ว ผู้ใช้สามารถค้นหาไฟล์ข้อมูลโดยการป้อนคีย์เวิร์ด ให้กับระบบนำมาใช้ในการค้นหา โดยในขั้นตอนแรก ระบบจะทำการแปลความหมายของคีย์เวิร์ด แล้วไปค้นหาคีย์เวิร์ดจากเอกสารอิเล็กทรอนิกส์บนฐานข้อมูล เมื่อพบจะนำผลลัพธ์เหล่านั้นมาทำการจัดอันดับในการแสดงผลให้ตรงกับความหมายที่แปลมาจากคีย์เวิร์ด แล้วจึงนำไปแสดงผลต่อผู้ค้นหาต่อไป โดยในการแสดงผลจะเหมือนกับการผลในหัวข้อ 4.5.5 ซึ่งผู้ใช้สามารถเลือกที่จะดาวน์โหลดหรือแสดงผลต่อไปได้ ส่วนการแปลความหมายของคีย์เวิร์ด และกระบวนการคำนวณจัดลำดับการแสดงผลจะกล่าวถึงในหัวข้อ 4.6

4.5.8 การลบข้อมูลสื่อและข้อมูลดิจิทัล

การลบข้อมูลสื่อและข้อมูลดิจิทัลผู้ใช้จะต้องทำการเข้าสู่ระบบก่อนเช่นเดียวกันแต่ผู้ที่ทำการลบข้อมูลสื่อได้นั้นจะต้องเป็นผู้ที่มีสิทธิในการลบข้อมูลสื่อกล่าวคือ ระบบจะทำการตรวจสอบสิทธิของผู้ใช้จากลูกที่ระบบได้สร้างไว้ในขั้นตอนการเข้าสู่ระบบหากเป็นผู้ใช้ระดับผู้ดูแลระบบจะมีลูกที่ที่มีข้อมูลสิทธิที่เป็นระดับผู้ดูแลระบบก็จะสามารถทำการลบข้อมูลสื่อได้ ส่วนติดต่อผู้ใช้ในขั้นตอนการลบนั้นก็คล้ายกับในขั้นตอนการเรียกดูสื่อและข้อมูลดิจิทัลแต่จะมีฟังก์ชันในการลบข้อมูลสื่อเพิ่มขึ้นมา การลบนั้นสามารถทำได้โดยการเลือกหมวดหมู่หลักที่ต้องการตามด้วยหมวดหมู่ย่อย หากมีข้อมูลสื่ออยู่ในหมวดนั้นก็จะสามารถที่จะเลือกได้โดยทำการเลือกเช็คบ็อก (Checkbox) แล้วเลือกทำการลบระบบก็จะทำการลบข้อมูลสื่อที่นั้นออกจากฐานข้อมูล และข้อมูลสื่อบนเซิร์ฟเวอร์

4.5.9 การลบหมวดหมู่ของสื่อและข้อมูลดิจิทัล

การลบหมวดหมู่ของสื่อและข้อมูลดิจิทัลเป็นส่วนที่สำคัญของผู้ดูแลระบบ ผู้ดูแลระบบต้องทำการตรวจสอบให้แน่ใจเสียก่อนการลบว่าหมวดหมู่นั้นไม่สำคัญอีกต่อไปแล้ว การลบหมวดหมู่นั้นผู้ดูแลระบบต้องทำการเข้าสู่ระบบเช่นเดียวกับผู้ใช้ทั่วไป ในการลบหมวดหมู่นั้นผู้ดูแลต้องเลือกฟังก์ชันการเรียกดูข้อมูลสื่อ การลบหมวดหมู่สามารถลบหมวดหมู่หลัก หมวดหมู่ย่อยที่1 และหมวดหมู่ย่อยที่2 การลบหมวดหมู่หลักจะทำได้ เมื่อเลือกฟังก์ชันการเรียกดูข้อมูลแล้ว จะมีส่วนแสดงเท็กซ์บ็อก (Textbox) ให้กรอกชื่อหมวดหมู่ที่จะทำการลบตามที่แสดงไว้ที่หน้าเว็บเพจ หากต้องการที่จะลบหมวดหมู่ย่อยที่1 ต้องทำการเลือกหมวดหมู่หลักเสียก่อน หลังจากที่ได้เลือกหมวดหมู่หลักแล้วหน้าเว็บจะแสดงหมวดหมู่ย่อยที่1 ผู้ดูแลทำการกรอกชื่อหมวดหมู่ย่อยที่1 ที่ต้องการลบออกจากระบบ หากต้องการลบหมวดหมู่ย่อยที่2 ก็

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สามารถทำได้เช่นเดียวกับการลบหมวดหมู่ย่อยที่1 โดยทำการเลือกหมวดหมู่หลักก่อนตามด้วยหมวดหมู่ย่อยที่1 หน้าเว็บจะแสดงผลหมวดหมู่ย่อยที่2 ออกมา ผู้ดูแลทำการกรอกชื่อหมวดหมู่ย่อยที่2 ที่ต้องการลบออกจากระบบ

หลังจากที่ทำการเลือกหมวดหมู่ที่ต้องการลบออกจากระบบแล้วระบบจะทำการลบชื่อหมวดหมู่ดังที่เลือกไว้ออกจากฐานข้อมูล และทำการลบไคเล็กทอรี่ (Directory) และข้อมูลสื่อต่างๆที่อยู่ภายใต้ไคเล็กทอรี่นั้นออกไปจากเซิร์ฟเวอร์

4.5.10 การสร้างหมวดหมู่ของสื่อและข้อมูลดิจิทัล

การสร้างหมวดหมู่ของสื่อและข้อมูลดิจิทัลเป็นส่วนที่สำคัญอีกเช่นเดียวกันกับการลบหมวดหมู่ เพราะผู้ดูแลต้องทำการตรวจสอบเนื้อหาข้อมูลสื่อของผู้ใช้ที่ทำการจัดเก็บมาเพื่อที่จะคัดเลือกว่าหมวดหมู่ใดที่ต้องทำการสร้างขึ้นมาเพื่อให้สอดคล้องกับความต้องการการจัดเก็บข้อมูลของผู้ใช้ ในการสร้างหมวดหมู่ของสื่อและข้อมูลดิจิทัลนั้นคล้ายกับการลบมา การสร้างหมวดหมู่นั้นผู้ดูแลต้องเลือกฟังก์ชันการเรียกดูข้อมูลสื่อ การสร้างหมวดหมู่สามารถสร้างหมวดหมู่หลัก หมวดหมู่ย่อยที่1 และหมวดหมู่ย่อยที่2ได้ การสร้างหมวดหมู่หลักจะกระทำได้เมื่อเลือกฟังก์ชันการเรียกดูข้อมูลแล้ว จะมีส่วนแท็กชื่อบ็อก (Textbox) ให้กรอกชื่อหมวดหมู่ที่จะทำการสร้างตามที่แสดงไว้ที่หน้าเว็บเพจ หากต้องการที่จะสร้างหมวดหมู่ย่อยที่1 ต้องทำการเลือกหมวดหมู่หลักเสียก่อน หลังจากที่ได้เลือกหมวดหมู่หลักแล้ว หน้าเว็บจะแสดงผลหมวดหมู่ย่อยที่1 ผู้ดูแลทำการกรอกชื่อหมวดหมู่ย่อยที่1 ที่ต้องการสร้างเข้าสู่จากระบบ หากต้องการสร้างหมวดหมู่ย่อยที่2 ก็สามารถทำได้เช่นเดียวกับการสร้างหมวดหมู่ย่อยที่1 โดยทำการเลือกหมวดหมู่หลักก่อนตามด้วยหมวดหมู่ย่อยที่1 หน้าเว็บจะแสดงผลหมวดหมู่ย่อยที่2 ออกมา ผู้ดูแลทำการกรอกชื่อหมวดหมู่ย่อยที่2 ที่ต้องการสร้างเข้าสู่ระบบ

หลังจากที่ทำการเลือกหมวดหมู่ที่ต้องการสร้างเข้าสู่จากระบบแล้วระบบจะทำการสร้างชื่อหมวดหมู่ดังที่เลือกไว้เข้าสู่ฐานข้อมูล และทำการสร้างไคเล็กทอรี่ (Directory) นั้นบนเซิร์ฟเวอร์

4.5.11 การออกจากระบบ

การออกจากระบบสามารถทำได้โดยเลือกฟังก์ชันออกจากระบบ ระบบจะทำการทำลายคุกกี้ที่ได้สร้างไว้ตอนการเข้าสู่ระบบโดยทำการตั้งค่าอายุของคุกกี้ให้เป็นศูนย์คุกกี้นั้นก็จะถูกทำลาย

4.6 ระบบค้นหา (Search engine)

