



RUNDER TISCH GIS E.V.

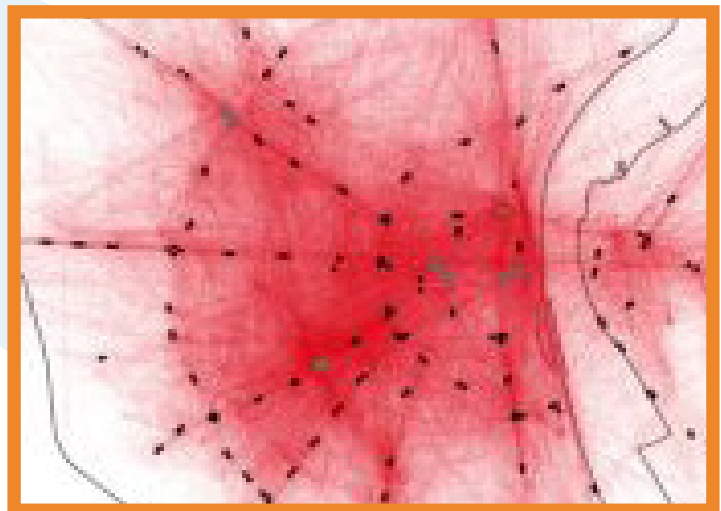


Technische Universität München

Kolbe / Bill / Donaubauer (Hrsg.)

Geoinformations- systeme 2021

Beiträge zur 8. Münchner GI-Runde



München, März 2021

Titelgrafik:

siehe Beitrag *Zusammenhänge zwischen Bikesharing und ÖPNV - eine geostatistische Analyse in Köln* von Katja Schimohr (Abb. 2)

Alle in diesem Werk enthaltenen Angaben, Daten, Ergebnisse, usw. wurden von den Autoren nach bestem Wissen erstellt und mit größtmöglicher Sorgfalt geprüft. Dennoch sind inhaltliche Fehler nicht völlig auszuschließen. Daher erfolgen die Angaben usw. ohne jegliche Verpflichtung oder Garantie der Autoren. Sie übernehmen deshalb keinerlei Verantwortung und Haftung für etwa vorhandene inhaltliche Unrichtigkeiten.

© 2021 Runder Tisch GIS e.V.

Erscheinungsweise

Digital als PDF unter www.rundertischgis.de/veranstaltungen/muENCHNER-GI-RUNDE.html

Urheberrecht

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Runden Tisches GIS e.V. unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Inhaltsverzeichnis

Vorwort	5
Keynote	
<i>Morich, Lutz:</i> Vernetzte Mobilität – notwendige neue Perspektive(n) für die Automobil-Industrie.....	6
Förderpreis Geoinformatik	
<i>Graf, Lukas:</i> Tiefe neuronale Netze und multi-temporale Statistik zur Fernerkundung bewässerter Landwirtschaft	9
<i>Lehmkuhl, Oliver:</i> Entwicklung einer Crowdsourcing-Anwendung zur automatisierten Detektion von Straßenschäden mit Hilfe von Methoden der Künstlichen Intelligenz als Webservice.....	13
<i>Schimohr, Katja:</i> Zusammenhänge zwischen Bikesharing und ÖPNV - eine geostatistische Analyse in Köln	18
<i>Schuldt, Carolin:</i> Meeresspiegelanstieg in Norddeutschland – Eine GIS-basierte Simulation und Visualisierung	22
<i>Herlé, Stefan:</i> A GeoEvent-driven Architecture based on GeoMQTT for the Geospatial IoT	25
<i>Kinne, Jan:</i> Webbasierte Innovationsindikatoren für mikrogeographische Analysen	28
<i>Kovács-Györi, Anna:</i> Defining and assessing urban livability – A GIScience perspective.....	38
Normung / Interoperabilität	
<i>Simonis, Ingo:</i> Neue OGC-Standards für Geo Web Services	44
<i>Clemen, Christian:</i> Trends der BIM und GIS Standardisierung – Bericht von der Joint ISO/TC 59/SC 13 - ISO/TC 211 WG: GIS-BIM....	45
<i>Baumann, Peter:</i> Coverage-Standards: Aktuelles aus OGC, ISO und INSPIRE.....	46
GIS und BIM: Vom Entwurf zur Nutzung	
<i>Jaud, Štefan:</i> Georeferencing	49
<i>Kaden, Robert:</i> BIM as Built	51
<i>Kutzner, Tatjana und Donaubaauer, Andreas:</i> CityGML 3.0 – Neue Konzepte und Perspektiven für einen Digitalen Lageplan in 3D	55
Geo-Algorithmen für Big Geospatial Geodata	
<i>Werner, Martin:</i> Next Steps for Big Geospatial Data	58
<i>Rieke, Matthes:</i> Verteilte und Effiziente Prozessierung von Geodaten mit KI-Methoden	61
<i>Hauert, Jan-Henrik:</i> Interaktive Karten für große Datensätze – eine algorithmische und kartographische Herausforderung.....	63
Vernetzte Mobilität	
<i>Bogenberger, Klaus:</i> Simulation und Steuerung von On-Demand Flotten.....	65
<i>Schön, Steffen:</i> Vernetzte dynamische Sensorplattformen: mehr Fahrzeuge - bessere Performance?.....	67

3D*Bauer, Wolfgang:*

Projekt Geomassendaten Nutzung von Geomassendaten der Bayerischen Vermessungsverwaltung durch Pilotanwender verschiedener Branchen..... 69

Engel, Frank:

Offene 3D-Geodaten in Thüringen 71

Coors, Volker:

Neue OGC-Standards als Katalysator für die Nutzung von 3D-Geoinformation 74

Geodateninfrastrukturen*Roscher, Ralf:*

Langzeitspeicherung von Geodaten 76

Stangl, Christian:

Digitalisierungsspeicher am IT-DLZ – Archiv oder Register?..... 77

Grau, Bernhard:

Kooperation der staatlichen Archivverwaltungen mit der AdV bei der Archivierung von Geobasisdaten 79

Mobile GIS*Franke, Vincent:*

Vorstellung des neuen Leitfadens mobile GIS – Von der GNSS-basierten Datenerfassung zu Mobile Mapping, Version 4.0 81

Schmid, Julian:

Neue Wege bei den Vor-Ort-Kontrollen zur Agrarförderung..... 83

Baran, Ramona:

HydroVISH & KomVISH on the Web: Visualisierung und Nutzung von 3D-Massendaten aus Vermessung und Modellierung sowie amtlichen Geodaten im digitalen Wasserbau und der digitalen Kommunalverwaltung 86

Vorwort

Der vorliegende Tagungsband ist eine Sammlung der erweiterten Abstracts zur *Münchener Geoinformatik-Runde 2021*, die vom *Runden Tisch GIS e.V.* am 23. und 24. März 2021 als virtuelle Konferenz ausgerichtet wurde.

Das Vortragsprogramm der *Münchener GI-Runde 2021* war gegliedert in fünf Wissenschaftsforen mit eingeladenen Vorträgen aus der Forschung, drei Sitzungen mit Beiträgen aus der Praxis sowie dem *Innovationsforum* des *Runden Tisch GIS e.V.*, für das sich Unternehmen mit kreativen, neuartigen Anwendungen und Lösungen rund um die Verwendung von Geoinformation bewerben konnten.

Die Beiträge in diesem Tagungsband entstammen den Wissenschafts- und Praxisforen, für die hochkarätige Vortragende aus der Geoinformatik und benachbarten Disziplinen wie Bauinformatik und Elektrotechnik gewonnen werden konnten. Die erweiterten Abstracts geben einen Überblick über aktuelle Arbeiten und Projekte aus den Bereichen *Normung / Interoperabilität, GIS und BIM, Geo-Algorithmen für Big Spatial Geodata, Vernetzte Mobilität, Geodateninfrastrukturen sowie 3D-Geodaten und -Methoden*.

Traditionell war eines der Wissenschaftsforen den Finalisten des *Münchener Förderpreises Geoinformatik* gewidmet, so dass wir Ihnen auch Kurzfassungen der ausgewählten Abschlussarbeiten und Dissertationen in diesem Tagungsband präsentieren können.

Wir wünschen Ihnen eine erkenntnisreiche Lektüre und hoffen, dass Sie zahlreiche Anknüpfungspunkte und Ideen für Ihre eigenen Arbeiten in den Beiträgen zur *Münchener GI-Runde* finden.

Unser Dank gilt *Frau Dr. Gabriele Aumann* für die Koordination und Erstellung dieses Tagungsbands, allen Autorinnen und Autoren, der Jury des *Münchener Förderpreises Geoinformatik*, den Leitern der *Kompetenzpools* am *Runden Tisch GIS e.V.*, die für die Programmgestaltung der Praxisforen verantwortlich waren und nicht zuletzt den Sponsoren sowie allen Teilnehmerinnen und Teilnehmern der *Münchener GI-Runde 2021*.

München, im März 2021

Thomas H. Kolbe, Ralf Bill und Andreas Donaubaue

Vernetzte Mobilität – notwendige neue Perspektive(n) für die Automobil-Industrie

Lutz MORICH

AUDI AG · 85045 Ingolstadt ·
E-Mail: Lutz.Morich@audi.de

1 Einführung ins Thema

Mobilitätsanbieter müssen die Systemgrenzen ihres Denkens und Handelns zukünftig sehr viel weiter öffnen als das heute der Fall ist. Es reicht nicht aus, nur in den eigenen Produkten und Service-Angeboten, die unmittelbar mit dem Produkt-verknüpft sind, zu denken. Stattdessen muss eine ganzheitliche Betrachtung der Mobilität aus der Benutzersicht erfolgen. Methoden des Systems Engineering, Design Thinking, und der Customer Journey können dem Anbieter helfen, dieses Ökosystem zu beschreiben und die eigene Rolle zu finden. Diese Rolle kann übergreifend unter Einbindung weiterer Partner dargestellt werden oder einen Ausschnitt aus dem Ökosystem bedienen mit klaren Schnittstellen zu den umliegenden Partnern. [...] Die Modellierung eines ganzheitlichen Ökosystems baut auf allgemeingültigen Zielsetzungen auf, die sich aus den Szenarien der Mobilität ableiten lassen. Aus den dort verwendeten Schüsselfaktoren der politischen, ökonomischen, ökologischen, technologischen, legislativen und sozialen Dimensionen ergibt sich ein Ordnungsrahmen. Dieser definiert sowohl die Rahmenbedingungen und deren weitere Entwicklung als auch die Partner des Ökosystems. Deren gemeinsames Verständnis und Zusammenwirken ist letztlich erfolgsentscheidend für die Akzeptanz des Systems bei den Mobilitätsnutzern. Eine Steigerung der Lebensqualität unter Ökologischen Rahmenbedingungen ist eine wesentliche Zielsetzung für das Ökosystem. Das umfasst emissionsfreie Antriebe (beschleunigter Markteintritt durch regulatorische Maßnahmen), Car sharing und flexiblere Nutzung von ÖPNV, produktivere Nutzung von Reisezeiten durch automatisiertes Fahren, Staufreiheit durch Verkehrsleitsysteme und Smart Data-Anwendungen sowie Steigerung der Verkehrssicherheit bis hin zur Vision von null Verkehrstoten. (SUHM & JUNG 2017)

Im vorliegenden Beitrag werden Möglichkeiten und Notwendigkeiten für die Einbettung des „Einzel-Systems Automobil“ im „Gesamtsystem (System of Systems) der urbanen Umgebung“ am Beispiel der Forschungsaktivitäten zum digitalen Zwilling der Stadt Ingolstadt dargestellt.

2 Zusammenfassung des Vortragsinhalts

2.1 Historische Betrachtung

94 Jahre war das Automobil lediglich über den Fahrzeugführer mit seiner Umgebung vernetzt. Es funktionierte überwiegend mechanisch, später wurde es mit elektrischen Komponenten angereichert. Von der ersten Patentschrift von Gottlieb Daimler 1886 musste fast ein Jahrhundert vergehen, bis BMW mit dem On-board-Computer in der Formel 1 Fahrzeugdaten an die Box weitergeben konnte (HARTWIG 2020). Es dauerte weitere 16 Jahre, bis GM und Motorola mit OnStar 1996 den ersten Telematik-Dienst auf den Markt brachten (WALFORD 2014). Die Perspektive der Automobilindustrie beschränkte sich auf den Kunden im Fahrzeug. Der Einfluss auf das Umfeld spielte lediglich im Rahmen der Verkehrssicherheit und der Emissionen der Fahrzeuge eine Rolle. Maßnahmen dazu waren vor allem getrieben von regulatorischen Vorgaben wie Anschnallpflicht (1976) oder Verpflichtung zu Katalysatoren für Benziner (1984) oder aus dem Kundenwert Insassensicherheit wie z.B. bei der ABS-Breiteneinführung bei Ford (1985). Der Schutz anderer Verkehrsteilnehmer kam erst auf mit den Aktivitäten von Verbraucherschützern wie EURONCAP (2005-2010).

Über 100 Jahre wurde das Automobil als das Medium für individuelle Mobilität nur aus der Perspektive des Fahrers betrachtet. Vernetzungen fanden auch in den letzten Jahren vornehmlich im Inneren des Fahrzeugs statt, die Kommunikation mit dem Umfeld erfolgte ausschließlich über den Fahrer. Erst ab der Jahrhundertwende setzen sich mit der Einführung von GPS, Mobilfunk und Internet of things schrittweise die ersten Verbindungen mit der „Außenwelt“ durch.

Daher fällt es der automobilen Industrie nach wie vor schwer, einen Perspektiv-Wechsel vorzunehmen. Trotz zunehmender Umfeldvernetzung findet erst langsam die Betrachtung aus der Perspektive des „System of Systems Verkehrsumfeld“ mit dem Bestandteil des Systems „Fahrzeug“ als einem unter vielen statt. Umfangreiche Kenntnisse über das Umfeld, den Verkehr, dessen Phänomene und Lösungsansätze zu den Problemen sind innerhalb der Automobil-Unternehmen aktuell nicht ausreichend vorhanden.

Die vom BMVI geförderten Verbundprojekte SAVe: (2018-2020) und SAVeNoW (2021-2023) versuchen, das Verständnis und die Lösungsfähigkeit für automatisierte und vernetzte Mobilität über interdisziplinäre Forschung zu verbessern. Beide aufeinander aufbauende Projekte verfolgen den Ansatz der Systembetrachtung anhand der Methodik des Digitalen Zwillings.

2.2 Anwendung am Bsp. Ingolstadt – Projekt „SAVe“

Das Forschungsprojekt SAVe: (Verkehrs- und Funktions-Sicherheit im Automatisierten und Vernetzten Fahren) hat sich zum Ziel gesetzt, die Voraussetzungen zu schaffen, urbane Mobilität von morgen durch Simulation intelligent planen, entwickeln und steuern zu können. Hierzu wurde die Methodik und Architektur für den Aufbau eines virtuellen Testfeldes entwickelt. Es enthält alle statischen Elemente einer Stadt wie Straßen und Wege mit zugehöriger Verkehrslogik und -elementen, Gebäuden, Vegetation und Objekten im Straßenumfeld. Zusätzlich werden alle dynamischen Elemente wie Fahrzeuge, schwache Verkehrsteilnehmer wie Fußgänger und Zweiradfahrer abgebildet. Die Modellierung der dynamischen Elemente erfolgte durch mikroskopische und submikroskopische Verkehrs-Simulationen, die über statistischen Daten (Quelle-Ziel-Beziehungen) und Verkehrsdaten (Zählungen, Ampel-Informationen) sowie Floating Car Data von Erprobungsfahrzeugen der AUDI AG kalibriert werden. Abb.1 zeigt die Architektur des Gesamtmodells. Hierbei konnte die Lücke zwischen den bislang überwiegend getrennten Domänen der Geoinformatik und der Fahr- und Verkehrs-Simulation weiter verringert und dem virtuellen Umfeld der Fahrsimulation eine umfangreiche Toolwelt erschlossen werden.

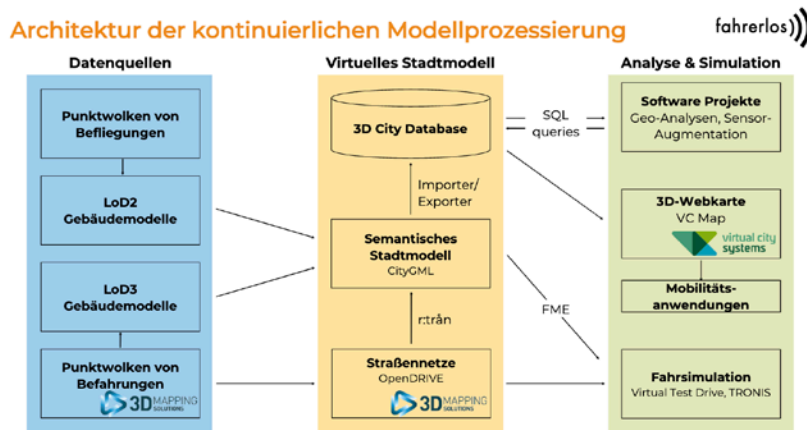


Abb. 1: Architektur des virtuellen Testfeldes

Über die Simulationsanwendungen konnten exemplarisch spezifische Fragestellungen der Verkehrseffizienz, - Sicherheit und der gesellschaftlichen Akzeptanz von automatisierten Fahrfunktionen beantwortet werden. Insbesondere über die genaue Betrachtung der komplexen Verkehrslogik der „Forschungskreuzung Hindenburg-Ringlerstraße“ (Abb. 2) wurde aufgezeigt, wie das orchestrierte Zusammenspiel unterschiedlicher Datenquellen und Simulationsmodelle zur Beantwortung verschiedener Fragestellungen des Gesamtsystems erfolgreich herangezogen werden kann (MORICH et al. 2020).

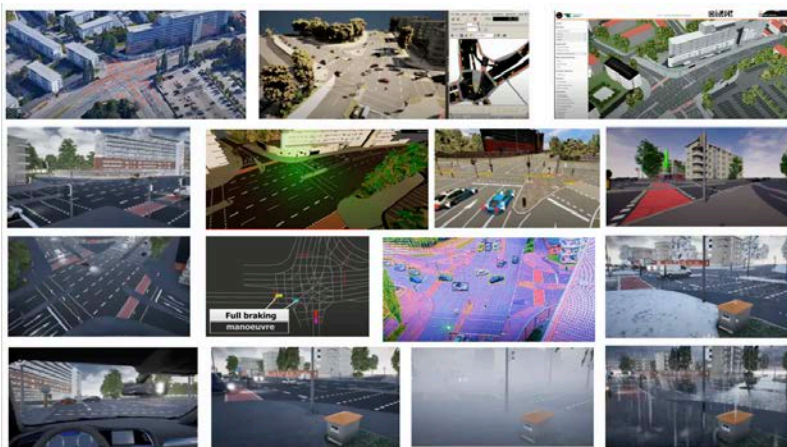


Abb. 2: Ergebnisse unterschiedlicher Simulations-Anwendungen der „Forschungskreuzung Hindenburg-Ringlerstraße“

2.3 Ausblick am Bsp. Ingolstadt – Projekt „SAVeNoW“

Das Verbundprojekt „SAVeNoW“ (Funktions- und Verkehrs-Sicherheit für Automatisierte und Vernetzte Mobilität – Nutzen für die Gesellschaft und ökologische Wirkung) baut auf den Ergebnissen von „SAVe“ auf. Die methodischen Grundlagen werden überführt in ein auf das Stadtgebiet skaliertes und betreibbares Gesamtmodell. Das Verbundprojekt

erforscht umfangreiche Facetten eines Digitalen Zwilling am Beispiel von Ingolstadt. Relevante statische Elemente (Straßen, Gebäude, Verkehrsinfrastruktur, Verkehrsregeln), dynamische Elemente (Verkehrsteilnehmer) und Rahmenbedingungen (Tageszeit, Wetter) werden inkl. umfangreicher Metadaten abgebildet. Mithilfe dieses Digitalen Zwilling werden Fragestellungen der Domänen Verkehrseffizienz, -sicherheit, -ökologie und gesellschaftlicher Akzeptanz analysiert und durch Lösungsszenarien mit zugehörigen Maßnahmen beantwortet. Abb. 3 zeigt das Modell in illustrierter und funktionaler Form. Das „Gesamtwerkzeug“ funktioniert in zwei ineinandergreifenden Regelkreisen aus Digitalem (reale Stadt) und Virtuellem Testfeld (virtuelles Stadtmodell), die die o.g. Domänen in Form von „Was-wäre-wenn“-Szenarien mit entsprechenden Maßnahmen simulieren und die Maßnahmenwirksamkeit optimieren. Basis ist eine IT-technisch umgesetzte Gesamtarchitektur aus Datenerfassung, Prozessierung, Modellbildung, Simulation, Auswertung, Anwendung und Optimierung. Das Gesamtmodell, das die Regelkreise in Funktion hält und weiterentwickelt, soll gesellschaftlich akzeptiert, betreibbar und wirtschaftlich sein.



Abb. 3: Digitaler Zwilling in illustrierter und funktionaler Darstellung

Über die Gesamtbetrachtung des „System of Systems“ schafft das Projekt die Möglichkeit, insbesondere für einen Automobilhersteller/Mobilitäts-Dienstleister verschiedene Perspektiven einzunehmen und ein Systemverständnis zu entwickeln. Das Modell hilft dabei, Lösungen für unternehmerische und technische Fragestellungen über die reine Mobilität hinaus zu erarbeiten und zu erproben.

Literatur

- SUHM, A. & JUNG, H. (2017), Vernetzte Fahrzeuge – Neue Geschäftsmodelle für Mobilität. In: Digital vernetzt. Transformation der Wertschöpfung, 225-238
- HARTWIG, M. (2020), Autonomes Fahren – Chancen für eine sichere, effiziente und nachhaltige Mobilität für alle?
- WALFORD, L. (2014) Definition of Connected Car – What is the connected car? Defined. In: auto connected car news. 22. APRIL 2014
- MORICH, L. et al. (2020), finissage@fahrerlos <https://fahrerlos-save.de/finissage/>
- CZWICK, C. & ANDERL, R. (2020), Cyber-physical twins - definition, conception and benefit

Tiefe neuronale Netze und multi-temporale Statistik zur Fernerkundung bewässerter Landwirtschaft

Lukas GRAF^{1,2}

¹Department für Geoinformatik - Z_GIS · Universität Salzburg · Schillerstraße 30 · 5020 Salzburg

²Vista Geowissenschaftliche Fernerkundung GmbH · Gabelsbergerstraße 51 · 80333 München

E-Mail: graf@vista-geo.de

1 Einführung ins Thema

Ein Großteil des globalen Wasserverbrauchs wird durch die Landwirtschaft verursacht und hierbei vor allem durch Bewässerung. WISSER ET AL. (2008) gehen davon aus, dass bis zu 70% der verfügbaren oberflächennahen Wasserressourcen für landwirtschaftliche Zwecke verwendet werden. Da die Landwirtschaft nicht nur eine stetig wachsende Nachfrage nach Lebensmitteln und Biomasse bedienen muss, sondern durch adverse Effekte des Biodiversitätsverlustes und des Klimawandels zusätzlich unter Druck gerät, ist eine nachhaltige Nutzung der verfügbaren Wasserressourcen unabdingbar.

Eine nachhaltige(re) Nutzung der Wasserressourcen erfordert eine möglichst genaue und raum-zeitlich hochaufgelöste Erfassung der Wasserverfügbarkeit und des Wasserbedarfs. In der Landwirtschaft stellen Beregnungskreise (*Irrigation Centre Pivots*) eine gängige Praxis der Bewässerung dar – insbesondere in ariden oder semi-ariden Gebieten mit größeren agrarisch genutzten Flächen wie den nordamerikanischen Great Plains oder Teilen Nordafrikas. Beregnungskreise bestehen aus einem beweglichen Arm, der mit Sprinklerdrüsen versehen ist, der um einen Fixpunkt rotiert. Daraus ergeben sich charakteristische kreisförmige Muster, die anhand von multispektralen Fernerkundungsdaten optisch gut erkennbar sind. Die satelliten-gestützte Kartierung der Beregnungskreise stellt daher ein traditionelles Anwendungsgebiet der optischen Fernerkundung dar, wie das - anhand von LANDSAT-TM Daten erstellte - „Nebraska Centre-Pivot Inventory“ mit über 50 000 kartierten Beregnungskreisen eindrucksvoll zeigt (CARLSON 1989).

Da die Kartierung von Bewässerungskreisen durch menschliche Experten eine langwierige Aufgabe darstellt, geht dieser Vortrag der Frage nach, ob tiefe neuronale Netze (*Deep Neural Networks*) für die automatisierte Kartierung von Beregnungskreisen geeignet sind. Dies wird anhand des U-NET Ansatzes (RONNEBERGER, FISCHER & BROX 2015) und Sentinel-2 Daten untersucht. U-NET wurde in der medizinischen Bildverarbeitung zur Segmentierung von Tumorgewebe aber auch in der Fernerkundung erfolgreich eingesetzt und zählt zur Klasse der *Convolutional Neural Networks*. U-NET verwendet neben der spektralen Information räumliche Beziehungen, weshalb der Ansatz für die Detektion und Lokalisation von Objekten geeignet ist. SARAIVA ET AL. (2020) setzten U-NET bereits erfolgreich für die Kartierung von Bewässerungskreisen in Brasilien ein: Daher wird der Ansatz auch in dieser Studie verfolgt.

Als Untersuchungsgebiet wurde ein Teil des spanischen Duero-Einzugsgebietes ausgewählt. Die Duero-Region ist bereits heute durch einen hohen Anteil an Bewässerungslandwirtschaft geprägt und wird nach gegenwärtigen Klimaprojektionen bis 2100 unter gehäuften Dürreperioden und Wasserknappheit leiden (GIL, GARRIDO & GÓMEZ-RAMOS 2011). Dies verleiht der Kartierung der Bewässerungslandwirtschaft in diesem Gebiet eine besondere Dringlichkeit. Neben der Kartierung der Beregnungskreise anhand monotemporaler Sentinel-2 Daten geht der Vortrag der Frage nach, ob multitemporale Statistiken des *Normalized Difference Water Index* (NDWI) eine Aussage über den Beregnungsstatus der Centre Pivots erlauben: Dies ist relevant, da nicht jeder Beregnungskreis ständig verwendet wird, und somit nicht zwangsweise zum landwirtschaftlichen Wasserverbrauch beiträgt.

2 Zusammenfassung des Vortragsinhalts

2.1 Methodik

Die Methodik ist schematisch in Abbildung 1 dargestellt und umfasst die Kartierung der Bewässerungskreise (a) mit dem U-NET Ansatz und die Ermittlung des Bewässerungsstatus (b).

Wie alle neuronalen Netzwerke muss das U-NET Modell trainiert werden. Hierfür wurden 615 Beregnungskreise im Untersuchungsgebiet von einem Experten anhand einer wolkenfreien, atmosphären-korrigierten Sentinel-2 Aufnahme vom 15. Juni 2018 kartiert. Die kartierten Bewässerungskreise wurden auf dem ESA Foodsecurity TEP (<https://foodsecurity-tep.net/>) veröffentlicht. Die Geometrien der Beregnungskreise zusammen mit der Sentinel-2 Aufnahme wurden als Trainingsdaten verwendet, wobei die Sentinel-2 Reflexionswerte den Modell-Input darstellen und die kartierten Beregnungskreise die Labels, die das Modell vorhersagen soll. Neben den Sentinel-2 Kanälen 2 (Blau), 3 (Grün), 4 (Rot) und 5 (NIR1) (U-NET SPECS) wurden die ersten drei Hauptkomponenten aller Sentinel-2 Kanäle als Input für U-NET (U-NET PCA) verwendet. Die räumliche Auflösung der Daten beträgt 10 m.

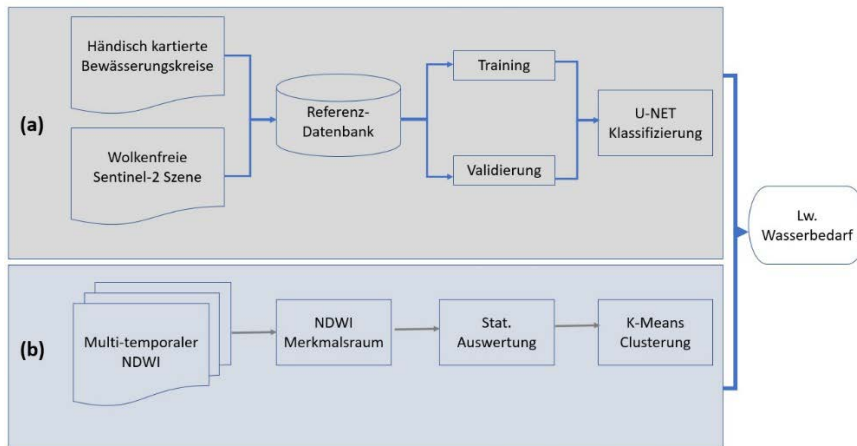


Abb. 1: Schematische Übersicht der Methodik. a) Kartierung der Bewässerungskreise mittels U-NET, b) Erfassung des Bewässerungsstatus mit k-Means Clustering.

Anhand von Erfahrungswerten aus der Literatur und systematischen Tests wurde die U-NET Architektur aufgebaut (Anzahl der neuronalen Schichten, Größe und Art der Filter, etc.) und 5/6 der vorhandenen Daten für das Training verwendet. Die genaue Spezifikation ist bei GRAF, BACH & TIEDE (2020) beschrieben. Nach dem Training wurde das U-NET Modell auf die bislang nicht verwendeten Daten angewandt und validiert. Hierfür wurden klassischen Metriken der binären Klassifikation verwendet (Abbildung 1a).

Für den Berechnungsstatus (Abbildung 1b) wurde der NDWI aller wolkenfreien Sentinel-2 Szenen des Untersuchungsgebietes zwischen Januar und Juni 2018 berechnet und in multi-temporalen statistische Kenngrößen (etwa Mittelwert, Minimum und Maximum über die Zeit pro Pixel) überführt. Einschränkend wurden nur Pixel der Berechnungskreise verwendet. Aus diesen wurde ein Merkmalsraum aufgespannt, der mittels k-Means geclustert wurde. Jeder Pixel eines Berechnungskreises konnte einem von drei Clustern zugeordnet werden, welche mittels Expertenwissens in drei semantische Klassen überführt wurden.

2.2 Ergebnisse

U-NET basierte Kartierung der Bewässerungskreise

Tabelle 1 zeigt die erzielten Validierungsergebnisse der beiden U-NET Modelle auf dem Teil des Eingangsdatensatzes, der nicht zum Trainieren verwendet wurde. Der F1-Score als generelles Klassifikationsergebnis zeigt, dass beide Modelle nur eine sehr geringe Klassifikationsgenauigkeit aufweisen und durch eine hohe Anzahl Falsch-Positiver wie Falsch-Negativer Klassifikationsergebnisse beeinträchtigt sind. Dies ist an den ebenfalls niedrigen Werten für Präzision und Recall erkennbar, welche die Fehleranfälligkeit eines binären Klassifikators beschreiben.

Tabelle 1: Klassifikationsgüte von U-NET SPECS (auf vier Sentinel-2 Kanälen trainiert) und von U-NET PCA (unter Verwendung der ersten drei Hauptkomponenten).

U-NET Modell	U-NET PCA	U-NET SPECS
Präzision (-)	0.12	0.52
Recall (-)	0.26	0.22
F1-Score (-)	0.16	0.31

Insgesamt weißt das U-NET SPECS Modell, welches auf den Sentinel-2 Kanälen 2, 3, 4, 5 trainiert wurde, eine höhere Klassifikationsgenauigkeit auf als U-NET PCA. Ein Blick auf die räumlichen Ergebnisse der Klassifikation mittels U-NET SPECS in Abbildung 2 offenbart analog zu den in Tabelle 1 aufgetragenen Metriken die Schwächen des Klassifikators: Nicht alle Bewässerungskreise wurden durch das Modell vollständig erkannt oder überhaupt nicht als Bewässerungskreis klassifiziert (Falsch-Negative). Andererseits finden sich in Abbildung 2 Flächen, die fälschlicherweise als Bewässerungskreis kartiert wurden (Falsch-Positive). Lediglich einzelne Kreise wurden vollständig richtig kartiert. Die in dieser Studie erzielten Ergebnisse sind deutlich weniger akkurat als Vergleichswerte von SARAIVA ET AL. (2020) in Brasilien. Als Erklärung kommt die hohe Komplexität der Landnutzung im Untersuchungsgebiet in Frage (im Vergleich zu Nord- und Südamerika). Auch eine ungenügende Parametrisierung und Architektur des U-NET Modells ist nicht auszuschließen: Das neuronale Netz könnte nicht tief genug sein und das rezeptive Feld zu klein. Das Verfahren eignet sich somit im Untersuchungsgebiet noch nicht für operationellen Einsatz. Weitere Forschung und das Testen anderer Algorithmen sind notwendig.

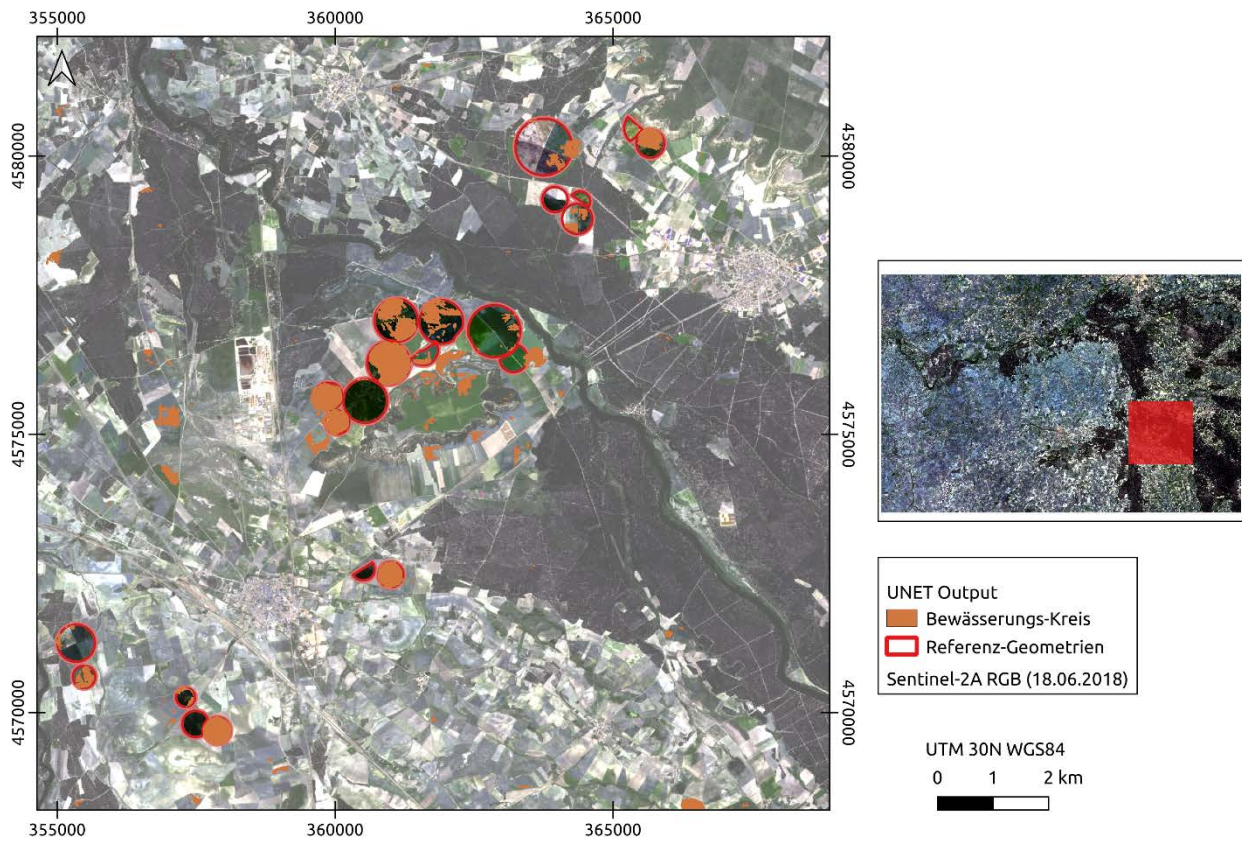


Abb. 2: Räumliche Ergebnisse der Kartierung der Bewässerungskreise mit U-NET SPECS auf einem Teil des Studiengebietes, der nicht zum Training verwendet wurde. Die von U-NET klassifizierte Flächen sind orange eingefärbt; die Referenz-Geometrien der Kreise sind rot umrandet. Im Hintergrund ist die verwendete Sentinel-2 Szene als Echtfarbbilderkarte erkennbar.

NDWI basierte Ermittlung des Bewässerungsstatus

Die Clusterung des multi-temporalen NDWI Merkmalsraumes lieferte räumlich wie zeitlich plausible Ergebnisse. Die drei ermittelten Cluster sind gut voneinander trennbar und lassen sich eindeutigen semantischen Klassen zuweisen, welche in Tabelle 2 beschrieben sind.

Tabelle 2: Beschreibung der drei NDWI Cluster, die innerhalb der Bewässerungskreise gefunden werden konnten.

Cluster	Merkmale
Cluster 1	Nicht permanent beregnete/ grüne Vegetation innerhalb der Bewässerungskreise. Dieser Cluster ist durch niedrige NDWI Werte zu Beginn des Jahres (Winter) gekennzeichnet, die zum Sommer ansteigen, was auf den Einsatz von Bewässerung in den Sommermonaten hindeutet.
Cluster 2	Ausschließlich trockene Vegetation, Brache und offener Boden innerhalb der Bewässerungskreise. Dieser Cluster ist durch fortwährend sehr niedrige NDWI Werte gekennzeichnet. Bewässerungskreise, die eine Vielzahl an Pixeln dieses Clusters enthalten, sind nicht in Betrieb.
Cluster 3	Durchgängig beregnete/ grüne Vegetation innerhalb der Bewässerungskreise. Die NDWI-Werte sind konstant sehr hoch, was aufgrund der klimatischen Gegebenheiten am Duero nur durch permanente Bewässerung zu erzielen ist.

Mittels dieses Ansatzes konnte gezeigt werden, dass bis zu 31% der Bewässerungskreispixel im Untersuchungszeitraum (Jan-Jun 2018) nicht beregnet wurden und somit nicht zum landwirtschaftlichen Wasserbedarf beitragen. Somit liefert das Verfahren wichtige Erkenntnisse über landwirtschaftliche Bewässerungspraxis, die jedoch in weiterer Forschung genauer aufgeschlüsselt und untersucht werden sollte, als es hier der Fall ist.

Literatur

- CARLSON, M. P. (1989), The Nebraska Center-Pivot Inventory: An example of operational satellite remote sensing on a long-term basis. In: *Photogramm. Eng. Remote Sens.*, 55: 587–590.
- GIL, M., GARRIDO, A. & GÓMEZ-RAMOS, A. (2011), Economic analysis of drought risk: An application for irrigated agriculture in Spain. In: *Agricultural water management* 98 (5): 823–833. <https://doi.org/10.1016/j.agwat.2010.12.008>.
- GRAF, L., BACH, H. & TIEDE, D. (2020), Semantic Segmentation of Sentinel-2 Imagery for Mapping Irrigation Center Pivots. In: *Remote Sensing* 12(23), 3937, 1-19. <https://doi.org/10.3390/rs12233937>.
- RONNEBERGER, O., FISCHER, P. & BROX, TH. (2015), U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, herausgegeben von NAVAB, N., HORNEGGER, J., WELLS, W.M. & FRANGI A.F., 234–241, Springer International Publishing, Cham.
- SARAIVA, M., PROTAS, E., SALGADO, M. & SOUZA, C. (2020), Automatic Mapping of Center Pivot Irrigation Systems from Satellite Images Using Deep Learning. In: *Remote Sensing* 12 (558), 1–14. <https://doi.org/10.3390/rs12030558>.
- WISSER, D., FROLKING, S., DOUGLAS, E.M., FEKETE, B.M., VÖRÖSMARTY, C.J. & SCHUMANN, A.H. (2008) Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets. In: *Geophysical Research Letters* 35 (24). <https://doi.org/10.1029/2008GL035296>.

Entwicklung einer Crowdsourcing-Anwendung zur automatisierten Detektion von Straßenschäden mit Hilfe von Methoden der Künstlichen Intelligenz als Webservice

Oliver LEHMKUHL

Universität Osnabrück · Neuer Graben 29 · 49074 Osnabrück

E-Mail: olehmkuhl@uni-osnabrueck.de

1 Einführung in die Thematik

Straßen sind eine der wichtigsten Bestandteile der heutigen Infrastruktur und bilden die primär genutzten Transportwege für den Personen- und Güterverkehr. Die Wahrung der Verkehrssicherheit dieser Transportwege stellt somit ein Hauptanliegen dar. Mit dem Erhalt eines guten Fahrbahnzustandes können Unfälle und folglich Todesopfer vermieden werden. Detaillierte und aktuelle Informationen über den Fahrbahnzustand bilden die maßgebliche Grundlage für eine sachgemäße Planung der Fahrbahnсанierung. Um eine Aufrechterhaltung des Straßennetzes eines gesamten Landes zu gewährleisten, detektieren und klassifizieren zahlreiche Unternehmen die Fahrbahnbeschädigungen zumeist mit aufwändigen, manuellen Vor-Ort-Inspektionen. Aufgrund dieser Vorgehensweise verzögert sich die Instandsetzung und Reparatur. Infolgedessen kann sich der Grad der Mängel zuspitzen und der dafür benötigte Sanierungsaufwand, in Form von Zeit sowie Kosten, wächst. Mit Hilfe von Deep-Learning-Techniken können neuronale Netzwerke trainiert werden, die genau diese zeit- und kostenintensiven Auswertungen übernehmen (SESSELMANN et al. 2020, S. 18). Insbesondere hat sich die Option etabliert, eine Auswertung von Fahrbahnbeschädigungen in Kombination mit Smartphone-Aufnahmen umzusetzen.

In diesem Beitrag wird mit einer hybriden Applikation für die gängigen Betriebssysteme iOS und Android der Erhebungsprozess der Smartphone-Aufnahmen auf die Mitbürgerinnen und Mitbürger ausgelagert, wodurch die entstandenen Beschädigungen frühzeitig lokalisiert und somit kostengünstiger in Stand gesetzt werden können. Die auf den Aufnahmen basierenden Auswertungen zur Schadensdetektion werden mit Methoden der Künstlichen Intelligenz (kurz KI) umgesetzt. Die zu entwickelnde KI soll dabei in einem Webservice agieren und die erhobenen Aufnahmen auf bestehende Beschädigungen untersuchen.

2 Zusammenfassung des Vortrag inhalts

2.1 Einführung in die künstlichen neuronalen Netzwerke

(Künstliche) Neuronale Netzwerke (engl. artificial neural networks - kurz ANN) beziehen sich in ihrer Funktionalität auf das Neuronennetz des menschlichen Gehirns. Sie bestehen aus einer Vielzahl parallel arbeitender Einheiten, den sogenannten Neuronen. Übereinander angeordnete Neuronen werden als Schicht bezeichnet. Gegenüber dem ANN stellt ein faltendes neuronales Netzwerk (engl. Convolutional Neural Network – kurz CNN) eine spezielle Art von neuronalem Netz dar, welches Faltungsoperationen (engl. Convolution) anstelle von Neuronenlayern in mindestens einer ihrer Schichten verwendet. Im Wesentlichen besteht ein CNN aus Convolutional- und Aggregations-Schichten (engl. Pooling Layer). Diese können paarweise verwendet und beliebig oft wiederholt werden. Am Ende dieser Wiederholungen befinden sich eine oder mehrere Schichten vollständig verbundener Neuronen (engl. Dense oder Fully Connected Layer) (HU et al. 2015, S. 2).

2.2 Verwendete Datengrundlage

Um einen möglichst erfolgreichen und effektiven Lernprozess zu erzielen, benötigt ein neuronales Netzwerk im großen Maße Trainingsdaten. Gegenwärtig besteht kein einheitlicher Datensatz, der für die Detektion von verschiedenen Straßenschäden auf Bilddaten frei zur Verfügung steht. Folglich wurde die Datengrundlage eigenständig aus diversen Quellen zusammengestellt. Der überwiegende Anteil der Trainingsdaten wurde über eine Bildersuche auf Google mit dem Schlagwort Pavement Distress am 19.11.2019 erhoben. Darüber hinaus wurden auch Bilddaten von bereits bestehenden umfangreichen Datensätzen genutzt. Dazu zählen das German Asphalt Pavement Distress Dataset (kurz GAPS) der Technischen Universität Ilmenau (STRICKER et al. 2019), das Road Damage Dataset 2018 (MAEDA et al. 2018) und der Datensatz CRACK500 (ZHANG et al. 2016). Bei der Zusammenstellung der Daten sowie bei der Unterteilung der Schadensklassen wurde sich an den GAPS-Daten orientiert. Der entstandene Datensatz involviert die vier verschiedenen Schadens- und Beeinträchtigungstypen aus Abb. 1 sowie die Klasse „Intakte Straße“. Jede Trainingsaufnahme umfasst dabei exakt eine Klasse.

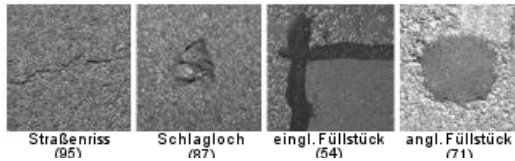


Abb.1:
Unterteilung der Schadensklassen mit Angabe der Bilddatenmenge, abgeändert von (EISENBACH et al. 2017, S. 4, Abb. 4)

Ergänzend zu der beschriebenen Differenzierung der Straßenbeschädigungen soll eine KI den Schaden auf den Aufnahmen verorten können. Die Datengrundlage muss dafür neben den Bildinformationen noch zusätzliche Annotations (Anmerkungen) involvieren. Konkret handelt es sich bei den Annotations um Ground-Truth Informationen, die angeben, an welchen Positionen des Bildes die Beschädigungen zu verorten sind (vgl. Abb. 2). Die Ground-Truth wurde manuell in Form einer binären Maske für jede einzelne Aufnahme der Datengrundlage selbstständig erhoben.

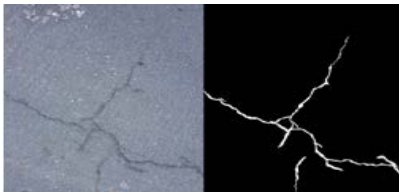


Abb.2:
Erweiterter Trainingsdatensatz, abgeänderte Aufnahmen aus dem CRACK500-Datensatz (ZHANG et al. 2016)

2.3 Entwicklung einer Künstlichen Intelligenz

Auf Basis einer eigenen CNN-Architektur erfolgt die Klassifikation der Schadenstypen. Es wurde von bekannten Architekturen abgesehen, da diese häufig zu komplex für diese Form der Aufgabenstellung sind. Die genutzte Architektur setzt sich aus insgesamt fünf Filter- (engl. Convolutional Layer) und jeweils fünf Aggregations-Schichten (engl. MaxPooling Layer) zusammen, die sich abwechselnd wiederholen. Nach jeder Wiederholung von Filter- und Aggregations-Schicht folgt eine Normalisierungsschicht (engl. Batch Normalization Layer). Unter der Verwendung einer ReLu- Aktivierungsfunktion (engl. Rectified Linear Unit) erfolgt in den Convolutional Layern die Skalierung der Eingabesignale in einen Ausgangswert (CHOLLET 2015). Um einer Überanpassung des Netzes an die Trainingsdaten entgegenzuwirken, wurden zusätzlich Dropouts eingefügt. Durch einen Dropout werden während der Trainingsphase eine bestimmte Menge zufälliger Neuronen ignoriert, sodass ein reduziertes Netz übrig bleibt (LEE & KWON 2017, S. 4). Die letzten Netzwerkschichten bestehen aus insgesamt drei neuronalen Layern (bzw. Dense Layer). Die Ausgangsschicht verwendet die Softmax-Aktivierungsfunktion (normalisierte Exponentialfunktion), welche für die fünf Zielklassen die Wahrscheinlichkeitsverteilung berechnet (CHOLLET 2015). Mit der CNN-Architektur konnte, in Kombination mit den gewählten Hyperparametern, eine Klassifikationsgenauigkeit von 80% auf den Evaluierungsdaten erzielt werden.

