

Gewinnung, Analyse und Visualisierung von geographischen Metadaten anhand des IGLIS-Datenbankmodells

Martina DÜRAUER, Johannes FORSTHUBER,
Roland GRILLMAYER und Christoph PERGER

Zusammenfassung

Metadaten liefern wesentliche Zusatzinformationen zu geographischen Ressourcen. Erst durch die Dokumentation dieser Ressourcen in Form von Metadaten wird eine Suche und Bereitstellung dieser, für bestimmte Anwendungen ermöglicht. Des Weiteren leisten Metadaten einen wesentlichen Beitrag zur Wertsicherung von geographischen Ressourcen. Obwohl der Nutzen von Metadaten systemen unumstritten ist, wird diesem Themenkomplex zum Teil zu wenig Aufmerksamkeit in der praktischen Umsetzung gewidmet. Das lässt sich einerseits auf die sehr komplexe Materie der Metadatenstandards und andererseits auf die oft fehlende bzw. nicht praxistaugliche Integration von Metadateninformationssystemen für die Ersterhebung, Wartung und Suche von geographischen Ressourcen in geographischen Informationssystemen zurückführen. Ein weiterer Schwachpunkt der zu einem geringen Nutzerinteresse in diesem Themenkomplex beiträgt ist die Tatsache, dass die Informationen die in den Metadaten abgebildet werden oft unzureichende bzw. nicht benutzerfreundlich aufbereitet und visualisiert werden. Diesen unterschiedlichen Schwachpunkten widmet sich dieser Beitrag und stellt verschiedene Funktionsmuster und Konzepte sowie deren praktische Realisierung und Integration in das Geoinformationssystem IGLIS der Firma Forsthuber GmbH dar. Die entwickelten Konzepte und Funktionsmuster wurden in weiterer Folge von IGLIS Nutzern hinsichtlich ihrer Praxistauglichkeit evaluiert und als praxistauglich bewertet.

Das grundlegende Visualisierungskonzept basiert auf regelmäßigen Rasterzellen die über die verschiedenen Datenschichten gelegt werden und die Bezugsgeometrie der Visualisierung darstellen. Für jede Bezugsgeometrie (Rasterzellen) wurden die für die Visualisierung relevanten Metadateninformationen automatisiert abgeleitet. Die Informationen und Ausprägungen der Metadatenelemente werden durch Transparenz und Farbe der Rasterzellen dargestellt. Die Rasterzellgröße ist variabel einstellbar. Die so erhaltene Visualisierung stellt für den Benutzer eine erste Interpretationsgrundlage für eine möglichst rasche und einfache Entscheidungsfindung dar.

1 Metadaten und Visualisierung

Geodaten sind eine Abstraktion der realen Welt. Es gibt keine perfekten, vollständigen und richtigen Daten. Der Grad der Näherung, der Vereinfachung und des Weglassens ist abhängig davon, wer die Daten für welche Aufgabe erfasst hat. Damit die Daten nicht zweckentfremdet verwendet werden und so zu falschen Schlussfolgerungen führen, werden die Annahmen und Einschränkungen, unter denen sie entstanden sind, dokumentiert. Unter dem Begriff Metadaten wird das Ergebnis dieser Dokumentation verstanden. Vereinfacht gesagt sind Metadaten Daten über die Daten (BARTELME 2005).

Auf Grund der hohen Verfügbarkeit von Geodaten in immer kürzeren Zeitintervallen und unterschiedlichen Qualitäten gewinnen Metadaten und die auf diesen aufbauenden Metadateninformationssysteme immer mehr an Bedeutung.

Innerhalb eines Unternehmens minimieren somit Metadaten den Kosten- und Zeitaufwand bei vielen Aufgaben und verhindern Informationsverluste durch Personalfluktuations. (WASSILIOS 2005) Mitarbeiter können die gewünschten Datensätze schneller lokalisieren und die Qualität des Datensatzes analysieren und bewerten. Somit ermöglichen Metadaten eine Abschätzung der Eignung von Datensätzen für bestimmte Fragestellungen und Anwendungsgebiete. Des Weiteren ermöglichen Metadaten eine Dokumentation von Datenbeständen und leisten somit einen wesentlichen Beitrag zur langfristigen Wertsicherung dieser (LANDESVERMESSUNG U. GEOBASISINFORMATION BRANDENBURG, SENATSVERWALTUNG F. STADTENTWICKLUNG 2009). Da die Erfassung von Metadaten eine Zeit- und somit auch kostenaufwändige Aufgabe darstellt, stellt die Gewinnung bzw. direkte Ableitung von Metadaten aus den tatsächlichen Datenbeständen eine wesentliche Herausforderung bei der Entwicklung und Integration von Metadateninformationssystemen in GIS-Applikationen dar.