ระบบค้นหาเป็นระบบที่ช่วยในการค้นหาไฟล์เอกสารในฐานข้อมูล โดยการรับคีย์เวิร์ดจากผู้ใช้งานนำมาพิจารณาแปลความหมายของคีย์เวิร์ด แล้วจึงนำไปค้นหาจากคอนเท้นต์ต่างๆของเอกสารอิเล็กทรอนิกส์ แอมแอลบนฐานข้อมูล เมื่อได้ผลลัพธ์จะนำเอกสารที่ตรงกับคีย์เวิร์ดมาทำการจัดลำดับว่าเอกสารใดมีความหมายตรงกับคีย์เวิร์ดมากที่สุด(ranking) แล้วจึงนำผลลัพธ์นั้นมาแสดงผลต่อผู้ใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.6.1 การแปลความหมายของคีย์เวิร์ด

จะพิจารณาเป็นลำดับขั้นดังนี้

- อันดับแรกรับคีย์เวิร์ดที่ต้องการนำไปใช้ค้นหาจากผู้ใช้
- พิจารณาจากสัญลักษณ์พิเศษที่ผู้ใช้ใส่เข้ามานอกเหนือจากข้อความที่ผู้ใช้ต้องการแล้วนำไปค้นหา โดยสัญลักษณ์พิเศษมีทั้งหมด 3 ตัว คือ
 - สัญลักษณ์ฟีนหนู “...” โดยใส่ครอบประโยคหรือวลีเพื่อให้ค้นหาจากทั้งกลุ่มคำนั้น
 - สัญลักษณ์บวก “+” โดยใส่ไว้หน้าคำหรือวลีเพื่อให้แสดงผลเฉพาะเอกสารที่มีคำหรือวลีนั้นๆอยู่
 - สัญลักษณ์ลบ “-” โดยใส่ไว้หน้าคำหรือวลีเพื่อให้แสดงผลเฉพาะเอกสารที่ไม่มีคำหรือวลีนั้นๆอยู่

จะพิจารณาความหมายคีย์เวิร์ดและสัญลักษณ์ในกรณีต่างๆดังนี้

1. คำหรือกลุ่มคำที่ไม่มีสัญลักษณ์พิเศษ เช่น

key : computer

ระบบจะค้นหาและแสดงผลเอกสารที่มีคำว่า “computer”

key : computer system

ระบบจะค้นหาแสดงผลเอกสารที่มีคำว่า “computer system”, “computer” และ “system” ตามลำดับ

key : computer system security

ระบบจะค้นหาแสดงผลเอกสารที่มีคำว่า “computer system security”, “computer”, “system” และ “security” ตามลำดับ

2. เมื่อใส่สัญลักษณ์ฟีนหนู ครอบกลุ่มของคำ จะมองว่า กลุ่มคำนั้นเป็นประโยคหรือวลีชุดเดียวกัน เช่น

key : “computer system” security

ระบบจะค้นหาและแสดงผลเอกสารที่มีคำว่า “computer system” และ “security” ตามลำดับ โดยไม่แสดงผลเอกสารที่มีแต่คำว่า “computer”, “system” หรือ “system computer”

key : computer “system security”

ระบบจะค้นหาและแสดงผล “computer” และ “system security” ตามลำดับ โดยไม่แสดงผลเอกสารที่มีแต่คำว่า “system”, “security” หรือ “security system”

key : “computer system security”

ระบบจะค้นหาและแสดงผล “computer system security” เท่านั้น โดยไม่แสดงผลเอกสารที่มีแต่คำว่า “computer”, “system” หรือ “security” คำใดคำหนึ่งเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. เมื่อใส่สัญลักษณ์บนหน้าคำ ระบบจะค้นหาและแสดงผลเอกสารที่มีค่านั้นเท่านั้น เช่น

key : +computer system

ระบบจะค้นหาคำว่า “computer” และ “system” แต่จะแสดงผลเอกสารเฉพาะที่มีคำว่า “computer” เท่านั้น ส่วนคำว่า “system” จะมีปรากฏหรือไม่ก็ได้

4. เมื่อใส่สัญลักษณ์บนหน้าคำ ระบบจะค้นหาและแสดงผลเอกสารที่ไม่มีค่านั้นเท่านั้น เช่น

key : computer –system

ระบบจะค้นหาคำว่า “computer” และ “system” แต่จะแสดงผลเอกสารเฉพาะที่มีคำว่า “computer” และไม่มีคำว่า “system” ปรากฏอยู่ในเอกสารนั้น

5. เมื่อใส่สัญลักษณ์บนหรือบนหน้าข้อความที่ถูกครอบงำด้วยเครื่องหมายฟันทู เช่น

key : +”computer system” security

ระบบจะค้นหาคำว่า “computer system” และ “security” แต่จะแสดงผลเฉพาะเอกสารที่มีคำว่า “computer system” ส่วน “security” จะมีปรากฏอยู่ด้วยหรือไม่ก็ได้

key : computer –“system security”

ระบบจะค้นหาคำว่า “computer” และ “system security” แต่จะแสดงผลเฉพาะเอกสารที่มีคำว่า “computer” และไม่มีคำว่า “system security” อยู่ด้วย

- หลังจากแสดงผลแล้วจะมีทางเลือกสำหรับผู้ใช้อีกว่าต้องการค้นหาเอกสารจาก แท็กอะไรเป็นพิเศษหรือไม่ เช่น ต้องการค้นหาจากเฉพาะแท็กชื่อเอกสาร หรือแท็ก <Title> นั่นเอง

4.6.2 หลักการคำนวณในการจัดอันดับของเอกสาร เพื่อแสดงผล

การจัดอันดับของเอกสาร คือ การจัดเรียงเอกสารในการแสดงผลการค้นหา มีหลักการคิดคำนวณหา

ค่าความคล้ายคลึงกันของคีย์เวิร์ดและเทอมในเอกสาร โดยใช้สูตรการหาค่าความคล้ายคลึงกันจากเวกเตอร์โมเดล ซึ่งเป็นโมเดลหนึ่งในวิชาอินฟอร์เมชัน รีทรีฟวอล(Information Retrieval) โดยมีหลักการคำนวณอยู่ คือ นำค่าน้ำหนักของเทอมของคีย์เวิร์ด และค่าน้ำหนักของอินเด็กซ์เทอมในเอกสารซึ่งใช้เทอมเดียวกันกับเทอมของคีย์เวิร์ดเท่านั้น มาสร้างเป็นเวกเตอร์ กำหนด

$\vec{q}$ $\vec{q} = (w_{1,q}, w_{2,q}, \dots, w_{i,q})$ เป็นเวกเตอร์บอกค่าน้ำหนักของแต่ละเทอมของคีย์เวิร์ด โดย i คือจำนวนทั้งหมดของคีย์เวิร์ด

$\vec{d}$ $\vec{d} = (w_{1,d}, w_{2,d}, \dots, w_{i,d})$ เป็นเวกเตอร์บอกค่าน้ำหนักของแต่ละเทอมของอินเด็กซ์เทอมในเอกสาร

k_i เป็นอินเด็กซ์เทอมลำดับที่ i

d_j เป็นเอกสาร ลำดับที่ j

$w_{i,k}$ เป็นค่าน้ำหนักแสดงความสัมพันธ์ของคู่เทอม k_i ในเอกสาร d_j

$q_{i,j}$ เป็นค่าน้ำหนักแสดงความสัมพันธ์ของคู่เทอม k_i ใน q

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เราสามารถกำหนดค่าให้แก่แต่ละ $w_{i,j}$ ได้ดังสูตร

$$w_{i,j} = f_{i,j} \times idf_i$$

โดย $f_{i,j}$ หาได้จากสูตร

$$f_{i,j} = \frac{freq_{i,j}}{\max_l freq_{l,j}}$$

ซึ่งค่า $freq_{i,j}$ คือ ค่าของความถี่ของการปรากฏของเทอมที่ i ในเอกสารที่ j และ $\max_l freq_{l,j}$ คือ ค่าความถี่ของการปรากฏที่มากที่สุดของเทอม l ในเอกสารที่ j และหาค่า idf_i ได้จากสูตร

$$idf_i = \log \frac{N}{n_i}$$

โดย N คือ ค่าจำนวนเอกสารทั้งหมดบนฐานข้อมูล และ n_i คือ ค่าจำนวนเอกสารที่มีเทอมของคีย์เวิร์ดอยู่ ส่วนการคำนวณหาค่าน้ำหนักของเทอมของคีย์เวิร์ดนั้น ทฤษฎีของซอลตันและบัคเกย์(Salton and Buckley) ได้เสนอแนะการกำหนดค่าน้ำหนักของเทอมของคีย์เวิร์ดในควายรี q ที่ได้นั้นหาได้จากสูตร

$$w_{i,q} = \left(0.5 + \frac{0.5 freq_{i,q}}{\max_l freq_{l,q}} \right) \times \log \frac{N}{n_i}$$

ซึ่ง $freq_{i,q}$ คือ ค่าความถี่ที่ปรากฏของเทอมของคีย์เวิร์ดในควายรี q และ $\max_l freq_{l,q}$ คือ ค่าความถี่ที่ปรากฏของเทอมของคีย์เวิร์ดที่มากที่สุด

