
Implementation systematischer Qualitätssicherungs- und Modellierungsroutinen bei der Verwendung von GIP-Daten

Martin Loidl, Bernhard Zagel und Robin Wendel

IFFB Geoinformatik, Universität Salzburg · martin.loidl@sbg.ac.at

Full paper double blind review

Zusammenfassung

Bedingt durch entsprechende Gesetzesinitiativen (IVS Gesetz) und ein langfristiges Förderregime, hat sich die Graphenintegrationsplattform (GIP) als Standard für die behördliche Haltung straßenbezogener Daten in Österreich etabliert. Ursprünglich konzipiert für die Verwendung in behördeninternen Verwaltungsabläufen, wird die GIP vermehrt als Grundlage diverser Anwendungen herangezogen. Die beiden bekanntesten Anwendungen sind die Verkehrsauskunft Österreich (VAO) und basemap.at. Daneben kommen Daten aus der GIP zunehmend in regionalen Projekten zum Einsatz. Eine breite Verwendung der GIP-Daten, außerhalb der Behörden und in unterschiedlichsten Domänen, ist durch Open Government Data(OGD)-Initiativen absehbar.

Um die Daten aus GIP auch für spezifische Anwendungen außerhalb der Behörden bzw. unabhängig vom Datenhalter nutzbar zu machen, erweisen sich die Implementation systematischer Qualitätskontrollen und Modellierungsroutinen als vorteilhaft. Dieser Beitrag zeigt anhand eines konkreten Fallbeispiels, welcher Mehrwert in einer behördenexternen Anwendung durch die Evaluierung der Qualität bzw. Eignung der Daten und räumliche Modellierungsschritte erzielt werden kann. Die Konzepte und Workflows werden generisch präsentiert, um die Übertragbarkeit auf weitere Anwendungen zu gewährleisten.

1 GIP-Daten

Über die Hintergründe hinsichtlich Organisation und technischem Aufbau der Graphenintegrationsplattform (GIP) liegen zahlreiche Berichte und Publikationen vor (z. B. HEIMBUCHNER 2014). An dieser Stelle werden deshalb nur einige Implikationen angeführt, die für NutzerInnen von GIP-Daten außerhalb der Behörden von Relevanz sind.

1.1 GIP in den Behörden

Bei der GIP handelt es sich um einen dezentral organisierten, behördlichen Datenbestand, der vordergründig für die Bewältigung hoheitlicher Aufgaben im Infrastruktur- und Verkehrsbereich dient. Die verschiedenen Behörden, Infrastrukturbetreiber und Tarifverbände, die eigene GIP-Instanzen verwalten, kommunizieren periodisch ihren aktuellen Datenbestand an einen zentralen Server, wo die Daten zusammengeführt und für Anwendungen, die

über die jeweiligen Hoheitsgebiete hinausgehen (z. B. nationaler Routingdienst) bereitgestellt werden. Verwaltet und bearbeitet werden die dezentralen GIP-Datenbestände mit einem eigenen, proprietären Client, der eine integrative Sicht auf die verschiedenen Informationsschichten der GIP (die zum Teil in getrennten Datenbanken hinterlegt sind) erlaubt. Diese „direkte“ Sicht auf den GIP-Datenbestand ist für Akteure außerhalb von Behörden bzw. GIP-Konsortialpartnern nicht möglich.

1.2 INTREST Export-Schnittstelle

Um Anwendungen und externen Stellen Zugriff auf die GIP-Daten zu ermöglichen, existiert eine definierte Export Schnittstelle, die sich des INTREST Data Formats (IDF) als Austauschformat bedient. Dabei handelt es sich um eine ASCII-Datei, die die Inhalte der relationalen GIP-Datenbank(en) wiedergibt. Details zur Schnittstelle finden sich in der entsprechenden Dokumentation¹.

Die Schnittstelle in ihrer aktuellen Version (Stand Dezember 2014) berücksichtigt nicht alle Daten, die in der GIP tatsächlich vorliegen. Das bedeutet, der exportierte Datenbestand ist in seinem Umfang nicht ident mit der GIP an sich. Da sich die Definition der Exportschnittstelle primär an den Erfordernissen einiger weniger Anwendungen (allen voran die Verkehrsauskunft Österreich) orientiert, kann es sein, dass für spezifische Anwendungen mitunter relevante Inhalte nicht übergeben werden können. Um diesem Umstand für radverkehrsspezifische Anwendungen zu begegnen, wurde vom GIP-Konsortium ein sogenannter „Export Korrektor“ in die Exportroutine integriert. Damit können einige radaffine Informationen aus der GIP mit in die Austauschdatei übergeben werden.

