บทที่ 2

ระบบสารสนเทศภูมิศาสตร์

2.1 ระบบสารสนเทศภูมิศาสตร์

ระบบสารสนเทศภูมิศาสตร์มาจากคำภาษาอังกฤษว่า Geographic Information System มักจะเรียกโดยใช้คำย่อว่า GIS ซึ่งเป็นเทคโนโลยีที่เริ่มนิยมนำมาประยุกต์ใช้กับชีวิตประจำวันในด้านต่าง ๆมากขึ้น ทำให้วงการต่าง ๆ ทั้งการศึกษา บริษัทผู้ผลิตและขายซอฟต์แวร์ในท้องตลาด รวมทั้งประชาชนทั่วไป ตื่นตัวหันมาให้ความสนใจกับคำว่า “ระบบสารสนเทศภูมิศาสตร์” กันมากขึ้น มีผู้ให้คำนิยามคำว่า “ระบบสารสนเทศภูมิศาสตร์” ต่าง ๆ อาทิ

บริษัทอีเอสอาร์ไอ ให้ความหมายว่า [ESRI, 2005] “ระบบสารสนเทศภูมิศาสตร์ เป็นระบบของอุปกรณ์คอมพิวเตอร์ ซอฟต์แวร์ ข้อมูลภูมิศาสตร์ และบุคลากร ที่ถูกออกแบบเพื่อให้สามารถจัดเก็บ ปรับปรุงแก้ไข ดูแลจัดการ วิเคราะห์ และแสดงข้อมูล ในรูปแบบสารสนเทศที่อ้างอิงกับภูมิศาสตร์ ได้อย่างมีประสิทธิภาพ”

ศูนย์วิจัยภูมิสารสนเทศ เพื่อทรัพยากรธรรมชาติ สิ่งแวดล้อม เศรษฐกิจ และสังคมของประเทศไทย ให้ความหมายว่า [สมบัติ, 2004] “ระบบสารสนเทศภูมิศาสตร์ คือ กระบวนการทำงานเกี่ยวกับข้อมูลในเชิงพื้นที่ด้วยระบบคอมพิวเตอร์ ที่ใช้กำหนดข้อมูลและสารสนเทศ ที่มีความสัมพันธ์กับตำแหน่งในเชิงพื้นที่ เช่น ที่อยู่ บ้านเลขที่ สัมพันธ์กับตำแหน่งในแผนที่ เส้นรุ้ง เส้นแวง”

นอกจากนี้ยังมีแหล่งความรู้ทางด้านระบบสารสนเทศภูมิศาสตร์อีกมากมายที่ได้กล่าว ถึงนิยามของระบบสารสนเทศภูมิศาสตร์ ซึ่งพอจะสรุปได้ว่า ระบบสารสนเทศภูมิศาสตร์เป็นเทคโนโลยีที่เกี่ยวข้องกับกระบวนการทำงานกับข้อมูลเชิงพื้นที่ ที่มีการอ้างอิงถึงระบบพิกัดภูมิศาสตร์ และในปัจจุบันมีการนำระบบคอมพิวเตอร์มาอำนวยความสะดวกในการนำเข้า จัดเก็บ วิเคราะห์ และแสดงผลข้อมูลในรูปแบบที่เข้าใจได้ง่าย

ระบบสารสนเทศภูมิศาสตร์จะทำงานได้อย่างสมบูรณ์ต้องประกอบด้วยองค์ประกอบ 6 อย่างด้วยกัน คือ ฮาร์ดแวร์ (Hardware) ซอฟต์แวร์ (Software) เทคโนโลยีการสื่อสารข้อมูลและระบบเครือข่าย (Data Communication and Network) ข้อมูล (Data) บุคลากร (People) และกระบวนการดำเนินงาน (Process)

1. ฮาร์ดแวร์ คือ อุปกรณ์ต่างๆที่ใช้ในการทำงาน ตั้งแต่ การนำเข้า การจัดเก็บ การประมวลผล และการแสดงข้อมูล เช่น Digitizer, Scanner, Computer, GPS (Global Positioning System), Plotter และ เครื่องพิมพ์ เป็นต้น

2. ซอฟต์แวร์ คือ โปรแกรมคอมพิวเตอร์ที่ใช้ดำเนินการเกี่ยวกับข้อมูลภูมิศาสตร์ ได้แก่ โปรแกรม ArcInfo, ArcView หรือ MapInfo เป็นต้น โปรแกรมระบบจัดการฐานข้อมูล เช่น Microsoft Access, Oracle, Informix และ DB2 เป็นต้น โปรแกรมระบบปฏิบัติการต่างๆ เช่น Windows และ UNIX เป็นต้น รวมถึงโปรแกรมช่วยงานต่างๆ (Utilities) เช่น Notepad และ Editor อื่นๆ

3. เทคโนโลยีการสื่อสารข้อมูล ระบบเครือข่าย และอินเตอร์เน็ต คือ เทคโนโลยีที่ใช้เพื่อการถ่ายโอนหรือแลกเปลี่ยนข้อมูลกันระหว่างผู้ส่งและผู้รับ โดยผ่านช่องทางสื่อสาร เช่น อุปกรณ์อิเล็กทรอนิกส์ หรือคอมพิวเตอร์ เพื่อให้ผู้ส่งและผู้รับเกิดความเข้าใจซึ่งกันและกัน ซึ่งในปัจจุบัน นิยมนำคอมพิวเตอร์มาเชื่อมต่อกันด้วยระบบเครือข่าย และอินเตอร์เน็ต

4. ข้อมูล คือ ข้อมูลที่จะนำเข้า ข้อมูลเหล่านี้อาจอยู่ในรูปของแผนที่ หรือได้จากการทำสำรวจภาคพื้นดิน หรือข้อมูลที่ได้มาจากการแปลงข้อมูลของโปรแกรมอื่น ประกอบด้วย 2 ส่วนหลักที่เชื่อมโยงกัน คือ ข้อมูลเชิงพื้นที่ และข้อมูลคุณลักษณะ ซึ่งจะอธิบายรายละเอียดในหัวข้อ 2.2

5. บุคลากร หมายถึง บุคลากรที่มีความรู้พื้นฐานทางด้านคอมพิวเตอร์ และระบบสารสนเทศภูมิศาสตร์ ซึ่งรวมถึงผู้เชี่ยวชาญเฉพาะด้านที่ทำหน้าที่ในการออกแบบ และดูแลรักษาระบบ ไปจนกระทั่งผู้ใช้งานทั่วไปที่สามารถนำระบบไปประยุกต์ใช้งาน

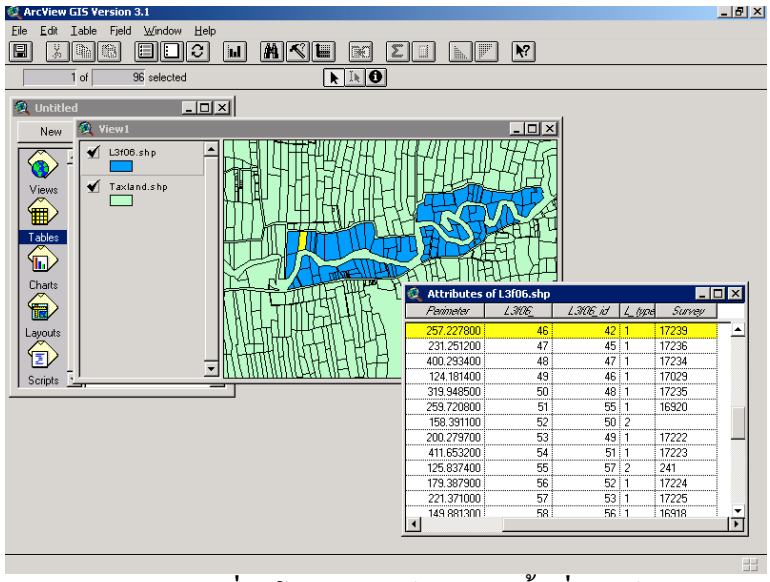
6. กระบวนการดำเนินงาน ประกอบด้วยหลายขั้นตอนการทำงานของระบบสารสนเทศ เช่น การออกแบบ และวางแผนพัฒนา นอกจากนี้ยังมีข้อกำหนดต่างๆที่แตกต่างกันไปของแต่ละหน่วยงาน/องค์กร ระบบสารสนเทศที่ทำงานได้ดีจะต้องผ่านการวางแผนและออกแบบ มีกฎเกณฑ์ซึ่งเป็นแบบจำลองการทำงานที่เหมาะสมสำหรับแต่ละองค์กร

2.2 ข้อมูลในระบบสารสนเทศภูมิศาสตร์

สิ่งที่ทำให้ระบบสารสนเทศภูมิศาสตร์แตกต่างจากระบบสารสนเทศโดยทั่วไป คือ ข้อมูล เนื่องจาก ระบบสารสนเทศภูมิศาสตร์สามารถจัดการกับข้อมูลเชิงพื้นที่ เชื่อมโยงกับข้อมูลคุณลักษณะได้ ดังตัวอย่างในภาพประกอบ 2-1 เพราะฉะนั้น ข้อมูลในระบบสารสนเทศภูมิศาสตร์แบ่งเป็น 2 ประเภท คือ ข้อมูลเชิงพื้นที่ และข้อมูลคุณลักษณะ