เมื่อได้ค่าน้ำหนักของทุกเทอมทั้งในเวกเตอร์ $\vec{d}$ และ $\vec{q}$ จะสามารถหาค่าความใกล้เคียงกันของทั้งสองเวกเตอร์ได้จากสูตร

$$\begin{aligned} sim(d_j, q) &= \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \times |\vec{q}|} \\ &= \frac{\sum_{i=1}^n w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^n w_{i,j}^2} \times \sqrt{\sum_{i=1}^n w_{i,q}^2}} \end{aligned}$$

เรานำค่า $sim(d_j, q)$ มาใช้ในการบ่งบอกค่าความคล้ายกันของทั้งสองเวกเตอร์ โดยเมื่อค่า $sim(d_j, q)$ เข้าใกล้ 1 หมายความว่าเอกสารมีความคล้ายคลึงกันมาก แต่ถ้าค่า $sim(d_j, q)$ เข้าใกล้ 0 หมายความว่าเอกสารมีความคล้ายคลึงกันน้อย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากคำนวณหาความคล้ายคลึงกันของเอกสาร จากค่าน้ำหนักที่กำหนดไว้ในแต่ละเทอมแล้ว ในการจัดลำดับเอกสารเพื่อแสดงผลการค้นหานั้น จะมีการคูณค่าคงที่เข้าไปเพื่อเป็นการทวีคูณ ความสำคัญของเทอมนั้นๆ

ที่ได้มาจากการแปลความหมายของคีย์เวิร์ด เช่น รับคีย์เวิร์ดมาจากผู้ใช้ที่ต้องการค้นหา

key : computer security

จากคีย์เวิร์ดนี้ระบบจะทำการค้นหาคีย์เวิร์ดทั้งหมด 3 เทอม คือ “computer security” “computer” และ “security” โดยระบบจะให้ความสำคัญกับคำว่า “computer security” เป็นอันดับแรก จะคูณค่าคงที่ให้กับเทอมนี้มากกว่าอีกสองเทอมที่เหลือ ดังค่าในเวกเตอร์ของควยรี q

$$\vec{q} = (2 \times w_{1,q}, 1 \times w_{2,q}, 1 \times w_{3,q})$$

จากเวกเตอร์ ค่าน้ำหนักของ $w_{1,q}$ ถูกคูณด้วย 2 ซึ่งอีก 2 เทอมถูกคูณด้วย 1 ทำให้ค่าน้ำหนักของเทอม $w_{1,q}$ มีค่าทวีคูณขึ้นในอัตราส่วนที่มากกว่าอีก 2 เทอม เพื่อบ่งบอกว่าให้ความสำคัญกับเทอมแรกมากที่สุด ในอัตราส่วน 2:1:1

4.6.3 การคำนวณในการจัดลำดับในเอกสาร

ขั้นตอนที่ 1 รับคีย์เวิร์ดจากผู้ใช้บริการระบบที่ต้องการใช้ในการค้นหาแล้วทำการแปลความหมายของคีย์เวิร์ดตามรูปแบบที่กล่าวไว้ในหัวข้อ 4.6.1 ในขั้นตอนนี้จะแบ่งคีย์เวิร์ดตามความสำคัญคือ วลีหรือประโยคที่มีความสำคัญเป็น 2 เท่าของคำเดี่ยวๆ นอกจากนี้จะทำกำกับคีย์เวิร์ดแต่ละตัวว่ามีสัญลักษณ์พิเศษใดบ้าง “+” หรือ “-”

ขั้นตอนที่ 2 นำคีย์เวิร์ดทุกคำไปค้นหาจากฐานข้อมูล โดยใช้คำสั่ง SQL เลือกเอกสารที่มีคำคีย์เวิร์ดต่างๆ ในขั้นตอนนี้จะเลือกสร้างคำสั่ง SQL ในการเลือกเอกสารตามเงื่อนไขของคีย์เวิร์ดแต่ละตัวด้วย โดยคีย์เวิร์ดที่มีสัญลักษณ์ “+” กำกับอยู่จะต้องปรากฏอยู่ในทุกเอกสารที่ถูกเลือก คีย์เวิร์ดที่มีสัญลักษณ์ “-” กำกับอยู่จะต้องไม่ปรากฏอยู่ในเอกสารใดที่ถูกเลือกอยู่ตามเงื่อนไขที่กำหนด ในคำสั่ง SQL ทั้งคีย์เวิร์ดที่มีทั้งสองสัญลักษณ์กำกับอยู่จะถูกเชื่อมกับคีย์เวิร์ดอื่นๆด้วยคำเชื่อม AND โดยมีเงื่อนไข เท่ากับ และ ไม่เท่ากับตามลำดับ

ขั้นตอนที่ 3 เมื่อได้เซตของเอกสารที่มีคำคีย์เวิร์ดต่างๆตามเงื่อนไขแล้ว จะนำแต่ละเอกสารมาจัดลำดับในการแสดงผล โดยใช้หลักการคำนวณตามสูตรในหัวข้อ 4.6.2 ในการทำงานนั้นระบบจะคำนวณทุกค่าที่ปรากฏในสูตรการคำนวณ แต่ในการอธิบายตัวอย่างการคำนวณจะแบ่งเป็นสองส่วน จากสูตรการคำนวณหาความคล้ายคลึง คือ ผลรวมของเวกเตอร์ของค่าน้ำหนักของแต่ละคีย์เวิร์ดในแต่ละเอกสารคูณกับเวกเตอร์ของค่าน้ำหนักของแต่ละคีย์เวิร์ดใน query แล้วนำหารด้วยผลคูณของขนาดของสองเวกเตอร์ พิจารณาจากสูตรแล้วค่าที่มีผลกับค่าความคลึงเปลี่ยนแปลงไปมากที่สุดคือค่าของน้ำหนักของแต่ละคีย์เวิร์ดในแต่ละเอกสาร กับ ค่าของน้ำหนักของแต่ละคีย์เวิร์ดใน query และพิจารณาต่อไปค่าน้ำหนักของแต่ละคีย์เวิร์ดในแต่ละเอกสารเกิดจากความถี่ที่ปรากฏในเอกสารคูณกับค่า idf_i ดังนั้นจะอธิบายหลักการคำนวณเป็นสองส่วน คือ ส่วนที่ค่าความคล้ายคลึงมีผลมาจากความถี่ที่ปรากฏในเอกสาร กับ ส่วนที่ค่าความคล้ายคลึงมีผลมาจากค่า idf_i

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างการคำนวณส่วนที่มีผลมาจากความถี่

คีย์เวิร์ด : computer security

จากการแปลความหมายคีย์เวิร์ดจะได้คีย์เวิร์ดทั้งหมด 3 ตัว คือ “computer security” “computer” และ “security” สมมติมีเอกสารที่มีคีย์เวิร์ดต่างๆปรากฏอยู่อย่างน้อย 1 คีย์เวิร์ด ทั้งหมด 3 เอกสาร

เอกสารคีย์เวิร์ด	“computer security”	“computer”	“security”	ความคล้ายคลึง
เอกสารที่ 1	0	6	1	7
เอกสารที่ 2	2	4	2	10
เอกสารที่ 3	2	2	3	9
Query	2	1	1	

ตารางที่ 4-21 แสดงการคำนวณค่าความคล้ายคลึงด้วยค่าน้ำหนัก

จากตารางการด้านบน ใน query มีคีย์เวิร์ดตัวที่ 1 คือ “computer security” ให้ความสำคัญมากกว่า คีย์เวิร์ดตัวที่ 2 และ 3 (“computer” และ “security”) สองเท่า ในเอกสารที่ 1 มีคีย์เวิร์ดที่ 1,2 และ 3 ปรากฏ อยู่ 0 ครั้ง,6 ครั้ง และ 7 ครั้งตามลำดับ ดังนั้น ค่าความคล้ายคลึงจะคำนวณอย่างง่าย ๆ ได้ $(0 \times 2) + (6 \times 1) + (1 \times 1)$ จะได้เท่ากับ 7 และด้วยวิธีการคำนวณแบบเดียวกันจะได้ค่าความคล้ายคลึงของใน เอกสารที่ 2 และ 3 คือ 10 และ 9 ตามลำดับ

ผลการจัดลำดับจะได้ว่าเอกสารที่ 2 มีความคล้ายคลึงกับ คีย์เวิร์ดมากที่สุด รองลงไปคือเอกสารที่ 3 และ 1 ตามลำดับ

พิจารณาจากตัวอย่าง ค่าความถี่ของคีย์เวิร์ดตัวที่ 1 ซึ่งเป็นวลี “computer security” มีผลต่อการค่า ความคล้ายคลึงมากที่สุด จากเอกสารที่ 2 และ 3 มีคีย์เวิร์ดที่ 1 ปรากฏอยู่เพียง 2 ครั้ง แต่มีค่าความ คล้ายคลึงมากกว่าเอกสารที่ 1 เพราะเอกสารที่ 1 ไม่มีปรากฏคีย์เวิร์ดตัวที่ 1 เลย ถึงแม้ในเอกสารที่ 1 จะมี คีย์เวิร์ดตัวที่ 2 ปรากฏอยู่ถึง 6 ครั้ง

