คุณภาพของข้อมูล GIS

104314 – ระบบสารสนเทศภูมิศาสตร์

คณะเกษตรศาสตร์ ฯ มหาวิทยาลัยนเรศวร

จรัณธร บุญญนุภาพ

คุณภาพข้อมูล ?

ระดับของความละเอียดและความถูกต้องที่มีอยู่ในตัวข้อมูลนั้น ๆ

“Degree of excellence” in a product, service or performance

เป็นสิ่งที่บ่งชี้และแสดงถึงความน่าเชื่อถือ สถานภาพ และการใช้ประโยชน์ของข้อมูลในระบบ GIS

ทำไมคุณภาพข้อมูลถึงมีความสำคัญ

- ปกติการสร้างข้อมูลภูมิศาสตร์ จะไม่มีการกำหนดมาตรฐานของคุณภาพข้อมูล (ข้อมูลจะถูกใช้ภายในองค์กรเท่านั้น)
- ปัจจุบันมีการนำข้อมูลภูมิศาสตร์และ GIS มาใช้ในการตัดสินใจและวางแผนในด้านต่าง ๆ เพิ่มมากขึ้น
- ดังนั้นข้อมูลภูมิศาสตร์ที่ผลิตจากองค์กรต่าง ๆ จะกลายเป็นแหล่งข้อมูลทุติยภูมิ (secondary data sources) ที่สำคัญ ซึ่งข้อมูลเหล่านี้มีมาตรฐานและคุณภาพข้อมูลที่แตกต่างกัน

ทำไมคุณภาพข้อมูลถึงมีความสำคัญ

- คุณภาพของข้อมูลที่นำมาใช้ จะเป็นสิ่งกำหนดคุณภาพและความน่าเชื่อถือของการตัดสินใจหรือผลลัพท์ที่ได้ (Decision making or Output)

- สรุปคือ ประสิทธิภาพของการตัดสินใจและวางแผนที่จะแก้ไขปัญหาหรือกระทำสิ่งใด จะแปรผันโดยตรงกับคุณภาพของข้อมูล

คุณภาพของข้อมูลภูมิศาสตร์

มีองค์ประกอบที่สำคัญ 4 อย่าง คือ

- ความถูกต้อง (Accuracy)
- ความแม่นยำ (Precision)
- ความสอดคล้องเชิงตรรกะ (Logical Consistency)
- ความสมบูรณ์ (Completeness)

ซึ่งองค์ประกอบเหล่านี้จะขึ้นอยู่กับมิติทางภูมิศาสตร์ด้านต่าง ๆ คือ

space (spatial), time (temporal), and theme (thematic)

ความถูกต้อง (Accuracy)

ระดับที่ตรงกันระหว่างข้อสนเทศบนแผนที่หรือในฐานข้อมูล กับสิ่งที่เป็นจริง ภายใต้ค่าที่ยอมรับได้

ความถูกต้อง (Accuracy)

ความถูกต้องเชิงพื้นที่ (Spatial Accuracy)

- ความถูกต้องขององค์ประกอบฐานข้อมูลเชิงพื้นที่
- พิจารณาถึงความถูกต้องของตำแหน่งทางภูมิศาสตร์ทั้งในแนวราบและแนวดิ่ง
- โดยความถูกต้องเชิงพื้นที่จะขึ้นอยู่กับ มาตราส่วนของแผนที่ต้นฉบับ
- ค่าความผิดพลาดเชิงพื้นที่สามารถแสดงโดย ค่า RMS (root means square) error

mathematically known as $RMS = ((x_1-x_2)^2 + (y_1-y_2)^2)^{½}$

ความถูกต้อง (Accuracy)

ตัวอย่างความถูกต้องเชิงตำแหน่ง

มาตรฐานความถูกต้องเชิงตำแหน่งในแนวราบ (Horizontal Accuracy) สำหรับแผนที่มาตราส่วนต่าง ๆ (USGS)

- 1 : 1,200 เท่ากับ 3.33 feet
- 1 : 2,400 เท่ากับ 6.67 feet
- 1: 10,000 เท่ากับ 27.78 feet
- 1 : 100,000 เท่ากับ 166,67 feet

ความถูกต้อง (Accuracy)

ความถูกต้องเชิงเวลา (Temporal Accuracy)

ความเข้ากันได้ของ วันเวลาที่ทำการผลิตแผนที่ ระหว่างเวลาที่ทำการบันทึกกับ เวลาจริง สาThe agreement (often implicit, e.g. refers to a map date) between the encoded and “actual” temporal coordinates for a feature

Note: Temporal Coordinate is the temporal limits within which the feature is valid

ความถูกต้อง (Accuracy)

ความถูกต้องของลักษณะเฉพาะ (Thematic Accuracy)

- เป็นความถูกต้องของค่าข้อมูล (attribute values) ภายในฐานข้อมูล
- ซึ่งระบบการเก็บค่าจะขึ้นอยู่กับมาตรส่วนในการวัดค่าของข้อมูล

ข้อมูลเชิงปริมาณ (Quantitative data) เช่น ปริมาณน้ำฝน อุณหภูมิ

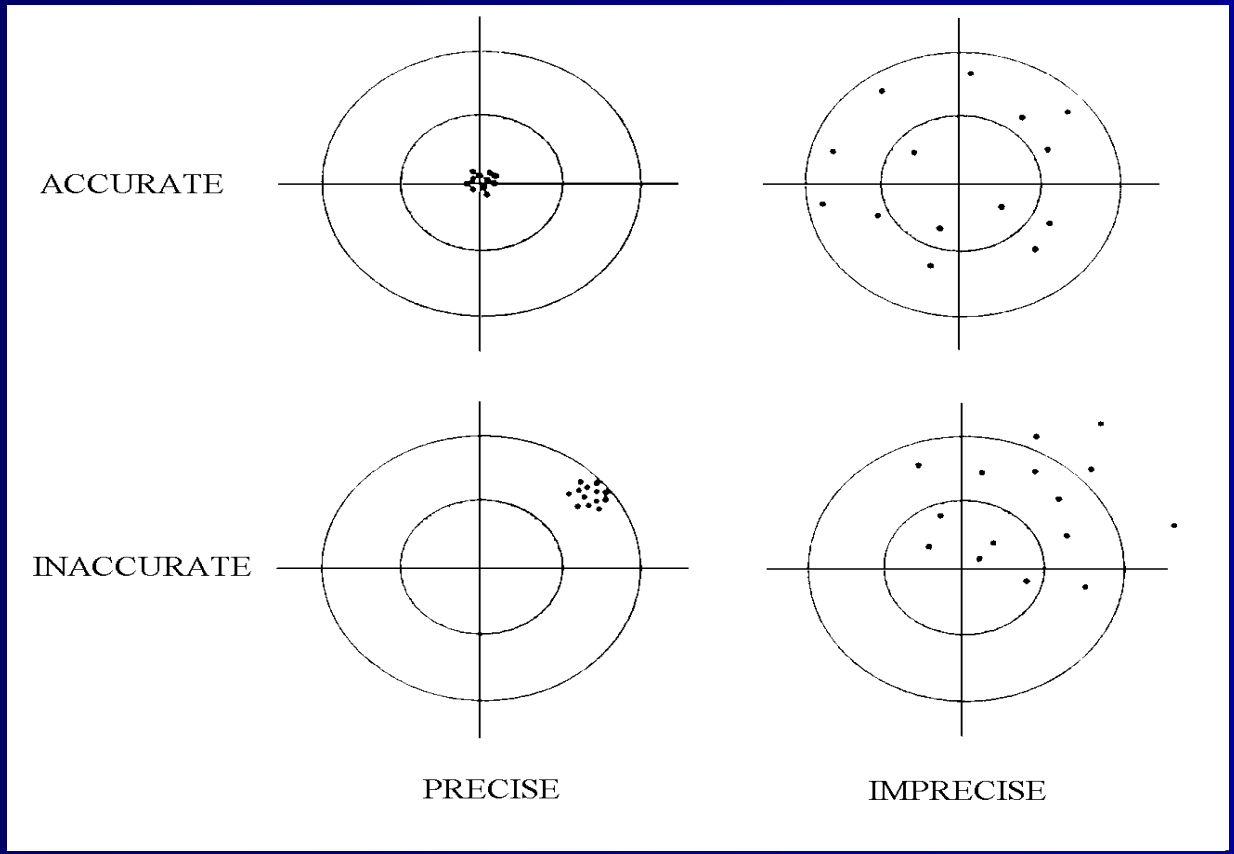
ข้อมูลเชิงคุณภาพ (Qualitative data) เช่น ประเภทการใช้ที่ดิน ประเภทชุดดิน

ความแม่นยำ Precision (Resolution)

- แสดงถึงระดับความแม่นยำหรือความเที่ยงตรงของการวัด (measurement) และความตรงกันของคุณลักษณะในฐานข้อมูล

ความถูกต้องและความแม่นยำเชิงพื้นที่

Spatial Accuracy and Precision



ความแม่นยำ Precision (Resolution)

ความละเอียดเชิงพื้นที่ (Spatial Resolution)

แสดงถึงสิ่งที่เล็กที่สุดที่สามารถสังเกตหรือแสดงออกมาได้

- In raster data, it refers to the **linear dimension of a cell** or grid

  ขนาดของช่องกริด

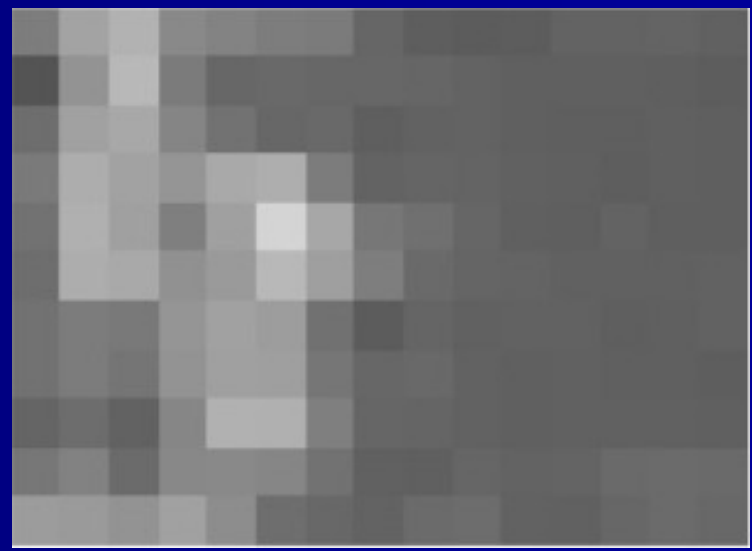
- In vector data, it refers to the **minimum mapping unit size**

