

تاریخچه سیستم‌های اطلاعات جغرافیایی

- اولین داده مکانی در ۲۰۰ سال قبل از میلاد توسط اراتوستن از جغرافیدانان قدیم یونان ارائه شد.
- از دهه ۱۹۷۰ به دلیل امکان دسترسی به رایانه‌های مناسب، تکنولوژی لازم برای کار با داده‌های مکانی بوجود آمد و سیستم‌های اطلاعات جغرافیایی برای افزایش توان تحلیل حجم زیادی از داده‌ها بوجود آمدند.
- پیشرفت GIS در دهه‌ی ۱۹۶۰ را می‌توان مدیون سه عامل دانست که عبارت‌اند از: پیشرفت سریع محاسبات کامپیوتری، پیشرفت در علوم فتوگرامتری، کارتوگرافی و تمایل به بررسی کمی داده‌ها در علوم و فنون مرتبط با زمین.

قانون اول جغرافیا: تمامی اشیاء با اشیاء دیگر در ارتباط‌اند ولی دو شیء نزدیک به هم (از نظر فضایی و زمانی) ارتباط بیش‌تری دارند تا دو شیء دور از هم.

تعریف سیستم: سیستم مجموعه‌ای از اجزاء محدود است که برای رسیدن به هدف خاصی با هم همکاری می‌کنند. یا به عبارتی دیگر سیستم شامل عناصر اصلی و روابط بین آن‌ها و مجموعه‌ای از روابط میان سیستم و محیط آن است.

تعریف داده و انواع آن

- داده پایه و اساس هر سیستم و کوچک‌ترین جزء یک سیستم است. داده‌های خام در سیستم‌ها پردازش می‌شوند تا برای تولید اطلاعات جهت تصمیم‌گیری‌های صحیح به کار روند.
- داده یا Data عبارت است از تشریح کمی و کیفی ویژگی‌های پدیده یا به بیانی دیگر، توصیف پدیده‌ها با عدد، متن، گرافیک، مختصات و ...
- داده‌های رقومی اطلاعات کد شده و ساختار یافته برای پردازش‌های خاص است که عموماً در یک سیستم رایانه‌ای شکل می‌گیرد.
- آگاهی‌های نامرتب اولیه را داده و آگاهی‌های مرتب شده را اطلاعات می‌گویند.
- داده‌های جغرافیایی، داده‌هایی‌اند که دارای ارزش مکانی‌اند، بدین معنا که این داده‌ها برای نمایش ویژگی‌های مکانی مانند موقعیت جاده‌ها، شهرها و روستاها و همچنین برای تعریف و توصیف ارزش‌های غیرمکانی مانند اندازه، وسعت، شدت، تراکم و ... به کار می‌روند.
- منظور از داده‌های جغرافیایی یا فضایی، داده‌هایی است که نمایشگر پدیده‌های مربوط به جهان واقعی بر حسب:
 ۱. موقعیت آن‌ها نسبت به سیستم مختصاتی شناخته شده
 ۲. ویژگی‌های توصیفی آن‌ها مانند ارتفاع از سطح دریا، نوع خاک و پوشش گیاهی
 ۳. روابط فضایی آن‌ها نسبت به هم که نشان‌دهنده چگونگی اتصال و ارتباط پدیده‌هاست. این ویژگی به عنوان توپولوژی شناخته می‌شود و مشخص کننده فضا و ویژگی‌های فضایی پدیده‌ها مانند متصل بودن است.

متادیتا: Data about Data

متادیتا یا فراداده اغلب به داده‌هایی گفته می‌شود که داده‌های دیگر را توصیف می‌کنند. در واقع داده‌های مرجع ساختاریافته‌ای هستند که به مرتب‌سازی و شناسایی صفات داده‌هایی که توصیف می‌کنند، کمک خواهند کرد. متادیتا عبارت است از داده‌هایی که در مورد داده‌های ما توضیح می‌دهند. متادیتا به سوالات زیر پاسخ می‌دهد:

- اطلاعات ما چی هستند؟ (what)
- این اطلاعات برای چی هست؟ (why)
- این اطلاعات را چه کسی و چه ارگانی جمع‌آوری کرده است؟ (who)
- زمان تهیه اطلاعات (when)
- اطلاعات چگونه تهیه شده‌اند؟ (who)
- مربوط به کجاست؟ (where)

نکته: متادیتا توصیف می‌کند که چه کسی، چه چیزی، چه زمانی، کجا، چرا و چگونه با داده‌ها تعامل داشته باشد.

مزایای متادیتا

۱. ماهیت داده‌ها مشخص می‌شود
۲. داده‌ها را مستقل از اشخاص می‌کند (یعنی نیازی به اشخاص برای توضیح داده‌ها نیست و خود متادیتا توضیح می‌دهد)
۳. امکان جست‌وجو روی داده‌ها را می‌دهد.

عناصری می‌توانند به عنوان داده‌ای جغرافیایی ثبت شوند که:

۱. دارای موقعیت مکانی باشد.
۲. قابل تعریف باشد.
۳. دارای مقیاس باشد.
۴. قابل طبقه‌بندی یا به عبارتی قابل پهنه‌بندی باشد.

Data Dictionary

اطلاعاتی را که در رابطه با نحوه ذخیره‌سازی داده‌ها به ما می‌دهد که این اطلاعات می‌تواند شامل نوع داده، حداکثر طول فیلد، اسم فیلد و ... باشد.