ข้อมูลเชิงพื้นที่ เป็นข้อมูลใด ๆที่เกี่ยวกับตำแหน่ง (Location) ที่มีความสัมพันธ์กับ รูปร่าง (Shape) ของคุณลักษณะทางภูมิศาสตร์ ข้อมูลเชิงพื้นที่แบ่งได้เป็น 3 ประเภท คือ จุด (Point) เส้น (Line) และรูปหลายเหลี่ยมปิด (Polygon) โดยข้อมูลประเภทจุด ใช้อ้างอิงตำแหน่งที่ตั้งสิ่งต่าง ๆ เช่น ที่ตั้งโรงเรียน ที่ตั้งเสา หม้อแปลง ข้อมูลประเภทเส้น ใช้แทนลักษณะของสิ่งที่แคบและยาว เช่น ถนน แม่น้ำ ส่วนข้อมูลประเภทรูปหลายเหลี่ยมปิด ใช้แสดงลักษณะของสิ่งที่เป็นบริเวณ เช่น เขตปกครอง (ประเทศ จังหวัด อำเภอ เป็นต้น) แหล่งน้ำ พื้นที่ป่า เป็นต้น [ESRI, 2005]

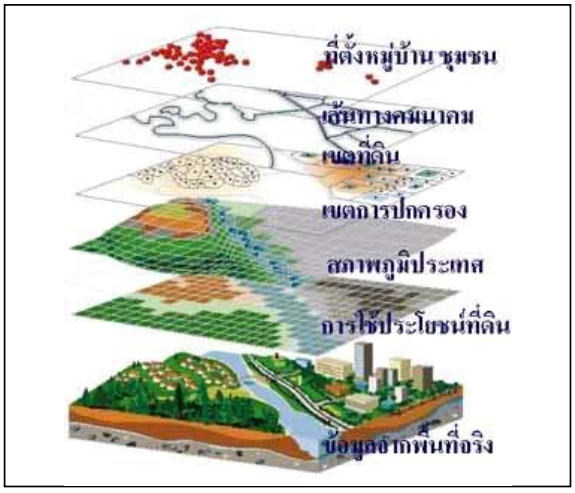
ข้อมูลคุณลักษณะ เป็นข้อมูลที่บรรยายคุณลักษณะทางภูมิศาสตร์ของข้อมูลเชิงพื้นที่ ซึ่งอาจเป็นข้อมูลเชิงคุณภาพหรือข้อมูลเชิงปริมาณ เช่น ชื่อถนน ชื่อจังหวัด จำนวนประชากร ประเภทแหล่งน้ำ ซึ่งมักจะเก็บข้อมูลไว้ในฐานข้อมูล [ESRI, 2005]



ภาพประกอบ 2-1 แสดงการเชื่อมโยงระหว่างข้อมูลเชิงพื้นที่ และข้อมูลคุณลักษณะของโปรแกรม ArcView

เนื่องจากข้อมูลในระบบสารสนเทศภูมิศาสตร์ เป็นข้อมูล “ภูมิศาสตร์” (Geography) ซึ่งพจนานุกรมราชบัณฑิตยสถาน ให้นิยามคำว่า “ภูมิศาสตร์” หมายถึง “วิชาที่ศึกษาถึงความสัมพันธ์ระหว่างสิ่งแวดล้อมทางธรรมชาติกับสังคมที่ปรากฏในดินแดนต่าง ๆ ของโลก” ภาควิชาภูมิศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ได้กล่าวถึง ภูมิศาสตร์ [ภาควิชาภูมิศาสตร์, ม.ป.ป.] ว่า “ภูมิศาสตร์ คือ ศาสตร์ทางด้านพื้นที่และบริเวณต่าง ๆ บนพื้น

ผิวโลก เป็นวิชาที่ศึกษาปรากฏการณ์ทางกายภาพ และมนุษย์ ที่เกิดขึ้น ณ บริเวณที่ทำการศึกษา รวมไปถึงสิ่งแวดล้อมที่อยู่บริเวณโดยรอบ”



ภาพประกอบ 2-2 ชั้นข้อมูลของระบบสารสนเทศภูมิศาสตร์ [จาก http://www.eric.chula.ac.th/gisthai/about-gis/gis.html, 1999-2005]

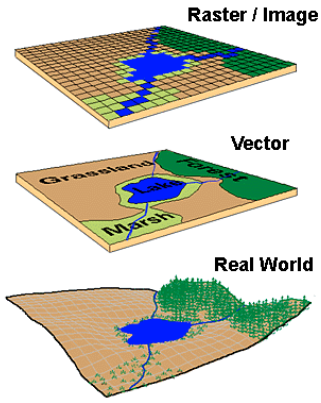
จากนิยามทั้งสองข้างต้นเห็นว่าข้อมูลภูมิศาสตร์มีจำนวนมากมาย การนำข้อมูลภูมิศาสตร์มาประยุกต์ใช้งานในระบบสารสนเทศภูมิศาสตร์ จึงต้องมีการจำแนกเป็นเรื่อง หรือแบ่งเป็นกลุ่ม เรียกว่า ชั้น (Layer) โดยที่แต่ละชั้นจะเป็นข้อมูลเกี่ยวกับเรื่องใดเรื่องหนึ่ง และชั้นข้อมูลสามารถเชื่อมโยงเข้าด้วยกันด้วยระบบภูมิศาสตร์ดังภาพประกอบ 2-2 จากสิ่งต่างๆ ที่เราสนใจบนโลก นำมาจัดเก็บในระบบสารสนเทศภูมิศาสตร์เป็นชั้นข้อมูล 6 ชั้น ได้แก่การใช้ประโยชน์ที่ดิน สภาพภูมิประเทศ เขตการปกครอง เขตที่ดิน เส้นทางคมนาคม และที่ตั้งหมู่บ้านชุมชน

2.3 รูปแบบข้อมูล

ดังที่กล่าวในหัวข้อ 2.2 ว่าข้อมูลในระบบสารสนเทศภูมิศาสตร์ แบ่งเป็น 2 ประเภท คือ ข้อมูลเชิงพื้นที่ และข้อมูลคุณลักษณะ ดังนั้นรูปแบบข้อมูล (Data model) ในระบบสารสนเทศภูมิศาสตร์จึงมีรูปแบบข้อมูลสำหรับการจัดเก็บข้อมูลเชิงพื้นที่ และรูปแบบข้อมูลสำหรับการจัดเก็บข้อมูลคุณลักษณะดังรายละเอียดต่อไปนี้

รูปแบบข้อมูลเชิงพื้นที่

รูปแบบข้อมูลเชิงพื้นที่ (Spatial data model) นิยมแสดงผลออกมาในรูปของแผนที่ และมีรูปแบบการเก็บข้อมูลอยู่ 2 แบบด้วยกันคือ รูปแบบข้อมูลแบบราสเตอร์ (Raster data format) และรูปแบบข้อมูลแบบเวคเตอร์ (Vector data format)



ภาพประกอบ 2-3 แสดงลักษณะข้อมูลบนโลก ข้อมูลแบบ Vector และข้อมูลแบบ Raster [จาก http://www.innovativegis.com/basis/primer/primer.html, 2005]

รูปแบบข้อมูลแบบราสเตอร์ เป็นรูปแบบการจัดเก็บข้อมูลในลักษณะช่องสี่เหลี่ยมจัตุรัส (Square cell หรือ Grid cell) ที่มีขนาดเท่า ๆกัน และต่อเนื่องกันไปตามแนวนอนและแนวตั้ง ดังตัวอย่างในภาพประกอบ 2-3 ในส่วนของ Raster/Image เป็นการแทนข้อมูลจริงที่เป็นแม่น้ำ ทะเลสาบ ป่า และทุ่งหญ้า ด้วยกลุ่มของช่องสี่เหลี่ยมที่ต่อเนื่องกัน จากภาพจะเห็นว่าความแม่นยำของข้อมูลจะขึ้นอยู่กับขนาดของช่องสี่เหลี่ยม นั่นคือถ้าช่องสี่เหลี่ยมมีขนาดเล็ก ก็จะมีความแม่นยำของข้อมูลมากกว่าช่องสี่เหลี่ยมที่มีขนาดใหญ่

รูปแบบข้อมูลแบบเวคเตอร์ เป็นรูปแบบการจัดเก็บข้อมูลในลักษณะเส้นที่มีทิศทาง ประกอบด้วยจุด และเส้นเชื่อมระหว่างจุด กำหนดจุดด้วยพิกัดด้วยคู่ลำดับ X และ Y ใช้สำหรับแสดงคุณลักษณะของข้อมูลเชิงพื้นที่ ตัวอย่างเช่นข้อมูลแบบจุด แทนด้วยคู่ลำดับ 1 คู่ ข้อมูลแบบรูปหลายเหลี่ยมปิด แทนด้วยชุดของคู่ลำดับ โดยคู่ลำดับแรก จะเป็นจุดเดียวกับคู่ลำดับสุดท้าย และเชื่อมโยงระหว่างจุดด้วยเส้น เป็นต้น ตัวอย่างในภาพประกอบ 2-3 ในส่วน Vector แสดงการแทนข้อมูลจริงที่เป็นแม่น้ำ ทะเลสาบ ป่า และทุ่งหญ้า ด้วยชุดของคู่ลำดับ (X,Y) ที่เชื่อมต่อกัน

อย่างไรก็ตามรูปแบบข้อมูลแบบราสเตอร์ และเวคเตอร์ สามารถแปลงไปแปลงมาได้ แต่อาจมีความคลาดเคลื่อนของข้อมูลในระหว่างการแปลงได้