  ขนาดของหน่วยที่เล็กที่สุด

ความแม่นยำ Precision (Resolution)

An Example of Spatial Resolution

Higher Spatial Resolution

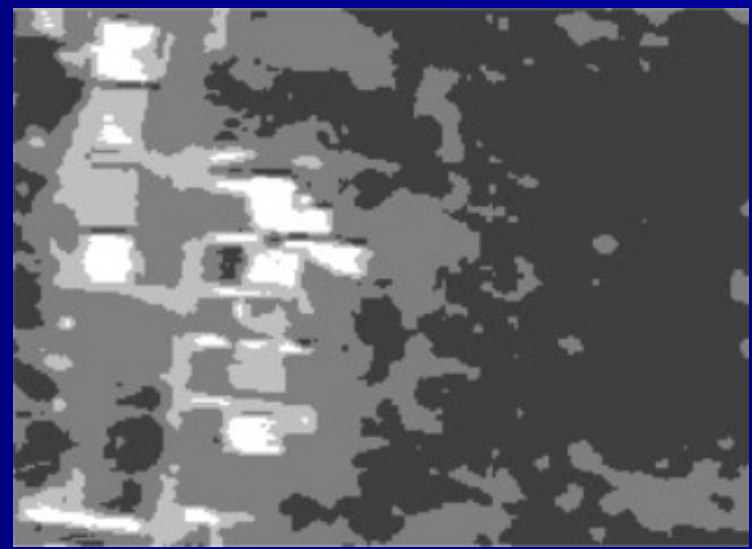


Lower Spatial Resolution

ความแม่นยำ Precision (Resolution)

An Example of Thematic Resolution

Higher Thematic Resolution



Lower Thematic Resolution

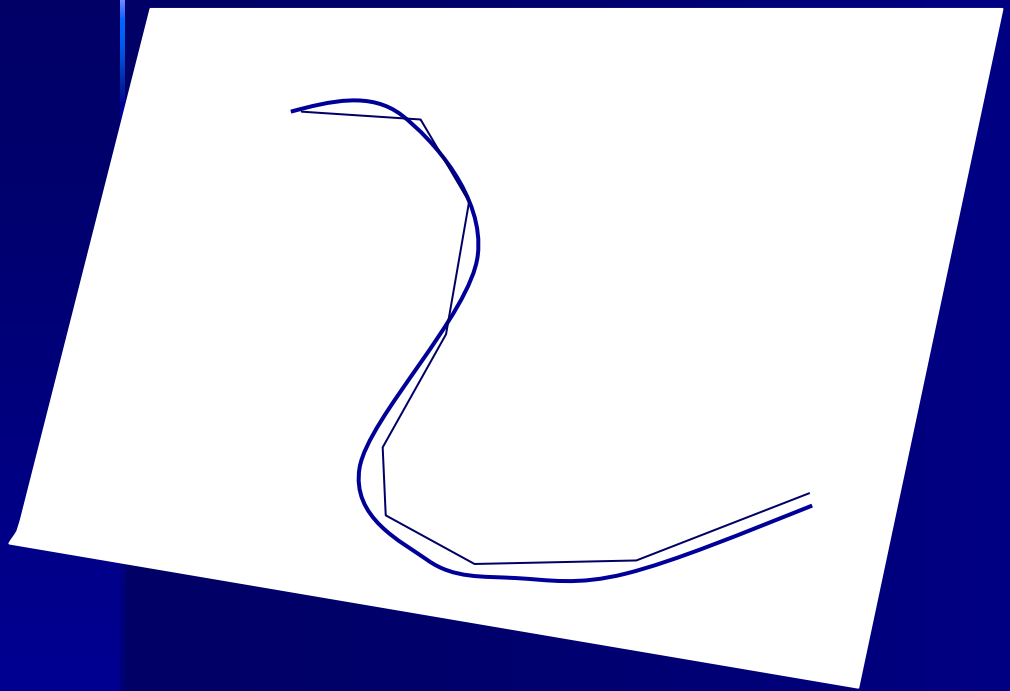
ความแม่นยำ Precision (Resolution)

Thematic Resolution

ความเที่ยงตรงของเกณฑ์การวัดหรือการจำแนกประเภท ในชั้นข้อมูลแต่ละประเภท

- สำหรับข้อมูลเชิงคุภาพ เช่น ระดับความละเอียดในการจำแนกประเภทการใช้ที่ดิน
- ถ้าเป็นข้อมูลเชิงปริมาณ อ้างถึงระดับของความแตกต่างที่น้อยที่สุดในข้อมูลเชิงปริมาณที่สามารถสังเกตุได้

ความเที่ยงตรงหรือความแม่นยำ (Precision)



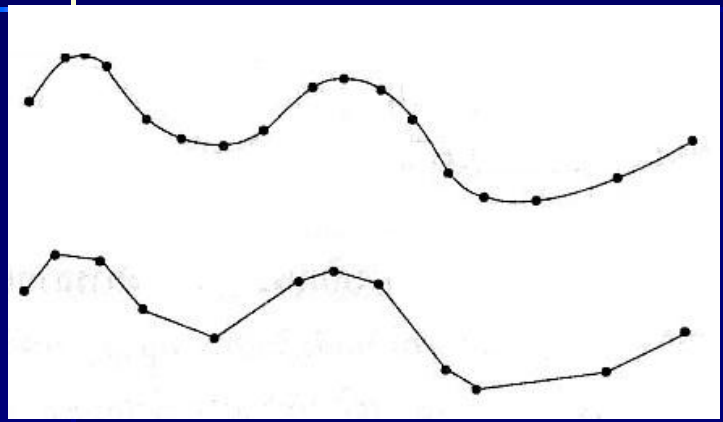
Digitizing Generalization

การดิจิไทซ์จะเป็นการลดทอนความเที่ยงตรงของลักษณะทางภูมิศาสตร์ (precision of feature) เนื่องจากความถูกต้องเชิงพื้นที่ (spatial accuracy) จะลดลงไป

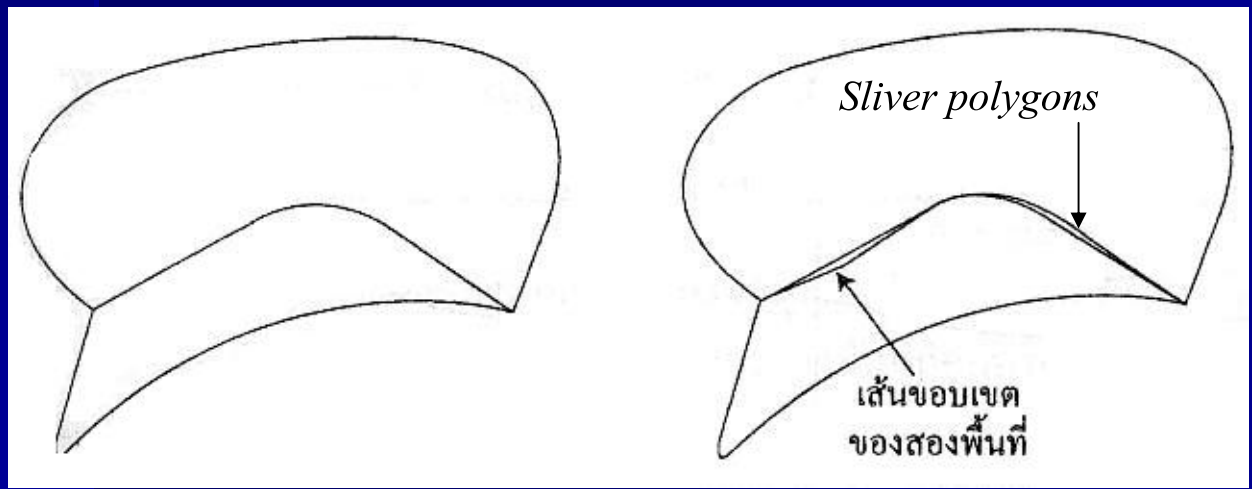
นอกจากนี้ความเที่ยงตรงยังเกี่ยวข้องกับมาตราส่วนของแผนที่ เพราะ features ทางภูมิศาสตร์จะถูกย่อมาตราส่วนให้น้อยกว่า 1:1

ความเที่ยงตรงหรือความแม่นยำ (Precision)

Data reduction



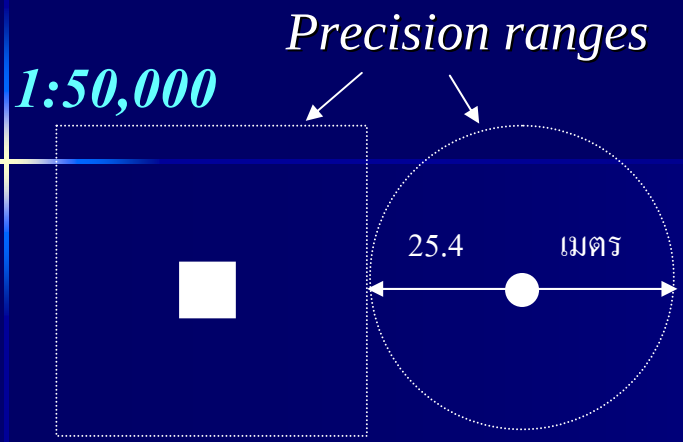
การดิจิไทซ์แบบต่อเนื่อง จำเป็นต้องลดทอนข้อมูล (data reduction) เพื่อกำจัดข้อมูลที่เกินความต้องการ