Um eine Objektdetektion umzusetzen, müssen weitere neuronale Netzwerke konstruiert und trainiert werden. Unter der Verwendung mehrerer U-Net-Architekturen findet die semantische Segmentierung der unterschiedlichen Schadenstypen statt (RONNEBERGER et al. 2015). Aus dem Umstand, dass ein U-Net auf eine Beschädigung trainiert wird und folglich für nur diese Beschädigung die höchsten Segmentierungsgenauigkeiten erzielt, begründet sich die Notwendigkeit einer vorgeschalteten Klassifikation. Anlässlich der Segmentierung verschiedener Beschädigungen besteht zwar die Option, ein einziges U-Net auf eine multiple Klassensegmentierung anzupassen, allerdings sind die erzielbaren Genauigkeiten nicht ausreichend. Diese Problematik entsteht insbesondere durch das hochgradige Klassenungleichgewicht zwischen den Pixeln, die einen Schaden markieren und denen, die den Hintergrund angeben (vgl. Abb. 2). Trotz der Einschränkung wurde, aufgrund der geringen Anzahl der lernbaren Parameter, eine U-Net-Architektur verwendet (1,95 Mio. Parameter) (RONNEBERGER et al. 2015). Ebenso ist das U-Net bei Segmentierungsaufgaben effizient und besitzt zeitgleich ein hohes Maß an Segmentierungsgenauigkeit. Auf Basis der U-Net-Architektur erfolgt der Trainingsprozess von insgesamt vier Modellen. Im Trainingsprozess der U-Net-Architekturen konnten Segmentierungsgenauigkeiten zwischen 81,6% und 90,5% auf den Evaluierungsdaten erzielt werden (abhängig vom Schadenstyp).

2.4 Einrichtung eines Webservices

Zielsetzung dieser Arbeit ist unter anderem die performante Implementierung einer KI in einen Webservice, mit dem eine automatisierte Detektion von verschiedenen Straßenbeschädigungen auf Bilddaten möglich ist. Dies wurde mit dem Flask Mikroframework umgesetzt. Der Flask-Webservice erhält die Aufnahmen von der hybriden App (vgl. Abschnitt 2.5) und nimmt mit dem Klassifikationsmodell eine Prognose der möglichen Schadensklasse vor (vgl. Abschnitt 2.3). Die Zielklasse mit der höchsten Wahrscheinlichkeit wird als Klasse der Vorhersage angenommen. Wenn das Netzwerk eine Beschädigung erkennt, wird die Aufnahme weiter an das entsprechende U-Net gesendet, welches auf die erkannte Beschä-

digung trainiert ist. Das jeweilige U-Net nimmt anschließend eine Segmentierung vor und erzeugt eine pixelgenaue Segmentierungsmaske. Der Webservice ist mit dem selbstgehosteten Parse mobile-backend-as-a-Service zur Datenspeicherung verknüpft. Um die Skalierbarkeit des Systems zu wahren und Anpassungen an der entstandenen Architektur durchführen zu können, wurde eine Middleware zwischen dem Webservice und der Parse-Datenbank integriert.

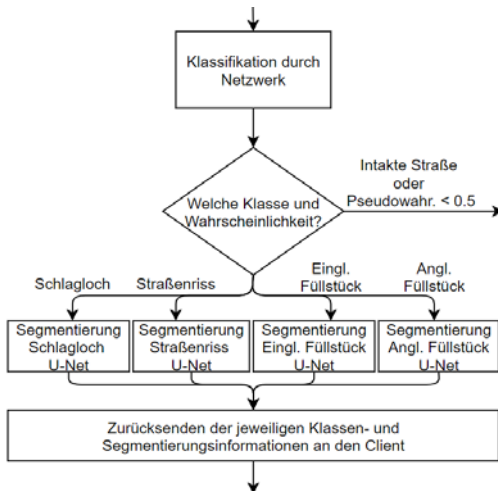


Abb. 3: Integration der KI in den Ablauf des Webservices, eigene Darstellung

2.5 Entwicklung einer Crowdsourcing-Anwendung

Wie zuvor beschrieben, soll die Datenerhebung den Crowdsourcing-Ansatz verfolgen. Um eine möglichst umfangreiche Zielgruppe zu erreichen, wird eine hybride Anwendung entwickelt. Folglich kann aus der entwickelten Anwendung auch für die gängigen iOS- und Android-Betriebssystem eine eigenständige App abgeleitet werden.

Bei der Umsetzung der Crowdsourcing-Anwendung wurden unterschiedliche Frameworks, Bibliotheken und APIs verwendet. Grundlegend basiert die Applikation auf dem Ionic-Framework und nutzt die Leaflet API zur Kartenvisualisierung. Eine Klassifizierung sowie Segmentierung der gemeldeten Beschädigungen erfolgt über einen Verbindungsaufwurf zur Middleware und dem damit verbundenen Webservice.

Die hybride Anwendung stellt ein Modul dar, das in dem Web-Framework HQ der IP SYSCON GmbH integriert wurde und somit den bestehenden Funktionsumfang erweitert. Das Modul namens Straßenschäden besteht dabei aus den zwei Hauptseiten „Kartendarstellung“ und „Formular-Straßenschaden“ (vgl. Abb. 4).

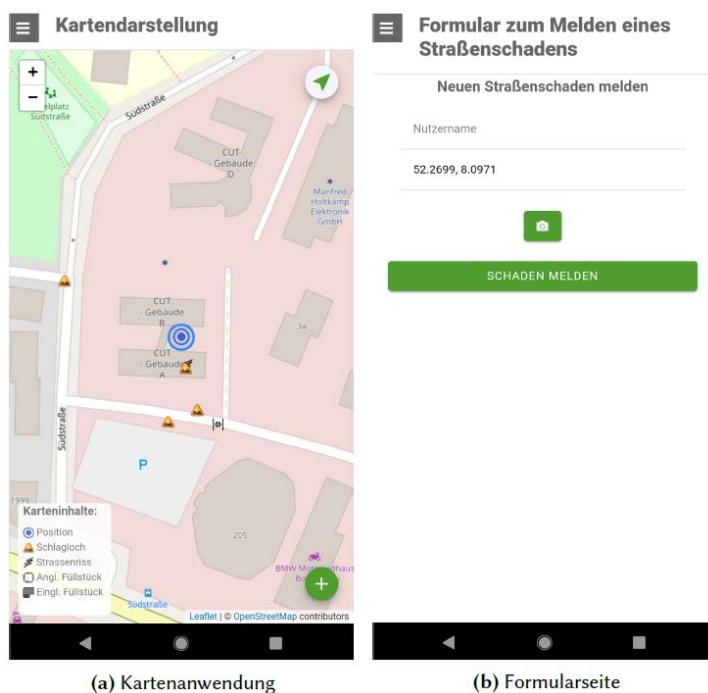


Abb. 4: Screenshots des Prototyps, eigene Darstellung

Zur Untersuchung der Einsetzbarkeit der Applikation wurde ein Praxistest im östlichen Stadtgebiet von Osnabrück durchgeführt. Bei der Teststrecke handelt es sich um Verkehrswege, die ein vermeintlich großes Teilspektrum der realen deutschen Verkehrswege abbilden. Auf den Verkehrswegen wurden die verschiedenen Beschädigungen sowie Füllstücke mit einem Nokia 5.1 und einem Galaxy Tab S4 (Modell SM-T835) an den Webservice übermittelt und analysiert. Im Anschluss wurden die Ergebnisse vergleichend gegenübergestellt. Diese zugrundeliegende Vorgehensweise gibt erste Eindrücke darüber, wie ausschlaggebend die Kameraqualität für die Erhebung mit dem Prototypen ist.

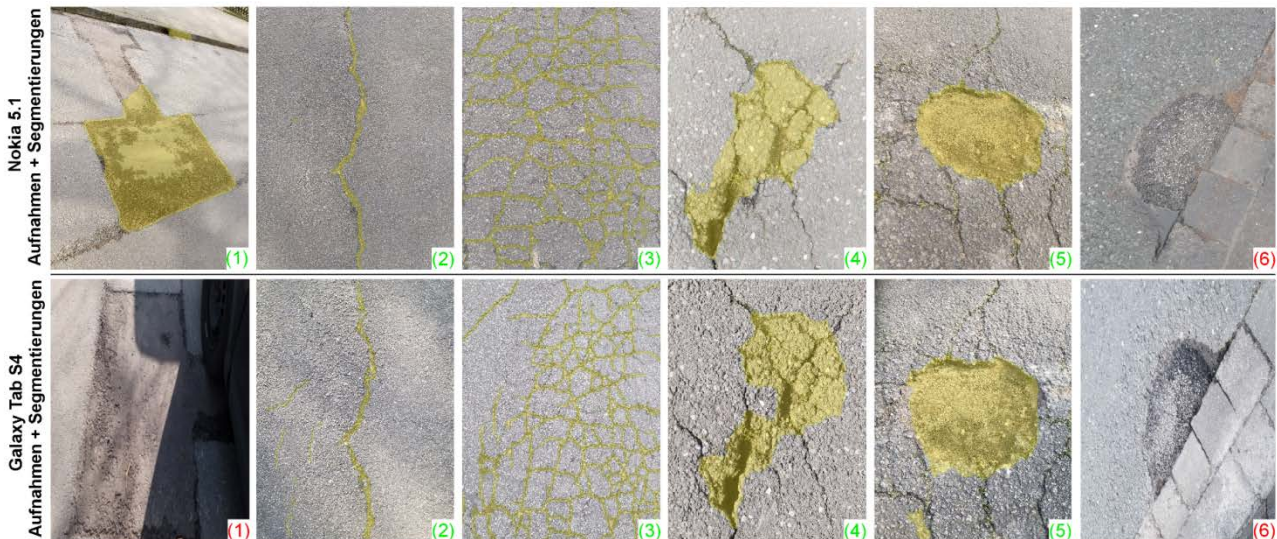


Abb. 5: Vergleich der Segmentierungsgenauigkeiten. Eine rote Nummerierung gibt eine Fehlklassifikation an, eigene Darstellung

2.6 Fazit und Ausblick

Im Zuge dieser Ausarbeitung konnten die manuellen Vorgehensweisen zur Erhebung der Fahrbahninformatoren automatisiert von einer Künstlichen Intelligenz in einem Webservice übernommen werden. Mit dem Prototyp auf iOS und Android besitzen die Mitbürgerinnen und Mitbürger die Möglichkeit Straßenbeschädigungen zu erheben und somit einen Beitrag für die Planung der Fahrbahnsanierung zu leisten. Die zugrundeliegende Technologie basiert auf den künstlichen neuronalen Netzen, die neben einer Vorhersage des Schadentyps auch eine Segmentierung der Beschädigung vornehmen.

Für den in dieser Arbeit entwickelten Webservice wurde auf das Flask Mikroframework zurückgegriffen. Der größte Vorteil bei der Verwendung einer KI als Webservice besteht darin, dass diese auch von anderweitigen Anwendungen über das Internet erreichbar ist und dementsprechend einen Mehrwert bieten kann. Der Praxistest mit zwei verschiedenen Android-Geräten im östlichen Stadtgebiet von Osnabrück überprüfte zum einen die Funktionalität des Prototyps und zum anderen die Segmentierungsgenauigkeit in Abhängigkeit von der integrierten Hardware. Es zeigt sich, dass der Auflösungsunterschied von drei Megapixeln nur marginale Auswirkung auf die Qualität der Segmentierung hat. Um dies jedoch tatsächlich zu belegen, müssen Testläufe mit weiteren Smartphones durchgeführt werden.

Während der Testphase kam es zu Fehlklassifikationen, die auf den geringen Umfang und somit auch der mangelnden Varianz des verwendeten Trainingsdatensatzes zurückzuführen sind. Trotz der erzielten hohen Zuweisungsgenauigkeit des Klassifikators können Fehleinschätzungen auftreten. Um die Genauigkeit des Klassifikators sowie der Segmentierungsnetzwerke weiter zu verbessern und künftig Fehlklassifikationen auszuschließen, können die übersendeten Aufnahmen in einen weiteren Trainingsprozess einfließen.

Durch die permanente Sicherung der Neuronennetze im Arbeitsspeicher des Servers, die zusätzliche Speicherbelastung der übermittelten Aufnahme sowie die Begrenzung von zwei Gigabyte Arbeitsspeicher ist die Nutzung von komplexeren Netzstrukturen gegenwärtig ausgeschlossen. Mit dem Praxistest wurde jedoch belegt, dass zur Detektion von Straßendefekten eine Netzstruktur von 0,2 Mio. Parametern zur Klassifikation sowie 1,95 Mio. Parametern zur Segmentierung ausreichen und folglich kein umfangreicheres Netzwerk erforderlich ist. Ebenso wurde mit dem Test bestätigt, dass durch Crowdsourcing und Smartphone-Aufnahmen eine sinnvolle Symbiose geschaffen wurde. Zukünftig kann mit dem Prototypen ein kostengünstiger und aktueller Beitrag zur Planung der Straßensanierung geleistet werden, ohne auf hochauflösende und kostenintensive Laserscanning-Daten zurückgreifen zu müssen.

Neben der Erweiterung des Netzwerkes um zusätzliche Beschädigungen (z.B. Defekte der Fahrbahnmarkierungen) zählt auch die Einstufung des Grades einer Beschädigung als nützliche Erweiterung. Um eine Priorisierung vornehmen zu können, müsste der Klassifikator mit weiteren Informationen über die Beschädigung (z.B. Tiefe, Breite und Höhe) trainiert werden, um den Schweregrad einer Beschädigung einstufen zu können. Damit das Netzwerk die Dimensionen der

Beschädigung zutreffend bestimmen kann, müssen die vom Nutzer getätigten Aufnahmen stets in gleichem Winkel und Abstand erfasst werden. Dieser Umstand ist als problematisch zu deklarieren.

Literatur

- M. SESSELMANN, R. STRICKER, T. NABER, & S. SCHELLE (2020), “Automatische Detektion und objektscharfe Georeferenzierung von Fahrbahnschäden aus Bilddaten eines Mobile-Mapping-System mithilfe von Deep Learning”. In: GIS.Science - Die Zeitschrift für Geoinformatik. ISSN: 1869-9391. url: www.gisPoint.de.
- H. MAEDA, Y. SEKIMOTO, T. SETO, T. KASHIYAMA, & H. OMATA (2018), “Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images”. In: Computer-Aided Civil and Infrastructure Engineering.
- M. EISENBACH, R. STRICKER, D. SEICHTER, K. AMENDE, K. DEBES, M. SESSELMANN, D. EBERSBACH, U. STOECKERT, & H.-M. GROSS (2017), “How to Get Pavement Distress Detection Ready for Deep Learning? A Systematic Approach”. In: International Joint Conference on Neural Networks (IJCNN). S. 2039–2047.
- L. ZHANG, F. YANG, Y. D. ZHANG, & Y. J. ZHU (2016), “Road crack detection using deep convolutional neural network”. In: IEEE International Conference on Image Processing (ICIP). S. 3708-3712.
- R. STRICKER, M. EISENBACH, M. SESSELMANN, K. DEBES, & H.-M. GROSS (2019), “Improving Visual Road Condition Assessment by Extensive Experiments on the Extended GAPs Dataset”. In: International Joint Conference on Neural Networks (IJCNN). S. 1-8.
- W. HU, Y. HUANG, L. WEI, F. ZHANG, & H. LI (2015), “Deep Convolutional Neural Networks for Hyperspectral Image Classification”. In: Journal of Sensors. 2015(2): 1–12.
- F. CHOLLET et al. (2015), Keras: The Python Deep Learning library. <https://keras.io/activations/>, Zugriff 01/20.
- H. LEE & H. KWON (2017), “Going Deeper With Contextual CNN for Hyperspectral Image Classification”. In: IEEE Transactions on Image Processing Nr. 26. S. 4843-4855.
- O. RONNEBERGER, P. FISCHER, & T. BROX (2015), U-Net: Convolutional Networks for Biomedical Image Segmentation.

Zusammenhänge zwischen Bikesharing und ÖPNV - eine geostatistische Analyse in Köln

Katja SCHIMOHR

Fakultät Raumplanung · Technische Universität Dortmund · August-Schmidt-Str. 10 · 44227 Dortmund
E-Mail: katja.schimohr@tu-dortmund.de

1 Einführung ins Thema

Bikesharing stellt eine verhältnismäßig neue Mobilitätsoption dar, die größtenteils im urbanen Kontext angeboten wird und die bereits bestehenden Verkehrssysteme ergänzt. Dabei kann zwischen stationsbasierten und sogenannten free-floating Systemen unterschieden werden. Letztere Organisationsform ermöglicht das flexible Ausleihen und Abstellen der Räder innerhalb des gesamten Bediengebietes. Die Nutzung solcher Bikesharingsysteme ist im europäischen Raum bisher nur wenig erforscht. Dabei bietet die Erfassung der genauen Standorte aller Leihräder, die im Kölner Bikesharingsystem von Nextbike öffentlich mittels einer API bereitgestellt werden, die Möglichkeit, die Nutzung des Systems mit einer hohen räumlichen und zeitlichen Genauigkeit zu analysieren. Somit können Erkenntnisse über das Zusammenwirken von Stadtstruktur und Bikesharing gewonnen werden, mittels derer Bikesharingsysteme optimal in das städtische Verkehrsnetz integriert werden können. Ein besonderer Fokus dieser Arbeit liegt in der Untersuchung der Interaktionen zwischen Bikesharing und ÖPNV, da sich Bikesharing besonders als Ergänzung zum ÖPNV anbietet, um Lücken im Netz zu schließen oder die letzte Meile zurückzulegen. Zur Untersuchung dieser Fragestellung wird eine Analyse der Nutzungsdaten des Bikesharingsystems in Köln im Zusammenhang mit vielfältigen Geodaten zur Stadtstruktur und zum ÖPNV-Angebot durchgeführt.

2 Zusammenfassung des Vortragsinhalts

2.1 Methodik

Es wird ein fünfwöchiger Untersuchungszeitraum vom 30.09. bis zum 04.11.2019 betrachtet, der eine zweiwöchige Baustelle im U-Bahn-Netz umfasst, durch die weitreichende Störungen verursacht werden. Während dieses Zeitraums werden die Standorte aller 1450 Leihräder in 15-min Intervallen aus der API mithilfe eines Excel VBA-Skripts automatisiert abgerufen und die Koordinaten inklusive Zeitstempel und Radnummer gespeichert. Dabei wird das Gebiet der sogenannten Flexzone des Bikesharing-Anbieters Nextbike in Köln betrachtet, welches auf 84 km² die zentralen Bereiche der Stadt umfasst. Die weitere Analyse wird in R sowie in ArcGIS Pro durchgeführt. Aus den 3300 erhobenen Datensätzen können alle getätigten Fahrten berechnet werden. Nach Ausschluss von Fahrten, deren Länge unterhalb 100 m liegt, umfasst der finale Datensatz 76 859 Fahrten.

Deskriptive Statistiken, die aus der räumlichen und zeitlichen Verteilung der Fahrten berechnet werden, geben einen ersten Eindruck über die Nutzung des Bikesharingsystems. Auch können Fahrten, die in die Umgebung von ÖPNV-Haltestellen führen, extrahiert und gesondert betrachtet werden.

Die Standorte und Fahrten werden in ArcGIS mit unterschiedlichen räumlichen Faktoren in Verbindung gebracht, wie etwa Flächennutzung, Anzahl von Points of Interest unterschiedlicher Kategorien, ÖPNV-Stationen sowie sozioökonomische Daten. Dazu wird das Untersuchungsgebiet in Zellen von 100 x 100m Größe aufgeteilt und die Anzahl der Fahrten, die dort starten, bestimmt. Außerdem werden für jede Zelle Werte für die unterschiedlichen räumlichen Indikatoren berechnet. Für die Flächennutzungskategorien werden die Anteile an der Gesamtfläche einer Zelle bestimmt. Die Anzahl der Points of Interest wird innerhalb 3 x 3 Zellen ermittelt, um auch den Einfluss naheliegender Angebote mit einzubeziehen. Der Abstand jeder Zelle zu nächstgelegenen Universitätsgebäuden, Bus- und U-Bahn-Stationen wird bestimmt. Dabei werden die von der Baustelle betroffenen Haltestellen zusätzlich separat in den Kategorien „gesperrte Haltestelle“, „Haltestelle mit Umleitungen“, „letzte Haltestelle auf dem normalen Linienweg“ sowie „Haltestelle mit Ersatzverkehr“ betrachtet. Als demographische Variablen werden die Einwohnerzahl jeder Rasterzelle sowie die Anteile unterschiedlicher Altersgruppen innerhalb 3 x 3 Zellen in die Untersuchung aufgenommen.

Als zeitliche Variablen werden der Wochentag, die Tageszeit, der Zeitraum der Baustelle und die Vorlesungszeit der Universität als Dummy-Indikatoren sowie die Wetterindikatoren Bedeckungsgrad, Lufttemperatur, Niederschlagsmenge und Sonnenscheindauer betrachtet.

Mittels bivariater Korrelationsanalysen nach Spearman werden die Fahrtenzahlen je Zelle mit den räumlichen Einflussfaktoren in Verbindung gebracht und solche Indikatoren ausgewählt, die möglichst stark mit der Anzahl an Fahrten und zugleich möglichst wenig mit anderen Indikatoren korrelieren. Die ausgewählten Variablen werden in einer Regressionsanalyse auf einen Einfluss auf die Anzahl der Fahrten untersucht. Dabei wird das Verfahren der negativen Binomial-

Regression angewandt, da die abhängige Variable nur ganzzahlige positive Werte inklusive der Null annehmen kann und Überdispersion vorliegt, d.h. die Varianz liegt deutlich über dem arithmetischen Mittel.

Zusätzlich wird dieselbe Vorgehensweise für die unterschiedlichen zeitlichen Variablen und Wetterindikatoren durchgeführt. Die Anzahl der Fahrten, die innerhalb jeweils einer Stunde starteten, wird ermittelt und als abhängige Variable in einer zweiten Korrelationsanalyse und negativen Binomial-Regressionsanalyse betrachtet.

2.2 Ergebnisse

Jedes der Räder wurde innerhalb des Untersuchungszeitraums mindestens einmal bewegt, 69,24-Mal im Durchschnitt. Der mittlere Abstand zwischen Start- und Endpunkt einer Fahrt liegt bei 1740,1 m. Die zeitliche Verteilung der Fahrten nach Stunde und Wochentag wird in Abbildung 1 dargestellt und deutet auf einen hohen Anteil an Arbeits- und Ausbildungswegen hin, da sich die Nutzungsspitzen mit den allgemeinen Zeiten der Rush-Hour im Berufsverkehr überschneiden.

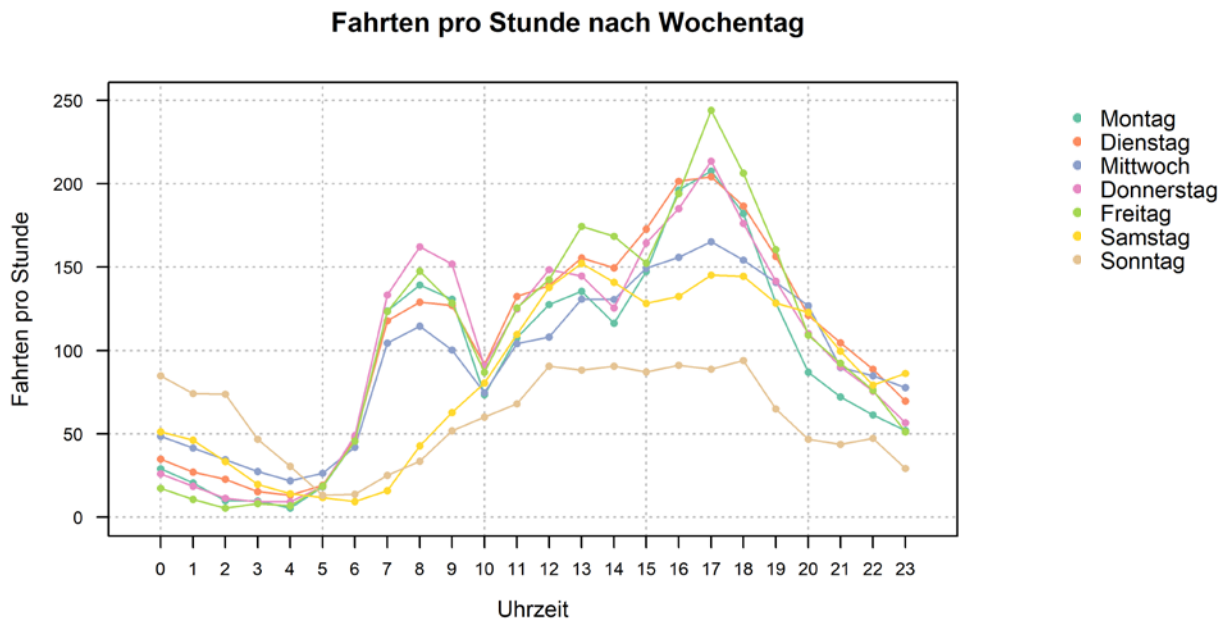


Abb.1: Nutzung im Tagesverlauf nach Wochentag

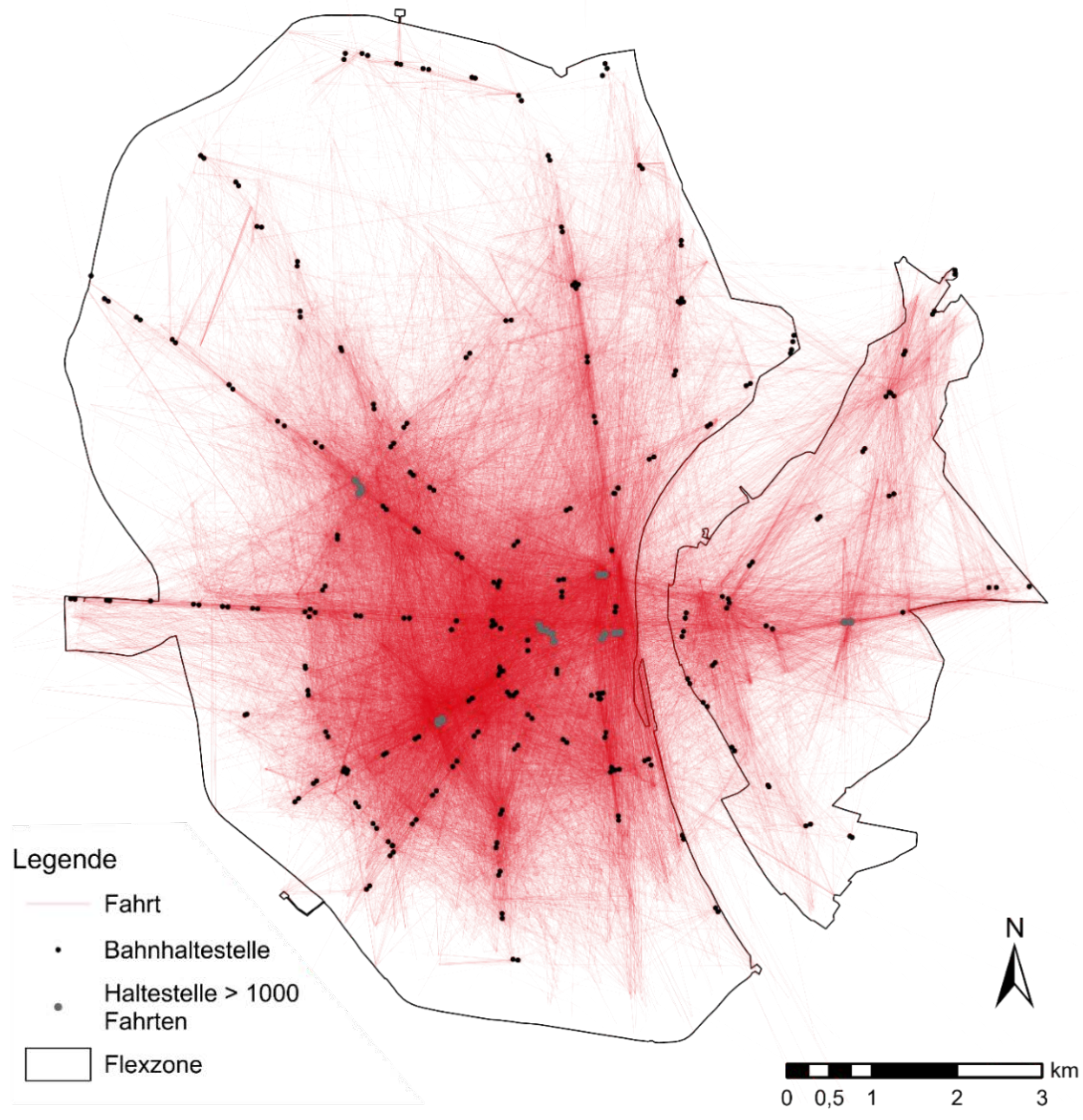
Die Nutzung von Bikesharing in Kombination mit einer ÖPNV-Fahrt hängt in großem Maße von der Verfügbarkeit von Rädern an Bus- und Bahnhaltstellen ab. Im Untersuchungszeitraum waren innerhalb von 300 m um U-Bahn-Haltstellen zu 50,9% und um Bushaltstellen zu 30,6% der Messzeitpunkte wenigstens ein Leihrad verfügbar. In Abbildung 2 wird zusätzlich die Struktur der Fahrten dargestellt, die zu einem Zielort innerhalb von 300 m um U-Bahn-Haltstellen führen.

Eine zeitliche Betrachtung der Anzahl an Bikesharing-Fahrten die vor, während und nach der Baustelle zu den vier unterschiedlichen Kategorien betroffener Haltestellen führen, zeigt, dass die durchschnittliche Fahrtenzahl während der Baustelle ansteigt und danach wieder ungefähr auf das Ursprungsniveau absinkt. Dabei liegt jedoch der Anteil der Fahrten, die zu zumindest teilweise gesperrten Haltestellen führen, während der Baustelle höher, während der Anteil der Fahrten, die zu Haltestellen mit zusätzlichen umgeleiteten Linien führen, niedriger ist.

Anhand der räumlichen Regressionsanalyse kann ein starker positiver Zusammenhang der Anzahl an Bikesharing-Fahrten mit der Anzahl an Geschäften, gastronomischen Einrichtungen und Bars/Kneipen sowie der Nähe zu Universitätseinrichtungen festgestellt werden. Auch weitere Points of Interest wie Gesundheitseinrichtungen, Sportanlagen oder touristische Attraktionen haben einen positiven Einfluss. Ein eindeutiger Zusammenhang mit den Elementen des ÖPNV-Systems ist ebenfalls nachweisbar: so sinkt die Anzahl der Fahrten mit steigendem Abstand zu Bus- und Bahnhaltstellen. Jedoch ist der gleiche Effekt bei Hauptverkehrsachsen zu beobachten. Hinsichtlich der Baustelle kann ein stärkerer Anstieg der Fahrten im Umfeld von gesperrten Haltestellen nachgewiesen werden. Es wird ein positiver Zusammenhang mit der Bevölkerungsdichte sowie mit dem Anteil der 18-29-Jährigen festgestellt.

Die zeitliche Regressionsanalyse führt zu den Erkenntnissen, dass insbesondere während der morgendlichen und abendlichen Rush-Hour viele Fahrten stattfinden. Auch der Semesterstart hat einen positiven Einfluss, während für die Baustelle kein signifikanter Zusammenhang nachgewiesen werden kann. Zusätzlich wird ein positiver Zusammenhang mit der Lufttemperatur und ein negativer mit der Niederschlagsmenge beobachtet.

Insgesamt werden in der Analyse vielfältige Zusammenhänge zwischen der Bikesharingnutzung und Elementen der Stadtstruktur nachgewiesen. Bezüglich der Interaktionen mit dem ÖPNV-System ist ein eindeutiger Zusammenhang erkennbar: Bei einer steigenden Distanz zur nächsten Haltestelle sinkt die Anzahl an Bikesharing-Fahrten. Die betrachtete Baustelle an mehreren innerstädtischen U-Bahn-Haltestellen hat geringfügige Auswirkungen auf die Nutzung des Systems, wobei die Bikesharingnutzung im Umfeld gesperrter Stationen tendenziell steigt. Damit kann vermutet werden, dass Bikesharing und Stadtbahn zwei Mobilitätsoptionen sind, die sich, zumindest in einem geringen Umfang, gegenseitig ersetzen können. Die Untersuchung zeigt zudem auf, wie Geodaten und frei verfügbare Standortdaten eines Bikesharingsystems kombiniert und innovativ ausgewertet werden können, um daraus für die Stadtplanung wichtige Erkenntnisse abzuleiten.



Impressum
Kartenerstellung: Katja Schimohr
Datum: 20.04.2020
TU Dortmund, Fakultät Raumplanung
Koordinatensystem: ETRS_1989_UTM_Zone_32N

Datenquellen:
Stadt Köln (2020) (www.offenedaten-koeln.de)

Abb. 2: Räumliche Verteilung der erfassten Bikesharing-Fahrten zu Bahnhaltestellen

Literatur

- KAVITI, S., VENIGALLA, M. M., ZHU, S., LUCAS, K., BRODIE, S. (2020), Impact of pricing and transit disruptions on bikeshare ridership and revenue. In: *Transportation* 47(2), 641-662.
- TRAN, T. D., OVTRACHT, N. (2018), Promoting sustainable mobility by modelling bike sharing usage in Lyon. In: *IOP Conference Series: Earth and Environmental Science* 143(1), 1-12.
- WANG, X., LINDSEY, G., SCHONER, J. E., HARRISON, A. (2015), Modeling Bike Share Station Activity: The Effects of Nearby Businesses and Jobs on Trips to and from Stations. In: *Journal of Urban Planning and Development* 142(1), 4-15.
- YOUNES, H., NASRI, A., BAIOCCHI, G., ZHANG, L. (2019), How transit service closures influence bikesharing demand; lessons learned from SafeTrack project in Washington D.C. metropolitan area. In: *Journal of Transport Geography* 76, 83-92.
- WANG, K., CHEN, Y.-C. (2020), Joint analysis of the impacts of built environment on bikeshare station capacity and trips attractions. In: *Journal of Transport Geography* 82, 1-10.

Meeresspiegelanstieg in Norddeutschland – Eine GIS-basierte Simulation und Visualisierung

Caroline SCHULDT

Geodäsie und Geoinformatik · HafenCity Universität Hamburg · Henning-Voscherau-Platz 1 · 20457 Hamburg
E-Mail: caroline.schuldt@hcu-hamburg.de

1 Einführung ins Thema

Der Meeresspiegelanstieg wird als eine der größten Bedrohungen für Küstenlebensräume weltweit erkannt (NICHOLLS et al. 2007). Die entstehenden Auswirkungen sind sowohl von ökologischer als auch von sozioökonomischer Bedeutung und treffen Küstenlandschaften, ihre Ökosystemleistungen und die Bevölkerung an der Küste (BRYAN et al. 2001). Die bestehenden Bedrohungen sowie die steigenden Risiken begründen das Interesse an möglichst genauen Einblicken in die Art und das Ausmaß möglicher Entwicklungen, die sich im Zuge des zukünftigen Meeresspiegelanstiegs zeigen. Unter Berücksichtigung verschiedener Szenarien kann ein Eindruck über die Größenordnungen der möglichen Veränderungen vermittelt werden. Um den bestehenden Küstenschutz miteinzubeziehen und bewerten zu können, sind außerdem Deiche und Sperrwerke zu berücksichtigen. Fraglich ist, ob die vorhandenen Deiche und Sperrwerke Überflutungen Einhalt gebieten können und wie die Situation nach einem Deichbruch aussehen würde.

Um die abstrakt wirkenden Risiken des Klimawandels zugänglich zu machen, müssen die wissenschaftlichen Erkenntnisse zum Meeresspiegelanstieg für die breite Öffentlichkeit visualisiert werden. Für die Verbreitung von kartographischen Informationen sind die Medien gesellschaftlich von größter und einflussreichster Bedeutung (SCHIEWE 2017). Eine Herausforderung liegt darin, die Inhalte schnell und leicht nachvollziehbar zu gestalten. Aus diesen Überlegungen entstand der Grundgedanke der vorliegenden Untersuchung. Das Ziel der Arbeit ist es, eine GIS-basierte Simulation des Meeresspiegels für die Küstenregion Deutschlands zu erstellen und diese für die Verwendung im journalistischen Kontext zu visualisieren.

2 Zusammenfassung des Vortragsinhalts

2.1 Einflussfaktoren auf die Simulation des Meeresspiegelanstiegs

Um das komplexe System eines ansteigenden Meeresspiegels so genau wie möglich zu modellieren, werden verschiedene Faktoren betrachtet. Neben der Genauigkeit spielen vor allem die regionalen Gegebenheiten eine Rolle, die in die Simulation integriert werden sollen. Zur Charakterisierung der Landoberfläche wird ein digitales Höhenmodell (DEM) genutzt. Das der Arbeit zugrunde liegende DEM besteht aus Daten der TanDEM-X-Mission, die sich durch ihre homogene Qualität und bisher unerreichte Genauigkeit auszeichnen (DLR 2018). Die Höhenwerte repräsentieren die ellipsoidischen Höhen von WGS84. Zusätzlich wird ein zweites DEM erstellt, das von Deichen bereinigt ist. Dafür werden alle Deiche, die im DEM enthalten sind, identifiziert und mittels einer Rasterberechnung in ihrer Höhe reduziert.

Da das DEM nur die Landoberfläche beschreibt, muss zudem die Wasserfläche modelliert werden, um als Grundlage für den ansteigenden Meeresspiegel zu dienen. Die Höhe des Meeresspiegels wird in Meter über Normalhöhennull angegeben und bezieht sich auf das Deutsche Haupthöhennetz DHHN92. Da der Meeresspiegel regional variiert, fließt in die Modellierung der mittlere höchste Wert der Wasserstände in einer Zeitspanne ein. Durch die Nutzung des Mittelwertes ist sichergestellt, dass keine monats-, tages- oder uhrzeitabhängigen Werte den aktuellen Wasserstand und damit die Simulation verfälschen. Für die Nord- und Ostsee liegen Messungen aus 165 Stationen vor, die mittels Inverser Distanzwichtung (IDW) zu einem flächendeckenden Meeresspiegel interpoliert werden.

Der nächste wesentliche Faktor zur Simulation des Meeresspiegelanstiegs ist die Vorhersage der Höhe der zukünftigen Meeresoberfläche. Basierend auf den Klimaszenarien des IPCC werden regionale Prognosen auf das deutsche Küstengebiet erstellt. Für die vorliegende Arbeit werden die fünf Eiskomponenten (dynamisches Eis und Oberflächenmassenausgleich in Grönland, dynamisches Eis und Oberflächenmassenausgleich in der Antarktis sowie Gletscher), die drei ozeanbezogenen Komponenten (dynamische Meeresoberflächenhöhe, globale thermosterische Anomalien und inverser Barometereffekt aus der Atmosphäre) sowie Landwasserspeicher zur Ermittlung der Prognosen für Nord- und Ostsee berücksichtigt.

Zur Änderung des relativen Meeresspiegelanstiegs gehören neben der ansteigenden Meeresoberfläche auch Änderungen in der Höhe der Landfläche. Diese vertikale Landbewegung kann auf der einen Seite durch Küstenerhöhungen den Klimawandel bedingten Meeresspiegelanstieg abschwächen. Auf der anderen Seite wird durch das Absinken der Landmassen der Meeresspiegelanstieg verstärkt. Aus diesem Grund weicht der Meeresspiegelanstieg lokal von den Prognosen ab.

Für die vorliegende Arbeit werden Daten der URL6a GPS Lösung verwendet, bei der GPS-Daten von 1995 bis 2014 analysiert wurden, und mittels IDW interpoliert.

Um die vorgestellten Faktoren in einer Simulation zusammenzuführen, müssen die Faktoren physisch sinnvolle Höhen vorweisen. Besonders für Höhenangaben eignet sich das Geoid, die Äquipotentialfläche des Erdschwerefelds, die den mittleren Meeresspiegel im Sinne der kleinsten Quadrate annähernd gleicht. Aus diesem Grund werden die unterschiedlichen Höhensysteme der Einflussfaktoren zu den Geoidhöhen des Modells EGM96 transformiert.

2.2 Identifikation von Überflutungsflächen

Auf Basis der vorgestellten Faktoren wird der Meeresspiegel für das Jahr 2100 simuliert. Die Simulation wird in ArcMap 10.6.1 und mithilfe von Python 2.7 durchgeführt. In einer Rasterberechnung werden der aktuelle Meeresspiegel (MS), die Prognose zum Meeresspiegelanstieg (MSA), die vertikale Landbewegung (VLB) und das DEM kombiniert, um die unter dem zukünftigen Meeresspiegel liegende Fläche zu erhalten. Dies erfolgt nach der Ungleichung

$$DEM + VLB \leq MS + MSA.$$

Nach dem entwickelten Raster werden alle Flächen, die unterhalb des prognostizierten Meeresspiegels liegen, potentiell überflutet. Das führt zu Seen im Landesinneren, die es aufgrund der fehlenden Verbindung zum Meer in der Realität nicht gibt. Aus diesem Grund wird im nächsten Schritt eine iterative Selektion durchgeführt, welche diejenigen Flächen des Datensatzes auswählt, die eine auf Basis des DEMs kreierte Küstenlinie berühren.

Aufbauend auf den identifizierten Überflutungsgebieten werden die sozioökonomischen Folgen anhand von Bevölkerungs- und Landnutzungsdaten ermittelt (vgl. Tab. 1). Bei der Berechnung der Überflutungsflächen zeigen sich gravierende Unterschiede – zum einen zwischen Nord- und Ostseeküste, zum anderen zwischen den verschiedenen Klimaszenarien, besonders aber zwischen der Annahme eines standhaltenden Küstenschutzes und eines Deichbruchs.

Tabelle 1: Landfläche und Bevölkerung, die dem Meeresspiegelanstieg in verschiedenen Szenarien ausgesetzt sind.

Szenario basierend auf IPCC-Prognosen	Nordsee		Ostsee	
	Fläche (km ²)	Bevölkerung	Fläche (km ²)	Bevölkerung
RCP4.5 (mittlerer Anstieg um 1,8 °C) mit Küstenschutz	476	1.400	585	4.077
RCP8.5 (mittlerer Anstieg um 3,7 °C) mit Küstenschutz	513	1.654	644	5.982
RCP8.5 (mittlerer Anstieg um 3,7 °C) ohne Küstenschutz	8.096	619.317	908	7.563

2.3 Visualisierung als Medienkarte

Ausgehend von den Anforderungen der journalistischen Kartographie werden statische Karten entwickelt, die als Sequenz beginnend mit den geringsten und abschließend mit den schlimmstmöglichen Auswirkungen gezeigt werden sollen. Die einzigen inhaltlichen Bestandteile der Karte sind das Festland und die Meeresoberfläche, die aufgrund ihrer Form von den Betrachtenden auf den ersten Blick erkennbar sind. Deutlich wird, dass die Deiche und Sperrwerke das Wasser aufhalten und nicht ins Festland eindringen lassen (vgl. Abb. 1). Vor den Deichen wird die Fläche allerdings überflutet. Das betrifft neben dem Deichvorland vor allem die Halligen, die keinen ausreichenden Deichschutz besitzen.

Mit einem höheren Wasserstand werden die Deiche stärker belastet und die Gefahr eines Deichbruchs steigt. Die damit einhergehende Gefahr wird in Signalfarben visualisiert. Beim Deichbruch-Szenario, das auf der IPCC-Prognose RCP8.5 basiert, fällt der große Unterschied zwischen Nord- und Ostsee auf (vgl. Abb. 2).



Abb. 1: Meeresspiegelanstieg bis 2100 nach RCP8.5 mit bestehendem Küstenschutz

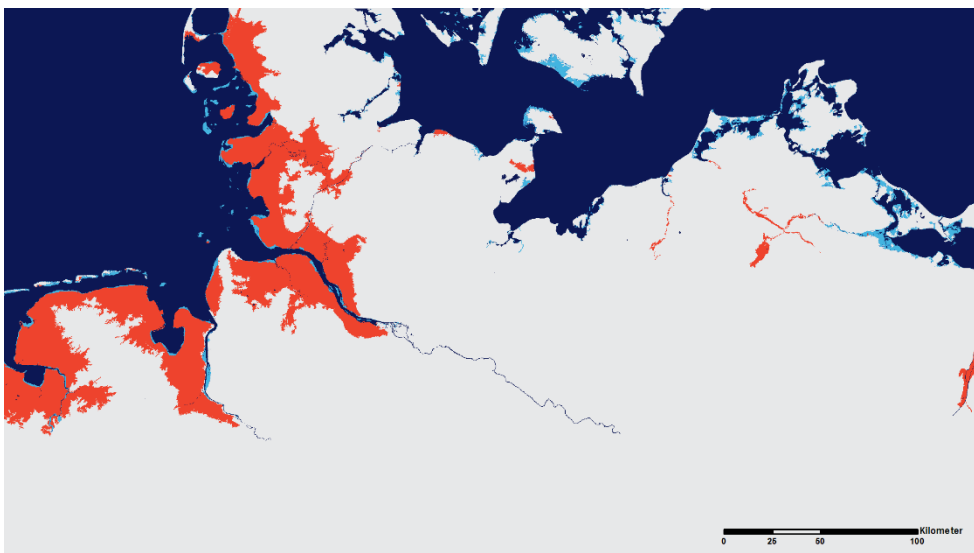


Abb. 2: Meeresspiegelanstieg bis 2100 nach RCP8.5 ohne bestehenden Küstenschutz

Literatur

- BRYAN, B., HARVEY, N., BELPERIO, T. & BURMAN, B. (2001), Distributed process modelling for regional assessment of coastal vulnerability to sea-level rise. In: *Environmental Modeling and Assessment*, 6 (1), 57-65.
- DLR (2018), TanDEM-X – Die Erde in drei Dimensionen. URL: <https://www.dlr.de/content/de/missionen/tandem-x.html> (abgerufen am 26.02.2021).
- NICHOLLS, R.J., WONG, P.P., BURKETT, V., CODIGNOTTO, J., HAY, J., MCLEAN, R., RAGOONADEN, S., WOODROFFE, C.D. (2007), Coastal systems and low-lying areas. In: PARRY, M., CANZIANI, O., PALUTKOF, J., VAN DER LINDEN, P. HANSON, C. (Hrsg.), *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, 315-356.
- SCHIEWE, J. (2017), Journalistic Cartography – Setting the Research Agenda. In: *Kartographische Nachrichten*, 67 (5), 245-253.