سیستم‌های اطلاعات

سیستم‌هایی هستند که داده‌ها یا اطلاعات را نگهداری و پردازش می‌کنند. نوع داده مهم‌ترین عنصر در سیستم اطلاعات است. برخی از این سیستم‌ها علاوه بر اعداد و حروف با نقشه نیز کار می‌کنند که این سیستم‌ها را سیستم اطلاعات مکانی می‌نامند.

تعاریف سامانه اطلاعات جغرافیایی به چهار دسته ابزار-مبنا، پایگاه داده، سیستم-مبنا و سازمان-مبنا تقسیم می‌شود.

در تعاریف سازمان - مبنا نقش سازمانها و مردم در داده‌های مکانی بیش تر از ابزارهای مورد نیاز، مورد تاکید است.

تعاریف GIS

- مجموعه‌ای سازمان یافته از سخت‌افزار، نرم‌افزار و نیروهای متخصص که جهت اخذ و ذخیره‌سازی، بهنگام‌سازی، بازیافت، تغییر و تحول، تلفیق یا یکپارچه‌سازی، نمایش و تجزیه و تحلیل داده‌های مکان مرجع و گرفتن خروجی مورد استفاده قرار می‌گیرد.
- یک سیستم اطلاعات مکانی (GIS)، به جمع‌آوری، تولید، نگهداری، بازیابی، تجزیه و تحلیل داده‌هایی می‌پردازد که در دنیای واقعی فضایی را اشغال نموده‌اند.
- سیستم‌های اطلاعات جغرافیایی یک سیستم رایانه‌ای است که مانند سایر سیستم‌های رایانه‌ای از مولفه‌های ورود داده، مدیریت داده، تجزیه و تحلیل داده و خروجی داده تشکیل می‌شود.
- سامانه اطلاعات جغرافیایی مجموعه‌ای از ابزار قدرتمند برای جمع‌آوری، ذخیره، استخراج، آنالیز و تبدیل و نمایش داده‌های مکانی از جهان واقعی برای حل مسئله یا مسائل معین است.

نکته: تلفیق داده‌های غیرمکانی و داده‌های مکانی از توانمندی‌های مهم GIS در نظر متخصصان انواع رشته‌های مختلف می‌باشد.

یک سیستم اطلاعات جغرافیایی دارای ۵ جزء اصلی است: سخت‌افزار، نرم‌افزار، داده، توابع، کاربران متخصص.

- سخت‌افزار: در یک GIS سه نوع سخت‌افزار داریم. سخت‌افزارهای ورود داده، سخت‌افزارهای ذخیره‌سازی و پردازش داده‌ها و سخت‌افزارهایی که خروجی را نمایش می‌دهند.
- نرم‌افزار: یک سیستم GIS خطی و کاغذی نیست بلکه دارای ماهیت رقومی و کامپیوتری است.
- داده‌ها: ۸۰ درصد داده‌های جهان واقعی مکانی هستند از بین ۵ جزء اصلی GIS داده مهم‌ترین جزء است ما دو نوع داده داریم: ۱. داده‌های مکانی و ۲. داده‌های غیرمکانی: که شامل اطلاعات توصیفی و زمانی است که ماهیت مکانی ندارند.
- توابع: شش نوع توابع داریم: جمع‌آوری داده‌ها (Data Acquisition)، حفظ و نگهداری داده (Data maintenance)، بازیافت داده‌ها (Data Retrieval)، پرسش و پاسخ (Query)، تجزیه و تحلیل (Analysis)، نمایش (Presentation).

- کاربران متخصص: کار آن دستور دادن، طراحی و کار کننده و دستور دهنده GIS هستند. کار کاربران متخصص تعریف توابع و الگوریتم‌ها، ساختار داده‌ها و نوع داده‌ها، طراحی سیستم و تعیین بهترین نرم‌افزارها و توابع لازم بر روی نرم‌افزارها.

نکته: در یک تقسیم‌بندی دیگر، سامانه اطلاعات جغرافیایی از ۶ جزء مشخص و قابل تفکیک تشکیل شده است که عبارتند از: شبکه، نرم‌افزار، داده‌ها، روش‌ها، سخت‌افزار، مردم و سازمان‌ها

نکته: رقومی‌کننده مهم‌ترین سخت‌افزار مورد استفاده در سامانه اطلاعات جغرافیایی برای ورود داده‌هاست.

عوامل مهم در یک سیستم GIS کارآمد:

۱. بدست آوردن داده‌های مناسب
۲. سازماندهی داده‌ها
۳. مدل تصمیم‌گیری
۴. اعتبارسنجی خروجی مدل (های) تصمیم‌گیری

فاکتورهای مهم در کیفیت داده‌ها:

۱. دقت: بیانگر این است که داده‌ها تا چه اندازه صحیح می‌باشند.
۲. کامل بودن: بیانگر بودن کامل بودن داده‌هاست.
۳. به روز بودن: بیانگر این است که داده‌ها تا چه اندازه به روز و جدید هستند.
۴. زمان: داده‌ها در چه زمان یا در چه دوره زمانی جمع‌آوری شده‌اند.

مولفه‌های اصلی داده‌های جغرافیایی

۱. بعد فیزیکی یا طبقه‌بندی: به عنوان مثال: جمعیت یک شهر یا پهناى یک جاده (بعد فیزیکی)، نوع صخره یا نوع پوشش گیاهی (طبقه‌بندی)
۲. موقعیت مکانی: معمولاً نسبت به یک سیستم مختصات مانند طول و عرض جغرافیایی تعیین می‌شود.
۳. زمان: اگر چه مولفه زمان به صورت صریح ذکر نمی‌شود، ولی دارای اهمیت حیاتی می‌باشد.