การดิจิไทซ์ซ้ำกัน 2 ครั้ง จะเกิดช่องว่างการเหลื่อมกันระหว่าง polygons ที่อยู่ติดกัน

Double digitizing

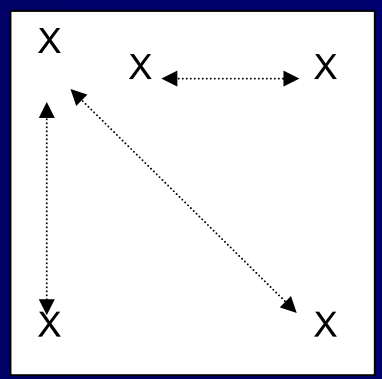
ความเที่ยงตรงหรือความแม่นยำ (Precision)



Map Imprecision

บนแผนที่ 1:50,000 วัตถุใด ๆ จะมีความถูกต้องทางตำแหน่งประมาณ 25.4 เมตร (83.3 ฟุต)

หมายถึงตำแหน่งต่าง ๆ บนแผนที่จะมีความคลาดเคลื่อนในแนวระนาบประมาณ 25.4 เมตร จะตำแหน่งจริงบนพื้นโลก

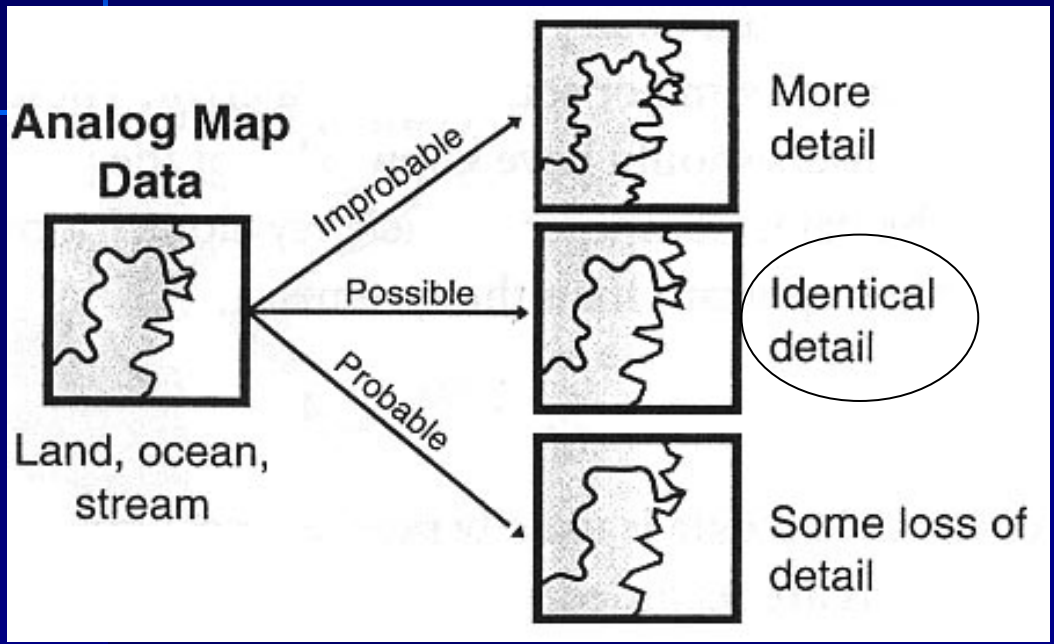


Raster cell Imprecision

สำหรับความถูกต้องทางตำแหน่งของข้อมูลแบบ raster จะขึ้นอยู่กับขนาดของกริด (grid cell size)

จรัณธร บุญญนุภาพ

ข้อมูลที่สร้างขึ้นใหม่ (Generation Data) - Second generation data

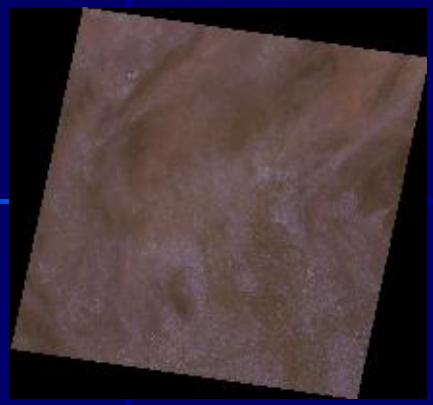


โดยปกติแล้วการทำสำเนาจะมีการสูญเสียคุณภาพของข้อมูล เนื่องจากไม่สามารถที่จะสร้างข้อมูลใหม่ให้เหมือนกับแผนที่ต้นฉบับ (analog map หรือ Photograph) ได้อย่างสมบูรณ์

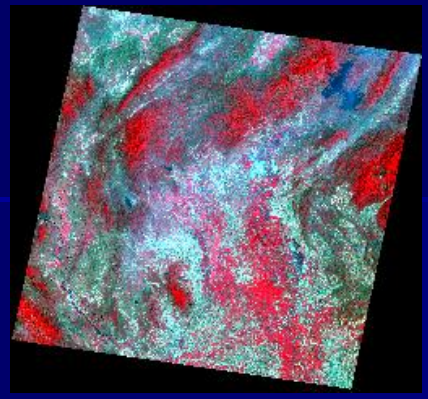
อย่างไรก็ตาม การทำสำเนาจากข้อมูลเชิงตัวเลข (digital data) จะได้ผลผลิตที่เหมือนกันข้อมูลต้นฉบับ (original data) อย่างสมบูรณ์

ข้อมูลที่สร้างขึ้นใหม่ (second generation) จะมีความถูกต้อง (Accuracy) น้อยกว่าข้อมูลต้นฉบับเสมอ

ข้อมูลที่สร้างขึ้นใหม่ (Generation Data) - Spline



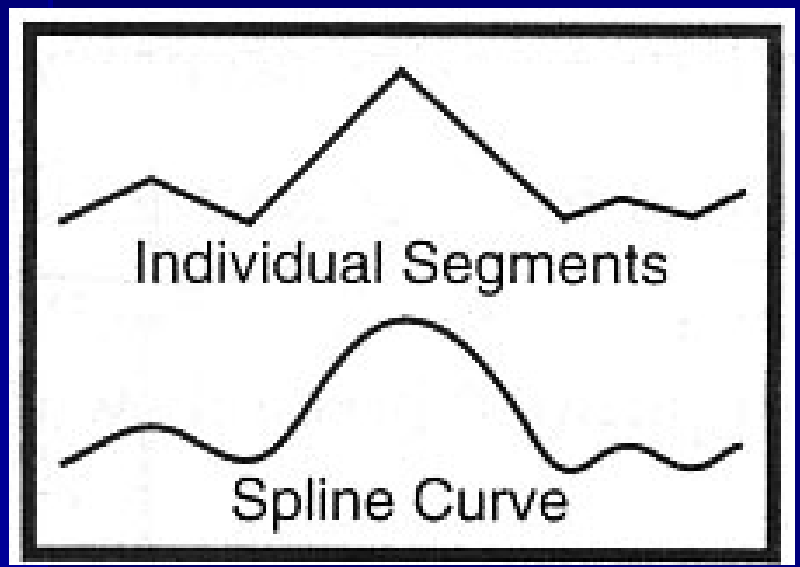
Original image



Enhanced Image

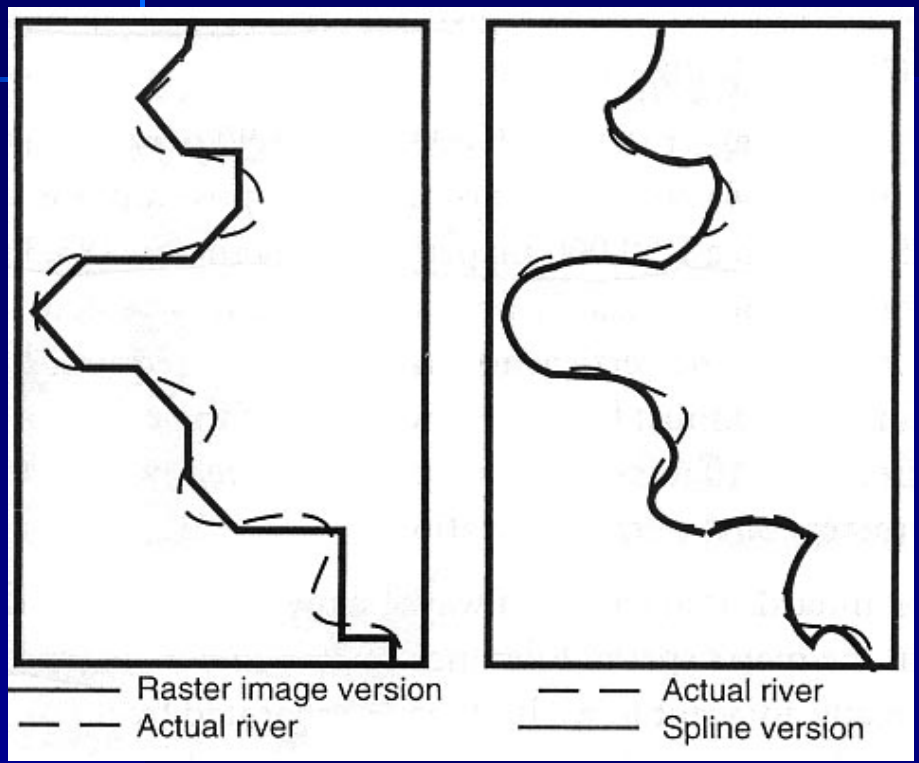
โปรแกรมคอมพิวเตอร์ สามารถช่วยให้ข้อมูลที่สร้างขึ้นใหม่แสดงผลได้ดีขึ้น

ตัวอย่างเช่น การเน้นภาพ (enhancement) จะทำให้ข้อมูลภาพมีความชัดเจนมากขึ้น ซึ่งช่วยให้การแปลตีความสะดวกขึ้น