A GeoEvent-driven Architecture based on GeoMQTT for the Geospatial IoT

Stefan HERLÉ

Geodetic Institute and Chair for Computing in Civil Engineering & Geo Information Systems · RWTH Aachen University · Mies-van-der-Rohe-Str.1 · 52076 Aachen
E-Mail: herle@gia.rwth-aachen.de

1 Introduction

The Internet-of-Things (IoT) extends the conventional Internet of computers by physical things. Its evolution is driven by innovations in sensor and actuator technology as well as by the miniaturization of processing and communication technologies. The IoT enables cyber-physical concepts such as smart objects or digital twins. Since their physical counterparts of the real-world possess also spatial attributes such as their position or orientation, geospatial data streams have to be exchanged between different systems of the IoT in real-time.

In this Geospatial IoT, applications are driven by geospatial data and spatiotemporal data streams that are issued by a great amount of sensors and smart objects. The huge volume and high velocity of geospatial data raise a multitude of new challenges and opportunities in different research areas for Geospatial IoT architectures and platforms (RIEKE et al. 2018). One of those challenges involves the integration and efficient dissemination of messages to recipients in real-time. Therefore, this contribution proposes a concept for an Event-driven Architecture (EDA) based on GeoEvents to facilitate and optimize the dissemination of spatiotemporal messages published by devices in the Geospatial IoT. With the spatiotemporal information inherent in each GeoEvent, the architecture enables the distribution of messages utilizing spatiotemporal filtering mechanisms. For instance, a dissemination process can be realized based on proximity. The proposed architecture is implemented by GeoMQTT, an extension to the wide-spread IoT protocol Message Queue and Telemetry Transport (MQTT). It introduces message types for modeling GeoEvents and GeoSubscriptions. The latter one is used by clients to specify spatiotemporal filters which are applied to the GeoEvents during the dissemination process. Besides IoT devices and services, GeoMQTT is integrated into contemporary GIS technologies and Spatial Data Infrastructures (SDI), so that these may also participate in the architecture and consume real-time GeoEvents.

2 Summary

2.1 A GeoEvent-driven Architecture for the Geospatial IoT

EDAs are widely used in IoT systems. In these systems the flow of controls is determined by events that include different properties such as a name, an id or a type. Based on the event's information the dissemination, processing and the resulting activities are controlled. Due to the push-based communication pattern, events can be exchanged in real-time. In the Geospatial IoT, spatiotemporal information can be included in the events, which may facilitate the dissemination based on spatial properties such as the position of a smart object or the orientation of a car. In our proposed GeoEvent-driven Architecture we define the message type GeoEvent as a 4-tuple that consists of a name, a spatial and a temporal component as well as the payload. GeoEvents are used in the architecture to determine the flow of controls (see Fig. 1).

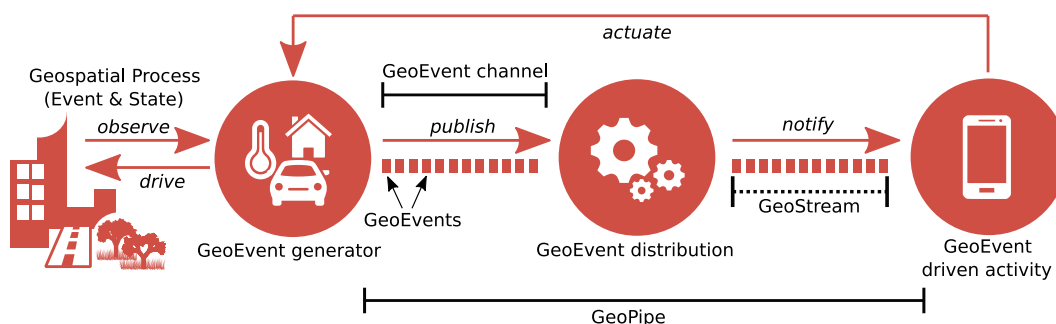


Fig. 1: GeoEvent-driven Architecture for the Geospatial IoT (Source: HERLÉ 2019)

GeoEvent generators (e.g. IoT devices) observe processes, events or states of the real-world. They create GeoEvents and publish these through channels. A GeoEvent distributor (or broker) analyses and disseminates the GeoEvents to consumers based on their interests which are registered in advance. The consumers parse the GeoEvents and invoke downstream activities. The architecture utilizes a publish/subscribe pattern with spatiotemporal subscriptions to register interests in GeoEvents with the distributor.

2.2 Dissemination using GeoMQTT

The proposed GeoEvent-driven architecture is implemented by GeoMQTT, an extension of the topic-based publish/subscribe protocol MQTT (ISO/IEC JTC 1 2016). In MQTT consumers specify their interest in events with topic filters that are applied by a broker to the topics of each event. GeoMQTT extends this mechanism by a GeoPublish and a GeoSubscribe type containing spatiotemporal components for modeling GeoEvents and GeoSubscriptions (Fig. 2).

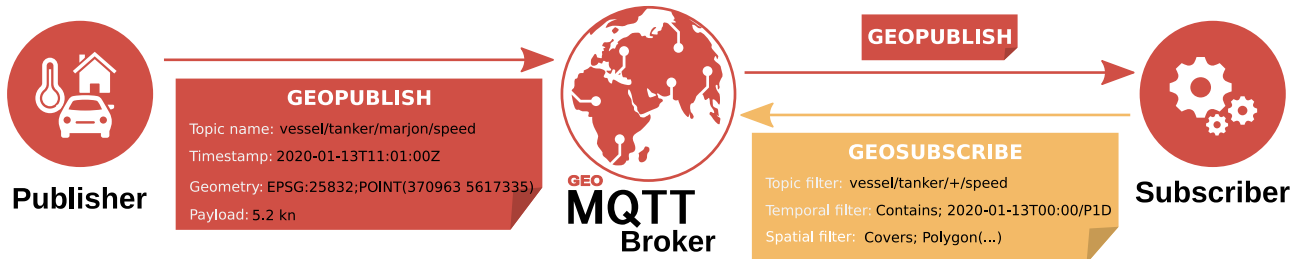


Fig. 2: GeoMQTT messaging and dissemination mechanism (Source: HERLÉ 2019)

The broker disseminates the GeoEvents based on the evaluation results of the filters specified in the subscriptions and the metadata of the events. Beside the topic filter, a GeoSubscription includes a temporal and as spatial filter. While the temporal filter consists of a time stamp/interval and a temporal relation, the spatial filter has a 2D geometry and a spatial relation according to the DE-9IM (HERRING 2011). The geometry can be defined with different encoding standards (e.g., GeoJSON) and any Coordinate Reference System (CRS). When forwarding to a specific client, the broker transforms the geometry of the GeoPublish message automatically into the requested CRS and encoding of the subscription.

2.3 GeoEvent Bus and GIS technologies

The mechanism of GeoMQTT allows the efficient real-time dissemination of GeoEvents between different producers and consumers in the Geospatial IoT based on spatiotemporal properties. A GeoEvent bus can be established to connect various GeoMQTT clients such as IoT sensors, data or processing services. The connected systems have access to the real-time spatiotemporal GeoEvents published by different producers. This way, we realized interfaces for Sensor Web Enablement (SWE) services to store and access historic GeoEvents, as well as Web Processing Services (WPS) to provide stream processing capabilities in real-time. Also contemporary GIS technologies such as desktop GIS or WebGIS portals can be connected to the bus and, thus, participate in consuming and producing GeoEvents. For example, Fig. 3 shows a fleet tracking system for public buses in the city of Aachen. The application is realized with GeoMQTT over WebSockets in a WebGIS to receive low-latency updates from the vehicles such as position or speed. The WebGIS uses the bounding box of the map view for the spatial filtering. If the user pans or zooms, the filter is adjusted accordingly.

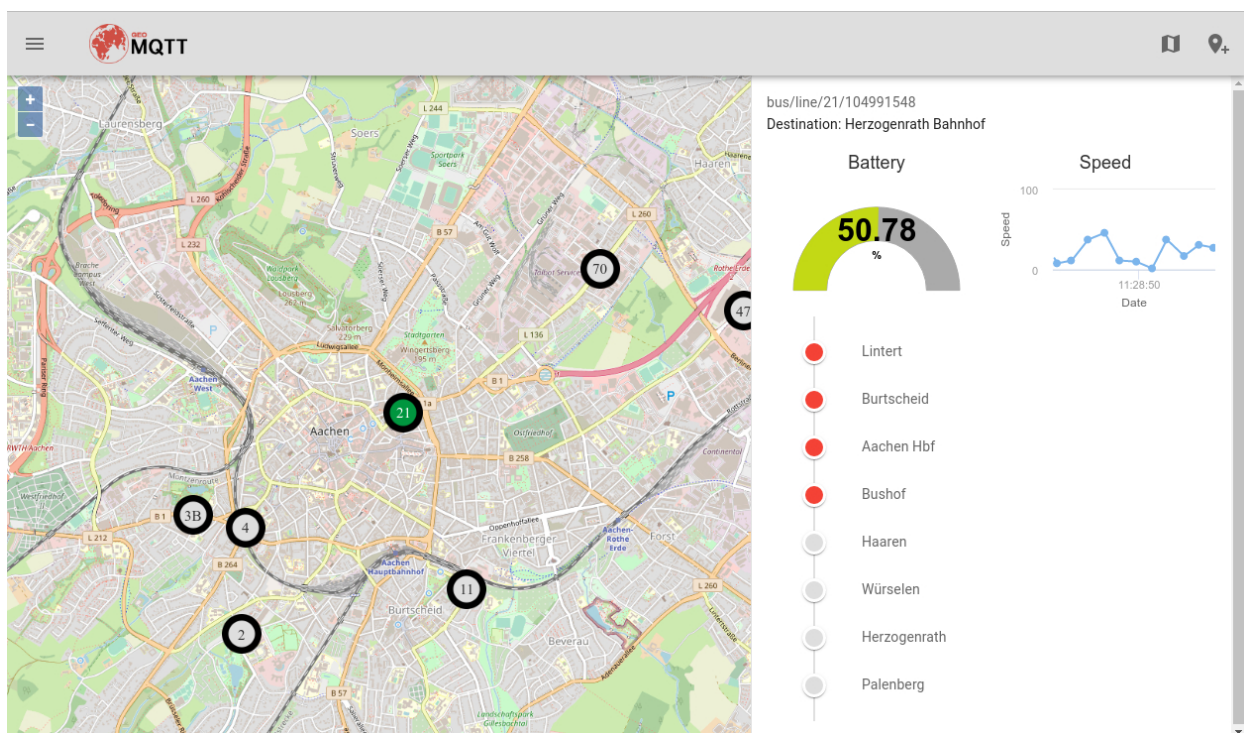


Fig. 3: Fleet tracking of public buses in a WebGIS using GeoMQTT over WebSockets

Literature

- HERLÉ, S. (2019), A GeoEvent-driven Architecture based on GeoMQTT for the Geospatial IoT (doctoral thesis). Publication series of the Geodetic Institute of RWTH Aachen University, No. 71. URL: <http://publications.rwth-aachen.de/record/772603> (last access: 23.02.2021).
- HERRING, J. R. (2011), OpenGIS Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture - Version 1.2.1," OpenGIS® Implementation Standard, 06-103r4.
- ISO/IEC JTC 1 (2016), Information technology – Message Queuing Telemetry Transport (MQTT) v3.1.1, ISO/IEC Standard, 20922.
- RIEKE, M.; BIGAGLI, L.; HERLÉ, S.; JIRKA, S.; KOTSEV, A.; LIEBIG, T.; MALEWSKI, C.; PASCHKE, T.; STASCH, C. (2018), Geospatial IoT—The Need for Event-Driven Architectures in Contemporary Spatial Data Infrastructures. In: ISPRS International Journal of Geo-Information, No. 7, pp. 1-29.

Webbasierte Innovationsindikatoren für mikrogeographische Analysen

Jan KINNE^{1,2,3}

¹Innovationsökonomik · Zentrum für Europäische Wirtschaftsforschung · L7, 1 · 68163 Mannheim, Deutschland

²Z_GIS - Department of Geoinformatics · Universität Salzburg · Kapitelgasse 4-6 · 5020 Salzburg, Österreich

³istari.ai UG (haftungsbeschränkt) · Julius-Hatry-Straße 1 · 68163 Mannheim, Deutschland

E-Mail: jan.kinne@istari.ai

1 Einführung ins Thema

Gegenstand der hier vorgestellten Dissertation ist die mikrogeographische Analyse von Unternehmensstandorten unter besonderer Berücksichtigung der Rolle von Innovationen. In den ersten beiden Beiträgen dieser Arbeit (Kapitel 2 und 3) werden detaillierte Geodaten zur mikrogeographischen Kartierung und ökonomischen Analyse von Unternehmensstandorten und relevanten Standortfaktoren verwendet. Die aufgezeigten Nachteile (mangelnde Aktualität, Abdeckung, Granularität und hohe Datenerhebungskosten) traditioneller, auf Umfragen und Patentdaten basierender Innovationsindikatoren motivieren den zweiten Teil dieser Arbeit, in dem ein neuartiger Ansatz zur Generierung webbasierter Innovationsindikatoren entwickelt, getestet und angewendet wird. Der Ansatz basiert auf einem Web-Mining-Framework, das sich auf eine speziell entwickelte Web-Scraping-Software (ARGUS) stützt, die zur Extraktion von Texten und Hyperlinks aus Unternehmenswebsites verwendet wird (Kapitel 4). Diese Webdaten werden dann analysiert, um mittels Data Mining innovationsbezogene Informationen zu finden. Die so gewonnenen Informationen dienen als Grundlage für einen neuartigen Typ von Innovationsindikatoren auf Unternehmensebene, die hochfrequent, granular und kostengünstig in großem Umfang erhoben werden können. Konkret wird ein "product innovator firm"-Vorhersagemodell auf Basis von Deep Learning und Website-Texten entwickelt, das Firmen aus einer traditionellen Innovationserhebung als Trainingsdaten verwendet (Kapitel 5). Das vorgeschlagene Framework wird auch für eine webbasierte Diffusionsanalyse verwendet, bei der die Verbreitung eines Informationssicherheitsstandards (ISO/IEC 27001) untersucht wird (Kapitel 6). In der letzten Studie (Kapitel 7) wird die Hyperlink-Vernetzung von innovativen und nicht-innovativen Unternehmen im Internet untersucht und deren Beziehungstypen bewertet. Diese Dissertation trägt zum Verständnis mikrogeographischer Wirtschaftsprozesse bei und entwickelt darüber hinaus einen neuen methodischen Ansatz zur Messung von Innovation. Gerade letzteres hat eine hohe gesellschaftliche Relevanz, da eine evidenzbasierte Politikgestaltung auf umfassende und aktuelle Indikatoren angewiesen ist, um Wirtschaftswachstum erfolgreich zu fördern sowie die Wirksamkeit wirtschaftspolitischer Maßnahmen zu evaluieren.

2 Zusammenfassung des Vortragsinhalts

2.1 Von der klassischen Standorttheorie zu einem mikrogeografischen Wahrscheinlichkeitsraster

Die Standorttheorie beschreibt, erklärt und bewertet wirtschaftlich-räumliche Systeme, deren geografisches Muster (Struktur), Interaktion (Beziehungen) und Dynamik. Es geht um die Frage, wo, wie und warum bestimmte Arten von wirtschaftlichen Aktivitäten angesiedelt sind. In Anbetracht der Tatsache, dass "das Standortmuster jeder Branche das Produkt einer großen Anzahl von Einzelentscheidungen ist" (SMITH 1981), besteht eines der Hauptziele darin, die Standortentscheidungen von Wirtschaftsakteuren zu untersuchen und jene Determinanten zu ermitteln, die diese Entscheidungen verursachen und beeinflussen. Diese Determinanten werden allgemein als Standortfaktoren bezeichnet. Die Anwendungsbreite der Standorttheorie kann von der einzelnen Betriebsstätte über Industriesektoren bis hin zu ganzen territorialen Produktionssystemen reichen. Jede dieser unterschiedlichen Analyseebenen geht mit einzigartigen Anforderungen an die verwendeten Daten, den theoretischen und methodischen Ansatz einher.

Smith schlug das Konzept einer kontinuierlichen räumlichen Variable vor, die als "spatial margins of profitability" oder "profit surface" bezeichnet wird und den geografischen Raum bestimmt, in dem ein Unternehmen einen Gewinn erzielen kann (oder nicht). In dieser Betrachtung können Unternehmen in der Tat einen Standort außerhalb der profitablen Bereiche wählen, aber sie werden weder in der Lage sein, einen Gewinn zu erzielen, noch auf lange Sicht wirtschaftlich lebensfähig zu sein. In Anbetracht der Tatsache, dass das Standortmuster jeder Branche das Produkt einer großen Anzahl individueller Entscheidungen ist, die von den Unternehmen dieser Branche getroffen werden, und dass sie "dazu tendieren, sich so nahe am optimalen Standort anzusiedeln, wie es ihr Wissen und ihre Fähigkeiten erlauben" (SMITH 1981), wird das gesamte Standortmuster der Unternehmen meist die *spatial margins of profitability* der betrachteten Branche nachzeichnen. Dieser Prozess der Ansiedlung von Unternehmen kann als ein stochastischer Prozess verstanden werden, der im *spatial probability framework* der Branche stattfindet, welches als ein Wahrscheinlichkeitsgitter (*probability grid*) modelliert werden kann, in dem jede Zelle durch eine bestimmte Wahrscheinlichkeit charakterisiert ist, ein Unternehmen anzuziehen. Diese Wahrscheinlichkeit lässt sich aus den in der jeweiligen Zelle vorhandenen Standortfaktoren ableiten.

In meiner ersten Studie "Analyzing and Predicting Micro-Location Patterns of Software Firms", die ich zusammen mit Bernd Resch verfasst habe und die im ISPRS International Journal of Geo-Information Band 7 Ausgabe 1 veröffentlicht wurde, haben wir offene Geodaten, Volunteered Geographic Information (VGI) und einen umfassenden Firmendatensatz (das Mannheimer Unternehmenspanel (MUP)) mit etwa drei Millionen Firmenbeobachtungen kombiniert, um ein branchenspezifisches Wahrscheinlichkeitsraster im Sinne von Smith empirisch zu schätzen. Auf der methodischen Ebene haben wir zunächst eine Explorative Spatial Data Analysis (ESDA) angewandt und dann in einem Folgeschritt ein Poisson-Regressionsmodell auf ein Ein-Kilometer-Raster gefittet, wobei jede Gitterzelle die Werte von 24 verschiedenen Standortfaktoren enthält. Diese Standortfaktoren umfassten Agglomerations-, Infrastruktur-, sozioökonomische, topografische und Annehmlichkeits-Standortfaktoren, die wir aus mikrogeografischen Daten abgeleitet haben. Die mikrogeografischen Daten basierten hauptsächlich auf OpenStreetMap (OSM) und offiziellen offenen Daten von statistischen Ämtern. Wir konzentrierten uns auf die Softwareindustrie, von der wir annehmen, dass sie in ihren Standortentscheidungen relativ frei ist und nur geringe Verzerrungen durch unbeobachtbare Standortfaktoren wie lokale Bebauungsbeschränkungen aufweist. Abbildung 1 veranschaulicht das Analyseraster und die von uns verwendeten Daten.

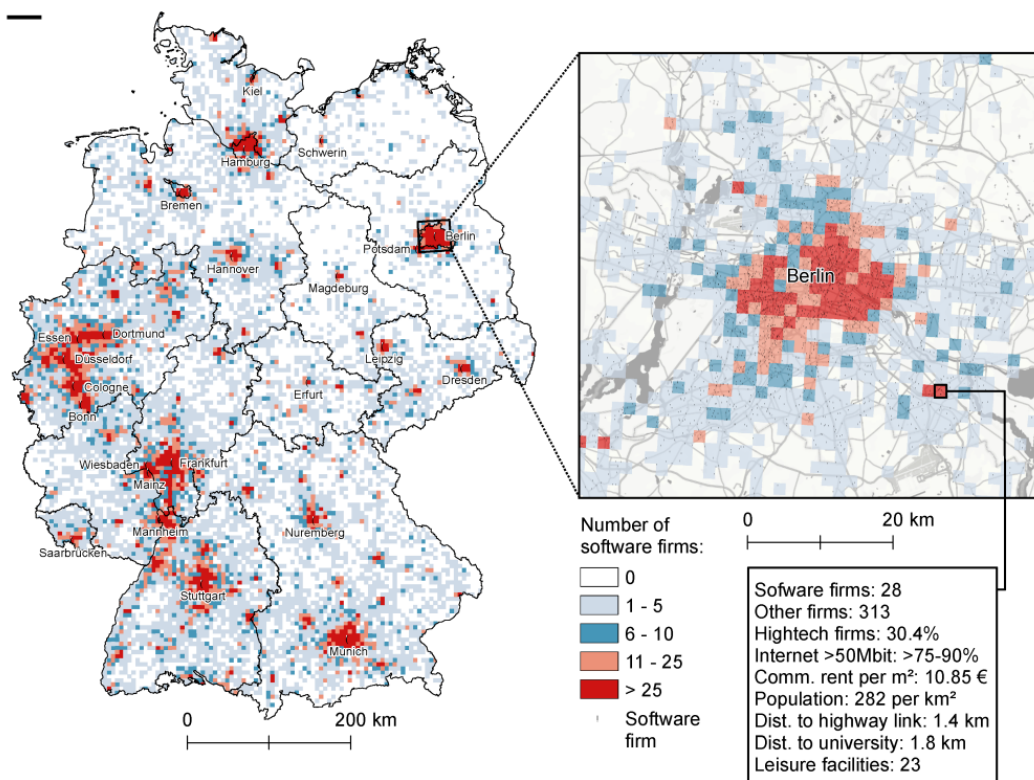


Abb. 1: Übersicht (5 km-Skala) und Zoom (1 km-Skala; mit Auswahl der Standortfaktoren für die Beispielzelle) des Standortmusters von Softwareunternehmen. Aus Kinne und Resch, 2018.

Wir schlussfolgerten, dass die mikrogeografische Analyseebene neue Einblicke in den Standortwahlprozess von Unternehmen liefert, aber auch, dass die meisten Standortfaktoren skalenrobust sind und dass unsere Ergebnisse in Bezug auf die Effekte von Standortfaktoren im Einklang mit früheren Studien mit aggregierten räumlichen Einheiten stehen. Wir kamen auch zu dem Schluss, dass unser mikrogeografisches Vorhersagemodell in der Lage ist, den Standort von Softwareunternehmen in einem zufriedenstellenden Maße vorherzusagen. Allerdings wiesen wir auch auf die besonderen Anforderungen an das statistische Modell und die Daten hin, die in einer mikrogeografischen Standortanalyse verwendet werden, wie zum Beispiel die Notwendigkeit von hochauflösenden Geodaten, die nicht in allen Bereichen verfügbar waren. Wir zeigten, dass dieses Problem in Städten am schwerwiegendsten ist, da diese oft segregierte Bevölkerungen und Stadtteile mit sehr unterschiedlichen sozioökonomischen Profilen aufweisen. Wir konnten feststellen, dass diese niedrig aufgelösten Geodaten zu systematischen Prognosefehlern in Städten führen. In der zweiten Arbeit meiner Dissertation habe ich mich daher speziell mit der Frage nach Unternehmensstandorten innerhalb von Städten beschäftigt und dabei vollständig disaggregierte Geodaten verwendet.

2.2 Wissensdichte und Unternehmensinnovation

In meiner zweiten Studie "Knowledge Proximity and Firm Innovation: A Microgeographic Analysis for Berlin", welche ich zusammen mit Christian Rammer und Knut Blind verfasst habe und die in Urban Studies Band 57 Ausgabe 5 veröffentlicht wurde, untersuchen wir das disaggregierte Standortmuster von innovativen und nicht-innovativen Firmen in Berlin und ihre Nähe zu Wissensquellen. Als Datengrundlage dienten ca. 8.000 auf Straßenebene geokodierte Berliner Unternehmen aus dem Mannheimer Unternehmenspanel (MUP), die an der Erhebung des Berliner Innovationspanels teilgenommen haben. In der Befragung werden den Unternehmen eine Reihe von Fragen zu ihren Innovationsaktivitäten gestellt. Die Fragen reichen von der Art und Weise der Forschung und Entwicklung (FuE) bis hin zur Einführung neuer oder deutlich verbesserter Produkte oder Dienstleistungen. Zusätzlich zu diesen Daten, aus denen wir die Standorte innovativer und nicht-innovativer Unternehmen in den Jahren 2012-2016 abgeleitet haben, haben wir eine Datenbank über die Lage und Größe von Hochschulen (Campus) und Forschungseinrichtungen in Berlin erstellt. Abbildung 2 zeigt die Standorte der innovativen und nicht-innovativen Unternehmen sowie der Hochschulstandorte und Forschungseinrichtungen, die wir in dieser Studie verwendet haben.

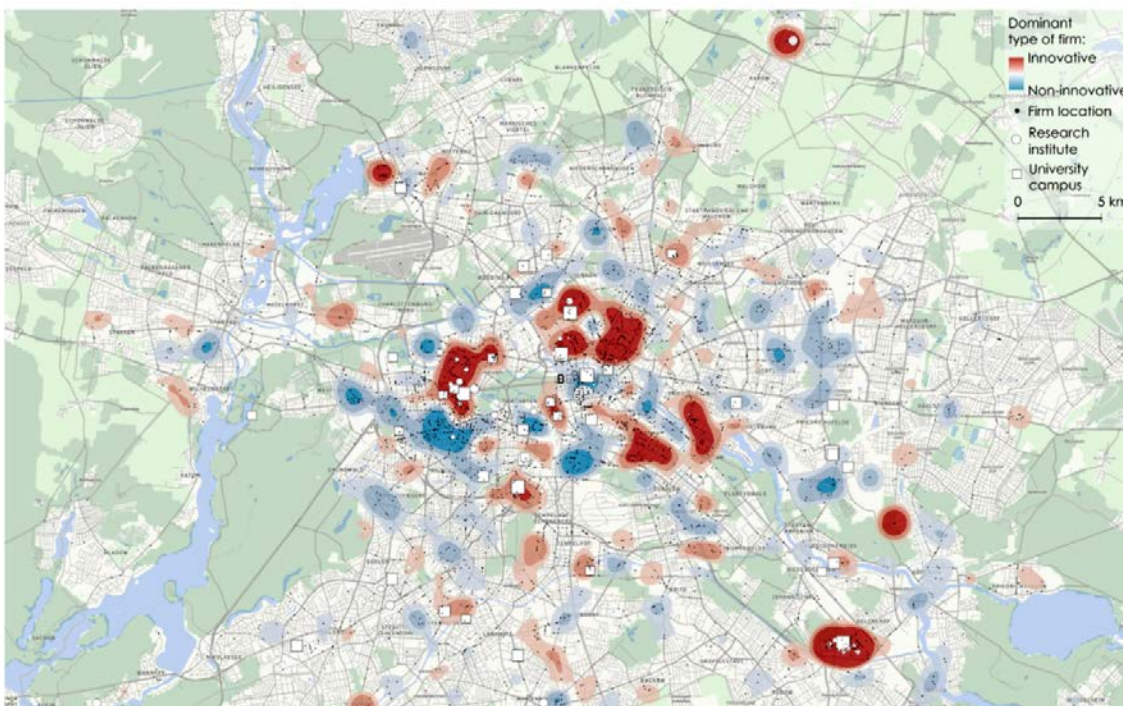


Abb. 2: Innovative und nicht-innovative Unternehmen und Wissensquellen in Berlin. Aus Rammer, Kinne, und Blind, 2020.

Neben einer deskriptiven statistischen Analyse haben wir auch ein Probit-Regressionsmodell verwendet, um die Beziehung zwischen dem Innovationsstatus eines Unternehmens (abhängige Variable in der Regressionsanalyse) und seinem Wissensumfeld (erklärende Variablen) zu analysieren. Wir fanden heraus, dass innovative, im Gegensatz zu nicht-innovativen Unternehmen, an Orten mit einem viel höheren *Local Buzz* und einer höheren Dichte an akademischen Wissensquellen angesiedelt sind. Eine unserer wichtigsten Erkenntnisse war, dass die geographische Reichweite dieser Unterschiede sehr begrenzt zu sein scheint. Was die Nähe zu Forschungsinstituten betrifft, nimmt die Konzentration innovativer Firmen beispielsweise bereits jenseits eines 50 m-Radius ab. Jenseits einer Entfernung von 1 km haben wir keinen signifikanten Zusammenhang mehr gefunden. Wir kamen zu dem Schluss, dass unsere Studie als ein erster Versuch gesehen werden kann, die Rolle von Wissensnähe für Innovation in der Stadt auf mikrogeografischer Ebene zu untersuchen. Wir gaben jedoch auch zu bedenken, dass wir davon abgesehen haben, ein umfassenderes Modell des städtischen Umfelds, in dem das Unternehmen tätig ist, zu untersuchen, beispielsweise indem wir die städtische Infrastruktur, Dichte oder Freizeiteinrichtungen in unsere Analyse einbeziehen. Wir haben auch betont, dass wir aufgrund unseres empirischen Settings nur Korrelationen zwischen Innovation und lokalen Wissensquellen beobachten konnten und keine streng identifizierten kausalen Beziehungen. Die Beschränkung unserer Studie auf die in der CIS erfassten Unternehmenstypen und die geografische Beschränkung auf eine einzige Stadt waren weitere Einschränkungen, die wir diskutiert und schließlich im zweiten Teil meiner Dissertation behandelt haben.

2.3 Ein Framework für webbasierte Innovationsindikatoren

In den ersten beiden Papieren dieser Dissertation haben wir gezeigt, dass die mikrogeografische Perspektive neue Erkenntnisse über die unternehmerische Standortwahl und die damit verbundenen Standortfaktoren bietet. Wir konnten auch zeigen, dass mikrogeographische Analysen mit besonderen Anforderungen an die verwendeten Daten und Methoden einhergehen. Für die Modellierung "harter" Standortfaktoren (d.h. Infrastruktur) konnten wir insbesondere auf hochauflösende VGI-Daten zurückgreifen. Im Gegensatz dazu erwies sich die Verfügbarkeit von hochauflösenden "weichen" Standortfaktoren teilweise als sehr begrenzt. So konnten wir beispielsweise Agglomerationsvorteile für ganz Deutschland noch mit der lokalen Dichte von Unternehmensstandorten aus dem MUP-Unternehmensdatensatz modellieren, während umfassende Informationen zur Innovationstätigkeit von Unternehmen nur für Berlin zur Verfügung standen (abgeleitet aus dem einzigartigen Berliner Innovationspanel). Eine Ausweitung unserer mikrogeografischen Studie auf andere Regionen war aufgrund der aufwändigen und kostspieligen Erhebung von Innovationsinformationen über fragebogenbasierte Befragungen nicht möglich. Patente (Anmeldungen, Zitationen und Lizenzen) werden häufig als alternative Innovationsindikatoren auf Unternehmensebene verwendet. Allerdings haben Patente als Innovationsindikatoren auch einige bekannte Unzulänglichkeiten. Dies gilt insbesondere für ihre sehr unterschiedliche Bedeutung innerhalb einiger Branchen. Während z.B. Unternehmen der Pharma- und Maschinenbaubranche häufig Patente anmelden, um ihre Erfindungen zu schützen, spielen Patente für Unternehmen der Informations- und Kommunikationstechnologie (IKT) kaum eine Rolle, da Software in Europa sehr schwer zu patentieren ist.

Im zweiten Teil meiner Dissertation habe ich mich daher auf die Entwicklung einer neuen Generation von webbasierten Innovationsindikatoren konzentriert, die überall zur Verfügung stehen und somit auch für umfassende mikrogeographische Analysen geeignet sind. Diese neuartigen Innovationsindikatoren basieren auf der Tatsache, dass Unternehmen in zunehmendem Maße digitale Spuren hinterlassen, die gesammelt und ausgewertet werden können. Eine effektive Nutzung dieser digitalen Daten ist erst in den letzten Jahren durch große methodische Fortschritte im Bereich des Data Mining, insbesondere des Natural Language Processing (NLP), möglich geworden. Die Adaption dieser neuartigen Methoden zum Zweck der Extraktion von innovationsrelevanten Informationen aus Webdaten, deren Transformation in Innovationsindikatoren auf Unternehmensebene und deren anschließende Nutzung in mikrogeographischen Analysen bilden den Kern des zweiten Teils dieser Dissertation.

In meiner dritten Studie "Web Mining for Innovation Ecosystem Mapping: A Framework and a Large-scale Pilot Study", die ich zusammen mit Janna Axenbeck verfasst habe und die in *Scientometrics* Band 125 Ausgabe 3 veröffentlicht wurde, entwickeln und testen wir ein Web Mining Framework für die Generierung von Innovationsindikatoren auf Unternehmensebene. Wir stellten zunächst ein kohärentes Web-Mining-Framework vor, das auf ARGUS (Automated Robot for Generic Universal Scraping) basiert, einem frei verwendbaren Web-Scraping-Tool, das wir für den Zweck des groß angelegten Broad-Web-Scraping (d.h. das Scraping mehrerer verschiedener Websites, im Gegensatz zum Focus-Web-Scraping, bei dem eine einzelne Website gescraped wird) entwickelt haben. Anschließend haben wir ARGUS in zwei Pilotstudien eingesetzt. Die erste diente dazu, Firmenwebsites als Datenquelle zu bewerten, indem eine große Anzahl von Firmenwebsites gescraped und analysiert wurde. In der zweiten Teilstudie verwendeten wir unser vorgeschlagenes Framework, um das Innovationsökosystem "KI in Berlin" abzubilden - ein Netzwerk von Berliner Firmen, die sich mit künstlicher Intelligenz beschäftigen.

Das von uns vorgeschlagene Framework ist in Abbildung 3 dargestellt. Es basiert auf einer Unternehmensdatenbank, die Informationen über Unternehmensmerkmale (z.B. Branche, Unternehmensgröße) und vor allem die Website-Adressen (URLs) der Unternehmen enthält. Idealerweise wurde die Unternehmensdatenbank mit Zusatzdatenbanken abgeglichen, die etablierte Innovationsindikatoren aus fragebogenbasierten Erhebungen, Patentdaten auf Unternehmensebene oder Literaturdaten enthalten, so dass traditionelle Innovationsindikatoren für eine Untergruppe der Unternehmen im Hauptdatensatz verfügbar sind. In einem ersten Schritt werden die Webadressen der Unternehmen an einen Web Scraper übergeben. Der Web Scraper wird dann verwendet, um Website-Inhalte (Texte, Hyperlinks etc.) von den Websites der Unternehmen herunterzuladen. In einem dritten Schritt werden mit Hilfe von Data-Mining-Techniken Informationen über die Innovationsaktivitäten der Unternehmen aus den heruntergeladenen Website-Inhalten extrahiert. Basierend auf diesen Informationen können neuartige Innovationsindikatoren konstruiert werden. In diesem Schritt können zusätzliche Metadaten über das Unternehmen verwendet werden, um die Analyse zu unterstützen (Vorklassifizierung, Auswahl eines Klassifizierungsmodells basierend auf Firmeneigenschaften, Informationen aus etablierten Innovationsindikatoren usw.). In einem letzten Schritt werden die neuen Innovationsindikatoren wieder mit der ursprünglichen Firmendatenbank zusammgeführt. In diesem letzten Schritt wird auch eine direkte Verbindung auf Unternehmensebene zwischen dem neuen Innovationsindikator und den etablierten Indikatoren aus den Hilfsdatenbanken hergestellt. Diese Verknüpfung kann dann genutzt werden, um die neuen Indikatoren mit traditionellen Indikatoren zu vergleichen.

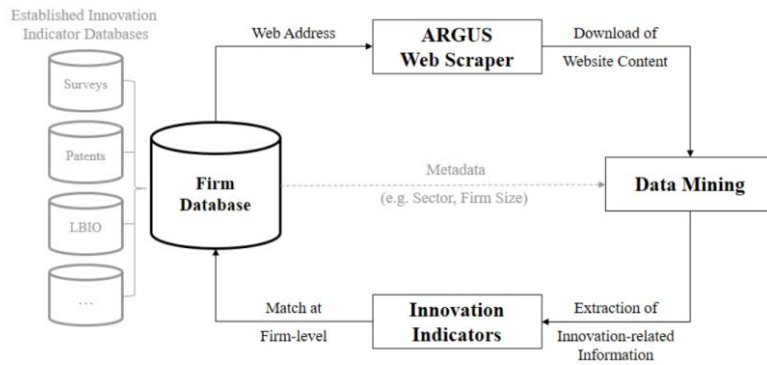


Abb. 3: Analyse-Framework zur Generierung webbasierter Innovationsindikatoren. Aus Kinne und Axenbeck, 2020.

Wir analysierten mit diesem Framework zuerst die grundlegenden Merkmale von Unternehmenswebsites in Bezug auf ihre Struktur und ihren Umfang. Dazu verwendeten wir ARGUS und eine umfangreiche Unternehmensstichprobe. Anschließend haben wir mit ARGUS die Websites aller in Berlin ansässigen Unternehmen gescraped. Mit einer Schlagwort-suche haben wir dann diejenigen Unternehmen identifiziert, die "Künstliche Intelligenz" auf ihren Websites erwähnen. Ziel war es, das "Innovationsökosystem KI Berlin" abzubilden, das aus Unternehmen, Interessengruppen und wissenschaftlichen Einrichtungen besteht, die "KI-engagiert" sind, d.h. sich in irgendeiner Form mit KI beschäftigen. Abbildung 4 zeigt den Anteil dieser Unternehmen an der lokalen Unternehmenspopulation für die Stadt Berlin.

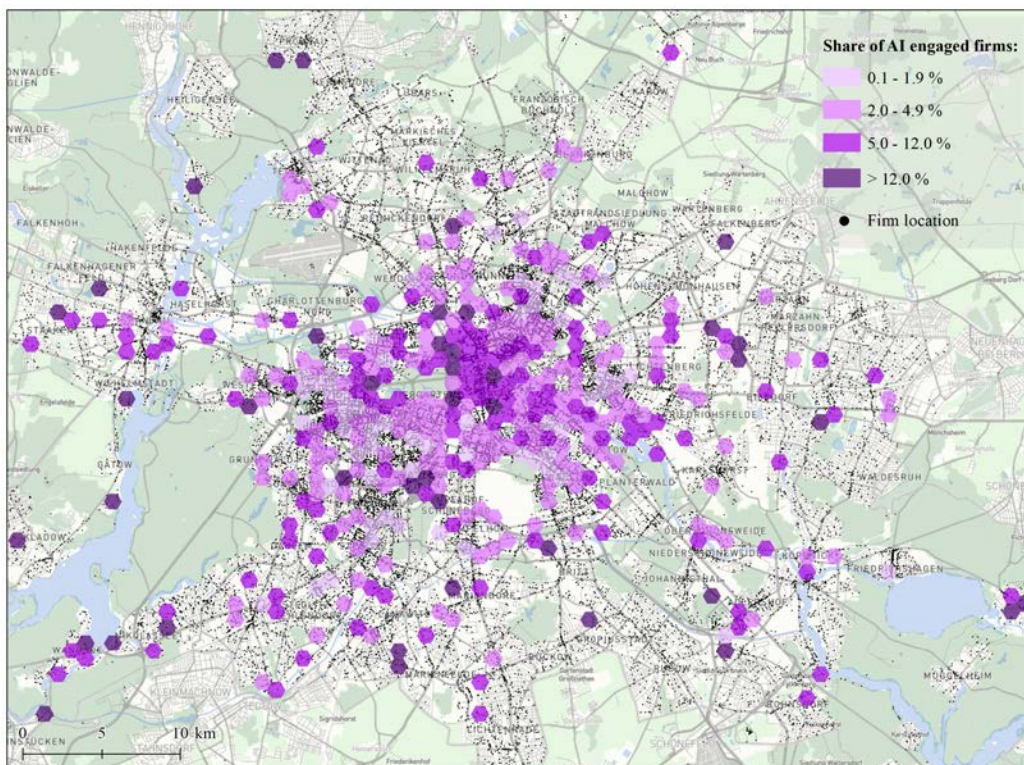


Abb. 4: Anteil der in Berlin ansässigen Unternehmen, die KI mindestens einmal auf ihrer Website erwähnen. Aus Kinne und Axenbeck, 2020.

Mit dieser Studie haben wir gezeigt, dass das von uns vorgeschlagene mehrstufige Web-Mining-Framework (Web Scraping, Data Mining, Indikatorengenerierung und Validierung) für die Kartierung von Innovationsökosystemen anwendbar ist und aussagekräftige Ergebnisse liefert. Allerdings wiesen wir auch darauf hin, dass ein ausgefeilterer Text-Mining-Ansatz notwendig wäre, um die verschiedenen Akteursgruppen (z.B. Firmen, die KI-basierte Produkte und Dienstleistungen anbieten, Universitäten, die sich mit Grundlagenforschung zu KI beschäftigen, und Interessengruppen, die KI-zentrierte Agenden vorantreiben) zu unterscheiden, die sich aus unserer einfachen Stichwortsuche ergaben. Dieses Problem haben wir in meinem vierten Beitrag behandelt, der im folgenden Kapitel vorgestellt wird.

2.4 Webbasierte Identifikation von innovativen Unternehmen

Im vorigen Kapitel haben wir einen ersten Ansatz für den Data-Mining-Schritt unseres entworfenen Web-Mining-Analyse-Frameworks vorgeschlagen, der auf einer einfachen Stichwortsuche basiert. Unsere Ergebnisse wiesen bereits in eine vielversprechende Richtung, aber wir stellten auch fest, dass ein ausgefeilterer Textanalyseansatz erforderlich wäre, um wertvolle Informationen aus den heruntergeladenen Webtexten zu extrahieren. Der Bereich der Natural Language Processing (NLP) hat in den letzten Jahren große Fortschritte bei der Analyse natürlicher Sprache gemacht, insbesondere durch den Einsatz von Deep Learning-basierten Algorithmen (siehe insbesondere MIKOLOV et al. 2011) und, in jüngerer Zeit, auf Transfer Learning basierenden Ansätzen. Diese methodischen Fortschritte gingen einher mit einer zunehmenden Verfügbarkeit von leistungsfähiger Hardware (insbesondere für Deep Learning geeignete Grafikprozessoren) und benutzerfreundlicher Software (z. B. die Python-Bibliothek Keras).

Das Ziel meiner vierten Arbeit "Predicting Innovative Firms using Web Mining and Deep Learning", die ich zusammen mit David Lenz konzipiert und geschrieben habe, war es, moderne Deep-Learning-Methoden für die Analyse von Webtexten und die Identifikation von innovativen Unternehmen (d.h. Produktinnovatoren) einzusetzen. Abbildung 5 skizziert unseren Ansatz. Die Websites von Unternehmen mit unbekanntem Produktinnovator-Status werden mit ARGUS abgefragt und alle Texte von diesen Websites heruntergeladen. Anschließend werden die heruntergeladenen Texte mittels tf-idf-Textvektorisierung als numerische Vektoren dargestellt. Diese Vektoren dienen dann als Input für ein künstliches neuronales Netzwerk (ANN), das zuvor mit Webtexten von Unternehmen trainiert wurde, für die traditionelle Innovationsindikatoren verfügbar sind. Hierfür wurden die in der MIP-Innovationserhebung befragten Unternehmen verwendet. Für jedes in dieser Umfrage befragte Unternehmen ist der Status des Unternehmens als Produktinnovator (oder Nicht-Innovator) bekannt und kann als "Label" für die Webtexte des Unternehmens (die ebenfalls mit ARGUS heruntergeladen wurden) verwendet werden. Während des Trainings "lernt" das ANN, welche Wörter oder Wortkombinationen einen Produktinnovator von einem Nicht-Innovator unterscheiden. Nach dem Training kann der Text eines Unternehmens ohne bekannten Innovationsstatus in das Modell eingegeben werden, das dann eine Vorhersage über die Wahrscheinlichkeit macht, dass das Unternehmen ein Produktinnovator ist ("Produktinnovator-Wahrscheinlichkeit"). Die Gesamtperformance des Modells, die sich an dem f1-Score von insgesamt 0,80 ablesen lässt, kann als sehr zufriedenstellend angesehen werden. Dies gilt insbesondere, da ein Modell, das nur auf traditionellen Unternehmensinformationen (Unternehmensgröße, Alter und Branche) basiert, nur einen f1-Score von etwa 0,70 erreicht. Vereinfacht ausgedrückt: Unser Modell ist in der Lage, einem Unternehmen in 80% der Fälle einen korrekten Innovationsstatus zuzuordnen, und das nur auf Basis der Unternehmenswebsite.

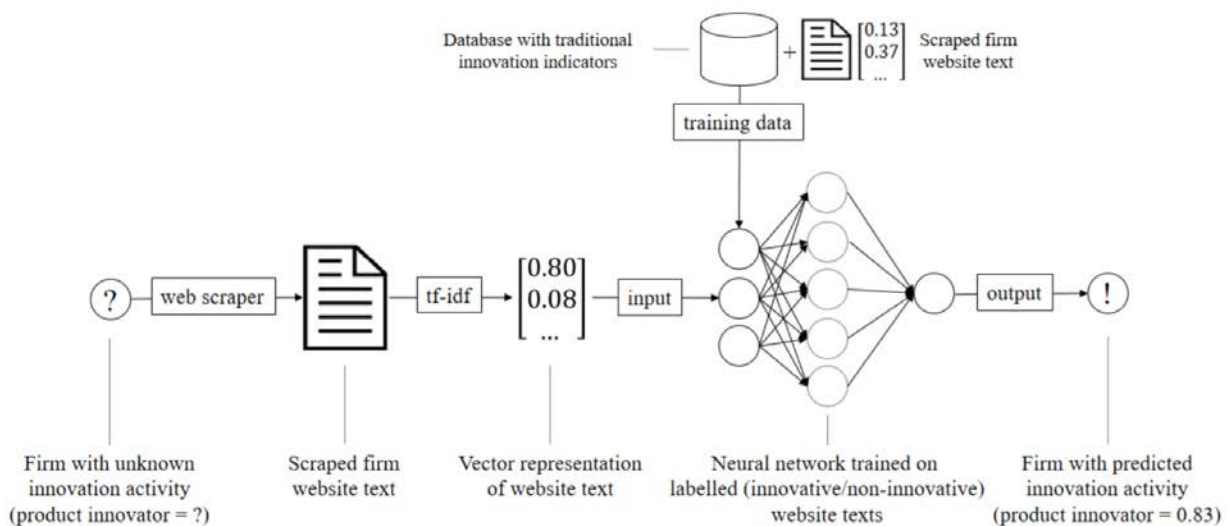


Abb. 5: Framework zur Vorhersage von Unternehmen mit Produktinnovationen anhand ihrer Website-Texte. Aus Kinne und Lenz, 2019.

Im zweiten Teil der Arbeit haben wir das trainierte Modell verwendet, um Produktinnovationswahrscheinlichkeiten für etwa 700.000 deutsche Unternehmen zu berechnen, die eine eigene Website haben. Diese Out-of-Sample-Vorhersagen haben wir dann für umfassende Robustheitstests verwendet, indem wir unseren neuen webbasierten Innovationsindikator mit etablierten Indikatoren verglichen haben. Eine dieser Robustheitsprüfungen war der Vergleich des mikrogeographischen Musters unseres webbasierten Indikators mit dem Muster einer Sondererhebung des MIP in Berlin (die Informationen aus dieser Sondererhebung wurden beim Training nicht verwendet). Dieser Vergleich ist in Abbildung 6 dargestellt, wo wir eine deutliche Übereinstimmung zwischen den beiden Mustern sehen können.