รูปแบบข้อมูลคุณลักษณะ

ข้อมูลคุณลักษณะของระบบสารสนเทศภูมิศาสตร์ สามารถเก็บโดยใช้ ซอฟต์แวร์ทางด้านระบบสารสนเทศภูมิศาสตร์โดยเฉพาะ หรือจะเก็บแยกต่างหากในระบบจัดการฐานข้อมูล ข้อมูลคุณลักษณะมีรูปแบบในการเก็บข้อมูลได้หลากหลายรูปแบบแตกต่างกันไป ได้แก่ แบบลำดับขั้น แบบเครือข่าย แบบเชิงสัมพันธ์ แบบเชิงวัตถุ และแบบความสัมพันธ์เชิงวัตถุ

รูปแบบข้อมูลลำดับขั้น (Hierarchical data model) เป็นรูปแบบข้อมูลที่จัดความสัมพันธ์ระหว่างข้อมูลเป็นระดับ ๆ เหมือนกิ่งก้านสาขาของต้นไม้ แต่เป็นลักษณะต้นไม้คว่ำ ถ้าอธิบายความสัมพันธ์ระหว่างข้อมูลสองระดับในรูปแบบความสัมพันธ์ระหว่างแม่และลูก กล่าวได้ว่าแต่ละระดับสามารถมีลูกได้อย่างไม่จำกัด แต่ลูกจะต้องมีแม่เพียงหนึ่งเท่านั้น ระบบจัดการฐานข้อมูลที่ใช้รูปแบบข้อมูลลำดับขั้น ไม่เป็นที่ยอมรับสำหรับการใช้เก็บข้อมูลระบบสารสนเทศภูมิศาสตร์เนื่องจาก โครงสร้างข้อมูลแบบลำดับขั้นเหมาะสมกับข้อมูลที่ไม่ค่อยมีการเปลี่ยนแปลงมากนัก และการที่มีแม่ได้เพียงหนึ่งนั้นก็ไม่เหมาะกับการเก็บข้อมูลทางภูมิศาสตร์

รูปแบบข้อมูลเครือข่าย (Network data model) เป็นรูปแบบข้อมูลที่คล้ายกับรูปแบบข้อมูลลำดับขั้น แต่ต่างกันที่ลูกสามารถมีแม่ได้มากกว่าหนึ่ง แต่ระบบจัดการฐานข้อมูลที่ใช้รูปแบบข้อมูลแบบเครือข่ายก็ยังไม่พบว่าได้รับการยอมรับในการนำมาใช้กับระบบสารสนเทศภูมิศาสตร์ มากกว่าระบบจัดการฐานข้อมูลที่ใช้โครงสร้างแบบลำดับขั้น เนื่องจากมีข้อจำกัดในเรื่องความยืดหยุ่นของรูปแบบข้อมูลเช่นเดียวกับรูปแบบข้อมูลลำดับขั้น และมีความซับซ้อนในการใช้งาน

รูปแบบข้อมูลเชิงสัมพันธ์ (Relational data model) จัดรูปแบบข้อมูลในรูปของตารางสองมิติ โดยแต่ละตารางประกอบด้วยแถวและสดมภ์ แต่ละสดมภ์ภายในตารางเดียวกันต้องมีชื่อไม่ซ้ำกัน ภายในสดมภ์เก็บค่าซึ่งเป็นสิ่งอธิบายคุณลักษณะแต่ละอย่าง ในระบบสารสนเทศภูมิศาสตร์ แต่ละแถวจะสัมพันธ์กับลักษณะข้อมูลของข้อมูลเชิงพื้นที่แต่ละชิ้น ข้อมูลสามารถเก็บไว้ในหลาย ๆตารางได้ แต่ละตารางสามารถเชื่อมโยงหรืออ้างอิงถึงกันได้ผ่านสดมภ์ที่เก็บข้อมูลความหมายเดียวกันและคุณลักษณะเดียวกัน ระบบจัดการฐานข้อมูลเชิงสัมพันธ์เป็นที่นิยมนำมาใช้จัดการข้อมูลคุณลักษณะของระบบสารสนเทศภูมิศาสตร์มากที่สุด เนื่องจากเป็นโครง

สร้างที่เข้าใจได้ง่าย มีความยืดหยุ่นสูง มีประสิทธิภาพในการจัดเก็บ และสามารถสืบค้นข้อมูลได้คล้ายธรรมชาติของมนุษย์ทำให้ระบบจัดการฐานข้อมูลเชิงสัมพันธ์เป็นที่นิยมใช้สำหรับจัดการข้อมูลทั้งที่เกี่ยวข้อง และที่ไม่เกี่ยวข้องกับระบบสารสนเทศภูมิศาสตร์ อย่างไรก็ตามในปัจจุบันเทคโนโลยีก้าวหน้าขึ้น จึงได้มีความต้องการจัดเก็บชนิดข้อมูลที่หลากหลายขึ้น นั่นคือ ข้อมูลแบบวัตถุที่มีความซับซ้อน เช่น แผนที่ ซึ่งประกอบด้วยจุด เส้น และรูปหลายเหลี่ยมปิด ประกอบกับได้มีแนวคิดเชิงวัตถุ (Object-Oriented) ของวิศวกรรมซอฟต์แวร์ที่เป็นที่ยอมรับ และนิยมนำไปใช้เพื่อการออกแบบ และการโปรแกรมกันอย่างแพร่หลาย ซึ่งรูปแบบเชิงสัมพันธ์ไม่สามารถรองรับความต้องการนี้ได้

รูปแบบข้อมูลเชิงวัตถุ (Object-Oriented data model) จัดรูปแบบข้อมูลโดยมองชิ้นส่วนเป็นวัตถุ แต่ละวัตถุประกอบด้วยข้อมูล และการทำงาน เป็นรูปแบบที่ออกแบบขึ้นเพื่อให้สอดคล้องกับแนวทางการโปรแกรมเชิงวัตถุ แต่ยังไม่นิยมนำมาใช้กับระบบสารสนเทศภูมิศาสตร์เนื่องจาก รูปแบบข้อมูลเชิงวัตถุทำให้มีการผูกติดการออกแบบฐานข้อมูลเข้ากับการโปรแกรมเชิงวัตถุมากเกินไป ทำให้การบำรุงรักษาและขยายระบบเป็นไปด้วยความยากลำบาก นอกจากนี้การทำงานด้วยภาษา SQL ยังต้องการพื้นที่ในการประมวลผลมาก ทำให้ประสิทธิภาพในการทำงานต่ำ เนื่องจากการสอบถามข้อมูลจะซับซ้อนมาก ซึ่งถึงแม้ว่าจะสามารถซื้ออุปกรณ์คอมพิวเตอร์เพิ่มเติมได้ แต่เวลาที่ใช้ในการคำนวณยังคงมากอยู่ [Brown, 2001]

รูปแบบข้อมูลความสัมพันธ์เชิงวัตถุ (Object-Relational data model) เป็นการเพิ่มความสามารถเชิงวัตถุให้กับรูปแบบข้อมูลเชิงสัมพันธ์ โดยที่ยังมีคุณลักษณะพื้นฐานของรูปแบบข้อมูลเชิงสัมพันธ์ครบถ้วน ทำให้รูปแบบข้อมูลความสัมพันธ์เชิงวัตถุสามารถรองรับข้อมูลทั่วไปและข้อมูลวัตถุที่ซับซ้อน เช่นข้อมูลภูมิศาสตร์เป็นต้น รูปแบบข้อมูลความสัมพันธ์เชิงวัตถุมีข้อดีคือรวมเอาจุดเด่นของรูปแบบข้อมูลเชิงสัมพันธ์และรูปแบบข้อมูลเชิงวัตถุเข้าด้วยกัน อีกทั้งองค์กรที่ใช้งานระบบจัดการฐานข้อมูลเชิงสัมพันธ์เดิม สามารถเปลี่ยนมาใช้ระบบจัดการฐานข้อมูลแบบความสัมพันธ์เชิงวัตถุไม่ยากนัก [Brown, 2001]

2.4 Shapefiles

ในระบบสารสนเทศภูมิศาสตร์มีรูปแบบการจัดเก็บข้อมูลแบบหนึ่ง เรียกว่า Shapefiles เป็นรูปแบบที่มีการจัดเก็บข้อมูลเชิงพื้นที่ในรูปแบบข้อมูลแบบเวคเตอร์ และจัดเก็บข้อมูลคุณลักษณะในรูปแบบข้อมูลเชิงสัมพันธ์ ซึ่งเป็นที่นิยมใช้เพื่อแลกเปลี่ยน หรือ ถ่ายโอนข้อมูลระหว่างซอฟต์แวร์ ในเอกสารของบริษัทอีเอสอาร์ไอ [An ESRI White Paper, 1998] ได้

อธิบายว่า Shapefiles เป็นการจัดเก็บข้อมูลรูปแบบหนึ่งที่สามารถจัดเก็บข้อมูลในลักษณะ จุด เส้น และพื้นที่ Shapefiles หนึ่งๆ ประกอบด้วยแฟ้มข้อมูลอย่างน้อย 3 แฟ้ม คือ แฟ้มข้อมูลหลัก (Main file) แฟ้มดัชนี (Index file) และแฟ้มตารางดีเบส (dBASE table)