1.3 INTREST nach GeoDB/GIS

Um aus den Tabellen, die in der ASCII-Datei des INTREST Exports enthalten sind, einen routingfähigen Graphen samt den beschreibenden Attributen für die Verwendung in einem geographischen Informationssystem (GIS) generieren zu können, müssen die Daten in den Tabellen in einem ersten Schritt zusammengeführt bzw. miteinander in Beziehung gesetzt werden. Im Anwendungsbeispiel, das in Abschnitt 4 beschrieben wird, wurde dazu ein Workflow definiert, mit dem die Zusammenführung in einer PostgreSQL/PostGIS Datenbank umgesetzt wurde. Für die weitere Modellierung und Analyse wurden Python bzw. ArcPy in einer QGIS und ArcGIS Umgebung verwendet.

2 Qualitätssicherung

Qualitätskontrollen räumlicher Daten umfassen mehrere, oftmals terminologisch unscharf getrennte Aspekte. Im Wesentlichen muss zwischen der Eignung von Daten („Fitness for use“ (DEVILLERS et al. 2007) und der eigentlichen Datenqualität, wie in BARRON et al. (2014) und VEREGIN (1999) zusammengefasst, unterschieden werden. Im Folgenden liegt der Fokus auf dem zweiten Aspekt. Zur Evaluierung der Geodatenqualität gibt es verschiedene Ansätze, wobei die ISO Norm 19157² als Synthese maßgeblich ist. Zur Beurteilung

¹ Dokumentation der Schnittstelle SST 4-10.

² <https://www.iso.org/obp/ui/#iso:std:iso:19157:ed-1:v1:en>

der Qualität straßenbezogener Daten sind folgende Aspekte von besonderer Bedeutung: geometrische Vollständigkeit, attributive Konsistenz (Widerspruchsfreiheit auf Attributebene) und die Topologie. Zusätzlich kann die Lagegenauigkeit eine Rolle spielen (z. B. bei der Visualisierung). Während die geometrische Vollständigkeit und die Lagegenauigkeit nur mithilfe von Referenzdatenbeständen beurteilt werden können, bietet sich bei der attributiven Konsistenz (LOIDL et al. 2014) und der Topologie eine intrinsische Evaluierung an.

2.1 Extrinsische Qualitätskontrolle

Bei der extrinsischen Qualitätskontrolle werden Referenzdatensätze oder eine Besichtigung vor Ort zur Feststellung der Datenqualität herangezogen („Ground truthing“). Hierbei ist es wichtig, dass die Qualität und Eignung der Referenzdatensätze gegeben ist.

Die geometrische Vollständigkeit lässt sich über die Anzahl der Elemente eines Netzwerks oder über die Netzwerklänge feststellen. Um räumliche Variabilitäten berücksichtigen zu können, ist es erforderlich das Netzwerk zu unterteilen; bei globaler Betrachtung kommt es ansonsten zu Glättungseffekten (vgl. KOUKOLETOS et al. 2012). Eine analytische bzw. statistische Feststellung der geometrischen Vollständigkeit im Vergleich zu Referenzdatensätzen, lässt sich nur dann umsetzen, wenn beiden Datensätzen das gleiche Datenmodell zugrunde liegt. Wird beispielsweise bei einem Straßendatensatz der Querschnitt rein attributiv beschrieben und nur die Straßenmittellachse geometrisch abgebildet, ergeben sich enorme Unterschiede, wenn der Vergleichsdatsatz einem anderen Datenmodell (z. B. eigene Geometrie je Fahrtrichtung) folgend aufgebaut ist. Unter anderem aufgrund dieser Limitation der analytischen Herangehensweise, wird die geometrische Vollständigkeit oftmals visuell evaluiert.

Ähnlich wie die geometrische Vollständigkeit lassen sich Aussagen zur Lagegenauigkeit mithilfe von Referenzdatensätzen bzw. einer Überprüfung vor Ort treffen. Während es für die Lagegenauigkeit von Punkten mit dem RMS eine etablierte statistische Größe gibt, existiert für die Bewertung der Lagegenauigkeit von Linien kein standardisiertes Verfahren (VEREGIN 1999). Die meisten Ansätze verwenden, ähnlich wie GOODCHILD & HUNTER (1997), Referenzdatensätze mit einer angenommenen, höheren Genauigkeit und generieren Maßzahlen basierend auf der Deckung der untersuchten Linie mit dem Puffer um die Referenzlinie (vgl. GRASER et al. 2013).

2.2 Intrinsische Qualitätskontrolle

Die intrinsische Qualitätskontrolle baut auf der inneren Struktur und internen Abhängigkeiten von Datensätzen auf. Referenzdatenbestände sind hierbei nicht notwendig.