نکته: توانایی اصلی یک سیستم اطلاعات جغرافیایی (GIS) ترکیب کردن داده‌های زمین مرجع با اطلاعات توصیفی است که عملیاتی نظیر جست‌وجوی مکانی و انطباق لایه‌های مختلفی از داده‌ها را شامل می‌شود.

ورودی داده‌ها

وارد کردن داده‌ها عبارت است از روند کدگذاری داده‌ها به یک فرمت یا شکل قابل خواندن توسط برنامه‌ها و نرم‌افزارهای مختلف و درج داده‌ها در پایگاه اطلاعاتی GIS.

نکته: داده‌ها در GIS به دو نوع کلی تقسیم می‌شوند: ۱. داده‌های مکانی ۲. داده‌های توصیفی غیرمکانی

داده‌های توصیفی را از نظر ماهیت می‌توان به ۴ دسته تقسیم نمود:

- داده‌های اسمی (Nominal): در این نوع داده فقط می‌توان گفت یک مقدار با مقدار دیگر یکی است یا خیر، مانند جنس خاک.
- داده‌های ترتیبی (Ordinal): این مقادیر دارای ترتیب بوده ولی تفاضل آن‌ها دارای معنی نیست، مانند رتبه شرکت‌کنندگان در یک مسابقه
- داده‌های فاصله‌ای (Interval): این مقادیر دارای ترتیب بوده و تفاضل بین آن‌ها معنی‌دار می‌باشد ولی فاقد صفر مطلق هستند، مانند درجه حرارت برحسب درجه فارنهایت.
- داده‌های نسبی (Ratio): این مقادیر علاوه برداشتن ترتیب و معنی‌دار بودن تفاضل دارای صفر مطلق نیز می‌باشند، مانند طول.

داده‌های توصیفی دو نوع می‌باشند:

- پیوسته (Continuous): مانند طول خیابان، درجه حرارت، رطوبت، بارش و ...
- گسسته (Categorical): مانند کد کلاس کاربری، نام خیابان، ...

منابع داده‌های جغرافیایی

- نقشه‌های کاغذی: نقشه تصویری است از قسمتی یا تمامی کره زمین با یک مقیاس معین بر روی صفحه‌ای افقی، که عوارض و پدیده‌های مختلف به وسیله‌ی یک سری علائم قراردادی بر روی آن نشان داده شده است، انواع نقشه‌ها با توجه به محتوی آن در دو گروه نقشه‌های عمومی و نقشه‌های موضوعی قرار دارند.
- نقشه‌های عمومی: مهم‌ترین نقشه عمومی، نقشه‌های توپوگرافی هستند که اهمیت بالاتری دارند در واقع نقشه‌های توپوگرافی مادر تمام نقشه‌های مربوط به زمین‌اند. این نقشه‌ها به دلیل نمایش جزئیات زیاد، مقیاس بالایی دارند که مقیاس آن‌ها بین ۱:۲۵۰۰ تا ۱:۵۰۰۰۰۰ است.
- نقشه‌های موضوعی یا ویژه: برای استفاده در این زمینه تخصصی خاص تهیه می‌شوند. نقشه‌های ناوبری دریایی، نقشه‌های راه‌ها، نقشه‌های هواشناسی، نقشه‌های ژئومرفولوژی، نقشه‌های بهره‌برداری از زمین، نقشه‌های پراکندگی و ...

- نقشه‌های پراکندگی نحوه‌ی پراکندگی یک موضوع خاص را نمایش می‌دهند. مثل نقشه‌های کروکروماتیک، کروپلیت، ایزوپلیت و ...
- منابع سنجش از دور
- عکس‌های هوایی: اولین روش تولید داده از طریق سنجش از دور، عکسبرداری هوایی بوده است که توسط دوربین‌های عکسبرداری تعبیه شده در هلیکوپتر یا هواپیما تهیه می‌شدند. عکس‌های هوایی در انواع مختلف به صورت سیاه و سفید معمولی، سیاه و سفید مادون قرمز، رنگی حقیقی و رنگی مجازی موجود است. متداولترین عکس‌های هوایی در ابعاد 23×23 سانتی‌متر است که در ایران در مقیاس‌های ۱:۵۵۰۰۰، ۱:۲۰۰۰۰، ۱:۴۰۰۰۰ موجود است.
- تصاویر ماهواره‌ای: تصاویر ماهواره‌ای توسط سنجنده‌های الکترونیکی که در ماهواره‌ها قرار دارد، تهیه می‌شوند و در دو گروه از ماهواره‌های منابع زمینی و ماهواره‌های هواشناسی قرار دارند. اکثر این ماهواره‌ها مدار خورشید آهنگ (Sun - Synchronous) و نزدیک قطبی دارند.