การตั้งชื่อแฟ้มข้อมูลทั้งสามนี้จะต้องใช้ชื่อหน้า ชื่อเดียวกัน กฎเกณฑ์การตั้งชื่อคล้ายกับการตั้งชื่อแฟ้มข้อมูลชนิดอื่นๆ นั่นคือจะต้องขึ้นต้นด้วยอักขระภาษาอังกฤษ a-Z หรือตัวเลข 0-9 ตามด้วยตัวเลขที่มากกว่า 0 หรืออักขระ a-Z เลข 0-9 อักขระพิเศษ _ - ส่วนชื่อท้ายของแฟ้มทั้งสามจะแตกต่างกัน ตัวอย่าง แฟ้มรูปทรงที่มีชื่อต้นว่า countries ประกอบด้วย

แฟ้มข้อมูลหลัก	countries.shp
แฟ้มดัชนี	countries.shx
แฟ้มตารางดีเบส	countries.dbf

แฟ้มข้อมูลหลัก เก็บข้อมูลรูปร่างเรขาคณิต ของข้อมูลเชิงพื้นที่โดย 1 รูปอธิบายด้วย 1 ระเบียน แฟ้มดัชนีเก็บระยะที่ห่างจากจุดเริ่มต้น (Offset) ของระเบียนในแฟ้มข้อมูลหลักที่เกี่ยวข้องกัน และแฟ้มตารางดีเบส เก็บข้อมูลคุณลักษณะที่มีความสัมพันธ์กับรูปร่างที่เก็บไว้ในแฟ้มข้อมูลหลัก โดยความสัมพันธ์ระหว่างระเบียนที่เก็บข้อมูลรูปร่างในแฟ้มข้อมูลหลัก และระเบียนที่เก็บคุณลักษณะในแฟ้มตารางดีเบส เป็นแบบหนึ่งต่อหนึ่ง ในปัจจุบัน นิยมเปลี่ยนการจัดเก็บข้อมูลคุณลักษณะจากตารางดีเบสไปเป็น ตารางของซอฟต์แวร์ระบบจัดการฐานข้อมูลเชิงสัมพันธ์ เช่น Microsoft Access, Oracle หรือ MS SQL Server เป็นต้น

แฟ้มรูปทรง สามารถสร้างขึ้นจากวิธีใดวิธีหนึ่งใน 4 วิธีต่อไปนี้

1. การส่งข้อมูลออก (Export) จากแหล่งข้อมูลโดยใช้ซอฟต์แวร์ ARC/INFO, PC ARC/INFO, Spatial Database Engine, ArcView GIS หรือซอฟต์แวร์ด้านแผนที่อื่นๆ
2. การดิจิไตซ์ (Digitize) คือการใช้เครื่องมือในการดิจิไตซ์รูปร่าง (Degitize shape)โดยใช้เครื่องมือที่มีอยู่ในซอฟต์แวร์อาร์คอินโฟร์
3. การโปรแกรม (Programming) เพื่อสร้างแฟ้มรูปทรงโดยใช้ภาษาการโปรแกรม เช่น ภาษา Avenue ในซอฟต์แวร์ ArcView GIS, MapObjects, ARC Macro Language (AML) ใน Simple Macro Language (SML) ใน PC Arc/Info
4. เขียนโดยตรงลงในแฟ้มข้อมูลที่เก็บรายละเอียดของ Shapefiles (Shapefiles specifications)

2.5 Oracle Spatial

ระบบจัดการฐานข้อมูล Oracle ตั้งแต่รุ่น 9i เป็นต้นไป จะมีส่วนหนึ่งสำหรับจัดการข้อมูลเชิงพื้นที่เรียกว่า Oracle Spatial โดยอาศัยรูปแบบข้อมูลความสัมพันธ์เชิงวัตถุ เรียกว่า SDO (Spatial Data Object) จากคู่มือที่อธิบายในเรื่อง Oracle Spatial ชื่อ “Oracle Spatial User’s Guide and Reference Release 9.0.1” [Murray, 2001] ได้อธิบายถึง SDO ว่า SDO ประกอบด้วย 4 ส่วนด้วยกันคือ

1. วิธีในการจัดเก็บข้อมูลเชิงพื้นที่
2. วิธีในการจัดทำดัชนีข้อมูลเชิงพื้นที่
3. กลุ่มการทำงานที่ทำหน้าที่ในการสอบถามข้อมูลเชิงพื้นที่
4. เครื่องมือในการจัดการข้อมูลเชิงพื้นที่

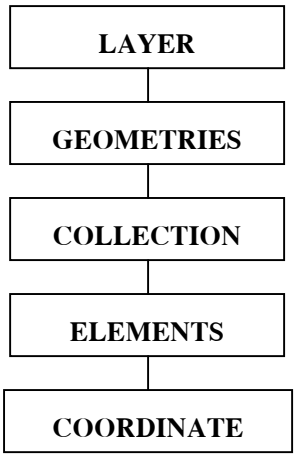
2.5.1 การจัดเก็บข้อมูลเชิงพื้นที่

ข้อมูลเชิงพื้นที่ที่ Oracle Spatial จัดเก็บ มีชนิดข้อมูลเป็น SDO_ GEOMETRY โดยเก็บรูปเรขาคณิตแต่ละรูปไว้ในแถวเดียว รูปแบบข้อมูลเป็นแบบโครงสร้างลำดับขั้น ดังแสดงในภาพประกอบ 2-4

Coordinate data คือข้อมูลคู่ลำดับ (X,Y) หรือข้อมูลละติจูด ลองจิจูด (latitude, longitude)

Elements เป็นหน่วยพื้นฐานของรูปเรขาคณิต ตัวอย่างข้อมูลได้แก่

- จุด จัดเก็บด้วยข้อมูลคู่ลำดับ 1 คู่
- เส้น ประกอบขึ้นจากชุดของคู่ลำดับ ตั้งแต่ 2 คู่ขึ้นไป
- รูปหลายเหลี่ยมปิด ประกอบขึ้นจากชุดของคู่ลำดับตั้งแต่ 2 คู่ขึ้นไป โดยคู่ลำดับคู่แรก มีค่าเท่ากับคู่ลำดับคู่สุดท้าย ในกรณีที่เป็นรูปหลายเหลี่ยมปิดที่มีช่องว่างตรงกลาง เช่น ทะเลสาบที่อยู่ในเกาะ จะแทนด้วย รูปหลายเหลี่ยมปิดสองรูปซ้อนทับกัน เรียกว่า รูปหลายเหลี่ยมเชิงซ้อน (Complex polygon)



ภาพประกอบ 2-4 โครงสร้างข้อมูลชนิด SDO_ GEOMETRY [จาก Oracle Spatial, An Oracle Technical White Paper, 2002 page 7]

Collection คือ กลุ่มของข้อมูลชนิดเดียวกัน เช่นกลุ่มของข้อมูลชนิดจุด กลุ่มของข้อมูลชนิดเส้น หรือกลุ่มข้อมูลชนิดรูปหลายเหลี่ยมปิด ในรูปเรขาคณิตรูปใดรูปหนึ่ง

Geometies เป็นวัตถุที่แสดงสมบัติของข้อมูลเชิงพื้นที่ โดยจำลองด้วยชุดลำดับของอิลิเมนต์พื้นฐาน ซึ่งเรขาคณิตแต่ละรูปอาจจะประกอบด้วยอิลิเมนต์เพียงอิลิเมนต์เดียว หรือจะประกอบด้วยหลาย ๆอิลิเมนต์ ที่เป็นอิลิเมนต์ประเภทเดียวกัน เช่น รูปหลายเหลี่ยมปิดหลาย ๆ รูป แทนกลุ่มของเกาะ หรือประกอบด้วยอิลิเมนต์ต่างประเภทกันก็ได้ เช่น รูปหลายเหลี่ยมปิดและจุด อยู่ด้วยกัน ตัวอย่างการแทนพื้นที่จริงก็เช่น รูปหลายเหลี่ยมปิดเชิงซ้อนแทนเมือง เมืองหนึ่งที่มีทะเลสาบและการแบ่งเขตการปกครองอยู่ข้างใน

Layers คือ ชั้นข้อมูล หมายถึง ชุดของรูปเรขาคณิต ที่แสดงคุณลักษณะแบบเดียวกัน หรือแสดงเรื่องเดียวกัน ดังภาพประกอบ 2-2 เช่น ข้อมูลของระบบสารสนเทศภูมิศาสตร์ชุดหนึ่ง ประกอบด้วย 3 ชั้นข้อมูล ชั้นข้อมูลแรกเกี่ยวข้องกับเขตการปกครอง ชั้นข้อมูลที่สอง แสดงถึงความหนาแน่นของประชากร และชั้นข้อมูลที่สามอธิบายถึงโครงข่ายถนนและสะพาน เป็นต้น ชั้นข้อมูลแต่ละชั้น รวมทั้งการจัดทำดัชนีเชิงพื้นที่ถูกจัดเก็บในฐานข้อมูลในรูปแบบของตารางมาตรฐาน

2.5.2 ชนิดข้อมูล สำหรับจัดเก็บข้อมูลเชิงพื้นที่

การเก็บข้อมูลใน Oracle Spatial จะเก็บชั้นข้อมูลไว้ในตาราง โดยชั้นข้อมูลหนึ่งอาจจะประกอบด้วยหนึ่งตาราง หรือหลายตาราง ส่วนรูปเรขาคณิตหนึ่งรูป จะเก็บไว้ในหนึ่งเซลล์ คือหนึ่งแถวและหนึ่งสดมภ์ของตาราง โดยสดมภ์ที่เก็บรูปเรขาคณิตจะมีชนิดข้อมูลเป็น MDSYS.SDO_ GEOMETRY