Die Sicherstellung der korrekten, geometrischen Topologie ist im Kontext der Netzwerkanalyse von größter Wichtigkeit. Topologische Fehler, die oftmals durch ungenaues Editieren entstehen, können beispielsweise mit einem heuristischen, semi-automatisierten Verfahren detektiert werden. Dabei werden Endpunkte in einem Netzwerk identifiziert, in deren unmittelbarer, räumlicher Nähe sich eine weitere Linie mit anderer ID befindet. Auf diese Art und Weise können Verdachtsfälle identifiziert und genauer überprüft werden³.

³ Nähere Ausführungen dazu auf <https://gicycle.wordpress.com/2014/08/29/data-quality-topology/> [Zugriff am 21.01.15].

Die attributive Konsistenz setzt voraus, dass, erstens überlappende bzw. redundante Attribute vorhanden sind und zweitens, dass diese (abhängigen) Attribute widerspruchsfrei sind (VEREGIN 1999). Die Überprüfung der attributiven Konsistenz kann unter anderem mithilfe von kombinierten Abfragen erfolgen (LOIDL et al. 2014).

3 Mehrwert durch Modellierung

Geographische Anwendungen, die auf Straßendatensätzen – unabhängig ob behördliche oder communitybasierte Daten – aufbauen, werden in der Regel für einen sehr spezifischen Zweck konzipiert. Dagegen sind die Datensätze selbst generischer konzipiert oder in ihrer Struktur und ihrem Inhalt aus einer Vielzahl von Datenerfassungen für unterschiedliche Anwendungsfälle gewachsen. Folglich gibt es nur in den seltensten Fällen eine hundertprozentige Eignung eines bestehenden Datensatzes für eine Anwendung. Im Fall der GIP dient die Datenbasis diversen behördlichen Anwendungen, die bei der Spezifizierung des Datenmodells zum Großteil bekannt waren. Folglich konnte bei der Konzeption des Datenmodells und der Attributstruktur auf diese ursprünglichen Anwendungskontexte Rücksicht genommen werden. Beispiele für die ursprünglich intendierte Verwendung der GIP reichen von Verordnungen im Straßenraum bis hin zu Förderagenden oder Baustellenmanagement. Da die Daten aus der GIP für viele Anwendungen außerhalb der eigentlichen Verwaltungstätigkeit eine gute Grundlage bilden, wurden in den vergangenen Jahren, und verstärkt seit diversen OGD-Initiativen neue, zum Teil sehr spezifische Anwendungen umgesetzt. Greifen nun spezifische Anwendungen auf generische oder für andere Zwecke intendierte Datenbanken direkt zu, wirkt sich dies entsprechend auf die Ergebnisse aus; z. B. derart, dass relevante Informationen nicht explizit, sondern verteilt über mehrere Attribute hinterlegt sind.

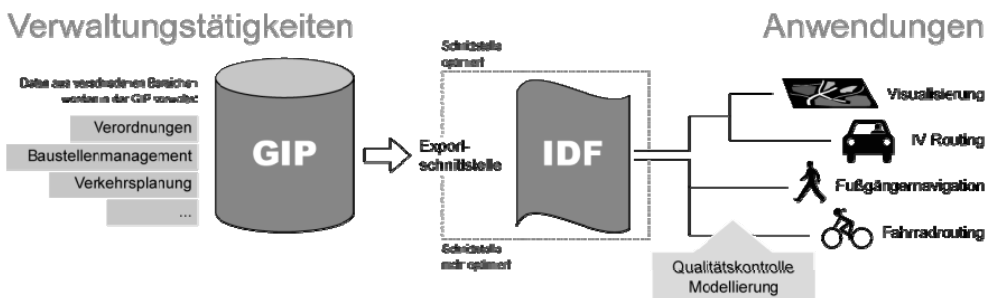


Abb. 1: Die aus der GIP exportierte IDF-Datei ist für bestimmte Anwendungen optimiert. Andere, spezifische Anwendungen, die ebenfalls die IDF-Datei unmittelbar nutzen, generieren z. T. wenig plausible Ergebnisse, da die Daten nicht entsprechend optimiert vorliegen. Implementierte Qualitätskontrollen und Modellierungsroutinen ermöglichen qualitative Anwendungen, trotz nicht optimierter Datenbanken und Schnittstellen.

