TEXTE 41/2012



Umweltinformationssysteme

Frühwarn- und Informationssysteme für den Hochwasserschutz

Umwelt Bundes Amt
Für Mensch und Umwelt

| TEXTE | 41/2012

Umweltinformationssysteme

Frühwarn- und Informationssysteme für den Hochwasserschutz

Workshop des Arbeitskreises „Umweltinformationssysteme" der Fachgruppe „Informatik im Umweltschutz", veranstaltet durch das Sächsische Landesamt für Umwelt, Landwirtschaft und Geologie in Zusammenarbeit mit dem Umweltbundesamt am 03. und 04. Mai 2012

UMWELTBUNDESAMT

Diese Publikation ist ausschließlich als Download unter
http://www.uba.de/uba-info-medien/4344.html
verfügbar.

Die im Workshopbericht geäußerten Ansichten
und Meinungen müssen nicht mit denen des
Herausgebers übereinstimmen.

ISSN 1862-4804

Herausgeber: Umweltbundesamt
Wörlitzer Platz 1
06844 Dessau-Roßlau
Tel.: 0340/2103-0
Telefax: 0340/2103 2285
E-Mail: info@umweltbundesamt.de
Internet: http://www.umweltbundesamt.de
http://fuer-mensch-und-umwelt.de/

Redaktion: Fachgebiet IV 2.1 Informationssysteme Chemikaliensicherheit
Dr. Gerlinde Knetsch

Sekretariat IV 2 Arzneimittel, Chemikalien und Stoffuntersuchung
Ute Zacharias

Dessau-Roßlau, August 2012

Vorwort

Den 24. Workshop des Arbeitskreises „Umweltinformationssysteme“ der Fachgruppe "Informatik im Umweltschutz" organisierte das Sächsische Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG). Er fand am 3. und 4. Mai 2012 in Dresden statt; mit mehr als 50 Teilnehmern und Teilnehmerinnen aus den Bereichen Wissenschaft, Verwaltung und Softwareentwicklung zeigte sich das große Interesse an der gewählten Thematik.

10 Jahre nach dem verheerenden Elbehochwasser 2002 wählte das Programmkomitee den Schwerpunkt „Hochwasserschutz“ aus. Die Stadt Dresden mit ihren Kulturgütern war besonders vom Hochwasser betroffen. Somit kam die Motivation für die Ausgestaltung eines Workshops mit dieser Fragestellung nicht von ungefähr. Die Umsetzung der Richtlinie 2007/60/EG des Europäischen Parlaments und des Rates vom 23. Oktober 2007 über die Bewertung und das Management von Hochwasserrisiken erfordert vielfältige Aktivitäten zur Bewertung von Hochwasserrisiken in gefährdeten Flusseinzugsgebieten, der Anfertigung von Hochwasserrisikokarten für alle Gebiete, in denen ein signifikantes Hochwasserrisiko besteht sowie zur Erstellung von Plänen für das Hochwasserrisikomanagement.

In einem Einführungsvortrag stellte das Sächsische Landesamt für Umwelt, Landwirtschaft und Geologie (Dr. Uwe Müller) verschiedene Umweltinformationssysteme für Sachsen vor. Speziell für den Bereich Wasser konnte eindrucksvoll die Relevanz des Hochwasserschutzes in Sachsen und eines Frühwarnsystems für Hochwasserereignisse im Landeshochwasserzentrum dargestellt werden.

Die Technische Universität Dresden präsentierte ein Modell zur Simulation des Wasserhaushalts in Erddämmen und Deichen als Basis für Stabilitätsuntersuchungen bei Hochwässern. Wesentliches Element dieses Modells ist die Verifikation der modellierten Ergebnisse mit gemessenen Daten bei Hochwassersituationen. Mit dem Beitrag zum Fachinformationssystem „Wasserrechtlicher Vollzug Sachsen“ konnte anschaulich gezeigt werden, wie Datenanforderungen aus der INSPIRE-Richtlinie in ein Fachverfahren „Digitales Wasserbuch“ integriert werden.

Die Realisierung des Hochwasser- und Sturmflutinformationssystems Schleswig-Holstein (HSI-SH) beleuchtete die Herausforderung eines Systems, in extremen Hochwassersituationen stark ansteigende Nutzerzahlen zu bewältigen und trotzdem Wasserstandsdaten und hochwasserrelevante aktuelle und zuverlässige Informationen über ein Portal für die Öffentlichkeit bereitzustellen.

Eine aus einem anderen Blickwinkel betrachtete Technologie zur Verschmelzung von realer Welt mit digitalen Informationen zu einer „Erweiterten Realität“ für Jedermann demonstrierte eine auf den mobilen Endnutzer ausgerichtete Smartphone - Entwicklung.

Die per Kamera des Smartphones aufgenommene Realität wird dabei mit hydrologischen Daten und Grafiken angereichert. Der Nutzer kann somit auf dem Smartphone Informationen für sich anbahnende Hochwasserereignisse abrufen. Der Begriff „Erweiterte Realität (Augmented Reality, kurz AR)" bezeichnet die visuelle Ergänzung der optischen menschlichen Wahrnehmung der Realität mit digitalen, kontextabhängigen Informationen in Echtzeit.

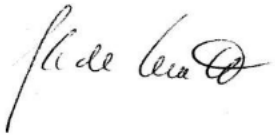
Das Thema Umwelt und Gesundheit stand – wie schon vor zwei Jahren – im Blickfeld von mehreren Vorträgen. Die GIS-gestützte Aufbereitung von Informationen zu gesundheitsrelevanten Umweltbelastungen im Land Berlin, die Untersuchung von Antibiotika und Antibiotikaresistenz im urbanen Abwasser von Dresden sowie die durch Verkehrslärm verursachten Umwelt- und Gesundheitsbelastungen in Berlin zeigten die engen Zusammenhänge von räumlichen Informationen und gesundheitlichen Risiken.

Weitere Beiträge des Workshops umfassten Methoden und Werkzeuge zum Management und der Auswertung von Daten zu persistenten organischen Stoffen (POPs) in der Umwelt, die von Bund und den Ländern in einem gemeinsamen Datenbanksystem zusammengeführt werden http://pop-dioxindb.de

sowie eines in der Umsetzung befindlichen Metadatenportals für Bodenmessdaten. Die Koordinierungsstelle des Informationsportals zu Umweltinformationen der Umweltverwaltungen von Bund und Ländern in Deutschland stellte ein Konzept des „PortalU" vor, welches Möglichkeiten der Darstellung und Abfrage von Umweltinformationen im Zusammenhang mit der Open Government und Open Data Strategie aufzeigte. Ein Beitrag zu neuen Wegen der Umweltberichtserstattung für das Pollutant Release and Transfer Register (PRTR) wies auf die Vorteile einer standardisierten XML-Schnittstelle für betriebliche Stamm- und Berichtsdaten hin https://www.xrepository.deutschland-online.de/

Der Workshop profitierte von der Aktualität des Themas Hochwasserschutz und zeigte, welche Bedeutung die Verknüpfung verschiedener Fachdisziplinen hat.

Der nächste Workshop des Arbeitskreises Umweltinformationssysteme findet in Berlin an der Hochschule für Technik und Wirtschaft (HTW) vom 23. bis 24. Mai 2013 statt. Das Hauptthema "Smarte Trends 4 U(I)S" lenkt den Blick auf Entwicklungen und Trends für Apps auf Smartphones und Tablets.



Dr. Gerlinde Knetsch
Umweltbundesamt, Fachgebiet Informationssysteme Chemikaliensicherheit
August 2012

Inhaltsverzeichnis

Simulation des Wasserhaushalts in Erddämmen und Deichen als Basis für Stabilitätsuntersuchungen mit dem Programm PCSiWaPro® 1

Issa Hasan, Martin Meyer, Jinxing Guo, Peter-Wolfgang Gräber

Hochwasser- und Sturmflutinformationssystem Schleswig-Holstein (HSI-SH) 9

Friedhelm Hosenfeld, Ralf Hach

Augmented Maps und Augmented Reality für den präventiven Hochwasserschutz 19

Frank Fuchs-Kittowski, Stefan Simroth, Bernd Pfützner, Frank Trosien

GIS-Komponenten zur Untersuchung von Antibiotika und Antibiotika-resistenz im urbanen Abwasser 33

Luise Hutka, Susanne Reinwarth, Daniel Kadner, Lars Bernard

Die GSBL-Kommunikationsplattform – ein gemeinsames Portal für Bund und Länder 41

Barbara Liebscher

Umweltgerechtigkeit als neue ressortübergreifende sozialräumliche Strategie im Land Berlin 51

Heinz-Josef Klimeczek

Verteilungswirkung von Verkehrslärm in Berlin – Methoden und Ergebnisse 57

Thilo Becker, Julia Friedemann

Aufbau und Verwendung von Fachterminologie 77

Thomas Bandholtz, Joachim Fock

Open Data Governance 91

Arne Leißner, Jo van Nouhuys

Web-Recherche und statistische Auswertung für Daten der Dioxindaten-bank des Bundes der Länder 107

Gerlinde, Knetsch, Philipp Gärtner, René Lehmann, Jo van Nouhuys, Gregor Raab

Einrichtung und Betrieb eines Metadatenportals für Bodenmessdaten 123

Jeannette Mathews, Thomas Bandholtz

Neue Wege in der Umweltberichterstattung – Das IT-Invest Projekt XUBetrieb des Umweltbundesamtes 135

Falk Hilliges

Bausteine zu einer Technologie für Participatory Sensing Anwendungen 141

Andreas Abecker, Valentin Zacharias, Julio de Melo Borges

PortalU und Open Data: Möglichkeiten der Darstellung und Abfrage von Umweltinformationen 157

Stefanie Haß, Fred Kruse

Digitale Biodiversitätsatlanten – Tools zur Unterstützung regionaler und internationaler Expertennetzwerke für die Erfassung und Bewertung biologischer Diversität 161

Heiko Brunken, Carl-Heinz Genzel, Heide-Rose Vatterrott, Martin Winkler

Simulation des Wasserhaushalts in Erddämmen und Deichen als Basis für Stabilitätsuntersuchungen mit dem Programm PCSiWaPro®

Issa Hasan, Martin Meyer, Jinxing Guo, Peter-Wolfgang Gräber
Technische Universität Dresden/ Institut für Abfallwirtschaft und Altlasten
issa.hasan@tu-dresden.de; martin.meyer@tu-dresden.de;
jinxing.guo@tu-dresden.de; peter-wolfgang.graeber@tu-dresden.de

Abstract

Construction, building materials, vegetation, atmospheric conditions and hydrogeological properties and especially the water balance of dams and dikes are the crucial factors for their stability in case of floods. It was found by means of physical model tests that slides could already occur under unsaturated conditions on the air side of dams, and thus the stability could be threatened. Using the program PCSiWaPro® one can simulate and describe the hydrogeological processes in the saturated and partially saturated zone under consideration of hydraulic and meteorological conditions. Several scenarios were simulated for this purpose using the program PCSiWaPro® and the simulation results were evaluated. The simulation results show clearly the impact of building materials and construction on the behavior of water saturation in dams' embankment and thus on the stability of the dams. The agreement between measured values and the computed ones using PCSiWaPro® is very good.

Kurzfassung

Konstruktion, Baumaterialien, Vegetation, atmosphärische Bedingungen und geo-hydraulische Eigenschaften und insbesondere der Wasserhaushalt von Dämmen und Deichen sind die entscheidenden Faktoren für deren Stabilität im Falle eines Hochwassereinstaus. Anhand von physikalischen Modellversuchen konnte gezeigt

werden, dass es bereits bei teilgesättigten Verhältnissen zu Rutschungen auf der Luftseite kommen kann und dadurch die Stabilität bedroht ist. Mit Hilfe des Simulationsprogramms PCSiWaPro® einschließlich der hydraulischen und meteorologischen Bedingungen kann man die geohydraulischen Prozesse im gesättigten und teilgesättigten Bodenbereich simulieren und beschreiben. Mehrere Szenarien wurden für diesen Zweck mit dem Programm PCSiWaPro® simuliert und ausgewertet. Die Simulationsergebnisse zeigen eindeutig die Auswirkung der Baumaterialien und der Konstruktion auf das Verhalten der Wassersättigung und dadurch auf die Standsicherheit des Dammes. Die Übereinstimmung zwischen gemessenen Werten und mittels des Simulationsprogramms PCSiWaPro® berechneten Werten ist als sehr gut zu bewerten.

1 Einleitung

Dämme und Deiche sind als Schutzbauwerke zum Zweck des Hochwasserschutzes zu bezeichnen. Das Programm PCSiWaPro® (entwickelt an der TU Dresden, Institut für Abfallwirtschaft und Altlasten) berechnet die instationäre Wassersättigungsverteilung anhand der 2D-Richards Gleichung unter Anwendung der Finite-Elemente-Methode. Die Anbindung eines Wettergenerators an PCSiWaPro® ermöglicht die instationäre Strömungsberechnung unter Beachtung von atmosphärischen Randbedingungen (Niederschlag, Verdunstung, Tagesmitteltemperatur und Sonnenscheindauer) und Wasserentzug durch Pflanzenwurzeln. Klimadaten können durch diese Anbindung mit Hilfe öffentlich zugänglicher Klimastationen des Webinformationssystems des DWD importiert werden. Zur Ermittlung der Auswirkungen von den oben genannten Faktoren auf die Durchströmung und die Rutschung des untersuchten Erddamms wird der teilgesättigte Bereich oberhalb der Sickerlinie, der Grenze zwischen vollgesättigter und teilgesättigter Zone im Dammkörper (siehe Abb.1), betrachtet. Mittels PCSiWaPro® wurde, nach den Angaben von Geometrie, Bodenparametern, Klimaparametern und geohydraulischen sowie zeitabhängigen Randbedingungen, ein 2D-Modell des Damms aufgebaut. Die Wassersättigung im Dammkörper sowie der Durchfluss aus der Luftseite des Damms wurden mit PCSiWaPro® berechnet. Hierfür wurde ein Beobachtungspunkt im Modell eingesetzt, um die gemessenen Werte mit den berechneten zu vergleichen.

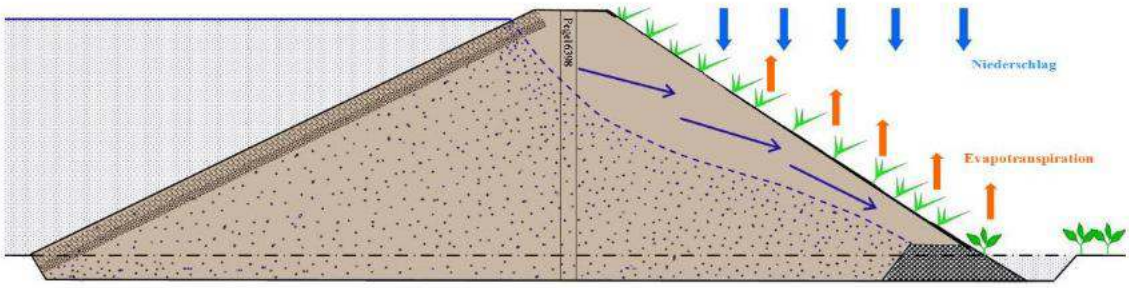


Abbildung 1. Wasserhaushalt im gesättigten und teilgesättigten Bereich

Die Konstruktion von Dämmen und Deichen ist für die Sickerlinie bei Hochwasser-Einstau sehr entscheidend. Am Institut für Wasserbau und Technische Hydromechanik der TU Dresden (Aigner, 2004) wurden mehrere physikalische Versuche durchgeführt, welche sich jeweils durch unterschiedliche Deichkonstruktionen auszeichnen (Abb. 2). Anhand dieser Versuche konnten die Rutschungserscheinungen auch für den teilgesättigten Bereich oberhalb der Sickerlinie (dunkel eingefärbt) nachgewiesen werden.

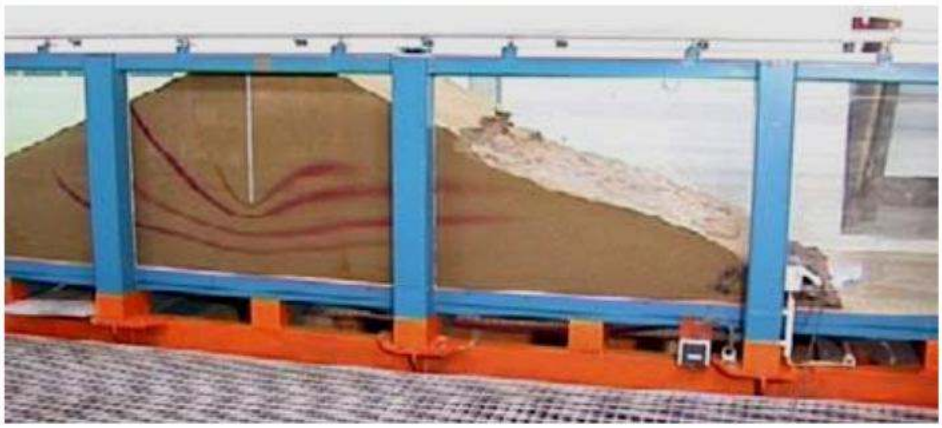


Abbildung 2. Ein physikalisches Damm-Model mit Rutschung auf der Luftseite (Aigner, 2004)

Mit dem Programm PCSiWaPro® wurden alle Versuche modelliert und mit den Aussagen aus (Aigner, 2004) verglichen. Dieser Modelldeich verfügt über eine Fußbreite von 3,38 m, eine Böschungsneigung von 1:2 bei einer Höhe von 77 cm. In die Deichkrone ist eine Gummiwand eingelassen, welche 100% der Deichhöhe einnimmt (Einbindetiefe). Unterhalb des Deiches wurde sowohl luft- als auch wasserseitig ein Untersatz (Berme) angefügt (Höhe 30 cm), um den Austausch zwischen einströmenden Wasser und Grundwasser zu ermöglichen. An der luftseitigen Böschung befindet sich eine Dränage, um überschüssiges Wasser wirkungsvoll abführen zu können.

2 Modellaufbau mit dem Simulationsprogramm PCSiWaPro®

2.1 Materialparameter

Die Parameter einzelner Materialien (Dammkörper, maschinell verdichteter Bereich unterhalb des Damms, unverdichteter Bereich außerhalb des Dammbereiches) müssen bei der Modellierung in das Programm PCSiWaPro® eingegeben werden. Diese können unter anderem aus vorhandenen GeODin-Datenbanken importiert werden. Die im Modell verwendeten Materialparameter sind in Abb. 3 zu sehen. Die bodenhydraulischen Parameter sowie die Van-Genuchten-Parameter (vGP) sind beim Modellaufbau erforderlich und wurden mit Hilfe diverser Informationsquellen (Kemmesies, 1995; DIN4220) abgeschätzt, da keine Messwerte zur Verfügung standen. Daher wurden mehrere Variantenrechnungen durchgeführt, um deren Sensitivität zu bestimmen. PCSiWaPro® bietet weitere Möglichkeiten für die Abschätzung der Materialparameter mit Hilfe der gemessenen Daten an, wie Pedotransferfunktionen (Vereecken, et al., 1989) und Parameteridentifikation.

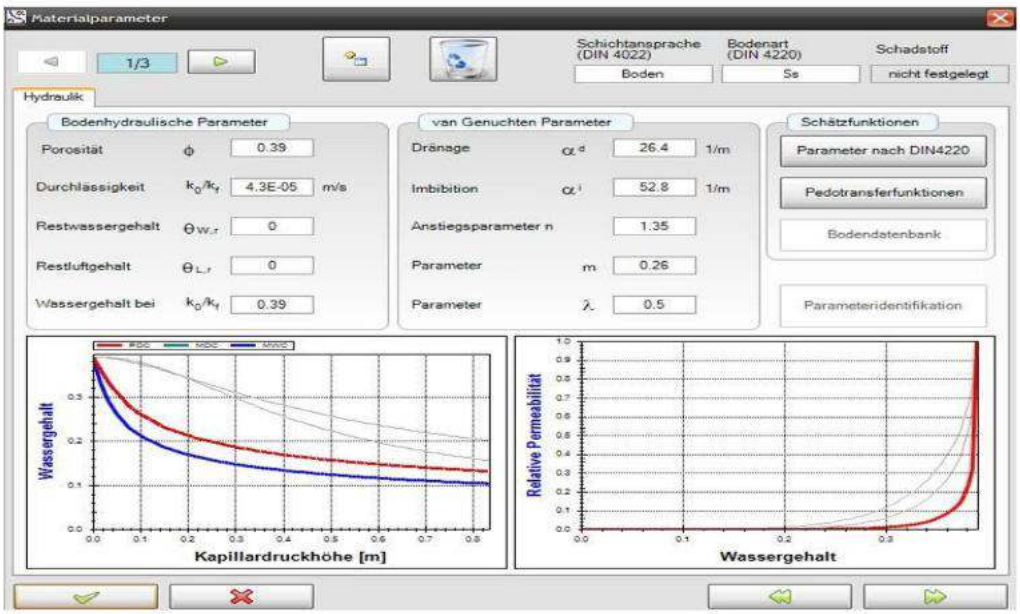


Abbildung 3. Materialparameter für das untersuchte Modell

2.2 Randbedingungen des Modells

Zur Modellierung benötigt man die Festlegung entsprechender Randbedingungen sowie Anfangsbedingungen des untersuchten Gebiets. Im Programm PCSiWaPro® ist als Anfangsbedingung für die Strömungsberechnung die Druckhöhenverteilung oder Wassersättigungsverteilung im gesamten Modellgebiet anzugeben. Die zeitabhängigen Randbedingungen im Modell werden durch die gemessenen Wasserstände auf der Wasserseite des Dammes über den entsprechenden Zeitraum definiert. Zudem können atmosphärische Randbedingungen, wie Niederschlag, Evaporation und Verdunstung, eingegeben werden. Diese werden nur für die nummerische Simulation des physikalischen Modells berücksichtigt. An der Luftseite des Damms wurde als Sickerfläche, eine Randbedingung 2. Art die bedeutet, dass solange der Bereich ungesättigt ist, ist kein Volumenstrom möglich, festgelegt.

2.3 Netzdaten

Für die Berechnung im Programm PCSiWaPro® ist die Diskretisierung des Modellgebiets erforderlich. Dabei können unstrukturierte dreieckige Netzelemente vorgegeben werden. Damit ist es möglich, auch unregelmäßige Modellgebiete adäquat abzubilden. Der im Programm implementierte Netzgenerator arbeitet nach dem Prinzip der „Boundary Representation Modelling Technique“ und erfordert somit die Vorgabe der Modellränder. Da insbesondere der Bereich des Dammkörpers für die Auswertung von Interesse ist, wurde dieser Modellabschnitt mit einer feineren Diskretisierungsschrittweite versehen. Sie beträgt hier 10 cm, damit ist auch eine bessere Darstellung der Änderung des Wassergehalts im ungesättigten Bereich oberhalb der Sickerlinie und ein genauerer Vergleich der berechneten mit den gemessenen Werten möglich.

3 Simulationsergebnisse und Auswertung

Ein weiteres Simulationsbeispiel behandelt einen Erddamm in Brandenburg. Zur Validierung des Programms wurden die Messdaten, die Wasserstände in einem Pegel im Dammkörper, mit den berechneten Werten verglichen. In diesem Fall wurde davon ausgegangen, dass kein Niederschlagswasser in den Damm strömt, da der Damm mit einer undurchlässigen Schicht aus Wasserbausteinen, Schotter und Geokunststoffen abgedeckt ist. Der Vergleich ist in Abb. 4 dargestellt.

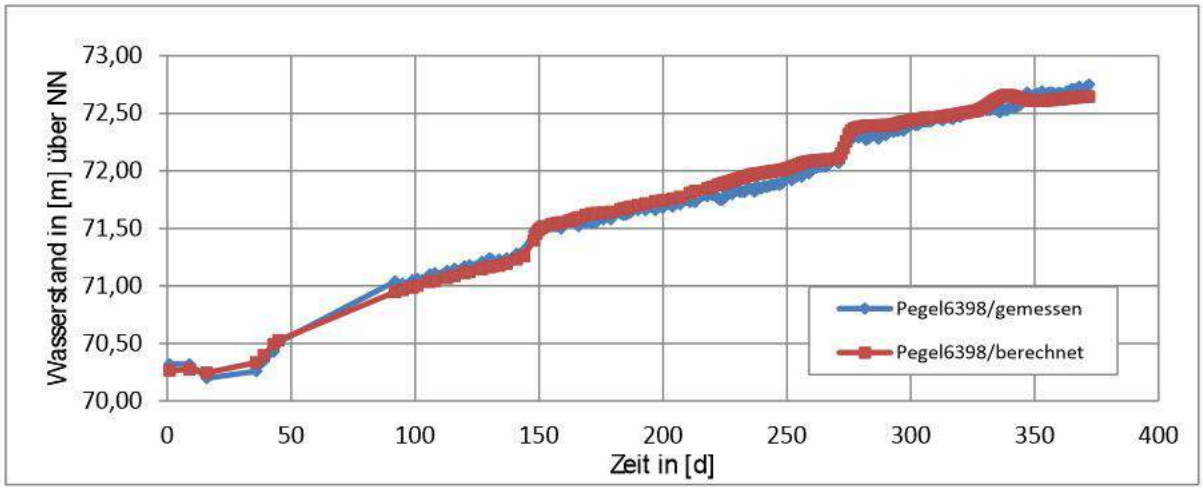


Abbildung 4. Vergleich der berechneten Werte mit den gemessenen Wasserständen im Dammkörper (Pegel 6398)

In Abb. 5 ist die Sensitivität des hydraulischen Materialparameters α dargestellt. Obwohl sich der Wasserstand innerhalb des Dammkörpers kaum ändert, zeichnet sich ein deutlicher Unterschied im teilgesättigten Bereich oberhalb der Sickerlinie ab. Dieser Bereich wird durch die höheren Kapillarkräfte bindigerer Materialien (größere kapillare Steighöhe des Wassers) aufgespreizt, was auf die luftseitige Standsicherheit des Damms einen negativen Einfluss haben kann.

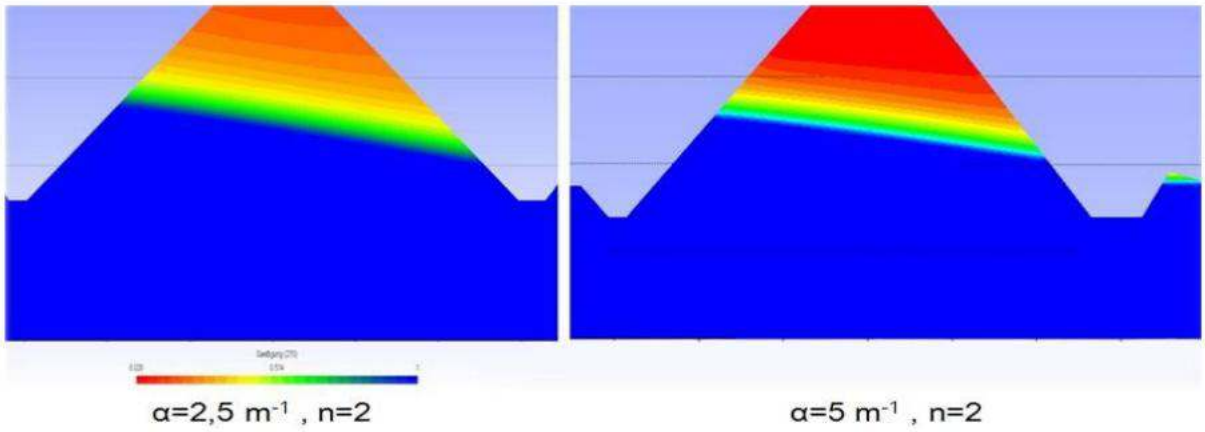


Abbildung 5. Auswirkung der Materialparameter (links liegt ein feineres Material vor) auf den Wassergehalt im Dammkörper

Das grafische Ergebnis der Simulation des zweiten Beispiels (physikalisches Damm-Modell, Abb. 2) ist in Abb. 6 dargestellt. Mit dem Simulationsprogramm PCSiWaPro® wurde eine typische Belastungssituation simuliert. Während der Wassergehalt im Bereich der Dränage abfällt, nimmt der Wassergehalt im ungesättigten Bereich des Dammkörpers zu und somit besteht eine Rutschungsgefahr auf der Luftseite. Durch

Niederschlag strömt zusätzliches Wasser in den Dammkörper und erhöht dadurch dessen Wassergehalt.

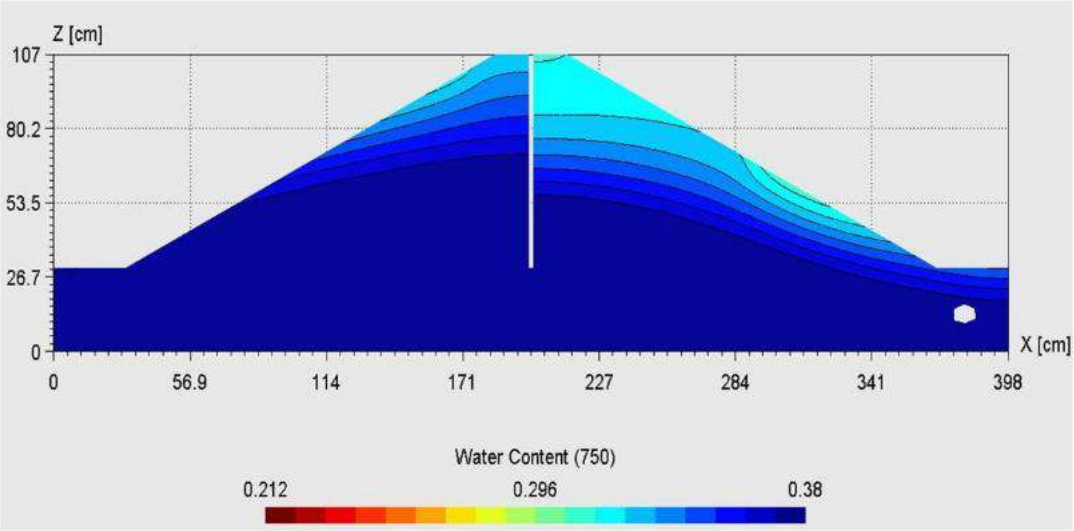


Abbildung 6. Verteilung des Wasserhaushalts im Dammkörper nach 750 Minuten unter Berücksichtigung der atmosphärischen Bedingungen

4 Schlussfolgerungen

Die Übereinstimmung zwischen den gemessenen Werten und den mittels des Simulationsprogramms PCSiWaPro® berechneten Werten ist als sehr gut zu bewerten. Die zu verzeichnenden Abweichungen sind vor allem auf Unsicherheiten der verwendeten Bodenparameter zurückzuführen.

Die Simulationsergebnisse zeigten bei dem physikalischen Dammmodell, dass bereits bei hoher Teilsättigung an der Luftseite Durchflussraten auftreten können und damit die Standfestigkeit des Deiches beeinträchtigt wird und es zu Rutschungen durch Aufschwimmen des Bodenmaterials (Auftrieb) und der damit verbundenen Verringerung des Auflagedrucks kommen kann.

Ein kontinuierliches Monitoring des Wassergehalts in Dämmen und Deichen ist für deren Stabilität vor allem bei Hochwasserereignissen ein guter Indikator.

5 Literaturverzeichnis

[Aigner, 2004]

Aigner, Detlef: Auswertung von Untersuchungen über den Einsatz einer Gummispundwand sowie einer Sickerleitung an einem durchströmten Modelldeich. Institut für Wasserbau und Technische Hydromechanik. Dresden : Technische Universität Dresden, 2004. Gutachten.

[Blankenburg, et al.]

Blankenburg, R., Nitsch, B. und Gräber, P.-W.: Simulation des Wasserhaushaltes in Deichen unter besonderer Berücksichtigung der Prozesse im teilgesättigten Bereich : Konferenzbeitrag zur ASIM STS-Tagung im März 2009 in Dresden.

[Geuzaine, et al. 2009]

Geuzaine, C. and Remacle, J.-F. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. International Journal for Numerical Methods in Engineering. Issue 11, 2009, Vol. 79, pp. 1309-1331.

[Gräber, et al., 2006]

Gräber, Peter-Wolfgang, et al.: SiWaPro DSS - Beratungssystem zur Simulation von Prozessen der unterirdischen Zonen. [Buchverf.] Jochen Wittmann und Mike Müller. [Hrsg.] Mike Müller. Simulation in Umwelt- und Geowissenschaften. Leipzig : Shaker Verlag, 2006, S. 225-234.

[Kemmesies, 1995]

Kemmesies, Oliver: Prozessmodellierung und Parameteridentifikation von Mehrphasenströmungsprozessen in porösen Medien. Dresden : Dresdner Grundwasserforschungszentrum e.V., 1995. ISSN 1430-0176.

[Vereecken, et al., 1989]

Vereecken, H.: Estimating the soil moisture retention characteristic from texture, bulk density and carbon content. Soil Science. 1989, Vol. 148, No. 6, pp. 389-403.

Hochwasser- und Sturmflutinformationssystem Schleswig-Holstein (HSI-SH)

Friedhelm Hosenfeld, hosenfeld@digsyland.de
Institut für Digitale Systemanalyse & Landschaftsdiagnose (DigSyLand)

Ralf Hach, ralf.hach@llur.landsh.de
Landesamt für Landwirtschaft, Umwelt und ländliche Räume
des Landes Schleswig-Holstein (LLUR-SH)

Abstract

The high-water and storm surge information system Schleswig-Holstein (HSI) is a portal site offering details about current water levels on digital maps as well as flood warnings and various background information about these topics including weather data, web cam positions, addresses and web links. The main challenge of the system design was to comply with different requirements: while during normal periods only an average number of users will visit the site, in extreme high-water situations the number of accesses will explode. Especially in these cases it is the system's essential task, to present highly up-to-date and reliable information. Thus, a solution was developed combining up-to-date data presentation with static HTML-pages reducing server load in extreme situations as far as possible in order to manage access for many concurrent users.

Kurzzusammenfassung

Das Hochwasser- und Sturmflutinformationssystem Schleswig-Holstein (HSI) ist ein Informationsportal, das unterschiedliche Angaben zu Wasserständen und Hochwasser in Schleswig-Holstein sowie Hintergrundinformationen zu diesen Themen präsentiert. Die Hauptherausforderung des Systems bestand darin, die in extremen Hochwassersituationen stark ansteigenden Nutzerzahlen zu bewältigen und trotzdem aktuelle und zuverlässige Informationen bereitzustellen. Die präsentierte Lösung besteht in der regelmäßigen Erzeugung statischer HTML-Seiten auf der Basis

aktueller Daten, deren Bereitstellung wenig Anforderungen an den Web-Server stellt, um hohe gleichzeitige Zugriffszahlen bewältigen zu können.

1 Einführung

Das Informationsportal *Hochwasser- und Sturmflutinformation* (HSI[1]) bietet sowohl aktuelle Wasserstandsangaben auf interaktiven digitalen Karten als auch Hochwasser- und Sturmfluthinweise in knapper Übersichtsform und weiterführende Hintergrundinformationen zu diesen Themenbereichen.