ชนิดข้อมูล MDSYS.SDO_GEOMETRY เป็นชนิดข้อมูลสำหรับจัดเก็บข้อมูลชนิด จุด เส้น และรูปหลายเหลี่ยมปิด หรืออาจเป็นกลุ่มของข้อมูลเหล่านี้ โดย Oracle Spatial มีการกำหนดข้อมูลชนิด MDSYS.SDO_GEOMETRY ดังตัวอย่างที่แสดงในภาพประกอบ 2-5 เป็นการสร้างวัตถุชื่อ SDO_GEOMETRY เพื่อรองรับการจัดเก็บข้อมูลเชิงพื้นที่

SDO_GTYPE เป็นตัวระบุชนิดของรูปเรขาคณิต เก็บเป็นตัวเลข 4 หลักในลักษณะ dltt โดย d คือจำนวนมิติของวัตถุ l คือระบบการอ้างอิงมิติการวัดแบบเชิงเส้น สำหรับวัตถุ 3 มิติ และ tt คือชนิดของรูปเรขาคณิต

SDO_SRID เป็นรหัสที่บอกถึงระบบพิกัดที่ใช้ในการอ้างอิงเชิงพื้นที่ เพื่อให้สัมพันธ์กับภูมิศาสตร์ ถ้าเป็นค่าว่าง (null) แสดงว่าไม่มีระบบพิกัดที่สัมพันธ์กับภูมิศาสตร์

SDO_POINT สำหรับกำหนดชนิดวัตถุที่เป็นแบบจุด และถูกกำหนดโดยใช้ชนิดวัตถุ ที่มีค่าของ X, Y และ Z ทุกตัวเป็นแบบตัวเลข (Number) โดยมีเงื่อนไขว่า วัตถุจะเป็นชนิดจุด ก็ต่อเมื่อค่าของ SDO_ELEM_INFO และ SDO_ORDINATES array เป็นค่าว่าง และค่าของ SDO_POINT ไม่เป็นค่าว่าง ถ้าเป็นกรณีอื่น ๆ ค่าของ SDO_POINT จะไม่มีความหมาย

```
CREATE TYPE SDO_GEOMETRY
AS OBJECT (
    SDO_GTYPE        NUMBER,
    SDO_SRID         NUMBER,
    SDO_POINT        SDO_POINT_TYPE,
    SDO_ELEM_INFO    MDSYS.SDO_ELEM_INFO_ARRAY,
    SDO_ORDINATES    MDSYS.SDO_ORDINATE_ARRAY);
```

ภาพประกอบ 2-5 Attributes of Oracle Spatial SDO

SDO_ELEM_INFO สำหรับกำหนดอาเรย์ขนาดต่าง ๆกัน ทำให้รู้ว่าจะแปลคู่ลำดับที่เก็บใน SDO_ORDINATES อย่างไร ประกอบด้วย

- SDO_START_OFFSET ค่าที่แสดงถึงระยะห่างภายในอาเรย์ของ SDO_ORDINATES ที่คู่ลำดับแรกของหน่วยข้อมูลนี้เก็บอยู่
- SDO_ETYPE เป็นค่าที่แสดงถึงชนิดของหน่วยข้อมูล แบ่งออกเป็น 2 กลุ่มใหญ่ ๆ คือ กลุ่มที่เป็นหน่วยข้อมูลพื้นฐาน (Simple element) ได้แก่ จุด เส้น รูปหลายเหลี่ยมปิดและกลุ่มที่เป็นหน่วยข้อมูลเชิงซ้อน (Compound element) ได้แก่เส้นเชิงซ้อน (Compound line string) วงแหวนรอบนอกของรูปหลายเหลี่ยมเชิงซ้อน (Exterior polygon ring) ซึ่งต้องกำหนดคู่ลำดับแบบทวนเข็มนาฬิกา วงแหวนรอบในของรูปหลายเหลี่ยมปิด (Interior polygon ring) กำหนดคู่ลำดับแบบตามเข็มนาฬิกา
- SDO_INTERPRETATION มีความหมายได้ 2 แบบขึ้นอยู่กับ SDO_ETYPE ว่าเป็นหน่วยข้อมูลแบบใด ถ้า SDO_ETYPE เป็นหน่วยข้อมูลพื้นฐาน จะแสดงถึงการจัดลำดับของคู่ลำดับของหน่วยข้อมูล และคู่ลำดับสุดท้ายจะน้อยกว่า ระยะของคู่ลำดับเริ่มต้นของหน่วยข้อมูลถัดไปอยู่ 1 หน่วยข้อมูล ถ้า SDO_ETYPE เป็นหน่วยข้อมูลเชิงซ้อน จะแสดงถึงจำนวนคู่ลำดับของหน่วยข้อมูลนั้น

SDO_ORDINATES เป็นอาเรย์ชนิดตัวเลข สำหรับเก็บค่าของคู่ลำดับที่ประกอบกันเป็นเส้นรอบวงของวัตถุ โดยอาเรย์เหล่านี้จะใช้ร่วมกันกับ SDO_ELEM_INFO และค่าในอาเรย์จะถูกเรียงลำดับด้วยมิติ เช่นรูปหลายเหลี่ยมปิดที่ประกอบด้วยจุดแบบ 2 มิติ 4 จุด จะถูกเก็บเป็น {X1, Y1, X2, Y2, X3, Y3, X4, Y4, X1, Y1} แต่ถ้าเป็นรูปหลายเหลี่ยมปิดที่ประกอบด้วยจุดแบบ 3 มิติ 4 จุด จะถูกเก็บเป็น {X1, Y1, Z1, X2, Y2, Z2, X3, Y3, Z3, X4, Y4, Z4, X1, Y1, Z1} อย่างไรก็ตามเนื่องจาก Oracle Spatial รองรับข้อมูลเชิงพื้นที่ได้เพียง 2 มิติ ดังนั้นการสร้างดัชนีเชิงพื้นที่ใน Oracle Spatial จะไม่สนใจค่า Z

2.5.3 ข้อมูลที่เกี่ยวข้องกับข้อมูลเชิงพื้นที่

การที่ Oracle Spatial สามารถอ่าน และวิเคราะห์ข้อมูลได้จะต้องอาศัยการเก็บข้อมูลเกี่ยวกับข้อมูลเชิงพื้นที่ (Metadata) ซึ่งเกี่ยวข้องกับ ระบบพิกัด (Coordinate System) และค่าความเชื่อมั่น (Tolerance)

ระบบพิกัด หรือ ระบบอ้างอิงเชิงพื้นที่ (Spatial Reference System) เป็นการกำหนดคู่ลำดับให้กับตำแหน่ง และสร้างความสัมพันธ์ระหว่างชุดของคู่ลำดับเหล่านั้น ทำให้สามารถแทนตำแหน่งในโลกด้วยชุดของคู่ลำดับเหล่านั้นได้ ข้อมูลเชิงพื้นที่ใด ๆ ก็จะมีระบบพิกัดที่เกี่ยวข้องด้วย โดยที่ระบบพิกัดนั้นอาจจะแทนตำแหน่งบนพื้นโลกได้ (Georeferenced) หรือแทนตำแหน่งบนพื้นโลกไม่ได้ ถ้าระบบพิกัดนั้นสามารถแทนตำแหน่งบนพื้นโลกได้ ก็จะมีหน่วยการวัดที่เกี่ยวข้องกับระบบพิกัดนั้น ๆ โดยอัตโนมัติ เช่น เมตร ข้อมูลเกี่ยวกับระบบพิกัดทั้งหมดเก็บไว้ในตาราง MDSYS.CS_SRS ซึ่งผู้ใช้งานสามารถเพิ่มระบบพิกัดของตนเองเข้าไปได้ โดยมีการอธิบายคำนิยามของระบบพิกัดในรูปแบบ WKT (Well Known Text) ซึ่งรูปแบบในการอธิบายกำหนดโดย OGC (Open GIS Consortium) และมีการกำหนดรหัสในการอ้างอิงที่ไม่ซ้ำกัน เรียกว่า SRID

ค่าความเชื่อมั่น เป็นค่าสำหรับการวัดระดับความแม่นยำของข้อมูลเชิงพื้นที่ มีค่าระหว่าง 0 ถึง 1 หมายถึง เศษส่วนของหน่วยของระยะทางที่ใช้งาน (ถ้ามีการกำหนดระบบพิกัด หน่วยของระยะทางจะถูกกำหนดให้กับระบบนั้นโดยอัตโนมัติ) ตัวอย่างเช่น ค่าความน่าเชื่อถือเป็น 0.005 หมายถึงความน่าเชื่อถือเป็น 1/200 ของหน่วยของระยะทาง โดยที่ถ้าเป็นข้อมูลทางภูมิศาสตร์ (ข้อมูลที่มีการกำหนดพิกัดด้วยค่าลองจิจูดและละติจูด) ค่าความเชื่อมั่นจะมีหน่วยเป็นเมตร และไม่ว่าในกรณีใดก็ตาม ตัวเลขของค่าความเชื่อมั่นยิ่งน้อย แสดงว่าความแม่นยำของข้อมูลยิ่งมากขึ้น

ข้อมูลที่ต้องเก็บสำหรับสดมภ์ที่มีชนิดข้อมูลเป็น SDO_GEOMETRY ที่สร้างขึ้นในตารางใด ๆ จะต้องมีข้อมูลเกี่ยวกับในตาราง USER_SDO_METADATA ดังนี้

