

# پرهیز از کاربرد نامناسب داده‌های مکانی

دو روش هشدار به کاربران

مؤلفین: Dr Richard T. Pascoe, Gerg M. Byrom

مترجم: اشرف عظیم زاده ایرانی

کارشناس ارشد جغرافیا

با فسن آوری کنونی، پردازش ترکیبی مجموعه داده‌های (datasets)<sup>(۱)</sup> مکانی از منابع متعدد امکان‌پذیر شده است. استفاده نابجا از مجموعه داده‌ها، غالباً کار پردازش آنها را پرهزینه، غیر دقیق و زمانبر می‌نماید. چگونه می‌توان چنین مشکلاتی را به حداقل رساند؟ نویسندگان این مقاله، نتیجه تحقیقات خود را درباره ساخت یک سیستم هشدار دهنده بر مبنای دو نوع راه حل، ارائه نمودند. اولین راه حل، محدود کردن استفاده از داده‌ها، تنها با یک عملیات مشخص می‌باشد. راه حل بعدی، لزوم تجزیه و تحلیل متادیتاهای (metadata)<sup>(۲)</sup> مربوط به مجموعه داده‌هاست تا از این طریق قابلیت کاربری مجموعه داده‌ها در موارد مختلف تعیین شود.

نگریده بودند. در بیشتر مواقع، یک کاربرد خاص که مورد نظر کاربر است، خصوصیتی از قبیل مقیاس و دقت مجموعه داده‌ها را تعیین می‌کند. این امکان نیز وجود دارد که خصوصیات فوق، داده‌ها را برای سایر کاربردها نامناسب نمایند. علم و آگاهی قبلی بر قابلیت‌ها و تواناییهای مجموعه داده‌ها، استفاده مناسب و به جا از آنها را ممکن می‌سازد. هر چند غالب افرادی که این علم را دارند، کسانی نیستند که کار پردازش را انجام می‌دهند. هر چه کاربر از قابلیت‌های مجموعه داده‌ها برای استفاده مناسب از

ارتباط پذیری بیشتری که به واسطه شبکه‌های کامپیوتری به وجود آمده، ما را به سمت و سوی کاربرد ترکیبی داده‌های استخراج شده از منابع گوناگون هدایت می‌کند. لیکن کاربران غالباً از منبع و سرچشمه مجموعه داده‌ها بی‌اطلاعند. و به همین دلیل به طور غیر عمد از داده‌ها استفاده نابجا می‌نمایند، که این منجر به حصول نتایج نادرست نیز می‌شود. علت بروز چنین مشکلی این است که معمولاً از داده‌هایی استفاده می‌شود که در اصل به آن منظوری که به کار برده می‌شوند، اخذ

تا سودمند واقع گردند.  
 کاربرد نامناسب از جنبه‌های اقتصادی، اجتماعی و یا قانونی می‌تواند تعریف شود. استانداردها یا قراردادهای ملی و بین‌المللی هم در این تعریف می‌گنجانند.

آنها مطلع‌تر باشد، فرصت بهتری خواهد داشت تا:  
 \* مجموعه داده‌ای را که امکان دارد برای کاربرد مورد نظرش مناسب‌تر باشد، جستجو نماید.  
 \* با درجه‌ای از اطمینان که بر پایه قابلیت این مجموعه داده‌ها در کاربرد مورد نظر می‌باشد، نتایج را تفسیر می‌نماید.



روش شی گرا (object-oriented) و روش متادیتا (metadata) در قسمت ذیل مورد بررسی قرار می‌گیرند.

### روش شی گرا و روش متادیتا

در روش شی گرا، کاربران به طور غیر مستقیم به واسطه عملیات تعریف شده، به داده‌ها دسترسی خواهند داشت. این عملیات مانند پوششی محافظ برای مجموعه داده‌ها عمل می‌کنند تا از به کار بردن نادرست آنها جلوگیری شود. به عنوان مثال، داده‌های مربوط به شبکه خیابانهای یک ناحیه، به نحوی تهیه شده که فقط بر روی مجموعه داده‌هایی با مقیاس کوچکتر از ۱:۴۰۰۰۰ قرار می‌گیرند. در نتیجه، مجموعه داده‌های شبکه خیابانها بر روی مجموعه داده‌های کاداستری در مقیاس ۱:۲۰۰۰۰ نمی‌تواند قرار گیرد. چون چنین عملیاتی برای آنها تعریف نشده است.

هنگامیکه یک توصیف (metadata) از مجموعه داده‌ها با یک کاربرد خاص مطابقت داده شود، بدین طریق، مشخص می‌گردد که آیا آن کاربرد برای مجموعه داده‌ها مناسب هست یا خیر؟

بنابراین سیستمی که کاربران داده‌ها را از نتایج کاربرد مجموعه داده‌ها آگاه نماید، می‌تواند کارایی مجموعه داده‌ها را افزایش داده و میزان خطا را در تجربه و تحلیل به حداقل برساند.

### استفاده مناسب از داده‌ها

علم و آگاهی قبلی از اینکه کاربرد مناسب داده‌ها چه مواردی می‌باشند، ضروری و لازم است. این علم و آگاهی برای عملیاتی که در مجموعه داده‌ها به کار برده می‌شوند (روش شی گرا) (۳) یا برای متادیتاها (روش متادیتا) (۴)، یک تعریف سازگار و منطقی، با اصول قواعدی نظیر قواعد زیر فراهم می‌نماید:

- \* مجموعه داده‌ها نباید از نظر مقیاس تفاوت‌های عمده داشته باشند.
- \* مجموعه داده‌ها باید در نواحی مورد نیاز دقت لازم را داشته باشند.
- \* مجموعه داده‌ها باید به اندازه کافی گویا باشند.
- \* مجموعه داده‌ها نباید برای کاربرد مورد نظر، قدیمی و کهنه شده باشند.
- این قواعد نیاز به تعریف جامع‌تر یا شاید نیاز به کاربردی صحیح دارند

می‌نمود.