Verantwortlich für den Betrieb des Informationssystems sind das Landesamt für Landwirtschaft, Umwelt und ländliche Räume Schleswig-Holstein (LLUR), der Landesbetrieb für Küstenschutz, Nationalpark und Meeresschutz Schleswig-Holstein (LKN) sowie das Ministerium für Energiewende, Landwirtschaft, Umwelt und ländliche Räume Schleswig-Holstein (MELUR). Die Entwicklung und technische Realisierung wurde vom LLUR durchgeführt, seit 2006 unterstützt von der Firma DigSyLand.

1.1 Anforderungen

Ein wichtiges Kriterium beim Aufbau des Informationsangebots war die Berücksichtigung des Nutzungsverhaltens der Web-Site. Während zu normalen Zeiten mit durchschnittlichen Zugriffszahlen von Interessierten gerechnet werden kann, die sich über aktuelle Wasserstandsentwicklungen auf dem Laufenden halten möchten oder über andere Hochwasser-relevante Aspekte informieren möchten, steigen die Zugriffszahlen in Hochwasser-Situationen punktuell sprunghaft an und führen bei Extrem-Ereignissen zu Spitzenbelastungen des Informationsportals. Genau in diesen Situationen kommt dem System aber eine entscheidende Bedeutung zu, so dass die gesamte Software-Architektur so ausgelegt sein muss, auch bei außergewöhnlich hohen Zugriffszahlen zuverlässig zu funktionieren.

1.2 Rahmenbedingungen

Da ein dynamischer Aufbau von Web-Seiten bei hohen Zugriffszahlen den Web-Server und die Basis-Infrastruktur (z.B. Datenbanken) deutlich stärker belastet als die Bereitstellung statischer Web-Seiten, andererseits aber die Aktualität der

[1] HSI: http://www.hsi.schleswig-holstein.de/

angebotenenen Wasserstandsdaten essentiell ist, musste in der technischen Realisierung ein Kompromiss gefunden werden, der beide Rahmenbedingungen erfüllt:

Die Wasserstandsinformationen sowie alle daraus folgenden Informationsdarstellungen werden in bestimmten Zeitintervallen, die automatisch bei Hochwassersituationen verkürzt werden, offline auf einem Preprocessing-System im Landesamt für Landwirtschaft, Umwelt und ländliche Räume (LLUR) aufbereitet. Dabei wird der komplette Web-Auftritt in Form statischer Web-Seiten erzeugt und anschließend automatisiert auf den Web-Server übertragen (siehe Abb. 4).

Auf diese Weise wird einerseits die hohe Aktualität der Seiten sichergestellt, andererseits werden aber technische Vorteile genutzt, um die Web-Server-Last beim Abruf der Informationen gering zu halten.

2 Informationsbereitstellung

Das Portal *Hochwasser- und Sturmflutinformation* bietet auf seiner Startseite eine Übersichtskarte mit einer Darstellung der Wasserstandspegel in Schleswig-Holstein sowie eine farbige Symbolisierung der zuletzt ermittelten Wasserstände, so dass auf einen Blick erkannt wird, welche Wasserstände im Land vorliegen (siehe Abb. 1). Sowohl die Uhrzeit der letzten Aktualisierung wie auch etwaige Messausfälle sind auf einen Blick sichtbar. Über eine Mouse-Over-Funktionalität sind die Zahlenwerte des gemessenen Wasserstands sowie der Durchfluss und weitere Angaben zum Pegel abrufbar. Eine Legende erleichtert die Einschätzung der dargestellten Informationen.

Auf Wunsch können detailliertere Karten der drei Teilgebiete Schleswig-Holsteins, der sogenannten Flussgebietseinheiten (FGE) abgerufen werden.

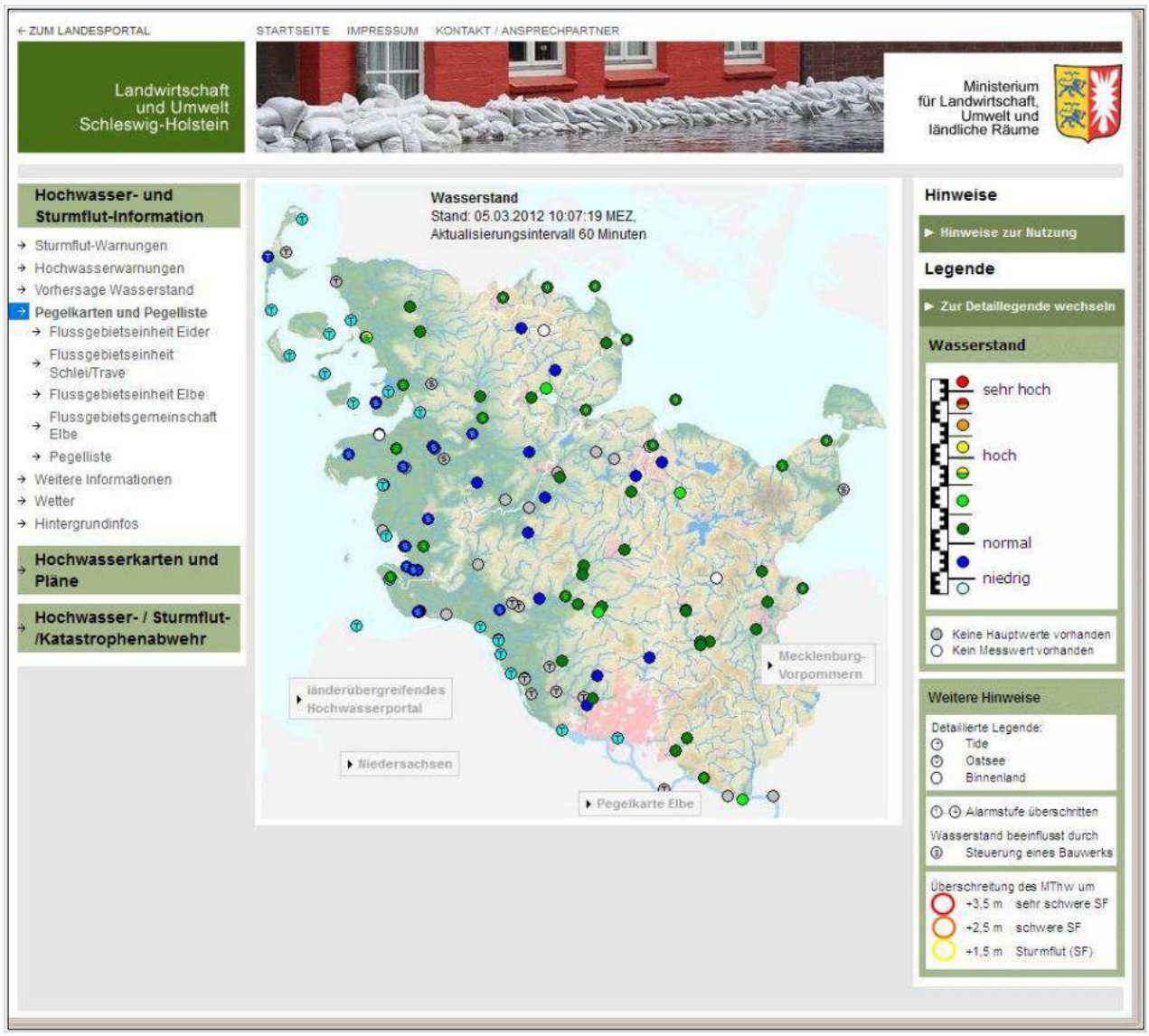


Abbildung 1: Startseite der Hochwasser- und Sturmflutinformation Schleswig-Holstein

Zu den weiteren Informationen, die über das Portal abgerufen werden, zählen

Aktuelle Sturmflut- und Hochwasserwarnungen

Wasserstandsvorhersagen (siehe Abbildung 2)

Listen der aktuellen Pegelmesswerte

Dynamische grafische Darstellungen der Wasserstandsganglinien sowie Hintergrundinformationen zu den Wasserstandspegeln. Diese Angaben werden nicht vor-prozessiert, sondern entweder dynamisch zur Aufrufzeit erzeugt oder bei

bestimmten Pegeln von der Wasser- und Schifffahrtsverwaltung in Pegel-Online[2] zur Verfügung gestellt.

Ergänzende Informationen zu Web-Cams, Karten und Listen mit Niederschlagsstationen und Niederschlagsdaten, Links zu wasserwirtschaftlichen Informationen aus Schleswig-Holstein

Hochwasserkarten und Pläne mit Karten und Landesverordnungen der Überschwemmungsgebiete

Links und Karten zur Hochwasser-, Sturmflut- und Katastrophenabwehr in Schleswig-Holstein

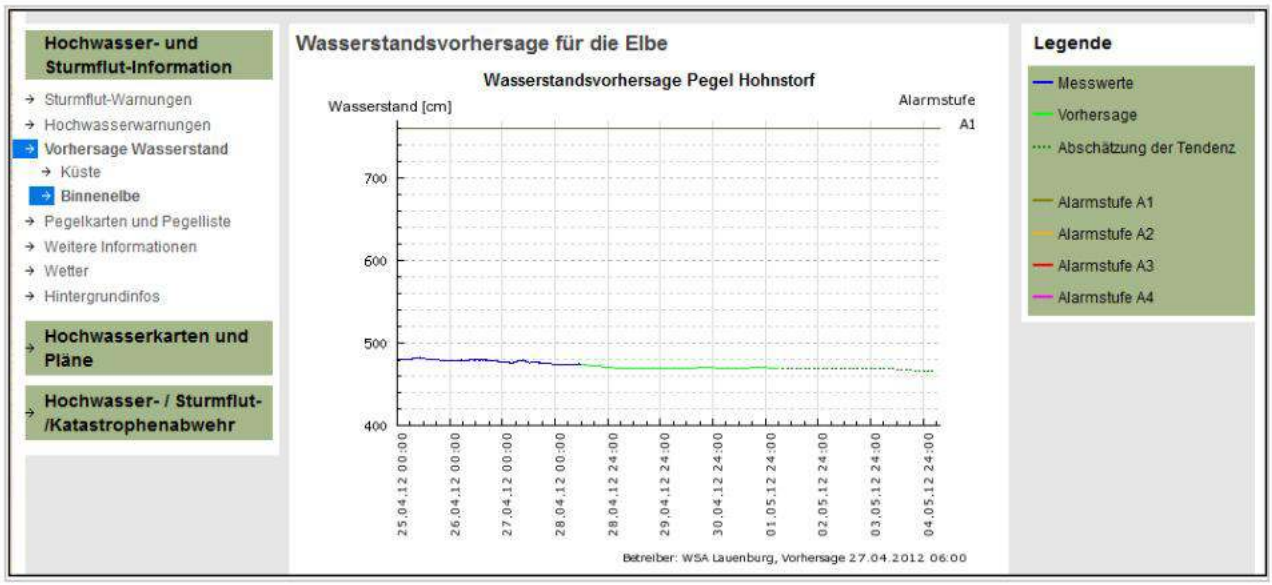


Abbildung 2: Vorhersagegrafik für den Wasserstand der Elbe

3 Technische Realisierung

Die aktuellen Messwerte der Pegel werden mittels Datenfernübertragung in die zentrale Datenbank im LKN übertragen. Dort werden alle Werte im Wasserwirtschaftlichen Informationssystem KISTERS (WISKI[3]) verwaltet.

Der komplette Aufbereitungsprozess zur Erstellung der Web-Seiten wird durch eine Cron-Zeitsteuerungssoftware ausgelöst. Während des Aufbereitungsprozesses wird der aktuelle Hochwasser-Zustand analysiert. In Abhängigkeit davon wird die

[2] Pegel-Online: http://www.pegelonline.wsv.de
[3] WISKI der Firma Kisters AG:
http://www.kisters.de/german/html/homepage.html&main=8D9B0E617478BDADC1257464003D3EE5

Zeitsteuerungstabelle (crontab) so modifiziert, dass das Wiederholungsintervall verkürzt (z.B. auf 15 Minuten) oder verlängert wird.



Abbildung 3: Steuerungsanwendung zum Einstellen von Warnungen

Auf dem Preprocessing-System im LLUR werden für den Aufbau der Web-Seiten skriptgesteuert alle relevanten Pegel-Informationen einschließlich ihrer Koordinaten aus WISKI extrahiert. Die Wasserstände werden bereits bei der Extraktion gemäß ihrer späteren farbigen Symboldarstellung klassifiziert.

Für Daten einiger Pegel, die nicht innerhalb des LKN verwaltet werden, wird auf den Webservice von Pegel-Online zugegriffen, um aktuelle Werte abzurufen.

Sobald alle Daten vorliegen, laufen Aufbereitungsskripte ab, die in der Skriptsprache Perl[4] erstellt wurden.

Anhand von Konfigurationsdateien, in denen die verschiedenen Kartendarstellungen beschrieben sind (z.B. Abmaße, räumlicher Ausschnitt, darzustellende

[4] Programmiersprache Perl: http://www.perl.org/

Pegel), werden die interaktiven Karten als *clickable Maps* erzeugt, die die Pegelsymbole sowie alle Mouse-Over-Informationen enthalten.

Daten, die mit einer HSI-Steuerungs-Anwendung (siehe Abb. 3) im LKN gepflegt werden, werden ausgelesen und analysiert. Mit dieser Anwendung kann der automatisch ermittelte Hochwasser-Zustand manuell übersteuert werden. Zudem können Alarmstufen für einzelne Landkreise aktiviert und Warnhinweise für die Darstellung im HSI-Web-Auftritt erstellt werden. Kreise, in denen Voralarm oder Alarm ausgelöst wurde, werden auf der Schleswig-Holstein-Karte farbig hervorgehoben, so dass der Status sofort erkennbar ist.

Auf der Basis der automatisch abgerufenen Pegel-Informationen und den manuell im LKN eingestellten Angaben wird der aktuelle Hochwasser-Status ermittelt. Als Konsequenz daraus wird das Aktualisierungsintervall für die Bereitstellung der Web-Seiten eingestellt. Zudem werden Warn-Hinweise generiert und in die Web-Seiten integriert.

Für die Hochwasservorhersage der Elbe werden Grafiken mit gemessenen und prognostizierten Daten erzeugt. Dafür werden aktuelle Daten mit Vorhersagedaten kombiniert, die im hydrologischen Normalfall von dem Wasser- und Schifffahrtsamt Magdeburg, im Hochwasserfall von der Hochwasservorhersagezentrale des Landesbetriebes für Hochwasserschutz und Wasserwirtschaft des Landes Sachsen-Anhalt mit dem Wasserstandsvorhersage Modell WAVOS Elbe[5] erzeugt und per automatisiert analysierbarer Email zugeschickt werden.

Wenn alle Web-Seiten-Teile erzeugt wurden, findet eine Kombination mit einem Rahmenlayout und weiteren statischen Web-Elementen statt, um den kompletten Web-Auftritt fertigzustellen.

[5] WAVOS Elbe: http://www.umweltdaten.landsh.de/public/hsi/WAVOS_Elbe_kolloquium_nov05.pdf

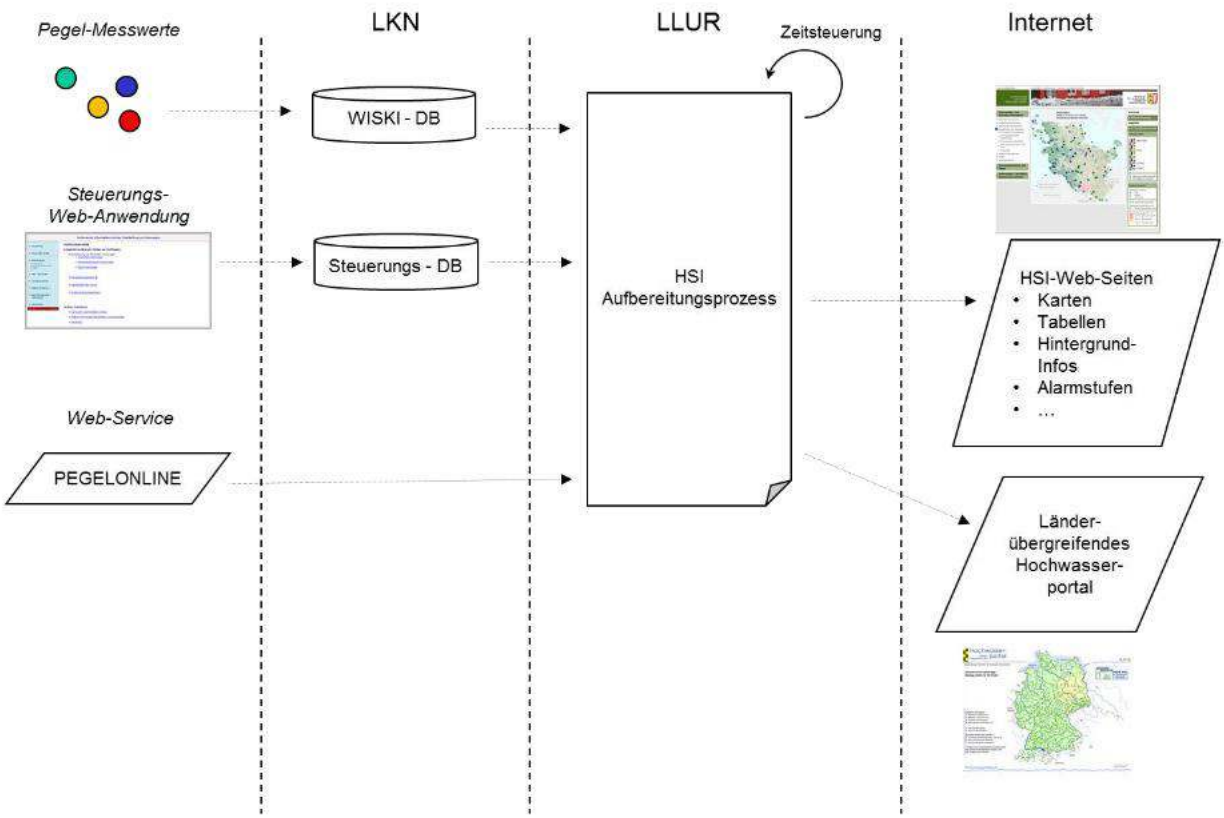


Abbildung 4: Schematische Darstellung des technischen Ablaufs

Zusätzlich werden einige XML-Dateien generiert, die vom länderübergreifenden Hochwasserportal[6] eingelesen werden. Dieses Portal zeigt auf einer Karten-darstellung die Hochwassersituation aller Bundesländer an und wertet dabei die per XML bereitgestellten Pegelinformationen sowie weitere Textbausteine mit verbalen Beschreibungen der Situation aus. Für Detailinformationen wird wiederum auf die ländereigenen Portale verlinkt.

4 Zusammenfassung und Ausblick

Das verwendete Verfahren hat sich seit seiner Inbetriebnahme aufgrund der speziellen technischen Realisierung bewährt:

Alle relevanten dynamischen Elemente und Daten werden automatisiert aus den entsprechenden Datenbanken bzw. Web-Diensten extrahiert und kombiniert.

[6] Länderübegreifendes Hochwasserportal: http://www.hochwasserzentralen.de

Statische Web-Seiten sowie das umgebende Rahmenlayout und die Menüsteuerung können leicht in entsprechenden HTML-Rumpfdateien zusammengestellt und gepflegt werden.

Notwendige Konfigurationsänderungen müssen jeweils nur an wenigen zentralen Stellen durchgeführt werden.

Auf der Basis der Erfahrungen während des Betriebes wurde das Verfahren kontinuierlich optimiert. Auch in naher Zukunft sind neue Erweiterungen geplant. So sollen zukünftig aktuelle Sturmflutwarnungen des BSH (Bundesamt für Seeschifffahrt und Hydrographie) automatisch ausgelesen und deren Information in das HSI-Portal integriert werden.

Augmented Maps und Augmented Reality für den präventiven Hochwasserschutz

Frank Fuchs-Kittowski, HTW Berlin, frank.fuchs-kittowski@htw-berlin.de
Stefan Simroth, HTW Berlin, stefan.simroth@htw-berlin.de
Bernd Pfützner, BAH Berlin, bah@bah.de
Frank Trosien, LUGV Brandenburg, frank.trosien@lugv.brandenburg.de

Abstract

In this paper possibilities of use of Smartphone-based mobile augmented maps (AM) and augmented reality (AR) in flood management are presented. Based on application fields in flood management mobile AM and AR applications are described and an information infrastructure supporting their realization is described.

1 Einleitung - Mobile Smartphone-basierte Erweiterte Realität

Wie zuvor das Internet revolutionieren nun mobile Geräte wie Smartphones den Informationsaustausch und die Kommunikation grundlegend. Sie ermöglichen die Verschmelzung von realer Welt mit digitalen Informationen zu einer „Erweiterten Realität“ für Jedermann. Der Begriff Erweiterte Realität (Augmented Reality, kurz AR) bezeichnet die visuelle Ergänzung der optischen menschlichen Wahrnehmung der Realität mit digitalen, kontextabhängigen Informationen in Echtzeit [Azuma, 1997]. Einem Benutzer einer AR-Anwendung werden in seinem Sichtfeld zusätzliche, virtuelle Informationen präsentiert, die in fester räumlicher Beziehung mit Objekten der realen Welt stehen. Ein verbreitetes Beispiel sind die Hilfslinien bei Sportübertragungen im Fernsehen, um z.B. beim Fußball die Entfernung zum Tor bei einem Freistoß oder die Position eines Spielers im Abseits anzuzeigen.

Bei der mobilen Augmented Reality (mAR) werden mobile Endgeräte zur Verschmelzung realer und digitaler Welt genutzt, um die gemeinsame Wahrnehmung von realen und digitalen Informationen möglich zu machen [Höllerer et al., 1999].

Damit werden die digitalen Informationen in ihren räumlichen Kontext gesetzt sowie die Sicht auf die reale Welt durch digitale Informationen und Objekte angereichert und dabei eine neuartige Wahrnehmung und Erfahrung der Wirklichkeit ermöglicht.

Bislang war die mobile Erweiterte Realität (mAR) vor allem Grundlagenforschung, denn die Anwendungen stellen hohe Anforderungen an Hardware und Software, was den breiten Einsatz in der Praxis behindert hat. Diese Geräte waren bisher komplex, teuer sowie unhandlich. Außerdem war ein hoher Aufwand für die Entwicklung von Software für die Darstellung der Inhalte auf diesen Geräten erforderlich.

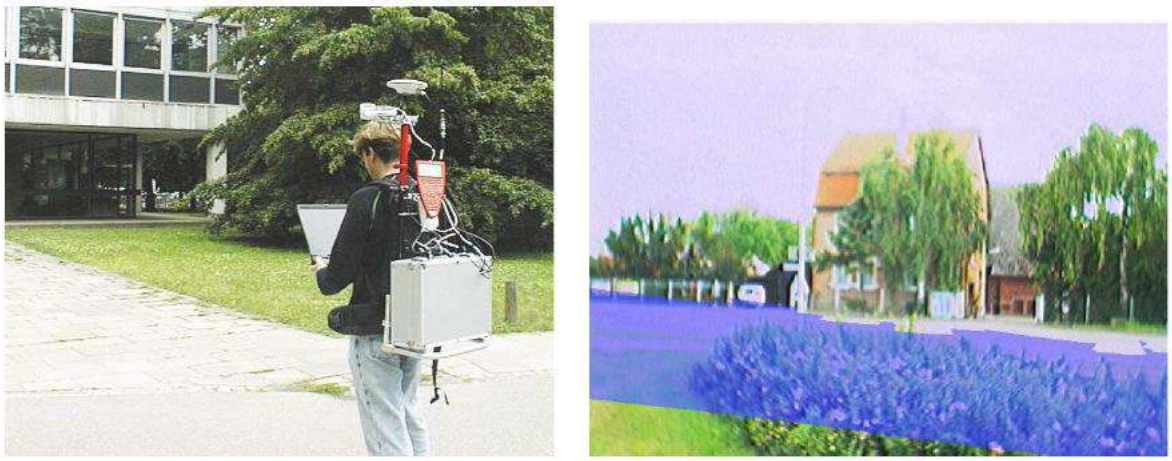


Abbildung 1: mAR-System zur Visualisierung von Hochwasserereignissen [Coelho, 2004]

Aufgrund des rasanten Fortschritts in der Entwicklung mobiler Endgeräte verfügen aber seit kurzem Smartphones der neueren Generation über die notwendige Rechenleistung, eine schnelle Internetverbindung sowie die erforderlichen Sensoren (GPS, Kompass, Beschleunigung) und Komponenten (Display, Videokamera), um Erweiterte Realität realisieren zu können. Außerdem sind sie kostengünstig, leicht handhabbar und massenhaft verbreitet. Neben dieser neuen Hardware existiert seit kurzem auch spezielle Software - sog. AR-Browser - (z.B. Layar, Wikitude, Junaio, Argon), die die Darstellung von Informationen kontextbezogen in Echtzeit im Kamerabild des mobilen Geräts (Smartphone etc.) ermöglicht. Dabei kann es sich um ganz unterschiedliche Arten von Informationen handeln, wie Text, 2D- oder 3D-Objekte, Video- und Audiosequenzen. Zudem ist auch das Ergänzen oder Erfassen von Daten durch die Nutzer möglich. Die AR-Browser lassen sich in andere Anwendungen integrieren, sind in der Regel kostenlos sowie auf vielen Smartphones bereits vorinstalliert.

2 Anwendungen für den präventiven Hochwasserschutz

Aufgrund dieser einfachen und weit verbreiteten technischen Basis sowie einer Vielzahl potenzieller Anwendungsszenarien für unterschiedliche Einsatzbereiche, z.B. in Tourismus, Bildung, Unterhaltung, Werbung, wird mAR ein großes wirtschaftliches Potenzial zugesprochen. Anwendungsmöglichkeiten im Bereich des Hochwasserschutzes sind z.B. Hochwasser-Gefahrenkarten, -Warnung, -Lehrpfade, die derzeit im Projekt MAGUN entwickelt und erprobt werden [MAGUN, 2012].

2.1 Hochwassergefahrenkarten

Einen entscheidenden Beitrag zur Hochwasservorsorge leistet die Informationsvorsorge [Müller, 2010], die u.a. durch die Umsetzung der EU-Hochwasserrisikomanagement-Richtlinie [EG/60/2007, 2007] hergestellt werden soll. Sie sieht die Erstellung von Gefahren- und Risikokarten für Gebiete mit potentiell signifikanten Hochwasserrisiko und darauf aufbauenden Hochwasserrisikomanagementplänen vor. Sie informieren darüber und verdeutlichen, welche Gefahr von Hochwasser grundsätzlich ausgehen kann.

Durch den Einsatz von mobiler Erweiterter Realität kann der durch die Hochwassergefahrenkarten gezeigte, virtuelle Wasserspiegel direkt in der Realität sichtbar gemacht werden. Damit wird es möglich, Hochwasserszenarien realitätsnah zu erzeugen und damit die Wahrnehmung und Analyse von Gefahren zu erweitern.

Hochwassergefahrenkarten lassen sich vor Ort über das mobile Endgerät als Erweiterte Karte bzw. Augmented Map [Reitmayr et al., 2006] (Vogelperspektive) oder als Erweiterte Realität bzw. Augmented Reality (im Kamerabild des Smartphones) darstellen. Dabei kann der Nutzer aus unterschiedlichen Jährlichkeiten des Hochwasserereignisses (HQ10, HQ100, HQExtrem, ...) wählen.



Abbildung 2: Hochwassergefahrenkarte in Kartenansicht (links) und in Kameraansicht als AR (rechts)

Solche mobilen Anwendungen dienen der Verbesserung des Informationsstandes und des Bewusstseins der Bevölkerung über die Hochwassergefahren, um individuelle Vorsorge zu fördern.

Ein Hochwasserrisikobewusstsein besteht beim Bürger oft nur während oder bis kurz nach dem Hochwasserereignis. Ohne ständig wiederkehrende Hochwasserereignisse und die Erinnerung an die bestehenden Gefahren, fällt das Risikobewusstsein relativ schnell wieder auf das Niveau vor dem Hochwasser ab [Müller, 2010]. Zur Förderung und Bewahrung des Risikobewusstseins dienen zusätzlich zu periodisch durchgeführten Informationsveranstaltungen auch o.g. mAR-Anwendungen.

Informationen zur Hochwasservorsorge kommen direkt vor Ort, im Risikogebiet bei den betroffenen oder interessierten Bürgerinnen und Bürgern an. Ein Bürger soll zum Beispiel darüber informiert werden, ob er sich zurzeit in einem von Hochwasser bedrohten Gebiet befindet bzw. wo potentielle Überschwemmungsgebiete liegen. Außerdem soll er wissen können, wie weit das Wasser im Hochwasserfall fließen kann und ob sein Haus und Grund ausreichend vor Hochwasser geschützt ist. Durch eine solche zielgerichtete Risikokommunikation kann eine Verbesserung der Eigenvorsorge erreicht werden.

2.2 Hochwasserwarnung

Damit die aktuell von einem Hochwasser bedrohten Bürger rechtzeitig Maßnahmen zu ihrem Schutz ergreifen können, müssen sie schnell aktuelle Informationen über

den aktuellen und erwarteten Wasserstand bzw. die aktuelle Hochwassergefahr erhalten [Hornemann, et al., 2006].

Aktuelle Hochwassermelde-Pegel lassen sich als Liste, Karte und Erweiterte Realität im mobilen Endgerät darstellen.



Abbildung 3: Hochwassermeldepegel auf Karte (links) und in Kamera als AR (rechts)

Eine solche mobile Anwendung dient der besseren Informationsversorgung, frühzeitigen Warnung und besseren Erreichbarkeit der Bevölkerung. Zusätzlich kann die Bevölkerung durch Hochwasserwarnungen über die aktuelle Hochwassergefahr unterrichtet werden. Solche Warnungen können auch aktiv an die Nutzer übermittelt werden (Push-Notification). So kann die Bevölkerung rechtzeitig etwas zu ihrem Schutz tun (z.B. Türen oder Fenster mit Dammbalken oder Sandsäcken sichern).

2.3 Bewusstseinsaktivierung (Historische Hochwassermarken)

Historische Hochwassermarken dienen dazu, die Wahrnehmung für die bestehende Hochwassergefahr an in der Vergangenheit überfluteten Orten zu stärken und wach zu halten. Sie zeigen Überflutungshöhen von historischen Hochwassern an und erinnern damit an diese vergangenen Überschwemmungen [Petrow, et al., 2003].

Hochwassermarken können von der interessierten Bevölkerung selbst erfasst (Foto inkl. Metadaten) und als Liste, Karte und Erweiterte Realität im mobilen Endgerät dargestellt werden.

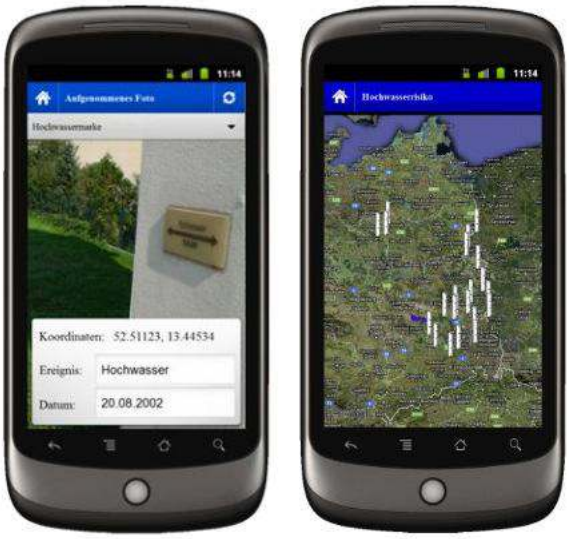


Abbildung 4: Erfassung einer histor. Hochwassermarke (links) und Darstellung auf Karte (rechts)

Eine solche mobile Anwendung dient der Involvierung und Aktivierung der Bevölkerung sowie der Steigerung des Bewusstseins der Bevölkerung über die Gefahren und Schadenspotenziale von Hochwasser. Es soll nicht nur die Erinnerung an vergangene Hochwasser-Ereignisse bewahrt werden. Darüber hinaus soll die Bevölkerung auch zu einem kontinuierlichen Engagement für Vorsorge veranlasst werden, in dem sie sich der spezifischen Hochwassergefahren bewusst werden.

2.4 Gewässerinformationen

Informationen über Gewässer lassen sich vor Ort abrufen und auf einer Karte oder als Erweiterte Realität darstellen. Sie können interessierten Bürgern als Informationsquelle (Name des Gewässers oder des Bauwerks, Einzugsgebietsgröße etc.) dienen. Sie können aber auch der Unterstützung von hydrologischen Fachexperten vor Ort dienen; z.B. durch Informationen wie Stationierung, in geeigneter Form aggregierte, gemessene oder berechnete Wasserstände und Abflüsse (Gewässerkundliche Hauptzahlen, Bemessungsabflüsse etc.), Ausbauzustand, Pflege-Rhythmus.

Derzeit müssen aber die zu einem Vor-Ort-Termin benötigten Daten entweder als analoge Karten, als manuelle Notiz oder Ausdruck auf Papier bzw. digital auf einem Laptop mitgebracht werden. Der Nutzer dieser Informationen muss sie weiterhin vor Ort interpretieren und verifizieren sowie sich in der realen Umgebung orientieren. Dabei ist sein Blick entweder auf die hydrologische Information oder auf das Gebiet gerichtet, niemals hat er eine gemeinsame Sicht auf das Gebiet und die zugehörige Informationen. Hierbei kann Erweiterte Realität (Augmented Reality, AR) einen

wesentlichen Beitrag leisten. AR bietet das Potenzial, die Modelldaten nicht nur über ihren Raumbezug vor Ort abzufragen, sondern diese auch direkt mit der realen Umwelt zu verbinden. Dies wird u.a. durch eine Überlagerung der Sicht auf die reale Umgebung mit Visualisierungen der Modelldaten erreicht.



Abbildung 5: Flusseinzugsgebiet als Karte (links) und Abfluss in Kamera als AR (rechts)

2.5 Hochwasser-Lehrpfad

Ein (Natur-) Lehrpfad dient verschiedenen Aufgaben. Dazu gehören u.a. die Umweltbildung, die Förderung der Regionalentwicklung, die Besucherlenkung und die Vermittlung einer spezifischen Thematik, z.B. der Sensibilisierung und Bewusstseinsbildung über Hochwassergefahren [Szekeres, 1999].

Mit einer mAR-Anwendung können Informationen über die Position der Objekte auf einem Lehrpfad sowie zugehörige Lern-Inhalte im Kamerabild angezeigt werden. Dadurch können die einzelnen Objekte leichter gefunden werden (Orientierung im Gelände, Wegweiser auf dem Pfad) und aktuelle (Zusatz-) Informationen kostengünstig bereitgestellt werden (multimedialer, selbstgesteuerter Lernprozess).

3 Technische Realisierung

3.1 Anforderungen

Wie die o.g. Beispiele zeigen, stellen mAR-Anwendungen ganz unterschiedliche Anforderungen an das unterstützende Backend-System.

Gemeinsam ist bei allen Anwendungen, dass Hochwasser-Daten auf einer Karte dargestellt (Augmented Map) oder im Kamerabild des mobilen Geräts eingeblendet (Augmented Reality) werden sollen. Der örtliche Kontext (Pose) ist durch GPS-Koordinaten und die Blickrichtung des Nutzers definiert. Der zeitliche Kontext ist immer die Darstellung in Echtzeit, wobei die Daten einen eigenen Zeitbezug (Echtzeit, Historisch, Aktuell, Zukünftig) besitzen können.