Des Weiteren besteht spezielles Interesse an mehrdimensionalen Visualisierungen von Metadatenelementen, wie zum Beispiel der Dimensionen Raum und Zeit und Qualitätskriterien wie Lagegenauigkeit in einer Abbildung. Derartige Visualisierungen sollten dem Benutzer von GIS-Systemen eine rasche und intuitive Interpretation der Dateneigenschaften ermöglichen und stellen ein hohes Innovationspotenzial für neu entwickelte Metadateninformationssysteme dar. Idealerweise werden diese Metadateninformationssysteme zur Visualisierung direkt in die GIS-Software integriert.

Visualisierung ist der Prozess der Erkundung, Transformierung und bildlichen Darstellung von Daten. Das Ziel der Visualisierung ist das Verständnis und den Einblick in Daten zu erhöhen. Abbildung 1 zeigt den Visualisierungsprozess. Der Prozess konzentriert sich auf die Daten. Als erster Schritt werden die Daten von einer Quelle erworben. Als nächstes werden die Daten mittels verschiedener Verfahren transformiert. Danach werden sie in einer für den Benutzer angebrachten Form abgebildet. Das Darstellen der Daten schließt den Prozess ab (SCHROEDER et al. 2006).

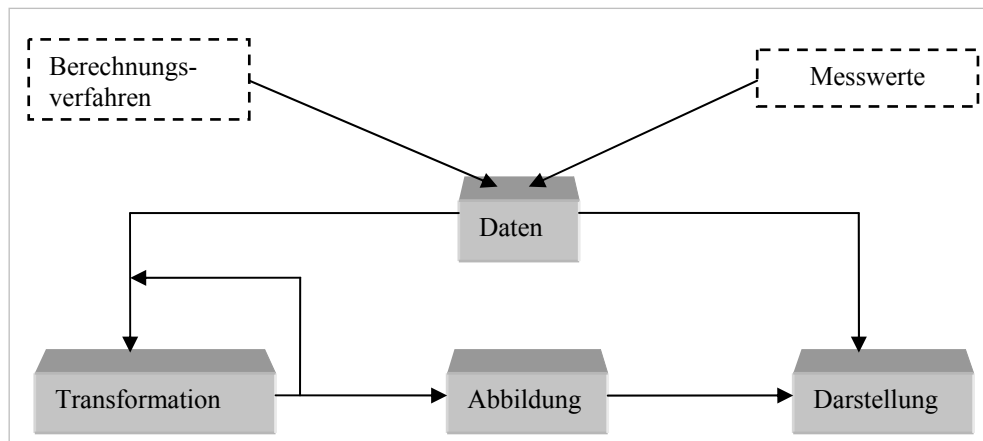


Abb. 1: Der Visualisierungsprozess (SCHROEDER et al. 2006)

Mittels Visualisierungen wird versucht, dem Betrachter die Entwicklung eines mentalen Modells zu ermöglichen. Dieses mentale Modell umfasst somit einen Kontext, in den der Betrachter die Darstellung einordnet. In erster Linie ist der Erfolg der Visualisierung davon abhängig, in welchem Maße der Betrachter dazu fähig ist, den Kontext der realen Welt aus der Abbildung zu rekonstruieren. Nur wenn es gelingt, die wahrgenommenen Strukturen in der Abbildung mit den tatsächlich existierenden Korrelationen zwischen Parametern in Verbindung zu bringen, lässt sich die Visualisierung als ein entscheidendes Hilfsmittel in der Datenanalyse einsetzen. (SCHUMANN et al. 2000)

Des Weiteren ist es bei Visualisierungen wichtig, dass ein dargestelltes Datenelement im Kontext des gemeinsamen Informationsraumes sichtbar bleibt, um Vergleich mit anderen Datenelementen durchführen zu können. Es muss aber darauf geachtet werden, es hinreichend detailliert darzustellen, um es interpretieren zu können. (REITERER et al. 2003)

