



บทที่ 5

การพัฒนาระบบ

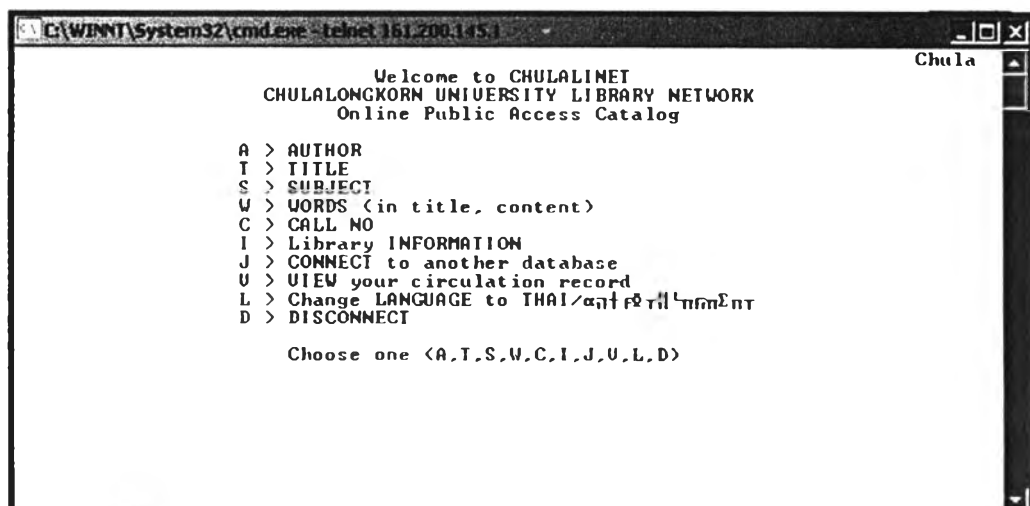
การพัฒนาระบบจะกล่าวแยกออกเป็น 2 ส่วนได้แก่ส่วนของการจัดเก็บและรวบรวมข้อมูล Traffic Log ด้วยโปรแกรมคอลเลคเตอร์ และส่วนของการแสดงภาพในเชิงกราฟิกที่บ่งบอกถึงปริมาณคุณภาพของข้อมูลที่ทำกรจัดเก็บตามแต่ละช่วงเวลา ซึ่งมีรายละเอียดดังนี้

5.1 การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลในส่วนนี้ได้มีการพัฒนาโปรแกรมคอลเลคเตอร์เพื่อรวบรวมและจัดเก็บข้อมูล สามารถแบ่งการทำงานออกเป็น 3 ส่วน ได้แก่ การกำหนดการทำงานของตัวโปรแกรม (Configuration) การประมวลผล (Execute Task) และการจัดเก็บลงฐานข้อมูลเอกซเอ็มแอล

1. การกำหนดการทำงานของตัวโปรแกรม

1.1 การกำหนดลำดับการสนทนากับระบบอินโนแพค (Configuration Chat Script) เป็นการกำหนดส่วนที่ใช้ในการติดต่อกันระหว่างโปรแกรมคอลเลคเตอร์ กับระบบโอแพคของอินโนแพคด้วยการ Telnet ซึ่งสคริปต์ที่ใช้ในการ Telnet จะเป็นรูปแบบคำสั่งที่เหมือนกับหน้าจอสืบค้นที่ใช้สำหรับสืบค้นข้อมูลบรรณานุกรมของห้องสมุดจุฬาลงกรณ์มหาวิทยาลัยดังภาพ



รูปที่ 5.1 แสดงหน้าจอสืบค้นข้อมูลของห้องสมุดจุฬาลงกรณ์มหาวิทยาลัย

สคริปต์การสนทนา (Chat Script) ที่ได้ทำการกำหนดไว้จะถูกบันทึกไว้ในรูปของเอกสารข้อความที่มีนามสกุลของแฟ้มข้อมูลเป็น ".cht" หลังจากบันทึกแฟ้มข้อมูลนี้ สามารถนำมาแก้ไขภายหลังได้ด้วยการแก้ไขผ่านทางโปรแกรมคอลเลคเตอร์ โดยไม่จำเป็นต้องทำการกำหนดสคริปต์การสนทนาใหม่อีกครั้ง

