



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

สืบเนื่องจากผลการวิจัยเรื่องการทำหนดรูปแบบบรรณานุกรมและการวางแนวทางการจัดเก็บแบบถาวรสำหรับข้อมูลต่อเนื่องที่ตัดมาเป็นกลุ่มให้สืบค้นผ่านห้องสมุดอัตโนมัติ [1] ได้กล่าวถึงการออกแบบและสร้างหน้าจอเป็นแบบเว็บ (Web) สำหรับการสรุปข้อมูลในลักษณะกราฟ และสามารถให้ผู้ใช้งานสืบค้นตามช่วงเวลาที่ต้องการได้จากระบบของห้องสมุดนั้น ได้มีการพัฒนาระบบเชื่อมต่อผ่านเว็บเข้าสู่ระบบอินโนแพค (INNOPAC) โดยการใช้สคริปต์ (Script) ในการติดต่อสนทนากับระบบอินโนแพค ที่เป็นระบบโอแพค (OPAC:On-Line Public Access Catalog) เสมือนเป็นการเรียกใช้จากผู้ใช้ปกติ

วิธีการดังกล่าวเป็นการแก้ปัญหาของการไม่ต้องแก้ไขในส่วนของโปรแกรมอินโนแพค ที่เป็นโปรแกรมสำเร็จรูปที่มีราคาสูง หากทำการแก้ไข ปรับเปลี่ยนหรือเพิ่มเติมโดยตรงเข้าไปที่ตัวโปรแกรมจะกระทำได้อย่างลำบาก หรือต้องเสียค่าใช้จ่ายสูง จากระบบเชื่อมต่องดกล่าวพบว่า สคริปต์มีการติดต่อกับระบบอินโนแพค เพื่อสืบค้นและดึงข้อมูลมาจัดเก็บในรูปแบบของเอกสารข้อความ (Text File) ซึ่งสามารถจัดเก็บข้อมูลได้เพียงรูปแบบเดียว จากการเขียนสคริปต์ไว้สำหรับตัดทอนข้อมูล ยังผลทำให้เกิดความยุ่งยาก หากต้องการดึงข้อมูลอื่น ๆ มาจัดเก็บที่มีรูปแบบต่างไปจากเดิมที่ระบุไว้ในสคริปต์ ซึ่งต้องเขียนหรือแก้ไขสคริปต์ใหม่เพิ่มเติมเมื่อต้องการสืบค้น และดึงข้อมูลมาเพื่อทำการจัดเก็บในรูปแบบอื่น อันเป็นความไม่สะดวก ซึ่งผู้ที่ทำการแก้ไขต้องมีความรู้เรื่องของการเขียนสคริปต์ถึงจะสามารถทำการแก้ไขเพิ่มเติมได้

ด้านส่วนของการแสดงข้อมูลในลักษณะกราฟผ่านทางหน้าจอบริบท เพื่อป้องกันคุณภาพของข้อมูล Traffic Log (Traffic Log Data Quality) ที่มีการจัดเก็บลงบนสื่อบันทึกถาวรซีดีรอม (CD-ROM) ข้อมูลที่จัดเก็บมีลักษณะเป็นแบบอนุกรมเวลาที่ต่อเนื่อง ซึ่งมีปริมาณของข้อมูลมาก และเพื่อการสืบค้นข้อมูลที่มีจำนวนมากจึงได้มีการนำมาจัดเก็บบนระเบียบบรรณานุกรมของระบบอินโนแพคอีกครั้ง โดยในงานวิจัยนี้คุณภาพของข้อมูล หมายถึง สัดส่วนปริมาณของข้อมูลในช่วงเวลาที่กำหนด อันแสดงถึงปริมาณของข้อมูล Traffic Log ที่มีการบันทึกไว้มีปริมาณมากน้อยเพียงไร เช่น ถ้าพิจารณาช่วงเวลาเป็นปีที่มี 365 วัน และมีข้อมูลเก็บสะสมไว้ 100 วัน คุณภาพของข้อมูลที่ได้ คือ 100/365 หรือ ประมาณ 27.40% การแสดงสัดส่วนปริมาณข้อมูลนั้น จะแสดงออกในเชิงกราฟอันเกี่ยวเนื่องกับเรื่องของเวลาที่ง่ายต่อการพิจารณา และสามารถนำค่าที่ได้นี้ไปใช้ในการวิเคราะห์ทางสถิติต่อไปได้

1.2 วัตถุประสงค์ของการวิจัย

เพื่อการวิเคราะห์ข้อมูล Traffic Log ที่จัดเก็บบนระเบียบนบรรณานุกรมห้องสมุดอัตโนมัติของจุฬาลงกรณ์มหาวิทยาลัย โดยแสดงออกในเชิงกราฟ ที่ง่ายต่อการพิจารณาถึงคุณภาพของข้อมูล

1.3. ขอบเขตของการวิจัย

ส่วนของการเตรียมข้อมูล

1.3.1. ออกแบบ และพัฒนาโปรแกรมสำหรับสนทนากับระบบอินโนแพค

1.3.2. สามารถกำหนดขอบเขตของข้อมูลในการตัดทอนข้อมูลระเบียบนบรรณานุกรมได้หลายรูปแบบโดยไม่ต้องเขียนสคริปต์