3 Umsetzung

Im Rahmen des von der Österreichischen Forschungsförderungsgesellschaft finanzierten Forschungsprojektes V4D wurden für die GIS-Software IGLIS Konzepte und Methoden für eine automatische Ableitung der für die Beurteilung der Datenqualität relevanten Metadatenelemente entwickelt. Vonseiten der Nutzeranforderungen wurden die Qualitätskriterien Aktualität und Vollständigkeit als wesentliche Eigenschaften für die Beurteilung der Datenqualität definiert. Für Geodaten, die mittels terrestrischer Vermessung erhoben werden, kann das Qualitätskriterium der Aktualität basierend auf dem Messdatum der einzelnen Koordinaten abgeleitet werden. Werden die Geodatensätze mit Methoden der Fernerkundung generiert, wird das Datum der Befliegung für die Visualisierung des Qualitätskriteriums der Aktualität herangezogen. Die Visualisierung der Vollständigkeit erfolgt mittels der Berechnung von Koordinatendichten für einen bestimmten geographischen Bereich.

Komplexe Visualisierungen sind, speziell für Menschen die auditiv (Gehör) und haptisch (Tastsinn) veranlagt sind, nur schwer intuitiv zu erfassen, und somit schwer interpretierbar. Mit diesem Hintergrund wurde das entwickelte Visualisierungskonzept bewusst einfach gehalten. Dadurch stellen die Visualisierungen auch für fachfremde Personen eine erste Interpretationsgrundlage für eine möglichst rasche und einfache Entscheidungsfindung dar. Das erarbeitete Visualisierungskonzept basiert auf der Darstellung eines Rasters und wurde bereits von den IGLIS Nutzern als praxistauglich evaluiert und prototypisch umgesetzt.

Die relevanten Informationen werden aus den Stützpunkten einer Geometrie rekonstruiert. Die Metadateninformationen können nicht für jeden Stützpunkt gesondert visualisiert werden. Stattdessen wird eine bestimmte Anzahl von Punkten aggregiert. Die Aggregation der einzelnen Punkte erfolgt basierend auf der Bezugsgeometrie, die durch ein regelmäßiges Raster gebildet wird. Jene Punkte, die innerhalb einer Rasterzelle liegen, werden aggregiert.

Die prototypische Umsetzung des Visualisierungskonzeptes ermöglicht eine freie Parametrierung verschiedener Eingangsvariablen, um für die unterschiedlichen Fragestellungen ein optimiertes Visualisierungsergebnis zu liefern. Für die auf einem Rasteransatz basierte Visualisierung kann zum Beispiel die Rasterzellengröße variiert werden. Durch diese freie Parametrierbarkeit des Raumbezuges können die Daten in einem unterschiedlichen räumlichen Kontext visualisiert und interpretiert werden.

Zusätzlich ist eine Parametrierung des Themenbezugs möglich. Die Geodaten sind thematisch in verschiedene Bereiche eingeordnet. Durch verschiedene Eingangsvariablen wird somit zusätzlich zu der räumlichen Filterung der Geodaten eine thematische Filterung ermöglicht.

In Abbildung 2 und Abbildung 3 sind verschiedene Visualisierungsergebnisse dargestellt, die mithilfe des entwickelten Prototyps erstellt wurden. Abbildung 2 zeigt die Visualisierung von Metadateninformationen mit einer Rasterzellgröße von fünfundzwanzig Metern. Die hier dargestellten Graustufen visualisieren das Qualitätskriterium Aktualität. Je heller die Rasterzelle ist, desto aktueller sind die vorhandenen Daten. In der Praxis wird das Qualitätskriterium Aktualität jedoch durch unterschiedliche Farben visualisiert. Aktuelle Daten werden grün dargestellt, veraltete Daten rot. Des Weiteren wird in der Praxis durch die Transparenz der Rasterzellen die Koordinatendichte, die als Maßzahl für das Qualitätskriterium der Vollständigkeit eines Themenbereichs interpretiert werden kann, visualisiert. Durch die schwarz-weiße Abbildung der Visualisierungsergebnisse wird die Transparenz hier jedoch schlecht bis kaum erkennbar.