گسترش استانداردهای متادیتاهای موجود، موجب گنجاندن مشخصه‌هایی در مجموعه داده‌ها می‌شود که با عملیات ویژه‌ای، شرایط را به توصیف درآورند. این عملکرد، گام مهمی در جهت تکمیل توصیف مجموعه داده‌ها می‌باشد.

به علاوه، گسترش استانداردهای متادیتاهای متادیتاهای تعیین می‌کند که برای مجموعه اطلاعاتی که به طور موقت در مدت پردازش داده‌ها ایجاد شده‌اند، یک سری از عملیات بهتر است یا متادیتا.

### نتیجه‌گیری

نویسندگان این مقاله، روش متادیتا را توصیه می‌نمایند. زیرا این روش به کاربر امکان می‌دهد هنگامیکه در دسترسی به داده‌ها محدودیتی نداشته باشد، درباره دلیل نامناسب بودن یک مجموعه از داده‌های خاص، به اطلاعات بیشتری دست یابد.

### پاورقی

۱- مجموعه‌ای از اطلاعات به هم مرتبط (پیوسته) که از اجرای جداگانه‌ای ساخته شده، اما در پردازش داده‌ها به عنوان یک واحد تلقی می‌شوند.

۲- متادیتا یا فوق داده، به اطلاعاتی درباره داده‌ها گفته می‌شود. به عنوان مثال: عنوان، موضوع، دقت، مقیاس و... متادیتاهای مربوط به داده‌های یک فایل نقشه‌ای می‌باشند.

3) object - oriented approach

4) metadata approach

5) Australia and New Zealand Land Information Council

6) Content Standard for Digital Geospatial Metadata

با رجوع به مثال قبلی، متادیتاهایی که دلالت بر مقیاس می‌کنند، برای مجموعه داده‌های کاداستری، ۱:۲۰۰۰۰ و برای داده‌های شبکه خیابانها، ۱:۴۰۰۰۰ می‌باشند. ترکیب این موضوع با آگاهی قبلی از اینکه مجموعه داده‌هایی با اختلاف مقیاس بیش از ۱:۱۵۰۰۰ نامناسب هستند و آنها را نمی‌توان روی هم قرار داد، این هشدار را برای کاربر به دنبال خواهد داشت که پوشش دادن دو نقشه با داده‌های ذکر شده، مناسب نمی‌باشد، هر چند، استانداردهایی برای متادیتاها وجود دارد. به عنوان مثال: انجمن اطلاعات استرالیا و نیوزیلند (ANZLIC)<sup>(۵)</sup>، به جای اینکه مشخصاتی نظیر دقت گرافیکی محدوده اطلاعات، توضیح درباره کپی رایت، مقیاس و موقعیت را شرح دهد، استاندارد برای متادیتاهای گرافیکی رومی (CSDGM)<sup>(۶)</sup> و نظایر آن تعریف کرده است که توسط آنها می‌توان کاربرد مناسب را تعیین نمود.

### یک نمونه کاربردی

برای به دست آوردن اطلاعاتی درباره هر دو نوع روش، به یک نمونه کاربردی توجه کنید.

کاربر، گروهی از مجموعه داده‌ها را انتخاب می‌کند. سپس یک سری عملیات را برای داده‌ها برمی‌گزیند. در روش شی گراف فقط عملیاتی که برای مجموعه داده‌های تعریف شده مشخص گردیده، توسط کاربر برگزیده می‌شود. در مورد روش متادیتا، یک سری عملیات برای همه مجموعه داده‌ها آماده شده است. این سری عملیات که توسط کاربر تعریف می‌شود، مورد تجزیه و تحلیل سیستم قرار گرفته و گزارشی مبنی بر اینکه چرا عملیات مربوط به برخی داده‌ها نامناسب هستند، ارائه می‌گردد.

### مقایسه روشها

در مورد روش شی گراف، اکثر مجموعه داده‌های موجود نه عملیات تعریف شده مناسب دارند، نه مکانیزی که بتواند این عملیات را تعریف کند. چنانچه در یک کاربرد خاص، استفاده از مجموعه داده‌ها نامناسب تشخیص داده شود، هم عملیات تعریف شده مناسب و هم مکانیزی که قادر به تعریف عملیات باشد، بایستی فراهم شوند.

استفاده از داده‌ها محدود به قابلیت کاربرد آنها به واسطه عملیاتی است که از قبل پیش بینی شده است، بدون اینکه توضیحی داده شود که چرا عملیات دیگری پیش‌بینی نشده و نباید مورد استفاده قرار گیرد. در این روش، وجود متادیتا اضافه است و هیچ کاربردی ندارد. در حالیکه اکثر متادیتاها، طبق استانداردهای گوناگون، برای هر دو نوع مجموعه اطلاعات تهیه شده‌اند از این رو، مشکل نمایش‌های مختلف از اطلاعات یکسان، پدیدار می‌گردد.

### مسائل حل نشده

خوب بود که هر یک از استانداردهای متادیتا، علاوه بر نشان دادن مقادیر مجموعه داده‌ها، مشخصه‌هایی از نحوه کاربرد اطلاعات را نیز ارائه