บทที่ 2
ทฤษฎีพื้นฐานที่ใช้ในโครงงาน

ในหัวข้อนี้จะกล่าวถึงทฤษฎีพื้นฐานต่างๆ ที่เกี่ยวข้องในการวิจัยโดยมีรายละเอียดของเนื้อหาดังนี้

2.1 ระบบสารสนเทศทางภูมิศาสตร์(Geographic Information System :GIS)

ระบบสารสนเทศภูมิศาสตร์ หรือ Geographic Information System : GIS คือกระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์ ที่ใช้กำหนดข้อมูลและสารสนเทศ ที่มีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ เช่น ที่อยู่ บ้านเลขที่ สัมพันธ์กับตำแหน่งในแผนที่ ตำแหน่งเส้นรุ้ง เส้นแวง ข้อมูลและแผนที่ใน GIS เป็นระบบข้อมูลสารสนเทศที่อยู่ในรูปของตารางข้อมูล และฐานข้อมูลที่มีส่วนสัมพันธ์กับข้อมูลเชิงพื้นที่ (Spatial Data) ซึ่งรูปแบบและความสัมพันธ์ของข้อมูลเชิงพื้นที่ทั้งหลาย จะสามารถนำมาวิเคราะห์ด้วย GIS และทำให้สื่อความหมายในเรื่องการเปลี่ยนแปลงที่สัมพันธ์กับเวลาได้ เช่น การแพร่ขยายของโรคระบาด การเคลื่อนย้าย ถิ่นฐาน การบุกรุกทำลาย การเปลี่ยนแปลงของการใช้พื้นที่ ฯลฯ ข้อมูลเหล่านี้ เมื่อปรากฏบนแผนที่ทำให้สามารถแปลและสื่อความหมายใช้งานได้ง่าย

GIS เป็นระบบข้อมูลข่าวสารที่เก็บไว้ในคอมพิวเตอร์ แต่สามารถแปลความหมายเชื่อมโยงกับสภาพภูมิศาสตร์อื่นๆ สภาพท้องที่ สภาพการทำงานของระบบสัมพันธ์กับสัดส่วนระยะทางและพื้นที่จริงบนแผนที่ ข้อมูลใน GIS ทั้งข้อมูลเชิงพื้นที่และข้อมูลเชิงบรรยาย สามารถอ้างอิงถึงตำแหน่งที่มีอยู่จริงบนพื้นโลกได้โดยอาศัยระบบพิกัดทางภูมิศาสตร์ (Geocode) ซึ่งจะสามารถอ้างอิงได้ทั้งทางตรงและทางอ้อม ข้อมูลใน GIS ที่อ้างอิงกับพื้นผิวโลกโดยตรง หมายถึง ข้อมูลที่มีค่าพิกัดหรือมีตำแหน่งจริงบนพื้นโลกหรือในแผนที่ เช่น ตำแหน่งอาคาร ถนน ฯลฯ สำหรับข้อมูลGIS ที่จะอ้างอิงกับข้อมูลบนพื้นโลกได้โดยทางอ้อมได้แก่ ข้อมูลของบ้าน(รวมถึงบ้านเลขที่ ซอย เขต แขวง จังหวัด และรหัสไปรษณีย์) โดยจากข้อมูลที่อยู่ เราสามารถทราบได้ว่าบ้านหลังนี้มีตำแหน่งอยู่ ณ ที่ใดบนพื้นโลก เนื่องจากบ้านทุกหลังจะมีที่อยู่ไม่ซ้ำกัน ภาระหน้าที่หลักๆของระบบสารสนเทศภูมิศาสตร์ควรจะมีอยู่ด้วยกัน5อย่างดังนี้

2.1.1.การนำเข้าข้อมูล(Input)

ก่อนที่ข้อมูลทางภูมิศาสตร์จะถูกใช้งานได้ในระบบสารสนเทศภูมิศาสตร์ ข้อมูลจะต้องได้รับการแปลง ให้มาอยู่ในรูปแบบของข้อมูลเชิงตัวเลข (digital format) เสียก่อน เช่น จากแผนที่กระดาษไปสู่ข้อมูลใน รูปแบบดิจิตอลหรือแฟ้มข้อมูลบนเครื่องคอมพิวเตอร์อุปกรณ์ที่ใช้ในการนำเข้าเช่น Digitizer Scanner หรือ Keyboard

2.1.2.การปรับแต่งข้อมูล(Manipulation)

ข้อมูลที่ได้รับเข้าสู่ระบบบางอย่างจำเป็นต้องได้รับการปรับแต่งให้เหมาะสมกับงาน เช่น ข้อมูลบางอย่างมีขนาด หรือสเกล (scale) ที่แตกต่างกัน หรือใช้ระบบพิกัดแผนที่ที่แตกต่างกัน ข้อมูลเหล่านี้จะต้องได้รับการปรับให้อยู่ในระดับเดียวกันเสียก่อน

2.1.3.การบริหารข้อมูล(Management)

ระบบจัดการฐานข้อมูลหรือ DBMS จะถูกนำมาใช้ในการบริหารข้อมูลเพื่อการทำงานที่มีประสิทธิภาพในระบบ GIS DBMS ที่ได้รับการเชื่อถือและนิยมใช้กันอย่างกว้างขวางที่สุด คือ DBMS แบบ Relational หรือระบบจัดการฐานข้อมูลแบบสัมพัทธ์ (DBMS) ซึ่งมีหลักการทำงานพื้นฐานดังนี้คือ ข้อมูลจะถูกจัดเก็บ ในรูปของตารางหลาย ๆตาราง

2.1.4 .การเรียกค้นและวิเคราะห์ข้อมูล(Query and Analysis)

เมื่อระบบ GIS มีความพร้อมในเรื่องของข้อมูลแล้ว ขั้นตอนต่อไป คือ การนำข้อมูลเหล่านี้มาใช้ให้เกิด ประโยชน์ เช่น

- ตำแหน่งที่เกิดเหตุอยู่ที่ไหน ?
- แต่ละอำเภอมีระยะห่างกันกี่กิโลเมตร ?
- ตำบลนี้อยู่ที่อำเภอไหนในจังหวัด ?