1. ชื่อตารางที่มีสดมภ์ มีชนิดข้อมูลเป็น SDO_GEOMETRY
2. ชื่อสดมภ์ที่เก็บข้อมูลแบบ SDO_GEOMETRY
3. จำนวนมิติ สำหรับสดมภ์ SDO_GEOMETRY
4. ขอบเขตต่ำสุด และสูงสุดสำหรับแต่ละแกน หรือมิติ โดยทั่วไปถ้าเป็นข้อมูลภูมิศาสตร์ จะหมายถึงค่าลองจิจูด และละติจูด ขอบเขตก็จะเป็น (-180,180) และ (-90, 90) ตามลำดับ
5. ค่าความเชื่อมั่น สำหรับแต่ละมิติ ซึ่งโดยทั่วไปมักจะใช้ค่าเดียวกันสำหรับทุก ๆแกนหรือมิติ
6. รหัสในการอ้างอิงข้อมูลเชิงพื้นที่ (SRID)

2.5.4 การนำเข้าข้อมูลเชิงพื้นที่

เพื่อให้การสอบถามข้อมูลเชิงพื้นที่มีประสิทธิภาพ การนำเข้าข้อมูลเชิงพื้นที่ประกอบด้วยสองขั้นตอน คือ ขั้นตอนการนำเข้าข้อมูลเชิงพื้นที่ และขั้นตอนการปรับปรุงดัชนีเชิงพื้นที่ การนำเข้าข้อมูล ทำได้ 2 แนวทาง คือการนำเข้าครั้งละมาก ๆ และการนำเข้าเป็นรายการ

การนำเข้าข้อมูลครั้งละมาก ๆ อาศัยโปรแกรมช่วยงานชื่อ SQL*Loader ซึ่งเป็นโปรแกรมช่วยงานที่สามารถนำเข้าข้อมูลคุณลักษณะ โดยการนำเข้าสามารถเลือกได้ว่าจะนำเข้าไปแทนที่หรือเพิ่มในตารางฐานข้อมูล การนำเข้าข้อมูลด้วยวิธีนี้ควรจะทำการลบดัชนี หรือยกเลิกการใช้งานดัชนีก่อน เพื่อให้การนำเข้าทำได้รวดเร็ว แล้วค่อยทำการสร้าง หรือปรับปรุงดัชนีเมื่อการนำเข้าข้อมูลเสร็จสิ้นลง เนื่องจากโปรแกรม SQL*Loader ยังไม่สามารถอ่านข้อมูลที่อยู่ในรูปแบบของระบบสารสนเทศภูมิศาสตร์ทั่วไปได้ ไม่ว่าจะเป็น ESRI shapefiles หรือ MapInfo Tab files หรือ AutoDesk DWG files เป็นต้น อย่างไรก็ตาม ESRI shapefiles ถือว่าเป็นรูปแบบข้อมูลที่นิยมนำมาใช้ในการแลกเปลี่ยนข้อมูลมากที่สุด โอราเคิลจึงสร้างโปรแกรมสำหรับเป็นเครื่องมือที่ทำหน้าที่ได้ สองอย่างคือ สามารถสร้างคำสั่ง SQL ในการสร้างตารางและข้อมูลที่เกี่ยวข้องในฐานข้อมูล และสามารถสร้างแฟ้มควบคุมสำหรับการนำเข้าข้อมูล ถ้าผู้ใช้งานนำเข้าข้อมูลด้วยวิธีนี้หลังจากนำเข้าข้อมูลเสร็จ ผู้ใช้ต้องใช้คำสั่ง SDO_MIGRATE.TO_CURRENT เพื่อปรับรูปแบบข้อมูลให้อยู่ในรูปแบบล่าสุดของฐานข้อมูลโอราเคิล

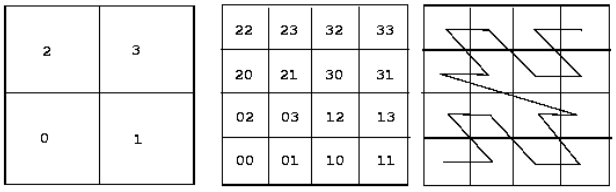
การนำเข้าเป็นรายการ อาศัยคำสั่ง Insert ของ SQL เหมาะกับการนำเข้าข้อมูลจำนวนน้อย การนำเข้าข้อมูลลักษณะนี้ ดัชนีเชิงพื้นที่จะถูกปรับปรุงโดยอัตโนมัติ เมื่อนำเข้าข้อมูลเชิงพื้นที่เรียบร้อยแล้ว เพื่อให้มั่นใจว่าข้อมูลเชิงพื้นที่ที่มีอยู่ มีความถูกต้อง Oracle Spatial มีแนวทางในการตรวจสอบความถูกต้องในเบื้องต้น เมื่อมีการสร้างดัชนีเชิงพื้นที่ และมีคำสั่งที่ทำหน้าที่ตรวจสอบความถูกต้องของข้อมูลเชิงพื้นที่สองคำสั่งคือ SDO_GEOM.VALIDATE_GEOMETRYWITH_CONTEXT และSDO_GEOM.VALIDATE_LAYER_WITH_CONTEXT

2.5.5 การจัดทำดัชนีข้อมูลเชิงพื้นที่

Oracle Spatial มีวิธีการจัดทำดัชนีสำหรับข้อมูลเชิงพื้นที่ให้เลือก 2 วิธี คือ การจัดทำดัชนีแบบ Quad-Tree และการจัดทำดัชนีแบบ R-tree

การจัดทำดัชนีแบบ Quad-Tree ใช้วิธีการแปลงวัตถุให้เป็นชุดของสี่เหลี่ยมและแบ่งช่องสี่เหลี่ยม (Tesselation) เป็นหลาย ๆช่องให้ครอบคลุมพื้นที่ของรูปเรขาคณิต โดยแบ่งในลักษณะลำดับขั้น เริ่มจากแบ่งเป็นสี่ช่องเท่า ๆกัน หลังจากนั้นแต่ละช่องจะถูกแบ่งต่อเป็นสี่ช่องเล็กไปเรื่อย ๆ จนกระทั่งได้ขนาดที่เหมาะสม หรือได้ตามเงื่อนไขที่กำหนดไว้ เช่น จำนวนช่องสี่เหลี่ยมถึงจำนวนสูงสุดที่กำหนดไว้ หรือได้ขนาดของช่องสี่เหลี่ยมเท่าที่กำหนดไว้ เป็นต้น ช่องสี่เหลี่ยมแต่ละช่องจะมีตัวเลขที่ไม่ซ้ำกันเรียกว่า Mortons codes หรือ z-values ดังภาพประกอบ 2-6

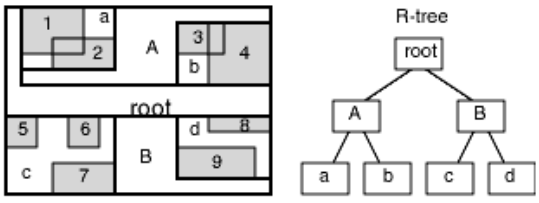
การจัดทำดัชนีแบบ R-Tree เป็นการสร้างดัชนีข้อมูลเชิงพื้นที่โดยไม่ได้ระบุว่าต้องการสร้างดัชนีแบบใด Oracle Spatial จะสร้างดัชนีข้อมูลเชิงพื้นที่แบบ R-Tree ให้โดยอัตโนมัติ การสร้างดัชนีข้อมูลเชิงพื้นที่แบบ R-Tree ทำโดยการประมาณรูปทรงเรขาคณิตให้เป็นรูปสี่เหลี่ยมขนาดเล็กที่สุด (Minimum Bounding Rectangle : MBR) ที่สามารถครอบคลุมพื้นที่รูปทรงเรขาคณิตได้พอดี ดังภาพประกอบ 2-7



ภาพประกอบ 2-6 Quad-Tree decomposition and Morton codes

ภาพประกอบ 2-7 Minimum Bounding Rectangle

ในแต่ละชั้นข้อมูล การจัดทำดัชนีเชิงพื้นที่มีโครงสร้างดัชนีแบบลำดับขั้น (Hierarchical index) ของ รูปสี่เหลี่ยมที่มีขนาดเล็กที่สุดที่ประมาณไว้ ดังภาพประกอบ 2-8



ภาพประกอบ 2-8 R-Tree Hierarchical index on MBRs

ภาพประกอบ 2-8 อธิบายได้ดังนี้

1. หมายเลข 1 ถึง 9 เป็นรูปทรงเรขาคณิตของชั้นข้อมูล
2. a b c และ d เป็นโหนดลูก (Leaf node) ของดัชนีแบบอาร์-ทรี และประกอบด้วยสี่เหลี่ยมที่เล็กที่สุดที่ครอบคลุมรูปทรงเรขาคณิต รวมทั้งพอยเตอร์ (pointer) ที่ชี้ไปยังรูปทรงเรขาคณิต เช่น a ประกอบด้วย MBR ของรูปทรงเรขาคณิตหมายเลข 1 และ 2 b ประกอบด้วย MBR ของรูปทรงเรขาคณิตหมายเลข 3 และ 4 เป็นเช่นนี้ไปเรื่อย ๆ
3. A คือ MBR ที่ประกอบด้วย a และ b ส่วน B คือ MBR ที่ประกอบด้วย c และ dและที่โหนดบนสุด (Root) ประกอบด้วย MBR ของ A และ B