Neben der Implementation von Prüfroutinen (Qualitätskontrolle) in die Datenaufbereitung kann durch eine zusätzliche Modellierung auf Attributebene ein Mehrwert für die Anwen-

dung erzielt werden, ohne, dass die Ausgangsdaten an sich geändert werden müssen. Dieser Modellierungsschritt fungiert als Schnittstelle zwischen der generischen Datenbank und den spezifischen Anforderungen der Anwendung. Wie LOIDL & ZAGEL (2014) am Beispiel von behördlichen und communitybasierten Daten zeigen, kann mit einem derartigen Vorgehen ein optimiertes Ergebnis, unabhängig vom Datenmodell und der Attributstruktur der verwendeten Datenbasis, erzielt werden. Wie in Abschnitt 1.3 angedeutet wurde, umfasst die Modellierung bei der Aufbereitung der Daten nicht nur die Attributebene, sondern auch die Geometrie bzw. den Graphen an sich. Dazu liegen, bezogen auf Daten aus der GIP bereits Arbeiten, beispielsweise von GEROE et al. (2011), vor.

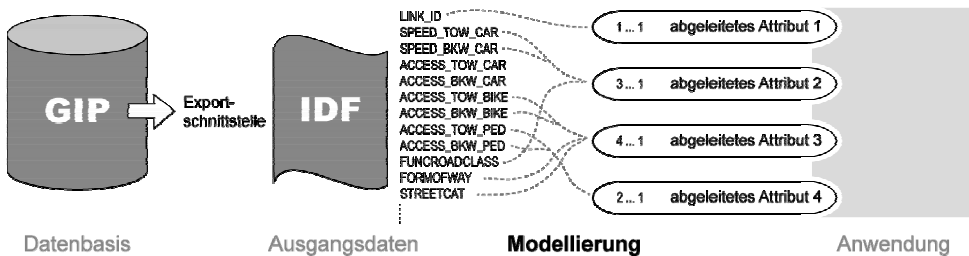


Abb. 2: Modellierung abgeleiteter Attribute für spezifische Anwendungen

Im nachfolgenden Anwendungsbeispiel, das auf Daten aus der GIP aufbaut, werden, für die Anwendung spezifizierte, Attribute aus den in der Datenbasis vorhandenen Attributen abgeleitet. Diese abgeleiteten Attribute stehen in einer 1-n-Relation zu den vorhandenen Attributen (siehe Abbildung 2).

4 Anwendungsbeispiel: sicherheitssensitives Fahrradrouting

Das Fahrradroutingportal radlkarte.info⁴ bietet, neben diversen Informationsangeboten für RadfahrerInnen, die Möglichkeit Routen nach der Verkehrssicherheit („empfohlene Route“) zu optimieren (LOIDL & ZAGEL 2014). Aktuell (März 2015) ist das Service für den Salzburger Zentralraum, inklusive der angrenzenden, bayrischen Gemeinden, verfügbar. Mit der Zusammenführung der straßenbezogenen Datenbestände der Stadt und des Landes Salzburg in einer gemeinsamen GIP, entstand die Anforderung, für die österreichische Seite des Projektgebiets ausschließlich die jeweils aktuellen Daten aus der GIP als Datengrundlage zu verwenden.

Die wesentlichen Herausforderungen, die sich vor der Umstellung der Datenbasis stellten, können wie folgt zusammengefasst werden:

- Die Datengrundlage und die Exportschnittstelle sind nicht für die Verwendung in einem Fahrradrouting optimiert (ein sogenannter „Export Korrektor“ adaptiert beim Generieren des INTREST-Exports den Datensatz um fahrradrelevante Informationen zu inkludieren). Da sich sowohl die benötigten Informationen an sich, als auch die Art

⁴ <http://www.radlkarte.info/> [Zugriff am 21.01.15].

der Abbildung, beispielsweise was die Detailtiefe anbelangt, von den Anforderungen eines MIV-zentrierten Routingangebots unterscheiden, wird zum Teil notwendige Information nicht, oder ungeeignet abgebildet.

- Die Qualität der Daten ist, bedingt durch unterschiedliche Stadien der Datenerfassung und -aufbereitung bei den verschiedenen Gebietskörperschaften, innerhalb des Projektgebiets heterogen.
- Die bestehende Anwendung ist für Daten mit einem anderen Datenmodell und einer abweichenden Attributstruktur konzipiert.

Vor diesem Hintergrund wurde ein mehrstufiger Workflow definiert, der sich aus der Datenaufbereitung, einer strukturierten Qualitätskontrolle, der Modellierung abgeleiteter Attribute und der Bereitstellung eines bewerteten, routingfähigen Netzes zusammensetzt (siehe Abb. 3).