1.2 การกำหนดเขตข้อมูลที่ต้องการดึงจากระบบอินโนแพค (Configuration Selected Fields) เป็นการกำหนด เพื่อเลือกเขตข้อมูล (Field) บนหน้าจอโปรแกรมคอลเลคเตอร์ สามารถทำการเลือกได้หลายเขตข้อมูล และทำการจัดเก็บเขตข้อมูลที่ได้จากการกำหนดอยู่ในรูปของเอกสารเอกซเอ็มแอล โดยรายละเอียดผู้ใช้งานไม่จำเป็นต้องทราบความหมายภายในเอกสาร เนื่องจากเป็นการใช้งานในส่วนของโปรแกรม ตัวอย่างของเอกสารเอกซเอ็มแอล สำหรับใช้ในการเลือกข้อมูลของ Traffic log บนระเบียบอนุกรมห้องสมุดเป็นดังรูปที่ 5.2

```

<XML>
  <CDNO>@P0@</CDNO>
  <TITLE>@P1@</TITLE>
  <LOCATION>@P2@</LOCATION>
  <PUBLISH>@P3@</PUBLISH>
  <EXPECTEDEXPIRY>@P4@</EXPECTEDEXPIRY>
  <CONTENT>
    <FILE>
      <BIBRECORD>@P5@</BIBRECORD>
      <FILESIZE>@P6@</FILESIZE>
      <LOGFILE PERIOD>
        <BEGIN_PERIOD1>
          <DATE>@P7@</DATE>
          <MONTH>@P8@</MONTH>
          <YEAR>@P9@</YEAR>
          <HOUR>@P10@</HOUR>
          <MIN>@P11@</MIN>
        </BEGIN_PERIOD1>
        <END_PERIOD1>
          <DATE>@P12@</DATE>
          <MONTH>@P13@</MONTH>
          <YEAR>@P14@</YEAR>
          <HOUR>@P15@</HOUR>
          <MIN>@P16@</MIN>
        </END_PERIOD1>
        <BEGIN_PERIOD2>
          <DATE>@P17@</DATE>
          <MONTH>@P18@</MONTH>
          <YEAR>@P19@</YEAR>
          <HOUR>@P20@</HOUR>
          <MIN>@P21@</MIN>
        </BEGIN_PERIOD2>
        <END_PERIOD2>
          <DATE>@P22@</DATE>
          <MONTH>@P23@</MONTH>
          <YEAR>@P24@</YEAR>
          <HOUR>@P25@</HOUR>
          <MIN>@P26@</MIN>
        </END_PERIOD2>
      </LOGFILE PERIOD>
      <FILENAME>@P27@</FILENAME>
    </FILE>
  </CONTENT>
</XML>

```

รูปที่ 5.2 แสดงเอกสารเอกซเอ็มแอลจากการกำหนดเขตข้อมูลที่ต้องการดึงจากระบบอินโนแพด

ความหมายของแต่ละส่วนย่อย (Element) ที่ทำการเก็บข้อมูลในเอกสารเอกซเอ็ม-
แวลมีดังนี้