คำสั่ง Spline operation ใช้ในการทำให้ข้อมุลเส้นที่ได้จากการดิจิไทซ์ มีความ smooth เพิ่มมากขึ้น ด้วยการเพิ่มจุด vertex ตรงตำแหน่งที่เป็นส่วนโค้ง ซึ่งทำให้เส้นมีความโค้งมากกว่าปกติจากที่เป็นอยู่

ข้อมูลที่สร้างขึ้นใหม่ (Generation Data) - Spline Application



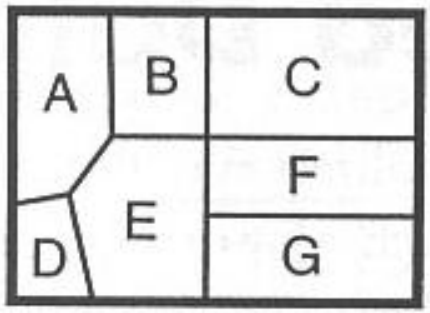
ภาพด้านซ้าย แสดงถึงข้อมูลเส้นแม่น้ำที่ได้จากการ scan (raster version) เปรียบเทียบกับเส้นแม่น้ำจริง (more sinuous)

ภาพด้านขาว เป็นการใช้ทำสั่ง Spline กันเส้นแม่น้ำที่เป็น Raster version ให้มีความโค้งเพิ่มมากขึ้น

อย่างไรก็ตาม เส้นแม่น้ำใหม่ที่ได้จากการทำ Spline ก็ไม่สามารถซ้อนกันกับเส้นแม่น้ำจริง เพียงแต่ทำให้แสดงผลเส้นแม่น้ำดีขึ้นเท่านั้น

ข้อมูลที่สร้างขึ้นใหม่ (Derived Data) - New Data

เป็นวิธีการปรับปรุงข้อมูลโดยการสร้างข้อมูลใหม่ขึ้นมาจากข้อมูลที่มีอยู่ ซึ่งอาจใช้วิธีการคำนวณทางคณิตศาสตร์ หรือ ทางสถิติ เป็นต้น



Districts

District	Avg Income
A	5,888
B	6,022
C	7,400
D	5,500
E	6,000
F	6,600
G	7,100

District income

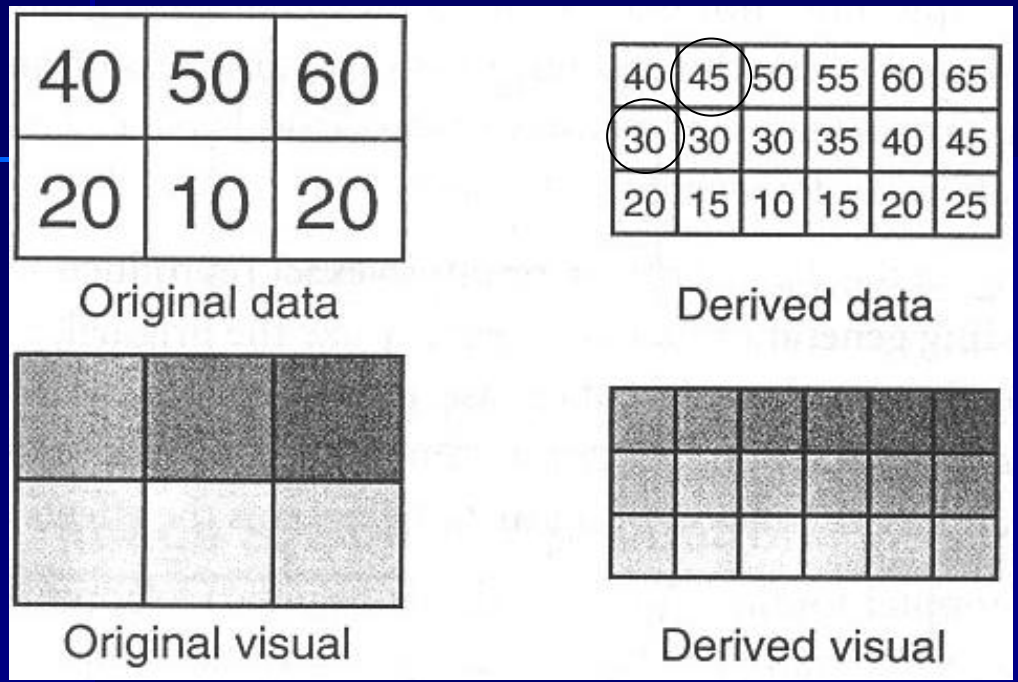
District	Avg Income	+- Mean	Class
A	5,888	-471	Low
B	6,022	-337	Low
C	7,400	+1341	High
D	5,500	-859	Low
E	6,000	-359	Low
F	6,600	+241	High
G	7,100	+741	High

Derived data

Class Map

แสดงการสร้างข้อมูลสารสนเทศใหม่ขึ้นมา จากการคำนวณหาค่าสูงกว่าและต่ำกว่าค่าเฉลี่ย (Mean) เพื่อจัดระดับชั้นของรายได้ (Low หรือ High) ของแต่ละอำเภอ

ข้อมูลที่สร้างขึ้นใหม่ (Derived Data) - Resampling

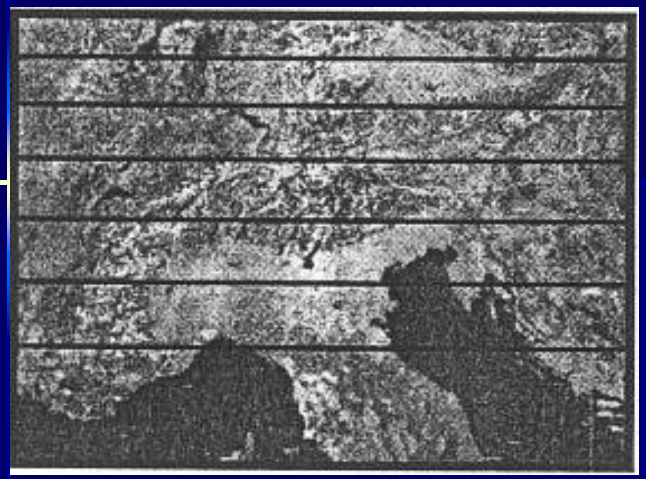


การเพิ่มจำนวนช่องกริด (raster cell) เป็นวิธีการหนึ่งในการปรับปรุงข้อมูล โดยเป็นการเพิ่มความละเอียดเชิงพื้นที่ (better resolution และ spatial detail)

Resampling

ค่า Value ของช่องกริดที่เพิ่มขึ้นมา จะเป็นค่า estimated value ที่ได้จากการคำนวณจากค่า value ของกริดที่อยู่ข้างเคียง

ข้อมูลที่สร้างขึ้นใหม่ (Derived Data) - Interpolation



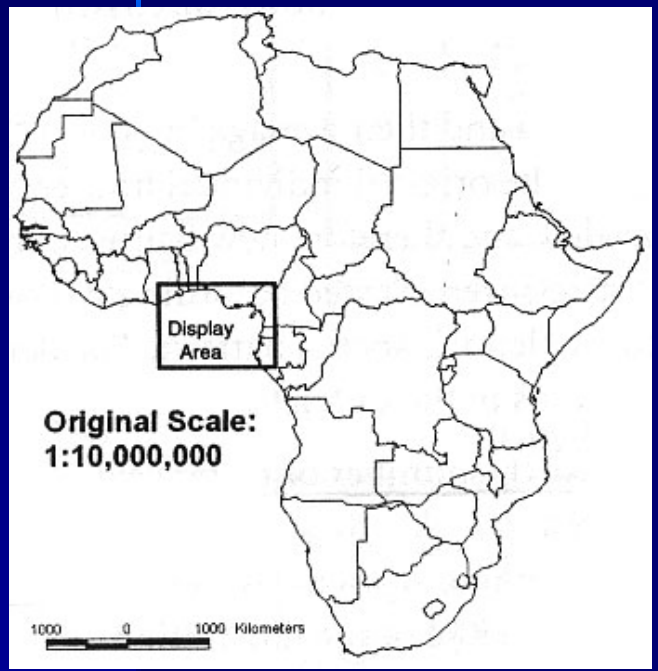
Missing scan line data

Interpolation repair

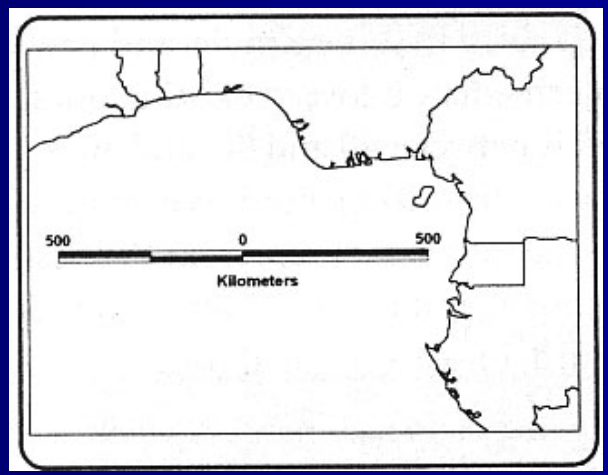
ข้อมูลภาพดาวเทียมที่เกิดความผิดพลาดแบบ miss line หรือ drop out ซึ่งจะทำให้ Pixel หายไปทั้งแถว (no-values lines) สามารถ สร้างข้อมูลที่สูญหายไปได้ โดยการใช้วิธี Resampling ในการคำนวณค่าของ Pixel ทั้งแถวที่หายไปขึ้นมาใหม่ จากการประมาณค่า value ของ Pixel ที่อยู่ข้างเคียง ซึ่งกระบวนการนี้เรียกกว่า **Interpolation**

มาตราส่วน (Scale)