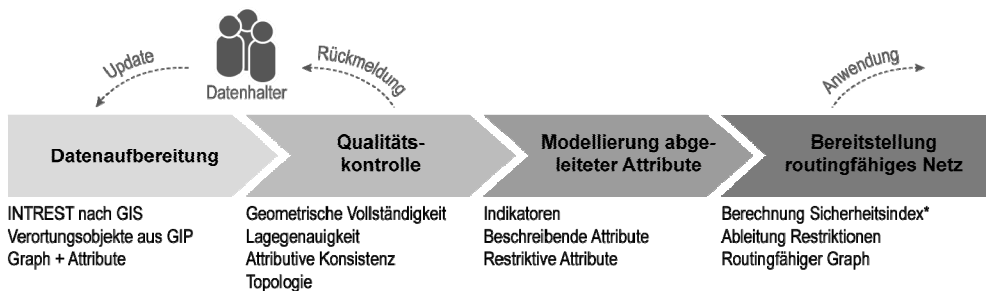


Abb. 3: Workflow für Verwendung von Daten aus der GIP für ein Fahrradroutingportal. Die Berechnung des Sicherheitsindex^{*} beruht auf LOIDL & ZAGEL (2014).

Die Datenaufbereitung wurde konzeptionell bereits in Abschnitt 1.3 behandelt; technische Details sind in diversen Schnittstellenbeschreibungen enthalten. Die Bewertung des Wegenetzes hinsichtlich der Verkehrssicherheit mithilfe eines indikatorenbasierten Bewertungsmodells, die als Widerstandswert im sicherheitssensitiven Fahrradrouting verwendet wird, wurde bereits an anderer Stelle ausführlich beschrieben (LOIDL & ZAGEL 2010, LOIDL & ZAGEL 2014). Deshalb liegt der Fokus der beiden folgenden Abschnitte auf den Methoden und Ergebnissen der Qualitätskontrolle und Modellierung im Projektgebiet. Insgesamt wurde ein Wegenetz mit einer Gesamtlänge von 1.960 km verarbeitet. Die angeführten Ergebnisse beziehen sich auf den Datensatz zu unterschiedlichen Zeitpunkten und werden hier lediglich exemplarisch angeführt.

4.1 Qualitätskontrolle GIP-Daten

Für eine Abschätzung der geometrischen Vollständigkeit wurde der Datensatz im direkten Vergleich mit zwei Referenzdatensätzen – einem zeitgleich extrahierten OpenStreetMap Datensatz und der letzten Version des städtischen Wegenetzes vor der Überführung der Daten in die GIP – evaluiert. Eine analytisch-statistische Qualitätskontrolle erwies sich, bedingt durch die unterschiedlichen Datenmodelle (Abbildung 4 (1)), in diesem Fall als nur eingeschränkt passend. Die visuelle Evaluierung ergab einen grundsätzlich hohen Vollständigkeitsgrad, wobei ein Aspekt beim Vergleich mit OSM Daten, gerade im Bereich des

Radverkehrs zu beachten ist: in OSM werden grundsätzlich alle Wege digitalisiert die sichtbar und/oder benutzbar sind (Abbildung 4 (3)), während in der GIP nur öffentliche bzw. allgemein benutzbare Wege hinterlegt sind. Für die Qualität bzw. Akzeptanz der Anwendungen ist aber oft das Vorhandensein direkter, mitunter informeller Verbindungen ausschlaggebend⁵. Dies ist allerdings weniger ein Qualitätskriterium als ein Aspekt der Eignung der Daten.

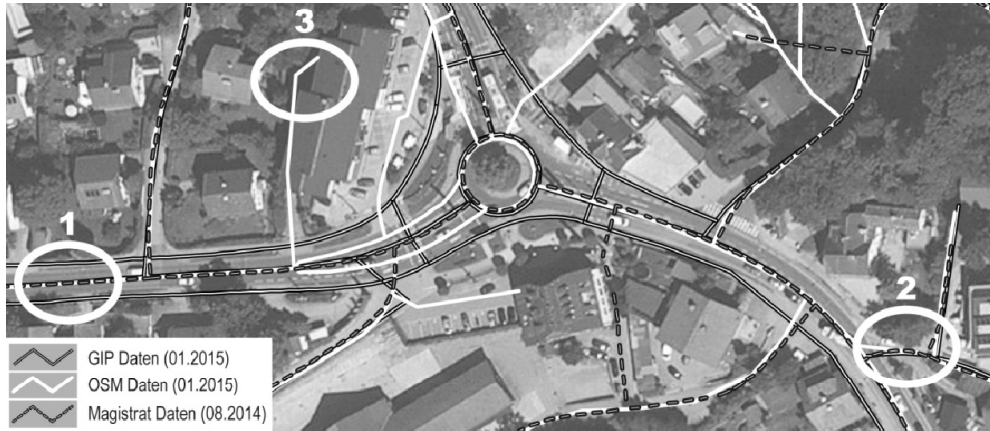


Abb. 4: Beispielausschnitt der visuellen Evaluation der Vollständigkeit und Lagegenauigkeit. Beschreibung der drei Details im Text.