ข้อมูลของดัชนีแบบ R-Tree ถูกเก็บไว้ในตารางชื่อ SDO_INDEX_TABLE ในวิวชื่อ USER_SDO_INDEX_METADATA ดัชนีแบบอาร์-ทรี ไม่เหมาะกับลักษณะงานแบบที่มีการใช้คำสั่งเพิ่ม และลบข้อมูลบ่อย ๆ เนื่องจากการสร้างดัชนีเชิงพื้นที่ จะต้องใช้พื้นที่บนโรลแบ็คเซกเมนต์ (Rollback segment) มากเพียงพอ เนื่องจากจะต้องใช้พื้นที่ถึง 100*n bytes โดยที่ n คือจำนวนแถวข้อมูล ถ้าพื้นที่บนโรลแบ็ค เซกเมนต์ มีไม่พอจะไม่สามารถสร้างดัชนีข้อมูลเชิงพื้นที่ได้

การเลือกใช้งานดัชนีข้อมูลเชิงพื้นที่

การพิจารณาว่าควรเลือกใช้ดัชนีข้อมูลเชิงพื้นที่แบบใด สำหรับการสอบถามข้อมูลแต่ละคำถามพิจารณา ได้ดังแสดงในตาราง 2-1

ตาราง 2-1 การพิจารณาเลือกใช้งานดัชนี

R-tree indexing	Quadtree indexing
- จัดทำดัชนีโดยอาศัยรูปสี่เหลี่ยมขนาดเล็กที่สุดที่ครอบคลุมข้อมูลเชิงพื้นที่รูปเรขาคณิต ทำให้มีความคลาดเคลื่อนของข้อมูลเชิงพื้นที่มากขึ้น	- จัดทำดัชนีโดยประมาณการข้อมูลเชิงพื้นที่แบบปูเป็นแผ่นกระเบื้อง ทำให้ความคลาดเคลื่อนน้อยลง
- สามารถจัดทำดัชนีข้อมูลเชิงพื้นที่ได้ถึง 4 มิติ	- สามารถจัดทำดัชนีข้อมูลเชิงพื้นที่ได้แค่ 2 มิติ

ตาราง 2-1 (ต่อ) การพิจารณาเลือกใช้งานดัชนี

R-tree indexing	Quadtree indexing
- การสร้างและการปรับปรุงดัชนีข้อมูลเชิงพื้นที่ทำได้ง่าย	- การปรับปรุงดัชนีข้อมูลเชิงพื้นที่มีความซับซ้อน การกำหนดค่าพารามิเตอร์ของการปรับดัชนีอาจมีผลกระทบกับประสิทธิภาพการทำงาน
- ต้องการพื้นที่เก็บข้อมูลน้อย	- ต้องการพื้นที่เก็บข้อมูลมาก
- ถ้าทำงานด้วยคำสั่งที่ต้องประมวลผลมาก เช่นการหาวัตถุที่อยู่ใกล้เคียง (ใช้คำสั่ง SDO_NN) ดัชนีแบบอาร์-ทรี ทำงานได้เร็วกว่า	- ถ้าทำงานด้วยคำสั่งที่ต้องประมวลผลมาก เช่นการหาวัตถุที่อยู่ใกล้เคียง (ใช้คำสั่ง SDO_NN) ดัชนีแบบควอดทรี จะทำงานได้ช้ากว่า
- ถ้ามีการปรับปรุงข้อมูลเชิงพื้นที่บ่อย ๆ ดัชนีแบบอาร์-ทรี ประสิทธิภาพด้อยกว่า	- ถ้ามีการปรับปรุงข้อมูลเชิงพื้นที่บ่อย ๆ ดัชนีแบบควอดทรี จะมีประสิทธิภาพดีกว่า
- ถ้าใช้คำสั่ง SDO_WITHIN_DISTANCE ในการสอบถามข้อมูล แนะนำให้ใช้ดัชนีข้อมูลเชิงพื้นที่แบบอาร์-ทรี	

2.5.6 การสอบถามข้อมูลเชิงพื้นที่

Oracle Spatial ใช้รูปแบบการสอบถามข้อมูลแบบสองระดับ (Two-tier query model) ในการสอบถามข้อมูลเชิงพื้นที่ นั่นคือการสอบถามข้อมูลจะประกอบด้วย 2 ขั้นตอนหลัก ๆ คือ การกรองข้อมูลหลัก (Primary Filter) และ การกรองข้อมูลรอง (Secondary Filter)

การกรองข้อมูลหลัก ทำหน้าที่เลือกข้อมูล จากข้อมูลทั้งหมด ทำการกรองโดยการเปรียบเทียบ ไม่มีการคำนวณที่ซับซ้อน ไม่เปลืองพื้นที่ใช้งาน และทำงานได้เร็ว เนื่องจากมีการใช้ดัชนีเชิงพื้นที่ในการทำงาน ผลลัพธ์ที่ได้จะส่งต่อให้เป็นข้อมูลนำเข้าให้กับการกรองข้อมูลรอง

การกรองข้อมูลรอง เป็นขั้นตอนที่จะทำการคำนวณที่ซับซ้อน โดยทำกับข้อมูลที่เป็นผลลัพธ์จากการกรองข้อมูลหลัก ดังนั้นจำนวนข้อมูลที่การกรองข้อมูลรอง ทำงานจึงมีจำนวนน้อยลง

ถึงแม้ว่าการสอบถามข้อมูล ประกอบด้วย 2 ขั้นตอนหลักๆ แต่ในบางกรณีสามารถใช้เฉพาะขั้นตอนแรก คือการกรองข้อมูลหลักได้ เช่น การย่อ/ขยาย แผนที่

คำสั่งในการทำงาน

- SDO_FILTER เป็นคำสั่งที่ใช้ในการกรองข้อมูลซึ่งทำงานเฉพาะการกรองข้อมูลหลัก
- SDO_RELATED เป็นคำสั่งที่ใช้สำหรับหาว่ามีรูปทรงเรขาคณิตใดบ้างที่มีบางส่วนซ้อนทับกัน คำสั่งนี้เป็นคำสั่งที่ทำงานทั้งสองขั้นตอน คือการกรองข้อมูลหลัก และการกรองข้อมูลรอง
- SDO_WITHIN_DISTANCE ใช้เพื่อเปรียบเทียบว่ากลุ่มของวัตถุหรือรูปเรขาคณิตอยู่ในระยะ n หน่วยจากตำแหน่งของวัตถุอ้างอิงหรือไม่
- SDO_NN เป็นคำสั่งใช้เพื่อค้นหาวัตถุ หรือรูปเรขาคณิตที่อยู่ใกล้เคียง
- SDO_GEOM.SDO_AREA คำนวณหาพื้นที่
- SDO_GEOM.SDO_LENGTH คำนวณหาระยะทาง
- ฯลฯ

โดยทุกๆคำสั่งที่กล่าวถึง ต้องส่งค่าอาร์กิวเมนต์แรก เป็นชื่อสดมภ์ที่ต้องการค้นหา และต้องมีดัชนีเชิงพื้นที่ อาร์กิวเมนต์ตัวที่สอง มักจะเป็นคิวรีวินโดว์ หรือพื้นที่ที่สนใจ อาร์กิวเมนต์ตัวที่สาม เป็นสตริงค์ที่ประกอบด้วยรายการของตัวแปรสำหรับแต่ละคำสั่ง

2.5.7 ข้อมูลสำหรับการแสดงแผนที่

Oracle Spatial มีการจัดเก็บข้อมูลสำหรับการแสดงแผนที่ และข้อมูลเชิงพื้นที่แยกออกจากกันอย่างชัดเจน ช่วยให้มีความเป็นอิสระระหว่างข้อมูลเชิงพื้นที่ และลักษณะการแสดงผลนั่นคือ ถ้ามีการปรับปรุงแก้ไขใดๆที่ข้อมูลสำหรับการแสดงแผนที่ จะไม่มีผลกับข้อมูลเชิงพื้นที่ และในทางกลับกันคือ ถ้ามีการแก้ไขข้อมูลเชิงพื้นที่ใดๆ ก็ไม่มีผลกับข้อมูลสำหรับการแสดงแผนที่

ข้อมูลสำหรับการแสดงแผนที่ (Map Metadata) เป็นส่วนสำคัญในการตัดสินลักษณะและรูปแบบแผนที่ที่จะสร้างขึ้นมา ประกอบด้วย สัญลักษณ์ของแผนที่ รูปแบบตัวอักษร รูปแบบของพื้นที่และเส้น เรียกว่า สไตล์ (Styles) กฎเกณฑ์ของรูปแบบที่เกี่ยวข้องกับตารางข้อมูลเชิงพื้นที่ ของชั้นแผนที่หรือเรียกว่า ธีม (Themes) ผู้ใช้งานแต่ละคนสามารถกำหนด Metadata เป็นของตนเอง หรือจะใช้งานร่วมกับผู้ใช้งานคนอื่นๆก็ได้ ในองค์กร ก็สามารถกำหนดชุดสัญลักษณ์ของแผนที่ที่สามารถใช้งานร่วมกันระหว่างแผนก แต่ละแผนกก็สามารถกำหนด ชั้นของแผนที่เป็นของตนเองได้ โดยใช้สัญลักษณ์ของแผนที่พื้นฐาน ที่เรียกว่า แผนที่ฐาน (Base maps) ร่วมกัน