روش‌های مختلف ثبت داده‌ها در یک سیستم اطلاعات جغرافیایی

۱. ثبت داده‌ها توسط صفحه کلید: اغلب داده‌های توصیفی توسط صفحه کلید ثبت می‌شوند. در برخی موارد این داده‌ها می‌توانند به شکل رقومی از یک پایگاه اطلاعاتی موجود که در آن داده‌ها توسط صفحه کلید به ثبت رسیده‌اند، بدست آیند. این روش نیاز به کار اپراتوری زیادی دارد، مهارت کمتری نیاز دارد، اساسا کم هزینه است.
۲. هندسه مختصات (COGO): این روش برای ثبت رکوردهای اطلاعاتی مربوط به زمین به کار می‌رود. دقت بسیار بالایی دارد. زمین را مانند نقشه به صورت دقیق همانگونه که در مدارک قانونی و کاداستری بیان شده نشان می‌دهد.
۳. تکنیک تبدیل نقشه‌های سنتی کاغذی یا سایر مدارک گرافیکی به داده‌های رقومی سازگار با کامپیوتر را رقومی‌سازی می‌نامند.
۴. رقومی کردن دستی: گسترده‌ترین روش ثبت داده‌های مکانی از روی نقشه‌ها می‌باشد. میزان کارایی این روش بستگی به کیفیت سخت‌افزار و نرم‌افزار به کار رفته و نیز مهارت اپراتور دارد. نیاز به محاسبات کمتری دارد. تجهیزات نسبتا ارزانتر سبب سهولت ساخت پایگاه داده در هر مکانی شده است. این روش وقت‌گیر است. کار مداوم اپراتور سبب پایین آمدن کیفیت کار می‌شود.
۵. رقومی‌سازی On Screen: ابتدا نقشه اسکن شده و سپس نقشه اسکن شده را رقومی‌سازی می‌کنیم. نسبت به حالت رقومی‌سازی دستی راحتتر، دقیق‌تر، سریعتر است. فرآیند بروزرسانی به سادگی انجام‌پذیر است.
۶. اسکن کردن: یک تصویر رقومی از نقشه اصلی با حرکت دادن یک نوار الکترونیکی در سطح نقشه تهیه می‌شود. داده‌ها در هنگام اسکن کردن به ساختار رستری در می‌آیند و در صورت لزوم باید به ساختار برداری تبدیل شوند.
۷. وارد کردن فایل‌های رقومی موجود: برای اشتراک‌گذاری داده‌های رقومی موجود باید یک استاندارد واحد تعریف شود.

نکته: برای سندهای ساده که نیاز به توضیحات کمی دارند از روش رقومی سازی نیمه اتوماتیک استفاده می شود. اما برای سندهای ساده یا مجزا که فقط دارای یک نوع اطلاعات هستند از روش رقومی سازی اتوماتیک استفاده می شود.

مراحل رقومی سازی دستی عبارتند از: ثبت، رقومی سازی عوارض نقطه‌ای، رقومی سازی عوارض خطی، رقومی سازی عوارض سطحی یا چندضلعی، افزودن اطلاعات توصیفی

نکته: اولین و مهم‌ترین مرحله در رقومی سازی عبارت است از تبدیل اطلاعات نقشه از حالت خطی و غیر رقومی به شکل رقومی و سازگار با کامپیوتر.

مزیای داده‌های رقومی بر داده‌های دستی و آنالوگ

- بروزرسانی کردن آن آسان تر است.
- نگهداری راحت تر
- حمل و نقل آسان تر
- کپی پشتیبانی Back up
- تکثیر و انتشار راحت تر
- طبقه بندی اطلاعات راحت تر است.
- سهولت ذخیره سازی بدون تخریب فیزیکی
- سهولت تغییر و تنظیم اندازه و مقیاس نمایش
- ارتباط بین اطلاعات راحت تر است.
- جست و جوی راحت تر
- دقت بالاتر
- مدل سازی راحت تر

نکته: زمان، مکان و ویژگی‌ها سه مشخصه مهم مورد بررسی در مدل‌های مجازی داده‌های جهان واقعی محسوب می شوند و با ارتباط منطقی داده‌های زمان و مکان و ویژگی‌ها، اطلاعات مجازی بسیار مفیدی در رابطه با پدیده‌های زمینی قابل حصول است.

انواع دستگاه‌های رقومی ساز

- میزهای دیجیتالیزر که شامل یک صفحه و یک Cursor دستی هستند.
- اسکنرها با آرایه‌های خطی می باشند جزء اصلی Scanner یک Sensor است که بصورت منظم تمام سطح تصویری را نقطه به نقطه و در راستای خطوط موازی Scan می کند.
- دسته دیگر دستگاه‌هایی می باشند که از روی صفحه مانیتور عوارض را رقومی می کنند.

مزایا و معایب رقوم سازی

۱. نیاز به تجهیزات گران قیمت نیست.
۲. رایج ترین روش برای کدگذاری عوارض مکانی، رقوم سازی دستی است.
۳. نیاز به آموزش کمی دارد.
۴. نیازی به نقشه های با کیفیت بالای نیست.
۵. نیاز به کار اپراتوری زیادی دارد.
۶. وقت گیر است.
۷. رقوم سازی دستی زمانی کم هزینه تر است که تعداد نقشه هایی که باید تبدیل شوند کم باشد. نقشه ها دارای شرایط مناسب و خوبی نباشند. نقشه ها دارای اطلاعات زیادی بوده و نیاز به تفسیر داشته باشند.
۸. برای شروع عملیات رقوم سازی، ابتدا باید حداقل ۴ نقطه مرجع مربوط به نقشه را ثبت کرد.

نکته: آخرین مرحله از یک عملیات رقوم سازی وارد کردن اطلاعات توصیفی عوارض می باشد. اطلاعات توصیفی در پایگاه داده ذخیره می شوند. این اطلاعات توصیفی می توانند قبل ما بعد و یا حتی همزمان با دیجیتالیز کردن عوارض در پایگاه داده قرار گیرند. دو روش برای برقراری اطلاعات توصیفی و داده های مکانی وجود دارد:

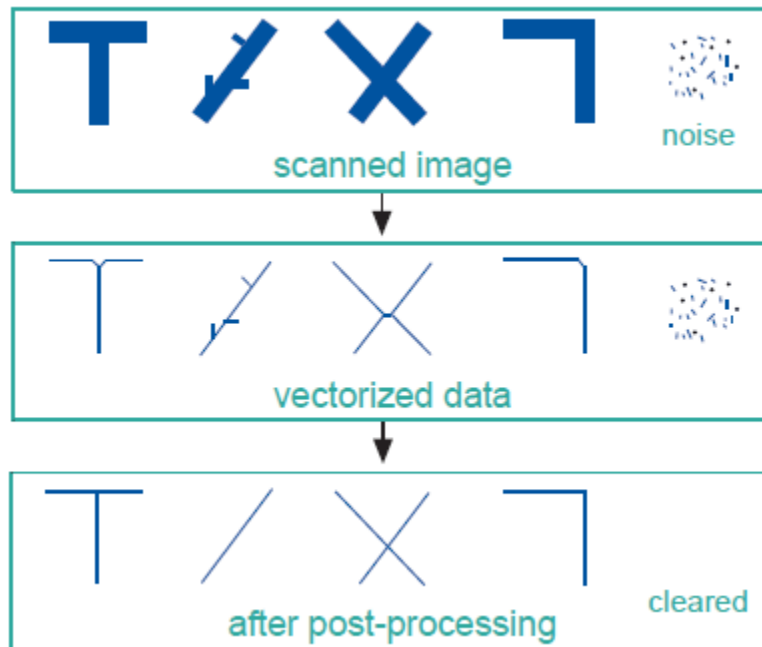
۱. در روش Logical یا منطقی هم اطلاعات توصیفی و هم داده های مکانی متناظر باید با استفاده از یک شناسه منحصر بفرد مشخص گردد. این شناسه منحصر بفرد برای هر عارضه باید در مرحله رقوم کردن و ورود اطلاعات توصیفی معرفی گردد.
۲. در روش Positional مجموعه اطلاعات توصیفی متعلق به یک عارضه به همراه یک نقطه در کنار عارضه ای که باید به آن متصل شود قرار می گیرد.

انواع خطاهای موقعیتی در رقوم سازی عبارتند از: اشتباه، حذف عوارض، عدم شناسایی درست نقاط کنترل، عدم بسته بودن پلیگون ها، اتصال غلط نقاط، به هم نرسیدگی، از هم ردشدگی و

مزایا و معایب اسکن کردن

۱. به راحتی انجام می شود.
۲. روشی سریع برای تولید اطلاعات رقوم است.
۳. فایل های رستری حجیم هستند و حجم زیادی از حافظه را اشغال می کنند.
۴. داده های توصیفی مستقیما نمی توانند به عوارض خطی یا پلی گونها متصل شوند.
۵. مسئله Line Scaling به راحتی حل نمی شود.
۶. نقشه های عاری از خطا و با خطوط واضح مورد نیاز است.
۷. تجهیزات و نرم افزار در این روش معمولا گران قیمت تر هستند.

۸. نیاز به کارمندان مجرب دارد.
۹. اسکن کردن زمانی کم هزینه‌تر است که نقشه‌ها تمیز و ساده بوده و اطلاعات کمی داشته باشند. نقشه‌ها تعداد زیادی پلی‌گون داشته باشند. نقشه‌ها دارای اشکال هندسی نامنظم متعدد باشند.



نکته: اسکنرها از سخت‌افزارهای ورودی داده در GIS هستند. دو دسته اسکنر وجود دارد: اسکنرهای تخت و اسکنرهای استوانه‌ای.

نکته: توان تفکیک اسکنر به نقطه در اینچ بیان می‌شود. انواعی از اسکنرها توانی تا ۶۰۰۰ نقطه در اینچ دارند. تفکیک ۳۰۰۰ رایج‌تر است. هر چه اندازه قدرت تفکیک کوچک‌تر انتخاب شود، جزئیات بیشتری از نقشه برداشت خواهد شد و برعکس با افزایش توان تفکیک حجم پرونده‌ی حاصل از عمل اسکنر بیشتر خواهد بود.

GIS Ready کردن داده‌های مکانی

دو نوع داده مکانی جهت ورود به محیط GIS وجود دارد:

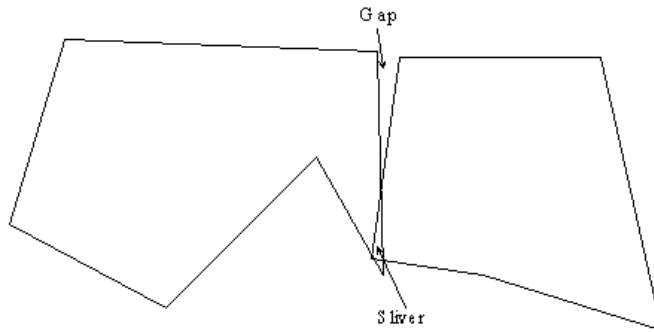
- داده‌های رستری (مانند تصاویر ماهواره‌ای یا عکس‌های هوایی یا تصاویر مربوط به اماکن موجود)
- داده‌های برداری (انواع نقشه)

خطاهایی که جهت ورود اطلاعات برداری به محیط GIS باید تصحیح شوند. این خطاها جزو خطاهای ظاهری موجود در نقشه می‌باشند.

۱. Sliver – Gap: یکی از مکان‌های ایجاد خطا در مرز مشترک بین دو پلیگون که در تشکیل هر دو پلیگون سهم دارد، می‌باشند. خطاهای موجود در مرز مشترک پلیگون‌ها که بر اثر رقومی‌سازی دوباره مرز بین پلیگون‌ها پدید می‌آیند شامل Sliver و Gap می‌باشد.