Dagegen ist die Art der Daten (z.B. Texte, 2D-Bilder, 3D-Objekte, Videos, Audio), genau wie die Thematik (Gefahrenkarten, Pegeldaten etc.) sehr unterschiedlich. Es müssen verschiedene Datenformate aus unterschiedlichen Quellen beherrscht werden. Damit diese heterogenen Daten aus verschiedenen Quellen jedoch auch in einem gemeinsamen System genutzt und ggf. neu kombiniert werden können, müssen die Input-Daten aufbereitet und zumindest intern in ein gemeinsames Format gebracht werden. Auch die Datenausgabe ist sehr vielfältig und erfordert unterschiedliche Datenformate für die Bereitstellung. Z.B. besitzt jeder AR-Browser sein eigenes Datenformat, und für die Darstellung spezieller Inhalte sind meist bestimmte Formate erforderlich (z.B. Gefahrenkarten etc.). Zudem sind unterschiedliche Nutzermengen (wenige bis viele) und Nutzergruppen (Fachleute, Bevölkerung etc.) zu unterstützen.

3.2 mAR-Informations-Infrastruktur

Zur Realisierung der o.g. mAR-Anwendungen wurde im Projekt MAGUN eine mAR-Informations-Infrastruktur für die Erstellung, Integration, Verwaltung und Bereitstellung von Inhalten für mAR-Anwendungen auf Basis von Open Source-Software (GeoServer, OSMDroid etc.) entwickelt.

Diese unterstützt zum einen die Integration von vorhandenen Daten aus verschiedenen Quellen in unterschiedlichen Formaten (multi-sourcing) und zum anderen die Erfassung neuer Inhalte vor Ort (in-situ), wobei unterschiedliche Methoden für die Erfassung der Inhalte realisiert werden, d.h. sowohl als offener kollaborativer

Prozess durch die Endnutzer (crowd sourcing, UGC, VGI) als auch als geschlossener, von spezialisierten Personen durchgeführter Prozess. Darüber hinaus wird die Wiederverwendung der Inhalte in verschiedenen Kontexten durch die Aggregation, Interoperabilität und Rekombination sowie die Bereitstellung für verschiedene AR-Browser und ggf. weitere Dienste über standardisierte Schnittstellen (Multi-Channel) ermöglicht.

Diese Infrastruktur soll Unternehmen dabei helfen, ihre vorhandenen Daten bzw. Inhalte auf mobilen Endgeräten und in mAR-Anwendungen nutzbar zu machen, zu erweitern und aktuell zu halten – auch mit Nutzern ohne spezielle technische Kenntnisse.

3.3 Architektur und Implementierung

Die mAR-Informations-Infrastruktur wurde als Client-Server-Infrastruktur konzipiert und implementiert, bei der mehrere Serveranwendungen miteinander interagieren, um Dienste (Webservices) für anfragende mAR-Clients zur Verfügung zu stellen.

Zu den aufeinander abgestimmten Serveranwendungen gehören Frontend-Webservices, Geodaten-Services und eine GIS-Datenbank. Ersterer bietet Webservices für Client-Anfragen und realisiert u.a. auch die Anbindung an die verschiedenen AR-Browser. Darüber hinaus ist die Ausgabe in übliche Geo-Datenformate (wie KML, GeoRSS) möglich. Der ebenfalls als Java Web App implementierte „GeoServer“ wird genutzt, um erforderliche Geodaten standardkonform über Webservices bereitzustellen (WMS, WFS, WCS, WTS). So können z.B. die Karten über WMS oder WTS an den Client oder auch als Vektordaten für das Rendering auf dem Client ausgeliefert werden. Für die Datenhaltung wird PostgreSQL mit dem PostGIS-Aufsatz eingesetzt. Es ist möglich verschiedenste Datenquellen, wie relationale Datenbanken, Webservices und Dateien (Shape-Dateien und andere Dateiformate), anzubinden und zu verwenden.

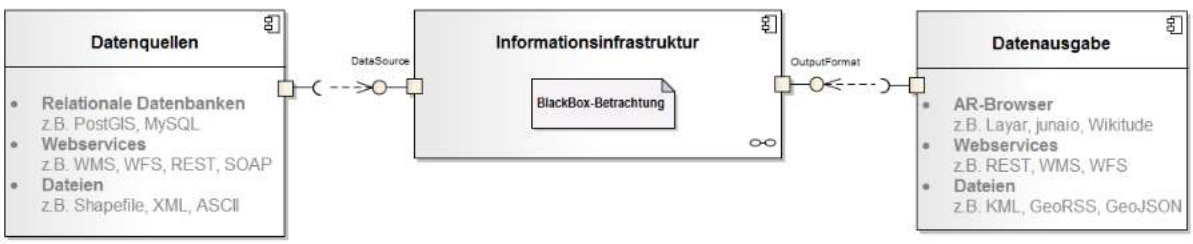


Abbildung 6: mAR-Informations-Infrastruktur - Schnittstellen für Datenquellen und Datenausgabe

Innerhalb der mAR-Infrastruktur erfolgt die Anbindung konkreter Datenquellen („Data Sources“), deren Abbildung auf Geoobjekte („Features“) und deren Transformation in konkrete Ausgabekanäle („Output Channels“).

Implementationen von Datenquellen lassen sich über die Schnittstelle „DataSource“ beitragen. Eine Datenquelle gibt für einen parametrisierten Aufruf eine Menge von DataEntity-Objekten zurück, die einen Datensatz in einer Tabellenzeile („row“) entsprechen. Zur Beschreibung der Datenstruktur muss eine Datenquelle ein Data-Schema-Objekt zurückgeben.

Über ein „Source Mapping“ werden aus den DataEntity-Objekten dann Geoobjekte („Features“), die durch einen Geoobjekttyp („FeatureTyp“) beschrieben werden. Dieses Quellen-Mapping wird in eine Tabelle zwischengespeichert, die die Struktur des FeatureType aufweist. Somit lässt sich diese Tabelle auch als FeatureType im GeoServer verwenden.

Die Ausgabe der Daten erfolgt in verschiedenen Formaten und wird über Ausgabekanäle („Output Channel“) gesteuert. Diese sind ähnlich zu Datenquellen aufgebaut, mit dem Unterschied, dass sie Geoobjekte („Features“) verarbeiten. Die Transformation erfolgt mittels eines Ausgabe-Mappings („Output Mapping“).

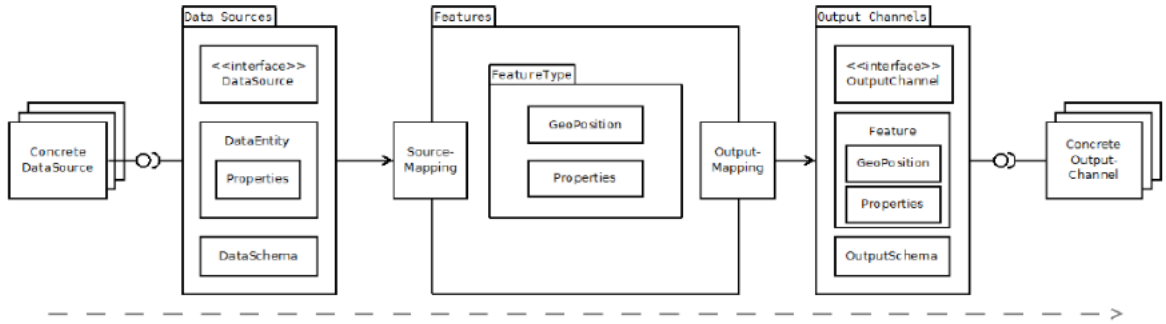


Abbildung 7: mAR-Informations-Infrastruktur – Datenintegration und –transformation

4 Zusammenfassung und Schlussfolgerung

Der Erfolg des World Wide Web (WWW) kann mit dem Begriff „Augmentieren“ (engl. augmentation) zusammengefasst werden. Im WWW ist ein Hypertext eine Methode, um relevante Informationen miteinander zu verbinden. Ein Hypertext augmentiert dabei die existierende Information mit anderen Informationen. Der Erfolg des Web 2.0 (und insbesondere von Wikipedia, der Blogosphäre etc.) basiert also nicht nur auf den nutzer-generierten Seiten, sondern auch auf der Möglichkeit, jeden Beitrag mit Links zu anderen Beiträgen anzureichern. Derzeit weitet die Erweiterte Realität dieses Konzept von der Online-Welt auf die natürliche Welt aus. Die zunehmende Leistungsfähigkeit von Smartphones bietet ein geeignetes Instrument, um digitale Informationen mit der natürlichen Welt zu assoziieren. Sie bieten darüber hinaus auch die Möglichkeit, Informationen der natürlichen Umgebung digital zu erfassen sowie über die digitalen Repräsentationen mit den natürlichen Objekten zu interagieren.

Mit der neuen Generation an weit verbreiteten, leistungsfähigen und kostengünstigen mobilen Endgeräten (Smartphones) sowie den plattformübergreifend vorhandenen AR-Browsern für die Darstellung von AR-Inhalten auf diesen mobilen Endgeräten steht nun eine technische Infrastruktur zur Verfügung, die eine kostengünstige Entwicklung mobiler AR-Anwendungen sowie eine massenhafte Nutzung dieser AR-Anwendungen durch Jedermann ermöglicht.

mAR hilft bei der Orientierung in der Realität und schafft im Ortskontext eine unmittelbare neue mediale Erfahrung. mAR erweitert die Wahrnehmung des Ortes durch die Anreichung mit Informationen aus Vergangenheit (Historische Hochwasser-Marken), Gegenwart (Hochwasser-Meldepegel) oder Zukunft (Hochwasser-Gefahrenkarten).

mAR bietet das Potenzial, Hochwasser-relevante Informationen (Messwerte, Modelldaten etc.) nicht nur über ihren Raumbezug abzufragen (ortsbezogene Dienste), sondern diese auch intelligent mit der „realen Welt“ zu verknüpfen. Die digitalen Informationen werden in ihren räumlichen Kontext gesetzt sowie die Sicht auf die reale Welt durch digitale Informationen und Objekte angereichert. Auf diese Weise ist zum einen eine bessere Interpretation und genauere Validierung der Daten sowie eine effizientere Entscheidungsunterstützung möglich. Zum anderen wird dabei eine

neuartige Wahrnehmung und Erfahrung der Wirklichkeit ermöglicht sowie die Interaktion mit der Realität verbessert.

Zudem bietet mAR die Möglichkeit, Hochwasser-Informationen in die Breite zu bringen. Smartphones sind inzwischen massenhaft verbreitet und die Entwicklung von mAR-Anwendungen aufgrund der Verfügbarkeit kostenloser AR-Browser effizient möglich. Somit lassen sich mAR-Anwendungen nicht nur für wenige Fachexperten entwickeln, sondern von einer riesigen Anzahl an Nutzern (z.B. „normale“ Bürger) nutzen. Sie dienen so zur Information über Hochwassergefahren oder auch zur Erfassung von relevanten Informationen, z.B. bei oder nach Hochwassern (Geschwemmsellinien).

Danksagung

Das Projekt MAGUN wird aus Mitteln des Europäischen Fonds für regionale Entwicklung (EFRE) sowie des Instituts für angewandte Forschung (IfaF) Berlin gefördert. Die Autoren danken den Mittelgebern für die Unterstützung sowie allen MAGUN-Projektpartnern für die fruchtbare Zusammenarbeit.

5 Literaturverzeichnis

[Azuma, 1997]

Azuma, Ronald T.A.: A Survey of Augmented Reality. In: Presence - Teleoperators and Virtual Environments. Vol. 6, No. 4 (August 1997), S. 355 - 385.

[Coelho, 2004]

Coelho, Alexandre Hering: Erweiterte Realität zur Visualisierung simulierter Hochwasserereignisse. Dissertation, Universität Fridericiana zu Karlsruhe, 2007.

[EG/60/2007, 2007]

Europäische Gemeinschaft: Richtlinie 2007/60/EG des Europäischen Parlaments und des Rates vom 23. Oktober 2007 über die Bewertung und das Management von Hochwasserrisiken. ABl. L 288 vom 06.11.2007.

[Höllerer, et al., 1999]

Höllerer, T.; Feiner, S.; Terauchi, T.; Rashid, G.; Hallaway, D.: Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System, Computers and Graphics, 23(6), Elsevier Publishers, 1999, S. 779-785.

[Hornemann, et al., 2006]

Hornemann, Corinna; Rechenberg, Jörg: Was Sie über den vorsorgenden Hochwasserschutz wissen sollten. Dessau: Umweltbundesamt, 2006.

[MAGUN, 2012]

MAGUN: Mobile Anwendungen auf Basis von Geoinformationen in einer In-situ Informationsinfrastruktur im Umwelt- und Navigationsbereich, URL: http://www.magun-projekt.de/, 2012.

[Müller, 2010]

Müller, Uwe: Hochwasserrisikomanagement - Theorie und Praxis. Vieweg & Teubner, 2010.

[Petrow, et al., 2003]

Petrow, Theresia; Thieken, Annegret; Kreibich, Heidi; Merz, Bruno: Vorsorgende Maßnahmen zur Schadensminderung. In: Hochwasservorsorge in Deutschland - Lernen aus der Katastrophe 2002 im Elbegebiet. Schriftenreihe des DKKV 29, Bonn: Deutsches Komitee für Katastrophenvorsorge e. V. (DKKV), 2003.

[Reitmayr, et al., 2005]

Reitmayr, Gerhard; Eade, Ethan; Drummond, Tom: Localisation and Interaction for Augmented Maps. In: Proc. IEEE ISMAR'05, October 5-8, 2005, Vienna, Austria. URL: http://mi.eng.cam.ac.uk/~gr281/augmentedmaps.html, 2005.

[Szekeres, 1999]

Szekeres, Peter: Naturlehrpfade. 3. Aufl., Marburg, Institut für Ökologie, URL: http://www.projektwerkstatt.de/download/texte_cd/reader/lehrpfade.pdf, 1999.

GIS-Komponenten zur Untersuchung von Antibiotika und Antibiotikaresistenz im urbanen Abwasser

Luise Hutka, Susanne Reinwarth, Daniel Kadner, Lars Bernard

Technische Universität Dresden

Luise.Hutka@tu-dresden.de
Susanne.Reinwarth@tu-dresden.de
Daniel.Kadner@tu-dresden.de
Lars.Bernard@tu-dresden.de

Abstract

The project ANTI-Resist will research the release of antibiotics and the development of antibiotic resistances in an urban sewage system using the example of the city of Dresden. The project is part of the program "Research for Sustainable Development" of the Federal Ministry for Education and Research (BMBF). It is a cooperation between the Technical University of Dresden (Forschungsverbund Public Health, Institute of Clinical Pharmacology, Institute for Urban Water Management, Institute of Hydrobiology, Professorship of Geoinformation Systems) and the Stadtentwässerung Dresden GmbH. The overall objective is to design strategies to reduce the release of antibiotics as well as the formation of antibiotic resistances and to develop corresponding monitoring and warning systems. (FVPH o.J.)

Einleitung

Das Projekt ANTI-Resist zielt auf die Ermittung von Antibiotikaeinträgen und die Untersuchung der Bildung von Antibiotikaresistenzen im urbanen Abwasser am Beispiel der Stadt Dresden. Das Projekt wird im Rahmen des Programms „Forschung für nachhaltige Entwicklungen“ des Bundesministeriums für Bildung und Forschung (BMBF) gefördert und in Kooperation der Technischen Universität Dresden (Forschungsverbund Public Health, Institut für Klinische Pharmakologie, Institut für Siedlungswasserwirtschaft, Institut für Hydrobiologie, Professur für

Geoinformationssysteme) sowie der Stadtentwässerung Dresden GmbH bearbeitet. Das generelle Ziel ist es, geeignete Strategien zur Minderung des Eintrags und der möglichen Resistenzbildungen zu konzipieren und Monitoring- und Frühwarnsysteme zu entwickeln. (FVPH o.J.)

1 Geoinformationssysteme zur Untersuchung von Antibiotika im Abwasser

Arzneimittelrückstände im Abwasser sind seit geraumer Zeit ein wichtiges Thema im Zusammenhang mit den Zielsetzungen der Abwasserbehandlung (Rohde 2011, Engelmann 2011). Es gibt zahlreiche Untersuchungen zum Wirkungsgrad verschiedener Verfahren, bei denen vereinzelt auch das spezifische Verhalten von Antibiotika und sich daraus entwickelnder Antibiotikaresistenz Untersuchungsgegenstand ist. In dem Projekt soll der Problembereich von Antibiotika und Antibiotikaresistenzen in den urbanen Abwassersystemen von der Medikamentenverschreibung über den Transport bis zur Abgabe in die Gewässer beschrieben werden, denn ein erhöhter Eintrag von Antibiotika in die Umwelt begünstigt die Bildung antibiotikaresistenter Bakterien. Ist ein Erreger resistent gegen ein bestimmtes Antibiotikum, so ist die Einnahme dieses Antibiotikums bei einer Infektion mit dem Erreger unwirksam.

Ein Werkzeug zur Analyse räumlicher und zeitlicher Verteilungsmuster stellen Geoinformationssysteme (GIS) dar. Heute verfügbare GIS bieten zahlreiche Methoden zur Modellierung der räumlichen Verteilung von Antibiotika-Einträgen, etwa zur Formulierung räumlicher Abfragen, zur räumlichen und thematischen Aggregation, zur Visualisierung sowie zur räumlichen Verschneidung. Die derzeit entstehenden Geodateninfrastrukturen (GDI) unterstützen Suche, Evaluierung, Bezug und Visualisierung benötigter Geodaten und damit die Realisierung effizienter Fachinformationssysteme zum Zugriff auf aktuelle, verteilt vorliegende Geodatenquellen. Geoportale (Bernard et al. 2005) dienen hierbei als Anwendungen, die die in GDI verfügbaren verteilten Geodaten und -dienste für spezielle Fachanwendungen und Nutzergruppen zusammenstellen.

Mit der Umsetzung der EU-Direktive INSPIRE (Infrastructure for Spatial Information in the European Community, EC 2007) findet derzeit der Aufbau einer europäischen GDI statt. Aktuelle GDI-Forschungsarbeiten zielen auf die Verbesserung verteilter

Prozessierung von Geodaten zur Ableitung ad hoc benötigter Geoinformationen, welche eine wesentliche Grundlage für den Aufbau von internetbasierten und breit nutzbaren Entscheidungsunterstützungs- und Frühwarnsystemen bilden (Kiehle et al. 2007, Bernard & Ostländer 2008, Müller et al. 2010). Der Umgang mit zeitvarianten Geodaten und mit unterschiedlichen (Zukunfts-)Szenarien in GIS- und GDI-Anwedungen ist im Kontext dieses Projekts eine weitere Herausforderung für die Geoinformatikforschung.

2 Fallstudie Abwassersystem Dresden

Rund 97 % des im Stadtgebiet von Dresden anfallenden Abwassers wird der Kläranlage Dresden-Kaditz zugeführt, davon ca. 75 % über ein Mischsystem. Da in einer mechanisch-biologischen Anlage wie der Dresdner Kläranlage Antibiotika kaum abgebaut werden, ist davon auszugehen, dass die weitgehend diffus in das Abwasser eingetragenen Antibiotika entweder über den Kläranlagenablauf oder durch die stoffliche Verwertung des Klärschlammes in die Umwelt gelangen. Zudem wird aufgrund des Mischsystems ein Teil des im Einzugsgebiet der Kläranlage anfallenden Abwassers über die Regenentlastungsbauwerke direkt bzw. nach mechanischer Behandlung in das Gewässer abgeschlagen.

3 GIS-Komponenten in ANTI-Resist

3.1 GIS-basierte Analyse

Im Rahmen des Projektes werden zunächst ambulante und stationäre Daten zur Verschreibung von Antibiotika statistisch und GIS-basiert aufbereitet, um den räumlich und zeitlich verteilten Eintrag in das Abwassersystem modellhaft zu beschreiben. Diese Ergebnisse fließen in die Entwicklung von entsprechenden GIS-Analysewerkzeugen ein. Die Verschneidung mit zusätzlichen relevanten Geodaten, z.B. zur Demografie, dient der Untersuchung der regionalisierten Antibiotika-Einträge auf korrelierende räumliche und zeitliche Verteilungsmuster. Ziele sind hierbei die räumliche Übertragbarkeit der ermittelten Antibiotika-Einträge sowie Aussagen zu Orten und Zeitpunkten mit potenziell besonders hohen Antibiotika-Aufkommen im Abwasser.

3.2 Konzeption eines Frühwarnsystems

Die aus den Verschreibungsdaten ermittelten Verteilungsmuster der Antibiotika-Einträge werden unter Berücksichtigung des Transportverhaltens mit den Ergebnissen der Messkampagnen fusioniert. Damit wird die Aussagekraft von Verschreibungsdaten im Hinblick auf eine Vorhersage kritischer Zustände evaluiert. Zudem ermöglicht dies, den Transport von Antibiotika im Abwassersystem und kritische Zustände für die Bildung von Antibiotikaresistenzen aufgrund von Verschreibungsinformationen vorherzusagen.

Basierend auf den Analysen zur Korrelation der ermittelten Antibiotika-Einträge mit weiteren Parametern (Klima, besondere Ereignisse, etc.) und der Prozessanalyse aus den vorangegangenen Teilschritten sollen mögliche Prediktorvariablen für den Eintrag von Antibiotika ermittelt und für den Aufbau eines Frühwarnsystems der zu erwartenden Antibiotikabelastungen genutzt werden. Dafür gilt es, entsprechende Warnlevel und Informationsanforderungen zu definieren und die Komponenten dieses Frühwarnsystems prototypisch umzusetzen. Ein solches, bisher nicht verfügbares Instrument kann sowohl der Unterstützung einer möglichst effektiven Abwasserbehandlung als auch zur Steuerung eventuell erforderlicher medizinischer Maßnahmen dienen.

3.3 ANTI-Resist-Geoportal

Die Ergebnisse der GIS-basierten Analyse sowie das Frühwarnsystem sollen Wissenschaftlern, Entscheidungsträgern und der Öffentlichkeit über ein webbasiertes Informationssystem verfügbar gemacht werden. Die zu konzipierende und zu entwickelnde ANTI-Resist-Geoportallösung soll eng mit der aktuell entstehenden GDI Sachsen und der GDI-DE verknüpft werden. Dies ermöglicht einerseits, andere Geoinformationsdienste über standardisierte Schnittstellen in das ANTI-Resist Geoportal einzubinden. Andererseits können die im Projekt entwickelten Geoinformationsdienste weiteren internetbasierten Informationsanwendungen zur Nutzung und Analyse bereitgestellt werden. Die Komponenten und Dienste des Geoportals werden prototypisch implementiert, um die Realisierbarkeit und den Nutzen der entwickelten Konzepte zu demonstrieren.

4 Ansatz zur GIS-basierten Ermittlung von Antibiotika-Einträgen

Ziel der GIS-basierten Analyse ist es, Antibiotika-Einträge in das Dresdner Kanalnetz unter räumlichen und zeitlichen Kriterien zu untersuchen. Daraus sollen später Prognosen über das zu erwartende Antibiotika-Aufkommen und die Bildung von Antibiotikaresistenzen abgeleitet sowie ein Frühwarnsystem entwickelt werden.

Dafür werden zwei verschiedene Modelle entwickelt: das Eintrags- und das Prognosemodell. Den Zusammenhang der beiden Modelle zeigt Abbildung 1. Das Eintragsmodell erfasst die Ursachen der Antibiotika-Einträge, d.h. welche Eintragsgruppen wann wo wie viel Antibiotika in das Dresdner Kanalnetz eintragen. Eintragsgruppen sind Quellen, durch die Antibiotika in das Abwasser gelangen, z.B. Patienten verschiedener Altersgruppen. Hauptgrundlage für das Eintragsmodell bilden die ambulanten AOK-Verordnungsdaten der letzten Jahre auf Stadtteilebene, die eine flächendeckende Betrachtung des Dresdner Stadtgebietes ermöglichen. Diese werden räumlich und zeitlich z.B. mittels Zeitreihen- und Clusteranalyse ausgewertet, um mögliche Muster der Antibiotikaverordnungen zu erkennen. In diesem Zusammenhang müssen auch verschiedene Anonymisierungsstrategien in Betracht gezogen werden, um ausreichend hohe Fallzahlen für die Durchführung von Analysen zu gewährleisten. Im Weiteren werden die Verordnungsdaten mit anderen Parametern verknüpft und im Hinblick auf Korrelationen untersucht. Als bedeutende Einflussfaktoren wurden demografische, soziale und meteorologische Parameter identifiziert und entsprechende Daten beschafft und aufbereitet.

Bevölkerungsdaten liegen auf Stadtteilebene nach Altersgruppen und Geschlecht für die betreffenden Jahre vor. Als sozialer Parameter wird die Beschäftigungsrate pro Stadtteil und Jahr herangezogen. Hinsichtlich der meteorologischen Einflussfaktoren werden z.B. Temperatur(schwankungen), Niederschlag, Luftfeuchte, Einstrahlung und Wind betrachtet. Diese Parameter werden nun mittels Korrelations- und Regressionsanalysen auf statistische Zusammenhänge mit den Verordnungsdaten untersucht. Zusätzlich sollen Diagnosedaten einbezogen werden, um die den Verordnungen zugrundeliegenden Krankheiten und wiederum deren Ursachen zu identifizieren. Daraus sollen weitere bedeutende Einflussfaktoren abgeleitet werden, die ggf. später in das Prognosemodell einfließen. Außerdem sollen stationäre Verordnungsdaten der städtischen Krankenhäuser als punktuelle Eintragspfade auf

zeitliche Muster hin untersucht werden. Als Werkzeug für die geostatistischen Analysen kommen hauptsächlich ArcGIS Desktop sowie die statistische Programmiersprache R zum Einsatz.

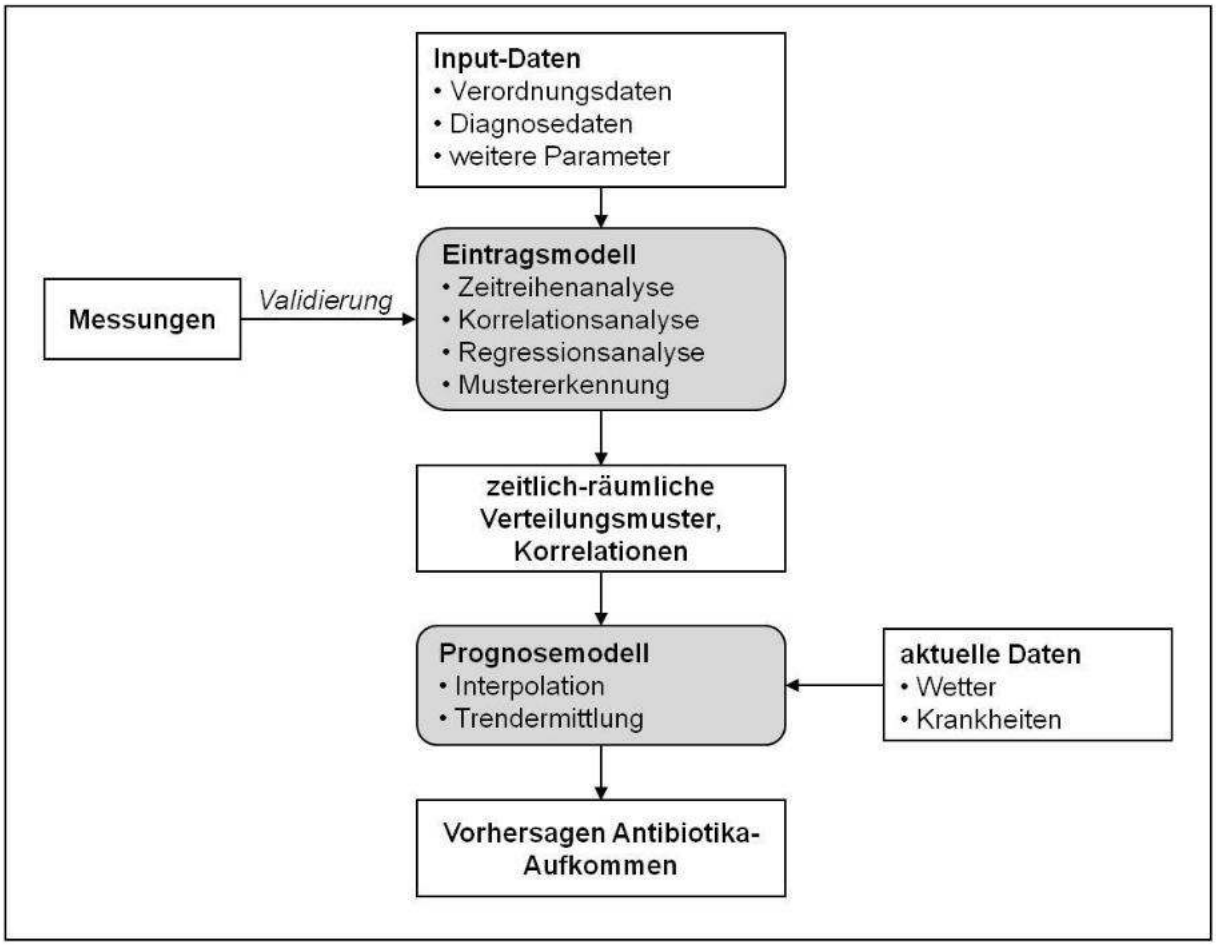


Abbildung 1: Modellübersicht zur GIS-basierten Ermittlung von Antibiotika-Einträgen

Um die Qualität des Eintragsmodells, auch im Hinblick auf dessen Prognosefähigkeit, zu bewerten, werden die modellierten Antibiotika-Einträge mit den im Verlauf des Projektes gemessenen Antibiotika-Aufkommen in der Kanalisation abgeglichen. Diese Messungen werden an ausgewählten Stellen im Abwasser innerhalb des Dresdner Stadtgebietes durchgeführt und haben eine zeitliche Auflösung von Stunden- bzw. Tageswerten. Darauf aufbauend wird ein Prognosemodell entwickelt, welches Vorhersagen zum zu erwartenden Antibiotika-Aufkommen im Abwasser bestimmter Stadtgebiete liefern soll. Prognosen werden durch die Kombination zeitnaher Informationen (z.B. Wetterdaten, Auftreten bestimmter Krankheiten) und den Ergebnissen des Eintragsmodells mit Hilfe von Inter-/Extrapolations- und Trendermittlungsverfahren ermittelt. Basierend auf den Vorhersagen zu Antibiotika-Ein-

trägen können kritische Zustände für die Bildung von Antibiotika-Resistenzen abgeleitet werden.

5 Literaturverzeichnis

[Bernard, et al., 2005]

Bernard, L., Kanellopoulos, I., Annoni, A. und Smits, P.: The European Geoportal - One step towards the Establishment of a European Spatial Data Infrastructure. *Computers, Environment and Urban Systems* (29), 15-31.

[Bernard, et al., 2008]

Bernard, L. und Ostländer, N.: Assessing climate change vulnerability in the arctic using geo-graphic information services in spatial data infrastructures. *Climatic Change* (87), 263-281.

[EC, 2007]

EC-2007: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Council, European Parliament and European, Official Journal of the European Union: L 108/1 - L 108/14.

[Engelmann, 2011]

Engelmann, U.: Befunde in sächsischen Kläranlagenabläufen. In: Statusseminar "Arzneistoffe in der Umwelt". Dresden, 2011. – URL http://www.umwelt.sachsen.de/umwelt/wasser/download/14_Befunde_in_saechsischen_Klaeranlagenablaeufen_Engelmann.pdf (Zugriff am 06.03.2012)

[Forschungsverbund Public Health Sachsen und Sachsen-Anhalt (FVPH, o.J.)]

ANTI-Resist - URL http://anti-resist.de/ (Zugriff am 05.03.2012)

[Kiehle, et al. 2007]

Kiehle, C., Greve, K. und Heier, C.: Requirements for Next Generation Spatial Data Infrastructures – Standardized Web Based Geoprocessing and Web Service Orchestration. *Transactions in GIS* 11(6), 819-834.

[Müller, et al., 2009]

Müller, M., Bernard, L. und Vogel, R.: Multi-Criteria Evaluation for Emergency Management in Spatial Data Infrastructures. In: Zlatanova, Sisi, Konecny, Milan, Bandrova, Temenujka (Hrsg.), Geographic Information and Cartography for Risk and Crisis Management, Lecture Notes in Geoinformation and Cartography (2010), S. 273-28.

[Rohde, 2011]

Rohde, S.: Untersuchungsergebnisse in Oberflächen- und Grundwasser in Sachsen. In: Statusseminar "Arzneistoffe in der Umwelt". Dresden, 2011. – URL http://www.umwelt.sachsen.de/umwelt/wasser/download/17_Untersuchungsergebnisse_in_OW_und_GW_in_Sachsen_Rohde.pdf (Zugriff am 06.03.2012)

Die GSBL-Kommunikationsplattform – ein gemeinsames Portal für Bund und Länder

Barbara Liebscher, Umweltbundesamt Dessau-Roßlau
barbara.liebscher@uba.de

Abstract

The coordination office of the Joint Substance Data Pool of the German Federal Government and the German Federal States (GSBL) is set up at the Federal Environment Agency (Umweltbundesamt, UBA). The coordination office updates the data of the GSBL database on the ORACLE database of the operator's centre twice a year. This set of updated data is the basis for the search databases provided to the users. The search database GSBLintern provides the totale data pool for authorities of the German Federal Government and the German Federal States. The version GSBLpublic holds an excerpt of the complete database, which is freely available to the general public.

In this research project a web-based communication platform was conceptualized and developed which supports the production process in the coordination office extensively and acts as a central contact point for the cooperation community for the exchange of information, data and software. The GSBL communication platform provides all partners involved in the production process with a software for production tracing which improves cooperation and protects the quality of the data stream.

Einleitung

Im Umweltbundesamt ist die Koordinierungsstelle des Gemeinsamen zentralen Stoffdatenpools des Bundes und der Länder (GSBL) etabliert. Hier erfolgt zweimal jährlich die Aktualisierung des Datenbestandes des GSBL. Die Recherche-Anwendung des GSBL mit dem Gesamtdatenbestand steht den Behörden von Bund und Ländern zur Verfügung, ein Teildatenbestand als GSBLpublic der interessierten Öffentlichkeit.

Im Rahmen eines vom BMU-finanzierten Vorhabens wurde eine Web-basierte Kommunikationsplattform konzipiert und entwickelt, die den Produktionsprozess in der Koordinierungsstelle umfassend unterstützt, die Produktionsverfolgung gewährleistet und die Qualität des Datenstroms sichert. Die Plattform fungiert zudem als zentraler Anlaufpunkt für die Kooperationsgemeinschaft zum Austausch von Informationen, Daten und Software.

1 Gemeinsamer zentraler Stoffdatenpool des Bundes und der Länder (GSBL)

Der gemeinsame zentrale Stoffdatenpool des Bundes und der Länder ist die größte und wichtigste Stoffdatenbank in Deutschland. Auf der Grundlage einer Verwaltungsvereinbarung arbeiten alle 16 Länder und der Bund bei der Erschließung, Pflege und Bereitstellung von einheitlichen und fachlich abgesicherten Informationen über Eigenschaften von Gefahrstoffen zusammen.

1.1 GSBL-Stoffdatenpool

Seit 1998 ist der GSBL im Produktionsbetrieb und stellt zu ca. 300.000 chemischen Stoffen in mehr als 480 Merkmalen Informationen zu Stoffeigenschaften und rechtliche Regelungen bereit.