Die Lagegenauigkeit der GIP-Daten ist innerhalb der Stadt – wo die Daten vom Magistrat übernommen wurden – sehr hoch. In den Stadtumlandgemeinden zeigen sich bei einer visuellen Evaluation zum Teil Ungenauigkeiten, die vor allem auf die Datengrundlage beim Digitalisieren der Daten (vielfach vom Orthophoto) zurückzuführen sein dürften. Abgesehen davon ist bei den vorhandenen Referenzdatensätzen nicht von einer höheren Genauigkeit auszugehen, was die quantitative Beurteilung der Abweichungen praktisch verunmöglicht (Abbildung 4 (2)). Die absolute Lagegenauigkeit spielt im konkreten Anwendungsfall ohnehin eine untergeordnete Rolle; für Routinganwendungen ist die korrekte Topologie in einem möglichst vollständigen Wegenetz von weitaus größerer Bedeutung.

Mittels des in Abschnitt 2.2 beschriebenen Verfahrens, wurden im Projektgebiet knapp 60 topologische Fehler identifiziert und in der Folge vom Datenhalter korrigiert. In der Mehrzahl der Fälle handelte es sich um Digitalisierfehler (Lücken bzw. Overshoots).

Die Qualitätskontrolle der attributiven Konsistenz machte sich die relativ hohe Redundanz der Attributdaten zunutze, indem kombinierte Abfragen an den Datensatz gestellt wurden. Dabei wurde ersichtlich, dass in einigen Fällen der funktionalen Abhängigkeit der Attribute nur bedingt Rechnung getragen wurde. So existierten beispielsweise Segmente die in der Attributspalte „FOW“ (Form of Way) den Wert 505 = ausschließlich Radweg und in der Attributspalte „Basetype“ den Wert 7 = ausschließlich Gehweg hinterlegt hatten. Neben

⁵ Siehe dazu https://gicycle.wordpress.com/2014/02/05/data_quality/ [Zugriff am 21.01.15].

solchen eindeutigen Widersprüchen zeigte sich auch, dass einzelne Attributausprägungen zum Teil uneindeutig spezifiziert und angewendet werden, was zu „verschwommenen“ Inkonsistenzen führte. Wird in einem Attribut eines Segments beispielsweise ein Wert für einen gemischten Geh- und Radweg hinterlegt und in einem weiteren der Wert für einen Gehweg, ist – ohne Zuhilfenahme weiterer Attribute – nicht eindeutig, wie diese Attribute zu interpretieren sind (Teilmenge oder Inkonsistenz). Insgesamt ist die Attributstruktur der GIP nicht für die Fahrradwendungen ausgelegt (gleiches gilt auch für Fußgänger). Abhilfe schafft hier zum Teil der „Export Korrektor“, der im Wesentlichen aus verschiedenen Attributkombinationen fahrradspezifische Attribute im INTREST Export anfügt. Einem ähnlichen Ansatz folgt die Modellierung abgeleiteter Attribute.

4.2 Modellierung abgeleiteter Attribute

Um das für die Bewertung des Wegenetzes verwendete indikatorenbasierte Bewertungsmodell parametrisieren zu können, ist es notwendig, die Verbindung zwischen den Indikatoren und deren Wertausprägung und den in der Datenbasis vorhandenen Attribute zu definieren. Dieses Schema ermöglicht eine, vom Datenmodell bzw. der Attributstruktur unabhängige Konzeption des Bewertungsmodells bzw. der Anwendung.

Aus den Attributen der IDF-Datei und den zusätzlich zur Verfügung gestellten Verortungsobjekten (z. B. für beschilderte Radrouten) wurden Indikatoren samt Wertausprägung, wie in Tabelle 1 zusammengefasst, mittels SQL-Abfragen abgeleitet. Für die Ableitung der Neigungsinformation wurde der Graph mit einem 1-m-DGM verschnitten. Je nach Datenverfügbarkeit kann das Indikatorenset erweitert werden; beispielsweise um die Information der KFZ Querschnittsbelastung.