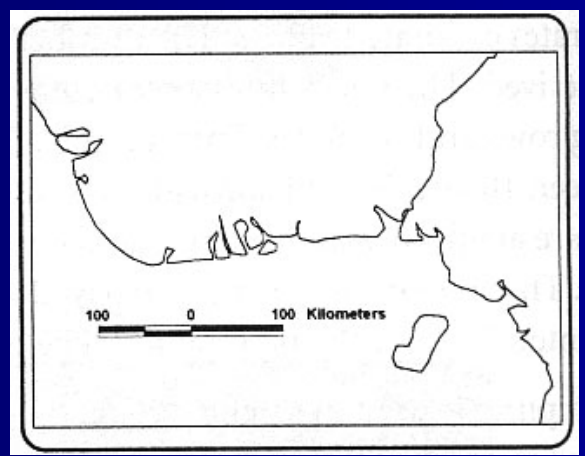
ถึงแม้ว่า GIS สามารถแสดงผลข้อมูลในมาตราส่วนที่หลากหลาย แต่ความเที่ยงตรง(Precision) ของข้อมูลจะขึ้นอยู่กับ มาตราส่วนของข้อมูลต้นฉบับที่ได้มา



ภาพต้นฉบับมาตราส่วน 1:10,000,000



ภาพที่แสดงในมาตราส่วนเดียวกันกับแผนที่ต้นฉบับ (1:10,000,000) ซึ่งแสดงให้เป็นถึงรายละเอียดของบริเวณชายฝั่งทะเล

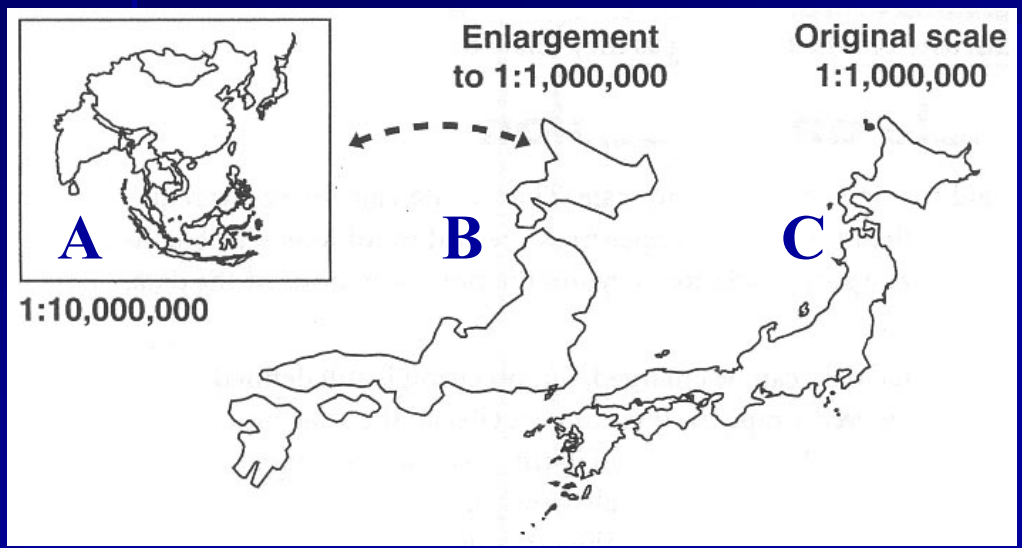


ภาพในมาตราส่วน 1:3,000,000 ขยายมากเกินไป เนื่องจาก feature ของชายฝั่งทะเลไม่สามารถแสดงรายละเอียดเพิ่มมากขึ้น

จรัณธร บุญญนุภาพ

ความแตกต่างกันของมาตราส่วน (Scale Differences)

ความเที่ยงตรง(Precision) ของข้อมูลจะถูกพิจารณาเพิ่มมากขึ้น เมื่อมีการนำข้อมูลที่มีมาตราส่วนแตกต่างกันมาใช้ใน GIS เนื่องจากจะมีผลต่อความถูกต้องเชิงพื้นที่ (Spatial accuracy) และคุณภาพของข้อมูล

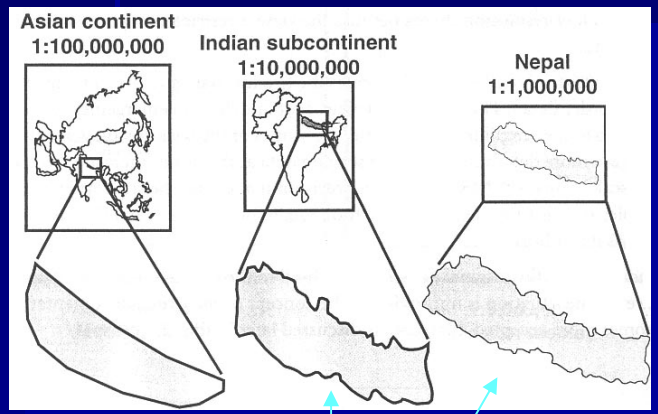


เปรียบเทียบภาพมาตราส่วน 1:1,000,000 (ภาพ B) ที่ถูกขยายมาจากข้อมูลต้นฉบับ (ภาพ A) ที่มีมาตราส่วน 1:10,000,000 และภาพต้นฉบับมาตราส่วน 1:1,000,000 (ภาพ C)

สังเกตได้ว่า ถึงแม้จะขยายมาตรส่วนในการแสดงภาพ แต่รายละเอียดและความถูกต้องของ feature ไม่ได้เพิ่มมากขึ้นตามมาตราส่วนที่ถูกขยาย

ความไม่เหมาะสมของมาตราส่วน (Scale Incompatibility)

ในการแสดงข้อมูลภาพหรือการผลิตแผนที่จำเป็นที่จะต้องคำนึกถึงมาตราส่วนของข้อมูลต้นฉบับ (original scale of data source) เนื่องจากข้อมูลบางอย่างมีรายละเอียดและคุณภาพ (resolution) มากเกินไป ในการแสดงผลด้วยมาตราส่วนขนาดเล็ก



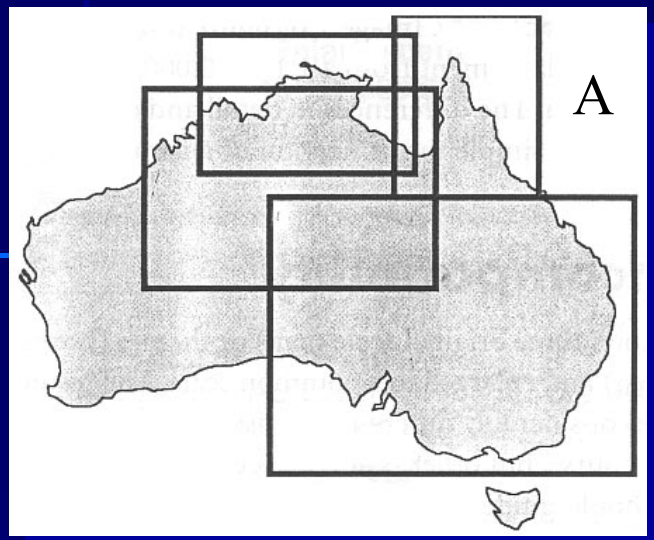
มาตราส่วน 1:50,000

ภาพมาตราส่วน 1:30,000

ภาพมาตราส่วน 1:3,000

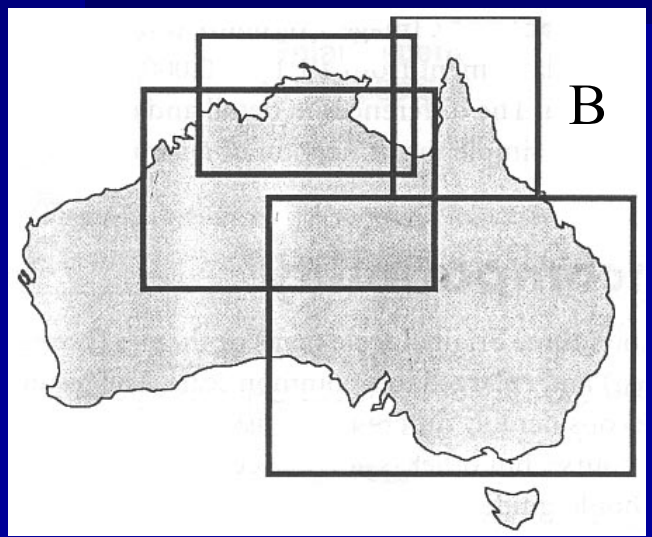
แผนที่ที่มีการย่อมาตราส่วน โดยไม่ผ่านการวางนัยทั่วไป (Generalization) หรือการลดทอนข้อมูล

Incomplete Area Coverage



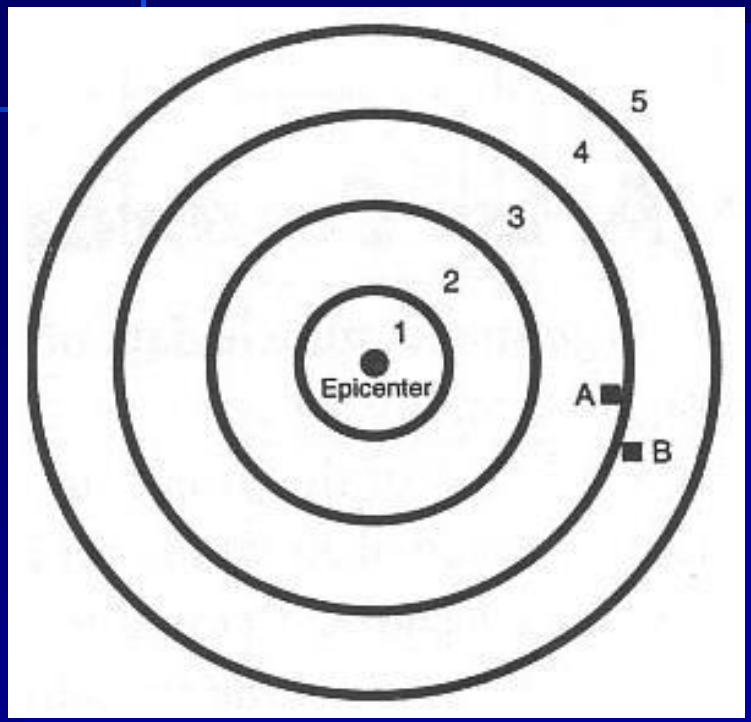
การดำเนินงานทาง GIS จะต้องมีข้อมูลครอบคลุมพื้นที่ศึกษาอย่างสมบูรณ์ โดยเฉพาะพื้นที่ศึกษาที่มีขนาดใหญ่ มักจะไม่สามารถหาข้อมูลได้คลอบคลุมทั้งหมด ดังนั้นจำเป็นต้องมีการเก็บรวบรวมข้อมูลที่ขาดหายไป หรือไม่ก็ต้องลดขนาดของพื้นที่ศึกษาลง โดยขึ้นอยู่กับการครอบคลุมพื้นที่ของมูลที่มีอยู่ ดังภาพ A