ตัวอย่างการคำนวณส่วนที่มีผลมาจากค่า *idf*, จากตัวอย่างเดียวกัน พิจารณาสูตร

$$idf_i = \log \frac{N}{n_i}$$

ค่า *idf*_i มีผลมาจากค่า *n_i* และ *N* โดยสมมติค่า *N* มีค่าเท่ากับ 100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารเดี่ยวเวิร์ด	“computer security”	“computer”	“security”	ความคล้ายคลึง
เอกสารใดๆ	1.6989	1.5228	1.5228	4.7445
Query	1	1	1	

ตารางที่ 4-22 แสดงการคำนวณค่าความคล้ายคลึงด้วยค่าidf,

ค่าที่ปรากฏในตารางของแต่ละเวิร์ดในเอกสารใดๆ คือค่า idf_i ของแต่ละเวิร์ด โดยในเวิร์ดตัวที่ 1 มีเอกสารที่ปรากฏเวิร์ดตัวที่ 1 อยู่ 2 เอกสาร ดังนั้น idf_i หาได้จาก $\log(100/2)$ ซึ่งมีค่าเท่ากับ 1.6989 และเวิร์ดตัวที่ 2 และ 3 มีเอกสารที่มีเวิร์ดปรากฏอยู่แค่เวิร์ดละ 3 เอกสาร ดังนั้นค่า idf_i หาได้จาก $\log(100/3)$ ซึ่งเท่ากับ 1.5228 ค่าความคล้ายคลึงเราจะไม่สนใจตัวอย่างนี้ ค่าที่น่าสนใจคือค่า idf_i ซึ่งจะนำไปใช้คูณกับค่าน้ำหนักในการหาค่าความคล้ายคลึงกันในตัวอย่างแรก

พิจารณาฟังก์ชัน $\log$ เมื่อค่า n ยิ่งมากขึ้นจะทำให้ค่า $\log$ หรือ idf_i มีค่าลดลงเรื่อยๆ ดังนั้นเวิร์ดที่ปรากฏอยู่ในเอกสารจำนวนมากๆ จะทำให้มีค่า idf_i น้อยกว่าเวิร์ดที่ปรากฏอยู่ในเอกสารจำนวนน้อยๆ ทำให้เวิร์ดที่ปรากฏในเอกสารจำนวนมากๆเราจะให้ค่าความสำคัญน้อยกว่า เพราะถือว่าเวิร์ดที่ปรากฏน้อยกว่าน่าจะมีความสำคัญมากกว่า เนื่องจากค่า idf_i นี้จะต้องนำไปคูณกับค่าความถี่ ก่อนนำไปคำนวณและส่งผลถึงค่าความคล้ายคลึงด้วย กล่าวให้เข้าใจง่ายๆคือ สมมติมีเอกสารที่ได้จากการค้นหา 100 เอกสาร โดยใช้เวิร์ด “computer” กับ “security” ปรากฏว่าใน 100 เอกสาร มีคำว่า “computer” อยู่ 95 เอกสาร และมีเอกสารที่มีคำว่า “security” อยู่ 10 เอกสาร จะเห็นได้ว่า “computer” เป็นคำที่ปรากฏอยู่เกือบทุกเอกสาร (95 จาก 100) เราจึงควรจะให้ค่าความสำคัญหรือค่าน้ำหนักกับ “security” มากกว่า ***หมายเหตุ ในขั้นตอนการจัดเรียงเอกสารนี้ เวิร์ดที่ถูกกำกับโดยสัญลักษณ์ “+” และ “-” ไม่มีผลใดๆต่อค่าความคล้ายคลึง

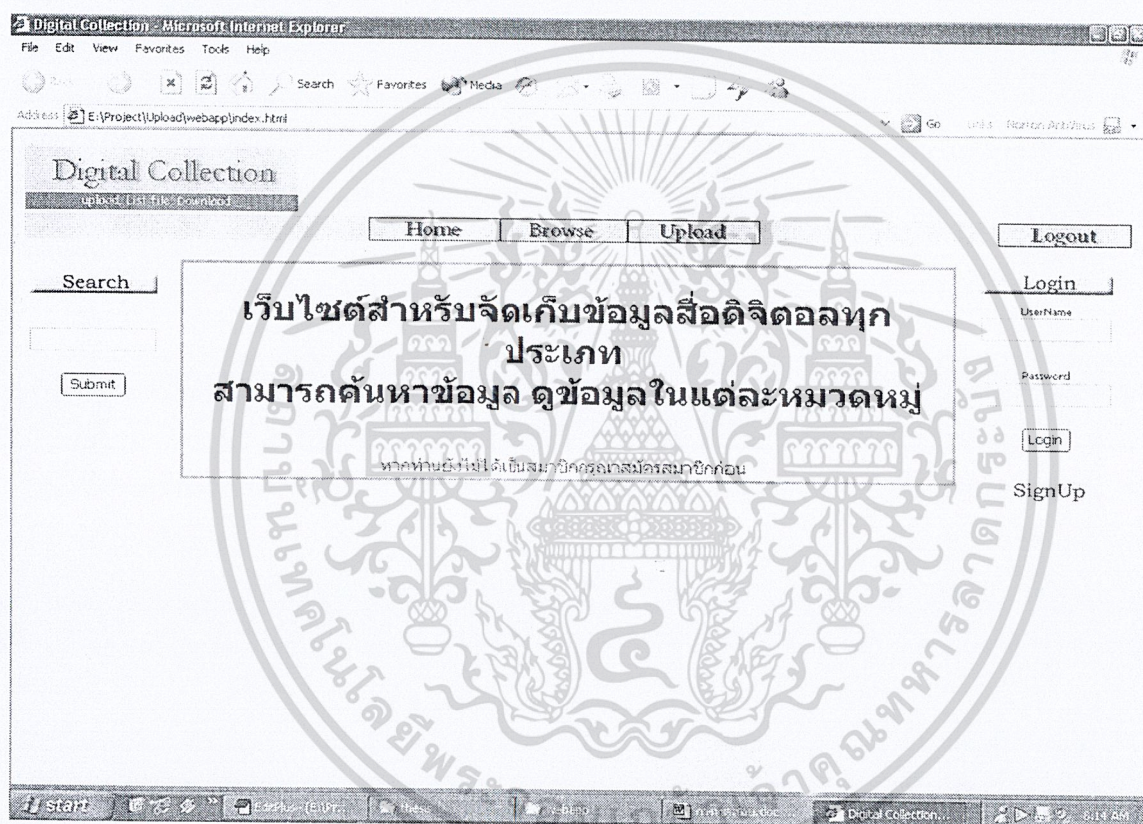
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

ผลการทำงานของระบบ

5.1 หน้าจอหลักของเว็บไซต์

ในรูป 5-1 คือหน้าแรกของเว็บไซต์ ผู้ใช้บริการต้องทำการล็อกอินเข้าระบบก่อน หรือถ้ายังไม่ได้เป็นสมาชิกก็ต้องทำการสมัครสมาชิก(Sign Up) ก่อน



รูปที่ 5-1 หน้าจอหลักของเว็บไซต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 การสมัครสมาชิกใหม่

เมื่อเลือกที่เมนูสมัครสมาชิกใหม่ (Sign Up) จะปรากฏหน้าให้กรอกชื่อผู้ใช้ อีเมลแอดเดรส และ กำหนดรหัสผ่าน ระบบจะนำข้อมูลเหล่านี้ไปเก็บบนฐานข้อมูล

Digital Collection - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites Media

Address: E:\Project\Upload\webapp\index.html

Go Link New Bookmarks

Digital Collection

upload List file Download

Home Browse Upload Logout

Search

Submit

Username: PCR

Email: Pcr@yahoo.com

Password:

Retype Password:

Submit Reset

Start

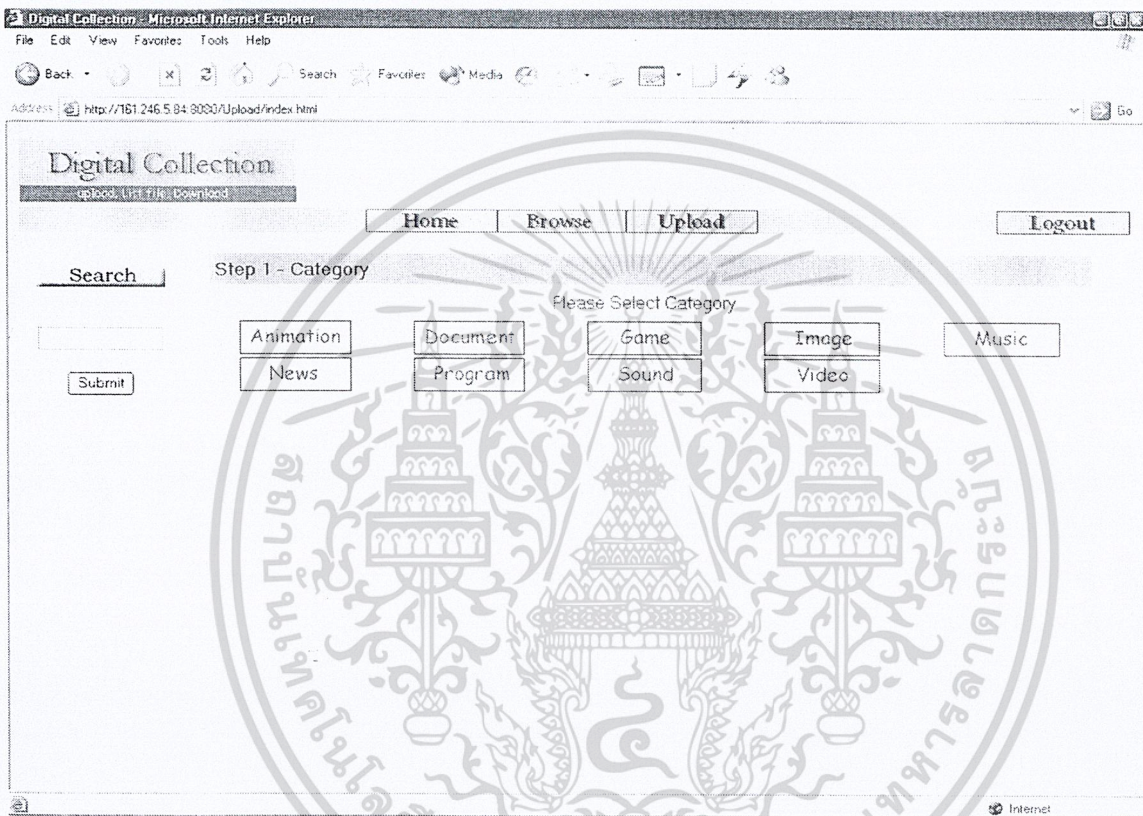
8:09 AM

รูปที่ 5-2 แสดงหน้าเว็บการสมัครสมาชิกใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 การจัดเก็บข้อมูลสื่อ และข้อมูลดิจิทัล

หลังจากทำการล็อกอินเข้าระบบแล้ว เลือกเมนู อัปโหลด(Upload) จะเข้ามาหน้าเว็บเพจดังรูปที่ 5-3 จะมีหมวดหมู่ให้เลือกสำหรับจัดเก็บไฟล์ข้อมูล



รูปที่ 5-3 แสดงหน้าเว็บการเลือกหมวดหมู่ในอันดับแรก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites Media Print View Source

Address: http://151.245.5.94:8080/Upload/index.html

Digital Collection

Upload File / File Download

Home Browse Upload Logout

Search

Step 1 - Category

Category: Music Back

Step 2 - Sub Category

Please Select Sub Category

- ☐ Alternative None
- ☐ Blues None
- ☐ classical None
- ☐ Dance None
- ☐ Electronica None
- ☐ HipHop None
- ☐ Jazz None
- ☐ Melody None
- ☐ Metal None
- ☐ Opera None
- ☐ Pop None

Submit

UserName Password Login

Done Internet

รูปที่ 5-4 แสดงหน้าเว็บการเลือกหมวดหมู่ในขั้นที่สอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Digital Collection - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media

Address: http://161.246.5.94:8080/Upload/index.html Go

Digital Collection

Upload List File Download

Home Browse Upload Logout

☐ Soul None
☐ Techno None
☐ Vocal None

Step 3 - Sender

Title : Cheek

Author : FOR

E-mail : papor@hotmai.com

Step 4 - File Description

Description : Cheeky song

Step 5 - Upload the file

File name : D:\song\cheek\in\cheek Browse

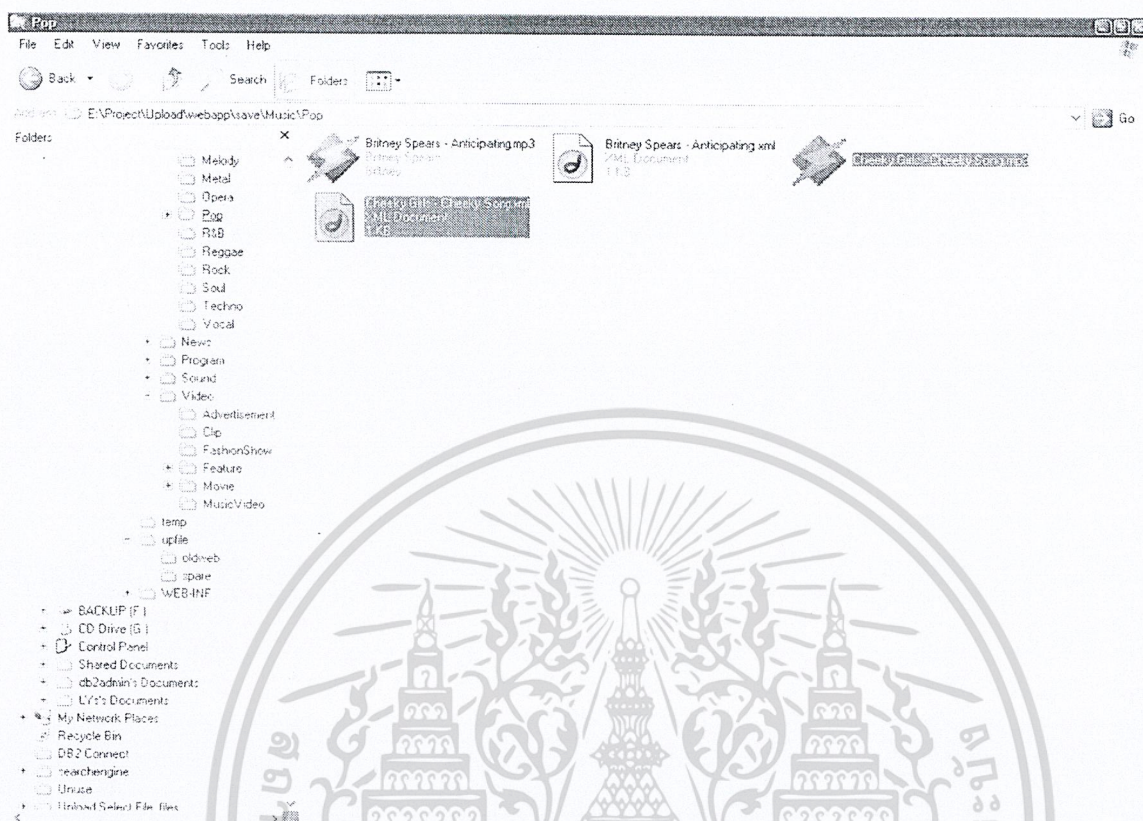
Submit Reset

Done Internet

รูปที่ 5-5 แสดงหน้าเว็บการเลือกหมวดหมู่ในขั้นที่สอง(ต่อ)

จากรูปที่ 5-3 และ 5-4 ซึ่งเป็นหน้าเว็บเพจเดียวกัน ผู้ทำการจัดเก็บจะต้องเลือกหมวดหมู่ให้กับไฟล์ในขั้นที่สอง โดยในบางหมวดหมู่มียังมีหมวดหมู่ย่อยที่ผู้ทำการจัดเก็บสามารถเลือกได้อีก เมื่อเลือกหมวดหมู่แล้วจะต้องกรอกข้อมูลตามแบบฟอร์มที่จัดไว้ให้ ประกอบไปด้วย หัวข้อ(Title) ชื่อผู้ทำการจัดเก็บ (Author) รายละเอียดอื่นๆเกี่ยวกับไฟล์ (Description) และ สุดท้ายเลือกไฟล์ที่ต้องการจะจัดเก็บในขั้นตอนที่ 5 เมื่อกรอกข้อมูลครบถ้วน ก็กดปุ่ม Submit

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5-6 แสดงไฟล์ข้อมูลที่ถูกจัดเก็บพร้อมกับเอกสารอิเล็กทรอนิกส์

Query Result		
	FILENAME	DOCROOT_ID
1	E:\Project\Uploadwebapp\save\Music\Pop\Cheeky Girls - Cheeky Song.xml	20030319085042557211000000

รูปที่ 5-7 แสดงฐานข้อมูลที่ทำกรจัดเก็บเอกสารอิเล็กทรอนิกส์ โดยเครื่องมือของดีบีทู(1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Query Result											
	D:\ROOT_ID	TITLE	FILENAME	FILETYPE	CATEGORY	SIZE	DATE	TIME	DESCRIPTION	AUTHOR	EMAIL
1	20030319085042557...	Cheeky	Cheeky Girls - Cheeky Song.m...	audio/mpeg	Music/Pop	3216614	March 19, 2	3:50:41 PM	Cheeky song POR		porpor@hotmail.c...