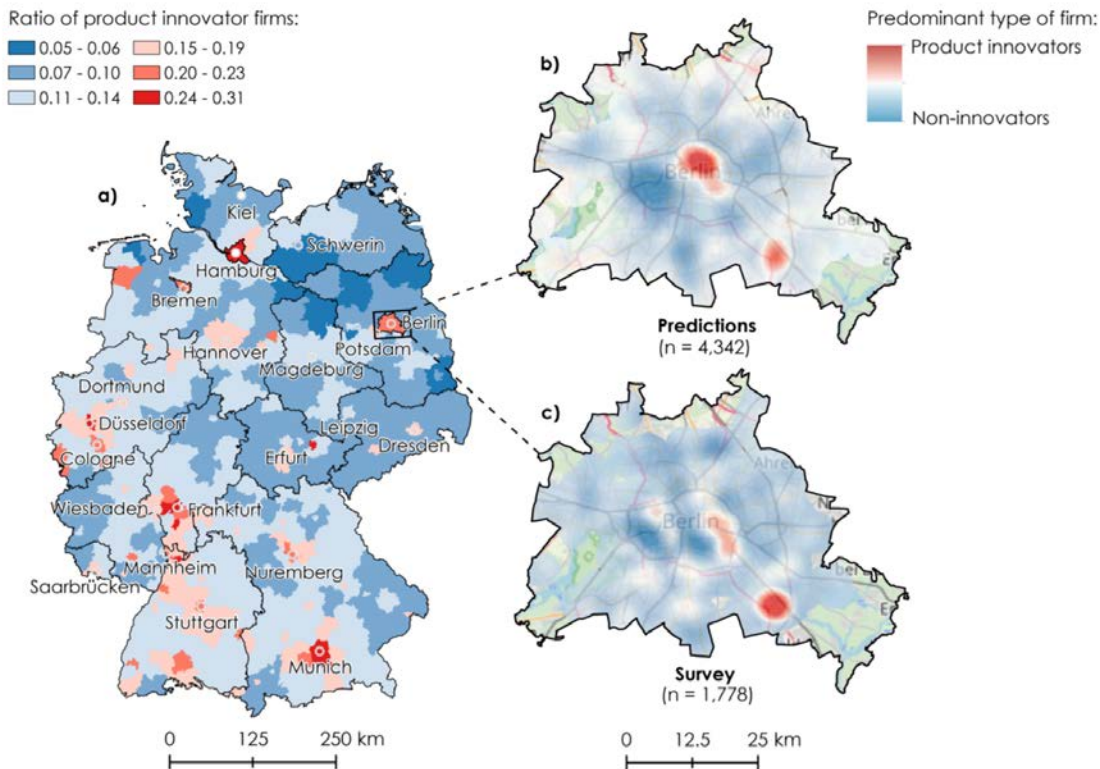


Abb. 6: Karte mit Produktinnovatoren. Links: Prognostizierter Anteil der Unternehmen mit Produktinnovationen nach deutschen Landkreisen. Rechts: Vorhergesagtes (oben) und erhobenes (unten) geographisches Muster von Produktinnovatoren und Nicht-Innovatoren in Berlin. Aus Kinne und Lenz, 2019.

Wir kamen zu dem Schluss, dass unser Modell zur Vorhersage von Produktinnovatoren (später InnoProb genannt) eine sehr gute Leistung erzielt, sowohl innerhalb des Testsatzes als auch im Vergleich zu den etablierten Indikatoren. Als Schwäche stellten wir die etwas schwache Leistung in Bezug auf den Recall von innovativen Unternehmen fest (d.h. recht viele innovative Unternehmen werden nicht erkannt, während nicht-innovative Unternehmen sehr zuverlässig erkannt werden). Insgesamt waren wir zuversichtlich, dass wir einen wertvollen Ansatz für die Analyse von Innovationen in Unternehmen geschaffen haben, der es Forschern ermöglicht, in jede beliebige Region "hineinzuzoomen" und mikrogeografische Analysen an einer umfassenden Stichprobe von Unternehmen durchzuführen. In den letzten beiden Arbeiten meiner Dissertation, die in den nächsten beiden Kapiteln vorgestellt werden, haben wir unseren Indikator für groß angelegte webbasierte Innovationsstudien verwendet.

2.5 Eine webbasierte Diffusionsanalyse

Mein fünftes Paper "Exploring the Adoption of the International Information Security Management System Standard ISO/IEC 27001: A Web Mining-Based Analysis", das ich gemeinsam mit Mona Mirtsch und Knut Blind verfasst habe, wurde in der IEEE Transactions of Engineering Management Volume 68 Number 1 veröffentlicht. Im Rahmen des Papiers haben wir einen Web Mining-basierten Ansatz entwickelt, um Unternehmen zu identifizieren, die nach einem bestimmten Standard zertifiziert sind. Wir validieren unsere Ergebnisse durch eine aufwändige manuelle Überprüfung der zuvor per Web Mining identifizierten, vermeintlich zertifizierten Unternehmen. Der Standard, den wir für diese Studie gewählt haben, ist der ISO/IEC 27001-Standard, der Unternehmen und Organisationen dabei helfen soll, "Informationssicherheits-Managementsysteme" zu entwickeln und zu pflegen. Dieser Standard wurde als "eines der effektivsten Risikomanagement-Tools zur Abwehr der Milliarden von Angriffen, die jedes Jahr stattfinden" beschrieben.

Als Datengrundlage wurde erneut das Mannheimer Unternehmenspanel (MUP) verwendet, das auf Basis des von uns vorgeschlagenen und getesteten Web-Mining-Frameworks um Webtextinformationen erweitert wurde. Zusätzlich steht für jedes der Unternehmen der webbasierte Innovationsindikator zur Verfügung, den wir im vorherigen Kapitel vorgestellt haben. In einem ersten Schritt haben wir Unternehmen identifiziert, die mindestens einen Textbezug zu den Managementnormen ISO 9001 (Qualitätsmanagement), ISO 14001 (Umweltmanagement), ISO 50001 (Energiemanagement) und ISO/IEC 27001 aufweisen. In einem aufwändigen manuellen Prozess haben wir dann alle per Web Mining identifizierten Unternehmen kategorisiert, die einen ISO/IEC 27001-Verweis auf ihrer Website haben. Unter den von uns identifizierten ISO/IEC 27001-zertifizierten Unternehmen befinden sich auffallend viele Unternehmen aus der IKT-Branche (47 % gegenüber den ca. 4 % in der gesamten Unternehmensstichprobe). Zertifizierte Unternehmen sind signifikant größer als nicht-zertifizierte Unternehmen und haben auch eine signifikant höhere Innovationswahrscheinlichkeit (basierend auf dem von uns entwickelten webbasierten Innovationsindikator) als nicht-zertifizierte Unternehmen. Während zertifizierte Unternehmen signifikant jünger sind als nicht-zertifizierte Unternehmen in der Gesamtstichprobe, sind zertifizierte IKT-Unternehmen signifikant älter als nicht-zertifizierte IKT-Unternehmen.

Wir kamen zu dem Schluss, dass unser Ansatz einen ersten erfolgreichen Versuch darstellt, Web-Mining-basierte Informationen zur Diffusion von Standards zu generieren. In der Diskussion unserer Ergebnisse haben wir die Bedeutung von zertifiziertem IT-Personal und zertifizierten Geschäftspartnern hervorgehoben, die wir in unserer Analyse gefunden haben. Wir haben auch verschiedene Implikationen für Manager, politische Entscheidungsträger und Standardisierungsbehörden identifiziert, die sich aus unserer Studie ableiten lassen. So haben wir beispielsweise die Möglichkeit eines politischen Anstoßes zur Umsetzung von ISO/IEC 27001-Maßnahmen in kleineren Unternehmen diskutiert, ohne dass diese tatsächlich eine Zertifizierung erlangen müssen, was erhebliche Kostenvorteile verspricht. Zusammenfassend hat unsere Studie gezeigt, dass Web Mining eingesetzt werden kann, um den aktuellen Stand der Diffusion von Standards in der Unternehmenspopulation zu ermitteln und die Treiber dieser Diffusion zu untersuchen. Im Gegensatz zu traditionellen Erhebungsmethoden ist unser Ansatz vor allem wegen des Kostenaspekts attraktiv, der häufige und umfassende Erhebungen ermöglicht.

2.6 Wie innovative Unternehmen im Web Beziehungen knüpfen

In den bisher vorgestellten Arbeiten zum Thema Web Mining haben wir den relationalen Aspekt unserer Daten, nämlich das Vorhandensein von Hyperlinks zwischen Unternehmenswebsites, nur angeschnitten, aber noch nicht im Detail betrachtet. Unternehmen verlinken auf ihren Websites u.a. auf Partnerunternehmen, Kundenunternehmen (z.B. zum Zwecke des Referenzmarketings) oder auf Verbände, in denen sie Mitglied sind. Auch Hyperlinks rein technischer Natur sind möglich, etwa um externe Inhalte abzurufen und auf der eigenen Website darzustellen. Ein bestehender Hyperlink zwischen zwei Unternehmens-Websites kann also aus den unterschiedlichsten Gründen vom verlinkenden Unternehmen erstellt worden sein, ist aber immer mit einer bewussten Entscheidung verbunden, diesen Hyperlink zu erstellen und öffentlich zu machen. Dabei hat das verlinkende Unternehmen selbst keinen Einfluss darauf, wer auf die eigene Website verlinkt. Insofern ist schon die Existenz eines Hyperlinks, seine Ausrichtung und Gegenseitigkeit (Mutuality) eine Manifestation von Unternehmensbeziehungen, die es zu untersuchen lohnt.

Hyperlink-Netzwerke als digitale Manifestation von realen Beziehungen zwischen Unternehmen wurden bereits vor fast zwanzig Jahren erkannt. Bislang gab es jedoch kaum Studien in der Innovationsforschung, die Hyperlink-Netzwerke für umfassende empirische Untersuchungen genutzt haben. Gerade für die Erforschung mehrdimensionaler Nähe, wie sie Boschma (BOSCHMA 2005) vorgestellt hat, haben webbasierte relationale Daten Vorteile gegenüber traditionellen relationalen Daten aus Umfragen und Patenten. Boschma beschreibt die verschiedenen Dimensionen von Nähe wie folgt: "Kurz gesagt, gibt die kognitive Nähe das Ausmaß an, in dem zwei Organisationen die gleiche Wissensbasis teilen; die organisatorische Nähe das Ausmaß, in dem zwei Organisationen unter gemeinsamer hierarchischer Kontrolle stehen, die soziale Nähe das Ausmaß, in dem Mitglieder zweier Organisationen freundschaftliche Beziehungen haben, die institutionelle Nähe das Ausmaß, in dem zwei Organisationen unter den gleichen Institutionen operieren, und die geografische Nähe die physische Entfernung oder Reisezeit, die zwei Organisationen trennt." Bei der Analyse solcher Nähen besteht ein Vorteil von Webdaten darin, dass Hyperlinks im Gegensatz zu den aus Patentdaten abgeleiteten Verknüpfungen (z.B. über Co-Applikationen) auch nicht-formale Beziehungen abbilden können. Es ist auch zu erwarten, dass Webdaten Vorteile in Bezug auf Aktualität, Granularität und Abdeckung haben (siehe vorherige Kapitel).

Unser Papier "The Digital Layer: How Innovative Firms Relate on the Web", das ich zusammen mit Miriam Krüger, David Lenz und Bernd Resch geschrieben habe, befindet sich derzeit im Status eines Arbeitspapiers. In dem Papier entwickeln wir einen Ansatz zur Abbildung von Firmennetzwerken auf Basis von Hyperlink Mining, wobei zusätzlich eine Bewertung der Hyperlinks vorgenommen wird (wir unterscheiden z.B. zwischen geschäftlichen und nicht-geschäftlichen Beziehungen). Wir bezeichnen den von uns erstellten Datensatz als "Digital Layer", der sozusagen ein digitales Modell der Beziehungen zwischen Unternehmen/Organisationen darstellt. In unserer Studie besteht der Digital Layer aus ca. 500.000 Unternehmen in Deutschland, die durch ihre Website-Texte beschrieben werden und durch ca. sieben Millionen Hyperlinks miteinander verbunden sind. Zusätzlich verfügen wir für jedes Unternehmen über Unternehmenscharakteristika (Größe, Alter, Standort) aus der MUP-Datenbank und drei Innovationsindikatoren auf Unternehmensebene: den Produktinnovator-Status aus dem CIS (verfügbar für ca. 2.400 Unternehmen), den Patentinhaber-Status der Firmen und unsere prognostizierten Produktinnovator-Wahrscheinlichkeiten (beide verfügbar für alle Unternehmen).

Der auf diese Weise erstellte Digital Layer ist in Abbildung 7 dargestellt. Das obere Panel zeigt die Standorte der erfassten Unternehmen und die durchschnittlichen Produktinnovator-Wahrscheinlichkeiten pro Rasterzelle. Das mittlere Panel zeigt eine aggregierte Darstellung der etwa sieben Millionen Hyperlinks zwischen den Unternehmen. Das untere Panel zeigt das individuelle Netzwerk eines einzelnen Unternehmens (das Zentrum für Europäische Wirtschaftsforschung in Mannheim) für ganz Deutschland (links) und innerhalb seiner Heimatregion (rechts).

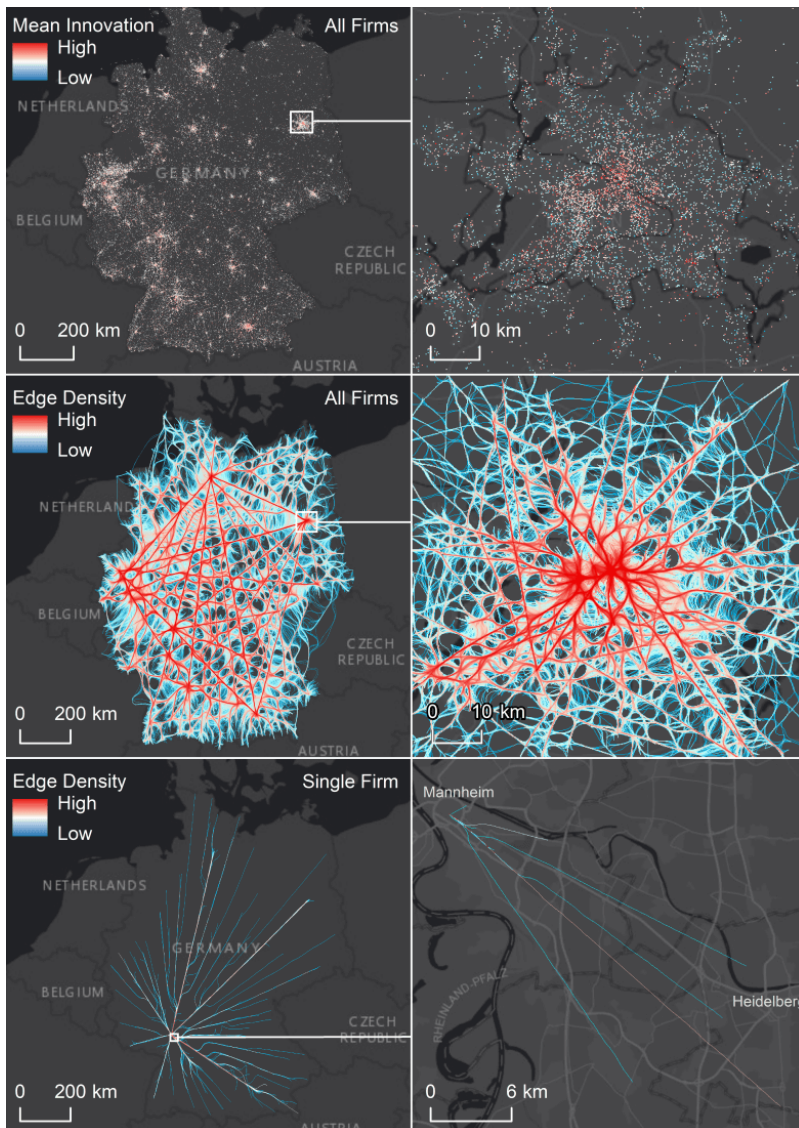


Abb. 7: Das Digital Layer für Deutschland. Aus Krüger et al. 2020.

Für jeden Knoten (Unternehmen) in unserem Netzwerk haben wir zunächst drei grundlegende Maße berechnet, die dessen Vernetzung beschreiben (Linkanzahl, mittlere Innovationsscores der Partnerunternehmen und die lokale Firmendichte). Zusätzlich berechnen wir für jedes Unternehmen im Digital Layer Maße für die durchschnittliche geografische, kognitive und organisatorische Nähe zu ihren (verlinkten) Partnern. Die geografische Nähe messen wir, indem wir die geografische Distanz zwischen den Standorten der verlinkten Unternehmen berechnen. Unser Maß für die kognitive Nähe soll Aufschluss darüber geben, ob verlinkte Unternehmen eine gemeinsame "Wissensbasis" teilen. Dazu berechnen wir die Ähnlichkeiten zwischen den Website-Texten der verknüpften Unternehmen mit Hilfe der Cosinus-Ähnlichkeit von tf-idf-vektorierten Texten. Wir approximieren die organisatorische Nähe, indem wir die Art der Beziehung (geschäftlich vs. nicht-geschäftlich) zwischen zwei verlinkten Unternehmen mit Hilfe eines Vorhersagemodells prognostizieren. Nicht-geschäftliche Beziehungen sind Beziehungen zwischen Firmen, die nicht direkt mit der Abwicklung von Geschäften miteinander verbunden sind und nicht-monetärer Natur sind. Geschäftsbeziehungen umfassen alle Hyperlinks zwischen Firmen, die miteinander Geschäfte machen oder gemacht haben. Anschließend berechneten wir die mittleren Proximity-Werte für jedes Unternehmen. Im Hauptteil der Arbeit untersuchten wir die Unterschiede zwischen innovativen und nicht-innovativen Firmen in Bezug auf ihre Vernetzungsmerkmale. Aus den Ergebnissen leiteten wir die folgenden Beobachtungen ab:

1. Innovative Unternehmen haben mehr Partner.

2. Innovative Unternehmen haben Partner, die innovativer sind.
3. Innovative Unternehmen haben überregionale Netzwerke (geografisch entfernte Partner).
4. Innovative Unternehmen pflegen eher Geschäftsbeziehungen.
5. Innovative Unternehmen verbinden sich eher mit Unternehmen mit ähnlicher Wissensbasis.
6. Innovative Unternehmen nutzen geographische Nähe, um kognitive Distanz zu vernetzten Partnern zu überwinden oder nutzen kognitive Nähe, um geographische Distanz zu ihren Partnern zu überwinden.

Zusammenfassend stellen wir fest, dass unser "Digital Layer"-Konzept ein interessanter Ansatz für umfassende empirische Studien zur Vernetzung von Unternehmen ist, der auch die Operationalisierung von Proximity-Maßen erlaubt, die mit traditionellen relationalen Daten nur schwer messbar waren. Besonders hervorgehoben haben wir die Möglichkeit, einzelne Branchen oder Regionen auf einer beliebigen geographischen Analyseebene zu analysieren. Für zukünftige Studien empfehlen wir insbesondere die kontinuierliche Erfassung von Webdaten, die auch eine Analyse des Digitalen Layers über einen längeren Zeitraum ermöglichen würde. Mikrogeografische Firm-to-Firm Knowledge Spillovers und die Diffusion von Technologie zwischen Firmen, Branchen und Regionen wären besonders interessante Forschungsbereiche.

2 Fazit

Rahmen dieser Dissertation wurde der Versuch unternommen, moderne GIScience-Methoden und -Daten mit einem Thema aus dem Bereich der Innovationsökonomie zu verbinden. Nach meinem Kenntnisstand stellt diese Arbeit einen ersten Versuch dar, eine umfassende, mikrogeographische Perspektive auf die Mechanismen hinter den Wechselwirkungen zwischen Unternehmen, Innovation und Standort zu liefern. Diese mikrogeographische Perspektive ist erst durch das Aufkommen von allgemein verfügbaren und detaillierten Geodaten, insbesondere von Volunteered Geographic Information, möglich geworden. In Bezug auf Unternehmen hingegen erwies sich die Datenlage als unzureichend, da traditionelle Innovationsindikatoren aufgrund ihrer unzureichenden Abdeckung und Granularität für diese Art der Analyse nicht geeignet sind. Dies veranlasste meine Mitautoren und mich dazu, einen neuartigen Ansatz zur Innovationsmessung auf Basis von Webdaten zu entwickeln, da es bisher keinen allgemein anwendbaren Ansatz zur Generierung von Innovationsindikatoren aus unstrukturierten Big Web Data gab. Mit unserer methodischen Arbeit haben wir ein Forschungsfeld erschlossen, das das Potenzial hat, auch außerhalb der Innovationsforschung einen relevanten Beitrag zur Generierung von Forschungsdaten sowie von Indikatoren für eine evidenzbasierte Politikgestaltung zu leisten. Die universelle Anwendbarkeit unseres Ansatzes ist auch die Basis für das akademische Spin-off-Startup *istari.ai*, bei dem ich einer der Mitgründer bin.

Literatur

- SMITH, DAVID M. (1981), *Industrial Location: An Economic Geographical Analysis*. 2nd. New York, Chichester, Brisbane, Toronto: John Wiley & Sons, S. 492.
- MIKOLOV, TOMAS ET AL. (2011), "Strategies for Training Large Scale Neural Network Language Models". In: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. DOI: 10.1109/ASRU.2011.6163930.
- BOSCHMA, RON (2005), "Proximity and Innovation: A Critical Assessment". In: *Regional Studies* 39.1, pp. 61–74. DOI: 10.1080/0034340052000320887.

Defining and assessing urban livability – A GIScience perspective

Anna KOVÁCS-GYÖRI

IDA Lab · Paris Lodron University of Salzburg · Jakob-Haringer-Straße 6 · 5020 Salzburg, Austria

E-Mail: anna.gyori@sbg.ac.at

1 Introduction

The term “livability” has become popular, mainly thanks to livability rankings and indices aiming to quantify urban quality (GIAP et al., 2014; OKULICZ-KOZARYN, 2013). However, “livability” – if interpreted appropriately – is more than just a statistical index or marketing tool. By facilitating the systematic and intersubjective evaluation of the person-environment relationship, livability can serve as a useful conceptual and analytical framework to assess and improve the quality of urban life and to investigate the influence of the urban environment on it. This evaluation should aim to measure specifically and systematically how well the built environment and the available services in a city fulfill the residents’ needs and expectations. In my doctoral thesis, I emphasized the relevance of livability assessment in urban quality improvement and also provided methodologies and insights to facilitate this assessment. As part of my research, I investigated and specified the conceptual background of livability, as well as designed and tested potential quantitative and qualitative analysis techniques to assess the livability of a place. By doing so, the representation of the person-environment relationship at different scales and contexts was of primary interest.

2 Summary of the presentation

2.1 Problem statement and objectives of the thesis

Although the intention to assess cities has a long history, the ubiquity of digital technology opened up new horizons also in this regard (BATTY, 2013). Due to this advancement, it is now at least technologically possible to analyze how people interact with the urban environment, however this analysis often focuses on rather quantitative approaches (CALABRESE et al., 2011; CROOKS et al., 2015; JACOBS-CRISONI et al., 2014; SAGL et al., 2012; THAKURIAH et al., 2017; WU et al., 2014). In recent years, there has been a shift to more qualitative aspects also observable, moving beyond the pure spatial or spatiotemporal analysis (COLLINS et al., 2013; FRANK et al., 2013; GIRARDIN et al., 2009; MITCHELL et al., 2013; QUERCIA et al., 2012; RESCH et al., 2016; SUN et al., 2015). Nevertheless, the question immediately arises: How can we evaluate urban quality, as the people perceive it, considering also the spatiotemporal settings and complexity of cities? The inclusion of qualitative aspects such as human perception or the preferences of the residents first require a firm theoretical concept, and as a follow-up a transferable assessment framework, that can be universally applied – at least to a certain degree – in various cities.

Although livability is a widely used term, especially when referring to urban environments, yet there is no general consensus on how to define it precisely (BLEČIĆ et al., 2013; KAMP et al., 2003). The existence of numerous similar and overlapping concepts (e.g., quality of life, well-being, sustainability) makes a consistent interpretation even more challenging (BROWN, 1975; KAMP et al., 2003; PACIONE, 2003; RUTH & FRANKLIN, 2014; VEENHOVEN, 2000; YOUNG & HERMANSON, 2012). Furthermore, existing urban quality (or livability) assessment approaches usually do not consider the person-environment relationship at different thematic and spatial scales, especially not in a transferable way. Their analyses often focus either on a city or a region (ANTOGNELLI & VIZZARI, 2018; CONTEH & OKTAY, 2016; SAILTUANGA, 2014) or provide frameworks for ranking cities through complex indices (GIAP et al., 2014; OKULICZ-KOZARYN, 2013). Works in the former group are usually quite thorough, but often not transferable; thereby their direct adaptation to other cities is limited. On the other hand, complex statistical indices for city ranking often ignore spatial aspects, especially regarding the intra-urban variance of the different factors, as well as, the varying socioeconomic background of the residents included in the analysis (CAPROTTI et al., 2017; COSTANZA et al., 2016).

As an approach to tackle these identified challenges of livability assessment, I had two objectives in my doctoral research. *Objective A* was more general, and encompassed the development of a conceptual framework for livability, also serving as a strong basis for the second objective. While in *Objective B*, I intended to consider various mixed-method analysis approaches to assess livability and its components in the urban environment.

- *Objective A*: Provide a conceptual framework for defining and assessing livability
- *Objective B*: Testing quantitative and qualitative geospatial analysis techniques to evaluate the quality of the person-environment relationship and its different aspects based on the conceptual livability framework

2.2 Methodology

During my doctoral studies, I published eight papers about the concept of livability (*Objective A*) and how the various aspects of the person-environmental relationship, as the conceptual interpretation of livability can be evaluated using GIS analysis techniques (*Objective B*). These aspects include among others the analysis of urban green (Kovacs-Györi et al., 2018) and large planned events (KOVÁCS-GYÖRI et al., 2018) using social media data, or assessing walkability by wearable biosensors (DÖRRZAPF et al., 2019). In my talk, I introduce the conceptual framework resulted from *Objective A* and present the key findings of an international livability survey, with respondents from eight different countries (KOVÁCS-GYÖRI et al., 2019; KOVÁCS-GYÖRI & CABRERA-BARONA, 2019).

To better consider the already mentioned high subjectivity of personal preferences and expectations in the person-environment relationship, our survey was designed to collect the view of residents in various cities on the interpretation of a livable environment. These livability factors were identified based on a conceptual framework (*Objective A*), and by asking residents we could estimate the relevance of each factor in perceiving higher satisfaction. The survey addressed five main livability categories such as sense of community, urban form, urban function, mobility and housing, as well as overall satisfaction. Overall satisfaction was interpreted as the self-reported livability of an area on a scale from 1 to 5, where 1 meant very low livability and 5 represented a very high satisfaction. The whole questionnaire consisted of 49 questions and was available in four languages (English, German, Spanish, Hungarian). Apart from the questions concerning demographic information, the response to each question reflected residential preference or self-reported evaluation. The questions in our survey addressed over a hundred different livability factors and not all of them have the equal importance in influencing people's satisfaction with living in a given city or a neighborhood. Therefore, we performed regression analysis to see which of these factors have the strongest statistical relationship with overall satisfaction (Figure 1). By performing this regression analysis, we reduced the number of key livability factors to 25, which were then used in further analysis steps to understand how the perception of livability is influenced by residential preferences (Figure 1).

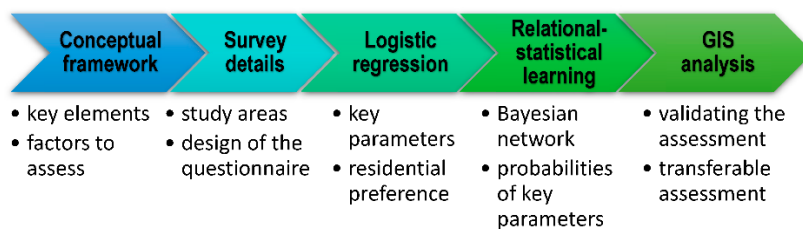


Fig. 1: Overview of the methodology for the global livability survey (KOVÁCS-GYÖRI et al., 2019)

After the regression analysis revealed the most important factors, we performed relational-statistical learning (also known as probabilistic modeling or Bayesian inference) to model the contribution of each livability factor to the overall satisfaction of a person, based on probability values. This method can help to discover how likely it was that if a factor was fulfilled (e.g., having trees along the street), a person would be satisfied with living in the city or that area. This modeling was based on a Bayesian network (a probabilistic graphical model that represents a set of variables and their conditional dependencies), and an algorithm calculated these probabilities using the data represented by the responses of the residents as input. We selected this method because it can be reliably applied to a relatively small sample size (440 responses), compared to other machine learning algorithms such as random forest. Furthermore, the algorithm that calculated these probability values was not a “black box”; the results can be derived and replicated by mathematical operations and these results were always the same, which could be a crucial characteristic in application for urban planning. An even more significant advantage of the method is the ability to provide these probability values not just for single factors or as a final value considering all parameters, but also the probabilities of each possible combination.

As a final step, GIS analysis was performed for Vienna as a study area to test whether the spatial analysis of selected livability variables can effectively represent overall livability. In other words, to see if we can model livability using only these parameters, without asking people directly about their preferences in details (for example in another city or in the future). First, we selected livability factors with spatial reference and the highest probability based on relational-statistical learning from our survey and collected the adequate data sets representing them from the open data portal of Vienna and Open Street Map (OSM). Next, we calculated the value of each parameter in every analysis unit (e.g. how many trees per kilometer along the street on average), and defined the overall probability of livability using the values from the relational-statistical learning. Finally, we compared the values from unique responses of the questionnaire and the values of these modeled intersubjective (spatial) livability parameters to see if the given person would be satisfied with living in the city.

2.3 The conceptual framework and the most important findings of the survey

Figure 3 shows the key elements and their relationship in livability (KOVÁCS-GYÖRI et al., 2019). The lower part represents the urban environment (i.e. actual state), which consists on the one hand of the built and natural environment along with the infrastructure, described as urban form. On the other hand, it also includes what this environment can provide for the citizens as urban functions, also reflecting how the people use this environment in light of the functionality. This

component of livability is more tangible, and spatial data for further analysis is available or can be most of the time acquired. At the same time, the upper part of the figure corresponds to the individual preferences (i.e. expectations) in livability and the person-environment relationship, consisting of human needs that the urban functions (or even urban form) need to satisfy and the personal values that define and influence the person's preferences. Due to the high level of subjectivity, individual preferences bring inherent limitation to livability assessment, as means or data to extract these preferences in an intersubjective way are often not available. Based on this framework, urban livability describes how well the environment can satisfy the citizens' needs and expectations, i.e. to what degree the actual state matches the goal state.

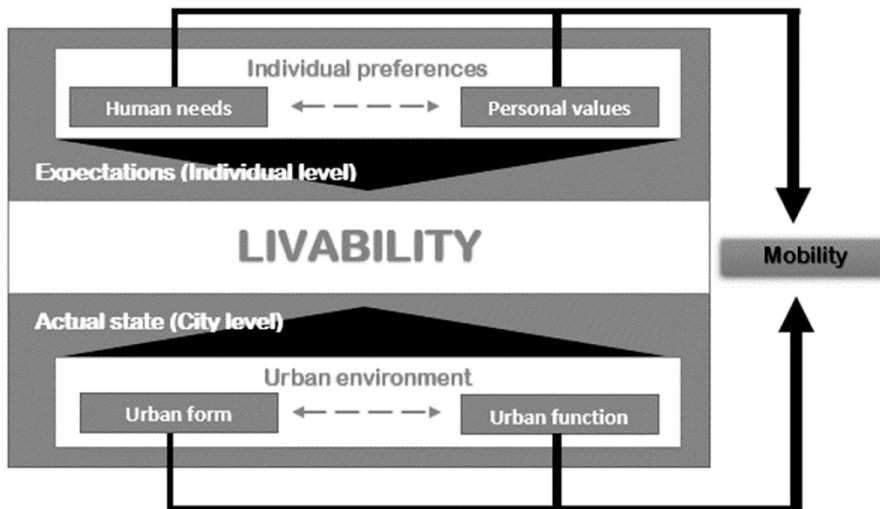


Fig. 3: The conceptual framework of livability (KOVACS-GYÖRI et al., 2019)

Mobility is highlighted in Figure 3, as the characteristics of mobility are interrelated with the key elements of the livability concept. Human needs and available urban functions define the destination that a person intends to reach, while personal values might affect the transportation mode. Moreover, urban form influences the actual route according to the infrastructure. This special role of mobility was considered in the survey as well, by investigating different aspects of mobility in the light of residential satisfaction.

As described above, in our survey, we collected responses from 440 people about over a hundred different potential livability parameters and selected the most significant ones using regression analysis (Figure 4). This resulted in a set of 25 parameters, for eight different thematic categories ranging from urban form to community or modifying factors.

Category (Acronym)	Variable	Estimate	p-Value
Modifying factors (MO)	Level of development (developing/developed)	1.71583	<0.001 ***
	Salary compared to basic needs	0.65158	0.02730 *
	Overall housing conditions ¹	2.20714	<0.001 ***
Community (CO)	Being a member of this community makes me feel good. ²	1.2085	<0.001 ***
	I trust people in this community. ²	0.9752	0.001618 **
	Most community members know me by face. ²	0.9979	0.008569 **
	Members of this community care about each other. ²	1.1183	0.008702 **
Urban form (FO)	Buildings in a street with similar style (architectural design). ³	-0.84488	0.089097 ·
	Street furniture (benches or chairs to sit, dustbins, shade, drinking fountains, etc.) ⁴	0.70985	0.089852 ·
	Trees along the streets. ⁴	1.16364	0.005797 **
	Overall quality of the built environment. ¹	1.47969	<0.001 ***
Transportation mode (TM) ⁷	Going to work by bicycle. ⁵	1.80005	0.01753 *
	Going to shop by public transport. ⁵	-0.81267	0.03276 *
	Going to shop by bicycle. ⁵	-2.13342	0.00474 **
Perceived mobility (PM) ⁸	Perceived cycling and walking safety. ⁶	0.5308	0.06222.
	Perceived cycling and walking quality. ¹	0.9728	0.00109 **
	Perceived overall transportation quality. ¹	1.2	<0.001 ***
Car (CA) ⁹	Car is the only way to reach the person's destination. ²	-0.81582	0.068534.
	Car is used because the person's destinations are too far. ²	0.98735	0.021055 *
Mobility infrastructure (M) ¹⁰	Special transportation services. ³	-3.227676	0.0118 *
	Easy to read traffic signs. ⁴	2.116724	0.0448 *
	Enforced speed limits. ⁴	1.194553	0.0372 *
Function & needs (FU, N)	I would let my children to walk/cycle alone. ²	0.8991	0.00227 **
	There is a lot of crime in the neighborhood. ²	-0.8912	0.01426 *
	Needs are fulfilled by the available urban functions. ²	0.4674	0.040538 *

p-value: **** < 0.001, *** < 0.01, ** < 0.05, · < 0.1. ¹ Quality: Poor/Fair/Good/Very good/Excellent; ² How well represents: Not at all/Somewhat/Mostly/Completely; ³ Importance: Not at all/Somewhat/ Important/Extremely important; ⁴ Binary: Have/Not have, Yes/No; ⁵ Frequency: Every day/Weekly/Monthly/Yearly/Never; ⁶ Safety: Very dangerous/Dangerous/Safe/Very safe; ⁷ What transportation mode (and how frequently) the people use for going to work, going shopping, or to reach their destination for free-time activities? ⁸ How do people perceive overall mobility conditions in terms of safety and quality for walking and cycling, or the transportation in general? ⁹ What are the reasons for a person to use a car as their main transportation mode (every day or almost every day)? ¹⁰ What factors or characteristics are the most relevant factors according to the people for the mobility infrastructure (e.g., road network, transportation services)?

Fig. 4: Statistically significant livability parameters based on the results of the regression analysis (KOVACS-GYÖRI et al., 2019)

In the next step, by using relational-statistical learning, we could identify the probability of higher perceived livability not just for each thematic groups (Figure 5) or individual parameters, but also for scenarios when parameters from a thematic group are only partially fulfilled. According to these probability values, urban form parameters have the highest contribution to the perception of higher livability (0.91935), if all conditions are satisfied. These conditions are:

- Buildings in a street with similar architectural style is important for the person
- Availability of street furniture (seating, dustbins, shade, drinking fountains, etc.)
- There are trees along the streets
- Good overall quality of the built environment (personal perception)

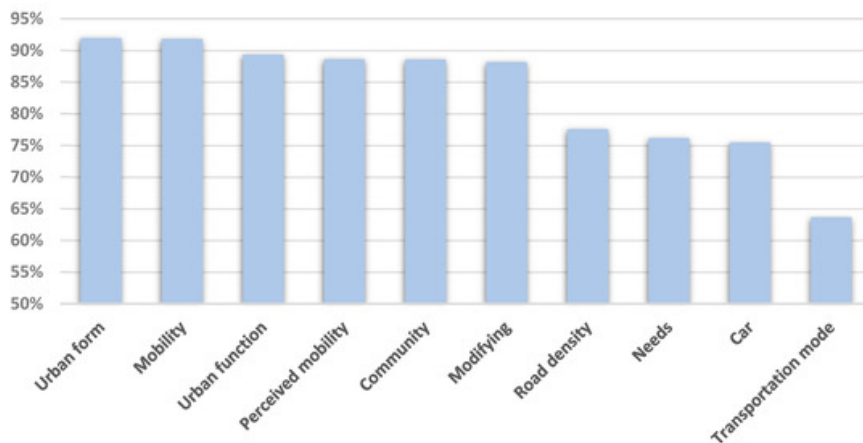


Fig. 5: Probability of each thematic group of representing high satisfaction with living in a given city (KOVACS-GYÖRI et al., 2019)

As the last step of the analysis, we modeled livability in Vienna using the identified urban form factors (trees, street furniture, overall perception of the built environment). Using the results of the relational-statistical learning, we assigned a probability value for each combination of the three factors based on whether they are fulfilled or not. This means how high is the probability of high satisfaction for a person in a given area within Vienna if one, two or all the three parameters are fulfilled. Overall, the prediction yielded an 80% accuracy for Vienna. This is a remarkable result for using only two calculated and one self-reported value, compared to the original set of over a hundred different livability parameters. A value of 80% can also be considered reliable because there will always be a segment of livability, which is highly subjective and cannot be grasped by intersubjective measures.

2.4 Conclusion

Living in cities certainly has its challenges even nowadays, but they also provide evident benefits by fulfilling the needs and expectations of the residents. The quality of life of urban dwellers mainly depends on the ability of the urban systems to fulfill these needs and expectations. In my work, I intended to illustrate that livability is a suitable and useful concept, while GIS supports the required analysis techniques to address these qualitative aspects of urban life in a systematic and intersubjective way. To achieve this goal, a firm conceptual framework had to be defined, as – although the concept of livability is popular and often referred to – the key elements of livability, their interrelationship and assessment options were not available in a single study yet. A conceptual framework is a crucial starting point for further, more detailed systematic assessment, and also essential for these assessment approaches to be transferable, as existing assessment approaches often only focus on specific study areas or use cases.

Beyond the introduction of a new conceptual framework, the main advancement of the thesis was the definition of assessment parameters considering the person-environment relationship on an intra-urban scale, instead of merely statistical indices, aggregated on city level, that are frequent in existing livability assessment approaches. Livability as a concept is able to consider the qualitative aspects of urban life (such as residential preference and satisfaction) in an intersubjective way by considering the person-environment relationship, however the inherent subjectivity due to the personal preferences is not and never will be negligible.

References

- BATTY, M. (2013), *The New Science of Cities*. MIT Press.
- CROOKS, A., PFOSE, D., JENKINS, A., CROITORU, A., STEFANIDIS, A., SMITH, D., KARAGIORGOU, S., EFENTAKIS, A., & LAMPRIANIDIS, G. (2015), Crowdsourcing urban form and function. *International Journal of Geographical Information Science*, 29(5), 720–741. <https://doi.org/10.1080/13658816.2014.977905>
- DÓRRZAPF, L., KOVÁCS-GYÖRI, A., RESCH, B., & ZEILE, P. (2019), Defining and assessing walkability: a concept for an integrated approach using surveys, biosensors and geospatial analysis. *Urban Development Issues*. <https://doi.org/10.2478/udi-2019-0008>
- GIAP, T. K., THYE, W. W., & AW, G. (2014), *A new approach to measuring the liveability of cities : the Global Liveable Cities Index*. 11(2), 176–196.
- GIRARDIN, F., VACCARI, A., GERBER, A., BIDERMAN, A., & RATTI, C. (2009), Quantifying urban attractiveness from the distribution and density of digital footprints. *International Journal of Spatial Data Infrastructures Research*, 4, 175–200. <https://doi.org/10.2902/1725-0463.2009.04.art10>
- KAMP, I. VAN, LEIDELMEIJER, K., & MARSMAN, G. (2003), Urban environmental quality and human well-being: Towards a conceptual framework and demarcation of concepts; a literature study. *Landscape and Urban Planning*, 65(July 2015), 5–18. [https://doi.org/10.1016/S0169-2046\(02\)00232-3](https://doi.org/10.1016/S0169-2046(02)00232-3)
- KOVACS-GYÖRI, A., CABRERA-BARONA, P., RESCH, B., & MEHA, M. (2019), Assessing and Representing Livability through the Analysis of Residential Preference. *Sustainability*, 11(18), 4934. <https://www.mdpi.com/2071-1050/11/18/4934>

- KOVÁCS-GYÖRI, A., RISTEA, A., HAVAS, C., RESCH, B., & CABRERA-BARONA, P. (2018), #London2012: Towards citizen-contributed urban planning through sentiment analysis of twitter data. *Urban Planning*. <https://doi.org/10.17645/up.v3i1.1287>
- KOVACS-GYÖRI, A., RISTEA, A., KOLCSAR, R., RESCH, B., CRIVELLARI, A., & BLASCHKE, T. (2018), Beyond Spatial Proximity — Classifying Parks and Their Visitors in London Based on Spatiotemporal and Sentiment Analysis of Twitter Data. *ISPRS International Journal of Geo-Information*, 7(9), 378. <https://doi.org/10.3390/ijgi7090378>
- OKULICZ-KOZARYN, A. (2013), City Life: Rankings (Livability) Versus Perceptions (Satisfaction). *Social Indicators Research*, 110(2), 433–451. <https://doi.org/10.1007/s11205-011-9939-x>
- RESCH, B., SUMMA, A., ZEILE, P., & STRUBE, M. (2016), Citizen-Centric Urban Planning through Extracting Emotion Information from Twitter in an Interdisciplinary Space-Time-Linguistics Algorithm. *Urban Planning*, 1(2), 114. <https://doi.org/10.17645/up.v1i2.617>
- THAKURIAH, P. (VONU), TILAHUN, N. Y., & ZELLNER, M. (2017), Introduction to seeing cities through big data: Research, methods and applications in urban informatics. In *Springer Geography*. https://doi.org/10.1007/978-3-319-40902-3_1

Neue OGC-Standards für Geo Web Services

Ingo SIMONIS

Open Geospatial Consortium

E-Mail: ingo.simonis@geospatialresearch.de

1 Zusammenfassung

Das Open Geospatial Consortium (OGC) ist eine weltweite Plattform zur Erforschung von Infrastrukturlösungen für digitale raumzeitliche Fragestellungen. Der Fokus des über 530 Mitglieder aus allen Bereichen der Industrie, öffentlichen Verwaltung und Forschungseinrichtungen zählenden Konsortiums ist dabei klar definiert. Ziel ist es, Interoperabilität zwischen beteiligten Datendiensten und Anwendungen zu erreichen. Dabei wird der gesamte Datenverarbeitungsprozess betrachtet, bestehend aus Datenbereitstellung, Datenbeschreibung und -entdeckung, Datenverarbeitung und -analyse, sowie Datenvisualisierungs- und -entscheidungssystemen. Ein Großteil der derzeitigen Arbeiten konzentrieren sich auf die Ergänzung bestehender Service Oriented Architecture (SOA)-basierter Standards, etwa Web Map Service, Web Feature Service, oder Web Coverage Service, mit REST-basierten Lösungen, die die Möglichkeiten und Vorteile des Internets optimal ausnutzen. In diesem Zusammenhang entstehen derzeit Open API basierte Schnittstellenbeschreibungen, die es erlauben, digitale Ressourcen in unterschiedlichen Repräsentationen bereitzustellen. Derzeit sind APIs zum Austausch von Vektordaten (OGC API-Features), Katalogdaten (OGC API-Records), Umweltdaten (Environmental Data Retrieval API), Kacheln (OGC API-Tiles), Rasterdaten (OGC API-Coverages), Landkarten (OGC API-Maps), Kartenstilen (OGC API-Styles) und Routen (OGC API-Routes) in Arbeit. Diese APIs werden die bestehenden SOA-basierten Instanzen nicht von heute auf morgen ablösen. Beide Ansätze werden weiter gepflegt, wobei ein schrittweiser Übergang über die nächsten 10-15 Jahre zu erwarten ist. In dieser Zeit werden die OGC APIs, welche sich derzeit bezüglich des Datenmodells und der Abfragemöglichkeiten generisch zeigen, auf einzelne Nutzergruppen angepasst werden, um die spezifischen Anforderungen dieser Gruppen zu erfüllen. Dabei werden typische Prozessierungsschritte, angefangen von der Kalkulation statistischer Größen bis zur Einbindung komplexer KI-basierter Modelle zunehmend auf die Serverseite verlegt und durch entsprechende API-Aufrufe zur Verfügung gestellt. Der Nutzer wird dadurch erheblich entlastet und kann auch große Datenmengen dank cloudbasierter Skalierungsmöglichkeiten nutzen. Die Prozessierung wandert vom lokalen Rechner in die Cloud. Dank spezieller APIs können dabei beliebige Prozesse cloud-seitig auf die Daten angewendet werden.

Trends der BIM und GIS Standardisierung – Bericht von der Joint ISO/TC 59/SC 13 - ISO/TC 211 WG: GIS-BIM

Christian CLEMEN

Hochschule für Technik und Wirtschaft Dresden · Friedrich-List-Platz 1 · 01069 Dresden
E-Mail: christian.clemen@htw-dresden.de

1 Zusammenfassung

Der Vortrag gibt einen kritischen Überblick zu den aktuellen Standardisierungsprojekten im BIM Kontext. Kritisch, weil die aktuelle BIM-Standardisierung aus der Perspektive der Vermessung und Geoinformation betrachtet wird. Es werden die vorläufigen Ergebnisse eines „Technischen Berichtes“ der ISO/TC 59/SC 13 - ISO/TC 211 Joint Working Group JWG14 „GIS-BIM interoperability“ vorgestellt. Es wird damit aufgezeigt wie vielfältig die Standardisierungsbemühungen zu Datenstrukturen, Diensten, Inhalten und Prozessen sind, wenn es um die Interoperabilität von digitalen Zwillingen der gebauten Umwelt geht. Als Fazit sollen auch zukünftige Trends der GIS-BIM-Integration antizipiert werden.