<CDNO>	เก็บข้อมูล หมายเลขซีดี
<TITLE>	เก็บข้อมูล ชื่อที่ทำการสืบค้น (Internet Traffic Log)
<LOCATION>	เก็บข้อมูลชื่อสถานที่
<PUBLISH>	เก็บข้อมูล วันที่ที่ทำการบันทึกซีดี
<EXPECTEDEXPIRY>	เก็บข้อมูล วันที่คาดว่าจะหมดอายุการใช้งานของซีดีรวม
<BIBRECORD>	เก็บข้อมูล หมายเลขระเบียบบรรณานุกรมบนระบบ อินโนแพค
<FILESIZE>	เก็บข้อมูล ขนาดของแฟ้มข้อมูล (MB)
<BEGIN_PERIOD1>	เก็บข้อมูล วันที่และเวลาที่เริ่มบันทึกข้อมูลในช่วงเวลา แรกบน 1 ระเบียบบรรณานุกรมของข้อมูล Traffic Log
<END_PERIOD1>	เก็บข้อมูล วันที่และเวลาที่สิ้นสุดบันทึกข้อมูลใน ช่วงเวลาแรกบน 1 ระเบียบบรรณานุกรมของข้อมูล Traffic Log
<BEGIN_PERIOD2>	เก็บข้อมูล วันที่และเวลาที่เริ่มบันทึกข้อมูลในช่วงเวลาที่ สองบน 1 ระเบียบบรรณานุกรมของข้อมูล Traffic Log
<END_PERIOD2>	เก็บข้อมูล วันที่และเวลาที่เริ่มบันทึกข้อมูลในช่วงเวลาที่ สองบน 1 ระเบียบบรรณานุกรมของข้อมูล Traffic Log
<DATE>	เก็บข้อมูลวัน
<MONTH>	เก็บข้อมูลเดือน
<YEAR>	เก็บข้อมูลปี
<HOUR>	เก็บข้อมูลชั่วโมง
<MIN>	เก็บข้อมูลนาที
<FILENAME>	เก็บข้อมูลชื่อแฟ้มข้อมูลที่อยู่ในซีดีรวม

2. การประมวลผล สามารถประมวลผลได้ 2 ลักษณะคือ แบบประมวลผลด้วยตัว
ผู้ใช้เองซึ่งในที่นี้คือบรรณารักษ์ และแบบอัตโนมัติ ตามที่ได้กำหนดใน Task Schedule ซึ่ง
สามารถกำหนดให้เป็นแบบประมวลผลเพียงครั้งเดียว หรือประมวลผลแบบประจำทุกเดือนตาม
วันที่ระบุ ซึ่งแล้วแต่ความสะดวกต่อการใช้งานของผู้ใช้ ข้อมูลที่ได้จากการประมวลผลจะถูกบันทึก

อยู่ในแฟ้มข้อมูลเอกซเซล (Excel File) ชื่อแฟ้มข้อมูลถูกกำหนดไว้เป็นชื่อเดียวกับ Task Name ข้อดีของการบันทึกแฟ้มข้อมูลเอกซเซล คือ ป้องกันการสูญหายของข้อมูลหากโปรแกรมคอลเลคเตอร์ไม่สามารถทำงานต่อได้หลังจากที่มีการประมวลผลข้อมูลเสร็จแล้ว และแฟ้มข้อมูลสามารถรองรับข้อมูลได้หากงาน (Task) มีปริมาณข้อมูลมาก

3. การจัดเก็บลงฐานข้อมูลเอกซเอ็มแอล ข้อมูลที่ได้จากการประมวลผลซึ่งถูกจัดเก็บในแฟ้มข้อมูลเอกซเซล จะถูกตรวจสอบ และ เปลี่ยนแปลงแก้ไขข้อมูลที่ผิดพลาดอีกครั้งโดยบรรณารักษ์ ก่อนที่มีการจัดเก็บข้อมูลลงฐานข้อมูลเอกซเอ็มแอล การจัดเก็บลงฐานข้อมูลเอกซเอ็มแอลจะจัดเก็บลงตู้คอนเทนเนอร์ (Container) ชื่อ บิบเรคคอร์ด ดีบี (BibRecord_DB.dbxml) รายละเอียดการใช้งานฐานข้อมูลเอกซเอ็มแอลสามารถดูได้จากภาคผนวก ข้อมูลที่จัดเก็บนี้จะถูกนำไปใช้ในการประมวลผลในส่วนของการแสดงคุณภาพของข้อมูลในเชิงกราฟิกผ่านทางเว็บเบราว์เซอร์ต่อไป

5.2 การแสดงข้อมูล

จากข้อจำกัดของระบบอินโนแพคในเรื่องของการแสดงข้อมูลออกทางหน้าจอคอมพิวเตอร์สามารถแสดงได้เฉพาะรูปแบบที่เป็นข้อความได้เท่านั้น จึงขอเสนอรูปแบบการแสดงผลข้อมูลบนระบบอินโนแพค ใหม่ในเชิงกราฟิกที่ง่ายต่อการใช้งานแทนรูปแบบเดิม ซึ่งในที่นี้จะทำการแสดงถึงคุณภาพของข้อมูล โดยมีนิยาม และเหตุผลของการนิยามดังนี้