รูปที่ 5-8 แสดงฐานข้อมูลที่ทำกรจัดเก็บเอกสารอิเล็กทรอนิกส์ โดยเครื่องมือของดีบีทู(2)

เมื่อผู้ใช้บริการดำเนินการเสร็จสิ้น ระบบจะทำการจัดเก็บไฟล์ข้อมูล สร้างเอกสารอิเล็กทรอนิกส์พร้อมกับเก็บลงบนฐานข้อมูล ในรูปที่ 5-6 แสดงไฟล์ข้อมูลที่ถูกจัดเก็บไว้ที่หมวดหมู่ตามที่ผู้ทำการจัดเก็บเลือกเอาไว้ รูปที่ 5-7 และ 5-8 แสดงฐานข้อมูลที่ทำกรจัดเก็บเอกสารอิเล็กทรอนิกส์ลงในตาราง MM_TAB และ TXML ตามลำดับ

Digital Collection
Upload List File Download

Home Browse Upload Logout

Search

กรณเลือกหมวดหมู่ทางด้านซ้ายมือ

FileName	Date	Author	Category	Description	Action
Cheeky Girls - Cheeky Song.mp3	March 19, 2003	POR	Music/Pop	Cheeky song	View Download
India arie - can i walk with you.mp3	March 19, 2003	POR	Music/Pop	india arie can i walk with you this is great song	View Download
Counting Crows - Big Yellow Taxi.mp3	March 19, 2003	POR	Music/Pop	big boy song	View Download
Pink - Family Portrait.mp3	March 19, 2003	POR	Music/Pop	live song and her first album	View Download
Simple Plan - Id do anything.mp3	March 19, 2003	POR	Music/Pop	news rookies	View Download
Tatu - All The Things She Said.mp3	March 19, 2003	POR	Music/Pop	Tatu the first position in chart	View Download
ATB - Bring It Back.mp3	March 19, 2003	POR	Music/Pop	Smart guy	View Download

Animation
Document
Game
Image
Music
Alternative
Classical
Dance
Electronic
Hip-hop
Jazz
Melody
Opera
Pop
R&B
Reggae
Rock

รูปที่ 5-9 แสดงการเรียกดูไฟล์ข้อมูลที่อยู่บนเซิร์ฟเวอร์โดยผู้ใช้บริการทั่วไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Figure 5-10: A screenshot of a web application titled "Digital Collection" running in Microsoft Internet Explorer. The browser address bar shows "E:\Project\Upload\webapp\TMPx3lhc2r0f.htm". The application has a navigation bar with "Home", "Browse", "Upload", and "Logout" buttons. A "Search" box is present on the left. The main content area displays a table of digital files under the heading "กรุณาเลือกหมวดหมู่ทางด้านซ้ายมือ". The table has columns for "FileName", "Date", "Author", "Category", "Description", and "Action". The "Action" column contains "View" and "Download" links for each file.

FileName	Date	Author	Category	Description	Action
☐ Cheeky Girls - Cheeky Song.mp3	March 19, 2003	POR	Music/Pop	Cheeky song	View Download
☐ India Arie - Can I Walk with You.mp3	March 19, 2003	POR	Music/Pop	India Arie can I walk with you this is great song	View Download
☐ Counting Crows - Big Yellow Taxi.mp3	March 19, 2003	POR	Music/Pop	big boy sing	View Download
☐ Pink - Family Portrait.mp3	March 19, 2003	POR	Music/Pop	live song and her first album	View Download
☐ Simple Plan - I'd Do Anything.mp3	March 19, 2003	POR	Music/Pop	news rookies	View Download
☐ Tatu - All The Things She Said.mp3	March 19, 2003	POR	Music/Pop	Tatu the first position in chart	View Download
☐ ATB - Bring It Back.mp3	March 19, 2003	POR	Music/Pop	Smart guy	View Download

รูปที่ 5-10 แสดงการเรียกดูไฟล์ที่อยู่บนเซิร์ฟเวอร์โดยผู้ดูแลระบบ

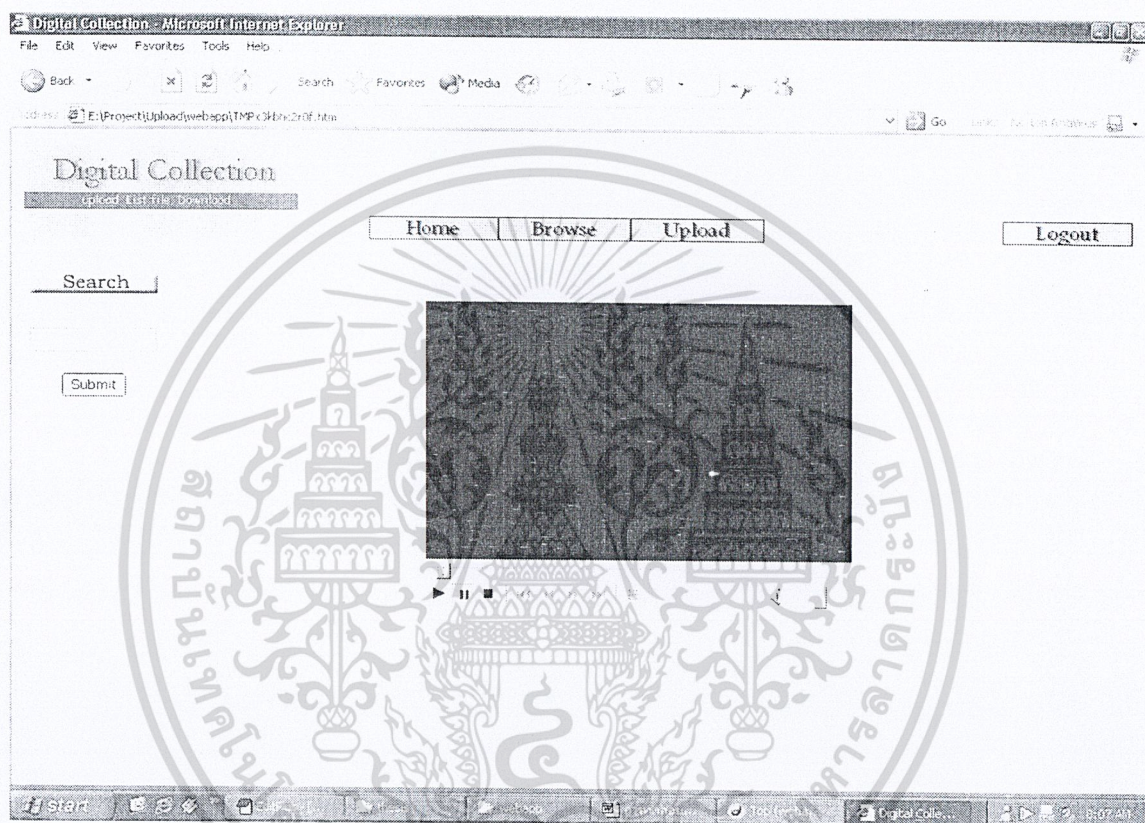
5.4 การเรียกดูไฟล์ข้อมูลบนเซิร์ฟเวอร์

จากรูปที่ 5-9 และรูปที่ 5-10 แสดงรายการของไฟล์ข้อมูลที่อยู่บนเว็บเซิร์ฟเวอร์โดยเลือกที่การเลือกเมนูเบราว์เซอร์ (Browse) หมวดหมู่ แล้วเลือกที่หมวดหมู่ดนตรีและแนวเพลงป๊อป (music/pop) โดยผู้ใช้บริการทั่วไป กับผู้ดูแลระบบตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.5 การแสดงผลข้อมูล

ทำได้โดยการขอเรียกดูไฟล์ข้อมูล (Browse) หรือจากการค้นหาไฟล์ข้อมูล (Search) แล้วเลือกวิว (view) ที่ไฟล์ข้อมูลที่ต้องการจะเล่น เลือกวิวกับไฟล์วิดีโอ (นามสกุล asf) หนึ่งจะปรากฏหน้าเว็บเพจดังรูป 5-11



รูปที่ 5-11 แสดงหน้าแสดงผลในการเล่นไฟล์ข้อมูลด้วยวินโดวมีเดียเพลย์เยอร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6

บทสรุป และวิจารณ์

6.1 สรุปและวิจารณ์

เป้าหมายของโครงการคือ พัฒนาเว็บเซิร์ฟเวอร์ให้เป็นแหล่งรวบรวมสื่อและข้อมูลดิจิทัลในรูปแบบของไฟล์ โดยอาศัยการให้บริการต่างๆแก่ผู้ใช้ให้สามารถจัดเก็บ ค้นหา และดาวน์โหลดไฟล์ข้อมูลจากบนเซิร์ฟเวอร์ได้ และไฟล์ที่ถูกจัดเก็บได้รับการกำกับความหมายให้ เพื่อสื่อถึงความหมายของตัวเองได้และจัดเก็บไฟล์ไว้เป็นหมวดหมู่ โดยสรุปเว็บเซิร์ฟเวอร์มีฟังก์ชันที่สามารถให้บริการแก่ผู้ใช้ได้ดังนี้