Smallest Scale Rule



โดยปกติแล้ว การทำงานใน GIS จะใช้ข้อมูลที่มีมาตราส่วนที่แตกต่างกัน ซึ่งรายละเอียดและความถูกต้อง (accuracy) ของผลลัพท์ที่ได้จะถูกกำหนดโดย ข้อมูลที่มีมาตราส่วนขนาดเล็กที่สุด (smallest applied scale)

Continuous Data Interpretation



Buffer Zone

ข้อมูลที่มีลักษณะต่อเนื่อง (continuous data) จะยากในการแสดงผลให้เห็นถึงความแตกต่างของ feature อย่างชัดเจน ดังนั้นจึงต้องแสดงผลในลักษณะของการแบ่งเขตที่ไม่ต่อเนื่อง (artificial discrete zone)

จากภาพตัวอย่างแสดงถึง จุด A และ B ที่ตั้งอยู่ภายในเขตระดับความรุนแรงของการเกิดแผ่นดินไหว ที่แตกต่างกัน ด้วยการทำ Buffering เพื่อกำหนดขอบเขตพื้นที่ระดับความรุนแรงจากจุดศูนย์กลางที่เกิดแผ่นดินไหว

ความสอดคล้องเชิงตรรกะ (Logical Consistency)

แสดงถึงความเป็นเหตุเป็นผลที่มีความสัมพันธ์กันระหว่างองค์ประกอบของข้อมูล

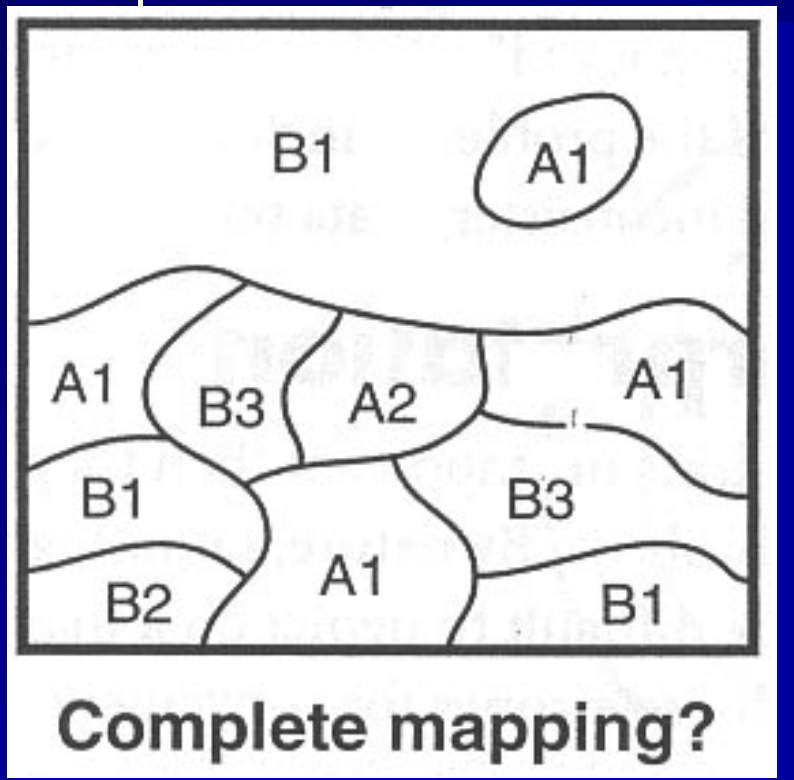
โดยทั่วไป ฐานข้อมูล GIS จะแสดงความสอดคล้องกันในรูปแบบที่สม่ำเสมอ ตัวอย่างเช่น จะถูกสร้างจาก แหล่งข้อมูลเดียวกัน และถูกเก็บรวมรวม แก้ไข จำแนกประเภท และประมวลผล ภายใต้ วิธีการเดียวกัน

ความสมบูรณ์ (Completeness)

- เกี่ยวข้องกับความผิดพลาดของข้อมูลที่หายไป หรือถูกละเลย (omission)

ความสมบูรณ์และความสอดคล้องของข้อมูล

(Complete and Consistent Data)



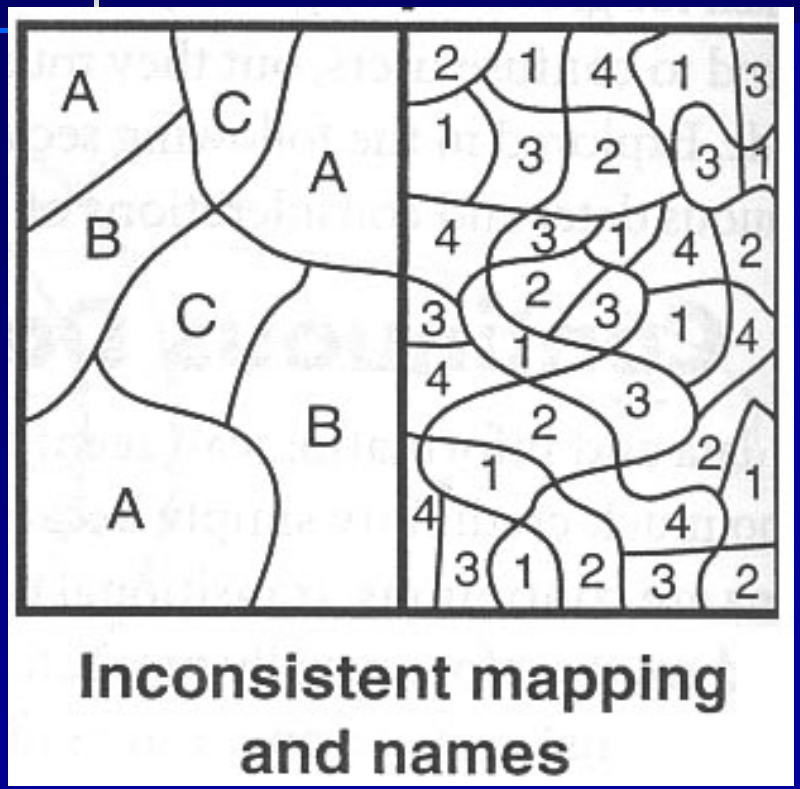
ความสมบูณ์ของข้อมูล อาจจะพิจารณาในด้านถูกต้องของข้อมูลเชิงคุณลักษณะ (attribute) รวมทั้งการมีข้อมูลครอบคลุมพื้นที่ศึกษาทั้งหมด

ความสมบูรณ์ทำให้ต้องเสียเวลาในการตรวจสอบความถูกต้องของข้อมูลที่ได้มา และอาจจะต้องทำการเก็บรวบรวมข้อมูลเพิ่มเติมอีก

ภาพตัวอย่าง แสดงให้เห็นถึงความแตกต่างของรายละเอียดในการจำแนกกลุ่มชุดดินในส่วนด้านล่างและด้านบนของภาพ ซึ่งทำให้ไม่แน่ใจว่าข้อมูลส่วนบนของภาพเกิดการสูญหาย หรือ มีการจำแนกที่ไม่สมบูรณ์ หรือ เป็นลักษณะที่แท้จริงที่เกิดขึ้นในสภาพพื้นที่

ความสมบูรณ์และความสอดคล้องของข้อมูล

(Complete and Consistent Data)



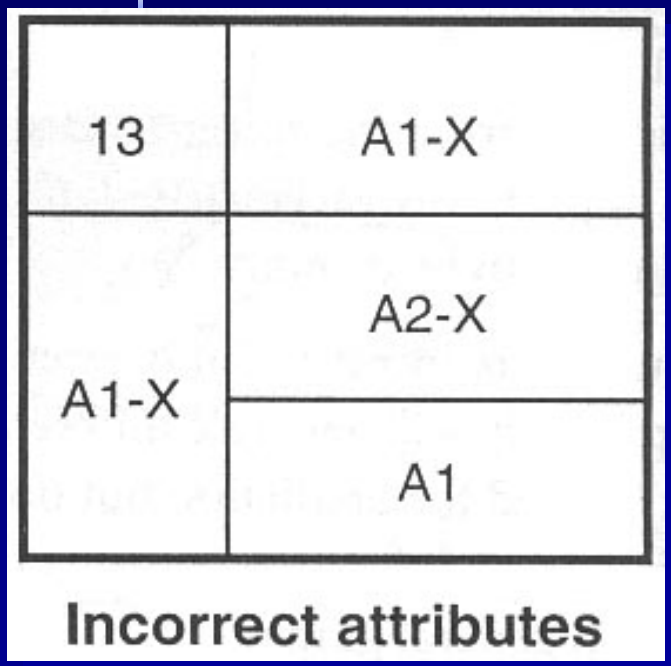
Inconsistent mapping and names

ความไม่สอดคล้องกัน (inconsistent) ของรายละเอียดข้อมูล อันเนื่องมากจากวิธีการเก็บข้อมูลแตกต่างกัน อาจรวมถึงระบบในการจำแนกประเภทข้อมูล (classification scheme) ที่แตกต่างกันด้วย