สไตล์ คือลักษณะที่มองเห็นได้ ใช้เพื่อแสดงคุณลักษณะเชิงพื้นที่ ประกอบด้วยสัญลักษณ์พื้นฐานของแผนที่ และคำอธิบาย สำหรับแทนจุด เส้น และพื้นที่ ซึ่งถูกกำหนดและเก็บไว้เป็นรูปแบบเฉพาะที่มีชื่อไม่ซ้ำกัน ข้อมูล Styles จะต้องมีการกำหนดไว้ล่วงหน้าเก็บไว้ในตารางฐานข้อมูล ชื่อ USER_SDO_STYLES ตาราง 2-2

ธีม คือกลุ่มของคุณลักษณะทางภูมิศาสตร์ ที่มีสมบัติเดียวกัน เปรียบเป็นภาพที่แสดงชั้นข้อมูลใดชั้นข้อมูลหนึ่งที่เกี่ยวข้องกับแผนที่ภูมิศาสตร์ สามารถกำหนดไว้ล่วงหน้าเก็บไว้ในตารางฐานข้อมูล ชื่อ USER_SDO_THEMES ดังแสดงโครงสร้างตารางไว้ในตาราง 2-3 หรือจะกำหนดเป็นส่วนหนึ่งของโปรแกรมในขณะที่มีการร้องขอแผนที่ ดังตัวอย่างในภาคผนวก ก ภาพประกอบ ก-2

แผนที่ฐาน คือแผนที่ที่เกิดจากการรวมกลุ่มของธีมเพื่ออำนวยความสะดวกในการเรียกใช้ Themes หลายๆ Themes ในการร้องขอแผนที่หนึ่งครั้ง ข้อมูลเกี่ยวกับการกำหนด Base map ถูกเก็บไว้ในตารางฐานข้อมูลชื่อ USER_SDO_MAPS ดังแสดงโครงสร้างตารางไว้ในตาราง 2-4

ตาราง 2-2 โครงสร้างตาราง USER_SDO_STYLES

แอตตริบิวท์	ชนิดข้อมูล	คำอธิบาย	เงื่อนไข
NAME	VARCHAR2(32)	ชื่อ Styles	NOT NULL
TYPE	VARCHAR2(32)	ชนิดของลักษณะ ได้แค่ COLOR, MARKER, LINE, AREA, TEXT, ADVANCED	NOT NULL
DESCRIPTION	VARCHAR2(4000)	คำอธิบาย	

ตาราง 2-2 (ต่อ) โครงสร้างตาราง USER_SDO_STYLES

แอตตริบิวท์	ชนิดข้อมูล	คำอธิบาย	เงื่อนไข
DEFINITION	CLOB	คำอธิบาย	NOT NULL
IMAGE	BLOB	ภาพ	
GEOMETRY	MDSYS.SDO_GEOMETRY	ข้อมูลสารสนเทศภูมิศาสตร์	

ตาราง 2-3 โครงสร้างตาราง USER_SDO_THEMES

แอตตริบิวท์	ชนิดข้อมูล	คำอธิบาย	เงื่อนไข
NAME	VARCHAR2(32)	ชื่อธีม	NOT NULL
DESCRIPTION	VARCHAR2(4000)	คำอธิบาย	
BASE_TABLE	VARCHAR2(32)	ตาราง	NOT NULL
GEOMETRY_COLUMN	VARCHAR2(2048)	แอตตริบิวท์	NOT NULL
STYLING_RULES	CLOB		NOT NULL

ตาราง 2-4 โครงสร้างตาราง USER_SDO_MAPS

แอตตริบิวท์	ชนิดข้อมูล	คำอธิบาย	เงื่อนไข
NAME	VARCHAR2(32)	ชื่อแผนที่	NOT NULL
DESCRIPTION	VARCHAR2(4000)	คำอธิบาย	
DEFINITION	CLOB		NOT NULL

2.6 การประยุกต์ใช้งาน และงานวิจัยที่เกี่ยวข้อง

ระบบสารสนเทศภูมิศาสตร์ สามารถนำมาประยุกต์ใช้งานอย่างหลากหลาย แทบในทุกสาขาอาชีพ ซึ่งพอจะสรุปได้ดังนี้

การประยุกต์ระบบสารสนเทศเพื่อใช้งานด้านสาธารณูปโภค ได้แก่ กิจการไฟฟ้า น้ำประปา ระบบการจราจร และระบบโทรศัพท์เป็นต้น โดยในกิจการไฟฟ้าสามารถนำไปประยุกต์ใช้งานตั้งแต่การขอใช้ การติดตั้ง การแก้ปัญหากรณีกระแสไฟตก หรือไฟฟ้าดับให้เกิดผลกระทบน้อยที่สุด ในกิจการน้ำประปา สามารถนำไปช่วยในการตรวจสอบและป้องกันท่อประปารั่วซึมเพื่อประหยัดทรัพยากรน้ำ ในระบบการจราจร สามารถช่วยในการกำหนดเส้นทางการเดินรถประจำทาง เพื่อให้บริการกับประชาชนได้แบบสะดวก ปลอดภัย ประหยัด และบริการได้มากที่สุด หรือช่วยในการเลือกเส้นทางที่สะดวกที่สุดสำหรับการขนส่งสินค้าจากโรงงานผลิตสู่แหล่ง

จำหน่าย และในระบบโทรศัพท์ สามารถนำไปวิเคราะห์หาจุดอับของสัญญาณโทรศัพท์ เพื่อปรับปรุงเครือข่ายการบริการให้มีความครอบคลุมและสัญญาณคมชัดยิ่งขึ้น

การประยุกต์ใช้งานด้านการสาธารณสุข ใช้เพื่อการศึกษาสภาพภูมิประเทศ ความสัมพันธ์ของนิเวศวิทยา และพาหะนำโรค ช่วยให้สามารถหาแนวทางการป้องกัน และแก้ไขการเกิดโรคระบาดได้ หรือนำมาใช้เพื่อศึกษาหาแหล่งในการกำจัดขยะที่เหมาะสม

การประยุกต์ใช้งานด้านโบราณคดี สามารถใช้เพื่อศึกษาถึงปัจจัยที่เป็นที่ตั้งของแหล่งโบราณวัตถุ เช่น แหล่งน้ำ สภาพดิน สภาพภูมิประเทศ เป็นต้น

การประยุกต์ใช้กับระบบธุรกิจ เช่นใช้ในการวิเคราะห์ด้านการตลาด การหาทำเลที่เหมาะสมในการขยายสาขา สำนักงานโดยอาศัยจำนวน และที่อยู่ของประชากร เพื่อเลือกที่ตั้ง ให้ได้ยอดจำหน่ายและผลตอบแทนทางเศรษฐกิจดีที่สุด และเชื่อมโยงข้อมูลปริมาณการขายกับแผนที่ ทำให้สามารถเพิ่มประสิทธิภาพวิเคราะห์วางเป้าหมายการตลาดในแต่ละพื้นที่สำหรับสินค้าแต่ละชนิด

ปัจจุบันองค์กรการปกครองส่วนท้องถิ่นหลาย ๆองค์กรได้มีการนำระบบสารสนเทศภูมิศาสตร์มาใช้ในระบบงานแผนที่ภาษีและทะเบียนทรัพย์สิน โดยมีการพัฒนาเป็นโปรแกรมประยุกต์ ตัวอย่างเช่น โปรแกรมระบบงานแผนที่ภาษีและทะเบียนทรัพย์สินเทศบาลตำบลบ้านพรุ อำเภอหาดใหญ่ จังหวัดสงขลา [อาทิตย์ เทอดสุวรรณ, 2543] แบ่งโปรแกรมเป็น 2 โปรแกรมหลัก

- **โปรแกรมส่วนแรก** คือโปรแกรมฐานข้อมูลภาษี พัฒนาด้วย VBA (Visual Basic for Applications) บนระบบการจัดการฐานข้อมูล Microsoft Access โปรแกรมส่วนนี้เป็นโปรแกรมที่ทำงานกับฐานข้อมูลทางทะเบียนทรัพย์สิน
- **โปรแกรมส่วนที่สอง** คือโปรแกรมแผนที่ภาษี พัฒนาด้วยภาษา Avenue บนซอฟต์แวร์ ArcView 3.x ใช้เป็นโปรแกรมที่ทำงานกับฐานข้อมูลเป็นที่เป็นหลัก และเป็นส่วนที่จะทำให้เห็นความสัมพันธ์ระหว่างข้อมูลแผนที่และข้อมูลทางทะเบียนทรัพย์สิน

ในปัจจุบันได้มีระบบจัดการฐานข้อมูลแบบความสัมพันธ์เชิงวัตถุที่สามารถจัดเก็บข้อมูลเชิงพื้นที่ และข้อมูลคุณลักษณะไว้ในฐานข้อมูลเดียวกัน จึงได้มีงานวิจัยเพื่อศึกษาถึงแนวทางการพัฒนาโดยใช้ระบบจัดการฐานข้อมูลแบบความสัมพันธ์เชิงวัตถุสำหรับการจัดเก็บ จัด

การ และดึงข้อมูลเชิงพื้นที่ทางน้ำ (Marine Spatial Information) ของ Australian Oceanographic Data Center (AODC) เพื่อช่วยให้ผู้ใช้งานกลุ่มต่าง ๆเข้าถึงข้อมูลได้ง่ายขึ้น และพัฒนาฐานข้อมูลเพื่อเก็บข้อมูลเชิงพื้นที่ทางน้ำ บนระบบจัดการฐานข้อมูลแบบความสัมพันธ์เชิงวัตถุ และพัฒนาเป็นโปรแกรมประยุกต์ที่ยอมให้เข้าถึงข้อมูลผ่านเว็บ (WWW) ได้ [Syafi, Searle, Masters, 1998]