- การสมัครเป็นสมาชิก
- การล็อกอินเข้าสู่ระบบ
- การจัดเก็บไฟล์ข้อมูล
- การค้นหาไฟล์ข้อมูล
- การดาวน์โหลดไฟล์ข้อมูล
- การขอดูไฟล์ข้อมูล
- การเล่นไฟล์เสียงและภาพบนเว็บ
- การย้ายหรือลบไฟล์ข้อมูล
- การสร้างหรือลบหมวดหมู่

จากบริการข้างต้นแม้ระบบจะไม่สามารถให้บริการแต่ละฟังก์ชันได้อย่างสมบูรณ์ในการสนองต่อความต้องการของผู้ใช้ แต่เว็บเซิร์ฟเวอร์ก็สามารถที่จะให้บริการหลักๆ ได้ในระดับหนึ่ง กล่าวคือ เว็บเซิร์ฟเวอร์สามารถให้บริการจัดเก็บ สร้างความหมายให้กับไฟล์ข้อมูล และให้ผู้ใช้บริการสามารถค้นหา ขอดู และดาวน์โหลดไฟล์ไปใช้งานได้

6.2 ปัญหาที่เกิดขึ้น

ปัญหาที่เกิดขึ้น ในรายละเอียดปลีกย่อยนั้นมีมากมาย ส่งผลให้การให้บริการแก่ผู้ใช้อยังไม่สมบูรณ์ ดังเช่น

- การรับค่าข้อมูลต่างๆจากผู้ใช้ จากฟอร์มที่กำหนดให้นั้น ข้อมูลที่ผู้ใช้บริการป้อนเข้ามาให้ นั้นไม่สามารถนำไปสร้างความหมายที่ดีมีประสิทธิภาพแก่ไฟล์ได้ ซึ่งปัญหานี้เกิดมาจากทางด้านผู้เข้าใช้บริการเอง เมื่อผู้ใช้ป้อนข้อมูลที่ไม่สื่อความหมาย หรือสื่อความหมายได้ไม่ดี ความหมายที่ใช้ในการกำกับไฟล์ก็ไม่สามารถนำไปใช้ประโยชน์ได้อย่างสูงสุด
- เว็บเซิร์ฟเวอร์ยังไม่ได้ให้บริการส่วนที่รองรับภาษาไทยได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6.3 แนวทางการพัฒนา

- สามารถที่จะพัฒนาในส่วนของภาษาไทยซึ่งในปัจจุบันมีมาตรฐานรองรับอยู่แล้ว แต่จะมีความยุ่งยากมากขึ้นในเรื่องการสร้างระบบค้นหา เพราะ โครงสร้างของภาษานั้นแตกต่างกับภาษาอังกฤษอยู่มาก

- ในเรื่องการรับข้อมูลที่จะนำไปใช้สร้างความหมายให้แก่ไฟล์ข้อมูลจากผู้ให้บริการทั่วไปนั้น วิธีแก้หนึ่งคือ ผู้ดูแลระบบต้องคอยดูแลติดตามและแก้ไขด้วยตัวเอง ซึ่งไม่ใช่วิธีแก้ปัญหาที่ดี อีกวิธีที่น่าจะให้ผลที่ดีขึ้น คือ ในการรับข้อมูลจากผู้ทำการจัดเก็บไฟล์ข้อมูล จะออกแบบให้มีข้อบังคับหรือรูปแบบที่ชัดเจนมากขึ้นกว่าเดิม ทำให้ผู้ใช้สามารถป้อนข้อมูลได้กระชับและถูกต้องยิ่งขึ้น เพื่อให้ข้อมูลที่ได้รับมาสื่อความหมายมากของไฟล์ยิ่งขึ้น



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

ภาคผนวก ก

ไฟล์ข้อมูลที่สามารถแสดงผลทางหน้าเว็บ

ก.1 ไฟล์ที่เล่นได้บนควิกไทม์ 6

รูปแบบของไฟล์นำเข้า (Import file formats)

- 3DMF
- AIFF
- AU
- Audio CD Data(MACINTOSH)
- AVI
- BMP
- Cubic VR
- DLS
- DV
- FlashPix
- FLC
- GIF
- JPEG/JFIF
- Karaoke
- MacPaint
- Macromedia Flash5
- MIDI
- MPEG-1
- MP3
- M3U
- MPEG-2
- MPEG-4(MP4)
- Photoshop
- PICS
- PICT
- PLS
- PNG
- QuickTime 6 Image File

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- QuickTime 6 Movie
- SF2
- SGI
- Sound
- TARGA
- Text
- TIFF
- TIFF Fax
- Virtual Reality(VR)
- Wave

รูปแบบส่งออก (Export formats)

- AIFF
- AU
- AVI
- BMP
- DV Stream
- FLC
- Image Sequence movie exporters
- JPEG/JFIF
- JPEG 2000
- MacPaint
- MIDI
- MPEG-4
- Photoshop
- PICT
- PNG
- QuickTime 6 Image
- QuickTime 6 Movie
- SGI
- System 7 Sound
- Targa
- Text

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- TIFF
- WAV

ก.2 ไฟล์ที่สามารถเล่นบนวินโดวส์มีเดียเพลเยอร์

File type (format)	File name extension
CD audio	.cda
Intel Indeo video technology	.ivf
Audio Interchange File Format (AIFF)	.aif, .aifc, and .aiff
Windows Media audio and video files	.asf, .asx, .wax, .wm, .wma, .wmd, .wmv, .wvx, .wmp, and .wmx
Windows audio and video files	.avi and .wav
Windows Media Player skins	.wmz and .wms
Moving Picture Experts Group (MPEG)	.mpeg, .mpg, .m1v, .mp2, .mpa, .mpe, .mp2v*, and .mpv2
Musical Instrument Digital Interface (MIDI)	.mid, .midi, and .rmi
AU (UNIX)	.au and .snd
MP3	.mp3 and .m3u
DVD video	.vob

ตาราง ๔-1 ชนิดของไฟล์ที่ วินโดวส์ มีเดีย เพลย์เยอร์ เล่นได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

เวกเตอร์โมเดลสำหรับหาความคล้ายคลึงกันของเอกสาร

ข.1. บทนำ

เวกเตอร์โมเดลเป็นหนึ่งในโมเดลวิชาการอินฟอร์เมชัน รีทรีฟวอล (Information Retrieval) ซึ่งเป็นโมเดลที่ช่วยในการกำหนดค่าน้ำหนัก (weight) ให้แก่นิเค็ช์เทอม(index term)ต่างๆในเวกเตอร์หนึ่ง เพื่อที่จะนำค่าน้ำหนักนี้ไปคำนวณและเปรียบเทียบกับอีกเวกเตอร์หนึ่ง เพื่อพิจารณาว่ามีความคล้ายกันหรือตรงประเด็นมากแค่ไหน ซึ่งจะมีประโยชน์ในการพัฒนาการค้นหาข้อมูลในฐานข้อมูลและนำมาแสดงผล

ข.2. สัญลักษณ์

- เพื่อให้่ายต่อการเข้าใจและเข้าใจตรงกัน กำหนด
- D เป็นเซตของเอกสาร ในฐานข้อมูล
 - Q เป็นเซตของเทอมที่ผู้ใช้ query
 - $\vec{q}$ $\vec{q} = (w_{1,q}, w_{2,q}, \dots, w_{i,q})$ เป็นเวกเตอร์บอกค่าน้ำหนักของแต่ละเทอมของ query โดย i คือจำนวนทั้งหมดของเทอมใน query
 - $\vec{d}$ $\vec{d} = (w_{1,j}, w_{2,j}, \dots, w_{i,j})$
 - k_i เป็นนิเค็ช์เทอมลำดับที่ i
 - d_j เป็นเอกสาร (Document) ลำดับที่ j
 - $w_{i,j}$ เป็นค่าน้ำหนักแสดงความสัมพันธ์ของคูเทอม k_i ในเอกสาร d_j
 - $q_{i,j}$ เป็นค่าน้ำหนักแสดงความสัมพันธ์ของคูเทอม k_i ใน q