Coverage-Standards: Aktuelles aus OGC, ISO und INSPIRE

Peter BAUMANN

Large-Scale Scientific Information Systems Research Group · Jacobs University Bremen · Campus Ring 12 · 28759 Bremen
E-Mail: p.baumann@jacobs-university.de

1 Einführung

Nie zuvor waren Rasterdaten so einfach und kostengünstig zu gewinnen, sowohl durch Sensoren (etwa an Bord von Satelliten oder Drohnen) als auch computergeneriert (etwa bei Wettervorhersagen und Klimamodellierung). Allerdings geht das nicht einher mit einer gleichermaßen Verfügbarkeit – insbesondere die Big Data Merkmale Variety und Veracity sind technologisch nicht gelöst: wie komme ich an die Pixel heran, welche zum Datei-Download in Millionen von Szenen in oftmals unhandlichen Dateigrößen (z.B. 100 GB), technisch herausfordernden Formaten (etwas SAFE und GRIB2) und mit schwer durchschaubaren Metadaten-Formaten und –Platzierungen angeboten werden? Woher weiss ich, welche Daten für meine spezielle Aufgabe geeignet sind hinsichtlich örtlicher und zeitlicher Verfügbarkeit, Auflösung, Daten-güte, etc.? Diese Herausforderungen bereiten für Nicht-Experten oftmals unüberwindbare Hindernisse und gestalten die Arbeit von Experten sehr umständlich und zeitaufwändig.

Basierend auf jahrzehntelanger Arbeit eines breiten Spektrums an Domänen-Experten haben OGC und ISO den Coverage-Standard ständig weiterentwickelt, mit besonderem Augenmerk auf ständige Kompatibilität. Dazu trägt insbesondere die saubere Trennung vom Datenmodell *Coverage Implementation Schema* (CIS) und Dienstemodell *Web Coverage Service* (WCS) bei. Heute bilden Coverages die allgemein anerkannte Basis, um Zugriff, Extraktion, Analyse, Fusion und Föderation auf Rasterdaten in interoperabler Weise zu beschreiben, selbst auf anderen Planeten (OOSTHOEK ET AL 2014). Damit haben Coverages eine neue Ära an Big Earth Data Diensten eingeleitet: anstelle von Umnemgen von Dateien wird eine Handvoll Datenwürfeln angeboten, welche in einheitlicher Weise Zugriff und Kombination auf 1-D Zeitreihen, 2-D Bilddaten, 3D x/y/t Bild-Zeitreihen und x/y/z geophysikalische Voxeldaten sowie 4-D x/y/z/t atmosphärische Daten bieten. Abb. 1 zeigt ein Kaleidoskop von Coverage-Anwendungen, hier realisiert mit der offiziellen OGC-Referenzimplementierung rasdaman, welche in umfangreichen Benchmarks als mehrere Größenordnungen schneller als weitere Tools erwiesen hat (BAUMANN et al 2021).

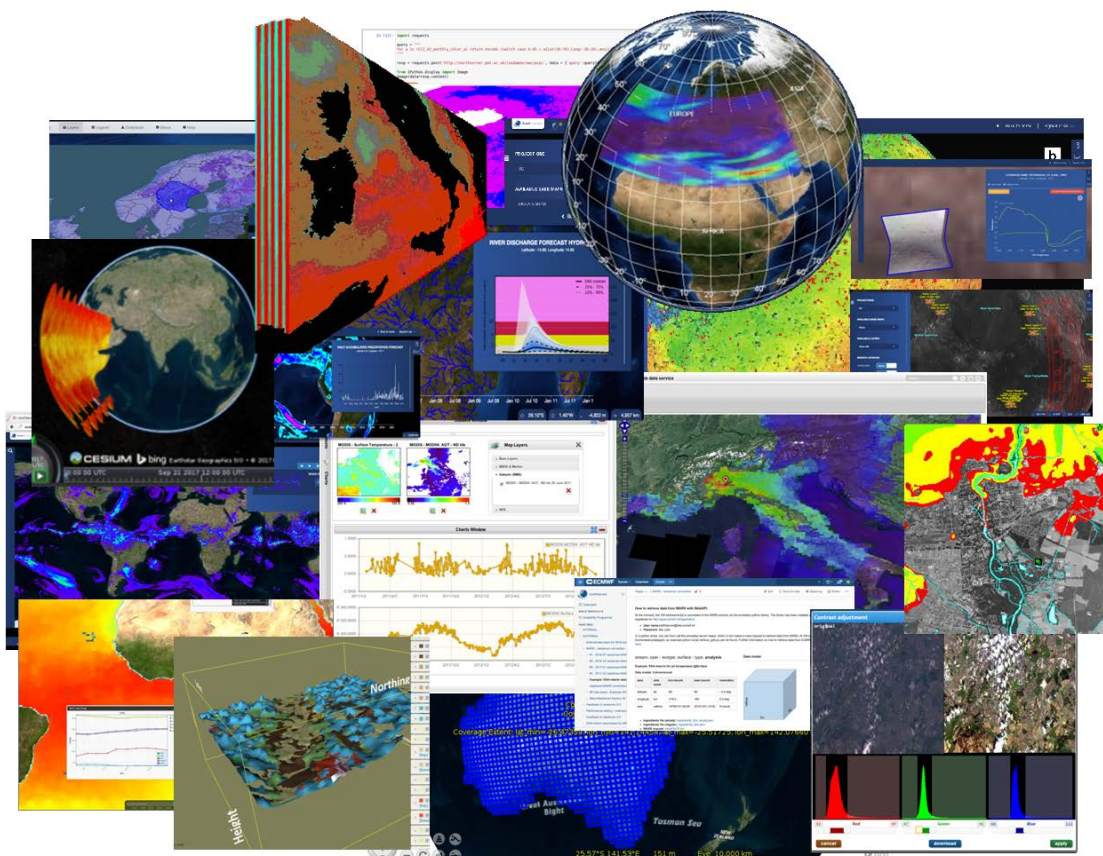


Abb. 1: Kaleidoskop von WCS/WCPS-Anwendungen, hier auf Basis von rasdaman (BAUMANN 2021)

Für Datenanbieter erlaubt das WCS-Ökosystem eine abgestufte Realisierung von Diensten, vom einfachen Basisdienst WCS-Core (Extraktion und Formatierung) bis zur flexiblen Datenanalyse mit *Web Coverage Processing Service* (WCPS). WCS/WCPS sind von einer ständig wachsenden Reihe von Open-Source und proprietären Werkzeugen unterstützt und haben sich auf Multi-Petabyte großen Datenarchiven bewiesen. Mit EarthServer existiert eine Multi-Petabyte Föderation von Copernicus-Datenwürfeln, welche Fusion bei völliger Ortstransparenz erlaubt.

Die beiden Organisationen ISO und OGC arbeiten eng zusammen – beispielsweise wurde OGC *Abstract Topic 6* (AT6) / *Coverages* früher bereits von ISO als 19123 übernommen, in jüngerer Zeit OGC Coverage Implementation Schema (CIS) als ISO 19123-2. Derzeit wird der überalterte 19123 Standard von ISO überarbeitet; das Ergebnis, 19123-1, befindet sich im fortgeschrittenen Stand der Abstimmung; OGC hat bereits angekündigt, die neue Version als neuen AT6 zu übernehmen.

INSPIRE hatte zuerst eine eigene Definition von Coverages kreiert, schwenkte dann jedoch über auf OGC Coverages – insbesondere im Zusammenhang mit den INSPIRE Download-Diensten, welche auf OGC Web Coverage Service (WCS) und damit dem OGC Coverage Modell beruhen. Damit hat INSPIRE eine sehr krispe Definition: INSPIRE Download Services = OGC CIS + OGC WCS Core + (optional) OGC WCPS.

WCS unterstützt eine Reihe von Dienste-Protokollen: GET/KVP, POST/XML und SOAP. In jüngerer Zeit treibt OGC OAPI-Coverages voran als eine annähernd REST-artige Schnittstelle, welche auf CIS und den wesentlichen WCS-Funktionalitäten basiert. In erster Näherung kann daher OAPI-Coverages als ein weiteres WCS-Protokoll aufgefaßt werden.

Im Bestreben, die technologischen Hürden bei Zugang und Auswertung von Daten zu verringern, ist der Begriff „Analysis-Ready Data“ (ARD) von USGS kreiert worden (ANDERSON et al 2019). Unter diesem – im ersten Ansatz recht generellen – Begriff werden seitdem Aktivitäten zusammengefasst, welche den Weg der Datennutzer vom Server in die Datenauswertung so schnell und effizient wie möglich machen. Ein erster Versuch, den Begriff speziell für Land-Daten präziser zu fassen, wurde mit CARD4L vorgestellt (CEOS 2016). Während der dort aufgestellte Forderungskatalog sinnvoll sind, erscheint er jedoch noch nicht sehr systematisiert und stringent motiviert. Die Coverage-Definition erscheint auch hier als sinnvolle Basis für eine Systematisierung des Feldes.

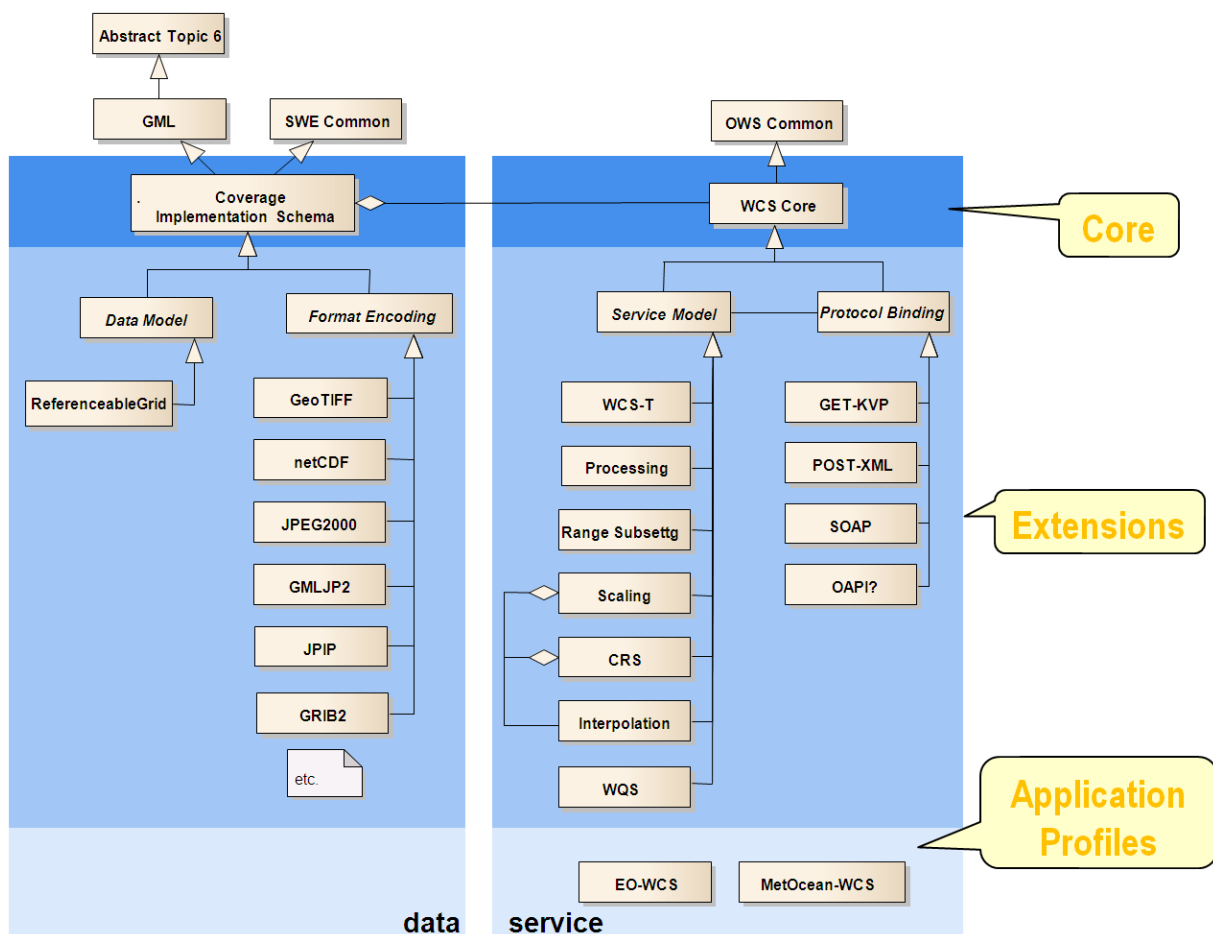


Abb. 2: Synopsis der OGC Coverage Daten- und Dienste-Standards (BAUMANN 2021)

2 Zusammenfassung des Vortragsinhalts

Im Vortrag stellen wir den aktuellen Stand der Standardisierung von Daten- und Dienstmodellen für Coverages vor, basierend auf eigenen Arbeiten in den Teams von OGC, ISO und INSPIRE. Neben einem kompakten Überblick über die Konzepte werden aktuelle Arbeiten vorgestellt, insbesondere die Zusammenarbeit zwischen ISO und OGC sowie die OGC-Arbeiten an OAPI-Coverages. Die Konzepte werden illustriert durch Live-Demos, von denen die meisten von Teilnehmern mit Internet-Anbindung selbst nachvollzogen und modifiziert werden können.

Literatur

- ANDERSON, C., LABAHN, S., HELDER, D., STENSAAS, G., ENGBRETSON, C., CRAWFORD, C., ... & BARNES, C. (2019, July), The US Geological Survey's Approach to Analysis Ready Data. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium (pp. 5541-5544). IEEE.
- BAUMANN, P. (2021), OGC Spatio-Temporal Coverage / Datacube Standards. <http://myogc.org/go/coveragesDWG>
- P. BAUMANN, D. MISEV, V. MERTICARIU, B.H. PHAM (2021), Array databases: concepts, standards, implementations. Springer Journal Big Data 8:28:2021. <https://doi.org/10.1186/s40537-020-00399-2>
- CEOS (2016), CARD4L (CEOS Analysis-Ready Data for Land). <https://ceos.org/ard/>
- N.N.(2021), EarthServer Datacube Federation. <https://earthserver.xyz>
- OOSTHOEK, J.H.P., FLAHAUT, J., ROSSI, A.P., BAUMANN, P., MISEV, D., CAMPALANI, P. & UNNITHAN, V. (2014), PlanetServer: Innovative approaches for the online analysis of hyperspectral satellite data from Mars. Advances in Space Research, 53(12)2014, pp. 1858 – 1871, <https://doi.org/10.1016/j.asr.2013.07.002>

Georeferencing

Štefan JAUD¹, Lee John GREGORY² and André BORRMANN¹

¹Chair of Computational Modeling and Simulation · Technical University of Munich · Arcisstraße 21 · 80333 Munich
E-Mail: {stefan.jaud}{andre.borrmann}@tum.de

²12D Solutions Pty Ltd. · PO Box 351 · Narrabeen, NSW, 2101 · Sydney · Australia

1 Introduction

Building Information Modelling (BIM) is being increasingly implemented within the Architecture, Engineering, and Construction (AEC) domain. Digital BIM models rich with geometric and semantic information about the asset are steadily replacing the conventional Computer-Aided Design (CAD) practises and getting implemented in every aspect of the very complex software and stakeholder landscape (JAUD et al. 2020).

One of the most important pieces of information about any physical AEC object is its location and orientation in the three-dimensional (3D) space. That is, the coordinate system (CS) of every object’s geometric data within the BIM model needs to be related to a defined geometric context (JAUD et al. 2021). This is usually done by providing the context of the element higher in the containment hierarchy, or – if such element does not exist – the ultimate context of the modelled asset: the Earth. Thus, every BIM object’s placement with regard to the Earth can be determined by tracing the connections in the placements’ dependency tree until the element with the reference to the ultimate context is found and consecutively applying the CSs’ placements of objects along the traced path.

The focus of this paper is the definition of the mentioned ultimate context, also called *georeferencing* in the context of BIM data modelling (JAUD et al. 2020). We take a closer look at the possible options and provide visual clues.

2 The ultimate context: the Earth

2.1 Placement definition

A geometric context defines the *placement* of object’s CS, by specifying (a) the position of CS’s origin within the geometric context, (b) the orientation of CS’s coordinate axes in the geometric context and (c) the scale factor between CS’s unit of measurement (UoM) and the UoM of the geometric context.

In the case of a Cartesian CS being used for both the geometric context’s CS as well as for the object’s CS, the parts (a)-(c) represent the *translation*, *rotation* and *scaling* of the object, respectively. The connection can be formalized using the well-known 7-parameter Helmert transformation. In other cases, the definitions of parts (a)-(c) may require more complex definitions, e.g. defining the position using a linear referencing method (JAUD et al. 2021).

2.2 Global vs. local

A usual confusion noted by the authors in the AEC domain are the notions of *global* and *local* contexts. These are interpreted differently by BIM experts and the surveyors and these collide at the construction site, as explained below.

A *surveyor’s global* context is defined to describe major parts (or all) of the Earth (keyword: coordinate reference system, CRS). There are two possible definitions of CRSs important to this paper that can be transformed between one another using complex algorithms (JAUD et al. 2020). (I) A *geocentric* CRS sets the CS’s origin in the Earth’s centre of mass, its vertical axis to coincide with Earth’s rotational axis and the horizontal axes to lie in the plane perpendicular to it. A reference meridian is chosen to define the first axis and the second is derived to span a right-handed Cartesian CS. (II) A *projected* CRS makes use of a map projection to piecewise “flatten” the curved surface of Earth. The CS’s origin and horizontal axes are defined within the projection algorithm. Vertical geometric information can be added using an independent vertical CRS to form a *composite* CRS, thus again defining a right-handed Cartesian CS.

A *surveyor’s local* CS is any other CS set in the context of any global CS, for example a construction site’s CS. This, in turn, represents the *BIM’s global* CS, i.e. the global context of all BIM objects placed in their corresponding *local* contexts as explained in Section 2.1. Therefore, the *ultimate* context of a BIM model is its placement within the surveyor’s global CRS, which is defined by the georeferencing metadata of a BIM model further explored in the next section.

2.3 The Earth as BIM’s ultimate context

We differentiate between three options for the definition of the ultimate context of a BIM model as presented in Figure 1 and Table 1. These differ in: (1) the type of CRS used for the definition of the context (see Section 2.2), as well as (2) the definition of the horizontal plane of the context’s CS. The latter can be based on two distinct possibilities: (i) using the map projection’s horizontal plane or (ii) using a topocentric CS spanned in a point on the Earth’s surface (hence the name

topo) with the direction of the gravity vector in that point defining the vertical direction. The horizontal plane is the plane perpendicular to that direction, passing through the topocentric CS's origin.

Option 1 uses a geocentric CRS and a topocentric CS to define the BIM's ultimate context. The conversion between the geocentric CRS and the placement of the BIM model (the topocentric CS) can be achieved using the aforementioned 7-parameter Helmert transformation, since both systems are Cartesian CSs. The limitation to this option is the fact that the curved surface of the Earth "sinks" below the defined horizontal plane when distancing from the CS's origin.

Option 2 uses a projected (or even a combined) CRS to represent the BIM's context. Consequently, the (irregular!) distortions introduced with the map projection remain intrinsic to the BIM model and must be considered accordingly. This is often overlooked in the practice due to lack of awareness and thus BIM dimensions wrongly treated as physical distances. This can become critical especially for long-stretched facilities such as mountain tunnels (WUNDERLICH 2020).

Option 3 is a combination of the former two: the definition of the context CS's origin follows option 2, while the definition of the horizontal plane stems from option 1. As such, the problem of "sinking" as intrinsic to option 1 remains; however, the definition of the CS' origin is more natural to practitioners since it is provided in a well-known CRS from option 2.

Table 1: Properties of the ultimate context options for a BIM model as presented in Figure 1.

Definition	Location	Horizontal plane	Horizontal scale factor
Option 1	Relative to Earth's centre	Perpendicular to gravity vector	1
Option 2	Relative to map projection's origin	Map projection's plane	Map projection's scale
Option 3	Relative to map projection's origin	Perpendicular to gravity vector	1

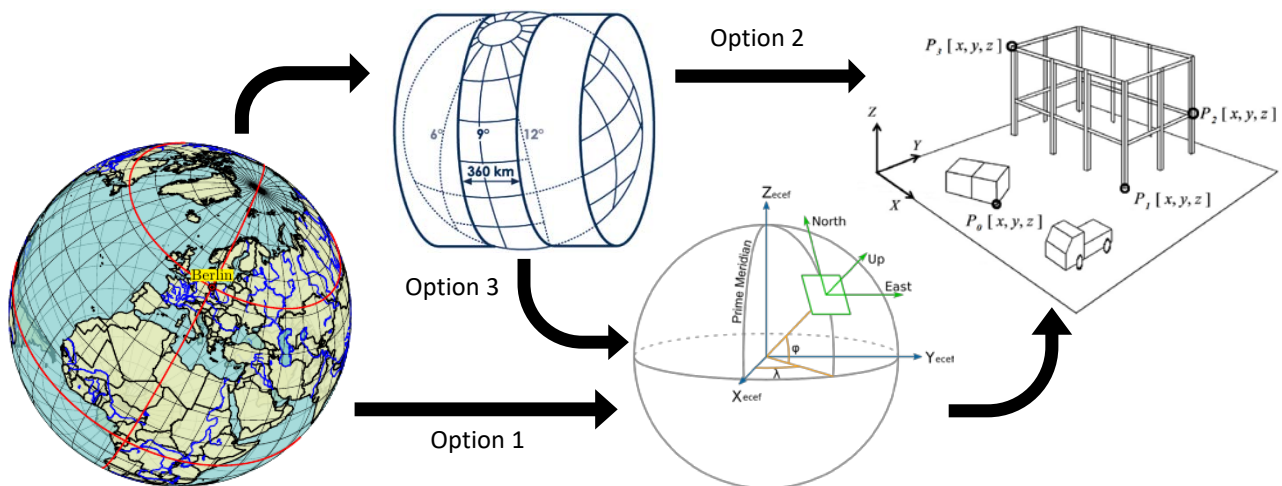


Fig. 1: A schematic representing the transformation path between the ultimate context (the Earth) and the global context of the BIM model (a construction site) for all three options.

3 Discussion

None of the three options described in Section 2.3 are perfect and be the general solution applicable to all cases. On the one hand, WUNDERLICH (2020) rightfully criticises interpretation of bigger BIM models' context according to option 3 and provides examples for the error from contemporary projects. He proposes option 1 as a possible solution, although he notes the difficulty in the meaning of the elevations within such a BIM model. On the other hand, JAUD et al. (2020) see option 2 as the only viable possibility for elongated infrastructure objects. We conclude that the topic needs a thorough analysis and a good understanding about the implications from all AEC stakeholders.

Literature

- JAUD, Š., DONAUBAUER, A., HEUNECKE, O. & BORRMANN, A. (2020), Georeferencing in the context of building information modelling. In: *Automation in Construction*, 118:103211.
- JAUD, Š., ESSER, S., WIKSTRÖM, L., MUHIČ, S., MIRTSCHIN, J. & BORRMANN, A. (2021), A critical analysis of linear placement in IFC models, In: *Proceedings of European Conference on Product and Process Modeling 2020-2021* [in press].
- WUNDERLICH, T. (2020), Misalignment – can 3D BIM overrule professional setting-out according to plane and height? In: *Contributions to International Conferences on Engineering Surveying*, A. Kopáček, P. Kyrinovič, J. Erdélyi, R. Paar, and A. Marenić, eds., pp. 3–12, Cham. Springer International Publishing.

BIM as Built

Robert KADEN und Katharina WECK

Fachgebiet Vermessung und Geoinformatik · Fachhochschule Erfurt · Altonaer Straße 25 · 99085 Erfurt
E-Mail: robert.kaden@fh-erfurt.de

1 Einführung ins Thema

Building Information Modeling (BIM) ist eine Arbeitsmethode, die im Rahmen der Digitalisierung der Baubranche immer öfter auch für den Bestand gefordert und angewandt wird. Grundlage einer BIM-konformen Planung für Sanierung, Umbau und Bewirtschaftung ist ein dreidimensionales, bauteilorientiertes Gebäudemodell des Bestandes. Doch inwiefern eignen sich die derzeit zur Verfügung stehenden BIM-Autorensoftware für die Modellierung von Bestandsgebäuden? Gibt es Unterschiede bei der Modellierung zum einen von Gebäuden mit regelmäßigen, rechtwinkligen und gradlinigen Strukturen im Bauhausstil und zum anderen von Gebäude mit unregelmäßigen, komplexen und krummlinigen Geometrien wie in der gotischen Architektur? Zu diesem Zweck wurden zunächst mit Hilfe eines 3D-Laserscanners (Leica BLK 360) eine moderne Stadtvilla im Bauhausstil sowie das Refektorium des mittelalterlichen Predigerklosters Erfurt erfasst und anschließend mit den BIM-Autorensoftware Allplan in Kombination mit dem Plug-In Scalypso und Autodesk Revit in Kombination mit der Erweiterung As-Built von Faro modelliert.

2 Zusammenfassung des Vortragsinhalts

2.1 BIM-Modellierung des Bestandes

Die BIM-konforme vermessungstechnische Erfassung und Modellierung des Gebäudebestandes stellt Geodäten vor Herausforderungen. Während bei den herkömmlichen Methoden der Modellierung in CAD die geometrische Repräsentation der Begrenzungs- bzw. Randflächen von Objekten im Vordergrund steht, werden in BIM die Bauteile einschließlich ihrer Eigenschaften in Form von regelmäßigen parametrisierten Volumenelementen modelliert (LEITFADEN GEODÄSIE UND BIM 2020). Das bedeutet, dass standardmäßig z. B. Wände idealisiert senkrecht stehen, die Außenflächen einer Wand parallel verlaufen, dass Bodenplatten und Decken horizontal platziert sind und sämtliche vordefinierte Bauteile genormten Standardmaßen entsprechen. Diese Art der Modellierung ist im Rahmen des Entwurfes und der Planung von Neubauten unproblematisch bzw. förderlich und bietet wesentliche Vorteile bei der Automatisierung der entsprechenden Prozesse. Bei der Erfassung des Bestandes mit geodätischen Verfahren, welche vornehmlich die beobachtbaren Oberflächen eines Bauwerkes erfassen, führt die BIM-konforme Modellierung zu Herausforderungen bei der Interpretation der Baukonstruktion und -materialien, bei der geometrischen Generalisierung unregelmäßiger Strukturen und der z. T. häufig notwendigen Erstellung individueller, generischer Bauteiltypen und -exemplare.

2.2 Testgebäude und Messmethode

Die gewählten Gebäude zur Untersuchung der Eignung der aktuellen BIM-Autorensoftware zur Modellierung von Bestandsgebäuden sind entsprechend unterschiedlich. Zum einen wurde ein Wohngebäude in Form einer Stadtvilla im Bauhausstil mit einer überwiegend regelmäßigen und rechtwinkligen Baukonstruktion und zum anderen ein unregelmäßiger und krummliniger gotischer Klosterbau gewählt.

Zur Erfassung der komplexen Geometrien wurde die Methode des 3D-Laserscannings gewählt, da diese sowohl bei einfachen als auch bei aufwändigen Raumstrukturen eine hohe Genauigkeit und Informationsdichte im Vergleich zum Handaufmaß oder Tachymetermessung liefert. Des Weiteren ist das Laserscanning im Gegensatz zur Photogrammetrie unabhängiger von äußeren Einflüssen wie Gegenlicht oder reflektierenden Oberflächen und kann somit jederzeit angewandt werden. Unter Berücksichtigung der Anforderungen bietet das Laserscanning die beste Lösung für die Bestandserfassung komplexer Raumstrukturen und gekrümmter Flächen und wurde deswegen im weiteren Verlauf der Arbeit für die Erfassung der Beispiellobjekte genutzt.



Abb. 1: links: Stadtvilla im Bauhausstil mit überwiegend regelmäßiger und rechtwinkliger Baukonstruktion und rechts: Gotischer Klosterbau mit unregelmäßigen und krummlinigen Konstruktionsmerkmalen

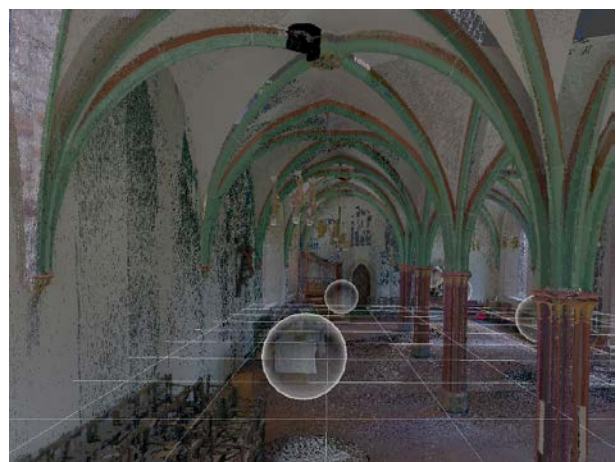
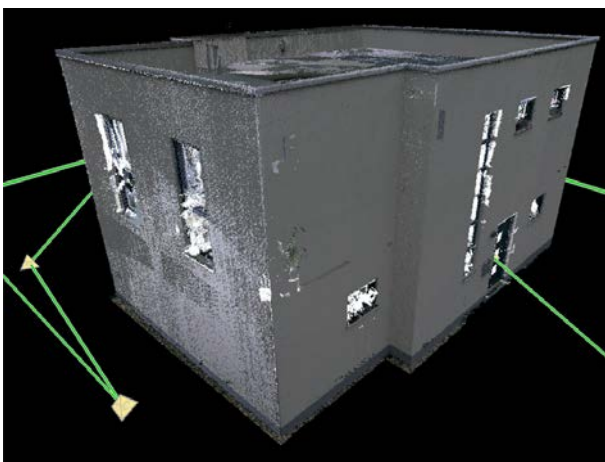


Abb. 2: links: Punktwolke der Stadtvilla visualisiert in Leica Cyclone Register 360 und rechts: Punktwolke des gotischen Klosterbaus visualisiert in Autodesk ReCap

2.3 BIM-konforme Modellierung der Testgebäude

Zur BIM-konformen Modellierung der beiden Testgebäude auf der Grundlage einer Punktwolke wurden zwei BIM-fähige Autorensoftware herangezogen und verglichen.

Allplan ist ursprünglich eine CAD-Software der Nemetschek Group und wurde in den letzten Jahren stetig weiterentwickelt, bis hin zu einer BIM-fähigen bauteilorientierten Software (ALLPLAN DEUTSCHLAND GMBH 2020). Allplan bietet zunächst keine native Möglichkeit, eine Punktwolke in ein Projekt zu importieren und direkt in dieser zu arbeiten. Jedoch erlaubt das Plug-In Scalypso die Analyse und Rekonstruktion terrestrischer und mobiler 3D-Laserscandaten in Allplan.

Revit (AUTODESK GMBH MÜNCHEN 2020) bietet im Gegensatz zu Allplan die Möglichkeit, eine Punktwolke direkt in ein Projekt zu laden und in dieser zu arbeiten. Allerdings ist ein 3D-Punktfang direkt in der Punktwolke in Revit nicht möglich. Des Weiteren kann die Punktwolke nicht im Familien-Editor zur Erstellung prototypischer Bauteile verarbeitet werden. Das Plug-In As-Built der Firma Faro ist eine Add-On-Anwendung, welche die Funktionen von Revit zur Punktwolkenverarbeitung erheblich erweitert. As-Built stellt Konstruktionshilfen sowie Werkzeuge zur Erstellung und Anpassung von Komponenten zur Verfügung. Zum einen kann dabei direkt in der in Revit eingelesenen Punktwolke gearbeitet werden, z. B. durch Funktionen wie das Einpassen von Wänden an die Punktwolke oder das automatische Erkennen von standardisierten Stahlprofilen. Zum anderen können mit As-Built Punktwolkenbereiche in den Familien-Editor eingefügt und durch exaktes Fangen von Punkten prototypische Bauteile, so genannte Familien, erzeugt werden.

Ziel des Projektes war es, auf Basis der Punktwolken jeweils ein BIM-Modell mit einem geometrischen Detaillierungsgrad LOG 300 und einem alphanummerischen Reifegrad LOI 200 zu erstellen. Die technische Gebäudeausstattung (TGA) sowie spezielle Statik wurde im Projekt nicht berücksichtigt.

Bei der Modellierung der Stadtvilla im Bauhausstil konnte in beiden BIM-Autorensoftware nahezu ausschließlich auf vordefinierte Bauteile zurückgegriffen werden. Es konnten zunächst sämtliche Bauteile des Rohbaus wie Wände, Attika,

Decken, Bodenplatte sowie der Fußbodenaufbau durch das Anpassen der Typen und Exemplare erstellt werden. Die Elemente Tür, Fenster und Fensterrahmen einschließlich Rollladenkästen und Rollladenschiene wurden teilweise mit Hilfe der Punktwolke im Familien-Editor individuell modelliert. In beiden BIM-Autorensoftware bzw. Plug-Ins stehen entsprechende Tools zur semiautomatischen Modellierung anhand der Punktwolkegeometrie zur Verfügung. Ein wesentlicher Unterschied bestand jedoch darin, dass im Gegensatz zu Allplan/Scalypso, in Revit/As-Built beispielsweise die Wandstärke automatisch durch die Angabe von Punkten auf beiden Bauteilseiten ermittelt werden kann. Diese Funktion ist in der Praxis jedoch auch nur bedingt anzuwenden, da in BIM-Modellen nicht selten nur standardisierte Rohbaumaße (z. B. Mauerwerk 17,5, 24, 36,5, etc.) modelliert werden und ein Putzabzug angenommen wird.



Abb. 3: BIM-Modell der Stadtvilla im Bauhausstil mit selektierter Wand im normierten Rohbaumaß (inkl. Putzabzug), visualisiert in der BIM-Autorensoftware Revit von Autodesk

Die Modellierung des Refektoriums des mittelalterlichen Predigerklosters in Erfurt gestaltete sich erwartungsgemäß wesentlich komplexer und dadurch zeitintensiver. Bis auf Teile der Außenmauern konnten keine vordefinierten Bauteile in den jeweiligen BIM-Autorensoftware Allplan und Revit genutzt werden. Die zur Unterstützung der Modellierung in den jeweiligen Plug-Ins Scalypso und As-Built zur Verfügung stehenden Tools konnten aus diesem Grund auch nur bedingt genutzt werden. Ein Großteil der Bauteile konnte lediglich händisch als generische Proxy-Elemente mit individueller Geometrie erzeugt werden. Besonders die Modellierung der Bauteile der Deckenkonstruktion Bündelpfeiler, Rippen und Kappen erwies sich als überaus zeitintensiv. Die Erstellung der Bündelpfeiler erfolgte auf der Basis von Schnitten entsprechend der jeweiligen Querschnitte aus der Punktwolke. Durch die Funktionen Extrusion bzw. Fahrweg-Sweep und Verschmelzen wurden dann die Pfeiler erstellt. Ähnlich wurden die Rippen der Kreuzgewölbe modelliert, wobei hier zunächst die Rippenverläufe als 3D-Linien aus der Punktwolke abgegriffen wurden. Durch Extrusion bzw. Fahrweg-Sweep des Rippenquerschnittes entlang dieser Linie wurden die Rippen erstellt. Die Kappen konnten in den beiden BIM-Autorensoftware nur auf unterschiedliche Weise modelliert werden. In Allplan/Scalypso wurden zunächst die entsprechenden 3D-Punkte der Kappe aus der Punktwolke in ein Digitales Geländemodell (DGM) überführt, welches im Anschluss in einen 3D-Körper umgewandelt wurde. In Revit/As-Built wurden die Kappen idealisiert, über zwei sich durchdringende Tonnengewölbe, modelliert. Erschwerend kam hinzu, dass ein Großteil der Bauteile wie Rippen und Kappen in jedem Gewölbeteil neu modelliert werden musste und nicht in Form eines Prototyps/einer Familie wiederverwendet werden konnte, da sich diese Elemente zwar zunächst augenscheinlich in der Deckenkonstruktion wiederholten, sich jedoch im Rahmen der angestrebten geometrischen Genauigkeit doch zu stark unterschieden.



Abb. 4: BIM-Modell des gotischen Klosterbaus mit Deckenkonstruktion bestehend aus Bündelpfeiler, Rippen und Kappe, visualisiert in der BIM-Autorensoftware Allplan von Nemetschek

2.4 Fazit

Die ausführlichen Testungen und Untersuchungen ergaben, dass beide Plug-Ins Scalypso und As-Built die jeweiligen BIM-Autorensoftware Allplan und Revit ertüchtigen, Punktwolken zu integrieren und zu interpretieren und BIM-konforme 3D-Gebäudemodelle von Bestandsgebäuden zu erstellen. Jedoch hat die Art der Baukonstruktion und die Bauepoche des zu erstellenden Gebäudes einen wesentlichen Einfluss auf den Modellierungsprozess und somit auf die Wirtschaftlichkeit. Regelmäßige und aus der heutigen Bauepoche stammende Bestandsgebäude wie die hier gezeigte Stadtvilla im Bauhausstil lassen sich unter weitestgehender Verwendung der nach aktuellen, normierten Standardmaßen vordefinierten Bauteile in verhältnismäßig kurzer Zeit (ca. 6h) in ein bauteilorientiertes BIM-Modell überführen. Im Gegensatz dazu gestaltet sich die Modellierung des unregelmäßigen und krummlinigen Refektoriums einer früheren Bauepoche bislang noch unverhältnismäßig (ca. 8d) in Bezug auf Kosten und Nutzen. Am aufwendigsten gestaltete sich die Erstellung der Deckenkonstruktion – also die Modellierung gewölbter bzw. gekrümmter Bauteile wie Rippen und Kappen. Dabei bieten beide BIM-Autorensoftware und Plug-Ins sowohl Vor- als auch Nachteile, sodass keine pauschale Aussage getroffen wird, welche Lösung für die Modellierung komplexer Strukturen die geeignetere Wahl darstellt. Wünschenswert wären hier Entwicklungen der Softwarehersteller, um die Erstellung von BIM-konformen Gebäudemodellen des Bestandes wirtschaftlicher zu gestalten und somit den Mehrwert eines BIM-Modells über den weiteren Lebenszyklus (Sanierung, Umbau, Bewirtschaftung) hinweg auch im Bestand nutzen zu können.

Literatur

- ALLPLAN DEUTSCHLAND GMBH (2020), Das Allplan Lösungsportfolio. URL: <https://www.allplan.com/de/>
 AUTODESK GMBH MÜNCHEN (2020), Revit. URL: <https://www.autodesk.de/products/revit>
 KADEN et. al. (2020), Leitfaden Geodäsie und BIM. URL: https://rundertischgis.de/publikationen/leitfaeden.html#a_bim_geo

CityGML 3.0 – Neue Konzepte und Perspektiven für einen Digitalen Lageplan in 3D

Tatjana KUTZNER und Andreas DONAUBAUER

Lehrstuhl für Geoinformatik · Technische Universität München · Arcisstr. 21 · 80333 München

E-Mail: {kutzner|andreas.donaubauer}@tum.de

1 Einführung ins Thema

Semantische 3D-Stadtmodelle werden häufig für eine exakte Darstellung von Stadt und Landschaft einschließlich der darin enthaltenen Objekte wie Gebäude, Brücken, Tunnel, Straßen und Vegetation verwendet. Ein weit verbreiteter Standard für die Modellierung, Speicherung und den Austausch dieser 3D-Stadtmodelle ist der vom Open Geospatial Consortium (OGC) herausgegebene internationale Standard CityGML (GROEGER et al. 2012). CityGML wird weltweit von vielen Städten und Ländern für die Verwaltung von 3D-Stadtmodellen und zur Harmonisierung von 3D-Gebäudedaten auf staatlicher Ebene genutzt.

Auf Grundlage der gewonnenen Erfahrungen bei der Nutzung von CityGML in vielfältigen Anwendungsbereichen sowie um CityGML für weitere Nutzergruppen und Anwendungsbereiche zu erschließen, wurde in den letzten Jahren die neue Version CityGML 3.0 (KOLBE et al. 2021) entwickelt, deren Verabschiedung durch das OGC für Mitte 2021 erwartet wird.

CityGML 3.0 bietet eine Vielzahl an Überarbeitungen und neue Funktionalitäten, welche die Verwendung von CityGML in Anwendungsbereichen wie Energie- und Umweltsimulationen, Stadtplanung, Verkehrsanalysen, autonomes Fahren, Smart Cities und Internet der Dinge (IoT) sowie die Interoperabilität mit anderen Standards im Geoinformationsbereich verbessern.

Der Vortrag gibt einen Überblick über die neuen und überarbeiteten Konzepte von CityGML 3.0 – mit Fokus auf neue Möglichkeiten im Bereich Planen und Bauen. Dies umfasst einerseits die verbesserte Interoperabilität mit Modellen aus dem Building Information Modeling (BIM), wie zum Beispiel die Industry Foundation Classes (IFC). Andererseits wird im Vortrag der Frage nachgegangen, wie ein zukünftiger Lageplan zum Baugesuch in 3D auf Basis von CityGML strukturiert sein könnte.

2 Zusammenfassung des Vortragsinhalts

2.1 Neue Konzepte in CityGML 3.0

Zentral in CityGML 3.0 ist das neu eingeführte Space-Konzept, das eine klare semantische Unterscheidung aller raumbezogenen Objekte in Raum (*Space*) und Raumgrenze (*SpaceBoundary*) vornimmt. Räume repräsentieren Objekte mit volumetrischer Ausdehnung (z.B. Gebäude, Gewässer, Bäume), Raumgrenzen dagegen Objekte mit flächenhafter Ausdehnung, welche Räume begrenzen und verbinden (z.B. Wand- und Dachflächen, Wasseroberfläche). Räume werden zudem danach unterschieden, ob sie durch physische oder virtuelle Raumgrenzen begrenzt sind (*PhysicalSpace* bzw. *LogicalSpace*) und ob sie in der städtischen Umwelt Raum belegen oder nicht (*OccupiedSpace* bzw. *UnoccupiedSpace*).

Das Level-of-Detail (LOD)-Konzept wurde dahingehend überarbeitet, dass Innenräume nun in jeder LOD-Stufe repräsentiert werden können (LÖWNER et al. 2016). Die Innen- und Außenrepräsentation von Stadtobjekten in unterschiedlichen LOD-Stufen kann kombiniert werden (z.B. die äußere Gebäudehülle in LOD1 mit dem Gebäudeinneren in LOD2 oder LOD3). Die Geometrien der Objekte können nun zusätzlich als 3D-Punktwolken repräsentiert werden.

Neu ist das Konzept der sogenannten anderen Konstruktionen (*OtherConstruction*). Damit können nun künstlich geschaffene Strukturen, die weder Gebäude noch Tunnel noch Brücken darstellen, besser repräsentiert werden wie beispielsweise große Kamine, Stadtmauern, Windturbinen oder Dämme.

Darüber hinaus ermöglicht CityGML 3.0 durch die neu eingeführten Module *Versioning* und *Dynamizer* die Repräsentation sich verändernder Eigenschaften von Stadtobjekten. Dazu zählen langsame Veränderungen, wie die Entwicklung von Städten oder Stadtmodellen im Laufe der Zeit, sowie regelmäßige und dynamische Veränderungen, wie Änderungen physikalischer Größen (z.B. Energiebedarf eines Gebäudes), Änderungen bezüglich der Geometrie sich bewegender Objekte oder Echtzeit-Sensorbeobachtungen. Zudem können Stadtobjekte mit (IoT-)Sensoren verknüpft werden

Die Entwicklung von CityGML 3.0 verfolgte einen modellbasierten Ansatz. Das konzeptuelle Modell wurde konform zu den relevanten ISO-Normen aus dem Geoinformationsbereich mit der Modellierungssprache Unified Modeling Language (UML) definiert (OGC 2020a). Anschließend wurden aus dem Modell automatisch GML-Schemata abgeleitet (OGC 2020b), weitere Formate wie Datenbankschema und JSON sollen später folgen.

2.2 Beitrag von CityGML 3.0 zur Interoperabilität mit BIM/IFC

Gebäudeelemente werden in IFC meist volumetrisch repräsentiert, zum Beispiel mittels Constructive Solid Geometry (CSG), wohingegen in CityGML die Gebäudeelemente bislang in Form ihrer Oberflächen dargestellt wurden (Boundary Representation, B-Rep). Bei der Überführung von IFC-Datensätzen nach CityGML 2.0 müssen deshalb die volumetrischen Objekte durch Extraktion der Oberflächen in B-Rep-Geometrien umgewandelt werden, was Probleme verursachen kann. Um die Überführung der volumetrischen IFC-Objekte nach CityGML zu erleichtern, wurde in CityGML 3.0 die neue Objektart *ConstructiveElement* eingeführt, welche gemäß dem Space-Konzept physische volumetrische Objekte repräsentiert. Die Extraktion der Oberflächen kann nun umgangen werden, indem die IFC-Objekte direkt auf die *ConstructiveElement*-Objekte in CityGML 3.0 abgebildet werden.

Durch die beiden neu eingeführten Objektarten Gebäudeeinheit (*BuildingUnit*) und Stockwerk (*Storey*) können Gebäude in logische Räume wie beispielsweise Wohnungen, öffentliche Gebäudebereiche oder thermische Zonen unterteilt werden. Zudem kann nun die IFC-Objektart *IfcBuildingStorey*, die ein Stockwerk in IFC-Datensätzen repräsentiert, explizit auf eine äquivalente Objektart in CityGML abgebildet werden. Das aktualisierte LOD-Konzept erlaubt darüber hinaus die Modellierung von Grundrissplänen für einzelne Stockwerke in LOD0 (KONDE et al. 2018).

Türen und Fenster können nun sowohl als volumetrische Objekte zum Füllen von Öffnungen in Bauwerken (Objektarten *Door* und *Window*) als auch als flächenhafte Objekte zum Verschließen der Öffnungen (Objektarten *DoorSurface* und *WindowSurface*) definiert werden. Dadurch können die volumetrischen IFC-Objekte *IfcDoor* und *IfcWindow* in entsprechende volumetrische CityGML-Objekte überführt werden. Ebenso können Innenräume, die in IFC durch die Objektart *IfcSpace* semantisch als physisch unbelegter Raum repräsentiert sind, in die Objektart *BuildingRoom* abgebildet werden, welches gemäß dem Space-Konzept ein äquivalentes Konzept in CityGML repräsentiert.

Durch die in CityGML 3.0 neu eingeführte Objektart *CityObjectRelation* können nun – wie in IFC – topologische, geometrische und semantische Beziehungen zwischen Gebäudeelementen (und auch zwischen anderen Stadtobjekten) explizit in CityGML beschrieben und die in IFC modellierten Beziehungen nach CityGML überführt werden.