علت ایجاد این خطاها عدم تطابق دقیق دو خطی است که در مرحله ترسیم مرز مشترک بین پلیگون‌ها بوجود می‌آید. برای اجتناب از این خطا بهتر است که مرز مشترک دو سطح مجاور تنها یکبار ترسیم شود و بعد در مرحله تعریف توپولوژی، همسایگی برای دو سطح مجاور تعریف گردد.

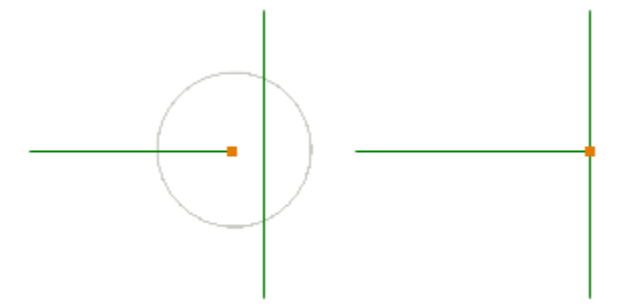
پلیگون‌هایی با مساحت خیلی کوچک غالباً مربوط به خطا در رقومی‌سازی خطوط است و می‌تواند به صورت اتوماتیک پاک شود.



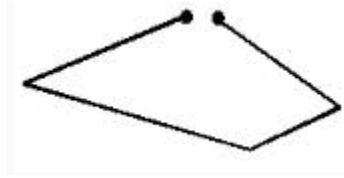
۲. Over-Shooting: خطای از هم رد شدن خطوط و عدم وجود Node در محل اتصال آن‌ها. برای مثال دیوار به عارضه خطی می‌رسد ولی در نقشه می‌بینیم که دیوار تا کمی بعد از محل برخوردش با عارضه موردنظر ادامه پیدا کرده است.



۳. Under-Shooting: خطای بهم نرسیدگی خطوط. در این حالت دیوار قبل این که به عارضه بعدی که با آن اتصال دارد، ناگهان در نزدیکی آن تمام شده است.



۴. بسته نبودن پلیگون‌ها: در این حالت سطوحی که می‌بایست بسته باشند، حداقل در یک محل بسته نیستند و دو خط به هم نرسیده‌اند.

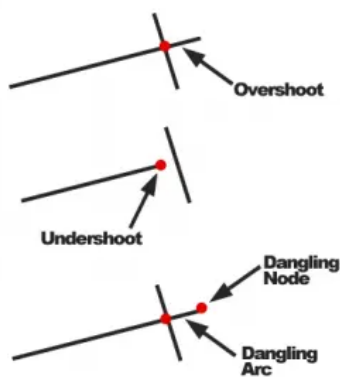


سه خطای عمود نبودن خطوطی که بر روی زمین عمود بوده‌اند، موازی نبودن خطوطی که بر روی زمین موازی بوده‌اند، غیر هموار بودن خطوط مستقیم معمولاً بدلیل عدم دقت اپراتور برداشت کننده یا ترسیم کننده رخ می‌دهند.

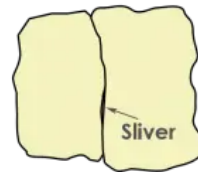
نکته: خطاهای از هم رد شدن و بهم نرسیدگی خطوط و همچنین بسته نبودن پلیگون‌ها معمولاً ناشی از عدم Snap کردن اپراتور در هنگام ترسیم عارضه رخ می‌دهند که یکی از دلایل این امر سرعت بخشیدن به کار می‌باشد که متأسفانه سبب می‌شود که وقت زیادی در مرحله ویرایش جهت یافتن و اصلاح این خطاها هزینه شود.

سایر خطاهای موجود در نقشه‌ها هنگام GIS Ready کردن داده‌های مکانی

Dangles, overshoots and undershoots are examples of errors.













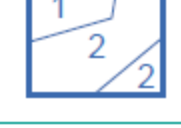





SLIVERS are small, spurious gap often as a result of imprecise digitization of features in GIS.

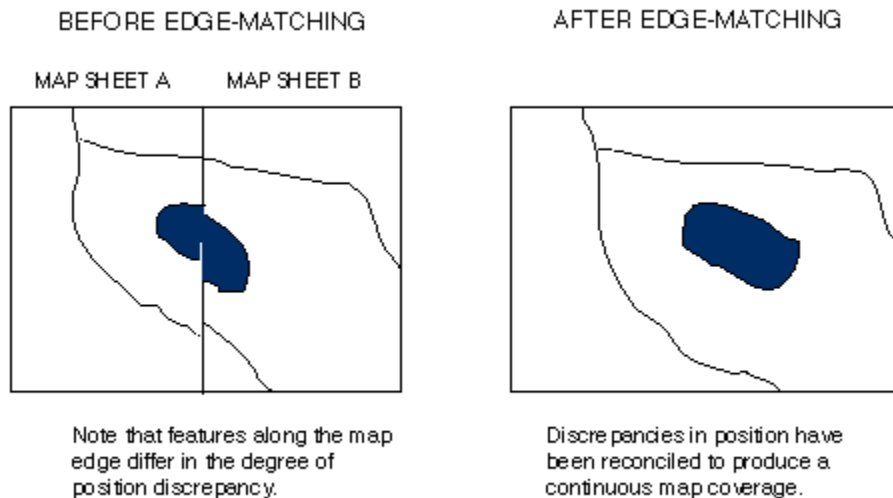


OVERLAPS are polygons or lines that are on top of each other which could result in double-counting areas.