หรือ ต้องมีการสอบถามอย่างง่าย ๆ เช่น ชี้เมาส์ไปในบริเวณที่ต้องการแล้วเลือก (point and click) เพื่อสอบถามหรือเรียกค้นข้อมูล นอกจากนี้ระบบ GIS ยังมีเครื่องมือในการวิเคราะห์ เช่น การวิเคราะห์เชิงประมาณค่า (Proximity หรือ Buffer) การวิเคราะห์เชิงซ้อน (Overlay Analysis) เป็นต้น

2.1.5.การนำเสนอข้อมูล(Visualization)

จากการดำเนินการเรียกค้นและวิเคราะห์ข้อมูลผลลัพธ์ที่ได้จะอยู่ในรูปของตัวเลขหรือตัวอักษร ซึ่งยากต่อการตีความหมายหรือทำความเข้าใจ การนำเสนอข้อมูลที่ดี เช่น การแสดงชาร์ต (chart) แบบ 2 มิติ หรือ3 มิติ รูปภาพจากสถานที่จริง ภาพเคลื่อนไหว แผนที่ หรือแม้กระทั้งระบบมัลติมีเดียสื่อต่าง ๆ เหล่านี้จะทำให้ผู้ใช้เข้าใจความหมายและมองภาพของผลลัพธ์ที่กำลังนำเสนอได้ดียิ่งขึ้นอีกทั้งเป็นการดึงดูดความสนใจของผู้ฟังอีกด้วย

2.2 Apache Web Server

Apache เป็นเว็บเซิร์ฟเวอร์ที่ใช้งานมากที่สุดในอินเทอร์เน็ตจุดกำเนิดของ Apache นั้นเกิดขึ้นจาก National Center for Supercomputing Applications (NCSA) HTTPd web server ซึ่งพัฒนาโดย Rob McCool ในช่วงปี 1990 และภายหลังจากที่โครงการ NCSA HTTPd ถูกยกเลิกไป ก็ได้มีนักพัฒนาหลายคนที่ได้นำ HTTPd มาปรับปรุงและใช้งาน

ในเดือน กุมภาพันธ์ 1995 ได้มีการจัดตั้ง Apache group ขึ้นโดยนักพัฒนา 8 คน และได้เผยแพร่เวอร์ชั่นแรกของ Apache คือ v 0.6.2 ในเดือนเมษายน 1995 และจากนั้น Apache 1.0 ก็ได้ถูกเผยแพร่เมื่อ 1 ธันวาคม 1995 และได้รับความนิยมอย่างรวดเร็วภายในเวลา 1 ปี กลายเป็นเว็บเซิร์ฟเวอร์ที่มีผู้ใช้งานมากที่สุด

ปัจจุบัน The Apache Software Foundation เป็นผู้ดูแลโครงการ Apache HTTP server ซึ่งมีจุดประสงค์เพื่อสร้างเว็บเซิร์ฟเวอร์ที่มีความทนทานต่อการใช้งาน มีคุณภาพในระดับของ commercial-grade มี feature ที่น่าใช้งาน และสามารถเปิดเผย source code ได้ ทั้งนี้สามารถใช้ Apache เว็บเซิร์ฟเวอร์ได้ฟรีภายใต้ข้อกำหนดของ Apache Software License

2.3 ภาษา Visual C#

ภาษา Visual C# เป็นภาษาที่ถูกออกแบบมาเพื่อการทำงานในยุค .NET โดยมีแนวของภาษาเป็นแบบเขียนโปรแกรมเชิงวัตถุสมัยใหม่(Modern Object Oriented Programming) เรียกสั้นๆ ว่า Modern OOP

สำหรับผู้ที่ได้ศึกษาภาษา C/C++ มาก่อนย่อมทราบดีว่า การเขียนโปรแกรมในเชิงวัตถุ มีความยุ่งยากและยากต่อการเข้าใจ เพราะต้องอาศัยความรู้หลายอย่างมาประกอบกัน ซึ่งสิ่งที่ทำให้เกิดความสับสนมากที่สุดคือ Pointer

สำหรับจุดยืนของ Visual C# จะอยู่ที่การอาศัยไวยกรณ์ที่ปรับปรุงมาจาก C/C++ ร่วมกับความง่ายของภาษา Visual Basic ร่วมเข้ามาเป็น Visual C#

แนวความคิดของการเขียนโปรแกรมแบบ Modern OOP เกิดจากการที่ไมโครซอฟท์พัฒนาคลาส(Class) ต้นแบบต่างๆ ขึ้นมา ที่เรียกว่า Base Class Library แล้วนำมาจัดหมวดหมู่ให้เป็นระเบียบ เมื่อต้องการเรียกใช้งานคลาสใด ก็จะอาศัยระบบ Namespaces เข้ามาช่วยในการระบุ Class ต้นแบบต่างๆ เพื่อให้ผู้พัฒนาสามารถนำ Object ต่างๆ ที่อยู่ใน Class นั้นๆ ออกมาใช้งานได้อย่างง่ายดาย

2.4 PostgreSQL

PostgreSQL เป็น Object-Relational DBMS โดยสามารถใช้รูปแบบของภาษา SQL ได้เกือบทั้งหมด และสามารถใช้ subselects transactions user-defined types และ functions ได้ อีกทั้งเป็น Database ซึ่งให้ Source code ฟรี

<u>ประวัติการพัฒนา</u>

-เป็น Project ของ Prof. Michael Stonebraker ที่ มหาวิทยาลัย Berkeley ซึ่งเดิมพัฒนามาจาก Ingres ซึ่งเป็นที่รู้จักในชื่อ CA-Ingres II ซึ่ง Ingres ใช้ ภาษา query QUEL