2.3 Digitaler Lageplan in 3D als CityGML 3.0 Application Domain Extension

Am Anfang eines Baugenehmigungsverfahrens steht die Erstellung eines Lageplans zum Baugesuch, der nach den heutigen Bestimmungen nach den Bauvorlagenverordnungen der Länder immer noch als zweidimensionaler Plan einzureichen ist, obgleich heute, auch im Zuge der Einführung von BIM, das Planen und Entwerfen mehr und mehr mittels 3D-Werkzeugen geschieht. Gemeinsam mit dem Bund der Öffentlich bestellten Vermessungsingenieure (BDVI) untersucht die Technische Universität München wie ein digitaler Lageplan in 3D beschaffen sein muss, um die angegliederten Prozesse rund um das Planen und Genehmigen von Gebäuden bestmöglich zu unterstützen. Im Rahmen dieser Untersuchungen wurde von MÜLLER (2021) ein Informationsmodell für einen 3D-Lageplan in Form einer CityGML 3.0 Application Domain Extension (ADE) entwickelt, wobei das Space-Konzept eine zentrale Rolle einnimmt. So werden die bereits in CityGML 3.0 vordefinierten Klassen für topographische Objekte genutzt, um Gebäude, Bäume, Straßen etc. als *PhysicalSpace* zu repräsentieren. Objekte aus Bereichen wie Liegenschaftskataster und Bauleitplanung sowie Abstandsflächen werden in der CityGML ADE als *LogicalSpace* repräsentiert und soweit vorhanden aus existierenden Standards wie AAA und XPlanung entlehnt. Anders als in den genannten Standards können die Objekte jedoch in der CityGML ADE dreidimensional repräsentiert werden. So können künftig beispielsweise Abstandsflächen nicht mehr nur in Form zweidimensionaler Polygone, sondern in Form der von Bebauung freizuhaltenden dreidimensionalen Räume visualisiert und analysiert werden. Die CityGML-Objektart *CityObjectRelation* wird genutzt, um Abstandsflächen mit den Objekten (beispielsweise Wandobjekte oder Balkone) zu verknüpfen, aus denen die jeweilige Abstandsfläche resultiert. Abstandsflächenberechnungen werden dadurch transparent und sind besonders bei komplexen Bausituationen leichter nachzuvollziehen. Die folgende Abbildung zeigt die grafische Darstellung eines 3D-Lageplans.

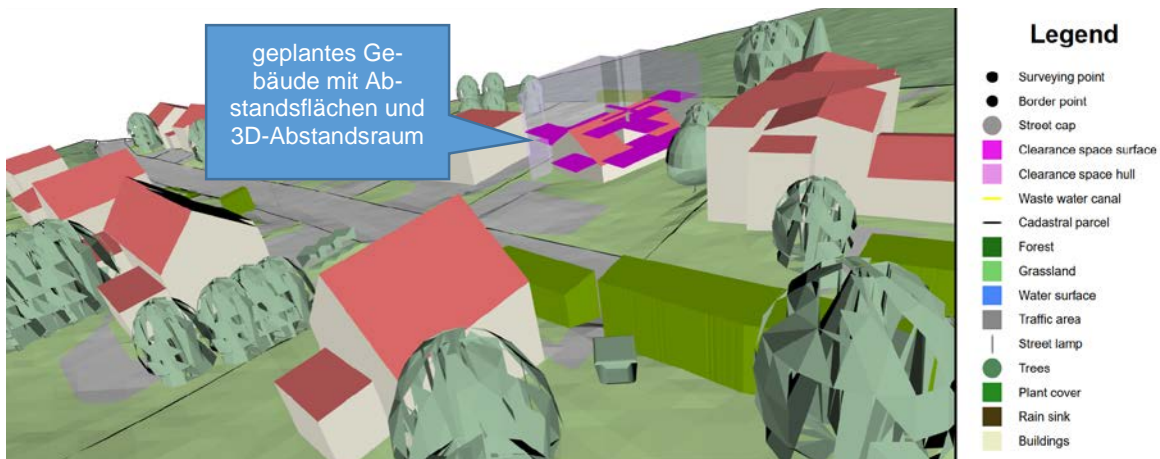


Abb. 1: Beispiel 3D-Lageplan zum Baugesuch (MÜLLER 2021)

Literatur

- GRÖGER, G., KOLBE, T. H., NAGEL, C., HÄFELE, K.-H. (2012), OGC City Geography Markup Language (CityGML) Encoding Standard Version 2.0.0. OGC Document 12-019. Open Geospatial Consortium.
- KOLBE, T. H., KUTZNER, T., SMITH, C. S., NAGEL, C., ROENSDORF, C., HEAZEL, C. (2021), OGC City Geography Markup Language (CityGML) Part 1: Conceptual Model Standard, Open Geospatial Consortium, noch nicht veröffentlicht.
- Konde, A., Tauscher, H., Biljecki, F., Crawford, J. (2018), Floor plans in CityGML. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences IV-4/W6:25-32.
- KUTZNER, T., CHATURVEDI, K., KOLBE, T. H. (2020), CityGML 3.0: New Functions Open Up New Applications. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 88, 43-61.
- LÖWNER, M.-O., GRÖGER, G., BENNER, J., BILJECKI, F., NAGEL, C. (2016), Proposal for a new LOD and multi-representation concept for CityGML. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences IV-2/W1:3-12.
- MÜLLER, F. (2021), Entwicklung einer CityGML Application Domain Extension: Der Amtliche Lageplan in 3D. Master Thesis am Lehrstuhl für Geoinformatik der Technischen Universität München.
- OGC (2020a): CityGML 3.0 Conceptual Model. <https://github.com/opengeospatial/CityGML-3.0CM>, Zugriff 2.3.2021.
- OGC (2020b): CityGML 3.0 Encodings. <https://github.com/opengeospatial/CityGML-3.0Encodings>, Zugriff 2.3.2021.

Next Steps for Big Geospatial Data

Martin WERNER

Department Aerospace & Geodesy, Professorship Big Geospatial Data Management · Technische Universität München · Willy-Messerschmitt-Str. 1 · 82024 Taufkirchen/Ottobrunn
E-Mail: martin.werner@tum.de

1 Introduction

Big Geospatial Data is generated every day and by everyone. For example, the use of smartphones generates large and complex traces of information about individuals and is collected by different entities such as the mobile network operators for understanding, monitoring, operating, and optimizing their cellular network, by search engines and other apps to place better services both in the main area of operation (e.g., better search result through spatial information) as well as in secondary areas such as personalized advertisements. In addition to this, we are surrounded by a multitude of sensor systems generating additional spatio-temporal signals. For example, satellites are equipped with multispectral cameras, high-resolution cameras, or synthetic aperture radar and give us a global, frequently updated collection of spatial observations. Furthermore, smart cities try to increase their efficiency (e.g. in terms of mobility, logistics, transport, etc.) by real-time observation linked with real-time control and management turning megacities into more sustainable places. Autonomous cars (and other robots) need to observe, track, understand, and communicate their knowledge about the surroundings and, therefore, are equipped with enormous amounts of sensors. Our governments also provide more and more reference data free of charge in various modalities, accuracies, resolutions, and formats. Finally, social networks with location-based components including Flickr, Twitter, or Facebook are creating huge traces of spatiotemporal information worth exploring.

These data sources altogether constitute a collection of big data. Big data, in this context, is understood as collections of data that are exceptional at least in the dimensions of the 3V definition of big data: volume, velocity, and variety. Spatial datasets consume large volume, especially raster data, the velocity in which data is generated is usually very high, and the variety in terms of syntactical and semantic representation of spatiotemporal data is unprecedented.

Though big data has been hyped in the last decade and is currently in a phase of consolidation, many open research issues remain. These range from communication (“how do we get the data to the right place for processing and decision making”), privacy (“how do we make sure that individual and private data is not compromised”) to overall ethical implications (“how do we make sure ethical boundaries stay”) both on an individual level as well as on a group level (e.g., discrimination of age, gender, race, etc.).

Tightly linked with this is the fact that spatial data is a very rich information source giving unprecedented and rising commercial power to those that have such data. People and western societies are comparably reluctant to support any innovations with a privacy impact, however, almost everyone (57.7 million individuals older than 14 years in Germany in 2019, according to (BITCOM RESEARCH 2019)) is actively using a smartphone leaving precise location traces with cellular operators, mobile device manufacturers and platforms, and, unfortunately, with advertisement networks as well.

Therefore, it is essential to develop integrated, efficient, and powerful big geospatial computing algorithms in order to support future development, to understand and mitigate current and future risk and to support sustainability goals such as the United Nations SDGs through the power of big geospatial data technology.

Currently, spatial data analysis takes place in silos often related to the application domain or the nature of the given data. For example, remote sensing is a domain in which raster representations of spatial knowledge dominate and time series are considered dense if they have a gap of less than two days. In contrast, trajectory computing and more generally human mobility analysis takes place in very different data landscapes in which time series are dense if they are collected at subsecond frequency and are largely represented as a set of timestamped points silently assuming that linear interpolation is performed during analysis. Despite an obvious potential, surprisingly few studies tried to bridge, for example, between these two domains of remote sensing and trajectory computing. One possible and plausible reason is the inability and difficulties to find and use data from other subdomains of spatial computing. To this end, the scientific data management community has largely adopted the FAIR principles: data collections shall become Findable, Accessible, Interoperable, and Reusable. The [Go-Fair Initiative] is a nice example of how this idea is taking shape. In this talk, we will discuss, how FAIR principles translate for spatial data and what additional research needs can be identified towards adapting these FAIR principles to spatial data.

2 Big Geospatial Data and the FAIR paradigm

In this section, we will quickly go through one possible implementation roadmap for FAIR data principles and discuss briefly how spatial data does fit or does not fit into this framework. In summary, we will see that FAIR is not as easy as

it might seem to be at a first glance. In fact, for geospatial data FAIRness involves the solution of long-standing and long-discussed research challenges such as entity linkage, content search, morally inoffensive data publication, and much more.

2.1 Findability

The findability of research data in the context of the FAIR paradigm is sometimes concretized as follows: “to be findable

- F1. (Meta)data are assigned a globally unique and eternally persistent identifier.
- F2. Data are described with rich metadata.
- F3. (Meta)data are registered or indexed in a searchable resource.
- F4. Metadata specify the data identifier.” (see GO-FAIR 2020)

While this is an interesting and convincing approach for atomic datasets that are generated once, consumed as a whole and truthfully characterized by their metadata, this approach falls short of spatial data, which is typically not characterized by its metadata alone. Therefore, we argue that findability is not supported from metadata alone. Instead, we need novel search algorithms for similarity search and entity linkage that can span different datasets or different versions of the same dataset.

2.2 Accessible

For FAIR datasets, once data is located through the Findability, it needs to be accessible. Similarly as above, the (GO-FAIR 2020) initiative spells out accessibility as follows:

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
- A1.1 The protocol is open, free, and universally implementable
- A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

With respect to these requirements, the spatial data community is exceptionally well-organized given all the OGC standards for data representation and transport via web services. In my opinion, however, this list falls short of enabling people to access in terms of guaranteed bandwidth. We need to incorporate the download durations into the overall assessment. When data is “theoretically” available, but cannot be accessed or processed within a reasonable timeframe, we did not reach a FAIR data management and open science is not possible. And this becomes a technical challenge when realizing that we are producing geospatial data at rates we currently cannot afford to communicate across the Internet.

2.3 Interoperable

We request data to be interoperable which is again largely supported by the decades of OGC activities. In spatial computing, we typically know how to represent data if the task is to share it among a wider group of applications. The (GO-FAIR 2020) initiative spells this out as

- I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (Meta)data use vocabularies that follow FAIR principles
- I3. (Meta)data include qualified references to other (meta)data

This sounds like a very convincing approach: everyone uses the same formal language and references each other's item. In reality, however, spatial data is different to catch in this way. First of all, references would be formed between each and every data object (“everything is related to everything else...”). But more importantly, it is very difficult to come up with a knowledge representation (for example an ontology or even just a labeling scheme) that is applicable on the whole globe. Spatial and cultural concept shifts make it very hard to formalize spatial data on a global, application-agnostic scale. In fact, we believe that interoperability will never be reached for spatial data as a whole. It is rather domain-specific and/or regional standards, ontologies, and conceptualizations that can successfully build a language for spatial knowledge representation.

2.4 Reusable

The aim of the FAIR principles is to allow data to be reused and the previous three aspects serve as prerequisites. The idea is that if we Findable, Accessible, and Interoperable data, it is just a small step to actually make use of this data. The (GO-FAIR 2020) initiative lists some additional aspects like metadata completeness, usage licenses, provenance, and community standards.

3 Conclusion

In this short abstract, but more in the associated presentation, we will have discussed concrete examples of how aspects of the FAIR paradigm are available in the spatial computing domain, but also where we fail to become fully FAIR and transparent. We try to highlight what we believe are the next open issues to resolve, the next important research questions to follow, and the next decisions that we need to take as a community in order to finally enable spatial computing communities to follow the FAIR principles or vice versa to refine the FAIR principles in a way such that spatial data and other ethically challenging datasets (e.g., from medicine, dual-use nature, etc.) can be supported in such a framework.

References

- BITCOM RESEARCH. (2019), Number of smartphone users in Germany from January 2009 to 2019 (in millions) [Graph]. In Statista. Retrieved March 08, 2021, from <https://www.statista.com/statistics/461801/number-of-smartphone-users-in-germany/>
- GO-FAIR INITIATIVE. (2020), FAIR Principles. <https://www.go-fair.org/fair-principles/>

Verteilte und Effiziente Prozessierung von Geodaten mit KI-Methoden

Matthes RIEKE^{1,2}, Benedikt GRÄLER^{1,3}, Simon JIRKA^{1,4}

¹52°North GmbH · Martin-Luther-King-Weg 24 · 48155 Münster

²E-Mail: m.rieke@52north.org

³E-Mail: b.graeler@52north.org

⁴E-Mail: jirka@52north.org

1 Geodaten für KI-Methoden

Methoden der Künstliche Intelligenz erfahren derzeit eine rasche Entwicklung und zunehmende Verwendung im Kontext von Geodaten aus den verschiedensten Fachdomänen. Dies geschieht gegenwärtig jedoch häufig in isolierten Lösungen. Viele Wissenschaftsfelder, in denen Geodaten eine zentrale Rolle spielen, haben den systemischen Einsatz moderner KI-Methoden noch nicht etabliert. Insbesondere besteht eine Diskrepanz zwischen den Anforderungen einer soliden und technisch fundierten Geodatenanalyse und der Anwendbarkeit moderner KI-Methoden wie Deep Learning in der Forschung. Zusätzlich werden effiziente Lösungen zur Vorbereitung und Bereitstellung der Geodaten hin zu einem “Analysis Ready Data”-Angebot benötigt (BAUMANN ET AL. 2019). Hierzu sind individuelle Lösungen zur Prozessierung und Publikation von Geodaten, die jedoch in breit anwendbaren Infrastrukturen eingebettet sind, unumgänglich.

Die Prozessierung von Geodaten bezieht sich auf die Verarbeitungsschritte, die durchgeführt werden, um Daten aus verschiedenen Quellen zu transformieren, zusammenzuführen, zu analysieren und zu visualisieren. Das wachsende Datenvolumen und die zunehmende Datenvielfalt sowie die Geschwindigkeit der Datenströme erfordern neue und fortschrittliche Methoden, Technologien und Architekturentwürfe, um diese Herausforderungen zu bewältigen (RIEKE ET AL. 2020). Die folgenden Aspekte spielen insbesondere eine wichtige Rolle:

- “Fit for purpose”: Bereitstellung von Geodaten in vorkonfektionierten Formaten und Referenzsystemen sowie über effiziente Zugriffsmethoden.
- Datenverarbeitung zur Erdbeobachtung: Effiziente und skalierbare Verarbeitung großer Mengen von EO-Daten.
- Cloud-Umgebungen (PaaS, IaaS, SaaS): Deployment-Strategien, Umgang mit Eingabedaten und Prozessierung-Ergebnissen.
- Standardisierung: Harmonisierung der Prozessierungsschnittstellen, Entwurf und Verwendung interoperabler Datenformate.
- Transparenz: Verbesserung der Reproduzierbarkeit, der Daten- und Metadatenqualität sowie der Auffindbarkeit von Prozessen deren Ergebnissen.

Im vorliegenden Beitrag sollen die Design-Konzepte aus zwei Forschungsprojekten, in denen die Verarbeitung von Geodaten mit KI-Methoden einen zentralen Baustein bilden, dargestellt werden

2 Infrastrukturen für die effiziente Prozessierung von Geodaten

2.1 KI-Methoden für Umweltdaten

Das Projekt “KI:STE” (KI STRategie für Erdsystemdaten) zielt darauf ab, die Verwendung von KI-Methoden zur Verarbeitung von Erdbeobachtungsdaten für verschiedene Anwendungen zu erleichtern und zu bewerten. Die im Projekt untersuchten Anwendungsfelder reichen von Luftqualität über Wolken und Strahlung bis hin zur Ausbreitung von Schnee, Eis und Wasser. Ein zentraler Baustein des Projekts wird die “Spatial Research Data Infrastructure” (SRDI) darstellen, welche die KI-Methoden und -Modelle mit Daten versorgen wird. Diese Infrastruktur muss flexibel auf die Anforderungen der KI-Algorithmen anpassbar sein und Daten in einem für die angestrebten Verarbeitungsschritte optimierten Format bereitstellen können.

2.2 KI-Methoden für die Planung von Schiffsrouten

MariData ist ein vom BMWi gefördertes Projekt, welches eine deterministische Analyse des Energiebedarfs von Schiffen auf der Grundlage von Umwelt- und physikalischen Bedingungen ermöglichen soll. Die verschiedenen Einflüsse auf den Energiebedarf werden analysiert, um Werkzeuge für ein energieoptimiertes Schiffsmanagement bereitzustellen. Dies führt beispielsweise zu Vorschlägen zu Routenalternativen oder Geschwindigkeits- und Trimmanpassungen.

Die Vielzahl verschiedener Datenquellen, wie z. B. Seekarten, Wetteraufzeichnungen und Vorhersagen (insbesondere Wind, Temperatur, Wellenhöhen und Strömungen), erfordern einen optimierten Ansatz für das Geodatenmanagement. Ein integrierter Datenspeicher liefert Daten sowohl für die Reederei als auch für die Schiffe auf See. Gerade aufgrund der besonderen Herausforderung zur Datenkommunikation in abgelegene Seegebiete ist es notwendig, die Daten sinnvoll zu unterteilen und vorzuverarbeiten, um die Datenlast zu verringern. Ein zentraler Aspekt besteht darin, Modelle zu entwickeln, die auf maschinellem Lernen (ML) basieren, um den Energiebedarf auf Grundlage der vorprozessierten Daten abzuleiten. Ein Routing-Service verwendet die zentral bereitgestellten Daten, um Routen mit verringertem Energiebedarf unter Berücksichtigung der nautischen Gegebenheiten, der Schiffssicherheit und der Lieferpläne bereitzustellen.

2.3 Infrastruktur-Konzepte

Für die Realisierung der Geodatenmanagement-Systeme wird in beiden Projekten ein Konzept entwickelt, welches die speziellen Anforderungen an Datensätze für Machine-Learning-Zwecke berücksichtigt. Insbesondere die Bereitstellung über einheitliche Schnittstellen (z.B. OGC Web Services und APIs) spielt eine wichtige Rolle mit Blick auf eine nachhaltige Nutzung der Daten. Neben diesen klassischen Bausteinen einer Geodateninfrastruktur werden Geo Data Cubes, z.B. "Open Data Cube" (KILLOUGH 2018) oder "gdalcubes" (APPEL UND PEBESMA 2019), für Raster-basierte Datensätze (z.B. Satellitendaten, Strömungsmodelle) in die Infrastruktur eingebunden. Dies ermöglicht eine effiziente Prozessierung "nahe an den Daten", da diese zugrundeliegenden Software-Lösungen auf einfache und effiziente Art und Weise durch eigene Methoden und Algorithmen erweitert werden können.

Die Skalierbarkeit der individuellen Lösung spielt ebenfalls in beiden Projekten eine zentrale Rolle. Hier werden diverse Technologien aus etablierten Cloud-Lösungen eingesetzt. Insbesondere die "Object Storages" (z.B. auf Basis des S3-Protokolls) zur Speicherung großer Raster-Datenmengen in Kombination mit effizienten Dateiformaten wie dem "Cloud Optimized GeoTIFF" (COG) oder "zarr" für die Speicherung von mehrdimensionalen Arrays sind wichtige Elemente der Lösungsstrategien.

Dieser Beitrag bietet einen Einblick in die frühe Phase der Konzeption und Umsetzung der beschriebenen Infrastruktur-Konzepte. Der Fokus liegt hierbei auf der Darstellung der spezifischen Anforderungen an die zu entwickelnden Lösungen. Zusätzlich werden erste Ergebnisse und potentielle Lösungsmuster dargestellt.

Literatur

- APPEL, M., PEBESMA, E. (2019), On-Demand Processing of Data Cubes from Satellite Image Collections with the gdalcubes Library. *Data* 2019, 4, 92. <https://doi.org/10.3390/data4030092>
- BAUMANN P., MISEV D., MERTICARIU V., HUU B.P. (2019), Datacubes: Towards Space/Time Analysis-Ready Data. In: Döllner J., Jobst M., Schmitz P. (eds) *Service-Oriented Mapping. Lecture Notes in Geoinformation and Cartography*. Springer, Cham. https://doi.org/10.1007/978-3-319-72434-8_14
- KILLOUGH B. (2018), "Overview of the Open Data Cube Initiative," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 2018, pp. 8629-8632, <https://doi.org/10.1109/IGARSS.2018.8517694>
- RIEKE, M., DROST, S., JIRKA, S., VOGT, A. (2020), Event-driven Processing of Earth Observation Data, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-16661, <https://doi.org/10.5194/egusphere-egu2020-16661>

Interaktive Karten für große Datensätze – eine algorithmische und kartographische Herausforderung

Jan-Henrik HAUNERT

Institut für Geodäsie und Geoinformation · Universität Bonn · Meckenheimer Allee 172 · 53115 Bonn
E-Mail: haunert@igg.uni-bonn.de

1 Herausforderungen

Die Verbreitung günstiger Positionierungssensoren und der Ausbau von Dateninfrastrukturen haben im letzten Jahrzehnt zu einem massiven Zuwachs an verfügbaren Geodaten geführt. Diese Daten stellen ein großes Potenzial für Aufgaben des alltäglichen Lebens sowie der Wirtschaft, Verwaltung und Wissenschaft dar – die Anwendungen reichen von der Suche nach einem Restaurant in fußläufiger Entfernung des aktuellen Standpunkts bis zu komplexen raum-zeitlichen Analysen über das Vorkommen von Starkwetterereignissen. Zur Interpretation von Geodaten durch Menschen haben sich Karten bewährt. Die Generierung von Karten auf Grundlage sehr großer Datensätze stellt die Kartographie jedoch vor neue Herausforderungen. Dieser Beitrag stellt Lösungsansätze für eine algorithmische und eine kartographische Herausforderung dar.

Die *algorithmische Herausforderung* besteht darin, *effiziente Algorithmen* zur schnellen Generierung übersichtlicher Karten aus sehr großen Datensätzen zu entwickeln. Insbesondere muss der Detailgrad der Daten an die Fragestellung der Nutzerin oder des Nutzers angepasst werden. Dabei sind klassische Aufgaben der Generalisierung (z. B. Selektion, Vereinfachung, Aggregation) und Schriftplatzierung zu lösen. Von zunehmender Bedeutung ist die *Vorberechnung von Datenstrukturen*, um eine schnelle Resonanz auf bestimmte Anfragen zu gewährleisten. Beispielsweise soll nach Anwendung thematischer oder zeitlicher Filter eine optimale Darstellung schnell angezeigt werden.

Die *kartographische Herausforderung* resultiert aus der hohen räumlichen Dichte der zu visualisierenden Objekte, die in der Regel aus einem großen Datenvolumen folgt. Kartographische Darstellungen ohne graphische Konflikte sind oft nur mit einem extrem großen Informationsverlust oder auf einem extrem großen Kartenmaßstab möglich. Bei der Suche nach einem geeigneten Hotel in einer Stadt ist man beispielsweise oft gezwungen, zu einem extrem großen Kartenmaßstab zu zoomen, um alle Suchergebnisse zu inspizieren. Insbesondere auf kleinen Displays geht dabei jedoch der geographische Kontext verloren. Dieses Dilemma erfordert neue Ansätze der Interaktion und graphischen Repräsentation.

2 Lösungsansätze

In meinem Beitrag stelle ich Lösungsansätze für die beschriebenen Herausforderungen vor, die an der Professur für Geoinformation des Instituts für Geodäsie und Geoinformation der Universität Bonn entwickelt wurden. Ich stelle Datenstrukturen zur Ermöglichung interaktiver Karten für raum-zeitliche Analysen vor (Abschnitt 2.1) und präsentiere neue Ansätze zur Karteninteraktion, welche die Navigation durch große Punktmengen erlauben (Abschnitt 2.2).

2.1 Karten zur Exploration raum-zeitlicher Daten

Visual Analytics bezeichnet die Analyse sehr großer raum-zeitlicher Daten mithilfe interaktiver Visualisierungen. Beispielsweise liegen sehr große Mengen an Vogelbeobachtungen mit Koordinaten und Zeitstempeln vor. Im Kontext der Biodiversitätsforschung werden diese Daten zur Analyse von Verbreitungs- und Bewegungsmustern von Vögeln visualisiert. Dichtekarten stellen ein geeignetes Mittel dar, um die räumliche Verbreitung einer Spezies in einem bestimmten Zeitraum zu visualisieren. Zur Analyse zeitlicher Veränderungen sollte dieser Zeitraum jedoch flexibel wählbar oder sogar interaktiv und kontinuierlich veränderbar sein. Ein Visualisierungssystem sollte *temporale Bereichsanfragen* ermöglichen und die Karte nach Spezifizierung einer Anfrage möglichst schnell aktualisieren, sodass kontinuierliche Animationen durch Variierung der Einstellungen eines Sliders möglich sind (siehe Abb. 1). Bonerath et al. (2019) und Bonerath et al. (2020) haben für diese Anwendung effiziente Datenstrukturen entwickelt und dabei die Idee von *Time-Windowed Data Structures* aus der algorithmischen Geometrie zum ersten Mal auf Probleme der Kartographie angewandt.

2.2 Karten für kleine und mobile Geräte

Zoomless Maps sind Karten, welche die Navigation durch große und detaillierte Datenmengen bei einem festen Kartenmaßstab erlauben. Das einfachste Modell einer Zoomless Map sieht die Verteilung von Punkten bzw. deren Signaturen oder Label über mehrere Seiten vor. Durch einfache Gesten bzw. Swiping können die Nutzerinnen und Nutzer der Karte von einer zur nächsten Seite navigieren und damit auf alle Daten zugreifen. Jede Seite muss dabei frei von graphischen Konflikten sein. Zudem sollen hoch gewichtete Objekte auf den vorderen Seiten erscheinen. Und schließlich soll die Information auf wenige Seiten verteilt werden. Gedicke et al. (2019) haben Optimierungsalgorithmen zur Lösung dieses

Probleme entwickelt. Ein ähnlicher Ansatz, bei dem Labels für Punktobjekte jedoch am Kartenrand platziert werden, wurde von Gedicke et al. (2020) vorgestellt; siehe Abb. 2. Noch offen ist die Frage, ob hinreichend gute Lösungen in Echtzeit während der Interaktion berechnet werden können oder sich eine Vorberechnung von Datenstrukturen anbietet. Zudem stellt die Kombination der entwickelten Interaktionsverfahren mit etablierten Verfahren (Zooming, Panning und Fokus+Kontext-Karten (Niedermann & Haunert 2019)) eine wichtige Aufgabe dar. Dabei gilt es insbesondere, die Stabilität der Karte zu berücksichtigen.

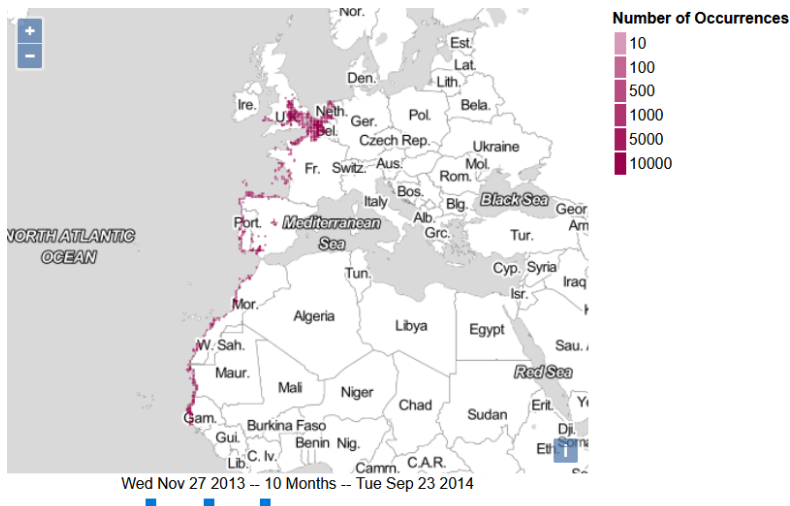


Abb. 1: Eine interaktive Karte, deren Inhalt auf ein gewähltes Zeitintervall begrenzt ist.

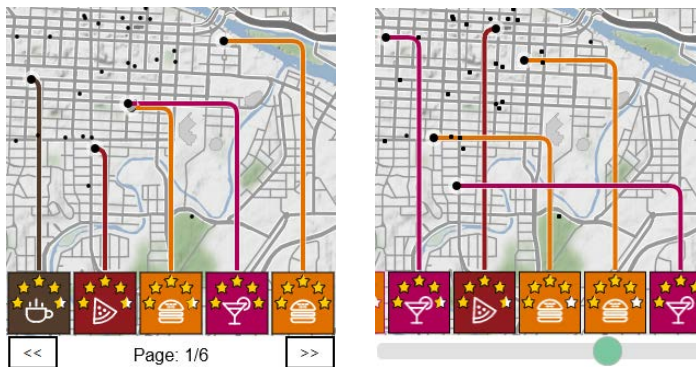


Abb. 2: Zwei Randbeschriftungsmodelle für Zoomless Maps, welche die Navigation durch große Punktmengen ohne Veränderung des Kartenmaßstabs erlauben. Links: Labels werden auf mehreren Seiten platziert. Rechts: Durch Sliding können Labels in die Karte gezogen werden.

Literatur

- BONERATH, A., NIEDERMANN, B., DIEDERICH, J., ORGEIG, Y., OEHRLEIN, J. & HAUNERT, J.-H. (2020), A time-windowed data structure for spatial density maps. In: Proc. 28th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL '20), Seiten 15-24.
- BONERATH, A., NIEDERMANN, B., & HAUNERT, J.-H. (2019), Retrieving alpha-shapes and schematic polygonal approximations for sets of points within queried temporal ranges. In: Proc. 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL '19), Seiten 249-258.
- GEDICKE, S., BONERATH, A., NIEDERMANN, B. & HAUNERT, J.-H. (2021), Zoomless maps: external labeling methods for the interactive exploration of dense point sets at a fixed map scale. In: IEEE Transactions on Visualization and Computer Graphics, 27(2):1247-1256.
- GEDICKE, S., NIEDERMANN, B. & HAUNERT, J.-H. (2019), Multi-page labeling of small-screen maps with a graph-coloring approach. In: Proc. 15th International Conference on Location Based Services (LBS 2019).
- NIEDERMANN, B. & HAUNERT, J.-H. (2019), Focus+context map labeling with optimized clutter reduction. In: International Journal of Cartography, 5(2-3):158-177.

Simulation und Steuerung von On-Demand Flotten

Klaus BOGENBERGER

Lehrstuhl für Verkehrstechnik · Technische Universität München · Arcisstr. 21 · 80333 München

E-Mail: klaus.bogenberger@tum.de

1 Zusammenfassung

Das Personenbeförderungsgesetz wurde vor kurzem überarbeitet, um einen rechtlichen Rahmen für On-Demand Mobilität zu schaffen (www.bmvi.de/SharedDocs/DE/Gesetze-19/entwurf-gesetz-personenbefoederungsrecht.html). Obwohl die Profitabilität solcher Mobilitätsdienste derzeit noch nicht gesichert ist, haben sie in den USA (Uber, Lyft) und China (Didi) bereits merkliche Marktanteile am modalen Split übernommen. Eine hohe Marktwerteinschätzung wird diesen Firmen aber vor allem im Hinblick auf autonome Fahrzeuge zugewiesen. Durch den Wegfall der Fahrerkosten, die irgendwo zwischen 50-75% der Gesamtkosten betragen, können autonome On-Demand Flotten wesentlich günstiger betrieben und angeboten werden, wodurch ein weiterer, womöglich drastischer Anstieg an Fahrten möglich ist. Daher ist es von großer Bedeutung, den Betrieb von On-Demand Flotten zu verstehen und die Wirkungen zu simulieren.

Bei unseren Simulationen erhält der Betreiber App-Anfragen, die sowohl Start als auch Ziel des Reisewunsches festlegen. Dann erstellt der Betreiber ein Angebot, das per App an den Nutzer kommuniziert wird. Auf Basis dieses Angebots entscheiden die Nutzer, ob sie das Angebot annehmen oder nicht. Wir unterscheiden On-Demand Mobilität vorwiegend in Hailing und Pooling. Im Hailing-Betrieb wird immer nur ein Fahrtwunsch gleichzeitig bedient und das Fahrzeug fährt zwischen zwei Fahrtwünschen ohne Passagier. Im Pooling-Betrieb können mehrere Fahrtwünsche gebündelt werden, wodurch sich die Flotte Fahrzeuge, die Kunden Geld, und das Straßennetzwerk Belastung einsparen können.

Die Steuerung von Flotten, vor allem im Ridepooling und bei großer Skalierung mit Hunderten bzw. Tausenden von Fahrzeugen, ist sehr komplex und lässt sich aufgrund der Dynamik des Problems praktisch nicht optimal lösen. Daher ist die entwickelte Simulationsumgebung sehr hilfreich, um verschiedene Flottensteuerungsalgorithmen zu testen und zu bewerten. Da aber neben den betrieblichen Kenngrößen auch die Kundenstatistik und die kompletten Fahrzeugtrajektorien Ergebnisse der Simulationen sind, können auch die Kundenseite evaluiert und verkehrliche und Umweltwirkungen geschätzt werden.

Die ersten Studien haben sich mit dem einfacheren Hailing-Betrieb beschäftigt. Die wichtigsten Komponenten des Betriebssystems dort sind die Kundenzuweisungs-Strategie sowie die Fahrzeugrepositionierungs-Strategie. Bei der einfachsten Kundenzuweisungsstrategie wird ein neuer Fahrtwunsch dem nächstgelegenen freien Fahrzeug zugewiesen; dieses fährt dann sofort los und holt den Kunden ab. Intelligentere Strategien berücksichtigen die Fahrzeuge, die derzeit andere Kunden ans Ziel bringen bei der Zuweisung. Bei der Fahrzeugrepositionierung werden Fahrzeuge ohne Passagier im Geschäftsgebiet bewegt. Anders als bei einer passagierlosen Fahrt um einen Kunden abzuholen, wird hier ein Gebiet als Ziel definiert, in dem mehr Nachfrage als verfügbare Fahrzeuge erwartet werden.

Eine Studie verglich ein autonomes Hailing-System mit einem Carsharing System, bei dem ebenfalls keine Fahrer benötigt werden, indem die Buchungsdaten des Carsharing Systems als Nachfrage an das autonome Hailing-System gegeben wurde. Die Leerfahrten erhöhten die Auslastung der Flotte beträchtlich, so dass nur ein Viertel der Fahrzeuge benötigt wurden, um die gleiche Nachfrage zu bedienen. Der Anteil der Leerfahrten betrug dabei in etwa 10% an der Gesamtfahrleistung der Flotte. In einer weiteren Studie wurde mit Hilfe einer mikroskopischen Verkehrsflusssimulation überprüft, wie sich diese Leerfahrten auswirken. Dazu wurden in einem Verkehrsmodell der Stadt München 10% aller PKW Fahrten durch Anfragen an das autonome Hailing-System ersetzt, das diese mit 4000 Fahrzeugen bediente. Wie erwartet stieg die Verkehrsbelastung an, allerdings hielt sich der Anstieg im Rahmen. Dies liegt daran, dass eine intelligente Flottensteuerung die Leerfahrten kurzhalten kann und aufgrund einer zeitlichen Komponente im Routing auch verhindert wird, dass die leeren Anfahrten zu Kunden lange Stauzeiten beinhalten.

Durch eine Variation der Nachfrage und der Flottengröße wurde anhand von Simulationen festgestellt, dass der Leerfahrtanteil mit der Skalierung sinkt. Dies ist dadurch erklärbar, dass es eine höhere Fahrzeugdichte gibt, wodurch das bestegelegene Fahrzeug im Schnitt näherliegt. Als nächstes wurde die Skalierung von Pooling-Systemen untersucht. Dort ergibt sich zusätzlich zu den kürzeren Anfahrten noch ein zweiter positiver Effekt bei größerer Skalierung: je höher die Nachfragedichte ist, desto höher ist die Wahrscheinlichkeit, dass die Quelle-Ziel Beziehungen einzelner Fahrten derart geartet sind, dass sich Fahrten bündeln („poolen“) lassen. Bei geringer Nachfrage (und Fahrzeugzahl) ist Pooling kaum möglich und es werden mehr Leerkilometer gefahren, als durch die Bündelung eingespart werden kann. Erst ab einer gewissen Skalierung erzeugt das System mehr Personenkilometer als Fahrzeugkilometer gefahren werden. Eine Auswertung der Fahrzeugtrajektorien hat ergeben, dass die größten Bündelungseffekte auf den Hauptverkehrswegen zustande kommen; da diese häufig die am stärksten belastenden Straßen im Verkehrssystem sind, ist dies positiv zu bewerten.

Bei der Auswertung der Trajektorien wurde auch festgestellt, dass zusätzliche Fahrleistung in kleineren Nebenstraßen entsteht. Dies ist notwendig, wenn ein Tür-zu-Tür Service angeboten wird und die Anfahrten zu Kunden daher Nebenstraßen benutzen müssen. Um diesem Effekt entgegenzuwirken wurde in einer weiteren Studie untersucht, wie viel Performancegewinn möglich ist, wenn man die Kunden einige Meter zum Abholpunkt und vom Ausstiegspunkt gehen lässt. Dadurch lassen sich Ein- und Ausstiege bündeln und Routen mit wesentlich geringerer Umwegzeit finden. Unsere Simulationen zeigten, dass sich die Systemeffizienz derart verbessern lässt, dass die die Flotte so viel weniger fahren muss, dass für Kunden die gleiche Gesamtreisezeit herauskommt. Das heißt, dass die Wartezeit und Fahrtzeit so viel geringer sind, dass die Zeit für die Fußwege wieder hereingeholt wird.

Weitere Studien beschäftigten sich mit Ladestrategien für On-Demand Flotten, der Kalibrierung der Flottensteuerungsparameter anhand Daten aus dem Realbetrieb von MOIA, der Verknüpfung von Flottensimulation mit sehr detaillierten agentenbasierten Nachfragenmodellen und der Regulierung von Verkehrssystemen mit On-Demand Flotten, zum Beispiel durch eine Limitierung der Flottengröße, eine Maut für PKW und ODM, sowie eine Erhöhung der Parkgebühren.

Vernetzte dynamische Sensorplattformen: mehr Fahrzeuge - bessere Performance?

Steffen SCHÖN

Institut für Erdmessung Leibniz Universität Hannover · Schneiderberg 50 · 30167 Hannover
E-Mail: schoen@ife.uni-hannover.de

1 Einführung ins Thema

In den letzten Jahren haben automatisierte und autonome Systeme Einzug in unser Leben gehalten wie Helfer in Haushalt und Garten, automatisierte Fahrzeugsysteme wie etwa Stau-, Spur-, Brems- und Kreuzungsassistenten im öffentlichen Straßenverkehr oder flexible, leicht anlernbare Assistenten in der Produktion im Kontext Industrie 4.0. Besondere Dynamik ist im Bereich des hochautomatisierten Fahrens hin zum autonomen Fahren zu beobachten.

Die Koexistenz einer zunehmenden Anzahl autonomer Systeme eröffnet die Möglichkeit zur Nutzung von Synergien (z.B. bei der Umgebungserfassung), erfordert aber auch Abstimmung (bspw. Kollisionvermeidung). An Stelle einer autarken Funktionsweise tritt die Einbeziehung externer Information, von anderen Sensorplattformen oder in Form von Vorwissen über die Umgebung. Die einzelnen autonomen Systeme können als Knoten eines Sensornetzes aufgefasst werden, wodurch Kollaboration zum Erreichen individueller und gemeinsamer Ziele möglich wird. Der entscheidende Mehrwert besteht aber darin, dass Aufgaben lösbar werden, die einzeln entweder gar nicht oder nur suboptimal bearbeitet werden können, wie beispielsweise eine Positionierung mittels GPS in einer engen Straßenschlucht oder die zeitgleiche Erfassung ausgedehnter Objekte durch mehrere Sensorplattformen.

Betrachtet man die erzielbare Performance, so stehen neben dem Qualitätsparameter der Genauigkeit (Präzision und Richtigkeit) der Positionierung, Geschwindigkeit und Ausrichtung auch weitere Parameter wie die Integrität und Kontinuität der Navigationslösung. Die Integrität stellt das Maß des Vertrauens dar, welches man technischen Systemen entgegenbringen kann. Das Integritätsrisiko ist dabei die Wahrscheinlichkeit, dass eine Fehlerschranke überschritten wird, ohne dass dies erkannt und der Nutzer innerhalb einer vorgegebenen Zeitspanne gewarnt wird. Abgesehen von der Luftfahrt sind Integritätskonzepte bisher kaum entwickelt und es besteht ein national wie international relevanter Forschungsbedarf.

An diesen Themen arbeitet das DFG geförderte Graduiertenkolleg i.c.sens: Integrität und Kollaboration in dynamischen Sensornetzen an der Leibniz Universität Hannover. Der Vortrag wird in einem ersten Teil die Forschungsfragen des Graduiertenkollegs und Ergebnisse der ersten Phase kurz zusammenfassen. In einem zweiten Teil beleuchten wir einen Aspekt genauer: Führt das Hinzufügen von mit Sensoren ausgestatteten Fahrzeugen im dynamischen Sensornetz zwingend zu einer Steigerung der Navigationsperformance aller Teilnehmer.

2 Zusammenfassung des Vortragsinhalts

2.1 Einführung in dynamische Sensornetze und das Graduiertenkolleg i.c.sens

Fahrzeuge sind heute mit einer Vielzahl an Sensoren zur Navigation und Umwelterfassung ausgestattet. Dies reicht von Kameras, RADAR oder LIDAR Sensoren über Odometer zu GNSS-Sensoren und Inertialmesseinheiten. Zusätzlich stehen externe Informationen in Form von Kartenmaterial oder 3D Stadtmodellen mit unterschiedlichen Layern wie Fahrspuren und Straßenmarkierungen, Verkehrszeichen oder Gebäuden und in variabler Güte und Detailgrad teilweise zur Verfügung oder müssen extra erfasst werden. Diese unterschiedlichen Messdaten werden dann in einem gemeinsamen Filter ausgewertet, um das Fahrzeug in seiner Umgebung zu positionieren. Bereits diese Integration sich ergänzender Sensorsysteme kann man als Kollaboration auf Plattformebeane auffassen.

In unserem Graduiertenkolleg gehen wir jedoch deutlich darüber hinaus und sehen den Schwerpunkt der Kollaboration auf der Interaktion zwischen mehreren Fahrzeugen oder allgemein Plattformen untereinander und mit ihrer Umgebung. Abb. 1 zeigt beispielhaft einen Ausschnitt aus den im Graduiertenkolleg durchgeführten Messkampagnen, (SCHÖN et al 2018, ICSENS 2021). Um die umfangreichen erhobenen Datensätze vielseitig nutzbar zu machen (zunächst für die Doktoranden und später nach einer Veröffentlichung der Daten dem Fachpublikum) ist eine durchdachte Datenstruktur und Abspeicherung nötig. Für das Graduiertenkolleg wurde extra eine solche Struktur entwickelt, die auch einen wesentlichen Beitrag zur aktuellen Diskussion zum Forschungsdatenmanagement leistet (FISCHER, SESTER & SCHÖN 2020).

Für die gemeinsam agierenden Systeme wurden spezielle Navigationsfilter entwickelt (GARCIA FERNANDEZ & SCHÖN, 2019), die die Vorteile der kollaborativen Navigation herausarbeiten und Messungen zwischen Sensorknoten berücksichtigen.

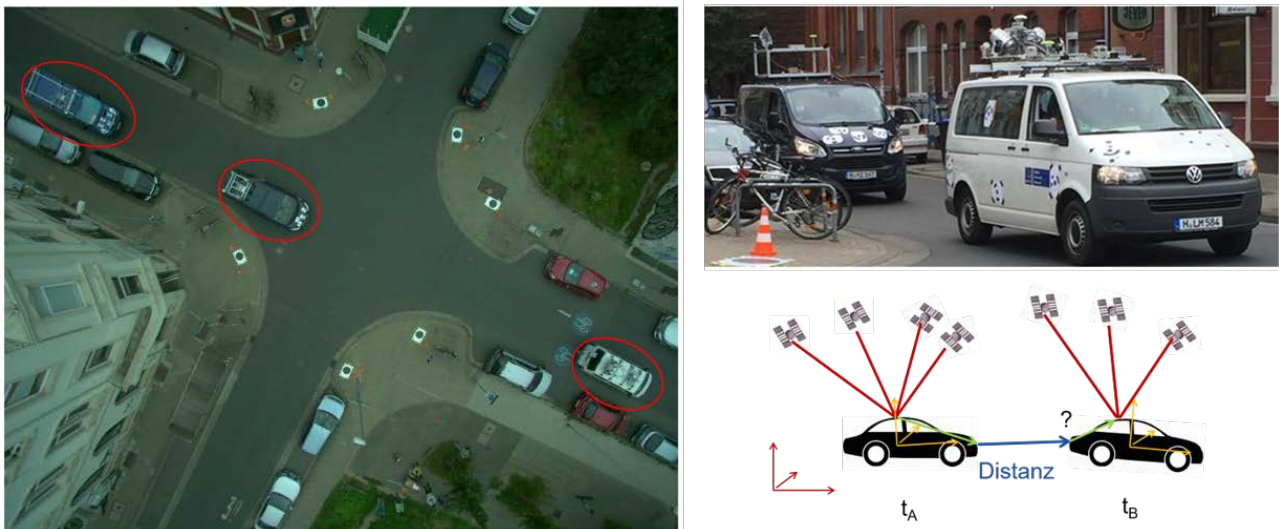


Abb. 1: Kollaborative Navigation und Umgebungserfassung im Graduiertenkolleg i.c.sens. Links: Blick vom UAV auf eine Kreuzung mit drei Messfahrzeugen. Rechts: Foto der Messfahrzeuge und schematische Darstellung der zu lösenden Identifikationsprobleme und Transformationen zwischen Koordinatensystemen und Zeitsystemen der einzelnen Fahrzeuge und Sensoren.

2.2 Mehr Fahrzeuge - bessere Performance?

In einem zweiten Teil des Vortrags werden wir auf eine wichtige Fragestellung eingehen: wie die Anzahl von Fahrzeugen die Performance der einzelnen Teilnehmer im dynamischen Sensornetz verbessert. Wenn man ein solches dynamisches Sensornetz intuitiv als Weiterentwicklung eines geodätischen Netzes mit zeitlich variierenden Messkonfigurationen auffasst, ist es naheliegend, dass eine Performancesteigerung erwartet werden kann. Dies gilt auf jeden Fall für den Qualitätsparameter der Präzision der Punktposition. Im Fall der Richtigkeit ist dies nicht einfach abschätzbar. Hier spielen die Anzahl, Größe und Verteilung nichtaufgedeckter grober Fehler im Datenmaterial und ihre Übertragung auf die Position eine entscheidende Rolle. Insbesondere bei Lowcost Sensoren aber auch bei LiDAR und GNSS im urbanen Bereich kann es zu vielen Ausreißern kommen. Ergebnisse einer Simulationsstudie werden aufzeigen, wo es zu Bruchpunkten in der Performancesteigerung kommt und welche Eingangsgrößen dafür ausschlaggebend sind.