Das Umweltbundesamt organisiert und betreibt den GSBL in Zusammenarbeit mit Behörden des Bundes und der Länder, die gemeinsam die Stoffdaten aufbereiten und pflegen.

Die gesicherten, umfassenden und aktuellen Stoffinformationen stehen den Mitarbeitern von Bundes- und Länderbehörden, der wissenschaftlichen Fachwelt sowie der interessierten Öffentlichkeit über das Internet zur Verfügung. Zweimal jährlich wird der Datenbestand aktualisiert.

1.2 Verbesserung der Zusammenarbeit von Bund und Ländern

Die stetige Zunahme des Datenvolumens und des zu aktualisierenden Datenumfanges sowie die damit verbundenen erhöhten Anforderungen an den Qualitätssicherungsprozess erforderten eine Unterstützung des Produktionsprozesses durch zukunftsorientierte Software.

Ziel des vom BMU-finanzierten Vorhabens war es, eine web-basierte technische Anwendung zu konzipieren und zu entwickeln, die den Produktionsprozess im Umweltbundesamt umfassend unterstützt und als zentraler Anlaufpunkt für die Kooperationsgemeinschaft des GSBL zum Austausch von Daten, Software und Informationen fungiert und so zur Verbesserung der Zusammenarbeit aller Beteiligten beiträgt.

Wesentliche Anforderungen:

Zentralisieren der Produktionsvorgänge sowie der Dokumentanablage,

Standardisieren der Vorgehensweisen in den Produktionsvorgängen,

Transparenz des Produktionsprozesses für alle Beteiligten,

Bereitstellen virtueller Räume für Arbeits- und Projektgruppen,

Unterstützung der Dokumentation (GSBL-Handbuch).

Die Konzeption und Umsetzung der Kommunikationsplattform erfolgte unter Beteiligung von Partnern aus Bund und Ländern im Rahmen eines Begleitkreises. Dies erlaubte eine umfassende kritische Betrachtung der Zusammenarbeit und führte zu einer erheblichen Verbesserung der Bund-Länder-Kooperation. Die iterative Entwicklung wurde in regelmäßigen Workshops dem Begleitkreis des Vorhabens vorgestellt, diskutiert und ggf. modifiziert.

2 GSBL-Kommunikationsplattform

Die GSBL-Kommunikationsplattform ist als Web-Portal realisiert. Der Anwendung liegt SharePoint Foundation 2010 als Entwicklungs-Plattform zu Grunde.

Innerhalb der für die Kommunikationsplattform bereitgestellten Webapplikation wird die Website des GSBL Portals als Stammwebsite der Stammwebsitesammlung

abgebildet. Abbildung 1 gibt einen Überblick zur Einstiegsseite der Kommunikationsplattform.

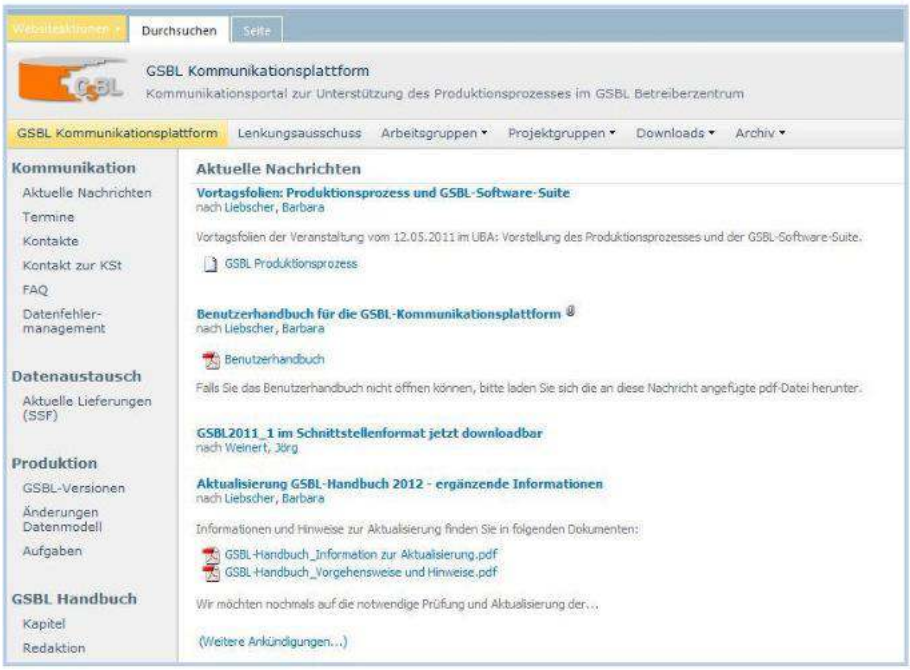


Abbildung 1: Kommunikationsplattform für die Prozesssteuerung des GSBL

Die Abbildung 2 zeigt die globale Sitemap, die den Gesamtaufbau der GSBL Kommunikationsplattform abbildet.

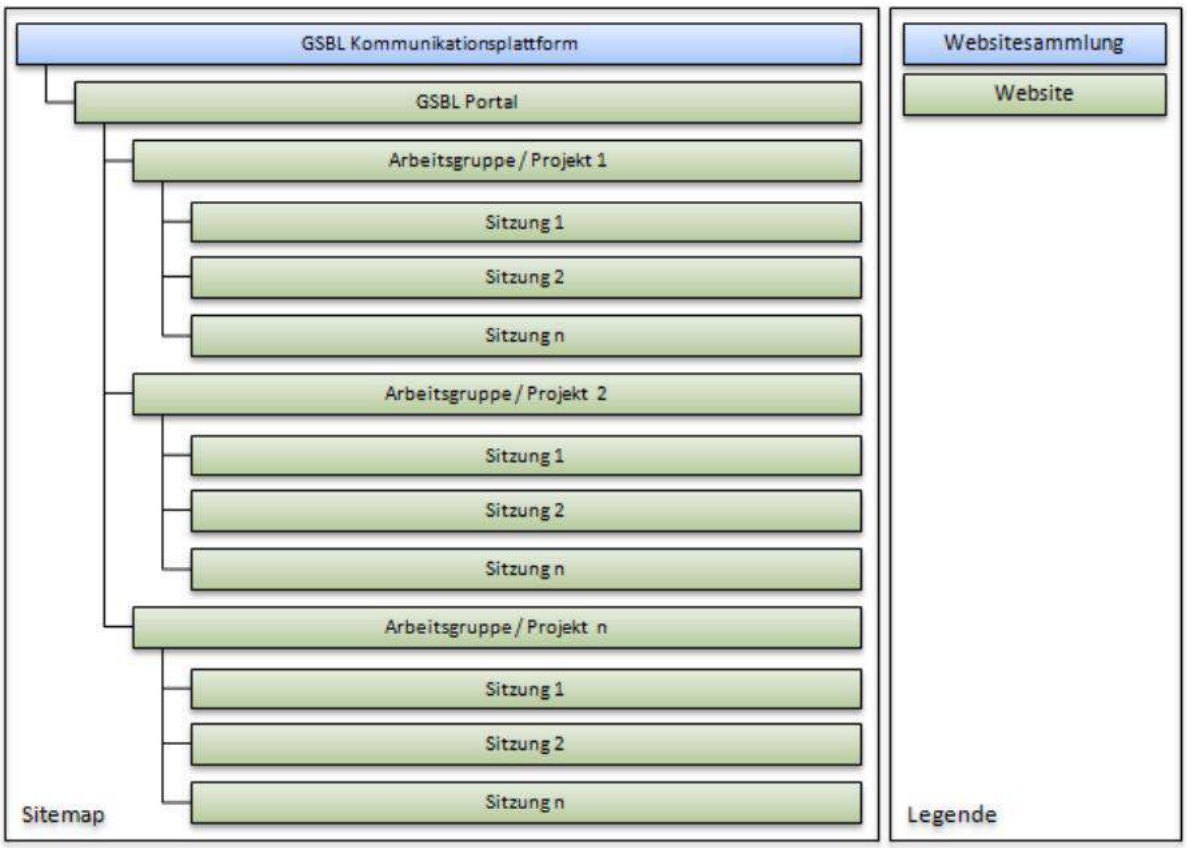


Abbildung 2: Globale Sitemap der GSBL-Kommunikationsplattform

Das GSBL Portal untergliedert sich in vier logische Themenbereiche, denen die jeweiligen Elemente wie Bibliotheken und Listen zugeordnet sind. Abbildung 3 zeigt die Struktur des GSBL Portals sowie dessen fachliche Kategorien und Elemente.

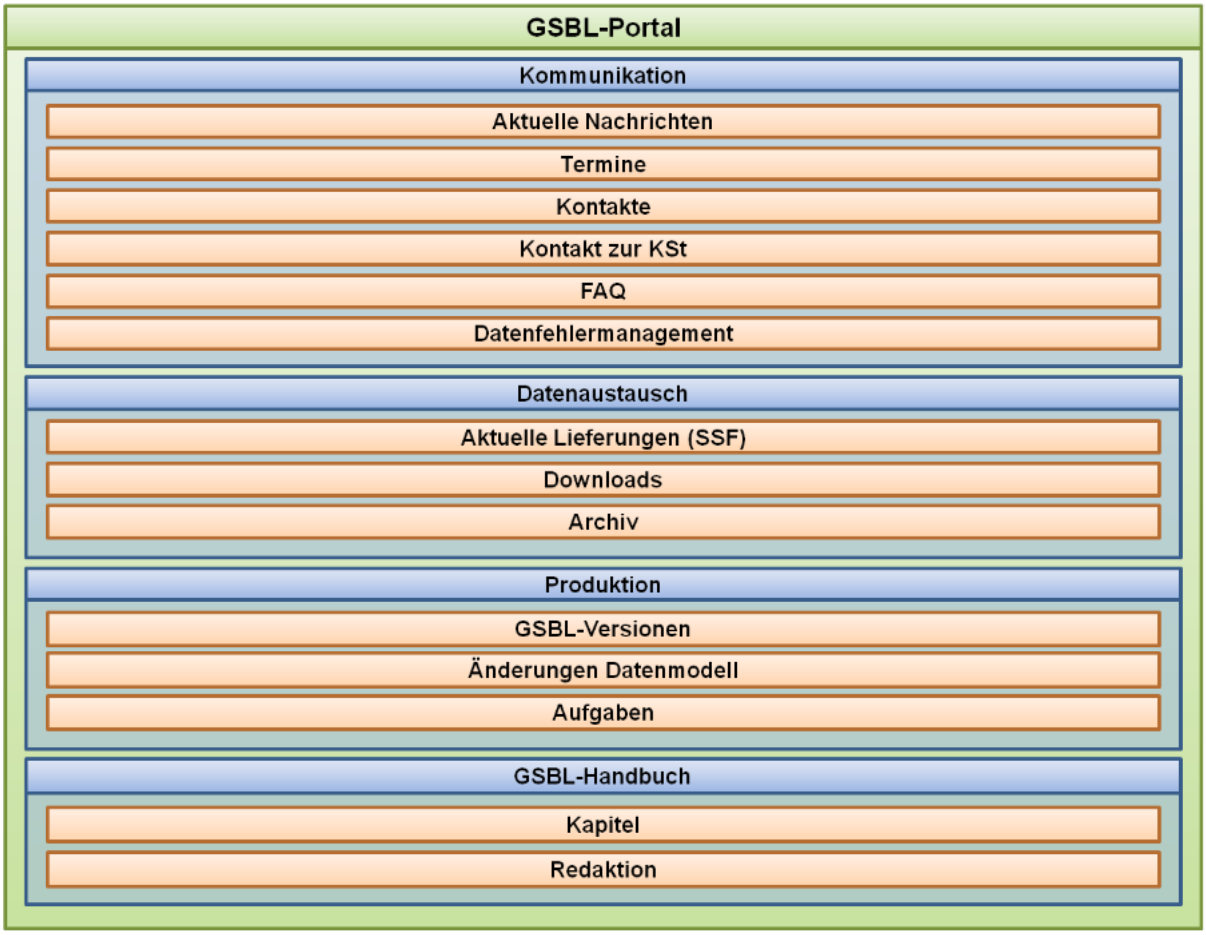


Abbildung 3: Sitemap des GSBL-Portals

Das Portal steht seit Dezember 2010 direkt unter https://www.gsblkoop.uba.de und als Link unter www.gsbl.de im Life-Betrieb bereit. Der personalisierte Zugriff kann sowohl aus dem Internet als auch dem Intranet des UBA erfolgen. Die Sicherheit wird durch Übertragungsverschlüsselung (http/SSL) sichergestellt.

Die Anwendung wird im Rechenzentrum des Umweltbundesamtes betrieben. Die Federführung für den Betrieb liegt in der Koordinierungsstelle des GSBL im UBA.

2.1 Funktionalitäten der Plattform

Das Portal stellt folgende Anwendungen bereit:

Produktionsunterstützung für das Betreiberzentrum und Nachvollziehbarkeit der Prozesse

Virtuelle Arbeitsräume zur Unterstützung der Arbeits- und Projektgruppen sowie des Lenkungsausschusses des GSBL für die Organisation von Sitzungen und

Diskussionsforum, dem Zugriff auf Kalender, Aufgabenlisten, Kontaktlisten und Dokumentenbibliotheken

Adressverwaltung und Berechtigungszuweisung

Verwalten von Software-Installationen und Updates sowie fachlicher Dokumente (Archiv)

Unterstützung zur Aktualisierung des fachlichen GSBL-Handbuchs mit der versionierten Ablage aller Kapitel und der Möglichkeit der Bearbeitung innerhalb des Portals, der Aktualisierung der Kapitelstruktur auf Grundlage des aktuellen Datenmodells, der Erzeugung eines Word- und pdf-Dokumentes sowie einer online-Hilfe.

Informationsangebot für aktuelle Nachrichten

E-Mail-Versand

anwendungsübergreifende Suchfunktionalität mit Volltextindizierung aller Datenbestände sowie Suche nach Metadaten.

Kern der integrierten Fachanwendungen ist die Produktionsunterstützung für das Betreiberzentrum. Diese Anwendung gewährleistet durch die integrierten Workflows eine vollständige Unterstützung des Produktionsprozesses einschließlich eines normierten Lieferprozesses mit der Erfassung von Begleitdaten (Abbildung 4).



Abbildung 4: Produktionsworkflow

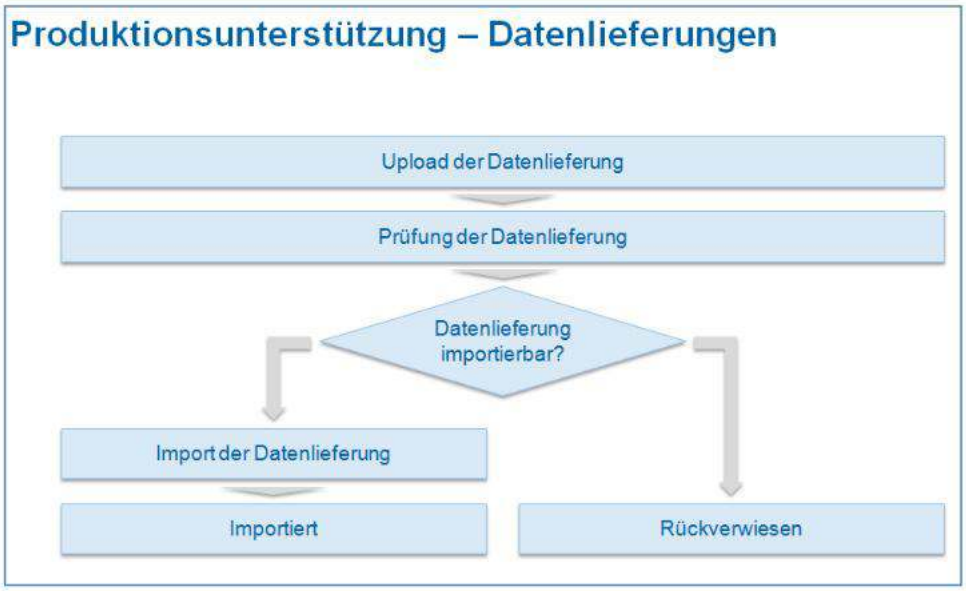


Abbildung 5: Importworkflow

Gesteuert wird der Produktionsprozess durch die Koordinierungsstelle, hier werden die einzelnen Prozessschritte gestartet und beendet. In die Ausführung der Prozessschritte selbst sind die Kooperationspartner aktiv eingebunden. Die von den Beteiligten bereitgestellten Dateien und Dokumente sowie die im Prozess erzeugten Dokumente stehen übersichtlich zur Einsicht und zum Download zur Verfügung (Abbildung 5).

Diese Arbeitsweise eröffnet den Partnern die Einsicht in den Prozess der Aktualisierung des GSBL und ermöglicht die Nachvollziehbarkeit der ablaufenden Prozesse (Abbildung 6).



Abbildung 6: Übersicht zu den GSBL-Versionen mit Angabe der Produktionsphase

Ein wesentlicher Bereich sind auch die virtuellen Arbeitsräume für Lenkungsausschuss, Arbeits- und Projektgruppen mit der Sitzungsorganisation, einem Diskussionsforum sowie Zugriff auf Aufgabenlisten, Kontaktlisten, Dokumentenbibliotheken und den zentralen Kalender.

Weitere Funktionalitäten sind die Verwaltung fachlicher Dokumente und Updates, die Bereitstellung aktueller Nachrichten, E-Mail-Versand und automatisierte Benachrichtigung, Adressverwaltung sowie Berechtigungszuweisung.

Ein eigens entwickelter Baustein zur Aktualisierung des fachlichen Handbuchs dient der Bearbeitung und versionierten Ablage aller Kapitel innerhalb des Portals (Abbildung 7) . Die Kapitelstruktur setzt jeweils auf das aktuelle Datenmodell auf und im Ergebnis kann per Knopfdruck ein Word- und PDF-Dokument sowie eine online-Hilfe erstellt werden.



Abbildung 7: Kapitel des fachlichen GSBL-Handbuches

Die anwendungsübergreifende Suchfunktionalität mit Volltextindizierung aller Datenbestände sowie die Suche nach Metadaten gewährleisten das schnelle Auffinden von Dokumenten und Informationen.

2.2 Was wurde erreicht

Mit der GSBL-Kommunikationsplattform steht den Kooperationspartnern des GSBL von Bund und Ländern ein wichtiges Instrument zur Verfügung, das den Produktionsprozess des GSBL umfassend unterstützt und die Kommunikation und Transparenz in der GSBL-Kooperation nachhaltig verbessert. Sie ist zentraler Anlaufpunkt und gewährleistet durch zahlreiche Fachanwendungen die Zusammenarbeit der Partnerinstitutionen. Gleichzeitig entlasten die zur Verfügung stehenden Funktionalitäten alle Partner von zeitaufwendigen behördeninternen Organisationsarbeiten und sichern die Qualität des Datenstromes und dessen Nachvollziehbarkeit.

Die zu Grunde liegende Software ermöglicht durch ihre Flexibilität die Anpassung und Erweiterung zu einzelnen Themenbereichen durch die Nutzer selbst.

Die entwickelte Web-basierte Kommunikationsplattform weist Modellcharakter auf. Der Portalaufbau, die Verknüpfung der SharePoint-Standardfunktionen und insbesondere die eigens entwickelten Fachanwendungen sind auch für den Einsatz in weiteren Bund-Länder-Projekten geeignet. Auf Basis dieser Plattform wird derzeit im Umweltbundesamt für das Informationssystem Qualitätssicherung Bund/Länder-Messprogramm Nord- und Ostsee ein vergleichbares Instrument entwickelt.

Umweltgerechtigkeit als neue ressortübergreifende sozialräumliche Strategie im Land Berlin

Heinz-Josef Klimeczek, Senatsverwaltung für Stadtentwicklung und Umwelt (Abt. IX – Umweltpolitik), Heinz-Josef.Klimeczek@SenStadtUm.Berlin.de

Abstract

There are hardly any studies in Germany which have examined the link between socio-economic status and health risks in the residential environment. The pilot project "Environmental justice in the State of Berlin" is the first study in Germany to have determined the link between social structure and environmental quality on a small spatial scale. This provides the substantive basis for the development of a new cross-cutting topic area at the interface of health, urban development and design, and the environment. The project focused on the health-related aspects of air pollution, noise, bio-climate, provision of green spaces, and urban development. The environmental data from these fields were related to "living environment areas" ("Lebensweltlich orientierte Räume", 447 planning areas). They represent the current spatial basis for planning, forecasting and monitoring demographic and socio-spatial developments in Berlin. As the last step, the spatially related findings were combined on a single (spatial) level of analysis. These initial investigations created the basis for an environmental urban monitoring system which complements the established monitoring systems by generating health-related information.

Einleitung

Gesundheitsverträgliche Umweltverhältnisse sind Voraussetzung für das Leben zukünftiger Generationen, dennoch findet in räumlichen und planerischen Zusammenhängen eine entsprechende Berücksichtigung der Umweltgüte nur am Rande statt. Vor allem mit Blick auf ressortübergreifende Ansätze im Rahmen integrativer städtebaulicher und stadtplanerischer Konzepte und Strategien sind neue

Beobachtungs- und Bewertungsgrundlagen zu entwickeln, mit denen die (stadträumliche) Verteilung von Umweltbelastungen praxistauglich untersucht werden kann. Vor diesem Hintergrund hat das Land Berlin in enger Kooperation mit mehreren Universitäten und Forschungseinrichtungen im Jahre 2008 ein ressortübergreifendes Modellvorhaben auf den Weg gebracht, um den Zusammenhang zwischen Sozialstruktur und Umweltgüte in der Hauptstadt zu ermitteln. Bundesweit erstmalig wurden für eine Gesamtstadt quartiersbezogene Umweltdaten ermittelt und mit Sozialdaten verschnitten.

1 Grundlagen / Themenfelder

Die raumbezogene Verschneidung der Umwelt- und Sozialdaten erfolgt auf Ebene der „Lebensweltlich orientierten Räume“ (LOR). Diese neue Raumhierarchie wurde 2006 durch Senatsbeschluss als neue Grundlage für das ressortübergreifende Planen und Verwaltungshandeln festgelegt. Mit den vorliegenden Untersuchungen wurden erstmalig Daten auf der untersten Ebene der LOR, der 447 Planungsräume, bereitgestellt. Im Vordergrund standen die 4 Kernthemen, die auf der Grundlage der Strategischen Umweltprüfung (SUP) ermittelt wurden:

Lärm,

Luftgüte,

Bioklimatische Belastung sowie

Grün- und Freiflächenversorgung.

Die Analysen und Aussagen zu diesen Themenfeldern wurden mit den kleinräumigen Aussagen zur sozialstrukturellen und sozialräumlichen Entwicklung des Monitorings Soziale Stadtentwicklung 2009 (siehe Abbildung 1) im Rahmen einer Basisberichterstattung zusammengeführt.

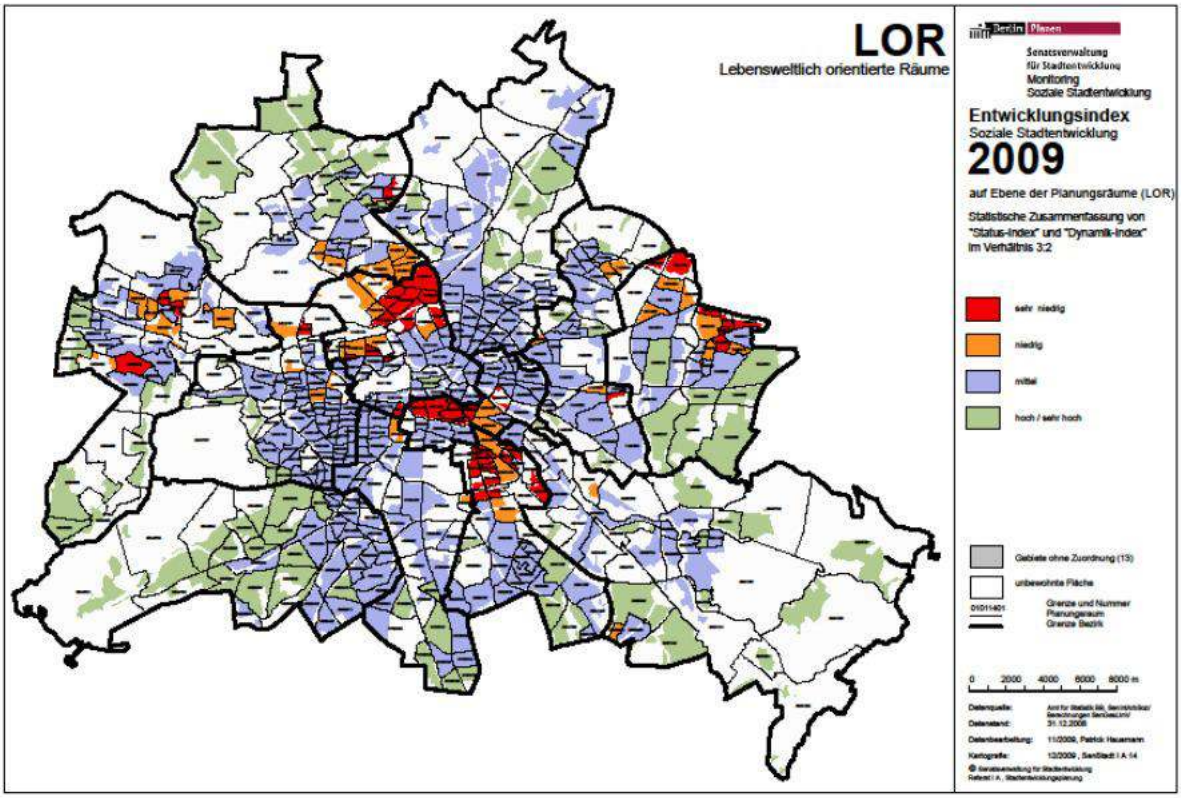


Abbildung 1: Entwicklungsindex Soziale Stadtentwicklung 2009

Weitere gesundheitsrelevante Bereiche, wie Gesundheitskosten, kleinräumige Aussagen zur gesundheitlichen Lage, Todesursachen, Stadtstruktur, Wohnlage, Lichtverschmutzung sowie die Klimaprojektionen (2046 – 2055) arrondieren diese ersten „orientierenden“ Untersuchungen. Darüber hinaus wurden erste Ansätze einer Betroffenenanalyse entwickelt.

2 Erste Ergebnisse

Die vorliegenden ersten Ergebnisse zeigen, dass sich im Bereich der hoch verdichteten Berliner Innenstadt der größte Teil der mehrfach belasteten Gebiete befindet. Die Untersuchungen belegen, dass ein Großteil der sozial benachteiligten Gebiete gleichzeitig auch von hohen gesundheitsrelevanten Umweltbelastungen betroffen ist. Die räumliche Verteilung dieser Umweltungerechtigkeit zeigt vor allem die Konzentration im Innenbereich Berlins. Hohe Mehrfachbelastungen durch Umweltfaktoren und das Zusammentreffen mit niedrigem sozialen Status zeigen sich vor allem im nördlichen Bereich des Bezirkes Friedrichshain-Kreuzberg, im Wedding, dem südlichen Bereich des Bezirks Reinickendorf und in Nordneukölln. (siehe

Abbildung 2) Eine sehr geringe / geringe Umweltbelastung und ein sehr hoher / hoher sozialer Status, finden sich dagegen vor allem im Außenbereich. Bereits jetzt ist erkennbar, dass die identifizierten mehrfach belasteten innerstädtischen Gebiete durch den Klimawandel künftig zusätzlich besonders betroffen sein werden, wodurch sich die Situation für die Bewohner in den Bezirken weiter verschlechtern wird. Die Ergebnisse ermöglichen erstmalig eine integrierte Betrachtung und damit eine neue Perspektive auf die Multidimensionalität der sozialräumlichen Verteilung von Umweltbelastungen und -ressourcen im städtischen Raum am Beispiel Berlins.

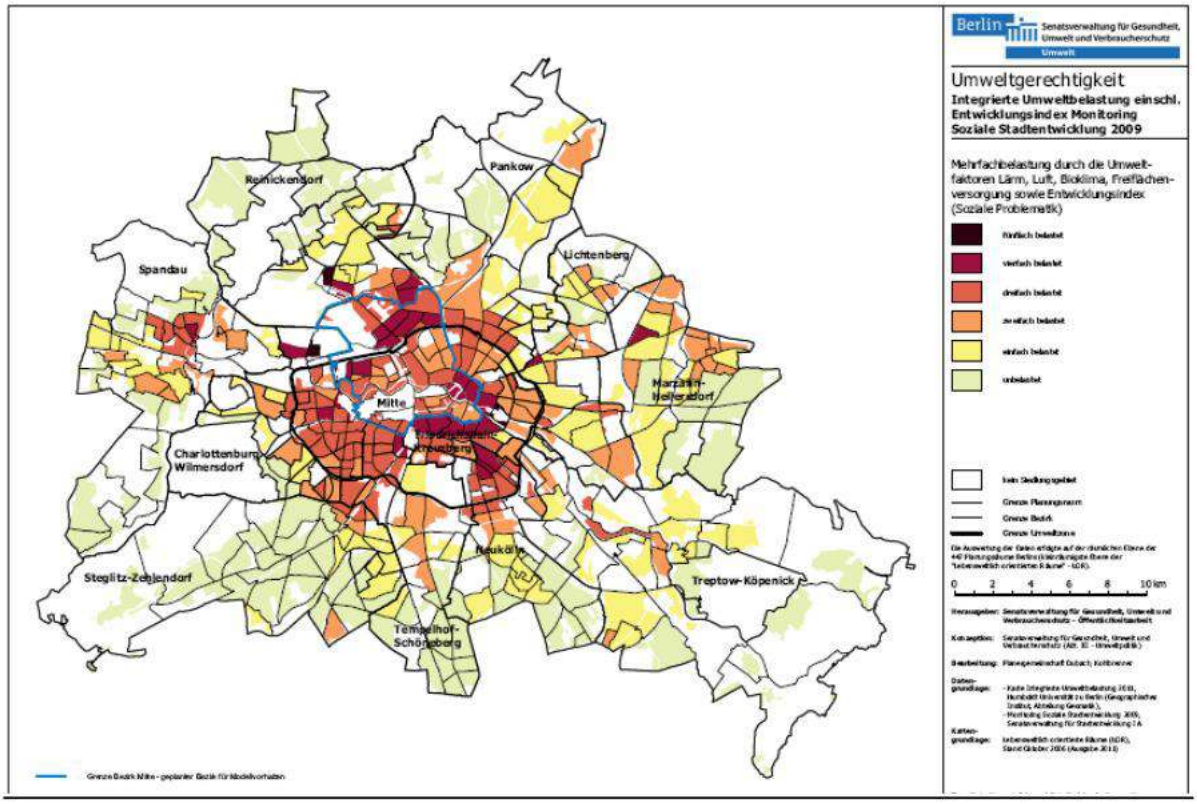


Abbildung 2: Integrierte Umweltbelastung einschl. Entwicklungsindex Monitoring soziale Stadtentwicklung 2009 (Bearbeitungsstand März 2009)

3 Umsetzung und weitere Entwicklung

Mit Blick auf die umwelt- und gesundheitspolitischen Zielsetzungen des Landes Berlin wird deutlich, dass vor allem integrierte Strategien und Konzepte der ministeriellen Politikfelder Gesundheit, Soziales, Stadtentwicklung, Städtebau und Umwelt stärker als bisher ressortübergreifend betrachtet und bearbeitet werden müssen. Hinsichtlich der weiteren Entwicklung ist es notwendig, die Bereiche im Sinne eines integrierten gesundheitsfördernden Gesamtkonzeptes besser miteinander zu

verzahnen. Nur so kann es gelingen, die Gesundheitsorientierung in den Bereichen Stadtentwicklung, Städtebau und Umweltplanung nachhaltig zu stärken. Mit Blick auf die künftige Stadtentwicklungs- und Umweltpolitik des Landes Berlin wird es eine wichtige Zielsetzung sein, diese umweltbezogenen Aussagen in das gesamtstädtische und bezirkliche Verwaltungshandeln strukturell und verfahrensorganisatorisch zu implementieren. Aufgrund neuer bzw. aktueller Daten werden derzeit die Ergebnisse zu den einzelnen Themenfeldern aktualisiert bzw. fortgeschrieben. In diesem Zusammenhang wird gleichzeitig die methodische Herangehensweise überprüft und - wissenschaftlich gestützt - weiterentwickelt. Darüber hinaus wird geprüft, wie die Aussagen zu einem Umweltbelastungsmonitoring weiterentwickelt und in bestehende Stadtbeobachtungssysteme integriert werden können.

4 Ausblick

Anfang 2011 wurde das Themenfeld „Umweltgerechtigkeit im städtischen Raum“ als Forschungsvorhaben in den Umweltforschungsplan des Bundesministeriums für Umwelt, Naturschutz und Reaktorsicherheit (BMU) aufgenommen. Hierdurch bekommen die Berliner Umweltgerechtigkeitsanalysen als ressortübergreifendes integriertes Gesamtkonzept zusätzliche (bundespolitische) Bedeutung. Im Vordergrund des Forschungsprojektes stehen vor allem die Erarbeitung von Instrumenten, Verfahren und Maßnahmen zur Herstellung von mehr (quartiersbezogener) Umweltgerechtigkeit. Darüber hinaus sollen Handlungsempfehlungen für das Verwaltungshandeln abgeleitet werden und die Implementierung eines umweltbezogenen sozialräumlichen Stadtbeobachtungssystems geprüft werden. Das Forschungsvorhaben (Laufzeit 2012 – 2014) wird vom Deutschen Institut für Urbanistik (Difu) durchgeführt.

Die ersten „orientierenden“ Ergebnisse sollen im Rahmen einer Landesberichterstattung zum Handlungsfeld „Umweltgerechtigkeit im Land Berlin“ im ersten Halbjahr 2013 veröffentlicht werden. Da es bundesweit bisher keinen vergleichbaren Ansatz auf sozialräumlicher bzw. quartiersbezogener Ebene gibt, wird das Land Berlin bei der Bearbeitung dieses Themenfeldes weiterhin eine Vorreiterrolle einnehmen.

5 Literaturverzeichnis

[Hornberg, et al., 2009]

Hornberg, C., Pauli, A.: Umweltgerechtigkeit – die soziale Verteilung von gesundheitsrelevanten Umweltbelastungen, Dokumentation der Fachtagung vom 27. bis 28. Oktober 2008 in Berlin, Universität Bielefeld

[Klimeczek, et al., 2008]

Klimeczek, H.-J.; Luck-Bertschat, G.: „(Sozial-)räumliche Verteilung von Umweltbelastungen im Land Berlin“ – Umweltgerechtigkeit als neues Themen- und Aufgabenfeld an der Schnittstelle von Umwelt, Gesundheit, Soziales und Stadtentwicklung. UMID: Umwelt und Mensch Informationsdienst, 2/2008, 26-29.

[Klimeczek, 2011]

Klimeczek, H.-J.: Umweltgerechtigkeit im Land Berlin – Entwicklung und Umsetzung einer neuen ressortübergreifenden Strategie. UMID: Umwelt und Mensch Informationsdienst, 2/2011, 19-21.