นิยามของคุณภาพของข้อมูล คือ สัดส่วนปริมาณของข้อมูล Traffic Log ในช่วงเวลาที่กำหนด ที่มีการจัดเก็บข้อมูลเป็นในลักษณะของอนุกรมเวลาลงบนสื่อถาวรซีดี-รอม (CD-ROM)

เหตุผลของการนิยาม คือ เนื่องจากข้อมูลเป็นข้อมูลเพื่อนำไปใช้ในทางสถิติต่าง ๆ เกี่ยวเนื่องกับเรื่องเวลา ดังนั้นถ้ามีการเก็บข้อมูลที่ไม่ครบในช่วงเวลาที่ต้องการหาค่าสถิติ สามารถกล่าวได้ว่าคุณภาพของข้อมูลนั้นไม่เหมาะสมในการคำนวณค่าสถิติ หรือเรียกว่ามีประสิทธิภาพไม่เต็ม 100% ซึ่งสัดส่วนของคุณภาพของข้อมูลขึ้นอยู่กับหน่วยนับของเวลานั้น โดยมีฐานอยู่ที่จำนวนวัน และชั่วโมงตามตารางดังนี้

ตารางที่ 5.1 แสดงฐานการคิดคุณภาพของข้อมูล Traffic Log

ช่วงเวลา	หน่วยนับ	ฐานของสัดส่วน
ปี (Year)	วัน	365 หรือ 366 วัน
เดือน (Month)	วัน	28 – 31 วัน
สัปดาห์ (Week)	วัน	7 วัน
วัน (Day)	ชั่วโมง	24 ชั่วโมง
0.00 น. - 12.00 น. (AM)	นาที	720 นาที
12.00 น. - 24.00 น. (PM)	นาที	720 นาที

การแสดงผลจะแสดงผ่านทางโปรแกรมเว็บเบราว์เซอร์ ตามช่วงระยะเวลาที่ได้นำมาแสดง ได้แก่ ปี เดือน สัปดาห์ วัน และช่วงระยะเวลาชั่วโมงใน 1 วัน เป็นในลักษณะของการลงลึก (Drill Down)

การประมวลผลเพื่อแสดงถึงคุณภาพของข้อมูลในแต่ละช่วงระยะเวลาจะใช้โปรแกรมที่ชื่อว่าเจนเนอเรตชาร์ต (GenerateChart) ในการประมวลผลออกมาเป็นแฟ้มข้อมูลรูปภาพที่เป็น JPG และจัดเก็บอยู่ในโฟลเดอร์อิมเมจ (Images Folder) ที่พร้อมสำหรับถูกเรียกใช้งานโดยเว็บเซิร์ฟเวอร์ ด้านการประมวลผลเพื่อให้ได้ข้อมูลที่ถูกต้องมากที่สุดจะทำการพิจารณาในระดับของหน่วยเวลาเป็นนาที เนื่องจากเวลาที่น้อยที่สุดที่มีการบันทึกไว้ของข้อมูล Traffic Log บนระเบียบบรรณานุกรมมีหน่วยเป็นนาที

สูตรที่ใช้ในการคำนวณของโปรแกรมมีดังนี้

$$\%LogTime = \text{TrafficLogTimeMin} * 100 / \text{TotalTimeMin}$$

%LogTime หมายถึง เปอร์เซนต์ของจำนวนเวลาที่ข้อมูล Traffic Log มีการบันทึกในช่วงระยะเวลาที่ทำการพิจารณาเทียบกับช่วงระยะเวลาที่ทำการพิจารณา