Literatur

- FISCHER, C., SESTER, M. & SCHÖN, S. (2020), Spatio-Temporal Research Data Infrastructure in the Context of Autonomous Driving, *ISPRS Int. J. Geo-Inf.* 2020, 9, 626
- GARCIA FERNANDEZ N. & SCHÖN S. (2019), Optimizing Sensor Combinations and Processing Parameters in Dynamic Sensor Networks, *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, Miami, Florida, September 2019, pp. 2048-2062 DOI: 10.33012/2019.16885
- ICSSENS (2021), DFG Graduiertenkolleg i.c.sens: Integrität und Kollaboration in dynamischen Sensornetzen www.icsens.uni-hannover.de
- SCHÖN S., BRENNER C., ALKHATIB H., COENEN M., DBOUK H., GARCIA FERNANDEZ N., FISCHER C., HEIPKE C., LOHMANN K., NEUMANN I., NGUYEN U., PAFFENHOLZ J.-A., PETERS T., ROTTENSTEINER F., SCHACHTSCHNEIDER J., SESTER M., SUN L., VOGEL S., VOGES R. & WAGNER B. (2018), Integrity and Collaboration in Dynamic Sensor Networks, *Sensors* 2018 DOI: 10.3390/s18072400

Projekt Geomassendaten

Nutzung von Geomassendaten der Bayerischen Vermessungsverwaltung durch Pilotanwender verschiedener Branchen

Wolfgang BAUER, Präsident Landesamt für Digitalisierung, Breitband und Vermessung

Landesamt für Digitalisierung, Breitband und Vermessung · Alexandrastraße 4 · 80538 München

E-Mail: wolfgang.bauer@ldbv.bayern.de

1 Einführung ins Thema

Amtliche Geobasisdaten werden schon seit geraumer Zeit zur Erledigung vielfältiger Aufgaben in allen Bereichen der staatlichen und kommunalen Verwaltung sowie in Wirtschaft und Hochschule eingesetzt. GIS-Dienstleister und Ingenieurbüros unterstützen die Anwender durch fachkundige Beratung und Bereitstellung von Softwareprodukten. Dabei ist festzustellen, dass Planungen vorwiegend im 2D-Bereich erfolgen. Schwerpunktmäßig werden Produkte wie die ALKIS-Flurkarte und das Digitale Orthophoto (DOP) genutzt.

Neue Aufgaben und Herausforderungen wie z. B. eine ständig stärker geforderte, transparente Bürgerbeteiligung in Planungsprozessen oder die mögliche Realisierung von Windkraft- und Solaranlagen erfordern zunehmend 3D-Planungsprozesse unter Verwendung hierfür geeigneter Geodaten.

Die Bayerische Vermessungsverwaltung (BVV) erfasst und stellt seit einigen Jahren sog. "Geomassendaten", die für 3D-Planungen verwendet werden können, bereit. Es ist feststellbar, dass diese Geomassendaten bisher nur in geringem Umfang in Planungsprozessen verwendet werden.

Der Vortrag stellt das Projekt Geomassendaten des Runden Tisch GIS e.V. (RTGIS) vor. Ziel des Projekts ist es, dass Pilotanwender an ausgewählten Projekten den Einsatz von Geomassendaten zeigen und die Ergebnisse dokumentieren. Die erarbeiteten Ergebnisse und Projektberichte sollen umfänglich vorgestellt und veröffentlicht werden. Anderen Anwendern sollen sie als Beispiele für mögliche Einsatzzwecke und zur Entscheidungsgrundlage dienen.

2 Zusammenfassung des Vortragsinhalts

2.1 Projekt

Die BVV erfasst seit mehreren Jahren Geobasisdaten, die neben der Lage- auch Höheninformationen enthalten. Aufgrund der hohen Auflösung und Genauigkeit zeichnen sich diese Daten durch eine hohe Datenmenge aus und können unter dem Begriff "Geomassendaten" zusammengefasst werden. Als Produkte sind hier das Digitale Orthophoto (DOP), das Digitale Oberflächenmodell (DOM), das Digitale Geländemodell (DGM 1), die Laserpunkte und die 3D-Gebäudemodelle (LoD1/LoD2) zu nennen.

Sowohl von GIS-Dienstleistern als auch der BVV selbst wurde festgestellt, dass diese Geomassendaten mit Ausnahme der DOP bisher nur in geringem Umfang in planerischen Prozessen v. a. in der kommunalen Verwaltung eingesetzt werden. Mögliche Gründe hierfür könnten sein, dass der Mehrwert von 3D-Planungen zu wenig erkannt und bekannt ist, dass ein notwendiges Know-how bei den Anwendern nicht vorhanden ist, dass es an technischer Ausstattung fehlt oder dass es an den Kosten scheitert.

Auf Anregung der Firmen Steinbacher Consult und GI Geoinformatik hat der RTGIS mit Beteiligung der BVV das Projekt Geomassendaten initiiert. Ziel ist es, an exemplarischen Anwendungsfällen Einsatzmöglichkeiten von Geomassendaten aufzuzeigen. Die projektbeteiligten Stellen bringen ihre Leistungen kostenfrei in das Projekt ein, so dass oben erwähnte Hindernisse nicht auftreten.

Der RTGIS hat in Zusammenarbeit mit der BVV die Pilotanwender ausgewählt, koordiniert das Projekt und sorgt für eine umfängliche Bereitstellung der Projektergebnisse. Die BVV stellt die Daten des DOP, DOM, DGM, die Laserpunkte, die 3D-Gebäudemodelle und mit dem WMTS BayernAtlas Hintergrundkarten für die Projektlaufzeit zur Verfügung. Die beteiligten GIS-Dienstleister und Technologie-Provider unterstützen die Pilotanwender durch ihr Know-how, durch technische Hilfestellung und Bereitstellung von Software.

Am Projekt nahmen 14 Pilotanwender aus der kommunalen und staatlichen Verwaltung, aus der Wissenschaft und aus der Wirtschaft teil. Die untersuchten Anwendungsfälle waren fachlich sehr breit gestreut und reichten von kleinräumigen Planungsvorhaben (z. B. Neuplanung von Gebäudeobjekten) bis hin zu großflächigen Trassierungsvorhaben (z.B. Visualisierung einer Bahntrasse).

Gemeinsam mit den GIS-Dienstleistern haben die Pilotanwender die Anwendungsszenarien pilotiert (09/2019-12/2020) und in einem strukturierten Abschlussbericht dokumentiert.

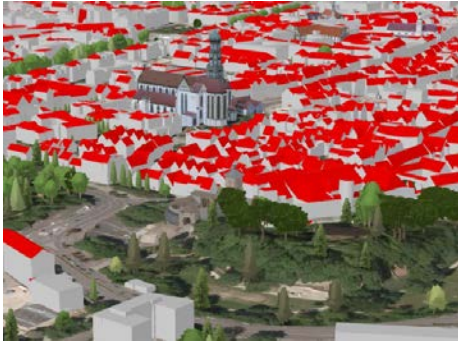


Abb. 1:

Kombination von Geodaten des LDBV (DOP, DGM1, LoD2) und der Stadt Augsburg (Landmarks, Baumkataster) in einer Web-Anwendung in ArcGIS Online

2.2 Ergebnisse

Aus den Abschlussberichten lassen sich in einer ersten Auswertung folgende Ergebnisse ableiten:

Anwendungen:

- Erfolgreiche Bearbeitung der fachlichen Aufgabenstellung durch zur Verfügung gestellten Geomassendaten möglich
- Mehrwert durch die Nutzung konnte erreicht werden
 - realitätsnahe Visualisierungen schaffen Verständlichkeit v.a. für Nicht-Experten
 - Unterstützung von Kommunikations- und Entscheidungsprozessen wird verbessert
- gute Übertragbarkeit auf andere Anwender v. a. bei kommunalen Anwendungen

Integration in Arbeitsprozesse:

- Anwendungen lassen sich gut und mit meist vertretbarem Aufwand in bestehende Verfahren integrieren
- hoher Nutzen bei Vorliegen eines Standardverfahrens
- hoher Aufwand, wenn Integration in neuen Arbeitsprozess erforderlich ist; Aufwand langfristig gerechtfertigt

Technik:

- leistungsfähige Hard- und Software sind Voraussetzung
- große Datenmengen stellen eine Herausforderung dar, Abhilfe könnte eine Reduzierung von flächenhaften Datensätzen auf projektbezogene Datensätze schaffen
- Skalierbare Systeme sind hilfreich, möglich durch Cloud-Lösungen

Geomassendaten:

- Geomassendaten der BVV ermöglichen eine erfolgreiche Bearbeitung der fachlichen Aufgabenstellung
- gute und einfache Integration in verwendete Software wegen Verwendung von Standardformaten
- Flächendeckung, laufende Aktualisierung und Kombinierbarkeit untereinander und mit Fachdaten sind Mehrwerte
- bundeseinheitliche Produktstandards ermöglichen Übertragbarkeit
- höhere und einheitliche Aktualität gewünscht, v.a. bei Kombination von LoD- und ALKIS-Daten, Aktualitätsunterschiede zwischen den unterschiedlichen Geomassendaten der BVV (v.a. Laserscanning vs. Luftbilder) sind nachteilig
- Laserpunkte zur Erstellung von Solarpotentialkataster besser als DOM-Daten geeignet
- Datenbereitstellung über Downloadverfahren gewünscht
- Angebot von 3D-Onlinediensten wäre hilfreich, aufwändige Datenhaltung und Aktualisierung entfällt

Personal/Organisation:

- Hohes technisches Know-how bei Anwendern notwendig, Unterstützung durch Dienstleister hilfreich
- Erarbeitung von neuen Prozessen in Planungsverfahren
- Arbeits- und Organisationskultur ist anzupassen

2.3 Öffentlichkeitsarbeit

Der Abschlussbericht und alle Projektberichte werden bzw. sind im "Aktivitätenkatalog" des Runden Tisch GIS e.V. eingestellt (<https://rtgkatalog.gis.lrg.tum.de/>). Zusätzlich wird auf verschiedenen Veranstaltungen zum Projekt berichtet. Geplant ist auch die Veröffentlichung der Ergebnisse durch die Projektbeteiligten. Ziel ist es, durch diese Pilotanwendungen und deren Veröffentlichung anderen Anwendern den Einstieg in die Verarbeitung dieser Daten auf Basis der entstandenen „use cases“ zu erleichtern und die Daten stärker in die Nutzung zu bringen.

Offene 3D-Geodaten in Thüringen

Frank ENGEL

Thüringer Landesamt für Bodenmanagement und Geoinformation · Hohenwindenstraße 13a · 99086 Erfurt
E-Mail: frank.engel@tlbg.thueringen.de

1 Einführung ins Thema

Als eine wesentliche Grundlage der Digitalisierung der Verwaltung wird die Bereitstellung von offenen Daten gesehen. Offene Geobasisdaten ermöglichen eine breite Nutzung in der Verwaltung, Wirtschaft, Forschung und Lehre sowie ebenso durch die Bürger. Die Herausforderung ist es, ein breites Spektrum an Bereitstellungsformen für Geobasisdaten anzubieten. Neben den verschiedenen Formen von Download- und Darstellungsdiensten wird für die 3-D Daten auf dem Geoportal der Geodateninfrastruktur Thüringens (GDI-TH) ein spezieller Darstellungsklient bereitgestellt. Unter Einbeziehung der dritten Dimension ermöglicht die Kombination von Geodaten eine Vielzahl von neuen Anwendungsmöglichkeiten und besonders wegen den unkomplizierten Nutzungsmöglichkeiten.

2 Zusammenfassung des Vortragsinhalts

2.1 Einführung von offenen Geodaten in Thüringen

Am 2. Februar 2016 hat das Kabinett der Thüringer Landesregierung das Landesprogramm „Offene Geodaten“ gebilligt. Es begann eine sehr kurzfristige Umsetzungsphase zur kostenfreien Bereitstellung von Geodaten. Bereits am 1. Januar 2017 standen Geobasisdaten auf dem Geoportal zum Download zur Verfügung. Großen Wert wurde auf eine einfache Nutzung der Downloadfunktionalitäten gelegt. Über die Geodateninfrastruktur in Thüringen werden seitdem die Dienste und Anwendungen frei nutzbar im Netz bereitgestellt. Zunächst wurden fast alle Geobasisdaten angeboten. Ausgenommen sind personenbezogene Daten des ALKIS und Daten, für die Rechte Dritter zu beachten sind. Inzwischen werden immer mehr Geofachdaten über die GDI nutzbar. Ein großer Vorteil ist die unkomplizierte Möglichkeit zur Kombination verschiedenster Datenbestände.

2.2 Bereitstellungsformen für 3D-Geobasisdaten

Insbesondere für die 3D-Geodaten wurden verschiedene Möglichkeiten der Bereitstellung umgesetzt, um einem möglichst breiten Anwenderkreis die Nutzung zu ermöglichen. Die Herausforderung bestand darin, die verschiedenen Möglichkeiten der Datenbereitstellung umzusetzen. Es galt dabei die Anforderungen der „GIS-Experten“ aber auch die Möglichkeiten der Gelegenheitsnutzer zu berücksichtigen. Die Darstellungsdienste der GDI-Th stellen einen WMS mit der Schummerung des DGM 5 bereit. Mit eigenen Methoden kann der hinterlegte Vektordatenbestand des DGM für Höhen- und Profilberechnungen im Geoclient der GDI genutzt werden. In Vorbereitung befindet sich der Dienst für das DGM 2. (vgl. Abb. 1)



Abb. 1: Darstellung des DGM 2 in Grauwertschummerung mit ALKIS-Gebäuden und den geltenden Überschwemmungsgebieten

Um in der GDI eine 3D-Darstellung der Geobasisdaten zu ermöglichen wurde ein spezieller Client beschafft und für die Belange des Thüringer Geoportals angepasst. Schwerpunkt ist die Visualisierung der landesweiten Gebäude im LOD2 oder LOD1 in Kombination mit den DGM und den DOP's. Es ist möglich die Gebäude in den unterschiedlichsten Darstellungsformen zu präsentieren und Gebäudeinformationen abzurufen. (vgl. Abb. 2)



Abb. 2: 3D Visualisierung Thüringens auf dem Geoportal (LOD2/DGM/DOP)

Für den Download der 3D-Geobasisdaten wurden eine spezielle, einfach zu bedienende Downloadkomponente im Geoportal-Th entwickelt. Die Daten stehen in einzelnen Kacheln zur Verfügung. Der Nutzer kann durch Auswahl mehrerer Kacheln ein größeres Gebiet zum Download selektieren. Zusätzlich wurde eine einfache Möglichkeit zur Selektion über Datenaktualisierungen entwickelt. Daneben ist auch die Nutzung von Atom-Feeds möglich. (vgl. Abb. 3)



Abb. 3: Downloadclient des Geoportals-Th für DGM-, DOM-, und LAZ-Daten

2.2 Nutzerreaktionen

Das seit dem Jahre 2017 bestehende Angebot zur kostenfreien Nutzung von Geobasisdaten fand bei den Anwendern ein durchweg positives Echo. Das hat sich ebenfalls in einem überraschend breiten Presseecho wiedergespiegelt. Bereits nach wenigen Tagen hatte eine Firma auf Basis der „offenen Geodaten“ ein komplettes 3D-Datenmodell des Landes im Netz visualisiert. GIS-Firmen haben die Daten über Downloadfunktionen für ihre Systeme aufbereitet und stellen sie den Nutzern in verschiedensten Anwendungen zur Verfügung. Bestätigt wurde uns vielfach, dass die frei verfügbaren Daten im Hochschulbereich vollkommen neue Möglichkeiten eröffnen. Wegen den open Data Kriterien der freien Zugänglichkeit und der Nichtdiskriminierung werden keine nutzerbezogenen Daten gespeichert. Eine Übersicht über alle auf den Offenen Geodaten des TLBG basierenden Anwendungen kann es daher nicht geben. Das Potential der frei verfügbaren Geodaten zeigt sich sehr deutlich an den Möglichkeiten einer effizienteren Arbeit in der Verwaltung. Ein Thüringer Landesamt stellt die für Werkverträge notwendigen Geodaten nicht mehr mühevoll zusammen, sondern verweist nur noch auf die Downloadmöglichkeit auf dem Geoportal. Die Methoden zur Datenbereitstellung über das Geoportal werden ständig verbessert. Dabei werden sowohl die Belange der „GIS-Experten“ als auch die der Gelegenheitsnutzer berücksichtigt. Die Downloadfunktionen des Geoportals nutzen heute etwa 11.000 verschiedene Anwender pro Monat. Das Downloadvolumen betrug im Jahr 2020 27 TB. Darunter waren beispielsweise 65.000 Topographische Karten und 115.000 DOP und Luftbilder.

Literatur

BUNDESMINISTERIUM DES INNEREN (2014), Nationaler Aktionsplan der Bundesregierung zur Umsetzung der Open-Data-Charta der G8
IKG-GIZ THÜRINGEN (2019), Landesprogramm „Offene Geodaten“
GEOPORTAL-TH, www.geoportal-th.de
GOVDATA, www.govdata.de

Neue OGC-Standards als Katalysator für die Nutzung von 3D-Geoinformation

Volker COORS

Institut für Angewandte Forschung · HFT Stuttgart · Schellingstr. 24 · 70174 Stuttgart
E-Mail: volker.coors@hft-stuttgart.de

1 Einleitung

Die web-basierte 3D-Visualisierung ermöglicht eine nahezu fotorealistische interaktive Darstellung von 3D-Stadt- und –Landschaftsmodellen im Web-Browser und auf mobilen Endgeräten, ohne dass hierzu zusätzliche Software installiert werden muss. Ermöglicht wurde dieser immense technologische Fortschritt neben der Weiterentwicklung der Browser-Technologie auch durch hierarchisch aufgebaute streaming-fähige Datenformate, die eine Übertragung von sehr großen Datenbeständen in Abhängigkeit vom Blickpunkt des Betrachters erlauben. Um die Interoperabilität zwischen Web-Client und Geodatenserver sicherzustellen, müssen Schnittstellen und Datenformate standardisiert werden.

In diesem Beitrag wird ein Überblick über die verschiedenen OGC-Standards und deren Rolle im Kontext der web-basierten Visualisierung gegeben. Darüber hinaus wird auf die OGC API für 3D-Visualisierung eingegangen, die im letzten Jahr in mehreren OGC Piloten entwickelt und getestet wurde.

2 Zusammenfassung des Vortragsinhalts

2.1 Geo-Visualisierungspipeline

Die Geo-Visualisierungspipeline beschreibt das grundlegende Prinzip der interaktiven web-basierten Darstellung von 3D-Stadt- und –Landschaftsmodellen. Sie besteht aus drei wesentlichen Schritten: Datenaufbereitung (Filtern), der Abbildung der Daten auf Darstellungselemente (Mappen) und Bildgenerierung (Rendern). Bei einer interaktiven web-basierten 3D-Visualisierung erfolgen Filtern und Mappen üblicherweise in einem Vorverarbeitungsschritt auf dem Server, das Rendern im Web-Browser auf dem Client, wie in Abb. 1 dargestellt.

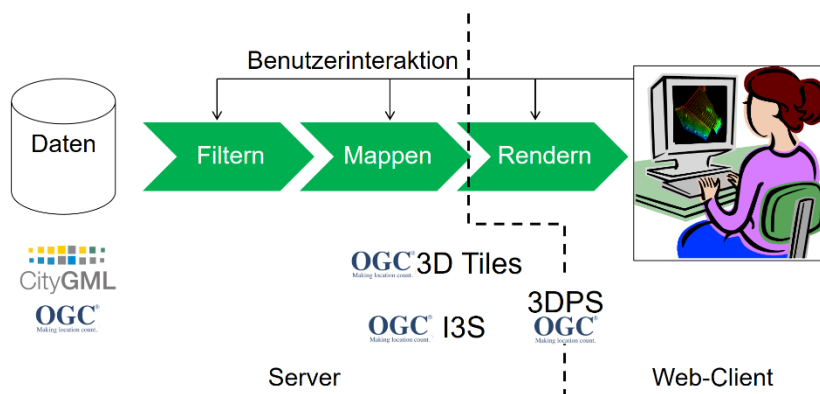


Abb. 1: Geo-Visualisierungspipeline für interaktive web-basierte 3D-Visualisierung

Die Ausgangsdaten, die i.d.R. als CityGML Datenbestand vorliegen, werden in eine hierarchische Komposition einer virtuellen Welt - dem sog. Szenengraph – transformiert. Dieser Szenengraph folgt der Metapher einer Filmkulisse und beinhaltet neben dem geometrischen Modell neben Beleuchtung und Kameraposition auch Information zum Verhalten einzelner Objekte. Der Benutzer kann die Szene durch Bewegungen der Kamera und Interaktion mit Objekten erkunden. Um das zu übertragende Datenvolumen vom Server zum Client zu reduzieren, werden im Idealfall nur die Daten, die nach dem Rendering auf dem Bildschirm sichtbar sind, in dem dazu notwendigen Detaillierungsgrad übertragen. Da sich die darzustellenden Objekte durch die interaktive Exploration der Szene ändern, müssen permanent Daten vom Server an den Client übertragen werden. Daher spricht man hier von einem Datenstrom zur Visualisierung des Szenengraphs. Im Gegensatz zu einem Video ist der Datenstrom bei der 3D-Visualisierung nicht linear, da die Interaktion des Benutzers nicht vorhersehbar ist. Um dennoch eine flüssige Darstellung mit einer Bildwiederholrate von mindestens 20 Bildern pro Sekunde sicherzustellen, wird die Szene als sog. Hierarchie von Hüllvolumen (Bounding Volume Hierarchy, BHV) aufgebaut. Die OGC community Standards 3D Tiles (COZZI et al. 2019) und Indexed Scene Layer (I3S) (REED & BELAYNAH 2020) sind Implementierungen solcher Hierarchien von Hüllvolumen.

2.2 Hierarchie von Hüllvolumen

Bei einer Hierarchie von Hüllvolumen wird ein großes 3D-Modell in Kacheln unterteilt, um eine schnelle Darstellung im Web-Browser zu realisieren. Im einfachsten Fall wird zur Erstellung der Kacheln eine Quadtree Datenstruktur verwendet. Dabei wird ein Gebiet regelmäßig in Quadranten unterteilt. Diese Methode ist auch bei web-basierten 2D-Karten üblich. Bei 3D-Modellen kommt hinzu, dass das Datenvolumen unregelmäßig verteilt ist. In einer dicht bebauten Innenstadt sind natürlich sehr viel mehr Gebäude vorhanden, als in ländlichen Regionen. Daher werden für die Visualisierung von 3D-Stadt- und Landschaftsmodellen auch komplexere Kachelungsstrategien wie adaptive Quadtrees und Octrees verwendet. Im OGC Testbed 13 wurden verschiedene Kachelungsstrategien hinsichtlich der Performance bei der webbasierten Visualisierung vergleichend untersucht (COORS 2018).

Die Datenstruktur einer Hierarchie von Hüllvolumen ist unabhängig von der verwendeten Kachelungsstrategie. Im Kern verbirgt sich hier eine Baumstruktur. Jeder Knoten des Baumes hat beliebig vielen Nachfolger. Jeder dieser Nachfolgerknoten hat eine vereinfachte Darstellungsgeometrie, die die Originalgeometrie abstrahiert. Das Hüllvolumen dieser Geometrie - in der Regel ein Quader oder eine Kugel - wird in den Metadaten des Nachfolgers vermerkt, ebenso der Fehler der Darstellungsgeometrie im Vergleich zur Originalgeometrie. Abhängig von dem geometrischen Fehler und der zu erwartenden Größe der Darstellungsgeometrie auf dem Bildschirm wird entschieden, welcher Teil des hierarchischen Szenengraphs zum Client übertragen werden muss. Im Datenformat 3D Tiles werden diese Metadaten in dem sog. TileSet.json gespeichert. Anhand der dort vorhandenen Metadaten kann ein Web-client entscheiden, welche Kacheln übertragen werden müssen. Der Speicherort dieser Kacheln ist im TileSet.json festgelegt.

2.3 3D Portrayal Service und 3D Geovolume API

Mit bestehenden Software-Werkzeugen kann aus einem CityGML-Modell eine Hierarchie von Hüllvolumen zur Visualisierung abgeleitet und als 3D Tiles oder I3S Modell auf einem Server bereitgestellt werden. Ist einem Web-client bekannt, wo das Modell zu finden ist, kann es im Web-Browser visualisiert werden. Nun verbleibt die Frage, woher ein Anwender weiß, auf welchem Server welche Modelle zu finden sind. Dies ist eine klassische Fragestellung einer Geodateninfrastruktur. Mit der 3D Geovolume API wird eine Art Katalogdienst aufgebaut, der es ermöglicht, auf einem Server, der verschiedene 3D-Szenen zur Verfügung stellt, nach Modellen in einem bestimmten Gebiet zu suchen. Dabei kann dieselbe 3D-Szene in verschiedenen Formaten bereitgestellt werden. Mit Deck.gl¹ wurde kürzlich eine Open Source Entwicklung vorgestellt, die sowohl 3D Tiles als auch I3S verarbeiten kann, aber das ist nicht bei allen Viewern der Fall. Daher ist es als Datenanbieter sinnvoll, die Daten in beiden Formaten bereitzustellen, um eine möglichst breite Datennutzung zu erzielen.

Die 3D Geovolume API definiert, wie die Struktur des Servers aufgebaut sein muss, damit ein Client auf dem Server nach bestimmten Modellen mit Hilfe einer RESTful API suchen kann, z.B. nach Modellen von New York. Das Ergebnis der Suche ist eine vordefinierte Hierarchie von Hüllvolumen. Die 3D Geovolume API wurde im OGC 3D Container Pilot spezifiziert und mit 6 verschiedenen Servern getestet, die unterschiedliche Modelle von einem detaillierten BIM Modell über Punktwolken bis hin zu einem weltweiten 3D-Gebäudebestand darstellen (MILLER et al. 2020). Es handelt sich dabei allerdings noch nicht um einen OGC Standard, sondern um eine Vorstudie, um ggf. einen neuen Standard zu entwickeln.

Im Gegensatz zur 3D Geovolume API erlaubt der 3D Portrayal Service (3DPS) zusätzlich die Selektion innerhalb einer Hierarchie von Hüllvolumen. Der 3DPS definiert eine Anfrageschnittstelle, die es z.B. erlaubt, alle Gebäude, die sich innerhalb eines bestimmten Gebietes befinden, darzustellen. Damit werden Anwendungsfälle wie die Integration eines Planungsmodells in ein bestehendes Stadtmodell unterstützt.

Literatur

- COORS, V. (EDITOR) (2018), 3D Tiles and I3S Interoperability and Performance Engineering Report, OGC Document 17-046, 2.3.2018, <http://docs.opengeospatial.org/per/17-046.html>
- COZZI, P., LILLEY, S., GETZ, G. (EDITORS) (2019), 3D Tiles Specification, Version 1.0, OGC Document 18-053r2, 31.1.2019, <https://www.ogc.org/standards/3DTiles>
- HAGEDORN, B., THUM, S., REITZ, T., COORS, V., GUTBELL, R. (EDITORS) (2017) KOUKOFIKIS, OGC 3D Portrayal Service, Version 1.0, OGC Dokument 15-001r4, <https://www.ogc.org/standards/3dp>
- KOUKOFIKIS, A., COORS, V., GUTBELL, R. (2018), Interoperable Visualization of 3D City Models using OGC's Standard 3D Portrayal Service, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. IV-4, S. 113–118. DOI: [10.5194/isprs-annals-IV-4-113-2018](https://doi.org/10.5194/isprs-annals-IV-4-113-2018).
- MILLER, T., TRENUM, G., LIEBERMAN, J. (EDITORS) (2020), 3D Data Container Engineering Report, OGC Document 20-029, 22.10.2020, <https://docs.ogc.org/per/20-029.html>
- REED, C., BELAYNEH, T. (EDITORS) (2020), OGC Indexed 3D Scene Layer (I3S) and Scene Layer Package Format Specification, Version 1.1, OGC Document 17-014r7, 8.2.2020, <https://www.ogc.org/standards/i3s>

¹ <https://deck.gl/>

Langzeitspeicherung von Geodaten

Ralf ROSCHER

Solution Manager Public Sector · AED-SICAD GmbH · Mallwitzstraße 1-3 · 53177 Bonn

E-Mail: ralf.roscher@aed-sicad.de

1 Einführung

Geodaten unterliegen einem beständigen Aktualisierungs- und Änderungsdruck, der dazu führt, dass bei den Geodatenproduzenten fortlaufend und in zunehmendem Maße Daten verwaltet werden müssen, die sich in den zuständigen Führungssystemen nicht im täglichen Zugriff und damit nicht mehr im Blick befinden. Hinzu kommt begleitend der technische Fortschritt der zur Ablage verwendeten Speichersysteme und -formate. Beide Aspekte begründen die Notwendigkeit, Geodaten in einem hierfür geeigneten System über lange Zeiträume sicher ablegen und dem digitalen Vergessen entziehen zu können.

LZS, das Langzeitspeichersystem der VertiGIS, derzeit erstellt in Kooperation mit dem LGV² Hamburg und der SER, dem Hersteller des europaweit führenden ECM³ DOXIS4, implementiert eine -anerkannten Leitlinien folgende- Datenerhaltungsstrategie zur langzeitstabilen Ablage von Geodaten.

2 Zusammenfassung des Vortragsinhalts

Der Vortrag legt zunächst die Motivation und Notwendigkeit zur Entwicklung eines Langzeitspeichersystems dar, indem die Problematik des digitalen Vergessens beleuchtet und erläutert wird. Es folgt die Ableitung einer Strategie zur langzeitstabilen Ablage von Geodaten, die sich zum einen aus den vielfältigen Erfahrungen des Projektpartners SER zur langzeitstabilen Dokumentenverwaltung ableitet und sich zum anderen auf den Empfehlungen der Arbeitsgruppe der AdV⁴ und KLA⁵ zur Archivierung von Geodaten abstützt.

Einer nachfolgenden Vorstellung des Projekts und der Projektpartner schließt sich die Vorstellung der Lösung sowie der zentralen Lösungseigenschaften an. Dabei folgt sowohl die initiale Skizzierung wie die spätere Evaluierung der gewünschten Eigenschaften entlang der Geodatensätze, die seitens des LGV heute und zukünftig in der LZS verwaltet werden sollen.

² Landesbetrieb für Geoinformation und Vermessung

³ Enterprise Content Management System

⁴ Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland

⁵ Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder

Digitalisierungsspeicher am IT-DLZ – Archiv oder Register?

Christian STANGL

Landesamt für Digitalisierung, Breitband und Vermessung - IT-Dienstleistungszentrum des Freistaats Bayern ·
St.-Martin-Straße 47 · 81541 München
E-Mail: christian.stangl@ldbv.bayern.de

1 Einführung ins Thema

Als Archivierung definiert das Bayerische Archivgesetz die Aufgabe, das Archivgut zu erfassen, zu übernehmen, auf Dauer zu verwahren und zu sichern, zu erhalten, zu erschließen, nutzbar zu machen und auszuwerten. Als Archivgut werden dabei archivwürdige Unterlagen einschließlich der Hilfsmittel zu ihrer Nutzung verstanden, beispielsweise Akten, Urkunden, Karten, Pläne, Bild-, Film- und Tonmaterial einschließlich der zu ihrer Auswertung erforderlichen Programme. Archivgut wird zeitlich unbegrenzt aufbewahrt und nicht mehr verändert.

Ein Register im Sinne einer Registratur (auf die im Registerrecht geregelten öffentlichen Register wie z.B. dem Grundbuch oder Handelsregister wird hier nicht eingegangen) dient hingegen der geordneten Ablage oder Speicherung von zunächst noch in Bearbeitung befindlichen Unterlagen. In deren weiterem Lebenszyklus kommen mit zunehmender Inaktivität der Informationen Einrichtungen wie Altregistratur oder die sogenannte Langzeitspeicherung zum Einsatz.

In der Praxis wird der Archiv-Begriff vor allem im IT-Bereich häufig verwässert. Es gibt Website-Archive und „Archivdateien“ (z.B. im TAR- oder ZIP-Dateiformat). Im Zusammenhang mit digitalen Registratur-Systemen wird von elektronischer Archivierung gesprochen und Datensicherungssysteme nennen eine längerfristige Aufbewahrung oft eine Archivfunktion.

Der Vortrag gibt einen Überblick, wie der Digitalisierungsspeicher am IT-DLZ in diesem Umfeld einzuordnen ist und Verwendung finden kann.

2 Zusammenfassung des Vortragsinhalts

2.1 Der Digitalisierungsspeicher am IT-Dienstleistungszentrum des Freistaats Bayern

Der Digitalisierungsspeicher am IT-DLZ ist weder ein Register noch ein Archiv, sondern ein Speichersystem vom Typ eines Object Storage. Er wurde im Jahr 2017 eingeführt, um die steigende Nachfrage nach einer Speicherlösung für sehr große Datenmengen im Petabyte-Bereich erfüllen zu können. Bei diesen Daten handelt es sich um überwiegend inaktive Bestände ohne sehr hohe Performance-Anforderungen. Der Digitalisierungsspeicher schließt die Lücke zwischen dem leistungsorientierten Block- und File-Storage (SAN-Speicher und NAS) und der Tape-Technik und bildet einen Kompromiss hinsichtlich Leistungsfähigkeit und Kosten. Mit seinen besonderen Eigenschaften kann er helfen, Anforderungen aus dem Archiv- und Registerbereich zu erfüllen. Die nachfolgende Abbildung 1 gibt einen Überblick, wie sich der Object Storage von einem Block Storage (SAN-Speicher) und einem File Storage (NAS) unterscheidet.

	Block Storage: „SAN-Speicher“	File Storage: „Fileservice“	Object Storage: „Digitalisierungsspeicher“
Zugriffsmethode	Über SAN (FC-P) mit SCSI-Kommandos	Über LAN/WAN (TCP/IP) mit SMB (CIFS) oder NFS	Über LAN/WAN (TCP/IP) mit HTTPS (S3, Swift, REST)
Metadaten	-	Systemmetadaten	plus Anwendungsmetadaten
Daten-Update	In-place	In-place	Neue Objektversion
Datenabsicherung	<ul style="list-style-type: none"> Synchrone Datenspiegelung Backup: klassische Datensicherung 	<ul style="list-style-type: none"> Synchrone Datenspiegelung (optional ungespiegelt) Backup: Remote-Snapshots 	<ul style="list-style-type: none"> 3-fache Speicherung (betriebsredundant) Versionierte Speicherung
Performance	Sehr hoch	Hoch	Mittel
Kosten	Hoch	Hoch	Niedrig
Anwendungsgebiete	<ul style="list-style-type: none"> Sehr hohe Leistungsanforderungen (z.B. Datenbank- und Transaktionsanwendungen) Lokaler Speicherzugriff (SAN) 	<ul style="list-style-type: none"> Gemeinsame Nutzung von unstrukturierten Daten mit Schutzmechanismus gegen simultane Änderungen an der gleichen Datei 	<ul style="list-style-type: none"> Unstrukturierte Daten, die nicht oder nur selten geändert werden (z.B. Bilder, Videos, Audiodateien, Logfiles, Sensordaten)

Abb. 1: Speicherarten (Disk) im Vergleich

2.2 Die Hitachi Content Platform (HCP)

Realisiert ist der Digitalisierungsspeicher im IT-DLZ mit der Hitachi Content Platform (HCP). Drei solche Systeme in unterschiedlichen Brandabschnitten an zwei Standorten bilden einen betriebsredundanten Verbund. Die Hardware besteht unter anderem aus untereinander vernetzten Storage Nodes, in denen Magnetplatten mit einer hohen Kapazität von bis zu 16 TB sehr kompakt verbaut sind, womit eine Speicherdichte von 12,8 PB/m² Rechenzentrumsfläche erreicht wird. Skalierbar ist das System bis zu einem Speichervolumen von 1,2 EB. Die Architektur entspricht in Grundzügen der der großen Cloud-Storage-Provider. Dazu passt auch, dass der bevorzugte Zugriff auf Basis von HTTPS mit dem S3-Protokoll erfolgt, welches von Amazon Web Services (AWS) für den Filehosting-Dienst Simple Storage Service (S3) entwickelt wurde und sich mittlerweile zum Industriestandard entwickelt hat.

Die Systemsoftware der HCP hat ihren Ursprung in der Hitachi Content Archive Platform (HCAP), die einen Fokus auf das Thema Compliance und die Unveränderbarkeit der gespeicherten Informationen hatte. Daher verfügt die HCP über viele Eigenschaften und Funktionen, die im Rahmen von Register- und Archivanwendungen relevant sein können:

- Die Mandantenfähigkeit des Systems und die Sicherstellung der Datenintegrität ist geprüft nach IDW PS 880, einem Standard des Instituts der Wirtschaftsprüfer in Deutschland e.V. (IDW). Unter anderem gibt es laufende Integritätsprüfungen anhand von Hash-Werten.
- Die Datenobjekte können mit umfangreichen anwendungsspezifischen Metadaten und Zugriffsregeln angereichert werden (siehe Abbildung 2).
- Das System stellt die Einhaltung von Aufbewahrungsfristen sicher. Im optionalen Compliance Mode können Objekte vor Erreichen des Verfallsdatums nicht gelöscht werden. Zusätzlich kann ein automatisches Schreddern der Daten (Überschreiben des Speicherplatzes mit Bit-Mustern) nach Löschung eingerichtet werden.
- Durch die dreifache und versionierte Objektspeicherung wird ein sehr hohes Maß an Datensicherheit erreicht.
- Spezielle Dienste zur Indexierung und Suche ermöglichen ein schnelles Auffinden von Informationen anhand der Metadaten.

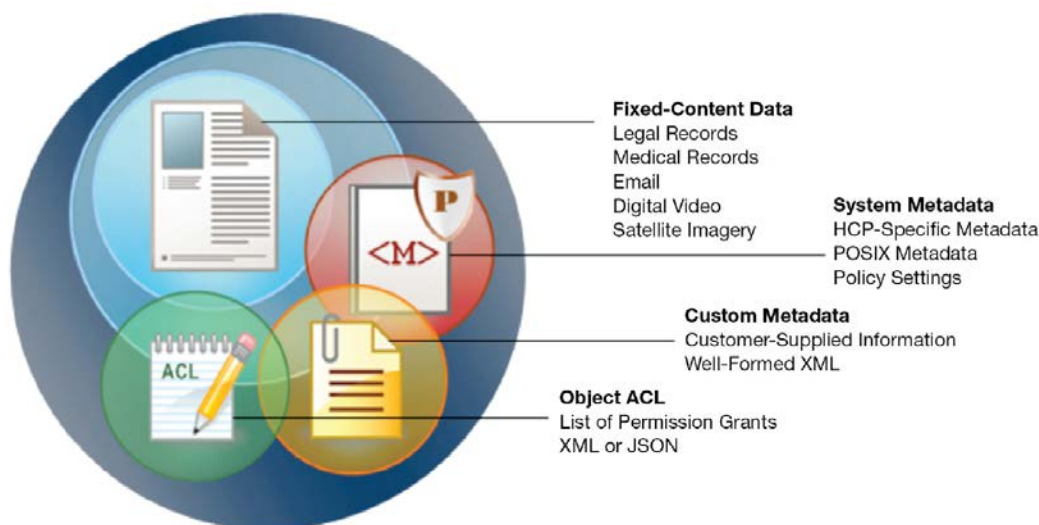


Abb. 2: HCP Object (Quelle: Hitachi Vantara)

Literatur

- BUNDESMINISTERIUM DES INNEREN (2014), Organisationskonzept elektronische Verwaltungsarbeit - Baustein E-Langzeitspeicherung, https://www.verwaltung-innovativ.de/SharedDocs/Publikationen/Organisation/e_langzeitspeicherung.pdf
- RATNER, M. (2014), The Fundamentals of Hitachi Content Platform, <https://community.hitachivantara.com/s/article/better-object-storage-with-hitachi-content-platform>
- HITACHI VANTARA (2020), Hitachi Content Platform Architecture Fundamentals, <https://www.hitachivantara.com/en-us/pdf/white-paper/content-platform-architecture-fundamentals-whitepaper.pdf>
- BEP SYSTEMS LIMITED (2019), Hitachi Content Platform and Compliance Obligations in the European Union, <https://www.hitachivantara.com/en-us/pdf/white-paper/hcp-compliance-obligations-in-european-union-whitepaper.pdf>
- HITACHI VANTARA (2017), Using Hitachi API for Amazon S3, https://knowledge.hitachivantara.com/Documents/Storage/Content_Platform/7.1.2/Manage_an_HCP_system/Using_Hitachi_API_for_Amazon_S3

Kooperation der staatlichen Archivverwaltungen mit der AdV bei der Archivierung von Geobasisdaten

Dr. Bernhard GRAU

Bayerisches Hauptstaatsarchiv · Schönfeldstraße 5 · 80539 München

E-Mail: bernhard.grau@bayhsta.bayern.de

1 Einführung ins Thema

Elektronische Geobasisdaten unterliegen wie alle Verwaltungsunterlagen der in den Archivgesetzen des Bundes und der Länder verankerten Anbietepflicht. Das bedeutet, dass Sie zum Zeitpunkt ihrer Entbehrlichkeit dem zuständigen Archiv zur Übernahme angeboten werden müssen. Ziel der elektronischen Archivierung ist die dauerhafte, das heißt zeitlich unbefristete Verwahrung der als archivwürdig eingestuften Informationen! Damit unterscheidet sich die Archivierung von der Langzeitspeicherung bei den abgebenden Stellen, die darauf abzielt, die Nutzbarkeit dieser Daten für den Zeitraum sicherzustellen, in dem sie für die laufenden Verwaltungsgeschäfte noch benötigt werden. Mit der Übergabe ans Archiv werden die ausgesonderten Geobasisdaten Archivgut und unterliegen den rechtlichen Bestimmungen der Landesarchivgesetze. Die abgebende Stelle behält jedoch ein Rückgriffsrecht auf die von ihr übergebenen Unterlagen.

Die Archivierung elektronischer Geobasisdaten stellt wegen der Art und der Qualität dieser Informationen, aber auch wegen des steten technologischen Wandels eine besondere Herausforderung dar. Zumindest bei der Aussonderung und Übergabe der Geobasisdaten ist ein enges Zusammenwirken der Vermessungsverwaltungen und der Archive von entscheidender Bedeutung. Hinzu kommt, dass sich die Herausforderungen bei Geobasisdaten, die als AdV-Produkte definiert sind, in allen Bundesländern gleichartig gestalten. Es lag daher nahe, die mit der Archivierung verbundenen Fragen und Probleme auf Bund-Länder-Ebene in einer gemeinsamen Arbeitsgruppe zu erörtern.

2 Zusammenfassung des Vortragsinhalts

2.1 Problemstellung

Die staatlichen Archive der Länder verwahren amtliche Karten und Pläne, die zum Teil bis zum Beginn der Frühen Neuzeit zurückreichen. Dabei handelt es sich zunächst um anlassbezogen erstellte handgezeichnete Pläne und gemalte „Augenschein“-Darstellungen. Mit Beginn der systematischen, mit modernen und exakten Methoden arbeitenden Landesvermessung nimmt die Überlieferung der Karten und Pläne in Ihrem Umfang erheblich zu. Die Umstellung auf elektronische Geoinformationssysteme bedeutet erneut eine erhebliche Steigerung der Informationsmenge. Sie wirft in Hinblick auf die Archivierung neuartige Fragen auf und stellt ganz neue Anforderungen an den dauerhaften Erhalt der elektronischen Kartenwerke.

2.2 Grundlagen und Methoden der digitalen Archivierung

Die Archive des Bundes und der Länder haben nach einer längeren Phase des konzeptionellen Austauschs begonnen „Digitale Archive“ aufzubauen und archivieren inzwischen auch elektronische Unterlagen. Die funktionale Ausgestaltung der Digitalen Archive richtet sich dabei durchwegs nach dem so genannten OAIIS-Modell. Für den dauerhaften Erhalt der Lesbarkeit setzen die öffentlichen Archive in Deutschland fast durchwegs auf die Migrationsstrategie. Demnach sind die archivierten Daten regelmäßig in neue Datenformate zu überführen, um deren Lesbarkeit und Interpretierbarkeit trotz des steten informationstechnologischen Wandels sicherzustellen. Dieses Vorgehen führt zu einer Veränderung der Dokumenteneigenschaften und birgt die Gefahr von Informationsverlusten. Durch eine detaillierte Dokumentation der Übernahme- und Erhaltungsprozesse soll die Nachvollziehbarkeit der zum Erhalt der Lesbarkeit erforderlichen Eingriffe gewährleistet und das Vertrauen in die Authentizität der elektronischen Überlieferung sichergestellt werden.

2.3 Die Archivierung elektronischer Geobasisdaten

Elektronische Geobasisdaten weisen spezifische Eigenschaften auf, die sie von üblichen Verwaltungsdaten unterscheiden. Sie werden in Fachverfahren vorgehalten, laufend aktualisiert und – als Vektordaten - in spezifischen, sonst nicht gebräuchlichen Formaten vorgehalten. Eine bundesweite Vereinheitlichung brachte der Umstieg auf ein gemeinsames Datenmodell und gemeinsame Fachanwendungen (AFIS, ALKIS, ATKIS) mit sich. Geobasisdaten, die länderübergreifend aktuell und nach einheitlich Standards bereitgestellt werden sollen, hat die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) als AdV-Produkte definiert und standardisiert. Es liegt nahe, diese auch länderübergreifend nach abgestimmten Prinzipien auszusondern und zu archivieren. Pendant zur AdV und

damit deren Ansprechpartner ist auf archivischer Seite die KLA (Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder).

2.4 Die Arbeitsgruppe „Archivierung von Geobasisdaten“

Auf Initiative des Präsidenten der Landesvermessung und Geobasisinformation Brandenburg haben AdV und KLA 2013 eine gemeinsame Arbeitsgruppe eingesetzt. Ihre Aufgabe war es vor allem, eine Bestandsaufnahme der amtlichen Geobasisdaten vorzunehmen, die einzelnen Produkte auf ihre Archivwürdigkeit hin zu prüfen sowie einheitliche Stichzeitpunkte und Datenformate für die Archivierung zu definieren. Ihre Ergebnisse hat die Arbeitsgruppe im Jahr 2015 in den „Leitlinien zur bundesweit einheitlichen Archivierung von Geobasisdaten“ niedergelegt. Sie wurden im Druck veröffentlicht sind aber auch über die Homepages von AdV und KLA abrufbar.

2.5 Die Leitlinien zur bundesweit einheitlichen Archivierung von Geobasisdaten

In der Arbeitsgruppe bestand Übereinstimmung, dass historische Geobasisdaten für Verwaltung und Forschung auf Dauer von unverzichtbarem Wert sind. Um eine sichere und dauerhafte Archivierung gewährleisten zu können, empfiehlt sich eine frühzeitige Anmietung und Übergabe, da nur so sichergestellt ist, dass keine Informationsverluste eintreten und die Daten in einem für die Archivierung geeigneten Datenformat ausgespielt und an das zuständige Archiv übergeben werden können. Als Kriterien für die Archivwürdigkeit von Geobasisdaten wurden die hohe inhaltliche Relevanz, ein möglichst feiner Detaillierungsgrad, der Nutzen für die Auswertung anderer Geodaten und die technikhistorische Bedeutung definiert. Um zu gewährleisten, dass die archivierten Geobasisdaten länderübergreifend nutzbar bleiben, hat die Arbeitsgruppe einheitliche Zeitschnitte definiert, zu denen sie von allen Landesarchiven übernommen werden sollen. Abhängig vom jeweiligen Produkt wurden außerdem Datenformate und Metadaten sets festgelegt, die bei der Aussonderung der Geobasisdaten an das jeweils zuständige Archiv zu übergeben sind.