-เป็นภาษาของตัวเอง ปัจจุบันได้หยุดพัฒนาไปแล้ว แต่ยังสามารถนำมาใช้ได้ฟรี

-ต่อมา Prof. Stonebraker ได้นำมาพัฒนาเป็น Postgres (มาจาก after Ingres) ซึ่งได้ใช้ภาษา query เป็น POSTQUEL เป็น Postgres version 4.2

- ต่อมาในช่วง ปี 1987 Postgres ได้เสนอ rules procedures time travel extensible types และ object-relational concepts

-Postgres ถูกนำมาใช้ เพื่อการค้า ในชื่อว่า Illustra (ปัจจุบัน ถูก Informix ซื้อไป และรวมเข้าไว้ใน Universal Server)

- ต่อมา นักศึกษาปริญญาเอก 2 คน คือ Andrew Yu และ Jolly Chen ได้พัฒนา Postgres ให้ใช้ภาษา query ตามรูปแบบที่เป็นมาตราฐาน แทนที่ ภาษา POSTQUEL เดิม ซึ่งได้เผยแพร่ในปี 1995 จึงเรียกเป็น Postgres95 หรือ version 5 หลังจากนั้น การพัฒนาต่อ โดย กลุ่มพัฒนาทาง Internet

- ปัจจุบัน เปลี่ยนชื่อเป็น PostgreSQL พัฒนาต่อเนื่องเป็น version 8.1

<u>ลักษณะโครงสร้าง</u>

ระบบที่ใช้ PostgreSQL จะติดตั้ง PostgreSQL ไว้ที่เครื่อง Server ซึ่งเป็นที่เก็บ database ด้วย และยังสามารถ ติดตั้ง PostgreSQL ได้มากกว่า 1 ชุดใน Server เครื่องเดียวผู้ดูแลระบบ PostgreSQL จะใช้ชื่อว่า *postgres* ซึ่งเป็นผู้ดูแลทั้ง ตัวโปรแกรม และ database ซึ่งสามารถทำงานกับบางคำสั่งเฉพาะ เพื่อจัดการ database และ ผู้ใช้บริการ (user) ซึ่ง ผู้ดูแลระบบ database (postgres) จะคล้ายการทำงานของ superuser ในระบบ Unix หน้าที่ของ *postgres* สามารถ สร้างชื่อ user และกำหนดสิทธิและระดับการใช้งานต่างๆได้

PostgreSQL ใช้รูปแบบการทำงาน แบบ Client/Server ซึ่งในการทำงานจะประกอบด้วย 3 process ทำงานร่วมกัน คือ

- Postmaster เป็น supervisory daemon process ซึ่งจัดการติดต่อระหว่าง Frontend กับ Backend process ในการ allocate share buffer จัดการค่าเริ่มต้นต่างๆในระหว่างเริ่มทำงาน และเก็บบันทึกการเข้าใช้ระบบและความผิดพลาดต่างๆที่เกิดขึ้น
- Postgres เป็น backend process เพื่อจัดการ database ถือว่าส่วนนี้เป็น process ที่ทำงานจริงๆ เช่น ทำงานตาม query โดย Postmaster จะสั่งให้สร้าง Backend process สำหรับทุกๆ การเชื่อมต่อกับ Frontend ดังนั้น Postgres นี้จะทำงานที่ server
- Frontend เป็น application ซึ่งจะทำงานที่เครื่อง client และจะส่งคำสั่งการเชื่อมต่อหรือคำสั่งต่างๆ มาที่ Postmaster แล้ว Postmaster จึงส่งต่อการทำงานไปที่ Postgres

<u>หลักการทำงานของ PostgreSQL</u>

การทำงานจะแบ่ง process ที่ทำงานดังที่กล่าวมาแล้ว คือ

- ในส่วนของ Supervisory daemon process คือ Postmaster
- ในส่วนของ User's Frontend application เช่น โปรแกรม psql หรือ CGI-Perl
- ในส่วน Backend database servers คือ Postgres

เมื่อโปรแกรม ทาง Frontend ต้องการข้อมูล หรือทำงานกับdatabaseโดยเรียกผ่านทาง library ibpq ซึ่ง library libpq นี้ จะส่ง requests ผ่านทาง Network ไปยัง Postmaster เมื่อ Postmaster ได้รับ request ดังกล่าว ทาง Postmaster จะสร้าง Backend process ขึ้นที่ server เพื่อติดต่อกับ Frontend แทน การทำงานนั้นจะเกิดขึ้นระหว่าง Frontend กับ Backend โดยไม่ผ่าน Postmaster อีก และ Postmaster ก็ทำงานต่อไป คือรอรับ request อื่นๆต่อไป Library libpq จะให้ หนึ่ง Frontend สามารถติดต่อได้หลาย Backend processes แต่การทำงานยังเป็นแบบ single threaded เนื่องจาก

library libpq ยังไม่สามารถทำ multithreaded ได้ ตามหลักการที่กล่าวมา ดังนั้น Postmaster กับ Backend จะต้องทำงานอยู่ที่ เครื่องเดียวกัน คือ database server แต่ Frontend จะทำงานที่เครื่องใดก็ได้

PostGIS

PostGIS เป็นการเพิ่มเติมในส่วนของฐานข้อมูลเชิงวัตถุสัมพันธ์ (Object–relational database system) ของ PostgraSQL ให้มีการรองรับวัตถุทางระบบสารสนเทศภูมิศาสตร์ (GIS object) เข้ามาเก็บไว้ในฐานข้อมูล มีการสนับสนุน GiST indexs และ R-tree indexs ซึ่งเป็นวิธีการค้นข้อมูลแบบตัวชี้ (Indexing) ที่ใช้ในฐาน
ข้อมูลเชิงพื้นที่สำหรับฐานข้อมูลที่มีขนาดใหญ่ ซึ่ง PostGIS เองมีการกำหนดการใช้งานโดย OpenGIS ที่เป็น ลักษณะพื้นฐานของ SQL (SFSQL)