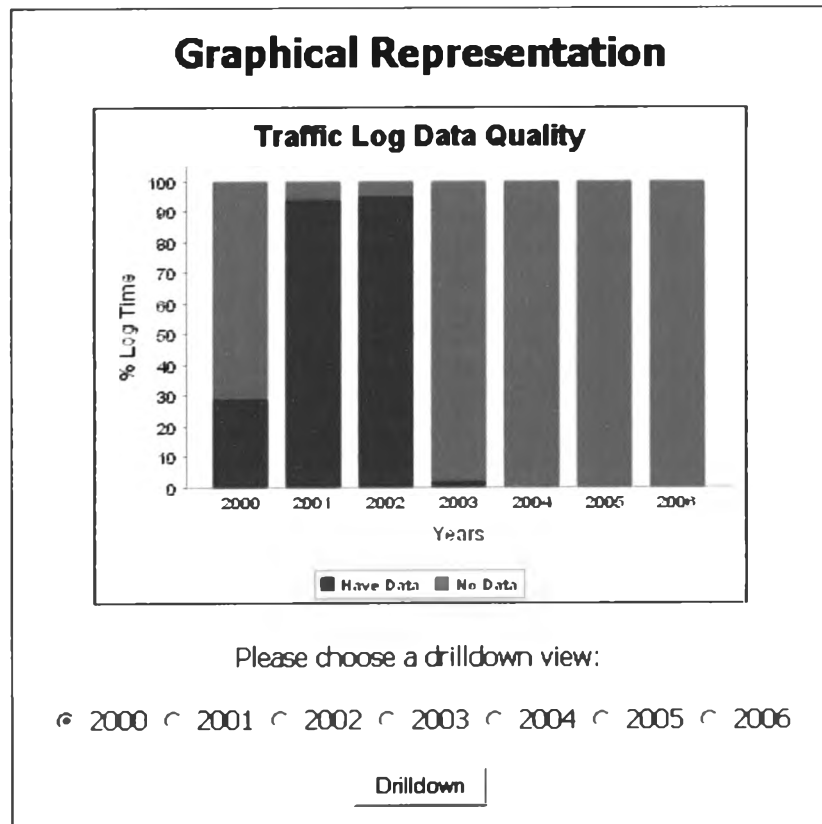
TrafficLogTimeMin หมายถึง จำนวนเวลาที่ข้อมูล Traffic Log มีการบันทึกในช่วง
ระยะเวลาที่ทำการพิจารณา มีหน่วยเป็นนาที

TotalTimeMin หมายถึง จำนวนเวลาทั้งหมดในช่วงระยะเวลาที่ทำการพิจารณา
มีหน่วยเป็นนาที เช่น พิจารณาเดือนกุมภาพันธ์ในปี 2000 จะมี
ค่าเท่ากับ 41,760 นาที (29 วันคูณกับจำนวนนาทีใน 1 วัน)

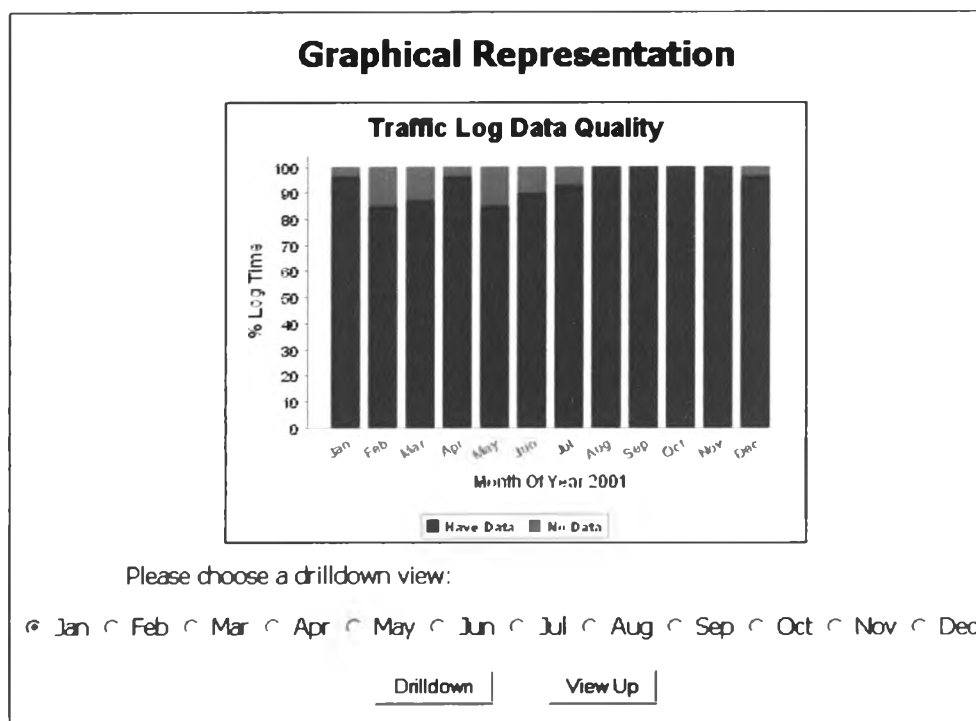
สัญลักษณ์ที่ใช้ในการแสดงถึงปริมาณคุณภาพของข้อมูลได้นิยามดังตารางที่ 5.1