ข.3. ทฤษฎีและการคำนวณ

จากค่าน้ำหนักของเทอมของ query q และค่าน้ำหนักของในแต่ละนิเค็ช์เทอมและในแต่ละเอกสาร d_j มีจำนวน i เทอม สามารถที่จะแสดงความสัมพันธ์หาค่าความคล้ายคลึงกัน(Similarity)มากที่สุด ระหว่าง $\vec{q}$ กับ $\vec{d}$ ได้ในรูปแบบของฟังก์ชันโคซายน์ ดังนี้

$$\begin{aligned} \text{sim}(d_j, q) &= \frac{\vec{d}_j \cdot \vec{q}}{|\vec{d}_j| \times |\vec{q}|} \\ &= \frac{\sum_{i=1}^I w_{i,j} \times w_{i,q}}{\sqrt{\sum_{i=1}^I w_{i,j}^2} \times \sqrt{\sum_{i=1}^I w_{i,q}^2}} \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ $|\vec{d}_j|, |\vec{q}|$ เป็นขนาดของเวกเตอร์ของเอกสารและเซตของ query ของผู้ใช้ตามลำดับ โดยค่า $w_{i,j}$ และ $w_{i,q}$ มีค่ามากกว่า 0 และ $0 \leq sim(d_j, q) \leq 1$ เราสามารถจัดอันดับของเอกสารโดยใช้ค่าของความคล้ายคลึงของเอกสารได้ โดยค่าที่ใกล้เคียงกับ 1 มากที่สุด จะหมายความว่า ในเอกสารนั้นมีเทอมที่ตรงกับเทอมที่ใช้ query มากที่สุด ในทางตรงกันข้ามถ้าเอกสารใดให้ค่าความคล้ายคลึงกับเทอม query เท่ากับ ศูนย์ หรือเข้าใกล้ศูนย์ แสดงว่าเอกสารนั้นไม่มีความคล้ายคลึงเลย หรือมีน้อยตามลำดับ

จากความสัมพันธ์การหาค่าความคล้ายคลึงกัน ปัจจัยที่เกี่ยวข้องโดยตรงในความสัมพันธ์นั้นคือ น้ำหนักของในแต่ละเทอมในแต่ละเอกสาร ดังนั้นปัญหาต่อไปคือทำอย่างไรจึงจะสามารถกำหนดค่าน้ำหนักของแต่ละเทอมได้อย่างเหมาะสมและมีประสิทธิภาพ ในเวกเตอร์โมเดลเราสามารถหาค่าความถี่ของจำนวนครั้งที่เทอมนั้นๆปรากฏ เป็นแฟกเตอร์หลักในการกำหนดค่าน้ำหนักของเทอม อีกแฟกเตอร์หนึ่งเป็นค่า ไอดีเอฟ หรือ อินเวอร์ส ค็อกคัมเมนซ์ ฟรีควเินซี (idf หรือ Inverse Document Frequency) ที่เป็นค่าที่บ่งบอกความสำคัญของเทอมนั้นเมื่อเปรียบเทียบกับค่าอื่นๆ โดยคำนวณจากจำนวนเอกสารที่มีเทอมปรากฏอยู่ ถ้าเทอมปรากฏอยู่ในเอกสารมากมายเทียบกับเอกสารทั้งหมดในฟังก์ชันลอการิทึมแล้ว ค่าไอดีเอฟจะให้ค่าออกมาน้อย และถ้าเทอมปรากฏอยู่ในเอกสารน้อยแล้ว ค่าไอดีเอฟจะให้ค่าออกมามาก ซึ่งแสดงว่าเทอมที่ปรากฏอยู่ในเอกสารจำนวนน้อยๆ จะให้ความสำคัญมากกว่าเทอมที่ปรากฏอยู่ในเอกสารจำนวนมาก ดังสูตร

$$idf_i = \log \frac{N}{n_i}$$

ที่ซึ่ง idf_i เท่ากับค่าไอดีเอฟของอินเด็กซ์เทอมที่ i ในทุกๆเอกสาร และกำหนด $freq_{i,j}$ เป็นความถี่คือจำนวนครั้งที่ปรากฏของเทอมที่ i ในเอกสารที่ j และ $\max_i freq_{i,j}$ คือค่าความถี่ของเทอมในเอกสารที่ j ที่มากที่สุด แล้วจะได้ค่าความถี่ $f_{i,j}$ (normalized frequency) ดังนี้

$$f_{i,j} = \frac{freq_{i,j}}{\max_i freq_{i,j}}$$

แล้วจะได้ค่าของน้ำหนักของเทอม ดังสูตร

$$w_{i,j} = f_{i,j} \times idf_i$$

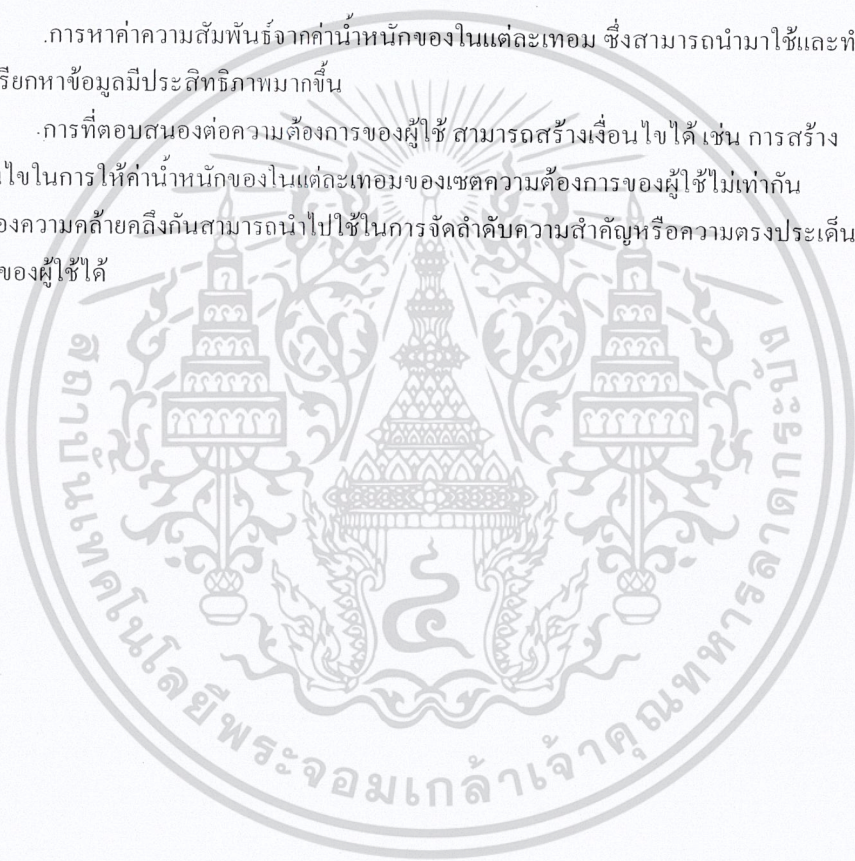
นอกจากนี้จากทฤษฎีของซอลตันและบัคเลย์ (Salton and Buckley) ได้เสนอแนะ ความสัมพันธ์หาค่าน้ำหนักของเทอมความต้องการของผู้ใช้ $w_{i,q}$ ไว้ว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$w_{i,q} = \left(0.5 + \frac{0.5 \, freq_{i,q}}{\max_l \, freq_{l,q}} \right) \times \log \frac{N}{n_i}$$

โดย $w_{i,q}$ คือค่าน้ำหนักของเทอมที่ i ในเซตของเทอมความต้องการของผู้ใช้ และ $freq_{i,q}$ คือค่าความถี่ที่วัดจากจำนวนครั้งที่เทอมปรากฏอยู่ในเซตของ query ของผู้ใช้ และ $\max_l \, freq_{l,q}$ คือค่าความถี่ของเทอมที่ปรากฏมากที่สุดในเซตของ query ของผู้ใช้

- จากเวกเตอร์โมเดล ได้ข้อดีของโมเดลนี้คือ
- การหาค่าความสัมพันธ์จากค่าน้ำหนักของในแต่ละเทอม ซึ่งสามารถนำมาใช้และทำให้การเรียกหาข้อมูลมีประสิทธิภาพมากขึ้น
 - การที่ตอบสนองต่อความต้องการของผู้ใช้ สามารถสร้างเงื่อนไขได้ เช่น การสร้างเงื่อนไขในการให้ค่าน้ำหนักของในแต่ละเทอมของเซตความต้องการของผู้ใช้ไม่เท่ากัน
 - ค่าของความคล้ายคลึงกันสามารถนำไปใช้ในการจัดลำดับความสำคัญหรือความตรงประเด็นกับความต้องการของผู้ใช้ได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- [1] Deitel & Deitel: "JAVA HOW TO PROGRAM THIRD EDITION", PRENTICE HALL, 2002
- [2] Marty Hall: "Core Servlets and Java Server Pages", Prentice Hall PTR, 2000
- [3] Ricardo Baeza-Yates & Berthier Ribeiro-Neto: "Modern Information Retrieval", Addison-Wesley, 1999, pp 27-30
- [4] Steven Holzner: "INSIDE XML", New Riders, 2001
- [5] "XML Extender Administration and Programming version",
<http://www.ibm.com/xsoftware/data/db2/library>
- [6] ดร. วีระศักดิ์ ชิงदार: "Fundamental of Java Programming Volume 1", บริษัท ซีเอ็ดดูเคชั่น จำกัด (มหาชน), 2543
- [7] ดร. วีระศักดิ์ ชิงदार: "Fundamental of Java Programming Volume 2", บริษัท ซีเอ็ดดูเคชั่น จำกัด (มหาชน), 2543
- [8] สุภชัย สมพานิช: "เข้าใจและใช้งาน XML ฉบับโปรแกรมเมอร์", นนทบุรี, อินโฟเพรส, 2544

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้