<u>ลักษณะทั่วไปของ PostGIS</u>

ดังที่กล่าวมาแล้วว่า PostGIS มีการกำหนดการใช้งานที่เป็นลักษณะพื้นฐานโดย OpenGIS Consortium (OGC) ซึ่งเป็นสถาบันที่ศึกษาเพื่อการสร้างอินเตอร์เฟสที่ทำโปรแกรมประยุกต์ซอฟแวร์ ให้มีการใช้งานได้กับระบบสารสนเทศภูมิศาสตร์ (GIS)

OpenGIS เป็นสิ่งกำหนดความชัดเจนในการเข้าถึงข้อมูลทางด้านธรณีที่แตกต่างกัน และประมวลผลทางธรณีของแหล่งทรัพยากรในสภาพแวดล้อมที่เป็นเครือข่าย (Network) ซึ่งลักษณะที่มีการกำหนด โดย OpenGIS มีดังนี้

• OpenGIS SFSQL Objects เป็นการกำหนดในส่วนของวัตถุเชิงพื้นที่ เช่น

- POINT
- LINESTRING
- POLYGON
- MULTIPOINT
- MULTILINESTRING
- MULTIPOLYGON
- GEOMETRYCOLLECTION

• OpenGIS SFSQL Representations เป็นการกำหนดในส่วนของการแสดงผลซึ่งเป็นมาตรฐานในการแสดงผลของวัตถุเชิงพื้นที่ มี 2 รูปแบบคือ

1) Well – Known Text (WKT) Form

เป็นการแสดงผลในรูปแบบที่เป็น String เช่น

POINT(1 1)

MULTIPOINT(1 1, 3 4, -1 3)

LINESTRING(1 1, 2 2, 3 4)

POLYGON((0 0, 0 1, 1 1, 1 0, 0 0))

MULTIPOLYGON((0 0, 0 1, 1 1, 1 0, 0 0), (5 5, 5 6, 6 6, 6 5, 5 5))

MULTILINESTRING((1 1, 2 2, 3 4),(2 2, 3 3, 4 5))

2) Well – Known Binary (WKB) Form

เป็นการแสดงผลในรูปแบบที่เกี่ยวกับบิตซึ่งจะดึงข้อมูลออกจากฐานข้อมูลโดยไม่มีการเปลี่ยนไปแสดงในรูปแบบที่เป็น String

<u>การใช้รูปแบบมาตรฐาน OpenGIS</u>

OpenGIS เป็นการระบุลักษณะพื้นฐานสำหรับ SQL ซึ่งจะกำหนดมาตรฐานระบบสารสนเทศภูมิศาสตร์ในส่วนที่เป็นชนิดของวัตถุ ฟังก์ชันที่ใช้ในการจัดการ และกลุ่มของตารางของ OpenGIS ที่ใช้มี 2 ตาราง ดังนี้

1) The SPATIAL_REF_SYS Table มีรูปแบบในการกำหนดดังต่อไปนี้

```
CREATE TABLE SPATIAL_REF_SYS (
SRID INTEGER NOT NULL PRIMARY KEY,
AUTH_NAME VARCHAR(256),
AUTH_SRID INTEGER,
SRTEXT VARCHAR(2048),
PROJ4TEXT VARCHAR(2048) )
```

SRID

ตัวเลขที่ใช้ในการระบุระบบการอ้างอิงเชิงพื้นที่ภายในฐานข้อมูล

AUTH_NAME

ชื่อมาตรฐาน หรือ ส่วนที่เป็นมาตรฐาน สำหรับระบบการอ้างอิง

AUTH_SRID

หมายเลขของระบบการอ้างอิงเชิงพื้นที่ ซึ่งมีการกำหนดโดยการอ้างอิงจาก AUTH_NAME

SRTEXT

เป็นการแสดงผลแบบ Well – known Text ของระบบการอ้างอิงเชิงพื้นที่

PROJ4TEXT

เป็น Proj4 library ซึ่งใช้กำหนด String ให้กับ SRID

2) The GEOMETRY_COLUMNS Table มีรูปแบบในการกำหนดดังต่อไปนี้

```
CREATE TABLE GEOMETRY_COLUMNS (
    F_TABLE_CATALOG VARCHAR(256) NOT NULL,
    F_TABLE_SCHEMA VARCHAR(256) NOT NULL,
    F_TABLE_NAME VARCHAR(256) NOT NULL,
    F_GEOMETRY_COLUMN VARCHAR(256) NOT NULL,
    COORD_DIMENSION INTEGER NOT NULL,
    SRID INTEGER NOT NULL,
    TYPE VARCHAR(30) NOT NULL )
    F_TABLE_CATALOG, F_TABLE_SCHEMA, F_TABLE_NAME
```

ลักษณะพิเศษของชื่อตารางที่ประกอบด้วยคอลัมน์เชิงเรขาคณิต F_GEOMETRY_COLUMN ชื่อของคอลัมน์ทางเรขาคณิต ซึ่งมีลักษณะเป็นตาราง COORD_DIMENSION มิติเชิงพื้นที่ของคอลัมน์ SRID หมายเลขของระบบอ้างอิงเชิงพื้นที่ ใช้สำหรับเรขาคณิตในตาราง ซึ่งเป็นคีย์นอกที่อ้างอิงถึง SPATIAL_REF_SYS TYPE ชนิดของวัตถุเชิงพื้นที่

2.5 XML

XML (eXtensible Markup Language) ถูกออกแบบมาเพื่อให้ผู้สร้างเอกสารสามารถนำไปใช้งานในรูปแบบวิธีการที่ง่าย มีความชัดเจนและเป็นเซตย่อยของ SGML (Standard Generalized Markup Language) ซึ่งเป็นภาษาที่นิยมใช้และได้รับการพัฒนาให้มีประสิทธิภาพสูงสุดในการทำงานบนเว็บ โดย XML จะประกอบด้วย 3 ส่วนพื้นฐานด้วยกัน คือ เอกสารข้อมูล (Data document) เอกสารนิยามความหมาย (definition document) และนิยามภาษา (definition language)