ตารางที่ 5.2 แสดงสัญลักษณ์ของแผนภูมิแบ่งตามประเภทช่วงเวลา

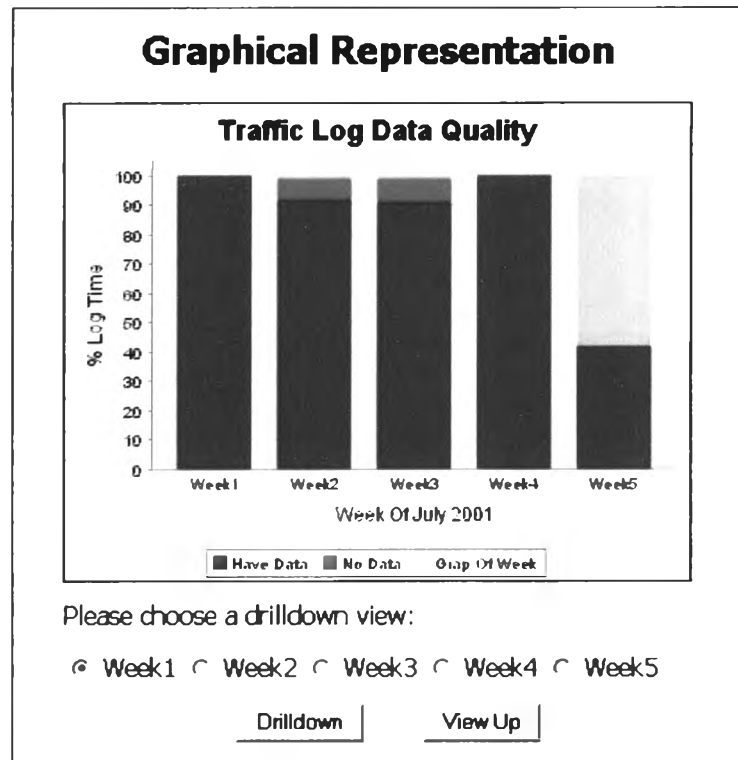
ช่วงเวลา	ประเภท Chart	สัญลักษณ์ที่ใช้แสดง
ปี (Year)	Stack bar chart	สีแดง แสดงถึงไม่มีข้อมูล Traffic Log สีน้ำเงิน แสดงถึงมีข้อมูล Traffic Log
เดือน (Month)	Stack bar chart	สีแดง แสดงถึงไม่มีข้อมูล Traffic Log สีน้ำเงิน แสดงถึงมีข้อมูล Traffic Log สีเทา แสดงถึง ช่วงเวลาที่เว้นว่างในแต่ละสัปดาห์
สัปดาห์ (Week)	Stack bar chart	สีแดง แสดงถึงไม่มีข้อมูล Traffic Log สีน้ำเงิน แสดงถึงมีข้อมูล Traffic Log
วัน (Day)	Stack bar chart	สีแดง แสดงถึงไม่มีข้อมูล Traffic Log สีน้ำเงิน แสดงถึงมีข้อมูล Traffic Log
0.00 น. - 12.00 น. (AM)	Pie chart	สีแดง แสดงถึงไม่มีข้อมูล Traffic Log สีน้ำเงิน แสดงถึงมีข้อมูล Traffic Log
12.00 น. - 24.00 น. (PM)	Pie chart	สีแดง แสดงถึงไม่มีข้อมูล Traffic Log สีน้ำเงิน แสดงถึงมีข้อมูล Traffic Log