Verteilungswirkung von Verkehrslärm in Berlin - Methoden und Ergebnisse

Thilo Becker, Julia Friedemann, TU Dresden, Lehrstuhl für Verkehrsökologie,
thilo.becker@mailbox.tu-dresden.de

Abstract

Noise causes a variety of negative impacts on human wellbeing and health. The main impacts are considered to be annoyance due to noise and health impacts. Traffic noise emission levels may vary greatly on a very small scaled level due to differences in traffic flows, distances between source and receiver, as well as reflection and attenuation. This may lead to a rather unequal distribution of noise exposition among the inhabitants of an urban environment. In our analysis, we examine whether this unequal noise distribution reflects the social distribution of a city, leading to a situation where most deprived social groups are affected most by noise.

For this purpose, we use data from strategic noise maps in order to compute external noise costs per exposed person on the level of street sections. We thereby transfer noise exposition to a single monetary unit, allowing us to directly compare different noise levels and to quantify the economic damage associated with high noise exposition. In a next step, noise costs are overlaid with differentiated information about the social composition of neighborhoods. The indicators used are the share of inhabitants with migration background as well as the share of inhabitants receiving social welfare on the level of street sections. The results show that noise exposure differs significantly between the analyses socio-demographic groups. Noise costs are on average twice as high for the most advantaged groups compared to the least advantaged groups. The results clearly state, that noise mitigation measures are not only a matter of environmental protection but also of social planning and public health.

1 Einleitung

Der motorisierte Verkehr verursacht eine Vielzahl an Umweltwirkungen. Die Belastung für Anwohner durch Lärm zählt dabei zu den zentralen Problemen. So fühlen sich in Deutschland 11 % der Bevölkerung äußerst oder stark durch Straßenverkehrslärm belästigt, weitere 44 % fühlen sich mittel oder etwas gestört. [Umweltbundesamt, 2010] Wenig bekannt sind hingegen die gesundheitlichen Auswirkungen von Lärm, die bis zu vorzeitigen Todesfällen in Folge von chronischer Lärmbelästigung führen können.

Die Lärmemissionen, die an einer Straße gemessen werden, sind abhängig von der Anzahl, der Art und der Geschwindigkeit von Fahrzeugen, aber z.B. auch von der Oberflächenbeschaffenheit der Straße. Die Lärmwirkung auf den Menschen hängt darüber hinaus von Faktoren, wie dem Abstand zwischen Quell- und Empfängerort und der akustischen Dämpfung und Reflexion u.ä. ab. Alle diese Faktoren erklären, warum Lärmimmissionswerte bei den Empfängern bzw. Belasteten auf kleinräumiger Ebene hochgradig variieren. Aus diesem Grund ist es wichtig, die Lärmbelastung auf einer sehr lokalen Ebene zu bewerten.

Eine derartige kleinräumige Bewertung wird im Folgenden anhand einer Fallstudie von Berlin vorgestellt.[7] Dabei erfolgt zunächst eine monetäre Bewertung der in Berlin durch Lärm von Straße, Schiene und Luftverkehr verursachten Schäden. Das ermöglicht zwei verschiedene Analyseschwerpunkte:

Erstens können die oft unberücksichtigt wirtschaftlichen Auswirkungen von Lärm quantifiziert und die Notwendigkeit von Lärmschutzmaßnahmen besser nach außen kommuniziert werden.

Zweitens können bestimmte sozio-demografische Gruppen in Hinblick auf die Betroffenheit durch Lärmbelastung miteinander verglichen werden. Schwerpunkt dieses Beitrags ist die Klärung der Frage, ob die Lärmbelastung in Berlin zwischen verschiedenen sozio-demografischen Gruppen ungleich verteilt ist.

Beim zweiten Schwerpunkt werden Ansätze der Umweltgerechtigkeitsforschung auf das Themenfeld Verkehrslärm übertragen. Diese Art der Analyse kann helfen, die

[7] Die hier vorgestellte Analyse ist Teil des Dissertationsvorhabens von Thilo Becker. Die Ergebnisse wurden in ähnlicher Form auf Englisch auf der 26. Internationalen Konferenz für Umweltinformation ENVIROinfo 2012 in Dessau vorgestellt.

ungleiche soziale Verteilung der Umweltbelastungen aufzuzeigen. Dieser Schwerpunkt ist auch der Fokus der hier vorgestellten Studie.

2 Stand der Forschung

Umweltgerechtigkeit kann unter verschiedenen Aspekten betrachtet werden. Neben der Forderung nach einer gleichberechtigten Bürgerbeteiligung in umweltrelevanten Planungen ist die gleichmäßige Verteilung von Umweltbelastungen innerhalb einer Gesellschaft ein Kernthema. „Keine Gruppe von Menschen soll einen unverhältnismäßigen Anteil an negativen Umweltwirkungen tragen“ [eigene Übersetzung, nach EPA, 2012]. In den USA wird das Thema Umweltgerechtigkeit schon seit den 1980ern Jahren in Forschung und Politik debattiert. In der europäischen Forschung lässt sich erst in den letzten Jahren ein wachsendes Interesse an der Analyse der sozialen Verteilung von Umweltbelastungen feststellen [Fairburn, 2009]. In Deutschland hat sich die Forschung auf diesem Gebiet bisher hauptsächlich auf Themen der öffentlichen Gesundheit konzentriert [Helmert, 2000; Mielck, 2000]. Erst seit Kurzem haben Sozialwissenschaften, Planungsbehörden und Politiker den Zusammenhang zwischen Umweltplanung und Sozialplanung erkannt und das Thema aufgegriffen [Klimeczek, 2011].

Allerdings liegen für Deutschland nach wie vor wenige empirische Studien vor. Sichtbar ist trotzdem die zunehmende Methodenvielfalt im Forschungsgebiet. Frühe empirische Studien für Deutschland arbeiteten zumeist mit Sekundärdaten aus bundesweiten repräsentativen Umfragen. Sie untersuchten z.B. die Zusammenhänge zwischen der wahrgenommenen Betroffenheit durch Umweltbelastungen und den sozio-demografischen Merkmalen [Hoffmann, 2003; Mielck, 2004; Kohlhuber, 2006]. Diese werden mittlerweile durch Forschungsansätze auf Basis geographischer Informationssysteme (GIS) ergänzt. Hier werden objektive, modellierte Lärmbelastungsdaten mit räumlich kodierten Informationen zu sozio-demografischen Merkmalen der Bevölkerung verschnitten. Für Berlin und auch für Hamburg wurden entsprechende Pilotstudien durchgeführt [Gaffron, 2010; Lakes, 2011; Becker, 2011]. Alle bis auf eine [Lakes, 2011] der hier zitierten Studien fanden eine Beziehung zwischen sozio-demografischen Merkmalen und der wahrgenommenen Lärmbelastung, der Wohnungsnähe zu Hauptstraßen oder der „objektiv“ modellierten Lärmexposition.

Allerdings betonen die meisten Autoren die Notwendigkeit für weitere Forschung und den explorativen Charakter ihrer Studien [Gaffron, 2011].

Zusammenfassend lässt sich feststellen, dass die empirische Evidenz bezüglich der Korrelation zwischen sozio-demografischen Gruppen und Lärmbelastung nach wie vor gering ist. Es werden zunehmend modellierte Daten eingesetzt, aber bisher erfolgte die Analyse maximal auf Ebene vergleichsweise großer statistischer Bezirke. Speziell für die sehr kleinräumig wirkenden Lärmbelastungen könnten allerdings Analysen mit einer deutlich feineren räumlichen Auflösung notwendig sein, um Belastungsunterschiede zwischen verschiedenen sozio-demografischen Gruppen wirklich sichtbar zu machen. Die im Folgenden vorgestellte Studie versucht diese Schwäche zu beseitigen, indem Lärmbelastung und sozio-demografische Daten auf Straßenabschnittsebene betrachtet werden.

3 Methodik

3.1 Einführung in die Thematik der externen Kosten

Lärm ist genauso wie Unfälle oder die Belastung durch Luftschadstoffe eine der offensichtlichen negativen Begleiterscheinungen des Verkehrs. Lärm stört bei der Kommunikation und Erholung, führt zu Stress, Schlafstörungen bis hin zu Herz-Kreislauf-Erkrankungen. All diese Folgewirkungen können mit Kosten hinterlegt werden: entweder mit Kosten für die Schadwirkung durch Lärm an sich, aber auch für ihre Vermeidung bzw. Verringerung durch Lärmschutzmaßnahmen. Es handelt sich zum allergrößten Teil um externe Kosten. Diese Kosten werden nicht direkt von den Verkehrsteilnehmern, also den Schadensverursachern getragen und deshalb als „extern“ bezeichnet. Es gibt keinen Marktmechanismus zur Anlastung dieser Kosten. Stattdessen sind im marktwirtschaftlichen Sinne unbeteiligte Dritte (z.B. Anwohner, Passanten) von den Kosten betroffen, deren Verluste nicht kompensiert werden. [Friedemann, 2010]

Grundsätzliches Ziel bei der Berechnung externer Lärmkosten ist die monetäre Bewertung des Nutzenverlustes, den betroffene Individuen durch die vom Verkehr verursachten Lärm erleiden (Schadenskosten). [Umweltbundesamt, 2007] Dabei werden sowohl materielle Kosten z.B. für Krankenhausaufenthalte wegen lärmbe-

dingten Gesundheitsschäden als auch immaterielle Kosten z.B. für den Verlust an Lebensqualität durch das Fehlen der Ressource „Ruhe“ einbezogen [Friedemann, 2010].

3.2 Kostensätze für externe Kosten

Externe Kostenschätzungen, unter anderem die des Lärms, finden bereits seit Jahrzehnten Eingang in die Verkehrsplanung. Seit den 1990er Jahren sind eine Reihe umfangreicher Forschungsprojekte durchgeführt worden, in denen die Methodik zur Ermittlung der externen Kosten maßgeblich weiterentwickelt worden ist. Besonders auf europäischer Ebene fanden nennenswerte Aktivitäten statt: Zu nennen sind speziell die Projekte EXTERNE und UNITE, in denen zum einen methodische Grundlagen erarbeitet und zum anderen die Höhe der externen Kosten in den EU-Mitgliedsstaaten ermittelt wurden. Die Ergebnisse sind in [Maibach, 2008] zusammengefasst. Wie Abbildung 1 zeigt, werden für Lärm in der Regel zwei Hauptwirkungen betrachtet.

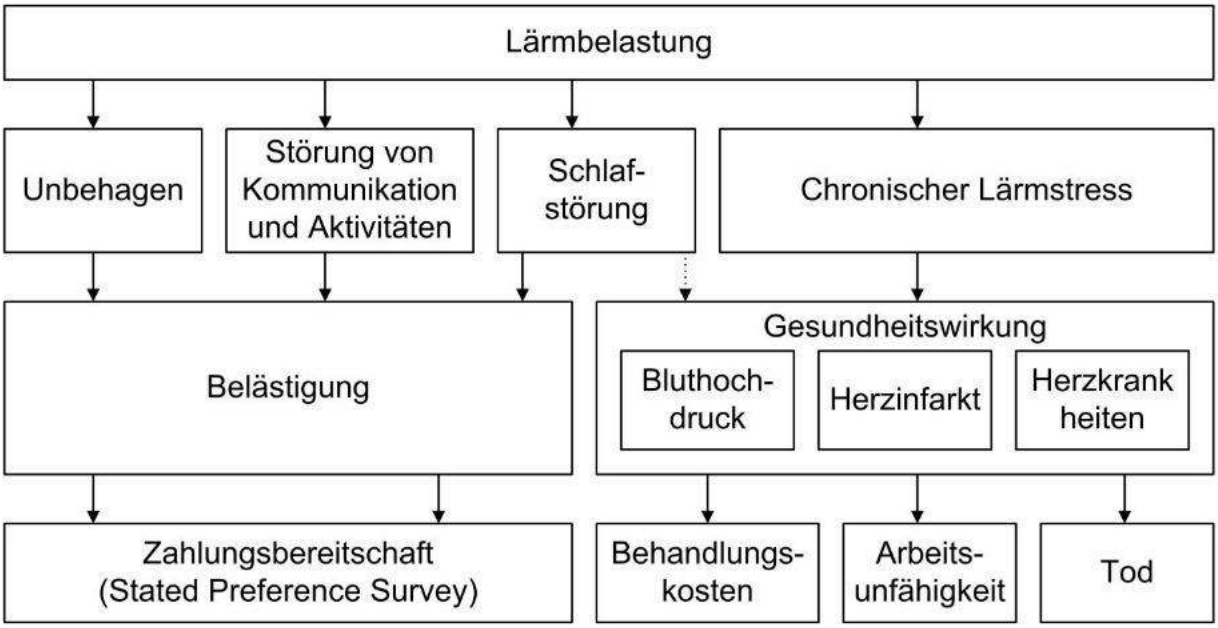


Abbildung 1: Darstellung der Lärmwirkungen und damit verbundenen externen Kosten. Quelle: eigene Darstellung, nach Babisch, 2006

Zum einen wird Lärm bei vielen Aktivitäten als störend empfunden. Der gefühlte Nutzenverlust für die Betroffenen kann in Befragungen erhoben werden. Konkret wird dabei die Zahlungsbereitschaft für eine gewisse Lärmreduzierung ermittelt. Für die

Schätzung externer Kosten gibt es für diesen Bereich mittlerweile konsolidierte Kostensätze, welche auch in dieser Studie verwendet werden [Maibach, 2008].

Neben den Kosten der Lärmbelästigung werden vor allem lärmbedingte Gesundheitswirkungen betrachtet. Grundlage der Kostensätze für Gesundheit sind Dosis-Wirkungs-Relationen, die den Zusammenhang zwischen dem Auftreten einer Krankheit und der Lärmbelastung am Wohnort beschreiben. Derartige Dosis-Wirkungs-Relationen werden in epidemiologischen Studien von Medizinern bestimmt [Babisch, 2006]. Die mit den Krankheitsfällen verbundenen Kosten wurden in dieser Studie dann mit Hilfe der Krankheitskostenstatistik [Destatis, 2010b] ermittelt. Berücksichtigt wurden dabei die Gesundheitswirkungen Herzinfarkt, Bluthochdruck und sonstige Herzerkrankungen [Babisch, 2006]. Tabelle 1 zeigt die Summe der externen Kosten durch Lärmbelästigung sowie Gesundheitswirkungen. Eine detaillierte Beschreibung der Methodik, mit der in dieser Analyse die externen Kosten berechnet wurden, findet sich in [Gerike, 2012; Becker, 2012].

L_{DEN} in dB(A)	Gesamte externe Lärmkosten in Euro (2008)
<55	0
55 - 60	76
61 - 65	315
66 - 70	639
71 - 75	1.004
76 - 80	1.431

Tabelle 1: Kostensätze externer Lärmkosten nach Pegelklasse in Euro/Betroffenen und Jahr. Quelle: eigene Berechnung.

3.3 Gebäudefeine Berechnung der externe Kosten

Daten zur Lärmbelastung konnten in dieser Studie der strategischen Lärmkarte von Berlin entnommen werden. [SenStadt, 2008] Derartige Lärmkarten müssen entsprechend der Umgebungslärmrichtlinie mittlerweile in allen Großstädten der Europäischen Union erstellt werden. [Europäisches Parlament, 2002]. Die Karten werden mithilfe von Lärmmodellen erstellt und dienen als Grundlage der städtischen Lärmminderungsplanung. Für die Modellierung werden eine Vielzahl von Eingangsparametern, wie die 3-D-Gebäudeform, der Straßenbelag und die Anzahl der täglichen Fahrzeuge auf einer Straße berücksichtigt. Modelliert werden zunächst die Lärmemissionen einer Schallquelle, z.B. eines Straßenabschnitts. Anschließend erfolgt

die Berechnung von Lärmimmissionen an den Häuserfassaden, d.h. bei den Betroffenen. Zur Berechnung der Lärmimmissionen werden dabei fiktive Immissionspunkte gleichmäßig auf der Fassadenaußenseite verteilt (siehe Abbildung 2). Sie befinden sich stets 4 m über dem Boden und ihre Anzahl variiert in Abhängigkeit von der Gebäudegröße [SenStadt, 2008]. In Berlin wurden 2008 Lärmimmissionen für insgesamt 3.910.391 Immissionspunkte berechnet. Die Ergebnisse können zwar offensichtlich nicht alle Spezialfälle wie Lärm in höheren Stockwerken oder innerhalb der Wohnräume berücksichtigen, aber es sind die besten momentan verfügbaren Daten und die Methodik entspricht dem gesetzlichen Standard.

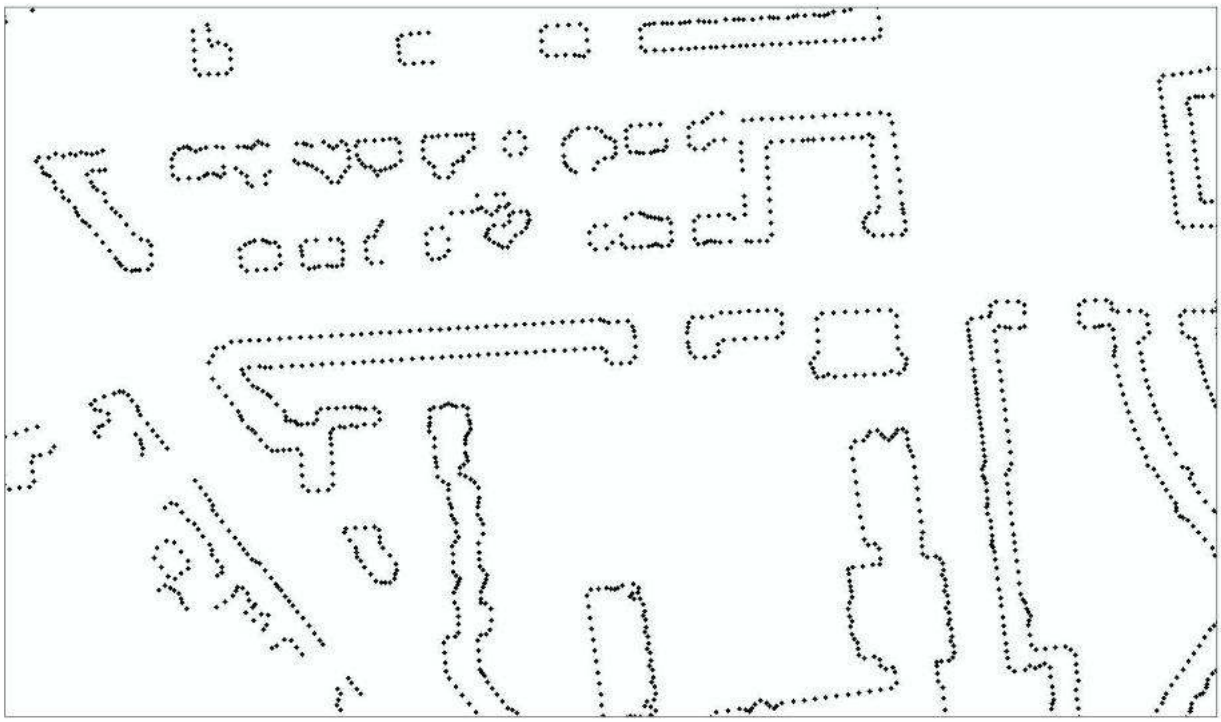


Abbildung 2: Beispiel für den Aufbau der Strategischen Lärmkarte mit Immissionspunkten an den Fassaden. Quelle: SenStadt, 2008

Jedem Immissionspunkt der strategischen Lärmkarte wurden im Modellierungsprozess unterschiedliche Attribute zugewiesen, darunter separate Lärmpegel für alle Verkehrsträger und ein Gesamtlärmpegel (energetisch addiert) sowie eine Gebäudekennung und die Anzahl der Bewohner je Immissionspunkt. Die Anzahl der Gebäudebewohner wurde dabei anhand der Einwohnerstatistik auf Ebene von Blöcken (Bezugsjahr 2005) ermittelt. Die Einwohner der statistischen Blöcke wurden dann anhand der Wohnfläche gleichmäßig auf die Gebäude verteilt. Innerhalb der Gebäude sind die Einwohner gleichmäßig auf alle Immissionspunkte verteilt worden.

Die Berechnungen dieser Studie basieren auf den angegebenen Gesamtverkehrslärmpegeln (Lärmindikator Tag-Abend-Nacht L_{DEN}) und der Anzahl der Bewohner je Immissionspunkt aus der strategischen Lärmkarte von 2008.

Die externen Kosten je Immissionspunkt wurden durch Multiplikation der Anzahl der Belasteten und den Kostensätzen in Tabelle 1 ermittelt. Danach sind die externen Kosten auf Ebene der Gebäude zusammengefasst worden. In einem letzten Schritt wurden die externen Kosten mit Hilfe der Gebäudekennung auf die Ebene von Straßenabschnitten aggregiert. Für diese Zuordnung wurde durch das Statistische Landesamt eine Datenbank zur Verfügung gestellt, welche eine Zuordnung der Gebäudekennungen zu Straßenabschnitten enthält. An Straßenkreuzungen orientiert sich die Gebäudezuteilung an der Adresse des Gebäudes [Statistik, 2011a]. Mithilfe dieser Datenbank konnte außerdem die Anzahl der Einwohner je Straßenabschnitt und somit auch die durchschnittlichen externen Lärmkosten je Einwohner für jeden einzelnen Straßenabschnitt berechnet werden. Die Aggregation auf Ebene der Straßenabschnitte war notwendig, da auch die verfügbaren sozio-demografischen Merkmale auf Ebene von Straßenabschnitten vorlagen.

3.4 Ermittlung der sozio-demografischen Merkmale

Sozio-demografische Merkmale von Bevölkerungsgruppen können durch eine Vielzahl verschiedener Indikatoren beschrieben werden. Oft werden in empirischen Studien Geschlecht, Alter, Staatsbürgerschaft, Bildungsniveau, Ausstattung mit technischen Geräten oder Angaben zum Haushaltseinkommen u.ä. verwendet [Destatis, 2010a]. Manchmal werden auch verschiedene einzelne Indikatoren zu einem aggregierten Gesamtindikator, z.B. der sozialen Schicht kombiniert. Für Berlin gibt es einen derartigen Indikator, in den unter anderen Informationen zur Arbeitslosigkeit einfließen. Dieser wurde bereits in früheren Lärmkostenstudien auf Ebene von statistischen Bezirken (in Berlin: Lebensweltlich Orientierte Räume) eingesetzt [Becker, 2011]. Allerdings sind derart breite Indikatoren nicht mit einer höheren räumlichen Auflösung verfügbar.

In der hier vorgestellten Analyse war vor allem die Beziehung zwischen dem verfügbaren Einkommen eines Haushalts und den externen Lärmkosten am Wohnort von Interesse. Es sollte untersucht werden, ob die Lärmbelastungen für finanziell bessergestellte Haushalte geringer sind, da sie sich eher eine Wohnung in ruhigeren,

oft teureren Wohngebieten leisten können. Zudem sollte der Einfluss der ethnischen Herkunft auf die Belastung durch Lärm ermittelt werden. Hier stellt sich die Frage, ob bestimmte ethnische Gruppen systematisch höheren Lärmkosten ausgesetzt sind als andere.

Die für diese Fragestellung benötigten sozio-demografischen Indikatoren Nettoäquivalenzeinkommen und ethnische Herkunft liegen für Berlin auf Straßenabschnittsebene nicht vor. Allerdings konnten in der Analyse zwei andere Variablen als (etwas unscharfe) Indikatoren für die genannten Eigenschaften genutzt werden. Verfügbar war zum einen der Anteil der Personen je Straßenabschnitt, der von SGB-II-Leistungen abhängig ist und der Anteil der Personen mit Migrationshintergrund.

Migrationshintergrund ist ein spezifischer Begriff der (Berliner) Bevölkerungsstatistik, der ausländische Staatsbürger, aber auch deutsche Staatsbürger der ersten oder zweiten Einwanderergeneration umfasst [Statistik, 2010]. SGB-II-Leistungsempfänger (umgangssprachlich „Hartz IV“) erhalten staatliche Unterstützung, die ein menschenwürdiges Existenzminimum gewährleisten soll.

Die hier verwendeten Indikatoren beschreiben den Anteil der Menschen mit Migrationshintergrund bzw. den Anteil der SGB-II-Leistungsempfänger an der Gesamtbevölkerung. Mit diesen Informationen ist es nicht möglich, einzelne Anwohner direkt einer sozio-demografischen Gruppe zuzuordnen. Dennoch können diese Informationen als Indikatoren für die zugrundeliegenden sozio-demografischen Gruppen dienen, da sie die Wahrscheinlichkeit ausdrücken, mit der einzelne Bewohner eines Straßenabschnitts zu einer der für uns relevanten sozio-demografischen Gruppe gehören.

Konkret verwendet wurden Daten zur Einwohnerzahl (2010), zum Anteil der Bewohner, der unter 65 Jahre alt und von SGB-II abhängig ist (Basisjahr 2009) und zum Anteil von Personen mit Migrationshintergrund (Basisjahr 2010) auf der Ebene von Straßenabschnitten. Die Anteile wurden vom Statistischen Landesamt zu 10 % Intervallen aggregiert. Für Straßenabschnitte mit weniger als 50 Einwohnern liegen aus Datenschutzgründen keine soziodemografischen Daten vor [Statistik, 2011b].

3.5 Analyse der Umweltgerechtigkeitssituation

Abbildung 3 zeigt die gewählte Analysemethode sowie die jeweils verwendete Datenbasis noch einmal grafisch. Aus den Informationen der strategischen Lärmkarte werden externe Lärmkosten basierend auf den Lärmimmissionen und der Einwohnerzahl berechnet. Anschließend wird der Zusammenhang zwischen der Betroffenheit durch Lärmkosten und den sozio-ökonomischen Merkmalen auf Straßenabschnittsebene analysiert. Wie in Abschnitt 3.1 bis 3.4 beschrieben, erfordert die gemeinsame Analyse von Daten zu Lärmpegeln, externen Kostenfaktoren und sozio-demografischen Daten die Nutzung verschiedener Datenquellen, welche auf Ebene der Straßenabschnitte zusammengeführt werden.

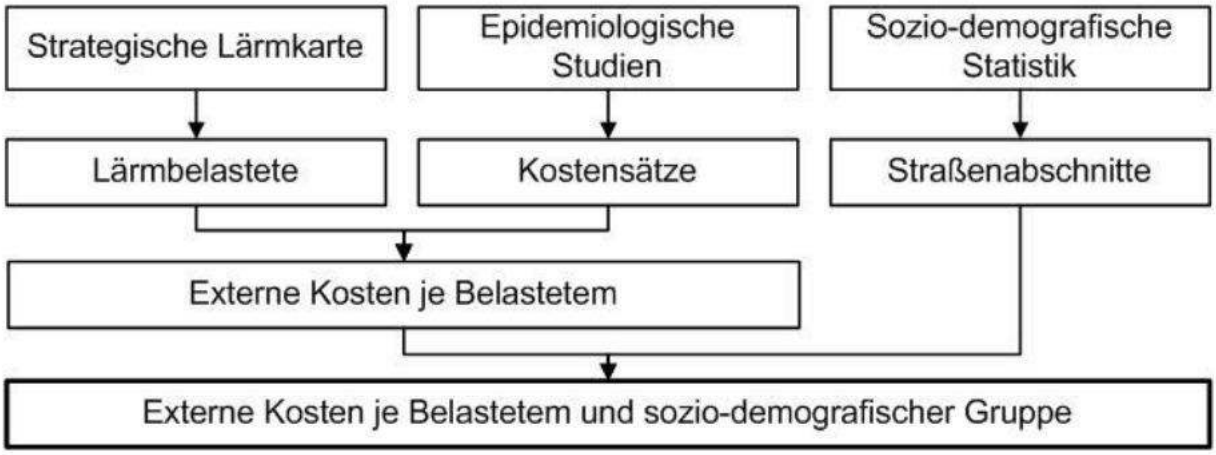


Abbildung 3: Methodischer Ansatz der Analyse und verwendete Datenquellen. Quelle: eigene Abbildung

Dabei wurden die externen Kosten je Straßenabschnitt direkt mit den sozio-demografischen Daten verknüpft. Natürlich konnten nicht alle Gebäude und deren Kosten zu Straßenabschnitten zugeordnet werden: für 0,68 % der Kosten war dies nicht möglich. Die Gründe dafür sind wahrscheinlich Änderungen der städtischen Infrastruktur über die Jahre. Weitere 12,31 % der Kosten (18,02 % der Einwohner, 57 % der Straßenabschnitte) konnten Straßenabschnitten zugewiesen werden, aber es stehen keine sozio-demografischen Daten für sie zur Verfügung. Ursache hierfür sind wahrscheinlich fehlende Daten für Straßenabschnitte mit weniger als 50 Einwohnern.

Einwohnerzahlen standen in dieser Analyse aus zwei verschiedenen Quellen zur Verfügung:

die ursprünglichen blockbasierten Daten aus der strategischen Lärmkarte

die Statistik der sozio-demografischen Merkmale für Straßenabschnitte.

Die Einwohnerzahl aus der strategischen Lärmkarte ist aufgrund der oben genannten Gründe etwa 20 % höher als die entsprechende Zahl aus der sozio-demografischen Statistik. Da die Lärmkosten anhand der Einwohnerzahl der strategischen Lärmkarte berechnet wurden, ist diese Datenquelle auch für die statistischen Berechnungen verwendet worden.

4 Ergebnisse

Nach der Statistik zur Grundsicherung wohnen 593.079 SGB-II-Leistungsempfänger in Berlin (2009) [Arbeitsagentur, 2012]. Dies entspricht einem Bevölkerungsanteil von 17,6 %. Die Anzahl der Menschen mit Migrationshintergrund beträgt 872.132 (2010) [Statistik, 2011c], was einem Bevölkerungsanteil von 25,7 % entspricht.

Tabelle 2 stellt (neben anderen Aspekten) die Verteilung der Bezieher von SGB-II-Leistungen auf Ebene von Straßenabschnitten dar. Die Tabelle zeigt, dass in einem gewissen Umfang Segregation zwischen den sozio-demografischen Gruppen existiert:

Die Hälfte der Bevölkerung lebt in Straßenabschnitten mit entweder einem unverhältnismäßig hohen oder niedrigem Anteil an SGB-II-Empfängern.

Beim Merkmal Migrationshintergrund leben sogar zwei Drittel der Bevölkerung in einem Straßenabschnitt mit unverhältnismäßig hohem oder niedrigem Anteil.

Anteil	Einwohner	Lärmkosten - arithmetisch	Lärmkosten - Median
< 10 %	481.891	80,44	18,06
10 bis < 20 %	789.271	110,51	48,41
20 bis < 30 %	618.887	136,20	64,50
30 bis < 40 %	434.361	147,55	75,56
40 bis < 50 %	243.483	154,75	83,20
50 bis < 60 %	119.465	154,50	79,85
> 60%	50.263	155,09	102,35
nicht zuordenbar	601.623	88,40	5,14

Tabelle 2: Durchschnittliche externe Kosten in Euro (2008) je Einwohner (arithmetischer Durchschnitt bzw. Median) nach Anteil der SGB-II-Leistungsempfänger auf Straßenabschnittsebene.
Quelle: eigene Berechnungen

Im nächsten Schritt wird die Verteilung der Lärmbelastung analysiert. Durchschnittliche externe Lärmkosten je Einwohner sind auf Ebene der Straßenabschnitte

berechnet worden. Die durchschnittlichen Lärmkosten pro Einwohner variieren zwischen 0 und 1.300 Euro. Das arithmetische Mittel für die Gesamtstadt liegt bei 117 Euro, der Median liegt bei 41 Euro. Abbildung 4 zeigt den Anteil der gesamten externen Lärmkosten, der kumuliert für Bevölkerungsanteile dargestellt ist. Insgesamt sind die Lärmkosten sehr ungleich verteilt. Ungefähr ein Viertel der Berliner Bevölkerung trägt keine Lärmkosten. Hier liegt die Lärmbelastung entweder unter 55 dB(A) oder ist nicht verfügbar. Auf der anderen Seite ist ein Viertel der Bevölkerung sehr stark durch Lärm belastet. Hier tragen 25 % der Bevölkerung etwa 75 % der Lärmkosten.

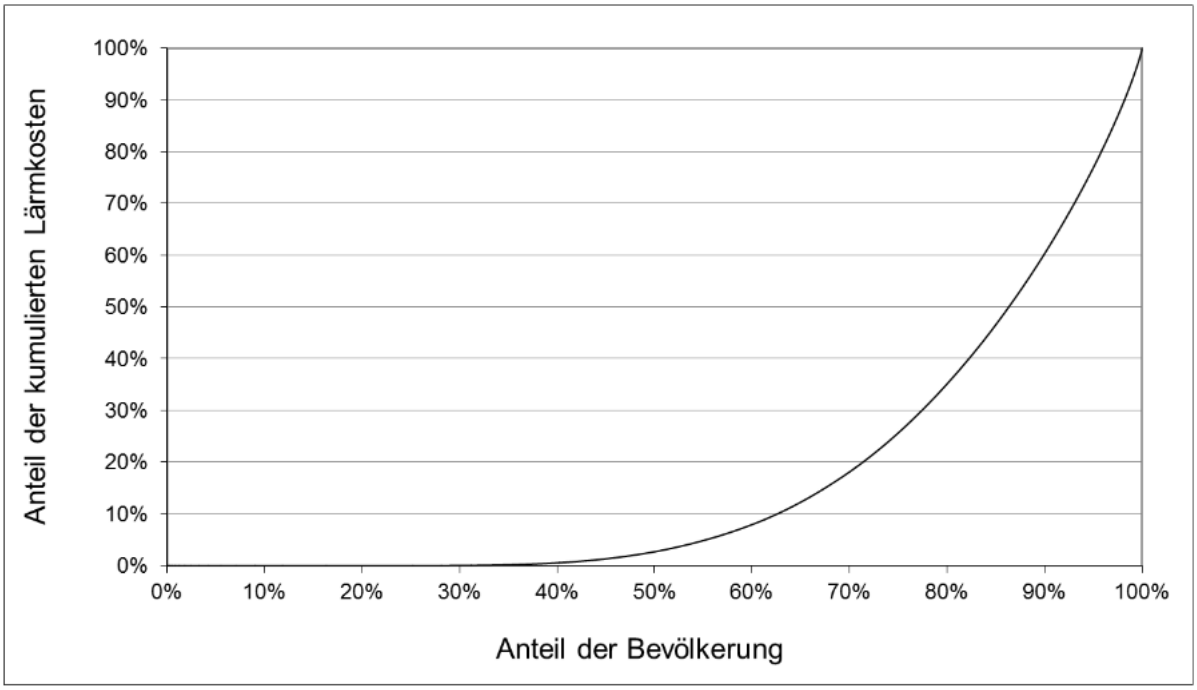


Abbildung 4: Kumulierte externe Kosten als Anteil an der Bevölkerung. Quelle: eigene Berechnungen

Bei der Analyse zeigte sich ein Zusammenhang zwischen der Betroffenheit durch hohe Lärmkosten und dem Anteil der Bewohner mit Migrationshintergrund bzw. dem Anteil der SGB-II-Leistungsempfänger. Es erfolgte ein Vergleich des arithmetischen Mittels und des Median der externen Kosten pro Einwohner unterteilt nach sozio-demografischen Gruppen. Wie in Abbildung 5 und Tabelle 2 erkennbar ist, unterscheiden sich beide Mittelwerte erheblich und sind höher für Straßenabschnitte mit einem hohen Anteil an Einwohnern mit Migrationshintergrund. Gleiches gilt für SGB-II-Leistungsempfänger. Für beide sozio-demografischen Merkmale gilt, dass sich die externen Kosten je Einwohner von der Gruppe mit den geringsten Anteilen zu der Gruppe mit den höchsten Anteilen verdoppeln.