Tabelle 1: Abgeleitete Indikatoren, inklusive Wertausprägungen

Indikator	Wertausprägung
Anzahl der Fahrspuren	[1 ... n]
Art der Fahrradinfrastruktur	[Radweg; Geh- und Radweg; Rad- oder Mehrzweckstreifen; Busspur für Radverkehr geöffnet]
Belagsart	[Asphalt; Schotter; unbefestigt; Kopfsteinpflaster]
Breite der Fahrbahn	[1 ... n] m
Höchstzulässige Geschwindigkeit	[0 ... 100] km/h
Neigungsklasse	[Bergauf 4; Bergauf 3; Bergauf 2; Bergauf 1; Eben; Bergab 1; Bergab 2; Bergab 3; Bergab 4]
Radroute (beschildert)	[Behördenempfehlung; nationale Route; Regionale Route; Lokale Route]
Straßenkategorie	[Hauptstraße Kategorie 1; Hauptstraße Kategorie 2; Gemeindestraße; Erschließungsstraße; verkehrsberuhigte Straße; Autofreie Straße; Für Radverkehr nicht passierbarer Weg]

Im indikatorenbasierten Bewertungsmodell (die Modellkomponenten und die Modellerstellung wird ausführlich in LOIDL & ZAGEL (2014) beschrieben) wird mit jedem Indikator ein relatives Gewicht (W_i) assoziiert. Die Gewichtung kann durch eine interaktive Eingabemaske je nach Anforderungen adaptiert werden. Die Wertaussprägungen je Indikator erhalten ebenfalls ein relatives Gewicht (S_i im Intervall $[0 \dots 1]$). Die standardisierte Bewertung je Segment ($0 =$ beste Eignung bzw. geringste Gefährdung, $1 =$ schlechteste Eignung bzw. höchste Gefährdung) für das gesamte Netzwerk ergibt sich dann wie folgt:

$$Index = \frac{\sum_{i=1}^n S_i * W_i}{\sum_{i=1}^n W_i}$$

Neben der Bewertung der einzelnen Segmente, die in weiterer Folge die Grundlage für ein Routing darstellt, werden aus den Ausgangsdaten noch restriktive (Fahrverbote, Einbahnen, Schiebegebote) und beschreibende Attribute (Belag, Straßenname) abgeleitet. Wie groß der Effekt des hier beschriebenen Vorgehens auf aufbauende Anwendungen hat, zeigt der Vergleich zweier Routinginformationssysteme, die beide auf ein und den gleichen Datensatz aufbauen.

5 Ergebnisse und Diskussion

Aktuell bieten zwei verschiedene Fahrradroutingportale ihre Dienste auf Basis der GIP-Daten an. Zum einen das österreichweite Service der Verkehrsauskunft Österreich (VAO) mit ihren jeweiligen Mandanten und zum anderen das Portal radkarte.info für den Salzburger Zentralraum. Da die exakt gleichen Daten zugrunde liegen, wird anhand der Routenvergleiche der Effekt der Modellierung deutlich (die Effekte aus der Qualitätskontrolle werden für beide Portale schlagend, da sie die Datenbasis an sich betreffen).

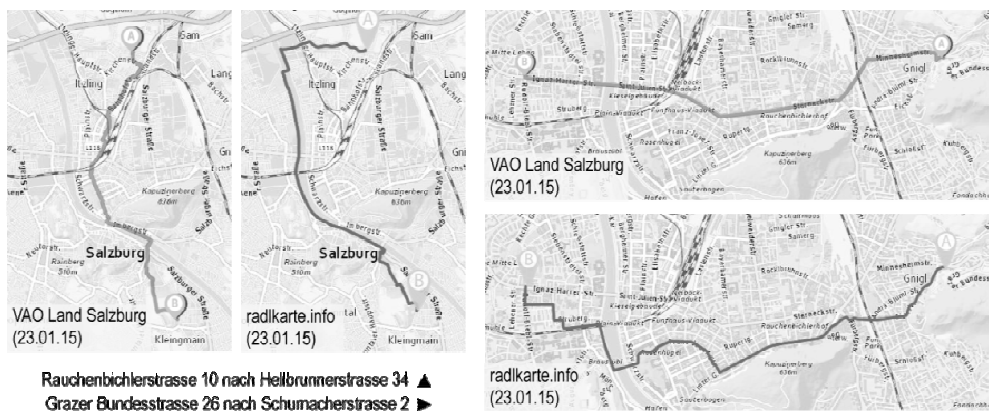


Abb. 5: Unterschiede im Fahrradroutingergebnis anhand zweier exemplarischer Start-Ziel-Relationen

Der in Abbildung 5 gezeigte Vergleich eignet sich nicht zur Bewertung der Plausibilität der Routingergebnisse (dafür fehlt die empirische Grundlage, wie Experteninterviews oder

User-Feedbacks), zeigt aber, wie sich die Routenempfehlungen der VAO für Radfahrer an der kürzesten Verbindung orientieren, während hinter den Vorschlägen der radkarte.info eine offensichtliche Präferenz für Radinfrastruktur und Straßen des niederrangigen Netzes zu erkennen ist.