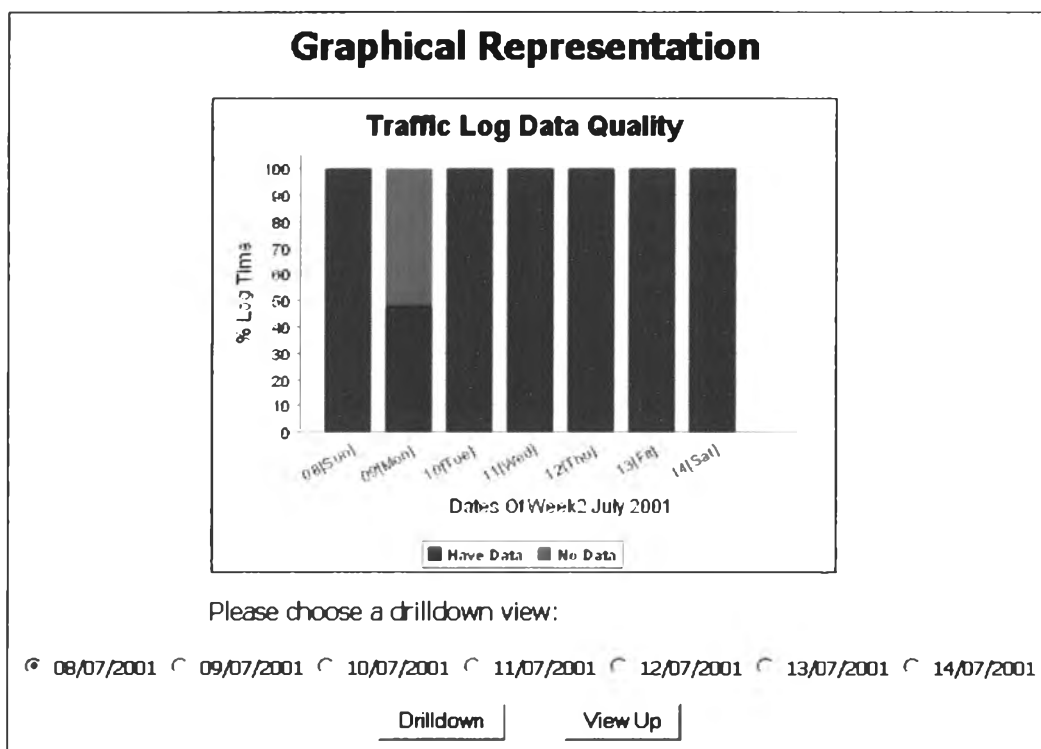
รูปที่ 5.3 แสดงคุณภาพของข้อมูล Traffic Log บน Web Browser ตามช่วงระยะเวลารายปี



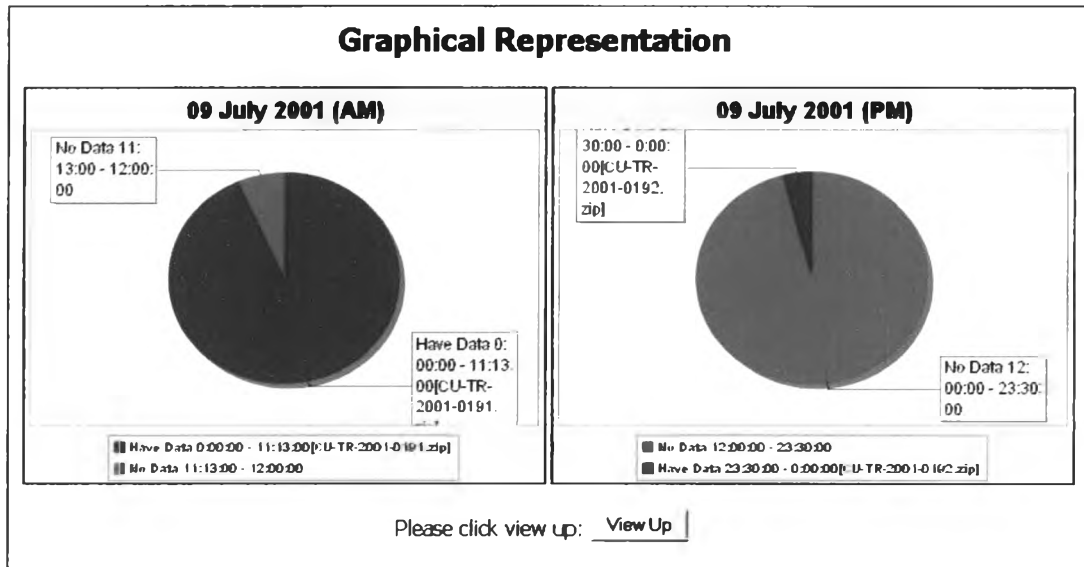
รูปที่ 5.4 แสดงคุณภาพของข้อมูล Traffic Log บน Web Browser ตามช่วงระยะเวลาแบบเดือน