Before cleanup	After cleanup	Description	Before cleanup	After cleanup	Description
		Erase duplicates or sliver lines			Extend undershoots
		Erase short objects			Snap clustered nodes
		Break crossing objects			Erase dangling objects or overshoots
		Dissolve polygons			Dissolve nodes into vertices

در انتهای مرحله ویرایش و اصلاح نوبت به ادغام و ورود اطلاعات به پایگاه داده مکانی می‌رسد. در این مرحله می‌بایست انطباق لبه‌های برگ نقشه‌های مجاور در ادغام اطلاعات مورد لحاظ گردد. عمل انطباق لبه‌ها (Edge matching) می‌تواند بصورت دستی یا خودکار انجام شود. عوارض خطی که در دو یا چند نقشه وجود دارند در زمان انطباق لبه‌های نقشه به یکدیگر می‌بایست تبدیل به یک عارضه منفرد شده و یکپارچه شوند، همچنین عوارض سطحی و پلیگون‌هایی که در مرز مشترک برگه‌ها قرار دارند باید ادغام و تشکیل یک پلیگون واحد را دهند.



نکته: به منظور اشتراک‌گذاری داده‌ها، تبادل اطلاعات و جلوگیری از تهیه مجدد آن‌ها از Clearinghouses استفاده می‌شود. قبل از هر کاری متادیتا نقشه‌ها را مطالعه می‌کنیم و جست‌وجو بر روی متادیتا خواهد بود. یک Clearinghouse یک ساختار فیزیکی است

که برای اندوختن و توزیع نمودن داده‌های رقومی و اطلاعات مربوط به این داده‌ها بکار می‌رود. متادیتا نقش مهمی را در Clearinghouses ایفا می‌کند.

اهداف Clearinghouses عبارتند از:

۱. به حداقل رساندن تلاش‌های تکراری برای جمع‌آوری داده

۲. استفاده از منافع به اشتراک‌گذاری داده‌ها

نکته: Clearinghouses به تهیه‌کنندگان داده اجازه ثبت داده‌های جغرافیایی و دسترسی به آن‌ها را می‌دهد. نقش Clearinghouses به صورت یک کاتالوگ به همراه پشتیبانی از لینک‌های داده‌های جغرافیایی و امکانات جست‌وجو می‌باشد.

کیفیت داده‌ها و متادیتا

به تمام ویژگی‌های یک محصول داده‌ای که نشان‌دهنده میزان تطابق آن با نیازمندیهای مورد نظر می‌باشد، کیفیت داده گویند. کیفیت به معنای مناسب بودن برای استفاده بیان می‌شود.

تعریف درستی و دقت

- درستی (Accuracy): عبارت است از میزان نزدیکی نتایج محاسبات و یا پیش‌بینی‌ها به مقدار واقعی کمیت مورد نظر، با میزان توافق بین یک آزمون و مقدار مرجع مورد قبول.
- دقت (Precision): عبارت است از میزان نزدیکی کمیت‌های مورد اندازه‌گیری نسبت به میانگین کمیت مورد نظر، (لزوماً دقت بالا به مفهوم درستی بالا نمی‌باشد).

مولفه‌های تعیین کیفیت داده‌های مکانی

- درستی موقعیت (Positional Accuracy)
- درستی داده‌های توصیفی (Attribute Accuracy)
- سازگاری منطقی (Logical Consistency)
- کامل بودن (Completeness)
- تاریخچه (Lineage)
- قدرت تفکیک (Resolution)

عوامل مهم در تعیین کیفیت داده‌ها:

۱. مولفه‌های ریز مقیاس (Micro Level Components): این مولفه‌ها مربوط به هر یک از المان‌های منفرد داده‌ها می‌باشند و معمولاً به وسیله تست آماری داده‌ها در مقابل منابع مستقل با کیفیت بالایی از اطلاعات ارزیابی می‌شوند.

- صحت موقعیت (Positional Accuracy): عبارت است از خطای موقعیت یک نقطه و یا انحراف موقعیت جغرافیایی یک نقطه بر روی نقشه نسبت به موقعیت واقعی آن بر روی زمین. صحت موقعیت معمولاً با دو جزء مشخص می‌شود که عبارتند از: Precision و Bias. بایاس عبارت است از اختلاف سیستماتیک بین موقعیت نشان داده شده و موقعیت واقعی یک نقطه و به طور ایده‌آل باید صفر باشد. بایاس معمولاً بوسیله تعیین خطای متوسط موقعیت یک سری نقاط نمونه اندازه‌گیری می‌شود.
 - صحت داده‌های توصیفی (Attribute Accuracy): مشخصات داده‌ها می‌توانند متغیرهای گسسته یا پیوسته باشند. یک متغیر گسسته فقط می‌تواند مقادیر مشخص و محدودی را داشته باشد. در حالی که یک متغیر پیوسته می‌تواند هر مقداری را بپذیرد.
 - همگونی منطقی (Logical Consistency): همگونی منطقی عبارت است از چگونگی حفظ روابط بین اجزاء داده‌ها، مثلاً مرز جنگل از خطوط وسط جاده نباید بگذرد و این نقص منطقی به حساب می‌آید.
 - قدرت تفکیک (Resolution): قدرت تفکیک یک مجموعه از داده‌ها عبارت است از کوچکترین واحد قابل تشخیص نمایش داده شده در آن مجموعه. در مورد عکس‌های هوایی و تصاویر ماهواره‌ای این قدرت تفکیک عبارت است از کوچکترین شی که می‌تواند تشخیص داده شود که این مطلب را قدرت تفکیک فضایی می‌نامند.
۲. مولفه‌های بزرگ مقیاس (Macro Level Components): مولفه‌هایی هستند که به مجموعه داده‌ها به عنوان یک کل نگاه کرده و آن را تحت یک مجموعه واحد بررسی می‌کنند.
- کامل بودن (Completeness): کامل بودن لایه، کامل بودن طبقه‌بندی، کامل بودن طبقه‌بندی بررسی و تحقیق.
 - زمان (Time): عبارت است از تاریخ تهیه داده‌های اولیه موجود در یک لایه. زمان فاکتور بسیار مهمی در تعیین کیفیت داده‌های جغرافیایی می‌باشد، زیرا بسیاری از داده‌های جغرافیایی در طول زمان سریعاً در حال تغییر هستند.
 - تاریخچه داده‌ها (Lineage): منظور از Lineage در یک مجموعه داده‌ها عبارت است از تاریخچه، سرچشمه و مراحل پردازش بکار گرفته شده در ایجاد مجموعه داده‌ها.
۳. مولفه‌های کاربردی (Usage Level Components): این مولفه‌ها که در کیفیت داده‌ها موثر می‌باشند به منابع سازمان برمی‌گردند. مثلاً تاثیر هزینه داده‌ها بستگی به منابع مالی یک سازمان دارد. یک مجموعه مشخصی از داده‌ها ممکن است برای یک سازمان پر هزینه باشد، در حالیکه برای سازمان دیگر هزینه کمتری را در بر می‌گیرد. قابلیت در دسترس بودن داده‌ها که عبارت است از میزان سهولت به دست آوردن و استفاده از داده‌ها یکی دیگر از مولفه‌های کاربردی محسوب می‌شود.