Es ist anzunehmen, dass unter anderem bedingt durch die Open Government Data(OGD)-Initiativen, die Anzahl der Anwendungen, die auf GIP-Daten basieren in Zukunft steigen wird. Dabei wird es, wie im vorgestellten Beispiel, sehr häufig der Fall sein, dass die Datenbasis ursprünglich für andere Zwecke als die einer spezifischen Anwendung generiert wurde. Wie hier gezeigt werden konnte, kann durch die Modellierung auf Basis des vorhandenen Datensatzes die spezifische Anwendung bzw. die Ergebnisse daraus optimiert werden, ohne, dass die Datensätze selbst manipuliert werden müssten. Zusätzlich bietet der Ansatz den Vorteil, dass Anwendungen unabhängig vom Datenmodell bzw. der Attributstruktur konzipiert werden können. Die Modellierung fungiert mit der integrierten Harmonisierung und etwaigen Transformation bzw. Translation der Datenbestände als Scharnier. Eine zusätzliche Qualitätskontrolle trägt zur Verbesserung der Datenqualität und der Anwendungsergebnisse bei. Eine solche Qualitätskontrolle muss allerdings, im Fall der GIP, in einen entsprechenden Kommunikations- und Arbeitsablauf mit den Datenhaltern eingebunden sein.

Literatur

- BARRON, C., NEIS, P. & ZIPF, A. (2014), A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS*, 877-895.
- DEVILLERS, R., BÉDARD, Y., JEANSOULIN, R. & MOULIN, B. (2007), Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. *International Journal of Geographical Information Science*, 21 (3), 261-282.
- GEROE, D., NAVRATIL, G. & FIBY, H. (2011), Entwicklung eines Graphenmodells als Grundlage für eine Routingapplikation für den motorisierten Individualverkehr basierend auf der GIP. In: STROBL, J., BLASCHKE, T. & GRIESEBNER, G. (Hrsg.), *Angewandte Geoinformatik 2011*. Wichmann Verlag, Berlin/Offenbach, 347-356.
- GOODCHILD, M. F. & HUNTER, G. J. (1997), A simple positional accuracy measure for linear features. *International Journal of Geographical Information Science*, 11(3), 299-306.
- GRASER, A., STRAUB, M. & DRAGASCHNIG, M. (2013), Towards an Open Source Analysis Toolbox for Street Network Comparison: Indicators, Tools and Results of a Comparison of OSM and the Official Austrian Reference Graph. *Transactions in GIS*, 510-526.
- HEIMBUCHNER, K. (2014), GIP Day. In: STROBL, J., BLASCHKE, T., GRIESEBNER, G. & ZAGEL, B. (Hrsg.), *Angewandte Geoinformatik 2014*. Wichmann Verlag, Berlin/Offenbach, 467-478.
- KOUKOLETSOS, T., HAKLAY, M. & ELLUL, C. (2012), Assessing Data Completeness of VGI through an Automated Matching Procedure for Linear Data. *Transactions in GIS*, 16 (4), 477-498.
- LOIDL, M. & ZAGEL, B. (2010), Wie sicher ist sicher? – Innovatives Kostenmodell zur Ermittlung des Gefährdungspotenzials auf Radwegen. In: STROBL, J., BLASCHKE, T. & GRIESEBNER, G. (Hrsg.), *Angewandte Geoinformatik 2010*. Wichmann Verlag, Berlin/Offenbach, 394-403.

- LOIDL, M., KRAMPE, S., ZAGEL, B. & PUCHER, G. (2014), Aufbereitung von OpenStreet-Map-Daten für GIS-Modellierungen und Analysen. In: STROBL, J., BLASCHKE, T., GRIESEBNER, G. & ZAGEL, B. (Hrsg.), *Angewandte Geoinformatik 2014*. Wichmann Verlag, Berlin/Offenbach, 505-514.
- LOIDL, M. & ZAGEL, B. (2014), Assessing bicycle safety in multiple networks with different data models. In: VOGLER, R., CAR, A., STROBL, J. & GRIESEBNER, G. (Eds.). *GI_Forum 2014*. Wichmann, Berlin/Offenbach, 144-154.
- VEREGIN, H. (1999), Data quality parameters. In: LONGLEY, P. A., GOODCHILD, M., MAGUIRE, D. J. & RHIND, D. W. (Hrsg.). *Geographical Information Systems – Principles and Technical Issues*. John Wiley & Sons, New York, 177-189.