รูปที่ 5.5 แสดงคุณภาพของข้อมูล Traffic Log บน Web Browser ตามช่วงระยะเวลาแบบสัปดาห์



รูปที่ 5.6 แสดงคุณภาพของข้อมูล Traffic Log บน Web Browser ตามช่วงระยะเวลาแบบวัน



รูปที่ 5.7 แสดงคุณภาพของข้อมูล Traffic Log บน Web Browser ตามช่วงระยะเวลาแบบชั่วโมง

หน้าจอบริการเว็บเบราว์เซอร์ได้สนับสนุนการสืบค้นข้อมูล Traffic Log ตามช่วงเวลา และตามหมายเลขของซีดี (CD No) แทนการสืบค้นผ่านหน้าจอของระบบอินโนแพค ที่ไม่สามารถสืบค้นตามช่วงเวลาและหมายเลขของซีดีได้ รายละเอียดที่ได้จากการสืบค้นได้แก่

1. CD No. คือ หมายเลขของซีดี-รวมที่จัดเก็บข้อมูล Traffic Log
2. File Name คือ ชื่อของ Log file ที่ถูกบีบอัดข้อมูลไว้ (Zip file)
3. Expire Date คือ วันที่หมดอายุของซีดี-รวม
4. Start Date – Time คือ วันที่และเวลาของการเริ่มบันทึก
5. End Date – Time คือ วันที่และเวลาของการสิ้นสุดบันทึก

สามารถแสดงตัวอย่างการสืบค้นข้อมูลดังรูปที่ 5.8

Search

From:Year Month Day — To:Year Month Day

CD No	File Name	Expire Date	Start Date-Time	End Date-Time
CD TR 2000-010	CJ-TR-2000-0031.zip	20051005	01/10/2000 23:59	03/10/2000 00:00
CD TR 2000-010	CJ-TR-2000-0032.zip	20051005	03/10/2000 00:00	04/10/2000 00:00
CD TR 2000-011	CJ-TR-2000-0033.zip	20051010	04/10/2000 00:00	05/10/2000 00:00
CD TR 2000-011	CJ-TR-2000.0034.zip	20051010	05/10/2000 00:00	06/10/2000 00:01

รูปที่ 5.8 แสดงข้อมูล Traffic Log ที่ได้จากการสืบค้นตามช่วงเวลาบน Web Browser

5.3 เครื่องมือที่ใช้ในการพัฒนา

1. VC++ Version 6.0 (Microsoft Visual Studio 6.0 Packet) ใช้สำหรับพัฒนาคอลเลคเตอร์โปรแกรม
2. Java v 1.4.2 ใช้สำหรับพัฒนาเซิร์ฟเลททางฝั่งเว็บเซิร์ฟเวอร์
3. Apache Tomcat Version 4.1 ใช้ทำหน้าที่ Web Server ที่ติดตั้งและทำงานภายใต้ระบบปฏิบัติการวินโดวส์
4. The JFreeChart Class Library Version 0.9.21 สำหรับใช้ในการสร้าง Chart ใช้พัฒนาร่วมกับภาษาจาวา
5. Berkeley DB XML Version 2.1.8 เป็น Database Application ที่ใช้สำหรับเก็บเอกสารเอกซเอ็มแอลที่มีส่วนต่อประสานโปรแกรมประยุกต์ (Application Program Interface) สนับสนุนทั้งภาษาซีพลัสพลัส และจาวา