Für manche Aufgabenstellungen ist es von Vorteil, eine sehr große Rasterzellgröße zu wählen. Abbildung 3 zeigt ein Visualisierungsergebnis mit einer Rasterzellgröße von 500 Metern. Hier erkennt man zum Beispiel, aufgrund der unterschiedlichen Graustufen des Visualisierungsergebnisses, in welchen Quadranten aktuelle Daten vorhanden sind. Da bei der Schwarz-Weiß-Abbildung weiße Rasterzellen aktuelle Daten und dunkle Rasterzellen veraltete Daten visualisieren, zeigt Abbildung 3, dass nur im linken oberen Quadranten aktuelle Daten vorliegen.



Abb. 2: Visualisierungsergebnis mit einer Rasterzellengröße von 25 m

Abb. 3: Visualisierungsergebnis mit einer Rasterzellengröße von 500 m

Der Benutzer muss entscheiden, welche Rasterzellgröße die für seine Fragestellung relevanten Informationen liefert. Damit zu klein gewählte Rasterzellgrößen, speziell bei fachfremden Personen, nicht zu Fehlinterpretationen führen, muss die Auswahlmöglichkeit der Rasterzellgrößen sinnvoll eingeschränkt werden.

4 Ergebnisse und Ausblick

Die Visualisierung von geographischen Metadateninformationen erschließt neue Inhalte und ermöglicht eine gezieltere Ressourcenplanung für die Überarbeitung der Geodatenbestände. Dies führt zu einer erheblichen Aufwand- und Kostenersparnis für das Unternehmen. Solche Visualisierungen ermöglichen eine rasche und zuverlässige Auskunft bezüglich der Datenqualität der Geodaten, da ein Bild oft mehr sagt als tausend Worte.

Visualisierungen von Metadateninformationen unterschützen bei der Beurteilung von EU-weiten heterogenen Datenbeständen und ermöglichen einen Vergleich zwischen ihnen. Erst durch das Überführen von existierenden Datenbeständen in das von INSPIRE vorgegebene Datenmodell, sind EU-weit homogene Datenbestände verfügbar. Für eine erste Evaluierung, ob bestimmte Datenbestände verfügbar sind und deren Qualität ausreicht um INSPIRE gerecht zu werden können Visualisierungen von Metadateninformationen herangezogen werden. Die Visualisierungsergebnisse helfen auch bei der Entscheidungsfindung für welche geographischen Bereiche keine Daten vorliegen und erhoben werden müssen.

Bei der Datenzusammenführung ist es von großem Interesse zu wissen, inwieweit die Datenbestände homogen sind. Dafür ist es sinnvoll Metadaten und Visualisierungen von Metadateninformationen heranzuziehen.

Eine nahtlose Integration von Metadaten systemen unmittelbar in die Produktionsketten würde einen wesentlichen Beitrag zur Erhöhung der Akzeptanz und Bereitschaft für die Erfassung und Wartung von Metadaten beitragen.

Literatur

- BARTELME, N. (2005): Geoinformatik. Springer-Verlag, Berlin/Heidelberg/New York.
- LANDESVERMESSUNG U. GEOBASISINFORMATION BRANDENBURG, SENATSVERWALTUNG F. STADTENTWICKLUNG (2009): <http://gdi.berlin-brandenburg.de/papers/Metadatenbroschuere.pdf>, (25.01.2010).
- REITERER, H, LIMBACH, T., MÜLLER, F., KLEIN, P. & JETTER, C. (2003): Ein visueller Metadaten Browser für die explorative Erkundung großer Datenmengen. – <http://mc.informatik.uni-hamburg.de/konferenzbaende/mc2003/konferenzband/muc2003-16-reiterer.pdf>, (20.04.2010).
- SCHROEDER, W., MARTIN, K. & LORENSEN B. (2006): The Visualization Toolkit. Pearson Education, Inc.
- SCHUMANN, H. & MÜLLER W. (2000): Visualisierung: Grundlagen und allgemeine Methoden. Springer-Verlag, Berlin/Heidelberg.
- WASSILIOS K. (2005): „Mit Metadaten Kosten sparen“. In: GeoBIT – Geoinformationstechnologie für die Praxis – Fachzeitschrift für raumbezogene Informationstechnologie, 10.