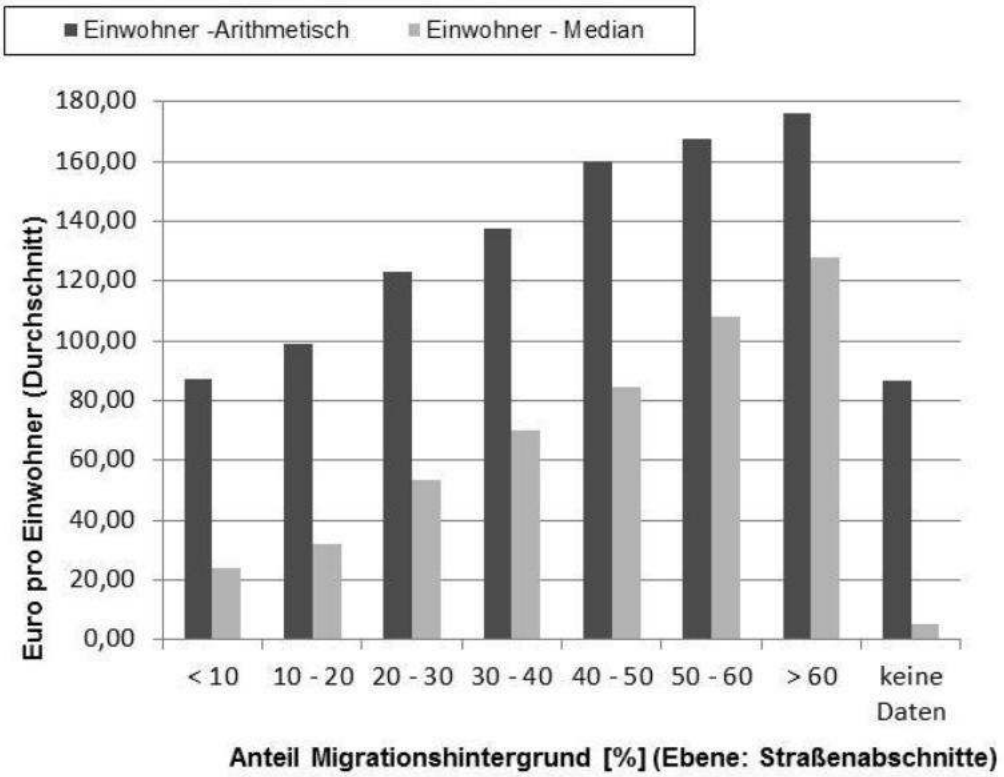


Abbildung 5: Externe Kosten je Einwohner in Euro (2008) nach Anteil der Einwohner mit Migrationshintergrund auf Straßenabschnittsebene.
Quelle: eigene Berechnungen

Diskussion

Die Ergebnisse der Analyse der Verteilung externer Lärmkosten auf sozio-demografische Gruppen wurden in Abschnitt 4 dargestellt. Es konnte gezeigt werden, dass die externen Lärmkosten in Berlin ungleich verteilt sind. Während ein Viertel der Bevölkerung gar nicht von Lärm betroffen ist, trägt ein weiteres Viertel etwa 75 % der Kosten. Außerdem wurde festgestellt, dass die Belastung durch Lärmkosten zwischen verschiedenen sozio-demografischen Gruppen stark variiert. Es gibt also einen Zusammenhang zwischen sozio-demografischen Variablen und den externen Lärmkosten pro Einwohner.

Wie bereits gezeigt, sind die durchschnittlichen Lärmkosten pro Einwohner in der Gruppe mit dem höchsten Anteil an SGB-II-Empfängern doppelt so hoch wie in der Gruppe mit dem geringsten Anteil. Ein ähnliches Ergebnis konnte beim Vergleich der sozialen Lärmkosten zwischen Gruppen mit hohem und niedrigem Anteil an Personen mit Migrationshintergrund beobachtet werden.

Diese Ergebnisse zeigen deutlich die Notwendigkeit, existierende Ungleichheiten in der Lärmbelastung zu thematisieren. Die beiden analysierten sozio-demografischen Indikatoren kennzeichnen Bevölkerungsgruppen mit einem begrenzten Freiraum bei der Wohnstandortwahl. Sie können hohen Lärmbelastungen vielfach nicht ausweichen. Somit sind Lärmminderungsmaßnahmen, hier speziell der an der Quelle ansetzende aktive Lärmschutz, nicht nur ein Element der Umweltplanung, sondern auch ein Mittel der Sozialplanung.

5 Schlussfolgerungen und Ausblick

Die präsentierte Analyse ist die erste ihrer Art für Berlin und auf Basis der Literaturanalyse auch in ganz Deutschland. Es konnten ungewöhnlich kleinräumige Daten auf der Ebene von Straßenabschnitten verwendet werden. Die Ergebnisse bestätigen frühere Studien, die einen schwachen Zusammenhang zwischen dem sozialem Status beziehungsweise der ethnischen Herkunft und Lärmbelastung nachweisen. In diesem Sinne ist die Studie ein wichtiger Beitrag zur Vertiefung des empirischen Wissens zum Thema Umweltgerechtigkeit im urbanen Raum. Die Verwendung der externen Lärmkosten als Indikator der Stärke der Lärmbelastung ist ein Novum im Bereich der Analyse zur Umweltgerechtigkeit. Eine monetäre Bewertung der Lärmbelastung ermöglicht zum einen eine Transformation der logarithmischen Lärmpegel auf einer arithmetischen Skala. Viel wichtiger ist aber, dass zum anderen auf diese Weise der mit der Lärmbelastung verbundene ökonomische Schaden verdeutlicht werden kann. Die Lärmbelastung der am meisten sozial benachteiligten Gruppen ist doppelt so hoch wie die der am wenigsten sozial benachteiligten Gruppen. Damit kann gezeigt werden, dass Maßnahmen zur Lärmminderung nicht nur ein Instrument des Umweltschutzes, sondern auch der Sozialplanung sind.

Um die auf Durchschnittswerten basierenden Ergebnisse zu bestätigen und zu konkretisieren, sind weitere Analysen der Datenbasis unter Zuhilfenahme statistischer Modelle geplant. Gleichzeitig soll die Datenlage bezüglich der sozio-demografischen Informationen weiter verbessert werden. So sollen die derzeit verwendeten Indikatoren zur Bestimmung der Finanzkraft und der ethnischen Herkunft durch treffsichere Variablen ersetzt werden. Zum anderen wurden weitere wichtige Kenngrößen, wie zum Beispiel Bildung oder die Altersverteilung (vgl. Gaffron 2012,

Mitchell 2003) bisher nicht in der Analyse berücksichtigt. Diese Defizite werden Gegenstand weiterer Forschung sein. Um alle Einwohner gleichermaßen vor übermäßiger Lärmbelastung zu schützen, müssen außerdem die genauen Ursachen für den aufgedeckten Zusammenhang zwischen sozialer Benachteiligung und Lärmbelastung untersucht werden.

Danksagung

Die Studie ist Teil des Dissertationsvorhabens von Thilo Becker und entstand in enger Kooperation mit dem Modellvorhaben „Umweltgerechtigkeit im Land Berlin“ (siehe Beitrag von Heinz-Josef Klimeczek, Senatsverwaltung für Gesundheit, Umwelt und Verbraucherschutz, Seite 51 bis 56). Wir bedanken uns bei Heinz-Josef Klimeczek, Horst Diekmann (Lärm) sowie Hartmut Bömermann (Sozialstatistik) und ihren MitarbeiterInnen für die Bereitstellung der benötigten Datenbasis und die intensiven Diskussionen zum methodischen Vorgehen. Finanziell unterstützt wurde die Arbeit durch Stipendien der Heinrich Böll Stiftung.

6 Literaturverzeichnis

[Arbeitsagentur, 2012]

Bundesagentur für Arbeit: Statistik der Grundsicherung für Arbeitssuchende nach dem SGB II. Zeitreihe Eckwerte SGB II nach Ländern. Nürnberg. Online verfügbar:

http://statistik.arbeitsagentur.de/Navigation/Statistik/Statistik-nach-Themen/Grundsicherung-fuer-Arbeitsuchende-SGBII/Bedarfsgemeinschaften-und-deren-Mitglieder/Bedarfsgemeinschaften-und-deren-Mitgliederhtml, letzter Abruf: 10.07.2012

[Babisch, W.]

Babisch, Wolfgang: Transportation Noise and Cardiovascular Risks. Review and Synthesis of Epidemiological Studies. Dose-effect Curve and Risk Estimation. Umweltbundesamt. Berlin, 2006.

[Becker, T.]

Becker, Thilo: Social distribution of external costs of noise impacts caused by transportation in Berlin. Proc. 2011 Internoise, Dokument Nr. 431268 der Proceedings DVD-ROM.

[Becker, T.]

Becker, Thilo.: Bewertung externer Effekte. Bewertung von Lärm. Dresden. Unveröffentlichter Entwurf, 2012.

[Destatis, 2010a]

Statistisches Bundesamt: Statistik und Wissenschaft – Demographische Standards. Band 17. Ausgabe 2010. Wiesbaden, 2010. Online verfügbar: https://www.destatis.de/DE/Methoden/DemografischeRegionaleStandards/Standards.html, Abruf: 03.06.2012

[Destatis, 2010b]

Statistisches Bundesamt: Gesundheit. Krankheitskosten 2002, 2004, 2006 und 2008. Fachserie 12, Reihe 7.2. Wiesbaden, 2010.

Europäisches Parlament / Europäischer Rat, Richtlinie über die Bewertung und Bekämpfung von Umgebungslärm, 2002.

EPA - Environmental Protection Agency: Definition Environmental Justice. Washington 2012. Online verfügbar: http://www.epa.gov/environmentaljustice/basics/index.html, Abruf: 03.06.2012

[Fairburn, J.]

Fairburn, J.: Environmental justice activities across Europe. proc. Fachtagung Umweltgerechtigkeit - die soziale Verteilung von gesundheitsrelevanten Umweltbelastungen, 72-74. Bielefeld, 2009.

[Friedemann, J., et al.]

Friedemann, Julia; Becker, Thilo; Becker, Udo: Wegekosten und externe Kosten - Analyse, Probleme, Bedeutung In: Handbuch der kommunalen Verkehrsplanung, Ordner 4; Kapitel 4.5.1; 58. Ergänzungslieferung. 2010.

[Gaffron, P.]

Gaffron, Philine: Urban transport, environmental justice and human daily activity patterns. Dokument-ID: 01479, Proceedings der 12. World Conference of Transport, Lissabon, Juli 2010.

[Gaffron, P.]

Gaffron, Philine: Umweltgerechtigkeit im Stadtverkehr - Status quo der Empirie in Deutschland und weiterführende Analysen UMID - II. Themenheft Umweltgerechtigkeit (2011) 81–88.

[Gaffron, P.]

Gaffron, Philine: Urban transport, environmental justice and human daily activity patterns. Transport Policy (2012), doi:10.1016/j.tranpol.2012.01.011

[Gerike, R., et al.]

Gerike, Regine et al.: Quantification and mapping external noise costs back to transport users - development of an integrated urban modelling approach. Proceedings 2012 Euronoise.

[Häußermann, H., et al.]

Häußermann, Hartmut et al.: Monitoring Soziale Stadtentwicklung Berlin 2009, Fortschreibung für den Zeitraum 2007 - 2008. Senatsverwaltung für Stadtentwicklung Berlin, 2010.

[Helmert, U. (ed.)]

Helmert, Uwe (ed.): Müssen Arme früher sterben? Soziale Ungleichheit und Gesundheit in Deutschland. Juventa-Verlag, Weinheim, 2000.

[Hoffmann, B., et al.]

Hoffmann, Britt; Robra, Bernt-Peter; Swart, Enno: Soziale Ungleichheit und Straßenlärm im Wohnumfeld - eine Auswertung des Bundesgesundheitssurveys. Gesundheitswesen 65 (2003) 393–401.

[Klimeczek, H.-J.]

Klimeczek, Heinz-Josef: Umweltgerechtigkeit im Land Berlin - Entwicklung und Umsetzung einer neuen ressortübergreifenden Strategie UMID - II. Themenheft Umweltgerechtigkeit (2011) 19–20.

[Köckler, H.]

Köckler, Heike et al.: Umweltbezogene Gerechtigkeit und Immissionsbelastungen am Beispiel der Stadt Kassel. CESR-Paper 1, 2008.

[Kohlhuber, M., et al.]

Kohlhuber, Martina; Mielck, Andreas; Weiland, Stephan K.; Bolte, Gabriele: Social inequality in perceived environmental exposure in relation to housing conditions in Germany. - Environmental Research 101 (2006) 246–255.

[Lakes, T., et al.]

Lakes, Tobia; Brückner, Maria: Sozialräumliche Verteilung der Lärmbelastung in Berlin. – UMID - II. Themenheft Umweltgerechtigkeit (2011) 26–28.

[Maibach, M., et al.]

Maibach, Markus et al.: Handbook on estimation of external costs in the transport sector. CE Delft, Delft, 2008.

[Mielck, A. (Hrsg.)]

Mielck, Andreas (Hrsg.): Soziale Ungleichheit und Gesundheit - empirische Ergebnisse, Erklärungsansätze, Interventionsmöglichkeiten. Huber, Bern, 2000.

[Mielck, A.]

Mielck, Andreas: Unterschiede bei Lärmbelastung und Luftverschmutzung nach dem Haushaltseinkommen. - In: Umweltgerechtigkeit. Die Soziale Verteilung von Umweltbelastungen. Gabriele Bolte, Andreas Mielck (Hrsg.). Juventa, Weinheim, 2004.

[Mitchell, G., et al.]

Mitchell, Gordon; Dorling, Danny: An environmental justice analysis of British air quality. Environment and Planning A 35 (2003) 909–929.

[SenStadt]

Senatsverwaltung für Stadtentwicklung: Umweltatlas Berlin. 07.05 Strategische Lärmkarte (Ausgabe 2008). Online verfügbar: http://www.stadtentwicklung.berlin.de/umwelt/umweltatlas/ei705.htm

Verwendeter Datensatz für das Projekt.

[Statistik, 2010]

Amt für Statistik Berlin-Brandenburg: Statistischer Bericht: Melderechtlich registrierte Einwohner im Land Berlin am 31. Dezember 2010. Online verfügbar: http://www.statistik-berlin-brandenburg.de/Publikationen/Stat_Berichte/2011/SB_A1-5_hj02-10_BE.xls, Abruf: 03.06.2012

[Statistik, 2011a]

Amt für Statistik Berlin-Brandenburg: Adressdaten für das Forschungsprojekt "Verteilungswirkung externer Kosten des Verkehrs in Berlin" zur Verfügung gestellt. Vertrag vom 26.04.2011.

[Statistik, 2011b]

Amt für Statistik Berlin-Brandenburg: Sozialdaten für das Forschungsprojekt "Verteilungswirkung externer Kosten des Verkehrs in Berlin" zur Verfügung gestellt. Vertrag vom 09.02.2011.

[Statistik, 2011c]

Amt für Statistik Berlin-Brandenburg: Statistischer Bericht A I 5 - hj 2 / 10: Melderechtlich registrierte Einwohner im Land Berlin am 31. Dezember 2010. Potsdam, 2011. Online verfügbar:

http://www.statistik-berlin-brandenburg.de/Publikationen/Stat_Berichte/2011/SB_A1-5_hj02-0_BE.xls, Abruf: 03.06.2012

[Umweltbundesamt]

Ökonomische Bewertung von Umweltschäden - Methodenkonvention zur Schätzung externer Umweltkosten. Dessau 2007. Online verfügbar: http://www.umweltdaten.de/publikationen/fpdf-l/3193.pdf, Abruf: 03.06.2012

[Umweltbundesamt]

Umweltbewusstsein in Deutschland 2010, Ergebnisse einer repräsentativen Bevölkerungsumfrage, Berlin, 2010.

Aufbau und Verwendung von Fachterminologie

Thomas Bandholtz, innoQ Deutschland GmbH, thomas.bandholtz@innoq.com
Joachim Fock, Umweltbundesamt, joachim.fock@uba.de

Abstract

Domain terminology is a network of hopefully shared conceptualisations in the area of discourse, designated by hopefully un-ambiguous terms and labels or - in terms of data representation - identifiers. In most cases those terms and identifiers are not managed on the enterprise level which causes misunderstandings and mismatches in all areas of communication. This article drafts a roadmap of evolving and exploiting domain terminologies based on experiences of the authors and further thoughts about the next steps to go.

Einleitung

Immer mehr Organisationen haben die Bedeutung einer zwischen den Beteiligten abgestimmten und zugänglichen Fachterminologie erkannt und setzen dies bereits ein oder bereiten einen solchen Einsatz vor. Der vorliegende Beitrag basiert auf den Erfahrungen aus einschlägigen eigenen Projekten seit 1998, geht aber teilweise über das dort schon Realisierte hinaus und entwirft ein aus heutiger Sicht umfassendes Szenario zur Aufgabenlösung. Dabei wirken intellektuelle Leistungen des Teams und die Nutzung eines „Automaten“ eng zusammen. Nicht alle beschriebenen Funktionen des Automaten sind in den bisherigen Projekten schon realisiert worden, die meisten von ihnen sind aber an anderer Stelle implementiert und können grundsätzlich integriert oder erweitert werden. In der Begrifflichkeit halten wir uns an [DIN 2342:2011-08] und verweisen an einigen Stellen auf den internationalen W3C Standard „Simple Knowledge Organisation System“ (SKOS)[8], der vor allem im Semantic Web verbreitet ist.

[8] http://www.w3.org/2004/02/skos/

1 Aufbau und Pflege von Vokabularen

Hier wird ein semi-automatisches Verfahren beschrieben, das bei Null ansetzt. Wenn in Ihrer Organisation bereits ein ausgebauter Thesaurus oder ähnliches vorliegt, kann dieses Vorgehen dazu dienen, dessen Qualität zu überprüfen und ihn systematisch fortzuschreiben – Sprache steht nicht still, Vokabulare erfordern deshalb fortgesetzte Pflege. Andernfalls entwickeln sich begriffliche Lücken. Der GEMET z.B. stammt aus dem Jahr 2000 und kennt weder den Emissionshandel noch die Feinstaubplakette, um nur zwei Beispiele zu nennen.

1.1 Vorliegende Ansätze prüfen

Welche Ansätze zu Terminologie gibt es bereits? Gibt es Glossare, Schlagwort- oder Schlüsselisten, Themenkataloge oder dergleichen?

Gibt es „Branchenstandards“ wie z.B. die Financial Industry Business Ontology (FIBO) in der Finanzwelt?

Sofern hier etwas vorliegt, werden aus der Gesamtmenge aller dort beschriebenen Begriffe die ca. 100 wichtigsten ausgewählt. Diese Begrenzung ist ein guter Trick, um die kontroverse Auseinandersetzung mit dem Wortgut zu provozieren, die das Team im weiteren Verlauf vor allem vertiefen muss. Die Ergebnisse auch der besten Automaten müssen immer intellektuell validiert und korrigiert werden.

1.2 Analyse eines repräsentativen Dokumentenkorpus

Hierzu werden ca. 100 Dokumente benötigt, die insgesamt die Organisation möglichst vollständig beschreiben. Die Dokumente werden Organisationseinheiten (Abteilungen oder Fachbereiche) zugeordnet. Nun generiert der Automat (z.B. „maschinengestützte Termextraktion und -Validierung“ [Prusseit, et al., 2010]) ein Ranking aller vorkommenenden Begriffe (für jede Organisationseinheit und insgesamt) und vergleicht dieses Ranking mit den zuvor identifizierten 100 wichtigsten Begriffen.

Beide Listen werden nun zusammengeführt und vom Team begutachtet.

1.3 Begriffe und ihre Benennungen

Dabei wird auffallen, dass zwischen Begriffen (skos:Concept) und Benennungen (Wörtern, rdfs:label in SKOS) unterschieden werden muss. Manche Wörter

verweisen auf denselben Begriff (z.B. „Müll“ und „Abfall“). Wenn man sie austauschbar benutzen kann, wird einer von ihnen als „Vorzugsbenennung“ ("skos:prefLabel") ausgewählt und die anderen werden zu Synonymen ("skos:altLabel") desselben Begriffs. Der Begriff ist eigentlich eine Abstraktion, die unabhängig von seinen (austauschbaren) Benennungen existiert. Wie genau man Begriffe unterscheiden will, ist Sache des Anwendungsfalls. „Müll“ unterschlägt irgendwie den Recycling-Aspekt von „Abfall“, aber lohnt sich die erhebliche Mühe solcher feinen Unterscheidungen im gegebenen Anwendungsfeld, ist sie mit den gegebenen Resourcen überhaupt zu leisten?

1.4 Hierarchie

Welche Begriffe sind die allgemeinsten Oberbegriffe, die "Top Concepts" dieser Schar? Gemeinsam sollen sie das gesamte Thema abdecken und hinsichtlich ihrer semantischen Ähnlichkeit möglichst gleich weit voneinander entfernt sein. Dies ist primär intellektuelle Leistung des Teams. Implementierbare Methoden zur Berechnung semantischer Ähnlichkeit (z.B. [Arweiler, 2008]) basieren auf gegebenen Pfadlängen zwischen den Begriffen, und diese liegen ja noch nicht vor.

Es ist nicht einfach zu entscheiden, wieviele es davon geben soll, und die Diskussion wird auch zu Vorschlägen neuer Begriffe führen. Vor allem wird sie nie enden, wenn man sie nicht willkürlich beendet, z.B. durch einen zeitnahen Stichtag. Später können jederzeit noch Änderungen vorgenommen werden, denn entscheidend ist lediglich die Nachhaltigkeit (Persistenz) der Begriffe, nicht so sehr die ihrer Beziehungen untereinander, wenn diese die Bedeutung des Begriffs nicht wesentlich verändern.

Aus den insgesamt vorliegenden Begriffen werden nun 2 bis 3 Hierarchie-Ebenen gebildet. Wenn die Top Concepts jetzt die Wurzeln von sehr unterschiedlich großen Bäumen sind, sollte geprüft werden, ob dies den Realitäten gerecht wird. Ausgeglichenheit der Teilbäume ist nicht das alleinige Qualitätskriterium. Wichtiger ist eine realitätsgerechte Abdeckung der Teilbereiche.

Das Team muss entscheiden, ob mehrere Hierarchien gebildet werden, klassisch wird unterschieden zwischen generisch, partitiv und instantiv. Instantiv im Sinne von „Micki ist eine Katze“, partitiv im Sinne von „ein Auto besteht aus Karosserie, Motor,

Getriebe usw.“, generisch im Sinne von Kategorien und Unterkategorien wie z.B. in einem Produktkatalog.

1.5 Verwandtschaft

Eine weitere klassische Beziehung ist die nicht näher spezifizierte Verwandtschaft („related Concept“). Der W3C Standard SKOS sagt dazu deutlich, dass Hierarchie und Verwandtschaft disjunkt sind. Die Hierarchie ist ja bereits selbst eine Art von Verwandtschaft, und deswegen geht es hier um alle nicht-hierarchischen Verwandtschaften wie z.B. „Rauchen erzeugt Lungenkrebs“ oder „Katze frisst Maus“. Wieder muss das Team entscheiden, ob solche Beziehungen weiter typisiert werden, also ob zwischen „erzeugt“ und „frisst“ unterschieden werden soll, und wenn ja, welche Beziehungen es außerdem geben soll. Solche Verfeinerungen sind zweifellos semantisch reizvoll, aber auch sehr aufwändig. Für die näher liegenden Anwendungsfälle (siehe nachfolgene Kapitel) werden sie nicht benötigt.

1.6 Befüllen

Alle in den Dokumenten gefundenen Begriffe, die bisher nicht für die Terminologie ausgewählt wurden, sind nun Kandidaten für deren Erweiterung. Man kann diese Liste zunächst verkleinern, z.B. anhand des erreichten Rankings oder auch manuell. Der Automat kann nun für jeden Kandidaten die Begriffe der Terminologie ermitteln, die am häufigsten mit ihm gemeinsam (im selben Dokument, Absatz, Satz) ermittelt wurden. Daraus ergeben sich Vorschläge für den Einbau des Begriffs in Hierarchie oder Verwandtschaft. Der Automat kann zumindest ermitteln, welche Knoten für den Einbau geeignet scheinen, aber wahrscheinlich nicht, ob es sich um einen Unter- oder Oberbegriff oder um eine bestimmte Art von Verwandtschaft handelt. Hier beginnt das Revier der „künstlichen Intelligenz“, das nach wie vor sehr akademisch und experimentell besetzt scheint. (Widersprüche sind willkommen.)

1.7 Definitionen

Im manchen Fällen sind die Begriffe klar definiert, wie z.B. in der Verdingungsordnung für Bauleistungen oder für die Abstufung der Paketgebühren. Dies hängt jedoch sehr von der jeweiligen Perspektive ab.

Genaue Definitionen der Begriffe sind nicht für alle Anwendungsfälle nötig, aber sie können beliebig viel Zeit verschlingen und das Vorhaben auch zermürben. Es ist

hilfreich, vorliegende Definitionen zu zitieren, aber man kann und muss keine Einstimmigkeit erzwingen. Es kann lediglich die Vielfalt der Fachsprache wiedergegeben, geordnet und bei (unserem) Anspruch der Kommunikation mit der Öffentlichkeit, in die Laiensprache/Alltagssprache/Pressesprache, also so etwas wie eine populärwissenschaftliche Vemittlung geschaffen werden.

In Sozialwissenschaften und vor allem Philosophie wird sogar jede Terminologiekontrolle abgelehnt und die Bedeutung eines Ausdruckes nur aus dem Kontext abgeleitet. Das ist wohl das andere Extrem gegenüber den Ideen des frühen Wittgenstein.

1.8 Schreibweisen und Zusammensetzungen

In Texten kommen die Wörter nicht immer so vor wie sie im Lexikon stehen. Zum Einen gibt es die bekannten Flexionsformen, zum Anderen die Zusammensetzungen. Während die Flexionsformen der meisten (90.000) Wörter im Deutschen Morphologie-Lexikon als Creative Commons vorliegen, bereiten die Zusammensetzungen größere Mühe, da z.B. „Abfallaufbereitungsanlagentechnik“ auf vielfältige Weise zerlegt werden kann. Das Open Source Werkzeug WordSplitter (Apache 2.0) kann deutsche Komposita mehr oder weniger vollständig zerlegen. Die Ergebnisse müssen zweifellos nachbearbeitet werden, denn nicht jede Zerlegung, wie zum Beispiel die Zerlegung von „Wasser-stoff“, ist hier sinnvoll.

1.9 Formalisieren

Für die weitere Anwendung wird dieses Vokabular in einer formalen, maschinenlesbaren Sprache benötigt. Dies ist SKOS, basierend auf RDF. Die Entwicklung einer komplexeren Ontologie ist nicht Bestandteil des hier beschriebenen Vorgehens, kann aber dadurch angestoßen werden.

1.10 Bekannt und verfügbar machen

Das Vokabular muss in der Organisation einfach zugänglich sein, und zwar sowohl für lesende Menschen als HTML Webseite als auch für maschinelle Anwendungen in RDF. Tools wie z.B. iQvoc leisten dies auf einfachste Weise.

2 Nutzung in der Kommunikation

Der offensichtliche, aber schwer messbare Nutzen einer abgestimmten und allen Beteiligten bewussten Terminologie liegt in einem besseren Verständnis bei jeglicher Kommunikation. Alle können sich auf die Begriffe der vereinbarten Terminologie beziehen und vermeiden damit Missverständnisse. Aus Sicht der Informatik müssen hierbei vor allem „softe“ Methoden helfen, die zu einer hinreichenden Bekanntheit und Akzeptanz der Terminologie führen.

Der Automat muss zunächst dafür sorgen, dass alle Begriffe mit ihren Definitionen und Beziehungen leicht zugänglich sind (im allgemeinen Web oder im Intranet) und dass sie eindeutig referenziert werden können (durch ihre Webadresse).

Ein bekannter und nützlicher Dienst ist der Abruf von „ähnlichen Begriffen“ bei der Suche oder beim Editieren von Texten.

Im Fall von Mehrsprachigkeit können spezialisierte Programme Übersetzungsvorschläge generieren [Massion, 2010].

Nach [Zerfaß, 2010] gehört dazu auch, zuvor ausdrücklich „verbotene Benennungen beim oder nach dem Schreiben eines Textes [zu] erkennen“. Ein praktisches Beispiel für faktisch verbotene Benennungen in Umfeld der deutschen Umweltbehörden ist z.B. „Waldsterben“ anstelle von „Waldschaden“. Eine entsprechende Kennzeichnung im UMTHES existiert allerdings bisher nicht und ebensowenig existiert eine entsprechene maschinelle Erkennung.

3 Nutzung im Wissensmanagement

Konventionelle Suchverfahren basieren auf einem Volltextindex und einem Zeichenkettenvergleich zwischen dem Suchwort und den indexierten Wörtern. Selbst wenn dabei schreibfehlertolerant vorgegangen wird (z.B. Levenshtein Verfahren oder „Kölner Phonetik“), bleibt die entscheidende Einschränkung bestehen: es werden immer noch Zeichenketten (engl. labels) verglichen und nicht Begriffe (engl. concepts).

Wenn ich also z.B. nach „Müll“ suche, werde ich keine Dokumente finden, in denen durchgängig das Wort „Abfall“ verwendet wird. In bestimmten Fällen mag genau dies gewollt sein, aber weitaus häufiger möchte man nicht zwischen begrifflich sehr ähnlichen Wörtern unterscheiden, sondern „alles zum Thema“ finden.

Eine einfache Art der Verbesserung besteht in der Hinterlegung von Synonymlisten, sodass der Anwender automatisch (oder auf Wunsch) mit allen hinterlegten Synonymen suchen kann. Noch immer aber werden Zeichenketten verglichen, es werden lediglich Zeichenketten in der Suchbedingung ergänzt.

3.1 Automatisches Indexieren

Die Thesaurus-basierte Erschließung geht hier einen wesentlichen Schritt weiter: der Thesaurus kennt Vorzugsbenennungen (skos:prefLabel) und Synonyme (skos:altLabel) und kann dies gleich bei der Indexierung einbringen. Das Dokument wird ausschließlich mit den Begriffen indexiert, auf die die gefundenen Benennungen verweisen. Wenn das gleiche Verfahren vor Anwendung der Suchwörter eingesetzt wird, werden automatisch immer Begriffe mit Begriffen verglichen, nicht möglicherweise abweichende Benennungen.

Ein professioneller Thesaurus geht aber weiter: er kennt auch die verschiedenen Schreibweisen (Beugungsformen, alte/neue Rechtschreibung, typische Falschschreibungen) jedes Worts und findet so z.B. auch „Bäume“ wenn nach „Baum“ gesucht wurde. Das gleiche gilt für Zusammensetzungen: die berühmte „Donaudampfschifffahrtskapitänsmütze“ kann sinnvoll auch auf unterschiedliche, aber nicht unbedingt jede Weise zerlegt geschrieben werden. Die Indexierung (und die Suche) müssen sich also auch mit dem gemeinsamen Vorkommen mehrerer Einzelbegriffe im selben Satz befassen, also prüfen, ob zwei oder mehr getrennte Wörter desselben Satzes gemeinsam ein zusammengesetztes Wort bedeuten können. Diese Prüfung kann nur die Wahrscheinlichkeit bewerten, so lange die Grammatik ignoriert wird.

So weit so gut – doch sehr viele Zeichenketten sind nicht eindeutig. Meint „Mast“ nun einen Schiffsmast oder ein Verfahren der Fleischproduktion, meint „Weide“ den Baum oder eine Wiese? Die Beispiele werden unzählig, wenn man Ortsnamen mit betrachtet: keine der fünf deutschen Gemeinden mit dem Namen „Berg“ liegt auf einem Berg. Zunächst kennt ein professioneller Thesaurus solche „Homonyme“ (hier genauer: Homographe) und erlaubt es, dieselbe Zeichenkette für unterschiedliche Begriffe zu verwenden. Jeder Begriff wird mit seinem Kontext vernetzt (Ober-/Unterbegriffe, verwandte Begriffe) und so kann aufgrund des im Satz/Abschnitt/Dokument dominanten Kontexts meist zutreffend entschieden werden.

Schließlich gehört zur semantischen Erschließung eine Entscheidung, welches die signifikantesten Begriffe eines Dokuments sind. Dies müssen nicht die am häufigsten verwendeten sein. Auch hier muss der Kontext berücksichtig werden, also die am stärksten vertretene „Wolke“ von untereinander vernetzten Begriffen, wobei auch zwischen Haupt- und Nebenthemen unterschieden werden kann.

3.2 Suchen und Vergleichen

iQvoc bietet mit autoClassify einen Service, der genau diese Art von Verschlagwortung implementiert und sowohl bei der Indexierung als auch bei der Vorbehandlung von Suchwörtern eingesetzt werden kann. So können auch ganze Textblöcke, News-Meldungen oder Dokumente als Suchbedingung verwendet werden – stets werden Dokumente mit differenziert ähnlicher Thematik gefunden.

3.3 Push Dienste

Es ist nicht immer notwendig oder wünschenswert, aktiv nach Dokumenten zu suchen. Jeder Anwender kann seine persönlichen Schlüsselbegriffe subskribieren und wird dann vom System benachrichtigt, sobald ein neues Dokument auftaucht, das diese Begriffe verwendet. Dies ist ähnlich wie Google Alerts, aber es basiert auf semantischer Erschließung der Nachrichten und nicht auf bloßem Zeichenketten-Vergleich. Beziehungen innerhalb der Terminologie wie z.B. die Abhängigkeit einer Aktie von bestimmten Rohstoffen können dabei auf vielfältige Weise vom Automaten berücksichtigt werden.

3.4 Themenspektrum analysieren

Während der automatischen Indexierung wird festgehalten, welche Begriffe als signifikant ermittelt wurden. So lässt sich zu jedem Zeitpunkt die Häufigkeitsverteilung einsehen. Zum einen sagt dies etwas aus über die Thematik der im betrachteten Zeitraum aufgetauchten Dokumente, zum anderen aber auch über die Brauchbarkeit des Wortguts, wobei selten gefundene Begriffe aber deshalb nicht weniger wichtig sein müssen. Wenn ein Begriff aber nie gefunden wird, ist daran aber sicher irgendetwas verkehrt.

Wie schon unter „1.6 Befüllen“ beschrieben, kann auch die Nähe der gefundenen Begriffe untereinander und zu bisher nicht aufgenommenen Wörtern weiter verfolgt und auf dieser Basis die Terminologie optimiert und erweitert werden.

4 Nutzung mit verlinkten Daten

Bisher liegen Wertevorräte oder Schlüssellisten für einzelne Attribute der verschiedenen Datenbestände verstreut vor und wer einen Datensatz liest, kann in der Regel nicht direkt auf dieses Wortgut zugreifen. Die Idee ist hier, alle Schlüssellisten im Web zu veröffentlichen, wobei jeder Begriff eine eindeutige Webadresse bekommt. Wenn in den Daten nun Links auf diese Webadresse verwendet werden, kann sofort nachgeschlagen werden, was der Begriff meint und welche Alternativen es gegeben hätte.