จากภาพตัวอย่าง เห็นว่าพื้นที่ด้านซ้ายมือมีลักษณะการจำแนกประเภทข้อมูลที่ละเอียดกว่าพื้นที่ด้านขวามือ ทำให้พื้นที่ทั้งสองไม่สามารถเชื่อมต่ออย่างสอดคล้องกันได้

นอกจากนั้นระบบของประเภทข้อมูลของพื้นที่ทั้งสองก็แตกต่างกัน (ตัวอักษร กับ หมายเลข)

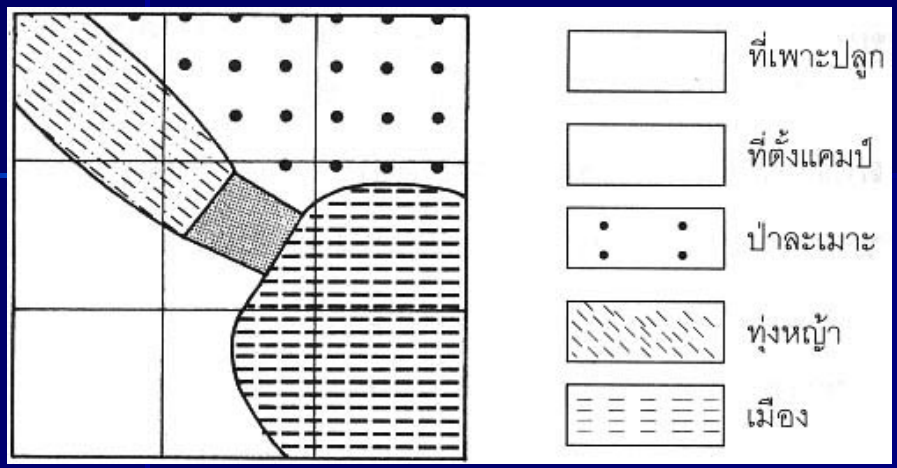
ความสมบูรณ์และความสอดคล้องของข้อมูล (Complete and Consistent Data)



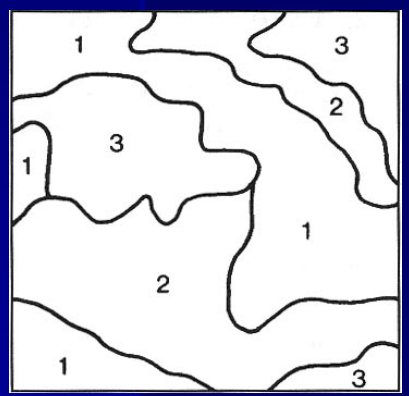
จากภาพแสดงถึง ความไม่สอดคล้องกันในการแสดงผลข้อมูลเชิงคุณลักษณะ (attribute label) โดย polygon 13 และ polygon A1 อาจจะแสดงlabel ที่ถูกต้องหรือไม่ถูกต้องก็ได้

ดังนั้นอาจทำให้เสียเวลาในการตรวจสอบความถูกต้องของข้อมูลเชิงคุณลักษณะของ polygon ทั้งสอง

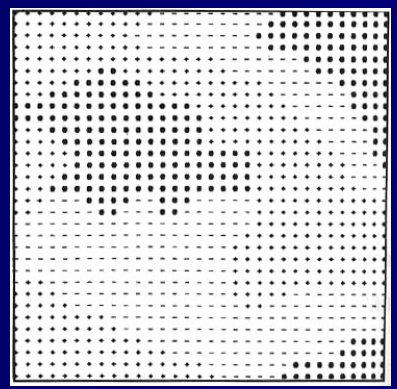
ความผิดพลาดที่เกิดจากการแปลงข้อมูล Raster เป็น Vectoe



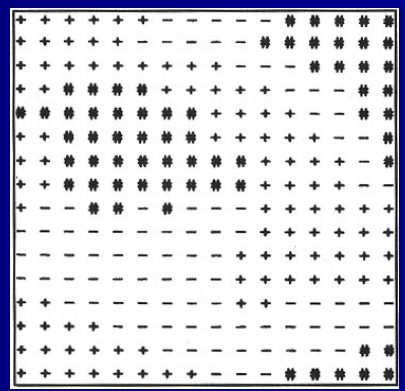
เนื่องจากช่องกริดแต่ละช่อง (grid cell) จะมีค่าของ feature ได้เพียงแค่ค่าเดียวเท่านั้น ดังนั้นค่าที่กำหนดให้กับกริดแต่ละช่องคือ ค่าเฉลี่ยของ feature ทั้งหมดที่อยู่ภายในช่องกริดนั้น ๆ ซึ่งทำให้เกิดการสูญเสียรายละเอียดของลักษณะทางภูมิศาสต์ (geographic features)



Vector data ต้นฉบับ



กริดขนาด 32 x 32



กริดขนาด 16 x 16

ในการแปลงข้อมูลจาก raster เป็น vector (rasterlization) ถ้าเพิ่มจำนวนของกริดหรือกำหนดให้กริดมีขนาดเล็กลง จะเป็นการลดข้อผิดพลาดหรือลดการสูญหายของรายละเอียดของลักษณะทางภูมิศาสตร์ลงได้

จรัณธร บุญญนุภาพ

ตัวอย่างของคุณภาพข้อมูล

ประเภท	นิยาม	ตัวอย่าง
ความถูกต้องเชิงตำแหน่ง	ระดับความถูกต้องในแนวตั้งและแนวระนาบของระบบพิกัดโลก	ความถูกต้องในรายละเอียดของพิกัดโลกในแนวตั้งและแนวระนาบ
ความถูกต้องเชิงคุณลักษณะ	ระดับความถูกต้องของข้อมูลเชิงคุณลักษณะ ซึ่งเชื่อมโยงกับข้อมูลเชิงพื้นที่	ความถูกต้องของคำอธิบายชนิดของดินในแผนที่ เมื่อเปรียบเทียบกับชนิดของดินที่เก็บตัวอย่างจากพื้นที่จริง
ความถูกต้องสมบูรณ์	ระดับความผิดพลาดและวิธีการในการจัดการกับข้อมูลที่ผิดพลาด	ความสามารถในการประเมินระดับของอาชญากรรมในพื้นที่เฉพาะ ซึ่งอาจก่อให้เกิดความผิดพลาด ถ้าข้อมูลไม่สมบูรณ์
ความสอดคล้องเชิงตรรกะ	ระดับความขัดแย้งในความสัมพันธ์ภายใต้ระบบฐานข้อมูล	ความแตกต่างกันในเรื่องของระดับชั้นของการจำแนกประเภทการใช้ที่ดิน

มาตรฐานและรูปแบบในการจัดสร้างและจัดเก็บข้อมูล

- เป็นปัญหาอันดับแรก ของการจัดการสารสนเทศขององค์กรขนาดใหญ่
- ปัจจุบันยังไม่มีมาตรฐานในการสร้าง บันทึก หรือ จัดเก็บข้อมูล ทำให้แต่ละหน่วยงาน ต่างก็ทำการสร้างและจัดเก็บฐานข้อมูลในรูปแบบที่ตนเองถนัด
  - รูปแบบ (format)
  - รหัส (Code)
  - สื่อบันทึกข้อมูล (Storage Device)

มาตรฐานของแผนที่ในระบบตัวเลข (Digital map)

- Format ใหม่ที่เรียกว่า Digest
- ประกอบด้วยข้อมูลแผนที่ที่เป็นแบบ Raster และ Vector
- เป็นมาตรฐานใหม่ที่ใช้ในกองทัพสหรัฐอเมริกาและกลุ่มประเทศNATO
- format digest ประกอบด้วยข้อมูล ที่เป็น Vector เช่น Vmap
- raster เช่น DTED,ADRG,CIB ซึ่งในการใช้งานในส่วนของการทหารของประเทศไทย
- เพื่อใช้ในการแลกเปลี่ยนข้อมูลกับทางพลเรือนและใช้งานร่วมกับพลเรือน
- format ของพลเรือน
- 7 format ดังนี้ Vmap,DTED,ADRG,Geotiff,Arc/info,Dgn,Bil(format ของดาวเทียม)

Metadata (data about data)

- **Metadata คืออะไร ?**

Metadata หรือ “ข้อมูลของข้อมูล” เป็นข้อมูลที่อธิบายรายละเอียดต่างๆ เกี่ยวกับชุดข้อมูล เช่น ชื่อเรื่อง, หน่วยงานที่เกี่ยวข้อง, ความเป็นมา, ขั้นตอนการจัดทำชุดข้อมูล, แหล่งที่มาของข้อมูล, ขอบเขตหรือพื้นที่ครอบคลุมของข้อมูล, วันเวลาของข้อมูล, การใช้และการเข้าถึงข้อมูล เป็นต้น

กำหนดโดยรูปแบบมาตรฐานการอธิบายข้อมูลระบบสารสนเทศภูมิศาสตร์ ของคณะกรรมการข้อมูลทางภูมิศาสตร์แห่งรัฐ (Federal Geographic Data Committee: FGDC)

Metadata

Metadata จะช่วยตอบคำถามให้เราในเรื่องต่อไปนี้

What ?	เป็นข้อมูลเกี่ยวกับอะไร มีรายละเอียดของข้อมูลอย่างไร
Why ?	เหตุผลหรือจุดประสงค์ในการสร้างหรือการใช้
When ?	ข้อมูลถูกจัดทำขึ้นเมื่อใด และมีระยะเวลาในการ Update อย่างไร
Who ?	ใครเป็นเจ้าของข้อมูล และได้รับการสนับสนุนข้อมูลจากใครบ้าง
Where ?	ขอบเขตของข้อมูลมีค่า latitude/longitude, พิกัด, ครอบคลุมพื้นที่ใด
How ?	มีขั้นตอนในการสร้างข้อมูลอย่างไร และสามารถเข้าถึงข้อมูลได้อย่างไร

Metadata มีประโยชน์อย่างไร ?