2.6 Offene Fragen und Probleme

Fünf Jahre nach der Veröffentlichung der Leitlinien hat die Arbeitsgruppe ihre Tätigkeit wie vorgesehen wieder aufgenommen, um das Papier aus dem Jahr 2015 noch einmal zu evaluieren. Die seinerzeit ausgesprochenen Empfehlungen wurden dabei weitgehend bestätigt.

Neu in den Blick genommen wurde die Frage der signifikanten Eigenschaften. Damit sind diejenigen Dateieigenschaften gemeint, die auf Dauer, das heißt auch über die zu erwartenden Formatmigrationen hinaus erhalten werden müssen, damit die archivierten Geodaten als integer und authentisch gelten können.

Weiter erörtert die Arbeitsgruppe nun auch die mit dem Themenkomplex „Urheberrecht, Datenschutz und doppelte Datenhaltung“ zusammenhängenden Rechtsfragen. Dies schließt die kontrovers diskutierte Frage ein, zu welchem Zeitpunkt, Geobasisdaten an die Archive abzugeben sind, und ob die abgebende Stelle verpflichtet ist, die ausgesonderten Informationen im Anschluss aus den eigenen Fachanwendungen zu löschen.

Uneinheitlich geregelt ist von Bundesland zu Bundesland die Frage, ob historische Geobasisdaten, die den Archiven zur Archivierung überlassen wurden, Geodaten im Sinne der INSPIRE-Richtlinien sind und demzufolge Inspire-konform zugänglich gemacht werden müssen oder nicht!

Schließlich bleibt festzustellen, dass sich die Arbeitsgruppe darauf beschränkt hat, Vorschläge für die so genannten AdV-Produkte zu erarbeiten. Darüber hinaus fallen bei den Vermessungsverwaltungen, aber auch bei vielen anderen Fachverwaltungen Geofachdaten an, die meist auf den Geobasisdaten aufsetzen, und ebenfalls der Archivierung harren!

Literatur

HANDBUCH - Archivierung von Geodaten des Bundes, herausgegeben vom Bundesamt für Landestopografie swisstopo und dem Schweizerisches Bundesarchiv BAR, Version 1.0 vom 13. Dezember 2016, abrufbar über folgenden Link: <https://www.swisstopo.admin.ch/de/wissen-fakten/geschichte-sammlungen/landschaftsgeadaechtnis/langzeitaufbewahrung/archivierung-geodaten.html>.

HANDREICHUNG zur Archivierung elektronisch vorliegender Geodaten. Gemeinsames Papier der ARK-Arbeitsgruppe AG ESys und des ARK IT-Ausschusses, o.O., 2009, abrufbar über folgenden Link: <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/handreichung-geodaten.pdf?blob=publicationFile>.

LEITLINIEN zur bundesweit einheitlichen Archivierung von Geobasisdaten. Abschlussbericht der gemeinsamen AdV-KLA-Arbeitsgruppe „Archivierung von Geobasisdaten“ 2014-2015, o.O., o.J. [2015]; abrufbar über folgenden Link: <https://www.bundesarchiv.de/DE/Content/Downloads/KLA/leitlinien-geobasisdaten.pdf?blob=publicationFile>.

Vorstellung des neuen Leitfadens mobile GIS – Von der GNSS-basierten Datenerfassung zu Mobile Mapping, Version 4.0

Vincent FRANKE

Institut für angewandte Informatik / Geographie · Universität Augsburg · Universitätsstraße 2 · 86159 Augsburg
E-Mail: vincent.franke@gi-geoinformatik.de

1 Einführung in das Thema

Der Leitfaden mobile GIS wird in Zusammenarbeit mit zahlreichen Forschungsinstituten, privatwirtschaftlichen Betrieben und öffentlichen Einrichtungen aus dem Netzwerk des Runder Tisch GIS e.V. erstellt. Die Herausgeber, bestehend aus der Technischen Universität München, dem Runder Tisch GIS e.V. und der GI Geoinformatik Augsburg GmbH bieten bereits seit 2013 regelmäßig aktualisierte Fassungen kostenlos und online abrufbar an.

Der Leitfaden bietet allgemeine und spezialisierte Fachbeiträge rund um die Themen der mobilen Datenerfassung und Geoinformationssysteme. Damit bildet er eine regelmäßige Markt- und Forschungsübersicht aktueller Themen und stellt Praktikern aus diversen Branchen relevantes Fachwissen zur Verfügung.

Die Inhaltliche Struktur setzt sich aus Forschungs- und Fachbeiträgen zu aktuellen und grundsätzlichen Themen der Geoinformatik zusammen. Hierzu zählen insbesondere GNSS-Positionierungstechnik, mobile Datenerfassung, IT-Plattformen und Sicherheit sowie Geobasisdaten und Open Data. Darüber hinaus bieten Entscheidungshilfen Unterstützung bei der Wahl der passenden Hard- und Software. Dies wird durch anschauliche Anwenderbeispiele aus Privatwirtschaft und öffentlichem Dienst ergänzt.

In der zweiten Hälfte finden sich umfangreiche Produktaufstellungen von einer breiten Auswahl an Händlern und Produzenten des deutschsprachigen Marktes. Aktuelle Hard- und Software kann komfortabel auf einen Blick verglichen werden, direkte Kontakte zu Hersteller und Vertrieb finden sich für jeden Eintrag.

Zuletzt sind auch Handlungsempfehlungen für Einsteiger und Fortgeschrittene insbesondere bezüglich Kosten und Investitionssicherheit enthalten.



Abb.1: Titelblatt des Leitfadens Mobile GIS in der 7. Auflage und Auszüge der Produktübersicht

2 Zusammenfassung des Vortragsinhalts

Zum September 2020, dem Auftakt der INTERGEO, wurde der Leitfaden hinsichtlich der dynamischen Entwicklung des Mobile GIS Umfelds umfassend aktualisiert. Die mittlerweile 7. Auflage erschien in der Version 4.0 mit maßgeblichen Neuerungen, insbesondere in den folgenden Bereichen:

- Positionierungstechnologien und Satellitennavigationssysteme
- Mobile Mapping, UAV, airborne Laserscanning
- Terrestrisches Laserscanning
- Open Data
- Neuprodukte aus dem Bereich GNSS und robuste Tablets

Dabei wurden im Rahmen der Aktualisierung alle Fachbeiträge überarbeitet. Die oben genannten Schwerpunkte fanden besondere Beachtung und bieten einen hochaktuellen Überblick. Die Themen sollen im Rahmen des Vortrags kurz vorgestellt werden.

Außerdem finden sich in der neuen Fassung insgesamt 20 Hard- und 49 Softwareprodukte aus den aktuellen Angeboten diverser Anbieter. Hier sind ebenfalls wichtige Neuerungen vertreten. Diese sind jeweils mit einheitlichen Kriterien bewertet worden und können untereinander verglichen werden.

Zuletzt sollen im Vortrag auch kurz die Vorüberlegungen zu Hard- und Software Produkten Erwähnung finden um Einsteigern die Marktübersicht zu erleichtern.

Literatur

https://rundertischgis.de/publikationen/leitfaeden.html#a_mobilegis

Neue Wege bei den Vor-Ort-Kontrollen zur Agrarförderung

Die FotoApp – eine Ergänzung zum künftigen Flächenmonitoring

Julian SCHMID

GI Geoinformatik GmbH · Morellstraße 33 · 86159 Augsburg
E-Mail: j.schmid@gi-geoinformatik.de

1 Hintergrund und Rahmenbedingungen

Landwirte von heute sind nicht nur Lebensmittelproduzent. Sie sollen nicht nur kosteneffizient, sondern auch nachhaltig arbeiten. Gleichzeitig sollen die Qualität unserer Böden und die Artenvielfalt erhalten bleiben. Mit der Unterstützung zum Erhalt der landschaftlichen Vielfalt und des Artenschutzes tragen die Akteure viel zur Gestaltung unserer Landschaft bei.

Diese Einhaltung von Umweltschutzstandards wird durch Förderungen im Zuge der EU-Agrarpolitik unterstützt. Die EU-Agrarförderung basiert auf der 1962 eingeführten Gemeinsamen Agrarpolitik (GAP), die den gesetzlichen Hintergrund für eine entsprechende finanzielle Unterstützung bildet. Um Missbrauch bei diesen Förderungen zu unterbinden sind entsprechende Kontrollvorgänge zwingend vorgeschrieben.

Die einhergehenden Kontrollen erfolgten bislang mittels Fernerkundung und Kontrollen durch Mitarbeiter der Agrarverwaltung vor Ort. Dabei wurden nach einem vorgegebenen Verfahren u.a. ein Teil der Antragsflächen kontrolliert, oder auch bauliche Prüfungen vorgenommen. Zukünftig soll eine flächendeckende Prüfung der Antragsflächen mittels Flächenmonitorings stattfinden wodurch sich auch die einzelnen Prüfvorgänge ändern werden. Neben der Intensivierung von automatisierten Fernerkundungsmethoden (z.B. mittels Sentinel-2 Daten) ist es das Ziel künftig die Landwirte selbst in den Prozess der Kontrolle und Klärung miteinzubeziehen. Um diese Aufträge einfach, schnell und effizient abwickeln zu können, wird für die Agrarverwaltung eine FotoApp benötigt, die auch die Landwirte bei der Antragsstellung unterstützt und begleitet. Im vorliegenden Abstract wird die Umsetzung dieser App vorgestellt.

2 Zusammenfassung des Vortragsinhalts

2.1 Technische Umsetzung

Das Land Niedersachsen hat bei der Entwicklung dieser App die Vorreiterrolle innerhalb der Bundesländer übernommen und startet mit der Beauftragung zur Entwicklung der FotoApp unter dem Namen FANi (= Fotos Agrarförderung Niedersachsen) in das Pilotprojekt ‚Flächenmonitoring‘.

Das niedersächsische Ministerium für Ernährung, Landwirtschaft und Verbraucherschutz beauftragte die Entwicklung der App. Die inhaltliche Ausarbeitung erfolgte in Zusammenarbeit mit der Landwirtschaftskammer Niedersachsen (LWK) sowie dem Servicezentrum Landentwicklung und Agrarförderung (SLA). Die technische Umsetzung und Entwicklung der App übernahm die GI Geoinformatik GmbH.

Bei der Entwicklung wird auf die Frameworks Xamarin und .Net gesetzt. Als Entwicklungsbasis dient das ArcGIS Runtime SDK for .NET (Esri Inc.). Durch das plattformübergreifende Xamarin-Framework kann die App mit verringertem Aufwand für verschiedene Plattformen (Android und iOS) den Landwirten bereitgestellt werden.

Bei der Umsetzung und Ausgestaltung der App wird auf einen agilen Ansatz, angelehnt an das SCRUM-Verfahren, gesetzt. Dies hat den Vorteil, dass kleinere thematische Einheiten umgesetzt und frühzeitig als lauffähige Version bereitgestellt werden können. Da die Umstellung hin zum Flächenmonitoring noch nicht abgeschlossen ist und dadurch auch für die App laufend neue Anforderungen entstehen, können diese bei Bedarf auch kurzfristig berücksichtigt und implementiert werden. Seit Mitte letzten Jahres steht die App einem ausgewählten Nutzerkreis zur Verfügung, um so erste Erfahrungen sammeln zu können. In Kürze erfolgt die Veröffentlichung der App und steht dann den Landwirten in Niedersachsen kostenfrei zur Verfügung.

2.2 Workflow

Sind bei Anträgen zur Agrarförderung seitens des Landwirts Nachweise erforderlich, so können diese zukünftig über die FotoApp (FANi) eingereicht werden. Mit Hilfe der App können verortete Fotos als Nachweis für den Zustand, die Nutzung oder eine Tätigkeit auf den eigenen Flächen bereitgestellt werden. Dazu zählen z. B. die angebaute Kultur, das Vorhandensein eines mobilen Stalls oder die zu einem bestimmten Zeitpunkt vorgeschriebene Mahd. Neben den Fotos von Flächen können Dokumenten als geeigneter Nachweis eingereicht werden.

Der Landwirt reicht seine Anträge wie gewohnt über die bestehenden Systeme ein. Diese werden durch die Fachstelle geprüft. Sind weitere Nachweise erforderlich, werden für den Landwirt diese als sogenannte Aufträge für die App bereitgestellt.

Über eine nutzerspezifische Kennung kann sich der Landwirt in der App anmelden und erhält seine Liste der zu erbringenden Aufträge. Dabei erfolgt auch ein Abruf der aktuellen Prüfregeln und Konfigurationen.

In einer interaktiven Kartenansicht werden die betroffenen Flächen und Standorte dargestellt und sind mit den jeweiligen Aufträgen verknüpft. Zur besseren Orientierung in der Karte werden Luftbilder und topographische Karte bereitgestellt. Innerhalb der Kartenansicht wird der aktuelle Standort angezeigt.

Der Anwender kann somit gezielt zwischen der Liste und Kartenansicht und den einzelnen Aufträgen wechseln und diese auswählen. Je nach Prüfauftrag muss der Landwirt eine unterschiedliche Anzahl an Fotos erfassen. Die Fotoerfassung kann nur erfolgen, wenn sich der Anwender innerhalb eines vorgegebenen Pufferbereichs zur jeweiligen Fläche befindet. Dabei gibt es verschiedene Vorgaben, z.B. Detailgrad – Nahaufnahme/Überblick, die berücksichtigt werden müssen.

Sind die Anforderungen erfüllt, kann der Landwirt die Fotos hochladen und im weiteren Verlauf einsehen, wie der Bearbeitungsstand zu seinen Aufträgen ist. Über die App können Antragsteller somit den Prozess der Antragsprüfung unterstützen, und aktiv bei Fragen mitwirken.

2.2 Besonderheiten

Bei der Umsetzung zur App gibt es neben einer Vielzahl an technischen Anforderungen, auch den Anspruch die Antragsteller im Verfahren zu beteiligen und hat somit die Anforderung, die Landwirte bei der Bedienung der App bestmöglich zu unterstützen und zu führen.

Nachfolgend eine Auswahl der wichtigsten technischen Anforderungen zur Umsetzung der App:

- Sicherheit und Datenschutz
- Offline-Nutzung
- Standort-Verifizierung
- Prüfung der Bildqualität (Schärfe, Kontrast, Belichtung)
- Speicherung von umfassenden Metadaten (z.B. Blickrichtung, Auflösung Bild, Standort, Zeit, Datum, etc.)
- Konfigurierbar und flexibel, z.B. bei geänderten Prüfregeln

Für den Erfolg einer solchen App spielt aber auch immer die Usability eine entscheidende Rolle. Ziel war es, dass die App anwenderfreundlich ist und sich ohne große Einarbeitung bedienen lässt. Deshalb lag bei der Entwicklung auch ein Schwerpunkt auf eine bestmögliche Anwenderunterstützung.

Die Landwirte werden bei der Bearbeitung der einzelnen Aufträge durch klare Menüstrukturen geführt und mittels Hinweistexte und Symbole innerhalb der App unterstützt. Zu jedem anfallenden Auftrag gibt es Kurzbeschreibungen zur Vorgehensweise und Hinweis zu den Anforderungen bei der Fotoerfassung. Eine klare Bildschirmaufteilung und selbst-erklärende Symbole vereinfachen die Orientierung und erleichtern die Ausführung von Aufgaben.

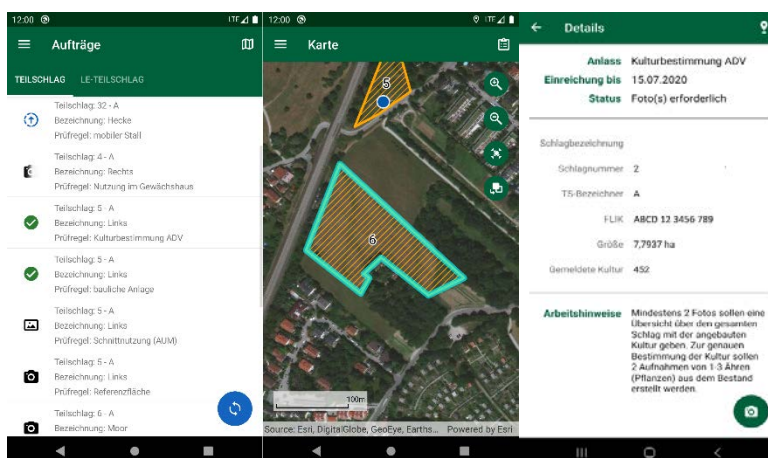


Abb. 1: Screenshots der Software FANi (links: Auftragsliste, mitte: Kartenansicht, rechts: Auftragsdetails)

2.3 Resümee

Die ursprüngliche Zielsetzung, der App, mit geotagged Fotos innerhalb eines vorgegebenen Zeitabschnitts (abhängig vom Reifegrad des Anbaus) den Antragsprozess durch die Beteiligung des Landwirts zu unterstützen, basierte auf den neuen EU-Vorgaben. Neben dem partizipativen Ansatz sollten auch die Aufwände der Vor-Ort-Kontrolle durch die Agrarverwaltung verringert werden. Bedingt durch die Pandemie bekam dieser Ansatz noch zusätzliche Bedeutung. Schnell wurden die Vorteile des agilen Entwicklungsansatzes deutlich, durch die die Projektbeteiligten früh ihre Interessen einbringen konnten. Dies führte dazu, dass der ursprüngliche Ansatz der reinen Fotodokumentation zu einer prozessbegleitenden App für den Antragsdialog erweitert wurde. Zukünftig noch weitere Themenfelder über die App bereitgestellt werden, bzw. bearbeitet werden können. Davon hängt sicherlich auch die Akzeptanz der App bei den Landwirten ab.

Stand: 12.03.2021

HydroVISH & KomVISH on the Web: Visualisierung und Nutzung von 3D-Massendaten aus Vermessung und Modellierung sowie amtlichen Geodaten im digitalen Wasserbau und der digitalen Kommunalverwaltung

Ramona BARAN¹, Werner BENGER¹ und Frank STEINBACHER¹

¹AirborneHydroMapping GmbH · Feldstraße 1b · A-6020 Innsbruck
E-Mail: info@ahm.co.at

1 Digitalisierung im Wasserbau und der Kommunalverwaltung

Im Rahmen der fortschreitenden Digitalisierung von Verwaltungs- und Kommunikationsprozessen gewinnen performante Datenbank- und Softwarelösungen schnell an Bedeutung: Es gilt Entscheidungsprozesse öffentlicher Behörden und Verwaltungen zu stützen und zu beschleunigen. Für den digitalen Wasserbau spielen noch weitere Aspekte eine wesentliche Rolle. Im Rahmen z.B. von Hochwasserschutzmaßnahmen müssen häufig Daten sowohl aus Vermessung, Planung, Unterhaltung und Modellierung zusammengeführt und im Kontext beurteilt werden. Allein die fortlaufende Entwicklung von Sensoren zur flächendeckenden 3D-Gewässervermessung führt zu hochdetaillierten und komplexen Massendaten, deren Größe inzwischen mehrere Terabyte ausmachen können. Am Beispiel: Die topo-bathymetrische Laserscanbefliegung 600 km Elbe entsprechen ca. 5.5 Terabyte an Rohdaten!

Vorgaben und neue technische Möglichkeiten machen Datenbasis immer komplexer – das fordert Behörden, Verwaltungen und Ingenieurbüros gleichermaßen. Damit steigen die Anforderungen an Hard- & Software-Module, die für die Speicherung, Verwaltung und Visualisierung der Daten benötigt werden, schon unter dem Gesichtspunkt der Gewässervermessung. Dazu kommt: Die Wasserrahmenrichtlinie der EU (EU, 2000) erfordert eine regelmäßige Neuvermessung der Gewässer zum Zwecke des Monitorings. Dies betrifft sowohl kommunale Behörden und Verwaltungen wie die Wasserwirtschaftsämter (WWAs) in Bayern als auch staatliche Institutionen wie die Wasserstraßen- und Schifffahrtsverwaltung des Bundes (WSV).

Aber auch jenseits des Wasserbaus steigen die Anforderungen für kommunale Verwaltungen und Ingenieurbüros bei raumbezogenen Aufgabenstellungen: Beispielsweise in der Infrastrukturplanung wie dem Straßenbau, sind die verfügbaren und einer Planung zu Grunde liegenden 3D-Geodaten ebenfalls komplex. So sind über die einzelnen Landesvermessungsämter verschiedene flächendeckende 3D-Geodaten erhältlich wie Gelände-, Oberflächen- und Gebäudemodelle (LoD1 & LoD2) oder eben die eigentlichen Vermessungsinformationen der Laserpunktwolke der jeweiligen Landesbefliegung.

1x1 m Kacheln vs. 3D-Datensätze: Erst Übergreifende 3D-Datensätze ermöglichen Auswertungen und Lösungsvorschläge. Bei den Vermessungsämtern liegen diese Daten nicht als einheitliche Gesamtdatensätze vor, sondern in 1x1 km großen Kacheln. Bei großräumigen Planungsvorhaben müssen deshalb oftmals viele kleinräumige Datensätze aufwändig zusammengeführt werden, um einen übergreifenden Planungsprozess zu ermöglichen. Zusammenhängende 3D-Datensätze sind essenziell, um z.B. bei verschiedenen möglichen Trassenverläufen einer Straße oder Bahnstrecke Konfliktpunkte mit bestehender Bebauung oder Naturschutzflächen zu evaluieren und Lösungsvorschläge zu erarbeiten (Abb. 1). Auch hier sind die Datenmengen beträchtlich. Aber genau dies ist der Normalfall bei der Verarbeitung hoch aufgelöster Laserpunktwolken. Am Beispiel: Die AHM hat für das Landesamt für Digitalisierung, Breitband und Vermessung Bayern (LDBV) die Laser-Daten aus der Landesbefliegung in eine Gesamtpunktwolke zusammengeführt (Abb. 1). 510 Milliarden Punkte, Datengröße 10 Terabyte!

2 KomVISH & HydroVISH für den effizienten Umgang mit 3D-Geomassendaten

Geoinformationssysteme (GIS) als lokal installierte Desktopanwendungen und online verfügbare Web Map Services (WMS) stellen einen etablierten Standard zur raumbezogenen und integrativen 2D-Geodatenanalyse dar, sind aber beim effizienten Umgang mit 3D-Geomassendaten und der integrativen 2D-3D-Analyse limitiert. KomVISH und HydroVISH sind kompakte Softwarelösungen für den Umgang mit 3D-Vermessungsdaten z.B. aus topo-bathymetrischen Laserscanbefliegungen oder 3D-Geodaten wie digitalen Geländemodellen.

- ⇒ **KomVISH** ist auf die Nutzung im kommunalen Umfeld zugeschnitten und erlaubt die effiziente Verknüpfung der 2D-GIS-Welt mit der 3D-Welt durch entsprechende GIS- & WMS-Schnittstellen.
- ⇒ **HydroVISH** bietet größeren Behörden und Verwaltungen sowie Ingenieurdienstleistern die Möglichkeit, ihre Datenprodukte in der Gesamtheit zu verwalten, zu visualisieren und zu verknüpfen.

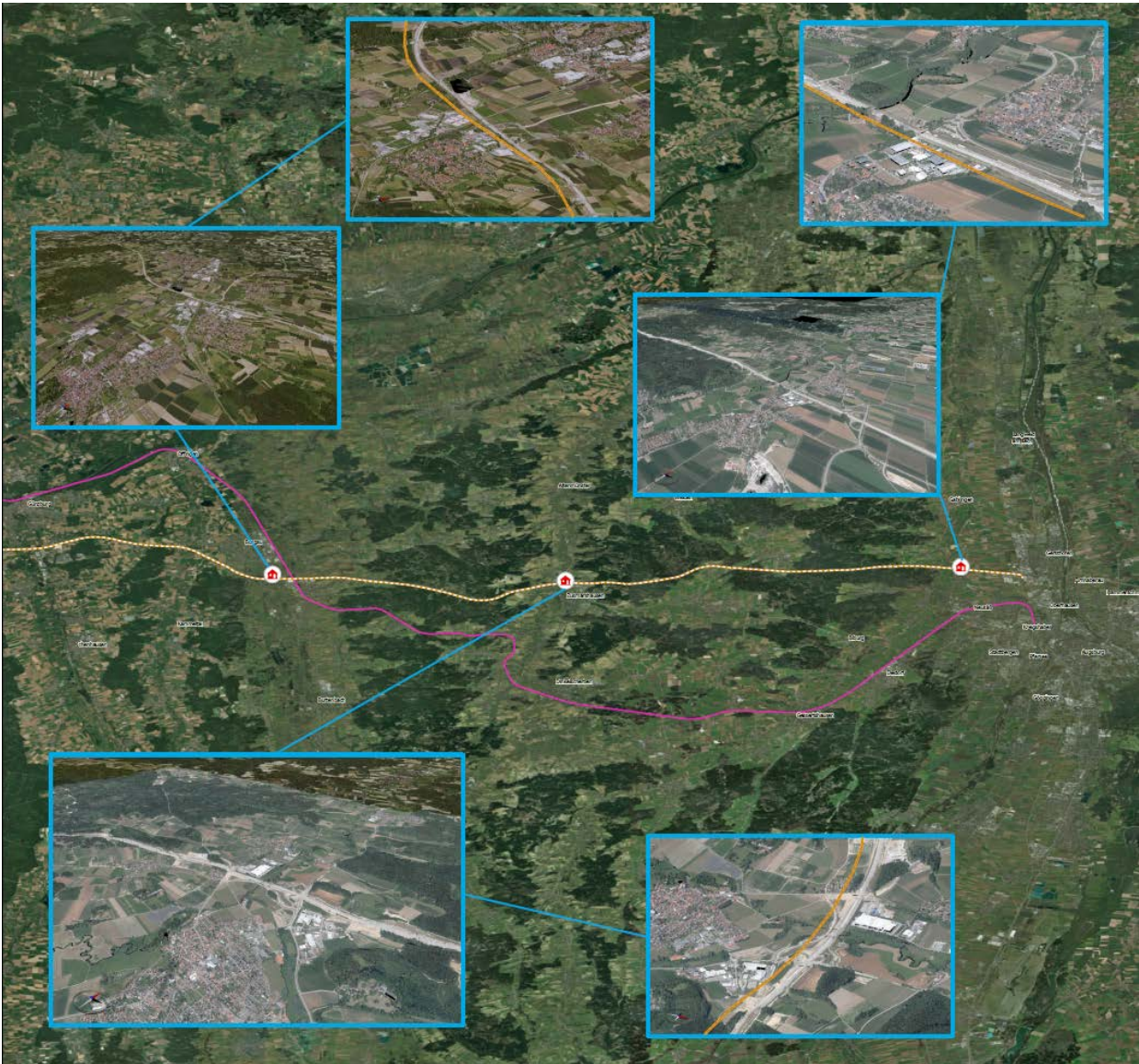


Abb. 1: Ausschnitt RGB-eingefärbte Laserpunktwolke Bayern: Konfliktpunkte zwischen aktueller Bebauung und eines Trassenverlaufs der Bahnstrecke zwischen Augsburg und Ulm parallel zur A8 (orange Linie; LDBV-Pilotprojekt Geomassendaten, Landkreis Augsburg; Koordination Runder Tisch GIS).

2.1 HydroVISH im digitalen Wasserbau für ganzheitliche Betrachtung räumlicher Skalen und zeitlicher Abläufe

In HydroVISH können komplexe wasserbaubezogene Datensätze parallel visualisiert und vertieft analysiert werden. Dazu gehören zu Grunde liegende Messdaten wie Punktwolken aus topo-bathymetrischen Befliegungen oder Echolotungen von Flussabschnitten, hydraulische Berechnungsnetze aus dem Bestand einer Behörde oder neu erstellte Netze (Abb. 2) und zeitabhängige hydraulische Berechnungsergebnisse für z.B. ein HQ100 (Abb. 2 & 3).

Dies bedeutet auch, dass unterschiedliche Datentypen bzgl. Geometrie miteinander betrachtet werden: Punkte vs. Linien wie Bruchkanten vs. Flächen wie ein hydraulisches Dreiecksnetz (Abb. 2). Der Vorteil liegt in der Möglichkeit der ganzheitlichen Betrachtung über räumliche Skalen und den zeitlichen Verlauf eines Hochwasserereignisses hinweg. So kann die Betroffenheit einzelner Gebäude und Grundstücke im Ereignisverlauf besser beurteilt und die Wirkung einzelner Schutzmaßnahmen (Überflutungsflächen in Flussauen, Deichplanung, Rückhaltebecken usw.) zum gleichen Zeitpunkt eines Hochwassers evaluiert werden (Abb. 3). Plausibilitätsprüfungen – beispielsweise die direkte Gegenüberstellung der eigentlichen Messinformation aus einer Punktwolke mit abgeleiteten Daten wie einem Geländemodell oder hydraulischem Berechnungsnetz – sichern die Datenqualität (Abb. 2).

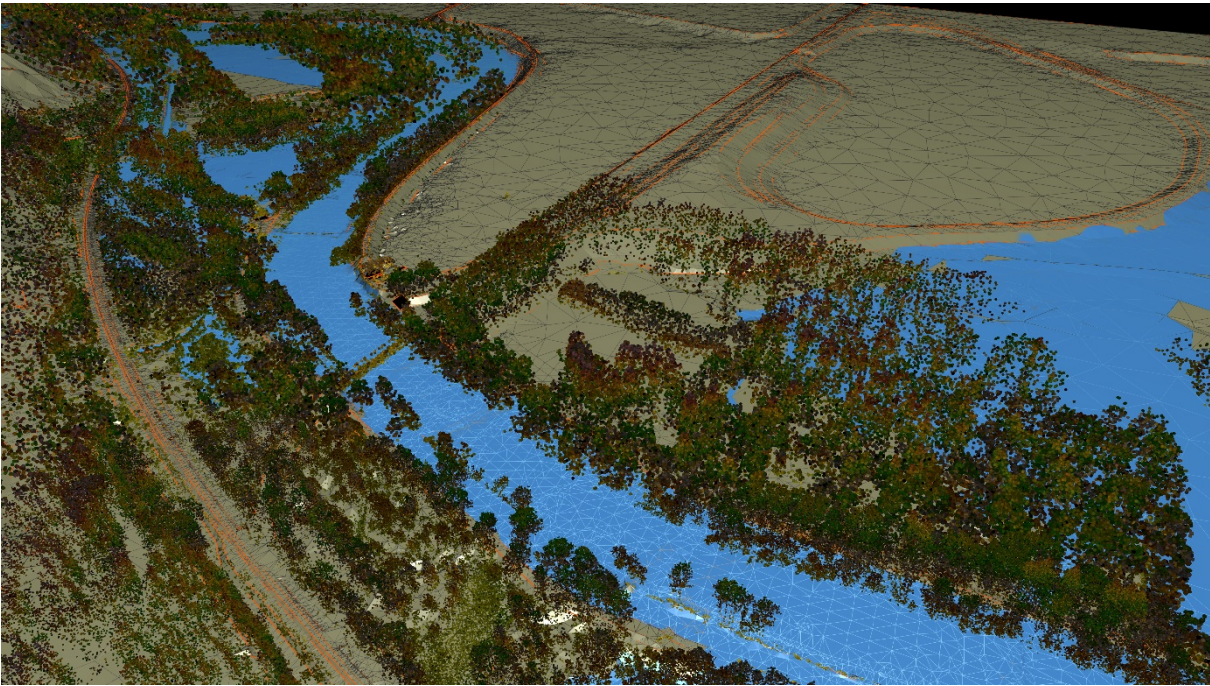


Abb. 2: Visualisierung HydroAS-2d Modellierungsergebnisse (blau) an der Mangfall in HydroVISH zusammen mit CFD-Netz, Punktwolke und Bruchkanten (orange Linien).



Abb. 3: Visualisierung eines HydroAS-2d Modellierungsschritte (hellblau) an der Mangfall in HydroVISH zusammen mit RGB-eingefärbten DOM zur verbesserten Beurteilung des zeitlichen Verlaufes und der räumlichen Ausbreitung eines HQ100-Ereignisses.

2.2 KomVISH für die digitale Kommunalverwaltung mit Echtzeitzugriff auf 3D-Geomassendaten

Geodaten sind für Kommunen und Behörden über die jeweiligen Landesvermessungsämter teils frei verfügbar wie in Thüringen und Nordrhein-Westfalen oder können zumeist im Zuge einer Rahmenvereinbarung wie in Bayern genutzt werden. Dabei orientieren sich Geodaten auch immer stärker an der dritten Dimension. So besteht in Bayern über das LDBV Zugriff auf die landesweiten 3D-Datensätze des LoD2, DGM, DOM oder die Laserpunktwolke (Abb. 1). Ohne eine entsprechende Software bleibt der Zugriff auf diese Massendaten allerdings schwierig – vor allem deren Ansicht in Echtzeit.

⇒ KomVISH ist als vereinfachte Version von HydroVISH die erste Softwarelösung zur Nutzung dieser 3D-Geodaten im kommunalen Umfeld mit der gleichzeitigen Kopplung an verschiedene GIS-Plattformen über entsprechende Schnittstellen (ESRI ArcGIS Pro & ArcMap sowie Kominfo, Abb. 4).

Dabei werden die 3D-Geodaten in Echtzeit parallel zu den im 2D-GIS vorliegenden Projekten (z.B. DFK, Orthobilder, Fachdaten) angezeigt und es lassen sich einfach räumliche Bezüge herstellen: Distanzen zwischen Gebäuden, Abmessungen von Baumkronen oder Gebäuden, der Abstand zu Freileitungen, Höhenunterschiede zwischen Dächern, Lichtraumprofile von Straßen oder Brücken sowie einfachen interaktiven geometrischen Schnittführungen.

- Die einzelnen 3D-Daten können separat oder zusammen visualisiert und analysiert werden (Abb. 4 rechts).
- 3D-Profillinien oder Höhenmessungen können von KomVISH aus in das gekoppelte GIS überführt werden (Abb. 4).
- Zusätzlich besteht eine direkte Verlinkung zu WMS-Diensten für eine verbesserte Orientierung (BayernAtlas 2D & 3D, OpenStreetMap), die durch die Einbindung des HERE-Kartendienstes mit der Möglichkeit der gezielten Suche nach Adressen weiter ergänzt wird.
- Über eine integrierte Kommentarfunktion kann der Informationsfluss sowohl zwischen einzelnen Datennutzern und Bearbeitern in den Kommunen und Behörden als auch zwischen der 3D- und 2D-Anwendung unterstützt werden.

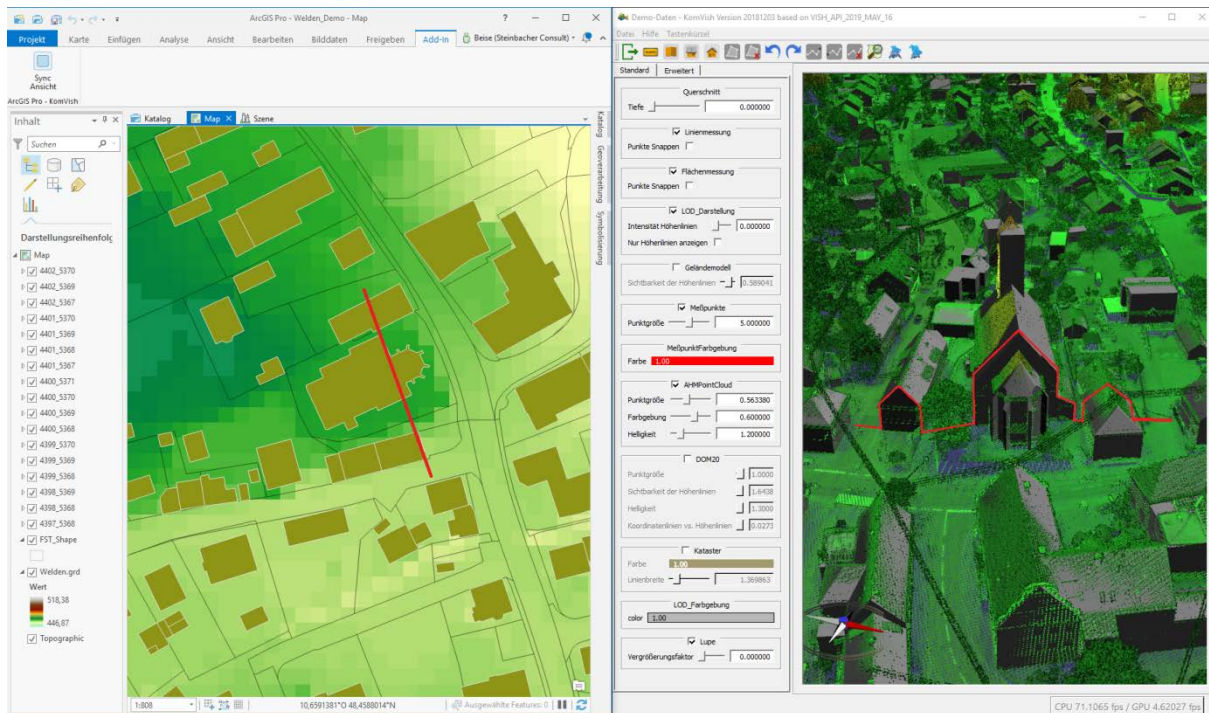


Abb. 4: Lage einer Profillinie (rot) in 2D (ArcGIS Pro links) und deren 3D-Geometrie (KomVISH rechts).

3 KomVISH & HydroVISH on the Web: Prozessierung und interaktive Datenverarbeitung aus dem Homeoffice oder im Außendienst

Als reine Desktopanwendungen erlauben HydroVISH und KomVISH den Zugriff und die Analyse von lokal gespeicherten oder auf einem Server im internen Netzwerk abgelegten Datensätzen. Damit verbunden sind entsprechende technische Anforderungen an Speicherplatz, RAM oder Grafikkarte wie auch bei anderen Desktopanwendungen, um eine effektive Nutzung der Daten am Arbeitsplatz zu gewährleisten. Im Pandemiejahr 2020 verstärkt sich nun unter gesundheitlichen und wirtschaftlichen Gesichtspunkten der Trend in Behörden und Verwaltungen aber auch in Unternehmen zu mehr Homeoffice mit Vernetzung von Endgeräten und digitalen Workflows. Die bereits zuvor steigende Tendenz zu Web-Anwendungen statt Desktop-Applikationen erleichtert bzw. beschleunigt den Umstieg von der Büroarbeit zum Home-

office, ist aber von der Verfügbarkeit entsprechender Web-Technologien abhängig. Letztere müssen sich nach dem kleinsten gemeinsamen Nenner der Fähigkeiten der Web-Browser richten, hinken also zwangsläufig dem technologisch Möglichen und der tatsächlichen Leistungsfähigkeit der eingesetzten Hardware hinterher.

Zwar gibt es gerade im GIS-Bereich zahlreiche Weblösungen, diese sind jedoch teilweise beschränkt auf zweidimensionale Kartendarstellungen, wie bei WMS-Diensten. Oder sie unterliegen massiven Datenreduktionen bei der Visualisierung von 3D-Daten. Bei der Prozessierung und interaktiven Datenverarbeitung im Homeoffice sind derartige Limitierungen schnell erreicht. Vor allen Dingen, wenn unter der Prämisse der höchstmöglichen Genauigkeit für bestmögliche Ergebnisse gearbeitet werden muss. Genau dies ist aber der Normalfall bei der Verarbeitung von Laserpunktwolken – wie beispielsweise die erwähnten Laserdaten von ganz Bayern mit 10 Terabyte. Der Download eines derart großen Datensatzes, unabhängig davon, wie gut die Datenverbindung auch sein mag, ist nicht praktikabel.

Im letzten Abschnitt 3.1 stellen wir daher zwei Möglichkeiten vor, die den Fernzugriff via Web-Browser auf die beiden Desktop-Anwendungen KomVISH (Abb. 5) und HydroVISH ermöglichen und vergleichen diese mit der eigentlichen Desktop- sowie einer WebGL-Anwendung. Damit unterstützen wir sowohl die effiziente Arbeit im Homeoffice als auch die nötige Vernetzung von Endgeräten, die beispielsweise im Außendienst von Bedeutung ist (Abb. 6).

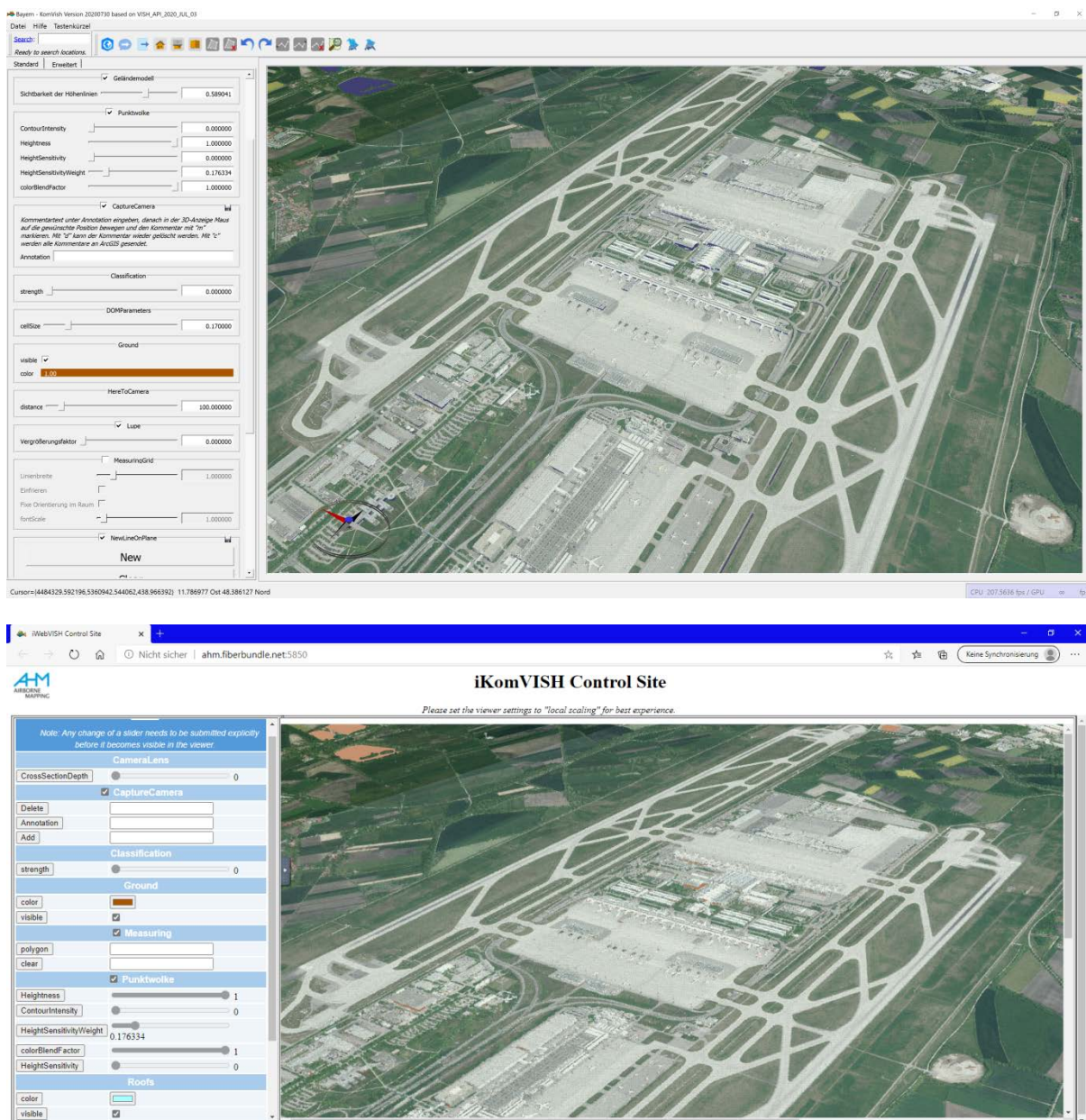


Abb. 5: KomVISH Desktop (oben) und via Webbrowser (unten).



Abb. 6:
Thinfinity Funktionalitäten und Vernetzung
von Anwendungen auf verbundenen Endgeräten.

3.1 Technologievergleich: Desktop- vs. Web-Anwendung und Hybridlösung

Die Anforderungen zur direkten Visualisierung hochaufgelöster Punktwolken sind deutlich größer als bei der flächenbasierten Darstellung von Geometrien oder der 2D-Darstellung von Karten. Tabelle 1 zeigt den Vergleich möglicher Technologien in Bezug zur *Maximalvariante* auf einem Desktoprechner. VNC (Virtual Network Computing) ist eine Lösung, die sich direkt in die Desktopsoftware integrieren lässt und somit den Fernzugriff auf eine bereits laufende Anwendung erlaubt. Meist wird VNC dann als fertige Applikation für den Fernzugriff auf den gesamten Desktop verwendet, was jedoch nicht unbedingt wünschenswert ist. Der Webanwender erwartet ja, dass mit dem Öffnen einer Webseite alles erledigt ist (*Webfeeling* in Tabelle 1). Eine Alternative zu VNC ist RDP (Remote Desktop Protocol), das bereits integrierter Teil des Windows-Betriebssystems ist. In dieser Grundausstattung bietet RDP ebenfalls nur vollen Desktop-Zugriff. Mit dem Thinfinity Remote Desktop Server von Cybele Software (<https://www.cybelesoft.com/thinfinity/remote-desktop/server/>) ist eine Beschränkung des Fernzugriffs auf nur bestimmte Anwendungen möglich. Auf dem Endgerät ist dafür keinerlei Programminstallation nötig, der Client läuft ausschließlich in einem Webbrowser basierend auf HTML5-Technologie. Zudem bietet Thinfinity RDP die Möglichkeit, dass mehrere Anwender gemeinsam dieselben Darstellungen anschauen (Kollaboration), was mit RDP in seiner Grundvariante an sich nicht möglich ist. VNC bietet diese Möglichkeit auch, kommt jedoch an die Performanz des auf Windows optimierten RDP nicht heran. Demgegenüber stünde die Variante eines eigens entwickelten Webclients, der wie die Desktopanwendung einen leistungsstarken Rechner benötigt (*Fat Client* vs. *Thin Client*), aber technologisch durch die Limitierungen von WebGL (z.B. keine sog. Geometrieshader möglich und daher eingeschränkte Darstellungsmöglichkeiten) und der Bandbreitenanforderungen derzeit für unsere Anwendungsszenarien wenig interessant ist.

Tabelle 1: Technologievergleich Desktop- vs. Web-Anwendung und Hybridlösung.

	Desktop	Integrierter VNC	Thinfinity RDP	Webclient
Client	Fat	Thin	Thin	Fat
Darstellungsqualität	Hoch	Hoch	Hoch	Beschränkt
Performanz	Maximal	Mittel	Hoch	Hoch
Darstellung	GPU Hardware	Videostream	Videostream	WebGL
Interaktivität	Maximal	Gut	Sehr gut	Sehr gut
Handhabung von großen Datenmengen	Sehr gut	Sehr gut	Sehr gut	Mäßig
Webfeeling	Nein	Ja	Ja	Ja
Lokale Installation nötig	Ja	Nein	Nein	Nein
Mehrfachnutzer	Nein	Ja	Möglich	Ja
Kollaboration	Nein	Ja	Möglich	Nein
Entwicklungsaufwand	-	Mäßig	Gering	Hoch