ปัจจุบัน มี 2 ภาษาด้วยกันที่มีการสร้างนิยามภาษาของเอกสารข้อมูลภาษาอื่นได้ คือ SGML และ XML เช่น ภาษา WML(Wireless Markup Language) ก็มีต้นกำเนิดมาจาก XML ที่ใช้ในการแสดงข้อความบนโทรศัพท์มือถือระบบ WAP (Wireless Application Protocol) โดยที่ XML ถือได้

ว่าเป็นส่วนหนึ่งของ SGML ที่เป็นข้อกำหนดในการสร้างหรือจัดทำเอกสารในรูปแบบเล็กทรอนิกส์ที่กำหนดโดย W3C หรือ World Wide Web Consortium ใน SGML มีกระบวนการเรียกว่า Information Analysis สำหรับใช้ตรวจสอบโครงสร้างและราย ละเอียดของข้อมูล กระบวนการดังกล่าวเรียกว่า DTD (document type definition) ซึ่งเป็นตัวชี้เนื้อหาของ ออบเจ็กต์ (Object) ในกลุ่มข้อมูล

ยุคแรกในช่วงปี ค.ศ. 1980 ภาษา HTML (Hypertext Markup Language) ได้ถูกพัฒนาขึ้นมาและกำหนดให้เป็นภาษาในระดับที่ง่ายสำหรับการสร้างเอกสารรายงานที่เป็นรูปแบบข้อมูลมาตรฐาน อย่างไรก็ตาม HTML ก็ยังมีจุดด้อยที่พบมากที่สุด คือ ไม่มีความสามารถในการจัดรูปแบบเอกสารหรือการนำข้อมูลกลับมาใช้ใหม่ที่ค่อนข้างทำได้ยาก และไม่สามารถแยกความแตกต่างระหว่างข้อมูลธรรมดาและที่อยู่ในรูปแบบคำสั่งได้ ส่วนใหญ่แล้วไม่เคร่งครัดทำตามกฎมากนักเพราะว่าข้อกำหนดไม่ค่อยรัดกุม แต่ภาษา HTML ก็ยังคงเป็นภาษาหลักที่ใช้บอกเบราเซอร์ให้รู้ว่าจะแสดงผลข้อมูลบนเว็บอย่างไร

ในห้องปฏิบัติการของบริษัท ไอบีเอ็ม ได้เริ่มต้นพัฒนาต้นแบบภาษามาร์คอัพ(Markup) เพื่อที่จะเอาชนะอุปสรรคและข้อจำกัด ภาษาดังกล่าวนี้จึงได้มีวิวัฒนาการเข้าไปใน SGML ซึ่งเป็นภาษาที่มีพื้นฐานมาจากแท็ก (Tag) ที่สามารถใช้กำหนดรูปแบบของข้อมูล จึงได้กลายเป็นภาษามาตรฐานสำหรับการกำหนดคำอธิบายรูปแบบโครงสร้าง ความแตกต่างของเอกสารอิเล็กทรอนิกส์ โดย SGML สามารถที่จะแยกข้อมูลออกจากคำสั่งและมีเครื่องมือการเขียนโปรแกรมที่มีประสิทธิภาพสูง

อย่างไรก็ตามการพัฒนา SGML ขึ้นมาก็ยังมีข้อจำกัดในด้านความยืดหยุ่น การนำไปใช้งานก็ยังมีความซับซ้อนอยู่ จึงจำเป็นต้องหาภาษาอื่นที่มีความยืดหยุ่น มีมาตรฐานที่ถ่ายโอนได้ง่าย มีค่าใช้จ่ายที่ต่ำ รวดเร็ว ใช้งานง่ายเหมือน HTML และต้องขยายความสามารถได้เหมือน SGML โดยจะต้องมีความเสถียรพร้อมกับสามารถปรับปรุงแก้ไขรูปแบบได้ง่าย ตลอดจนการจัดเก็บและการถ่ายโอนข้อมูลระหว่างแอพลิเคชันประเภทข้อความด้วยกัน ขณะเดียวกันก็ต้องสอดคล้องกับ HTML และ SGML ด้วยเหตุนี้เองจึงได้กำเนิดภาษา XML(eXtensible Markup Language) ขึ้นมา

XML สามารถที่จะจัดการได้หลายรูปแบบทั้งองค์ประกอบ โครงสร้างเอกสาร ลักษณะประเภท แอตทริบิวต์ และอิลิเมนต์ โดยเป็นภาษาที่ถูกออกแบบมาเฉพาะสำหรับการพัฒนาโปรแกรมเว็บเพื่อการจัดส่งข้อมูลสารสนเทศ ตลอดจนถูกนำมาใช้สร้างภาษามาร์คอัพ (markup) นั้นเอง ซึ่งตรงกันข้ามกับ SGML ที่มีความซับซ้อนมากกว่า ส่วน HTML ก็เป็นเอกสารไม่มีความยืดหยุ่นต่อการใช้งาน อย่างไรก็ตามในทางปฏิบัติเอกสาร XML มีกฎพื้นฐานเพื่อให้การสร้าง

เอกสารมีรูปแบบที่ถูกต้อง ในการใช้งานจริงโดยปกติแล้ว XML สามารถจัดเก็บฐานข้อมูล กำหนดโครงสร้างเอกสาร การนำเสนอมัลติมีเดียต่างๆ การจัดเก็บกราฟิกที่มีลักษณะแบบเวกเตอร์ ตลอดจนการสื่อสารระหว่างโปรแกรมต่างๆ และนอกจากนี้แล้ว XML ยังสามารถช่วยในการประมวลผลข้อมูลแล้วส่งผ่านให้โปรแกรมประยุกต์ไปยังแหล่งเก็บข้อมูล อย่างเช่น ข้อความหรือข้อมูล เป็นต้น