1.3.3. ทำการสืบค้นข้อมูลบนระบบโอแพค เพื่อดึงข้อมูลระเบียบนบรรณานุกรม มาทำการตัดทอนข้อมูล (Extract Data) ตามที่ได้มีการกำหนดขอบเขตของข้อมูลไว้

1.3.4. แสดงตัวอย่างของข้อมูลที่ได้จากการตัดทอนข้อมูลในรูปของตารางก่อนทำการจัดเก็บลงฐานข้อมูลเอกซเอ็มแอล

1.3.5. ทำการจัดเก็บข้อมูลลงฐานข้อมูลเอกซเอ็มแอล โดยใช้โปรแกรมสลีพแคท (Sleepycat Software)

1.3.6. โปรแกรมสามารถประมาณผลเพื่อเก็บรวบรวมข้อมูลจากระบบอินโนแพค ได้ 2 ลักษณะ คือ ประมวลผลตามเวลาที่กำหนด และประมวลผลโดยผู้ดูแลระบบเอง

1.3.7. ภาษาที่ใช้ในการพัฒนาโปรแกรมเก็บรวบรวมข้อมูล คือวิซวลซีพลัสพลัส (Visual C++) version 6.0

1.3.8. จัดทำคู่มือการใช้งานโปรแกรมสำหรับการสนทนากับระบบอินโนแพค และการกำหนดขอบเขตของข้อมูลระเบียบนบรรณานุกรมสำหรับใช้ในการตัดทอนข้อมูล

ส่วนของการแสดงผลในเชิงกราฟ

1.3.9. คุณภาพของข้อมูลที่นำมาแสดง คือ สัดส่วนปริมาณของข้อมูล Traffic Log ที่จัดเก็บบนระเบียบนบรรณานุกรมห้องสมุดอัตโนมัติของจุฬาลงกรณ์มหาวิทยาลัยในช่วงเวลาต่าง ๆ ที่กำหนดคือ ปี เดือน สัปดาห์ วัน และช่วงระยะเวลาชั่วโมงใน 1 วัน โดยนำเสนอออกมาในรูปแบบของแผนภูมิ

1.3.10. คุณภาพของข้อมูลที่นำมาแสดงจะใช้สัญลักษณ์สีในการบ่งบอกในแต่ละช่วงเวลาที่กำหนด เช่น

สีเขียว แสดงถึง มีข้อมูล Traffic Log

สีแดง แสดงถึง ไม่มีข้อมูล Traffic Log

1.3.11. ให้บริการผ่านทาง Web-base service โดยมีรายละเอียดดังนี้

1.3.11.1. ระบบปฏิบัติการ (Operating System) ที่ใช้คือ Windows 2000 Professional Service Package 4

1.3.11.2. Web Server ที่ให้บริการคือ Apache Tomcat 4.1

1.3.11.3. Web Browser ที่สามารถใช้งานได้คือ Internet Explorer v6.0, Netscape Navigator v8.1, Firefox 1.5

1.3.12. ใช้ภาษาจาวา (Java) ในการพัฒนาเป็นเชิงกราฟิก

1.4. ขั้นตอนการดำเนินงาน

1.4.1. ศึกษางานวิจัย [1] และ [2]

1.4.2. ศึกษาข้อมูลบรรณานุกรมบนระบบอินโนแพค

1.4.3. ศึกษาโปรแกรมการตัดทอนข้อมูลบนเอกสารข้อความจากเว็บไซต์ (Web Site)

ต่าง ๆ

1.4.4. ออกแบบโปรแกรมคอลเลคเตอร์ XML Database และเว็บเพจ (Web page)

1.4.5. พัฒนาโปรแกรมคอลเลคเตอร์ และเว็บเพจ

1.4.6. ทดสอบโปรแกรมคอลเลคเตอร์ และเว็บเพจกับระบบอินโนแพค

1.4.7. นำระบบขึ้นใช้งานกับระบบอินโนแพค

1.4.8. จัดทำเอกสารคู่มือการใช้งาน

1.5. ประโยชน์ที่คาดว่าจะได้รับ

1.5.1. เพื่อช่วยให้บรรณารักษ์สามารถตรวจสอบข้อมูล Traffic Log

1.5.2. เพื่อลดค่าใช้จ่ายในการพัฒนาระบบการสืบค้นข้อมูลอื่น ๆ บนระบบอินโนแพค

1.5.3. สามารถนำโปรแกรมคอลเลคเตอร์ไปพัฒนาต่อเพื่อช่วยวิเคราะห์การใช้งานการสืบค้นข้อมูลผ่านห้องสมุดอัตโนมัติในปัจจุบัน เช่น จำนวนครั้งการสืบค้นที่เกิดขึ้นต่อวัน หรือต่อสัปดาห์ ว่าเพียงพอต่อการให้บริการหรือไม่

1.5.5. เพื่อเป็นแนวทางที่จะพัฒนาระบบการสืบค้นห้องสมุดอัตโนมัติในรูปแบบที่เป็นกราฟิกต่อไปได้