1. ทำให้ผู้ใช้เกิดความเชื่อมั่นและเลือกใช้ข้อมูลได้อย่างถูกต้อง
2. มีความมั่นใจในการลงทุนในเรื่องของข้อมูล
3. การเปลี่ยนแปลงบุคคลากรในการปฏิบัติงานจะไม่ส่งผลกระทบต่อระบบงาน
4. เกิดการแลกเปลี่ยนการใช้ข้อมูลระหว่างหน่วยงานทั้งในและนอกประเทศ
5. ลดการสูญเสียงบประมาณและเวลาในการจัดทำข้อมูลที่ซ้ำซ้อนกัน
6. มีความเป็นมาตรฐานในการจัดทำและการจัดเก็บข้อมูล

Metadata สำหรับประเทศไทยจะถูกนำมาใช้เป็นมาตรฐานได้อย่างไร

หน่วยงานที่จัดทำข้อมูลและเป็นเจ้าของข้อมูลต้องให้ความสำคัญกับการจัดทำ Metadata และร่วมมือกันผลักดันให้เกิดการนำไปใช้จนกลายเป็นมาตรฐานนิยม

จรัณธร บุญญนุภาพ

GIS Metadata

การอธิบายข้อมูลระบบสารสนเทศภูมิศาสตร์ หรือ GIS Metadata เป็นการกำหนดแบบแผนหรือรูปแบบรายละเอียดในเรื่องของ

- การระบบ/จำแนก (identification)
- ขอบเขต (Extent)
- คุณภาพ (Quality)
- ข้อมูลเชิงพื้นที่ (Spatial)
- แบบแผนของข้อมูลเชิงเวลา (Temporal Schema)
- การอ้างอิงข้อมูลเชิงพื้นที่ (Spatial Reference)
- การเผยแพร่ของชุดข้อมูลทางภูมิศาสตร์เชิงตัวเลข (Distribution)

Report of Questionnaire (Metadata contents)

แบบสำรวจดัชนีข้อมูลระบบสารสนเทศภูมิศาสตร์ (**Metadata** ระดับที่ 1)

1.ชื่อเรื่อง/ชั้นข้อมูล (**Title**) * : เส้นทางคมนาคม

2.วันที่สำหรับอ้างอิง (**Reference Date**) * : 31 ธันวาคม 2540

3.ข้อมูลหน่วยงานรับผิดชอบ (**Responsible Party Information**)

3.1 ชื่อหน่วยงานรับผิดชอบ (**Responsible Party Organization Name**)

ฝ่าย/ส่วน :

กอง/สำนัก * : สารสนเทศสิ่งแวดล้อม

กรม * : ส่งเสริมคุณภาพสิ่งแวดล้อม

ทบวง/กระทรวง * : วิทยาศาสตร์ เทคโนโลยีและสิ่งแวดล้อม

3.2 บทบาทของหน่วยงาน (**Responsible Party Role**) * : ผู้เผยแพร่ (**Distributor**)

3.3 ที่อยู่ (**Address**): 60/1 ซ.พิบูลย์วัฒนา7

3.4 ตำบล/แขวง (**Tambon**): สามเสนใน

3.5 อำเภอ/เขต (**Amphoe**): พญาไท

3.6 จังหวัด (**Province**): กรุงเทพมหานคร

3.7 รหัสไปรษณีย์ (**Postal code**): 10400

3.8 ประเทศ * (**Country**): ไทย

3.9 โทรศัพท์ (**Telephone**): 0 2279 8955

3.10 โทรสาร (**Facsimile**): 2 2272 3038

3.11 การเชื่อมต่อทางสาย (**On-line Resource Linkage**):

4.ขอบเขตของชุดข้อมูล (**Dataset Extent**)

ชื่อขอบเขตทางภูมิศาสตร์ (**Geographic Extent Name**) * : จังหวัดปราจีนบุรี จังหวัดชลบุรี

วันที่สำรวจ/เก็บข้อมูล (**Temporal Extent Date/Time**) : 30 มิถุนายน 2538

5.ระดับของความละเอียด (**Resolution Level**) : 1:40K-1:199K

6.ภาษาของชุดข้อมูล (**Language of Dataset Code**) * : ไทย อังกฤษ

7.บทคัดย่อ (**Abstract**) * : เป็นเส้นทางคมนาคม ทางหลวงประเทศ ทางหลวงจังหวัด และเส้นทางที่ใช้เฉพาะบางฤดูกาล

8.จุดมุ่งหมาย (**Purpose**) : เพื่อใช้ในการศึกษาความเจริญและการติดต่อคมนาคมของชุมชนในจังหวัดพื้นที่ศึกษา

9.ความก้าวหน้า (**Progress**) : **Complete** (เสร็จสมบูรณ์)

10.ลักษณะข้อมูล (**Category**)

10.1 ชั้นข้อมูล (**Theme**) * : การสื่อสารและคมนาคม

10.2 ข้อมูลคำสำคัญ (**Keyword Information**)

ลำดับที่ 1

คำสำคัญ (**Keyword**) : เส้นทางคมนาคม

ประเภทคำสำคัญ (**Keyword Type**) : **Theme** (หัวข้อ)

ลำดับที่ 2

คำสำคัญ (**Keyword**) : **Trans**

ประเภทคำสำคัญ (**Keyword Type**) : **Theme** (หัวข้อ)

ลำดับที่ 3

คำสำคัญ (**Keyword**) : **transport**

ประเภทคำสำคัญ (**Keyword Type**) : **Theme** (หัวข้อ)

11.ข้อจำกัดในการเข้าถึง (**Access Constraints**) : ได้รับอนุมัติจากอธิบดี

12.ข้อจำกัดในการใช้ (**Use Constraints**) : อ้างอิงระเบียบการรักษาความปลอดภัยแห่งชาติ 2517

13.รายงานคุณภาพเชิงลึก (**Qualitative narrative report**) : ตรวจสอบความถูกต้องโดยใช้ **GPS** ในการสำรวจภาคสนาม

14.รายการของข้อมูลประวัติการเป็นมา (**Lineage**) : นำเข้าข้อมูลจากแผนที่ 1:50,000 ของกรมแผนที่ทหาร และข้อมูลจากภาพถ่ายดาวเทียม และการเก็บข้อมูลพิกัดด้วย **GPS**

15.ประเภทของการนำเสนอข้อมูลเชิงพื้นที่ (**Spatial Representation Type**) : **Text** (ตัวหนังสือ) **Vector** (ข้อมูลเชิงเส้น)

16.ประเภทของระบบการอ้างอิงข้อมูลเชิงพื้นที่ (**Spatial Reference System Type**) : **Geographic Identifiers** (การจำแนกทางภูมิศาสตร์)

17.การระบุ/จำแนกในการเผยแพร่ (**Distribution Identifier**) : **transport**

18.ชื่อรูปแบบของการเผยแพร่ (**Distribution Format Name**) : **Arc/Info**, **e**00, **dxf**

19.สื่อที่ใช้ในการเผยแพร่ (**Distribution Media**) : **CD-ROM** 4 **mm cartridge tape** 8 **mm cartridge tape**

20.ระดับมาตรฐานการอธิบายข้อมูล (**Level of Conformance**) * : ระดับที่ 1

21.ภาษาของการอธิบายข้อมูล (**Language of Metadata**) * : ไทย

22.วันที่กรอกแบบสอบถาม (**Metadata Date**) * : 5 เมษายน 2545

จรัณธร บุญญนุภาพ

การอธิบายข้อมูล GIS

ส่วนที่ 1 ส่วนข้อมูล (information Sections) ประกอบไปด้วย 8 ส่วนย่อย

1. ข้อมูลในการระบุ/จำแนก (Identification Information)
2. ข้อมูลที่เกี่ยวกับคุณภาพข้อมูล (Data Quality Information)
3. ข้อมูลที่เกี่ยวกับข้อมูลประวัติการเป็นมา (Lineage Information)
4. ข้อมูลการนำเสนอข้อมูลเชิงพื้นที่ (Spatial Data Representation Information)
5. ข้อมูลระบบอ้างอิง (Reference System Information)
6. ข้อมูลของรายการ/รูปแบบข้อมูล (Feature Cataloging Information)
7. ข้อมูลในการเผยแพร่ (Distribution Information)
8. ข้อมูลการอ้างอิงในการอธิบายข้อมูล (Metadata reference information)

การอธิบายข้อมูล GIS

ส่วนที่ 2 ส่วนสนับสนุน (Repeating Entities) ประกอบด้วย

9. ข้อมูลส่วนสนับสนุนในการอ้างอิงทั่วไป (Citation Information Entity)
10. ข้อมูลส่วนสนับสนุนเกี่ยวกับหน่วยงานรับผิดชอบ (Responsible Party Information Entity)
11. ข้อมูลส่วนสนับสนุนเกี่ยวกับที่อยู่ (Address Information Entity)
12. ข้อมูลส่วนสนับสนุนด้านขอบเขตของข้อมูล (Extent Information Entity)
13. ข้อมูลส่วนสนับสนุนในการเชื่อมต่อทางสาย (On-line Resources information Entity)

ดัชนีข้อมูลระบบสารสนเทศภูมิศาสตร์ (GIS Data Index)

- เพื่อบ่งชี้ว่าข้อมูลที่ต้องการนั้นมีอยู่ที่หน่วยงานใดบ้าง
- การจัดเก็บข้อมูลเป็นไปตามรูปแบบมาตรฐานการอธิบายข้อมูลระบบสารสนเทศภูมิศาสตร์ระดับที่ 1 ตามร่างมาตรฐาน ISO 19115 (Metadata)
- ดำเนินการโดย ศูนย์ข้อมูลข้อสนเทศ สำนักงานปลัดกระทรวงวิทยาศาสตร์ เทคโนโลยีและสิ่งแวดล้อม

การกำหนดชั้นข้อมูลพื้นฐาน

ประกอบด้วย

- ขอบเขตการปกครอง
- ลักษณะภูมิประเทศ
- การใช้ที่ดิน
- ทรัพยากรน้ำ
- ทรัพยากรดิน
- ทรัพยากรป่าไม้
- ทรัพยากรธรณี
- เส้นทางคมนาคม