4.1 Linked Data

In den letzten Jahren ist ein stetig wachsendes Netz aus Linked Data entstanden, das Informationen, die von unterschiedlichen Organisationen gepflegt werden, zu einem globalen Netzwerk verknüpft. Das Semantic Web, das selbst aufgeklärte Zeitgenossen bisher für ein eher akademisches Thema hielten, wird damit Realität. Teilnehmer am Linked Data-Netz sind z.B. BBC, New York Times, Wikipedia, MusicBrainz, die Deutsche Nationalbibliothek sowie zahlreiche wissenschaftliche Einrichtungen und Regierungsorganisationen. Aus der Umweltinformatik sind bereits EUNIS, GEMET, EARTh und UMTHES beteiligt.

Gemäß Tim Berners-Lee, Erfinder des Web und Chef des World Wide Web Consortium (W3C), liegen dem Linked Data-Konzept vier einfache Prinzipien zugrunde:

1. Verwende URIs als Namen für Dinge.
2. Verwende HTTP URIs, damit die Leute diese Namen nachschlagen können.
3. Wenn jemand einen Namen nachschlägt, biete nützliche Information und verwende dabei die Standards (RDF, SPARQL).
4. Biete von dort auch Links auf andere URIs, damit wir mehr Dinge entdecken können.

Damit können Menschen und Maschinen gleichermaßen die so vernetzte Information durchlaufen und systematisch durchsuchen. Gut gebaute Webanwendungen haben dies bisher für Menschen auch schon weitgehend ermöglicht. Aber nun kommen mit RDF und SPARQL Standards ins Spiel, die Maschinenlesbarkeit, schemagestützte Struktur und attributspezifische Suche ermöglichen.

4.2 Aus der Datenbank ins Linked Data-Netz

Ein einfacher Weg ins Linked Data-Netz geht direkt von Ihrer (relationalen) Datenbank aus. Mit der RDB-to-RDF Mapping Language R2RML kann beschrieben werden, wie die bestehenden Tabellen, Zeilen und Spalten in RDF abgebildet werden sollen. Open Source-Werkzeuge wie z.B. D2R Server übernehmen den Rest: Ihre Daten werden in RDF verfügbar und mit SPARQL durchsuchbar.

4.3 Die integrierte Webanwendung

Ein eleganterer Weg ist eine Webanwendung, die eine menschenlesbare HTML Darstellung der "Dinge" mit Linked Data verbindet. Dies setzt voraus, dass die Webanwendung wie ein semantisches Netz strukturiert ist. Dann ist nur noch die Unterstützung des Ausgabeformats nötig, das die Anforderung des Clients berücksichtigt (Content Negotiation). Eine solche Webanwendung hat innoQ z.B. für das Umweltbundesamt entwickelt: die Umweltprobenbank.

4.4 RDF Vokabulare

Die Verständlichkeit der eigenen Daten im globalen Linked Data Netz wird stark vereinfacht, wenn für das Schema bekannte RDF Vokabulare verwendet werden. Dies sind z.B. SKOS, Dublin Core, Data Cubes (vormals SCOVO), Geonames oder Darwin Core. Nicht immer ist dort alles genau so definiert, wie man es selbst getan hätte. Die Erweiterbarkeit von RDF durch Sub-Klassen und Sub-Properties bietet

gute Möglichkeiten, eigene Spezialisierungen erkennbar an die bekannten Vokabulare anzuhängen.

4.5 Das Unternehmens-Vokabular

Jede Organisation spricht eine spezifische Sprache. Zum Teil liegt dies an der spezifischen Thematik, die Wortwahl ist aber auch Ausdruck der Unternehmenskultur. Die Systematisierung eines solchen Vokabulars ist im ersten Abschnitt dieses Beitrages beschrieben. Am Ende liegt ein RDF Schema und ein Unternehmensthesaurus vor, den Sie z.B. mit der Open Source-Software iQvoc im Web weiter pflegen und verfügbar machen können. Beispiele sind der Semantic Network Service (SNS) des Umweltbundesamtes, an dessen Entwicklung wir beteiligt waren und den wir gerade auf iQvoc umstellen oder das SURF Projekt der Deutschen Bank.

4.6 Vernetzung

Bei diesem Vorgehen sollten Ihre eigenen Daten bereits weitgehend miteinander vernetzt sein. Was aber noch fehlt, ist die externe Vernetzung mit bekannten Datenbeständen wie Wikipedia, Geonames oder branchenspezifischen Beiträgen. Potenzielle Vernetzungspartner finden Sie z.B. im Linked Data Katalog.

Auch hierbei hilft Open Source, z.B. Silk - "A Link Discovery Framework for the Web of Data" – um diese Vernetzung teilautomatisiert herzustellen.

4.7 Zusammenführen von Datenbeständen

Wenn die Datenbestände in dieser Weise verfügbar sind, können sie auch mit der auf RDF spezialisierten Abfragesprache SPARQL durchsucht werden. Das bedeutet: Sie können die Aussagen aller Datenbestände zu einer bestimmten Spezies oder Substanz usw. leicht finden und zusammenführen. Dies gilt natürlich auch für den Dokumentenindex, der Ihnen gleich die einschlägige Fachliteratur mitliefert. Natürlich gilt dies auch umgekehrt: such alle Daten zu den wichtigsten Begriffen eines Dokuments.

5 Zusammenfassung

In der Umweltinformatik haben wir eine recht ungleiche Situation hinsichtlich der Verwendung von Fachterminologie. Auf Europäischer und nationaler Ebene (in Deutschland) liegen stabile allgemeine Thesauri vor, aber die darunter liegenden Fachdomänen haben keine vergleichbare Repräsentation.

Wir sind hier weniger direkt auf eine manchmal unzureichende Motivation eingegangen, sondern haben ein Paket von geigneten Methoden zu Aufbau und Nutzung von Fachterminologie präsentiert, das jeden bestehenden Ansatz in diese Richtung motivieren und unterstützen kann. Im Fall von Fragen kann das neue UFOPLAN 3712 12 100 Linked Environment Data project (Fock, et al., 2012) möglicherweise helfen.

6 Literaturverzeichnis

[12. DTT-Symposion, 2010]

12. DTT-Symposion 2010 - Best Practices in der Terminologiearbeit 15.-17. April 2010 in Heidelberg. http://www.dttev.org/tutorialsundvortraegedessymposions.html

[Arweiler, 2008]

Arweiler, Sven: Berechnung semantischer Ähnlichkeit: Ein Vergleich zwischen WordNet und FrameNet. Saarbrücken 2008.

http://www.coli.uni-saarland.de/projects/salsa/papers/svenBSc.pdf

[DIN 2342:2011-08]

DIN 2342:2011-08 Begriffe der Terminologielehre. Berlin 2011

[Drewer, 2010]

Drewer, Petra: Präskriptive Terminologiearbeit im Unternehmen – Bildung und Bewertung von Benennungen. In: [12. DTT-Symposion, 2010].

[Massion, 2010]

Massion, Francois: Qualitätsmanagement durch Terminologie im Übersetzungsprozess. In: [12. DTT-Symposion, 2010].

[Prusseit, et al., 2010]

Prusseit, Ines; Lehmann, Sabine; Tauter, Peter: Gleich und Gleich gesellt sich gern. Begriffsorientierte Termvalidierung. In: [12. DTT-Symposion, 2010].

[Schmidt, 2010]

Schmidt, Peter A.: Quasi das Gleiche mit anderen Worten: Einige terminologische und translatologische Allusionen und Illusionen. In: [12. DTT-Symposion, 2010].

[Schmitz, et al., 2010]

Schmitz, Klaus-Dirk; Straub, Daniela: Erfolgreiches Terminologiemanagement im Unternehmen. Praxishilfe und Leitfaden: Grundlagen, Umsetzung, Kosten-Nutzen-Analyse, Systemübersicht. Stuttgart 2010

[Zerfaß, 2010]

Zerfaß, Angelika: Werkzeuge und Technologien für die Terminologiearbeit. [12. DTT-Symposion, 2010].

Open Data Governance

Arne Leißner, Condat AG, arne.leissner@condat.de
Jo van Nouhuys, Condat AG, jo.van.nouhuys@condat.de

Abstract

Mit einer Vielzahl von Initiativen agieren Politik und Verwaltung in Sachen Open Data. Auf der Grundlage des Regierungsprogramms „Vernetzte und transparente Verwaltung" setzt sich die Bundesregierung das Ziel, bis 2013 eine bundesweite Plattform für Open Data zu schaffen. Auf Landesebene (z.B. Berlin, Bayern) und in Kommunen (z.B. Köln) wurden Initiativen gestartet und erste Plattformen eingerichtet. Auch die Europäische Union hat jüngst eine neue Open Data Strategie angekündigt, die der Öffentlichkeit die umfangreichen Archive und die gewaltigen Datenbanken der öffentlichen Verwaltungen der Europäischen Union zugänglich machen soll. Unter anderem die USA, Frankreich und Großbritannien stellen schon seit längerer Zeit Daten auf Open Data-Plattformen zur weiteren Verwendung bereit.

Der offene Zugang zu Daten der Verwaltung ist kein völlig neues Thema. In Deutschland unterhält zum Beispiel das Statistische Bundesamt die Plattform deStatis mit statistischen Daten zur Weiterverwendung. Im Umweltbereich gab es bereits sehr früh in Berlin mit dem FIS Broker ein Portal für Umwelt- und Geodaten. Neue technische Möglichkeiten, verstärkte partizipative Bestrebungen und ökonomische Erwägungen (Mehrfachnutzung von Daten) erfordern neue und übergreifende Strategien zur Bereitstellung und Nutzbarkeit von Daten aus der öffentlichen Verwaltung. Sind es an einem Ende Portale mit intelligenten Such- und Zugriffsmechanismen, die die Entwicklung beflügeln, so sind am Anfang und als Grundlage vor allem verlässliche Daten die Voraussetzung.

Praktische Erfahrungen beim Aufbau und der technischen Migration von Umweltinformations- und Fachinformationssystemen zeigen jedoch, dass vorhandene Datenbestände oftmals nicht den Anforderungen an Vollständigkeit, Fehlerfreiheit

und Eindeutigkeit genügen. Das betrifft nicht nur die Primärdaten, sondern auch die Metadaten, die insbesondere die Möglichkeiten und Grenzen einer Weiterverwendung aufzeigen.

Die Herstellung von Datenqualität vor der öffentlichen Bereitstellung liegt im genuinen Interesse von (Umwelt-) Verwaltungen und ist Element einer aktiven Open Data-Strategie, einer Open Data Governance. Was ist unter dem Gesichtspunkt der Datenqualität bei dieser Strategie zu beachten? Welche Grundsätze sind anzuwenden? Welche Methoden und Werkzeuge gibt es?

1 Open Data Governance in der Praxis

In den folgenden Ausführungen wird gezeigt, wie im Zusammenhang eines Migrationsvorhabens für ein zentrales System des Umweltbundesamtes (UBA) - dem Informationssystem Chemikaliensicherheit (ICS) - für ausgewählte Aspekte die Datenqualität analysiert und gemessen wird, wie Defizite ermittelt und bewertet werden, wie unter Abwägung von Aufwand und Effekt Zielgrößen und Strategien zur Verbesserung entwickelt und in Form von Datenbereinigungsmaßnahmen umgesetzt werden.

Die Erfahrungen aus diesem Vorhaben zeigen auf, dass hierbei Daten und Metadaten unter Einschluss von Stammdaten stets als Einheit aufgefasst werden müssen, um den Anforderungen als potentieller Open Data-Quelle rechtzeitig zu genügen. Mit den erzielten Verbesserungen wurden nachweisbar wesentliche Effekte zur besseren Navigation und Identifikation und somit zum sicheren Auffinden gesuchter Informationen erzielt.

2 Open Data und Data Governance

Die Veröffentlichung von Open (Government) Data im Rahmen der Wertschöpfungskette stellt stets eine verantwortungsbewusst vorgenommene kontrollierte Bereitstellung von Daten und Informationen dar. Sicherzustellen sind unter anderem:

- die Vertrauenswürdigkeit der Informationen als glaubwürdig und belastbar,
- die Eignung für den vorgesehenen Verwendungszweck bezüglich Relevanz und Angemessenheit,

die Integrationsfähigkeit, um Informationen untereinander in Beziehung setzen zu können,

die Transformierbarkeit je nach Verwendungskontext,

die Beschreibbarkeit und Recherchierbarkeit,

die Schutzwürdigkeit von personenbezogenen und sicherheitsrelevanten Daten.

Daten gewinnen in ihrer Rolle als Open Data an Gewichtigkeit und Relevanz und sind somit als wertvolles Gut zu behandeln. Datenschätze werden gehoben und durch die Zugänglichkeit für ein großes Publikum zwangsläufig aufgewertet.

Dies passiert nicht im Selbstlauf, sondern erfordert ein System von Regeln, Verantwortlichkeiten, Prozessen und Ressourcen um dieser Erwartungshaltung und Anforderung gerecht werden zu können. Das darunter liegende System und die dort eingesetzte Methodik sowie die dabei erreichte Stufe im Reifegradmodell werden synonym als Data Governance bezeichnet.

Data Governance leistet einen Beitrag zur gemeinsamen Nutzung von Daten und Anwendungen, Systeme, Prozesse und Anwender übergreifende Wiederverwendung. Die Herausforderung ist nicht alleinig technischer Natur. Die Behandlung von Daten als Asset ist gleichermaßen darauf gerichtet, ein gemeinsames Verständnis sowie Verhalten im Umgang und in der Interpretation von Informationen zu entwickeln.

Unmittelbar auf die Daten bezogene Teildisziplinen von Data Governance sind:

Datenqualität,

Datenintegration,

Master Data Management,

Meta Data Management,

Enterprise Data Connectivity & Data Federation.

Open Data funktioniert nicht ohne die Etablierung einer Data Governance. Das Thema „Datenqualität“ ist somit nicht nur Basis jeglicher Data Governance Initiative, sondern auch unverzichtbares Element jeder Open Data Initiative.

Datenqualität kann pragmatisch als die Eignung von Daten für den vorgesehenen Verwendungszweck oder als System von Datenqualitätsdimensionen wie nachfolgend dargestellt beschrieben werden.

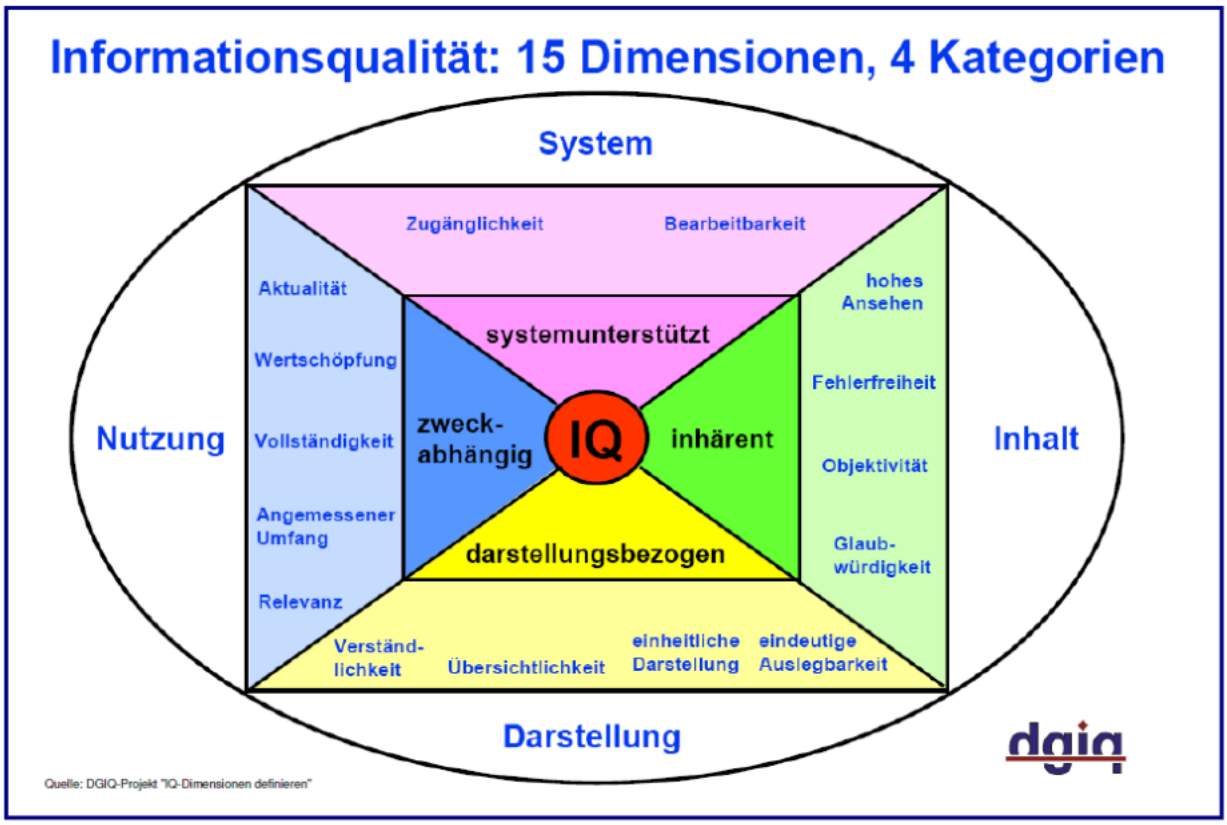


Abbildung 1: Datenqualitätsdimensionen Quelle: DGIQ-Projekt

Sind bei der Erzeugung und Verteilung von Informationen innerhalb einer Organisation oder in definierten Beziehungen zwischen verschiedenen Organisationen die Informationsempfänger in der Regel bekannt, ist in dem Umfeld von Open Data das Empfängerspektrum breit gestreut. Der vorgesehene Verwendungszweck und vorrangige Qualitätsmaßstäbe sind somit nur schwer eingrenzbar. Eine Reduktion von Datenqualität auf die Sicherstellung ausgewählter Einzelaspekte ist daher nicht zulässig. Im Kontext von Open Data ist somit der 360°-Ansatz zu verfolgen.

Was ist gute Datenqualität und was ist schlechte? Die Frage kann einerseits aus dem Gefühl und der Erfahrung heraus beantwortet werden – in Kenntnis bekannter Defizite und der mit ihnen verbundenen Nachteile sowie den resultierenden Aufwendungen zu ihrer Überwindung. Besser ist in jedem Fall ein systematischer Ansatz durch regelmäßige und wiederholte Messung des Datenqualitätsniveaus.

Wie kann eine derartige Messung und Quantifizierung vollzogen werden? Letztendlich steht immer die Frage im Mittelpunkt, in welchem Umfang die Erwartungen an die Daten erfüllt sind. Um diese Frage zu beantworten, werden Erwartungsregeln gesammelt und so formuliert, dass ihre Einhaltung mit technischen Hilfsmitteln überprüft werden kann. Wichtig dabei ist, auch scheinbar triviale Regeln aufzunehmen, weil bereits bei diesen oftmals unerwartete Regelverstöße aufgedeckt werden können. Beispiele dafür sind: „Das Feld xyz einer Datenbanktabelle muss immer gefüllt sein“ oder „Der vorgefundene Wert muss in einem bestimmten Intervall liegen oder einer vorgegebenen Notation entsprechen“. Die Erfüllung dieser einfachen Erwartungen stellt die Voraussetzung für die Sicherstellung einer Basisqualität dar, auf deren Basis komplexere Regeln aufgestellt und überhaupt erst sinnhaft überprüft werden können. Beispiele für derartige Regeln sind, spalten-, zeilen- und tabellenübergreifende Wenn-Dann-Beziehungen sowie Fragen des Matchings von Daten. Matching innerhalb einer Tabelle ist oftmals die Basis für die Identifikation und Bereinigung von Dubletten, Matching-Untersuchungen zwischen Tabellen dienen zumeist der Sicherstellung der Integrationsfähigkeit von Daten.

Je Regel lässt sich ermitteln, ob oder ob sie nicht erfüllt und wie groß das Maß der Abweichung ist. Diese Metriken können zu einem Datenqualitätsindex kombiniert und in ihrer Veränderung über die Zeit beobachtet werden. Eine stetige Verbesserung im Zusammenhang mit eingeleiteten Verbesserungsmaßnahmen manifestiert die Korrektheit des eingeschlagenen Weges und der hierbei getätigten Investitionen in Datenqualitätsinitiativen.

Was tun, wenn Regelverstöße, das heißt Abweichungen von den Erwartungen ermittelt werden? Zwischen Fachabteilung und IT ist je Regelverletzung eine Bewertung sowie – in Kenntnis geeigneter und finanzierbarer Maßnahmen – eine Zielvereinbarung vorzunehmen:

1. Die festgestellte Regelverletzung ist geschäftskritisch und nicht akzeptabel. Ziel ist die 100%ige Regelerfüllung.
2. Die Regelverletzung ist mehr oder weniger geschäftskritisch, eine 100%ige Regelerfüllung entweder nicht erforderlich oder nicht realistisch bzw. nicht wirtschaftlich erreichbar. Ziel ist ein höherer Erfüllungsgrad < 100%.

3. Die Regelverletzung ist zwar festgestellt, für das operative bzw. dispositive Geschäft ohne Belang. Die Einleitung von Verbesserungsmaßnahmen ist nicht vorzusehen. Eine Optimierung kann ggf. in der Richtung vorgesehen werden, eine zwar vorhandene, in der Praxis aber nicht verwendete Information aus den Systemen zu entfernen. Die Vermeidung von ungenutzten Datenfriedhöfen ist eine durchaus qualitätssteigernde und effizienzfördernde Verbesserung.

Was beinhalten Verbesserungsmaßnahmen? Sie zielen grundsätzlich in zwei Stoßrichtungen:

1. Datenbereinigung in den vorhandenen Datenfonds durch Korrektur oder regelbasierte Transformation an der Quelle oder in abgeleiteten Datenbeständen.
2. Präventive Prozess- und Systemverbesserungen zur Vermeidung des erneuten Auftretens von Abweichungen.

Innerhalb wirtschaftlicher Einheiten wird der Umfang und die Tiefe dieser Maßnahmen oftmals durch den erwarteten Return on Invest (ROI) bestimmt. Allerdings nicht als alleiniges Kriterium – sobald das Image des Unternehmens und das Vertrauen in seine Produkte und Dienstleistungen betroffen sind, haben diese Effekte Vorrang. Das deckt sich mit dem Anliegen von Behörden, die Daten für die Öffentlichkeit bereitstellen – auch sie möchten als vertrauenswürdig und verlässlich wahrgenommen und den hoch gesetzten Ansprüchen bei der Erfüllung des politischen Auftrages gerecht werden. Herausforderung ist die Sicherstellung von

Erwartungsgerechtigkeit,

Vertrauenswürdigkeit,

Belastbarkeit,

Regelkonformität (Compliance).

3 Datenqualität am Beispiel eines Migrationsprojekts für das Umweltbundesamt

Das Informationssystem Chemikaliensicherheit (ICS) des Umweltbundesamtes ist die dortige zentrale Vollzugsdatenbank. Sie unterstützt Aufgaben der Stoffbewertung

in unterschiedlichen Vollzügen (u.a. Pflanzenschutzgesetz (PflSchG), Arzneimittelgesetz (AMG), Chemikaliengesetz (ChemG)) und beinhaltet als Faktendaten strukturierte Daten zu

Kennzeichnung und Verwendung von chemischen Stoffen,

Physikalisch-chemische Eigenschaften dieser chemischen Stoffe,

Ökotoxikologische und toxikologische Eigenschaften,

Wirkungen auf die Umwelt

sowie weitere Klassen von Verwaltungsdaten.

Es handelt sich dabei um ein sogenanntes „gewachsenes System“ mit

einem hohen Anteil manueller Datenerfassung,

zusätzlichen diversen alten und neuen Importschnittstellen,

vielen Freitextfeldern,

Angaben in unterschiedlichen Sprachen.

Die Datenbank wurde in der Vergangenheit nicht mittels automatisierter Verfahren qualitätsgesichert. Es wurde somit zu Recht vermutet, dass inhomogene Datenlagen entstanden sind.

Der unmittelbare Anlass für die Initiierung des Vorhabens war die Migration der Daten aus der bestehenden Anwendung (ICSalt) in ein zu dieser Zeit in der Entwicklung befindliches Neusystem (ICSneu). Migrationsvorhaben sind typische Auslöser für Datenqualitätsprojekte. Zum einen muss grundsätzlich die technische Migrierbarkeit der Daten unter Beachtung der Constraints des Zielsystems sichergestellt werden, zum anderen stellt eine gute Datenqualität einen nicht zu unterschätzenden Faktor für die Akzeptanz des Neusystems dar.

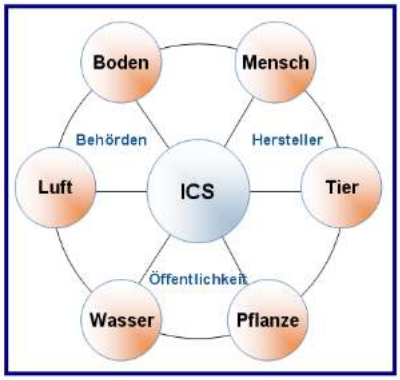


Abbildung 2: Zentrale Stellung des ICS

Das ICS nimmt beim Umweltbundesamt die Rolle eines Leitsystems ein und stellt daher bei Open Data Initiativen nicht nur selbst eine tragende Informations- und Auskunftsplattform dar sondern bildet für viele weitere Datenquellen eine wichtige und unverzichtbare Referenzquelle (Abbildung 2). Ein in seinen Daten und Metadaten konsistenter Datenbestand prägt somit auf direktem und indirektem Wege wesentlich die Qualität jeder zugänglichen Informationsbereitstellung und Datenabfrage.

Wie bei jedem DQ-Vorhaben galt es auch in diesem Fall, zunächst Erwartungen an die Daten zu formulieren, um deren Erfüllungsgrad zu überprüfen und bei festgestellten Abweichungen diese mit geeigneten Methoden zu beseitigen oder zu mindern. Für die Durchführung der Aufgabe wurde ein Ausschnitt aus den potentiell möglichen qualitätsrelevanten Fragestellungen gewählt.

Wesentliches Element von ICS ist die Verwendung von Katalogen und die Erwartung, dass in definierten Eintragsfeldern des Faktendatenbereiches ausschließlich in einem Katalog hinterlegte Textangaben gespeichert sind. Für die Untersuchungen wurden 444 Untermerkmale aus 58 Merkmalen sowie 78 mit ihnen verbundene Kataloge herangezogen. Das Ziel des Vorhabens bestand darin, den Hinterlegungsgrad zu analysieren und – soweit fachlich begründbar und vertretbar – durch Korrektur der Faktendaten zu verbessern. Abbildung 3 zeigt die Eckdaten des QS-Verfahrens am Datenbestand des Informationssystems Chemikaliensicherheit.

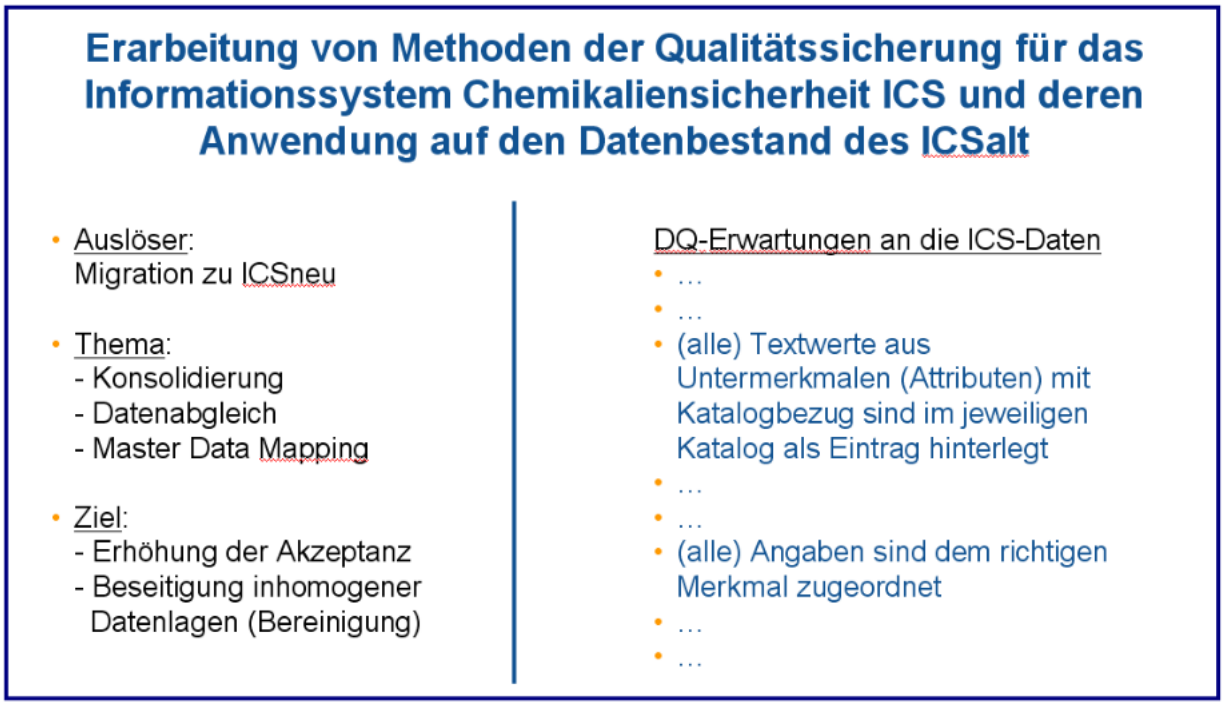


Abbildung 3: Eckdaten des DQ-Vorhabens

Für die Projektdurchführung wurde ein Vorgehen in vier Phasen gewählt:

1. Phase 1:

   Sichtung und Basisqualität

2. Phase 2:

   Analyse und Identifikation

3. Phase 3:

   Fachliche Zuordnung

4. Phase 4:

   Konsolidierung

In der Phase 1 „Sichtung und Basisqualität“ wurde vor dem eigentlichen Vergleich von Fakten und Katalogdaten zunächst eine Sichtung der einfließenden Daten vorgenommen, um daraus Rückschlüsse auf die Komplexität des Vorhabens zu ziehen und um Datenkonstellationen zu identifizieren, die sich auf die Qualität und das Ergebnis der Vergleichsoperationen auswirken können.

Zu berücksichtigen war insbesondere, dass die Daten aus verschiedenen Quellsystemen stammen und es sich weitestgehend um Freitextangaben mit unterschied-

lichen Schreibweisen und Kodierungen handelt. Es war somit von einem nur eingeschränkten Grad der Standardisierung und Vereinheitlichung auszugehen.

Zur Identifikation von Auffälligkeiten und zur Herstellung einer für den Datenvergleich geeigneten Basisqualität wird deshalb eine Reihe von Zustandsmessungen durchgeführt, die ein genaueres Bild von den einfließenden Daten liefern. Im Ergebnis dieses Profilings wurden Aussagen quantitativer und qualitativer Art über die Originaldaten gewonnen.

Zur Erzielung einer höheren Genauigkeit und zur Eliminierung vorhandener Unschärfen wurden die beschriebenen Messungen auf verschiedenen Varianten schrittweise transformierter Daten wiederholt. Beispiele derartiger Normierungstransformationen sind:

Umwandlung der Schreibweise zu Großbuchstaben,

Entfernung mehrfach aufeinanderfolgender Leer- und Trennzeichen,

Umwandlung und Korrektur von Umlautdarstellungen (u.a. Behebung von Konvertierungsfehlern aus unterschiedlichen Zeichensatzkodierungen),

Entfernung von Steuerzeichen aus den Texten,

Entfernung von Sonderzeichen.

Durch dieses Vorgehen lässt sich bestimmen, in welchem Umfang ähnliche Datenwerte vorliegen, die inhaltlich identisch sind und sich lediglich in der Schreibweise unterscheiden. Durch die Identifikation derartiger Dubletten (positiv formuliert: Von Mehrfacheinträgen) vereinfacht sich die Betrachtung der Zuordnung von Untermerkmalstexten zu den jeweiligen Katalogeinträgen: Kann ein Repräsentant der Dublettengruppe zugeordnet werden, so kann diese Zuordnung auch für die anderen Untermerkmalswerte dieser Gruppe mit alleinig technischen Mitteln vorgenommen werden.

Weiterhin wurde in dieser Phase erster Korrekturbedarf zur Sicherstellung einer hinreichenden Basisqualität der in die Analyse einfließenden Daten ermittelt. Durch Transformation wurde diese Basisqualität für die Folgephasen abgesichert.

In der Phase 2 „Analyse und Identifikation“ wurden anschließend die Zuordnungen von Untermerkmalswerten zu Katalogen im Detail analysiert und vervollständigt.

Fokussiert wird dabei auf die in den Untermerkmalen vorgefundenen unterschiedlichen (distinkten) Datenwerte. Die Häufigkeit des Auftretens eines Wertes innerhalb eines Untermerkmals liefert ein Indiz, in welchem Umfang sich ein zugeordneter bzw. nicht zugeordneter Wert auf die Zuordnungsquote eines Untermerkmals auswirkt.

Durch diese Art „Wertevorratsvergleich" lassen sich sehr schnell Aussagen ableiten, welcher Grad der Kataloghinterlegung bei jeder Untermerkmal-Katalog-Beziehung zu erwarten ist. Übersteigt die Zahl der unterschiedlichen Textwerte im Untermerkmal die des Kataloges, dann ist eine 100%ige Hinterlegung ausgeschlossen. Ähnliche Aussagen lassen sich über Längenbestimmungen erzielen. Sind Textangaben im Untermerkmal kürzer oder länger als die minimale bzw. maximale Textlänge im Katalog, so scheiden auch diese als Kandidaten für eine bereits gegebene Hinterlegung aus. In letzter Instanz wird der Hinterlegungsgrad dann über einen Direktvergleich zwischen den Merkmalsangaben und den Katalogeinträgen ermittelt.

Es stellte sich heraus, dass sich die Hypothese einer „inhomogenen Datenlage" bestätigte. Lediglich ein Viertel der untersuchten Untermerkmale wies eine 100%ige Kataloghinterlegung auf, bei fast 10% der Untermerkmale war kein einziger Eintrag hinterlegt. Insgesamt wiesen 45% der Untermerkmale einen Hinterlegungsgrad <50% auf.

Festzustellen waren dabei sehr unterschiedliche Fallkonstellationen. In einigen Fällen erschien das Prinzip der Kataloghinterlegung sachlich nicht begründet – zum Beispiel bei langen Bemerkungsfeldern, deren Freitextinhalte eine für eine Kataloghinterlegung viel zu hohe Variabilität aufweisen. In anderen Fällen waren Kataloge ungenügend gepflegt, so dass die Zahl der vorhandenen Einträge der tatsächlichen Wertemenge nicht annähernd adäquat entsprach. Als problematisch haben sich insbesondere Inkompatibilitäten zwischen den Wertemengen aus Untermerkmalen und denen aus den Katalogen herausgestellt. Dies war immer dann der Fall, wenn der Katalogeintrag sich nur auf eine Teilinformation aus den Untermerkmalswerten bezog. Beispiel sind hier insbesondere alle Formen von Messwerten in der Kombination von Maßzahl und Maßeinheit, zu denen der Katalog lediglich die anteilige Maßeinheit beinhaltet. Ein weiteres Beispiel sind Ja/Nein-Informationen, die in den Untermerkmalen und Randbedingungen oftmals mit im gleichen Feld stehenden

Freitext-Zusatzinformationen verknüpft sind, zu denen keine Katalogentsprechung definiert ist.