XML เป็นเอกสารที่เขียนด้วยข้อความปกติธรรมดา คุณจึงสามารถสร้างเอกสารหรือแก้ไขไฟล์ XML ได้อย่างง่ายดาย โดยการใช้โปรแกรมแก้ไขข้อความ (Text Editor) ซึ่งเป็นมากกว่าเครื่องมือการเขียนโปรแกรมที่ไม่ซับซ้อน หรือถ้าหากต้องการใช้โปรแกรมที่มีความสามารถพิเศษมากกว่านี้ ก็ต้องใช้โปรแกรมแก้ไขข้อความที่อยู่ในชุดโปรแกรม Microsoft Visual Studio เช่น Microsoft Visual C++ Microsoft Visual Basic และ Microsoft Visual Foxpro เป็นต้น ภาษา XML ใช้แท็กเริ่มต้นและแท็กปิดเสมอเช่นเดียวกับ HTML ซึ่งเรียกว่า อิลิเมนต์ (Element) ซึ่งเป็นการแบ่งแยกระหว่างข้อมูลและคำสั่ง เพื่อระบุว่าข้อมูลที่อยู่ระหว่างแท็กดังกล่าวคือ ข้อมูลอะไร

ส่วนประกอบในเอกสาร XML มีอยู่ 2 ส่วนหลักด้วยกัน คือ Prolog Element และ Document Element (หรือ Root Element) ในส่วนของเอกสาร XML คือ Element เดี่ยว ซึ่งสามารถบรรจุ Element เพิ่มเติมในเอกสาร XML ได้ โดยในเอกสาร XML นั้น Element จะแสดงลักษณะโครงสร้างของเอกสาร และจะแสดงส่วนประกอบเนื้อหาของเอกสารอยู่ภายในสัญลักษณ์ Element ประกอบด้วยแท็กเริ่มต้น(start-tags) เนื้อหาภายใน Element และแท็กสิ้นสุด (end-tags) ส่วนเนื้อหาภายใน Element สามารถเป็นได้ทั้งข้อมูลหรือ Element อื่นๆ ที่ซ้อนอยู่ภายในหรือทั้งสองแบบ

2.6 UMN MapServer

UMN MapServer เป็น MapServer ของ University of Minenesota ซึ่งเป็น Internet Web Application ที่เป็น Open Source ที่สามารถ Compile และใช้งานได้หลาย Platform ไม่ว่าจะเป็น Unit/Linux Microsoft Windows และแม้แต่ MacOS ซึ่ง Map Server ดังกล่าวได้ถูกเตรียมการไว้ให้ใช้ได้กับหลายๆรูปแบบไม่ว่าจะเป็น vector format ที่ support พวก ESRI shapefiles PostGIS ESRI ArcSDE และอื่นๆอีกมากมายผ่าน ORG raster formats ที่ support พวก TIFF/GeoTIFF EPPL7 and many และอื่นๆซึ่งสามารถศึกษาได้จาก http://mapserver.gis.umn.edu/

เริ่มแรกนั้น MapServer ได้ถูกพัฒนาที่ UNM ผ่าน ForNet Project ซึ่งได้รับการสนับสนุนจาก NASA โดยความพยายามของภาควิชา Natural Resource ของ Minnesota ต่อมายังคงได้รับการ

สนับสนุนผ่าน TerraSIP project ของ NASA ซึ่งต่อมา Sorftware ดังกล่าวได้เจริญเติบโตขึ้นและถูกควบคุมดูแลโดยผู้พัฒนาจากทั่วโลกและได้รับการสนับสนุนทุนหลายแหล่ง

MapServer Overview

ในรูปแบบพื้นฐานของ MapSever เป็น CGI Program ที่ปกติจะอยู่นิ่งๆบน Wep Server ของเรา แต่เมื่อ Request ถูกส่งไปยัง MapServer โดยจะใช้ Information ผ่าน Request URL และ Map File เพื่อสร้างรูปภาพแผนที่ตามที่เราส่ง Request มา ซึ่ง Request อาจคืนค่ารูปภาพของสัญลักษณ์ Scal Bars แผนที่อ้างอิง และค่าต่างๆผ่านทาง CGI Variables ซึ่งเป็นหลักการทำงานโดยทั่วไปของ MapServer

MapServer สามารถขยายขีดความสามารถและปรับแต่งได้ยากมากมันสามารถเพิ่มความสามารถเพื่อสนับสนุน Input และ Output Data ชนิดต่างๆที่แตกต่างกันได้หลายๆชนิด ซึ่งเป็นความสามารถของ MapServer ที่ถูก Compile แบบเต็มรูปแบบ

MapScipt

MapScipt จะเตรียม Scipt เชื่อมต่อสำหรับ MapServer ซึ่งเป็น Module ที่สามารถใส่เข้าไปได้ ซึ่งเป็นการเพิ่มความสามารถให้ MapServer เพื่อให้เราสามารถใช้ภาษา Script เช่น PHP Perl Python Ruby Tcl Java และ C#

โครงสร้างพื้นฐานของ MapServer Application จะประกอบด้วย

Map File

Map File เป็น configuration File ของ MapServer Application ซึ่งมีโครงสร้างเป็น Text โดยจะกำหนดพื้นที่ใน map ของเราบอก MapServer program ว่าข้อมูลเราอยู่ที่ไหนและภาพ Output แสดง บริเวณไหนนอกจากนี้ยังกำหนด Map Layers ของเราด้วยรวมถึงกำหนดแหล่งข้อมูล Projections และสัญลักษณ์ที่ใช้ต่างๆซึ่งโดยทั่วไปจะมีนามสกุล .map

Geographic Data

MapServer สามารถใช้ชนิดของข้อแหล่งมูลทางภูมิศาสตร์ได้อย่างหลากหลายซึ่งรูปแบบโดย Default คือ ESRI Shapefile ซึ่งรายละเอียดของ Geographic Data ที่ MapServer Support สามารถดูได้ที่ Website ของ UMN