منابع خطاها معمول در استفاده از GIS

- جمع‌آوری داده: خطا در زمینه جمع‌آوری داده، خطای نقشه‌های موجود، خطا در آنالیز داده‌ها
- وارد کردن داده‌ها: بی‌دقتی در دیجیتایز کردن، بی‌دقتی ناشی از عوارض جغرافیایی
- ذخیره‌سازی داده‌ها: دقت عددی نامناسب، دقت فضایی نامناسب

- تغییر و تحول داده‌ها: کافی نبودن فاصله کلاس‌ها، خطای ناشی از همپوشانی، خطای Sliver که در مرحله همپوشانی پلی‌گون‌ها بوجود آمده است.
- خروجی داده‌ها: دقت نداشتن در معرفی مقیاس، خطای ناشی از دقت نداشتن دستگاه مورد استفاده برای خروجی خطای ناشی از پایدار نبودن (medium).
- خطا در استفاده از نتایج: فهم غلط اطلاعات، استفاده نامناسب از اطلاعات

نکته: هدف از بررسی خطاها، حذف این خطاها نبوده بلکه چگونگی اداره کردن آن‌ها می‌باشد.

نکته: همواره برای اشتراک‌گذاری داده‌ها باید متادیتا داشته باشیم. $Meta\ Data = Data\ about\ data$ متادیتا حاوی اطلاعاتی همچون: چه داده‌هایی موجود است؟، کیفیت داده‌ها، فرمت داده‌ها، ترتیب جزئیات و قیمت، محدودیت‌ها در حق چاپ.

انواع مدل‌های داده‌های مکانی

دو روش کلی برای نمایش اجزاء فضای اطلاعات جغرافیایی وجود دارد:

- مدل برداری: در این مدل اشیاء با موقعیت‌ها در جهان واقعی به وسیله مجموعه نقاط و خطوطی که مرزهای آنرا تعیین می‌کنند نمایش داده می‌شوند. مدل برداری دقت بیشتری در موقعیت عوارض در فضا را امکان‌پذیر می‌سازد. در حقیقت سطح دقت به تعداد بیت‌های بکار برده شده برای یک مقدار در کامپیوتر محدود می‌گردد. مدل داده برداری دو ساختار اسپاگتی (ساختار نامنظم) و توپولوژی (ساختار منظم) است. در این مدل بازیافت، بهنگام‌سازی و تعمیر کلی بهتر انجام می‌شود و خروجی‌های دقیق‌تری ارائه می‌کند، اما ساختار پیچیده‌تری نسبت به رستر دارد، فناوری آن گران است بخصوص زمانی که دقت بالای سخت‌افزاری و نرم‌افزاری ضروری است. مدل برداری بعلاوه اینکه می‌تواند هر عارضه را بصورت مستقل و جداگانه داشته باشیم، Object Based Model گویند. این مدل یک مدل مجزا است که هر Object را بصورت مستقل و مجزا داریم.
- دلائل مهم استفاده از مدل برداری در GIS دقیق بودن ارائه و نمایش داده‌ها، کارایی از نظر ذخیره داده‌ها، کیفیت کارتوگرافیک خروجی و امکان دسترسی با ابزارها و روش‌های عملیاتی مانند همپوشی، تحلیل و سیستم‌های تصویری نقشه‌هاست.
- مدل رستری: در این مدل فضا به طور منظم به سلول‌هایی (که معمولا دارای شکل مربعی هستند)، تقسیم می‌گردد. موقعیت اشیاء یا شرایط جغرافیایی به وسیله سطر و ستون سلول‌هایی که اشغال می‌کنند، تعیین می‌گردد. کوچک‌ترین عنصر تشکیل‌دهنده رستر، پیکسل یا سلول نامیده می‌شود که ارزش هر یک از این سلول‌ها، نمایانگر اطلاعات طیفی یا توصیفی عارضه‌ی زمینی است. بین سلول‌های یک داده رستری هیچ رابطه‌ی منطقی وجود ندارد. داده‌های حاصل از اسکن کردن و تصاویر ماهواره‌ای دارای ساختار رستری‌اند. مشکل عمده در این مدل این است که هر سلول همگن و مستقل از سایر سلول‌هاست و دارای یک ارزش