HTML Pages

HTML Pages เป็นตัวเชื่อมระหว่างผู้ใช้และ MapServer ซึ่งโดยทั่วไปจะอยู่ใน Root ของ Wep โดยในรูปแบบที่ง่ายที่สุด MapServer สามารถถูกเรียกเพื่อวางรูปภาพของ Map ที่เป็นแบบ Static บน HTML Pages ซึ่งเพื่อทำให้ Map มีการตอบสนองกับเราภาพจะถูกวางไว้บน HTML Form บน Webpage

CGI Programs

CGI Programs เป็น Stateless คือทุกครั้งที่ Request ถูก Get จะไม่จำ Request ทุกๆ อย่างที่ผ่านมาซึ่งสิ่งเหล่านี้เป็นเรื่องปกติของ Application ด้วยเหตุนี้ทุกๆครั้งที่ Application ของเราส่ง Request ไปยัง MapServer จึงต้องการ Context Information(Layer อยู่บนอะไรอยู่ใน Map ที่ไหน Application Mode อะไร เป็นต้น)

MapServer CGI

MapServer CGI เป็น Binary File(.EXE) ที่รับ Request แล้วส่งคืนเป็นภาพหรือข้อมูล โดยปกติจะอยู่ที่ cgi-bin Directory หรือ Scripts Directory ของ HTTP Server และด้วยเหตุผลในเรื่องความปลอดภัยจึงไม่ควรจะอยู่ใน Web Root

HTTP Server

HTTP Server จะให้บริการ HTML Pages เมื่อถูกติดต่อโดยผู้ใช้งาน Browser เราต้องอาศัยการทำงานของ HTTP Server เช่น Apache และ IIS บนเครื่องที่เราติดตั้ง MapServer

2.7 Java Server Pages : JSP

Java Server Pages หรือ JSP เทคโนโลยีที่คิดค้นโดยบริษัท Sun Microsystems โดยพัฒนาบนพื้นของภาษาจาวาเพื่อเพิ่มประสิทธิภาพให้หน้าเว็บเพจมีความยืดหยุ่นสูงขึ้น โครงสร้างของ JSP นั้นเป็นลักษณะของแท็ก (tag) ชนิดพิเศษที่แทรกเข้าไปในเอกสาร HTML และ เปลี่ยนนามสกุลของเอกสารเป็น .jsp แทนที่จะเป็น .html หรือ .htm โดยแท็กเหล่านี้เว็บบราวเซอร์จะไม่สามารถตีความหมายได้ จะต้องนำไปประมวลผลก่อนที่เว็บเซิร์ฟเวอร์เท่านั้น แล้วนำผลลัพธ์ทั้งหมดส่งกลับมายังเว็บบราวเซอร์ในลักษณะของเอกสาร HTML ซึ่งเว็บบราวเซอร์สามารถตีความหมายและนำมาแสดงผลได้

การทำงานโดยรวมของ JSP จะเริ่มจากบราวเซอร์ร้องขอ (HTTP Request) เอกสารที่มีนามสกุลเป็น JSP ไปยังเว็บเซิร์ฟเวอร์ผ่านทางโพรโตคอล HTTP เว็บเซิร์ฟเวอร์ก็จะนำเอกสาร JSP ที่

ได้รับมานั้นส่งต่อไปให้ JSP Engine จากนั้น JSP Engine ก็จะประมวลผล และ ส่งผลลัพธ์กลับมายังเว็บเซิร์ฟเวอร์ แล้วเว็บเซิร์ฟเวอร์ก็จะส่งผลลัพธ์กลับมายังบราวเซอร์อีกที ในลักษณะของเอกสาร HTML บราวเซอร์ก็จะสามารถแสดงผลได้

ข้อดีของการใช้ JSP คือ

1. Write Once Run Anywhere JSP มีพื้นฐานมาจากภาษาจาวา ข้อดีประการหนึ่งที่มีตามมาด้วยคือการทำงานได้ในหลายระบบปฏิบัติการ

2. Component Reusable ด้วยความสามารถในการนำจาวาบีน (Java Bean) มาใช้ ซึ่งมีลักษณะเป็นคอมโพเน็นต์ ทำให้คอมโพเน็นท์เหล่านี้สามารถรนำกลับมาใช้ได้ใหม่และใช้ร่วมกันระหว่างผู้พัฒนาเว็บไซต์ทำให้การพัฒนาทำได้เร็วขึ้น

3. Java Extension เมื่อ JSP พัฒนาบนพื้นฐานของภาษาจาวา ซึ่งมีคุณสมบัติหลายอย่างไม่ว่าจะเป็น การโปรแกรมเชิงวัตถุ (Object Oriented Programming) การทำงานกับ Thread

2.8 Pocket PC

PocketPC เป็นอุปกรณ์คอมพิวเตอร์มือถือ ที่มีการเพิ่มเติมใหม่ล่าสุด เพื่อให้อุปกรณ์ที่ใช้ Windows มีความสามารถสูง มีเสถียรภาพเข้ามาแทนที่ คอมพิวเตอร์ที่มีขนาดเท่าฝ่ามือ (Palm-Size PC) ที่ ปฏิบัติการอยู่บน Windows CE 2.0 หรือ 2.11 และ Pocket Outlook โดย Pocket PC นี้จะอ้างอิงถึงอุปกรณ์ใช้ Windows ที่มีความสามารถสูง ที่ไมโครซอฟต์ขนานนามว่า Windows Powered Devices ซึ่งอุปกรณ์นี้ จะปฏิบัติการอยู่บน Windows CE 3.0 และรวมโปรแกรม ฉบับกระเป๋าของ Microsoft Desktop เช่นPocket Outlook และ Pocket Office เข้าไป นอกจากนี้ Pocket PC ยังมีความแตกต่างจาก Palm-Size PC คือ จะมีตัวประมวลผลที่ไวกว่า มีหน่วยความจำที่เยอะกว่า และปรับปรุงให้มีอายุ การใช้งานแบตเตอรี่ ที่ยาวนานกว่า ระบบปฏิบัติการใหม่ที่ Pocket PC ใช้นี้ แสดงให้เห็นถึงการให้ความสำคัญกับอุปกรณ์เคลื่อนที่ (Mobile Devices)

2.9 การวิเคราะห์ข้อมูลทางสถิติที่เกี่ยวข้องกับโครงงาน

สำหรับการวิเคราะห์ข้อมูลทางสถิติที่เกี่ยวข้องกับโครงงานจะใช้เรื่องสถิติเชิงอนุมานที่เป็น Parametric เนื่องจากข้อมูลที่จะนำเข้าเมื่อผ่านการทดสอบแล้วพบว่าการแจกแจงแบบปกติ ซึ่งสถิติแต่ละตัวที่ใช้ในโครงงานมีรายละเอียดดังต่อไปนี้

<u>Independent Sample t-test</u>

Independent Sample t-test เป็นการทดสอบความแตกต่างของค่าเฉลี่ยของกลุ่มตัวอย่างสองกลุ่ม เป็นการทดสอบเมื่อต้องการทราบค่าเฉลี่ยของทั้งสองกลุ่มว่ามีค่าเท่ากันหรือไม่ โดยมีข้อตกลงเบื้องต้นดังนี้

-กลุ่มตัวอย่างทั้งสองได้มาจากการสุ่ม(Random) จากประชากรที่มีการแจกแจงเป็นโค้งปกติ

-ไม่ทราบค่าความแปรปรวนของประชากร $\left(\sigma_1^2, \sigma_2^2\right)$ ใช้ค่าความแปรปรวนของตัวอย่างแทนแทน $\left(S_1^2, S_2^2\right)$

-กลุ่มตัวอย่างต้องเป็นอิสระต่อกัน(Independent)

สูตรที่ใช้ในการคำนวณมีดังนี้

ก.ค่า t เมื่อพบว่าความแปรปรวนของประชากรไม่แตกต่างกัน

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{S_p{}^2((1/n_1) + (1/n_2))}}$$

เมื่อ df=n1+n2-2

และ

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n1 + n2 - 2}$$

ข. ค่า t เมื่อพบว่าความแปรปรวนของประชากรแตกต่างกัน

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{(S_1^2 / n_1) + (S_2^2 / n_2)}}$$

$$df = \frac{(S_1^2 / n_1) + (S_2^2 / n_2)}{[(S_1^2 / n_1)/(n_1 - 1) + (S_2^2 / n_2)/(n_2 - 1)]}$$

เมื่อ $\overline{x_1}, \overline{x_2}$ แทนค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1 และกลุ่มที่ 2

S_p^2 แทนความแปรปรวนรวม

S_1^2, S_2^2 แทนความแปรปรวนของกลุ่มตัวอย่างที่ 1 และกลุ่มตัวอย่างที่ 2

n_1, n_2 แทนจำนวนขนาดของกลุ่มตัวอย่างที่ 1 และกลุ่มตัวอย่างที่ 2

สำหรับขั้นตอนการทดสอบความแตกต่างของค่าเฉลี่ยของกลุ่มตัวอย่างสองกลุ่มที่เป็นอิสระต่อกัน มีขั้นตอนการดำเนินการดังนี้

1.ตั้งสมมติฐาน

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

2. ทดสอบว่าความแปรปรวนประชากรของกลุ่มตัวอย่างทั้งสองเท่ากันหรือไม่

3. เลือกสถิติที่ใช้ทดสอบค่า t

4. พิจารณาขอบเขตยอมรับหรือปฏิเสธ H_0

จะปฏิเสธ H_0 เมื่อค่า t ที่คำนวณได้อย่างน้อยหรือเท่ากับ $-ta/_{2(df)}$ หรือ ค่า t ที่คำนวนได้น้อยกว่าหรือเท่ากับ $-ta/_{2(df)}$ จากการเปิดตารางแจกแจงแบบ t

<u>One Way ANOVA(F-test)</u>

One Way ANOVA เป็นการวิเคราะห์ความแปรปรวนแบบทางเดียวเป็นการทดสอบที่ผู้สนใจตัวแปรอิสระเพียงตัวเดียวเท่านั้น โดยจะใช้ในกรณีตัวแปรอิสระมีค่ามากกว่า 2 (ถ้าตัวแปรอิสระมีค่าเท่ากับสองก็ใช้ Independent Sample t-test)

โดยขั้นตอนการทดสอบ One Way ANOVA มีดังนี้

1.ตั้งสมมติฐาน

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = ... = \mu_k$$
$$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq ... \neq \mu_k$$

2. คำนวณค่าทางสถิติที่ใช้ทดสอบ F-test

$$F = MS_b / MS_w$$

เมื่อ MS_b แทน ความแปรปรวนระหว่างกลุ่ม = SS_b / df_b

MS_w แทน ความแปรปรวนภายในกลุ่ม = SS_w / df_w

$SS_t - SS_b - SS_w$

โดยที่ $SS_t = \sum_{i=1}^{n}\sum_{j=1}^{k} X_{ij}^2 - (T^2 / N)$

$$SS_b = \sum_{j=1}^{k}[T_j^2 / n_j] - (T^2 / N)$$

และคำนวณ SS_w จาก $SS_w = SS_t - SS_b$

3.กำหนดนัยสำคัญ α

4.พิจารณาขอบเขตวิกฤติโดยจะปฏิเสธ H_0 เมื่อ F ที่คำนวณได้มากกว่าหรือเท่ากับ $F_\alpha(k-1, k-2)$ ในกรณีเช่นนี้ผลการเปรียบเทียบจะมีค่าความแตกต่างอย่างมีนัยสำคัญทางสถิติ หมายความว่าค่าเฉลี่ยของประชากรเหล่านั้นมีความต่างกันอยู่อย่างน้อยหนึ่งคู่

5.หากพบว่ามีความแตกต่างกันอย่างน้อยหนึ่งคู่ก็ทำการทดสอบรายคู่โดยใช้วิธีการของ Duncan ซึ่งมีสูตรการคำนวณดังนี้

$$Dun = q\sqrt{(MS_w / n)}$$