Das zum Hinterlegungsaspekt in der Phase 2 „Analyse & Identifikation“ gewählte Vorgehen war auf eine schrittweise Reduktion der nicht hinterlegten Restmenge ausgerichtet. Hierbei werden werkzeugbasierte Techniken (Wortlisten, Teilzeichenketten, Textmuster) mit visuellen Inspektionen kombiniert und – falls sich fachliche Entsprechungen identifizieren ließen – Vorschlagslisten für die Transformation zu einem Katalogwert entwickelt. Diese Vorschlagslisten beinhalteten Einzelvorschläge bzw. Transformationsstrategien für identifizierte Erfassungsmuster.

In der Phase 3 „Fachliche Zuordnung“ wurde je nicht hinterlegtem Eintragswert auf Basis der in der Vorphase erarbeiteten Vorschlagslisten eine Finalentscheidung bezüglich Umwandlung oder Erhalt getroffen. Zu unterscheiden war dabei zwischen Fällen, bei denen eine Zuordnung zu einem Katalogeintrag prinzipiell nicht möglich ist und denjenigen, bei denen ein Katalogeintrag als geeigneter Kandidat identifiziert werden konnte. Selbst wenn ein potentiell geeigneter Kandidat zur Verfügung stand, wurde dieser nicht zwangsläufig zur Veränderung des Originalwertes herangezogen. Insbesondere dann, wenn fremdsprachliche Informationen vorlagen sowie dann, wenn eine vorgefundenen Information mit über die Katalogentsprechung hinausgehende Zusatzangaben versehen war, wurde dem Prinzip des Quellenerhalts der Vorrang eingeräumt und auf eine Abänderung des Originalwertes verzichtet. Für die Abstimmung zwischen IT und Fachabteilung wurden Datenstrukturen vereinbart, die zum einen die erarbeiteten Vorschläge, wie auch die getroffenen Finalentscheidungen nachvollziehbar dokumentierten, für einen rekursiven Feedbackprozess geeignet waren und die technische Grundlage für die vorzunehmende Konsolidierung innerhalb der Datenbank bildeten.

In der Phase 4 „Konsolidierung“ wird zunächst der vereinbarte Sollzustand in einer Arbeitsumgebung simuliert. Die in den Phasen 1 und 2 definierten Messungen bzgl. der Hinterlegungsquoten werden wiederholt und mit den Ausgangsergebnissen verglichen. Je Untermerkmal-Katalog-Beziehung war dann zu entscheiden, ob die erzielbare Verbesserung dem Sollzustand genügt oder eine weitere Zuordnungsiteration erforderlich ist. Nach der Fixierung des Änderungsumfanges wurden Datenbankskripte generiert und mit Qualitätssicherung nacheinander im Test- und im

Produktivsystem zur Ausführung gebracht. Die eingangs angesprochenen Messungen wurden auf die transformierten Datenbestände angewendet und gegen die erwarteten Sollergebnissen aus der Simulation geprüft.

Eine schematische Darstellung des Vorgehens ist der nachfolgenden Abbildung 4 zu entnehmen.

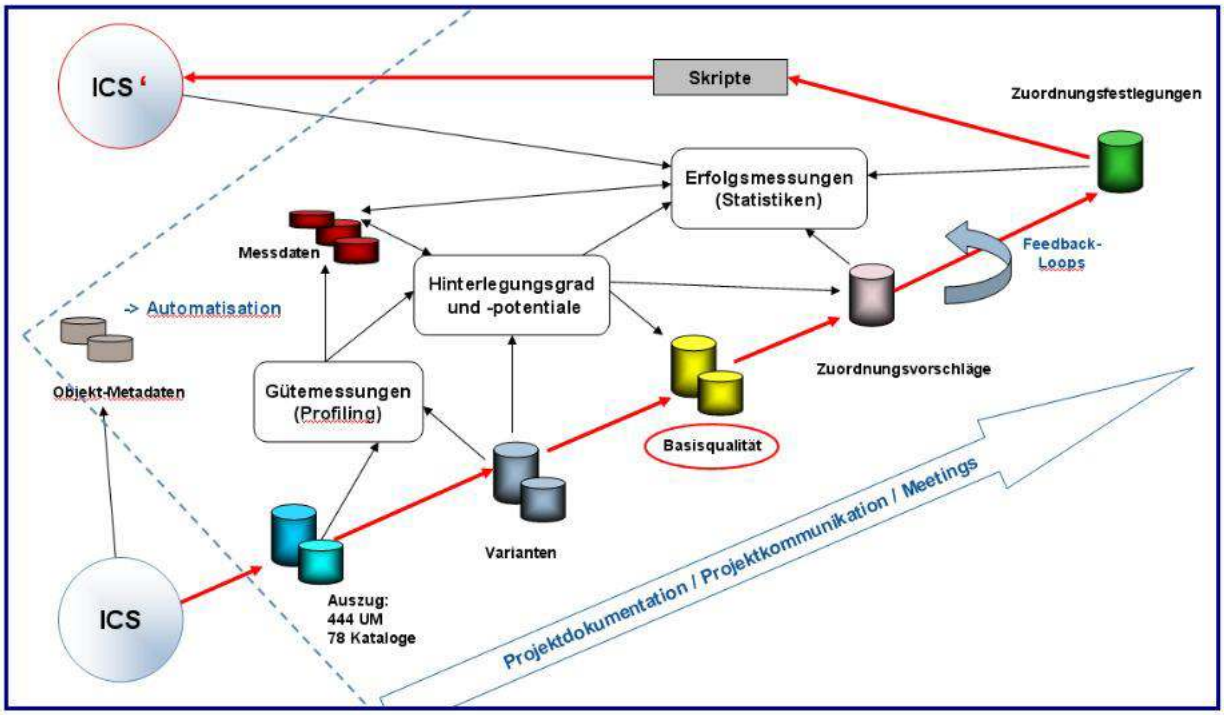


Abbildung 4: Gesamtvorgehen

Bei 444 Untermerkmalen und 78 Katalogen galt es, das Volumen der Daten und zur Vermeidung von individuellen Fehlern, den Anteil manueller Tätigkeiten soweit wie möglich zu reduzieren, ohne ihn gänzlich vermeiden zu können und zu wollen (Finalentscheidungen waren durch Fachexpertise zu treffen). Es wurde ein Framework erstellt, um einen hohen Grad der Automatisation, aber auch der Selbstdokumentation und Nachvollziehbarkeit zu erzielen. Kern dabei sind:

Objektmetadaten

Je betrachtetem Untermerkmal und Katalog (Objekt) wird eine Metadatenstruktur bereitgestellt und gefüllt, die den Auszug der Daten aus der Datenbank, die Durchführung von Texttransformationen zur Herstellung der Vergleichbarkeit und der Basisqualität, zur Durchführung fortgesetzter Messungen und Analysen, zur Durchführung der fachlichen Abstimmungsprozesse, der Konsolidierung und

deren Qualitätssicherung steuert. In den Metadaten enthaltene technische Informationen werden verwendet, um für alle Untermerkmale und in Beziehung stehende Kataloge die wiederkehrenden Standardaufgaben mittels Programmgeneratoren vollautomatisiert ablaufen zu lassen.

Messdaten

Je Untermerkmals- bzw. Katalogobjekt wird metadatenbasiert eine Struktur zur fortgeschriebenen Speicherung von Messdaten angelegt und im Rahmen der Untersuchungen fortlaufend gefüllt. Die Messergebnisse umfassen Profiling-Resultate, Zählungen aus Wertevorratsuntersuchungen, Aussagen zum Hinterlegungs- und Verwendungsgrad. Die Struktur ist derart angelegt, dass fallbezogene neue Messungen und deren Ergebnisse problemlos ergänzt werden können. Über Level ist eindeutig identifiziert, ob die Ergebnisse gleicher Messungen auf Basis von Originaldaten, mehrstufig transformierten Daten, simulierten Zuordnungsvorschläge, finaler Zuordnungsentscheidungen oder der modifizierten Zielsysteme ermittelt wurden. Somit ist auf jeder Arbeitsstufe der Arbeitsfortschritt mit dem erzielten Grad der Qualitätsverbesserung sowie die Ergebnisse der Qualitätssicherung selbst dokumentiert.

Ein in den Messdaten abgebildetes Ergebnis ist, dass die im ICS-Datenbestand durchgeführte Bereinigung zu einer signifikanten und nachgewiesenen Verbesserung der Hinterlegungsquote geführt hat. Hatten zuvor nur 55% der betrachteten Untermerkmale einen Hinterlegungsgrad >50%, wurde dieser Anteil auf 77% nach der Konsolidierung gesteigert. Wiesen vorher nur 26% der Untermerkmale eine Hinterlegungsquote von 100% auf, waren es anschließend 48%, was fast einer Verdoppelung entspricht.

Mit diesem Ergebnis wurden verbesserte Voraussetzungen geschaffen, um in einem Open Data Szenario die Navigation und Informationsrecherche unter Einsatz von Kataloginformationen maßgeblich zu stützen.

4 Open Data Governance: Best Practic

Klarheit zur Motivation (Pflichterfüllung versus Bedürfnis)

Einbettung in Strategien (Data Governance Initiative) und Sponsoren finden

Start small and focused („Klasse statt Masse“)

Überprüfbare Erwartungen an die Daten zu formulieren (auch wenn sie noch so trivial erscheinen) => Aufstellen von einfachen bis komplexen Regeln

Messbarkeit sicherstellen Erfüllungs- bzw. Abweichungsgrad pro Regel (DQ-Index über die Gesamtheit)

Immer mit dem Data Profiling beginnen

Zunächst stets Basisqualität und Vergleichbarkeit herstellen bzw. simulieren

Grenzen von Excel und SQL akzeptieren

Tools für hochwertige unscharfe Verfahren (Parsing, Standardisierung, Matching) einsetzen

Weiterhin auch visuelle / manuelle (Prüf-)Tätigkeiten vorsehen

In der Lage und gewillt sein, Entscheidungen zu treffen

Zusammenarbeit IT, Fachabteilungen, Data Stewards organisationsübergreifend organisieren

Erfüllung von Erwartungshaltungen bestätigen / Defizite identifizieren, quantifizieren und bewerten realistische Verbesserungsziele und -maßnahmen erneut messen

Web-Recherche und statistische Auswertung für Daten der Dioxindatenbank des Bundes der Länder

Gerlinde Knetsch, Umweltbundesamt, gerlinde.knetsch@uba.de
Philipp Gärtner, Umweltbundesamt, philipp.gaertner@uba.de
René Lehmann, Umweltbundesamt, rene.lehmann@uba.de
Jo van Nouhuys, Condat AG, jo.van.Nouhuys@condat.de
Gregor Raab, Condat AG, gregor.raab@condat.de

Abstract

Dioxins and PCBs are persistent organic pollutants (POPs), a term used to designate organic chemicals that accumulate in the bodies of humans, animals and plants and that have the potential to be transported across long distances. POPs may cause disturbances of the immune system, nervous system, reproductive functions and enzyme systems, with all its consequences. It is very important to monitore the contamination of the environment, foodstaff and feeding staff with dioxins permanently. As a measure to counteract the risk posed by POPs to man and the environment, a number of international environmental treaties have been agreed to in the past. For the sake of public information about the contamination of man and the environment by dioxins, the Federal Environment Agency collects data on the concentrations of dioxins and others POPs in the Federation and *Laender* cooperation Dioxin Information database www.pop-dioxindb.de. The data base includes a lot of metadata e.g. sampling site, sampling method, analysis method, laboratory data, and the analytic results with specific congeneres. An online Website with information and evaluation of the data is available. On the other hand a directly request to the database is possible getting timetrendas and congenere profiles.

1 Problemstellung

Nachrichten über dioxinbelastete Lebensmittel erschrecken Verbraucher in Deutschland immer wieder: In jüngster Vergangenheit war es dioxinbelastetes Tierfutter, aber auch Haltungsbedingungen von Tieren können Ursache von Belastungen der Lebensmittel sein. Oft wird der Verkauf potenziell betroffener Lebensmittel dann gestoppt. Ihre Auswirkungen nahmen zum Teil erhebliche Ausmaße an, die bis zum Sperren von landwirtschaftlichen Betrieben führte, wie beim Dioxinskandal mit technischen Fetten in Futtermitteln Ende 2010. Allerdings kann auch die allgemeine diffuse Belastung der Umwelt mit Dioxinen und PCB zu einer erhöhten Belastung von Lebens- und Futtermitteln führen. Ursachen fallen dabei nicht immer auf den ersten Blick ins Auge. Die Konsequenz ist jedoch, dass gleichwohl viele Verbraucherinnen und Verbraucher, Landwirtinnen und Landwirte verunsichert sind.

Doch bereits in den 1980er Jahren wurden In Deutschland eine Reihe von gesetzlichen Regelungen geschaffen, die den Ausstoß von Dioxinen aus Verbrennungsanlagen drastisch gesenkt und die Verteilung dioxinbelasteter Chemikalien unterbunden haben. Wenn in den letzten Jahren mit Dioxinen und PCB hoch kontaminierte Lebensmittel auf dem Markt entdeckt wurden, waren dafür oft punktuelle und von krimineller Energie gesteuerte Kontaminationsquellen irgendwo in der Produktionskette verantwortlich.

Aufgrund der Langlebigkeit dieser persistenten organischen Stoffe – auch POPs **P**ersistent **O**rganic **P**ollutants genannt – kommen sie in der Umwelt ubiquitär vor und werden nur sehr langsam abgebaut. Seit einigen Jahren stagniert der Eintrag in die Umweltmedien auf einem niedrigen Level. Langlebigkeit und Remobilisierung insbesondere von PCB aus Altlasten, fahrlässige Entsorgung von PCB und noch heute existierende alte sowie teilweise noch wenig erforschte neue Quellen für Dioxine und PCB sind Gründe für den weiteren Eintrag in die Umwelt. Wie Abbildung 1 zeigt, ist jedoch ein sehr deutlicher Rückgang von 1990 bis 2004 zu erkennnen.

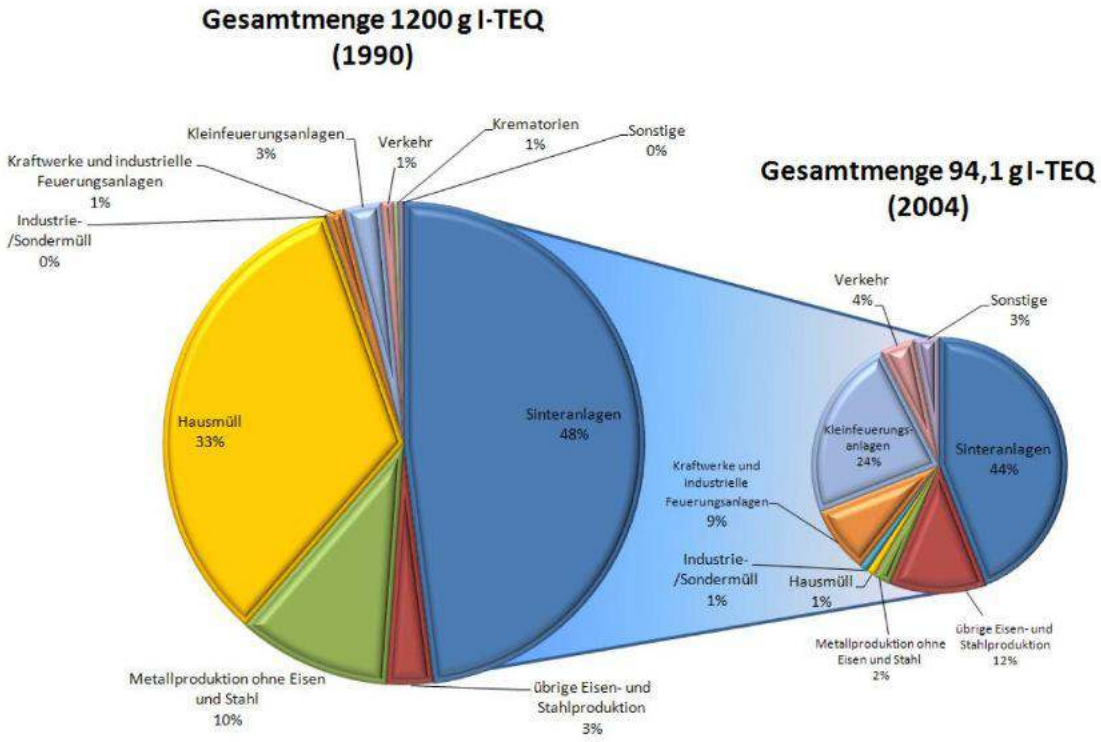


Abb. 1 Entwicklung der Gesamtemmissionen von Dioxinen in der Luft von 1990 zu 2004, Quelle: Umweltbundesamt

Der Mensch nimmt täglich Dioxine und PCB hauptsächlich über die Nahrung in Form tierischer Fette auf, auch unabhängig von besonderen Belastungssituationen. Dioxine reichern sich im Fettgewebe und in der Leber an; sie können Störungen des Immunsystems, des Nervensystems, des Hormonhaushalts, der Reproduktionsfunktionen und der Enzymsysteme mit all ihren Folgen hervorrufen.

Es ist von besonderer Wichtigkeit, die Belastung durch Dioxine in der Umwelt und in Lebensmitteln permanent zu beobachten und zu kontrollieren. Bund und Länder führen in ihren jeweiligen Zuständigkeiten Untersuchungsprogramme durch, deren Ergebnisse in Datenbanken und Informationssystemen dokumentiert werden.

Mit der Dioxin-Datenbank des Bundes und der Länder, in der die Ergebnisse von mehr als 100 Untersuchungsprogrammen des Bundes und der Länder zusammengefasst dokumentiert und strukturiert sind, wurde ein zuständigkeits- (Bund/Länder) und ressortübergreifendes (Umwelt/Lebensmittel/Gesundheit) Instrument zum Managment und Auswertung der Daten geschaffen.

2 Die Dioxin-Datenbank des Bundes und der Länder

Die Einrichtung und der Aufbau der Dioxin-Datenbank begannen 1991 mit dem Beschluss der 37. Umweltministerkonferenz. Darin wurde die Bund-Länder-Arbeitsgruppe DIOXINE beauftragt, die zentrale Dokumentation und Bewertung von Ergebnissen aus Untersuchungsprogrammen bezüglich der Belastung unterschiedlicher Kompartimente mit polyhalogenierten Dibenzo-p-dioxinen (PXDD) und polyhalogenierten Dibenzofuranen (PXDF) sowie polychlorierten Biphenylen (PCB) und weiteren chlororganischen Verbindungen in der Bundesrepublik Deutschland zu übernehmen.

Die Datenbank DIOXINE konzentrierte sich dabei zunächst auf Daten aus Untersuchungsprogrammen, die von Seiten des Bundes und der Länder initiiert wurden. Aber auch Ergebnisse, die im Rahmen von Forschungsvorhaben des Bundes und der Länder erhoben wurden, führte die Datenbank zusammen. Der ständige Datenaustausch zwischen Bund und Ländern wurde 1997 durch eine entsprechende Verwaltungsvereinbarung geregelt.

Die Datenbank wird vom Umweltbundesamt und dem Bundesinstitut für Verbraucherschutz (BVL) gemeinsam verwaltet. Aufgabe der Dioxin-Datenbank ist es, auf der Grundlage der dort dokumentierten Ergebnisse aus Messprogrammen zeitliche und räumliche Belastungsaussagen zu den Kompartimenten Boden, Wasser, Luft, und Biota zu erstellen. Standen zu Beginn der Dioxin-Datenbank die Dokumentation der Programme und ihrer Ergebnisse, die Hinzunahme weiterer Messergebnisse aus anderen Programmen und die Auswertung und Berichterstattung zur Belastung im Vordergrund, so wurden die Aufgaben mit der Einrichtung eines Webportals in 2009 auch um die Information für die Fachöffentlichkeit und die Bürger erweitert.

Der Web-Service schafft über einen Web-Client den Zugang zur zentralen Dioxin-Datenbank[9] im Umweltbundesamt. Ein Alleinstellungsmerkmal dieser Internetanwendung ist die Zusammenführung von Daten der verschiedenen Umweltkompartimente, dem Futter- und Lebensmittelbereich sowie der menschlichen Belastung in einem gemeinsamen Datenbanksystem.

[9] www.pop-dioxindb.de

3 Neue Aufgaben für die Dioxin-Datenbank

3.1 Mustervergleich von Datenprofilen

Die immer wieder in den vergangenen Jahren aufgetretenen Dioxin-Skandale haben dazu motiviert, den wachsenden und umfassenden Datenbestand der Dioxin-Datenbank nicht nur für die Berichterstattung[10], sondern auch für „Ursachenverfolgung“ zu nutzen, d.h. im Falle negativer Befunde zum Beispiel in Lebensmitteln anhand der Daten herauszufinden, woher möglicherweise die Belastung stammt und zumindest Suchhypothesen über mögliche Ursachen oder Verursacher formulieren zu können.

Dioxine – genauer PCDD, PCDF, dioxinähnliche PCB – sind Mulitkompenentengemische, die jeweils als spezifische Kongenere auftreten. Sie weisen jeweils ein spezifisches Muster in der Kombination der Einzelkomponenten auf und sind jeweils unterschiedlich toxisch. Für PCB sind theoretisch 209 PCB-Einzelverbindungen möglich. Zur Beurteilung der Toxizität wurden sogenannte Toxizitätszsäquivalentfaktoren herangezogen. Die einzelnen Kongenere werden relativ zum am stärksten toxisch wirkenden 2,3,7,8-TCDD (dem sog. „Sevesogift“) eingestuft.

Die in akuten Dioxin-Belastungssitiuationen aufgefundenen Kongenere können Hinweise auf mögliche Eintragspfade und Verursacher geben, wenn kompartimentspezifische Kongenere als Muster oder Profile in der Datenbank hinterlegt sind und das in der aktuellen Belastung aufgefundene Muster auf Ähnlichkeiten mit diesen hin untersucht werden kann.

Die Datenbank ermöglicht bereits heute den Vergleich von Verteilungsmustern mit dem archivierten Probenbestand. Konzeptionell basiert die Datendokumentation auf dem Ansatz der Einzelkongenere. Dieser Konsens der Bund-Länder-Arbeitsgruppe DIOXINE bietet die Grundlage für den Aufbau einer Kongenerenmusterdatenbank. Das Gesamtsystem verfügt derzeit über komplexe Auswertungsalgorithmen, insbesondere zu verschiedenen Toxizitätsäquivalenten (1998, 2006) und Merkmalen, die ein Kompartiment charakterisieren [Knetsch, 2012].

[10] http://www.umweltbundesamt-daten-zur-umwelt.de/umweltdaten/public/find.do;jsessionid=7E704EE682422833D6CFB430F81B0562

In der aktuellen Weiterentwicklung soll diese Berechung der Kongenerenmuster nun automatisiert erfolgen. So können Kontaminationsquellen schnell und belastbar identifiziert werden. Mit Hilfe von Expertenwerkzeugen und einem auf die Aufgabenstellung hin ausgewähltem Datenbestand werden Referenzprofile berechnet. Diese geben das typische Muster der Verteilung von Dioxinen oder PCB in einem Umweltkompartiment an. Anhand dieses Referenzprofils können Abgleiche mit kontaminierten Proben durchgeführt und strukturähnliche Umweltkontaminationen rechnerisch ermittelt werden.

Abbildung 2 zeigt den methodischen Ansatz zur Erstellung der Referenzprofile anhand des verfügbaren Datenbestandes im Dioxin-Datenbanksystem. Die mathematische Berechnung der Profile von Dioxinen und PCB erfolgt derzeit noch ausserhalb der Datenbank mit dem Statistik-Programm von MS Excel. Durch eine Exportschnittstelle aus der Webapplikation heraus können die recherchierten und aggregierten Daten in Excel weiterverarbeitet werden.

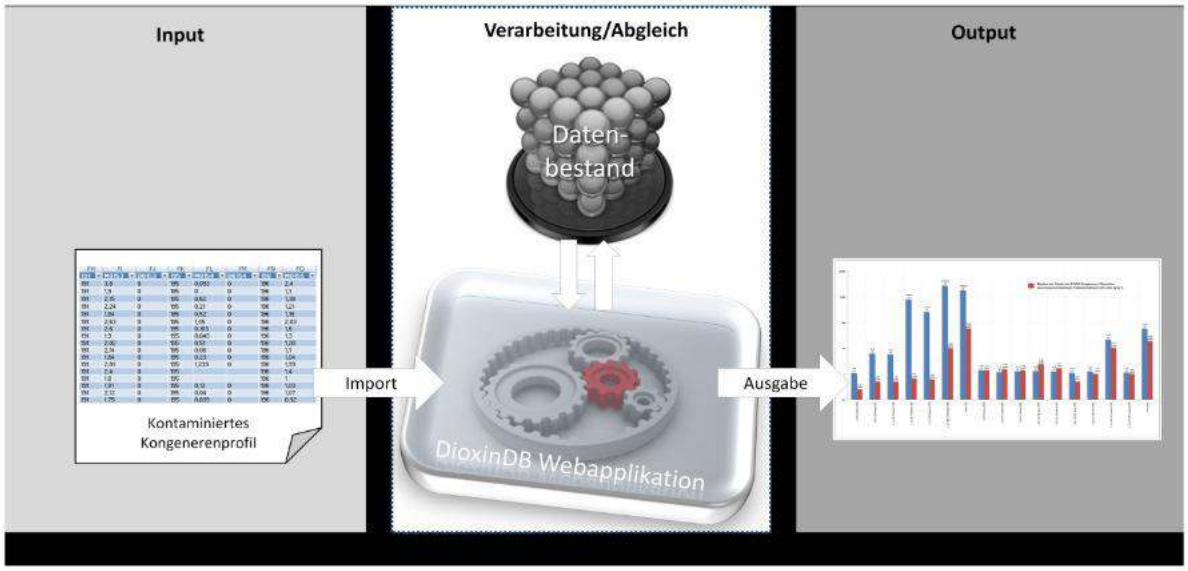


Abb. 2 Methodik zur Erstellung von Mustervergleichen, Quelle: Umweltbundesamt

3.2 Methodischer Ansatz zur automatisierten Berechnung der Datenprofile

Zukünftig sollen innerhalb der Webanwendung mit Hilfe eines Moduls der Skriptsprache R die Daten interpretiert werden. Auf diese Weise werden Kongenerenprofile gemäß aktueller mathematisch-statistischer Ansätze analysiert. In der Datenbank sind Referenzprofile für verschiedene Umweltkompartimente hinterlegt. Der Abgleich einer kontaminierten Probe mit einem Referenzprofil erfolgt automatisiert durch bestmögliche Skalierung der Referenzprobe, mit dem Ziel, beide Profile optimal in einer Abbildung visualisieren zu können. Dadurch werden Ähnlichkeiten und Unterschiede zwischen beiden Profilen leichter sichtbar.

Die Bewertung der Ähnlichkeit bzw. Unterschiedlichkeit zweier Kongenerenprofile kann anschließend vom Nutzer anhand der visualisierten Profile erfolgen. Darüber hinaus wird auf Wunsch im Hintergrund eine mathematische Bewertung der Ähnlichkeit beider Profile durchgeführt, um dem Nutzer ein Hilfswerkzeug bei der Bewertung an die Hand zu geben. Dies ermöglicht es dem Nutzer, bei Abgleich einer kontaminierten Probe mit vielen verschiedenen Referenzprofilen, sich die aus statistischer Sicht ähnlichste(n) und/oder unähnlichste(n) Referenzprofile anzeigen zu lassen.

Die statistische Methodik ist dabei als Unterstützung beim Vergleich von Kongenerenprofilen zu sehen und kann wertvolle Hinweise geben, jedoch ersetzt sie nicht den visuellen Abgleich der Profile. Diese Bewertung bleibt in der Verantwortung des Experten. Mit Hilfe der BigData-Komponente von R ist es möglich eine sehr große Anzahl Referenzprofile zu skalieren und für die graphische Darstellung aufzubereiten.

Diese Funktionalität wird dem Experten-Nutzerkreis über eine leistungsstarke, skalierbare Web Services- Plattform bereitgestellt. Durch die Einbindung leistungsstarker Mathematikbibliotheken können große Datenmengen ohne Nutzerlimitierung verarbeitet werden.

3.3 Flexiblere Datenrecherche, statistische Auswertungen und Visualisierungen

Die vorgenannten Anforderungen führen zur Forderung nach einer höheren Flexibilität der Recherchemöglichkeiten für Experten. Die gegenwärtige, ausschließlich

Kompartiment-zentrierte Recherche schränkt die Suche für Pfad- und/oder raumbezogene Betrachtungen zu sehr ein. Gewünscht werden eine freie und keine die Reihenfolge der Suchkriterien vorgebende Recherche für Experten.

Die unterschiedlichsten Aufgabenstellungen sollen dem Fachnutzer die Möglichkeiten bieten, Daten nach verschiedenen Kriterien auszuwerten. So steht manchmal die Frage im Raum, können wir nach Fanggebieten von Fischen Muster der Belastung mit PCB erkennen? Gibt es zum Beispiel Korrelationen hoher Belastungen der Fische mit bestimmten Fanggebieten? Aber auch Fragen zu technischen Produkten werden zunehmend an die Datenbank gestellt. Jüngste Beispiele zeigten, dass brüchiger Asphalt eine Quelle der Kontamination von Eiern sein kann. All diese sehr unterschiedlichen Sichten auf den Gesamtdatenbestand erfordert ein hochflexibles Recherchesystem mit vielfältigen statistischen Auswertungs- und Darstellungsmöglichkeiten.

3.4 „Offene Daten“ f r B rger und Wissenschaft

Die steigenden Zugriffe auf die Daten der Dioxin-Datenbank seit der Inbetriebnahme des Web-Portals 2003 stehen für das große Interesse der Bürger nach verlässlichen Informationen, besonders natürlich in aktuellen Belastungssituationen. Verständliche Informationen, Darstellung von Belastungssituationen und zeitliche Trendentwicklungen geben dem Bürger die Möglichkeiten, sich über die Ergebnisse zu informieren. Hierbei stehen zumeist zusammengefasste Informationen im Vordergrund, die bewertet sind und somit eine Einordnung für den Bürger zulassen.



Abb. 3: Abnahme der Dioxinbelastung in Milch von 1987 bis 2011, Quelle: BMU, 2011[11]

[11] www.bmu.de/pressemitteilungen/aktuelle_pressemitteilungen/pm/47702.php

Darüber hinaus ist die Wissenschaft an Detaildaten für eigene Untersuchungen und Analysen interessiert. Zunehmend wird der Wunsch nach sogenannten qualitätsgeprüften Einzeldaten im Rahmen der Open Data Bewegung nicht nur von der Wissenschaft, sondern auch von interessierten Bürgern und Fachnutzern geäußert. Dabei ist natürlich zu berücksichtigen, dass keine personenbezogenen Daten veröffentlicht werden dürfen und die datenschutzrechtlichen Bedingungen Beachtung finden. Auch kann den externen Nutzern – Wissenschaft und Bürgern – nicht zugemutet werden, sich in den doch sehr komplexen Strukturen der Dioxin-Datenbank selbst zurecht zu finden. Ihnen müssen Hilfsmittel für die Recherche angeboten werden.

4 Neu-Konzeption der Rechercheanwendung der Dioxindatenbank

4.1 Zweistufiges Nutzungskonzept

Die Anforderung „einfache Bedienung bei hoher Flexibilität“ impliziert in der Software-Entwicklung oft einen unauflösbaren Widerspruch: Je mehr Flexibilität und Variabilität ein Software-Produkt anbietet, desto komplexer wird die Bedienung. Bei der Konzeption der Rechercheanwendung POP-Dioxindatenbank wurde dieser Widerspruch durch ein zweistufiges Nutzungskonzept aufgelöst.

Stufe 1 Anwendung für Experten

- Unterstützung freier Recherchen an die Datenbank
- Freie Nutzung von Werkzeugen für Statistik und Visualisierung
- Definition von Anfragen, statistischen Auswertungen und Diagrammen durch Experten

Stufe 2 Anwendung für Nicht-Experten: „Ausführung von parametrisierbaren Recherchen“

Zentrales Anliegen ist hierbei,

1. die Experten von technischem Wissen zu entlasten und ihnen die Möglichkeit zu geben, die neu entstehenden Freiheitsgrade der Anwendung weitestgehend aus fachlicher Sicht nutzen zu können,
2. die Kenntnisse der Experten in den Dienst aller Nutzer der Rechercheanwendung zu stellen. Das Experten-Wissen, welche Informationen zur Darstellung eines bestimmten Sachverhalts oder zur Untersuchung einer vorliegenden Situation besonders wertvoll und aussagekräftig sind und wie diese Informationen am besten aufbereitet, aggregiert und präsentiert werden können, soll eingesetzt werden. Informative und leicht interpretierbare Recherchen können definiert und für Anfragen zur Verfügung gestellt werden.

4.2 Abstraktion vom technischen Datenmodell

Die Rechercheanwendung POP-Dioxin bietet eine nach fachlichen Kriterien konzipierte Nutzeroberfläche zur Definition von Anfragen, also zur Festlegung von Ergebnismengenstruktur und Filterbedingungen.

Die Vorteile dieses Ansatzes sind:

die Nutzer benötigen wenig Wissen zum Datenmodell,

die Erstellung einer Selektion kann aus Fachsicht erfolgen.

Die Nutzeroberfläche wird konfiguriert. Die Abstraktion vom technischen Datenmodell wird also letztlich mit Hilfe von Konfigurationsdateien vorgenommen.

4.2.1 Konfiguration des Selektionsmoduls

Bei der Konfiguration des Selektionsmoduls werden zunächst die für die Recherche relevanten Ausschnitte des Gesamtdatenmodells identifiziert und zur Definition von Anfragen bereitgestellt. Dies ist bereits ein sehr wichtiger Abstraktionsschritt. Anschließend werden die zur Selektionsdefinition bereitgestellten Merkmale mit fachlich sinnvollen Namen versehen, gruppiert und bei Bedarf mit zusätzlichen Metainformationen versehen. Die Konfiguration kann auch im laufenden Betrieb des Systems erweitert und angepasst werden. Dadurch wird auch für den Betrieb der Rechercheanwendung POP-Dioxin ein hohes Maß an Flexibilität und Erweiterbarkeit gewährleistet.

4.2.2 Auswahl von Ergebnis und Filterkriterien (Stufe 1)

Eine Recherche kann aufbauend auf diesem Selektionsbaum mit wenigen Klicks erstellt werden. Um Nutzern eine on-the-fly Recherche mit wenigen Pflichteingaben zu ermöglichen, jedoch auch individuelle Suchkompositionen zuzulassen, werden im ersten Schritt voreingestellte Kataloge angeboten. Eine Erweiterung dieser Kataloge ist jederzeit möglich und eine Recherche über diese Merkmale kann durchgeführt werden. Es können abhängig vom Datentyp verschiedene Vergleichchsoperatoren eingesetzt werden. Abbildung 4 zeigt die Auswahl möglicher Filterkriterien wie Kompartiment, chemische Stoffe, Raum- und Zeitbezug.

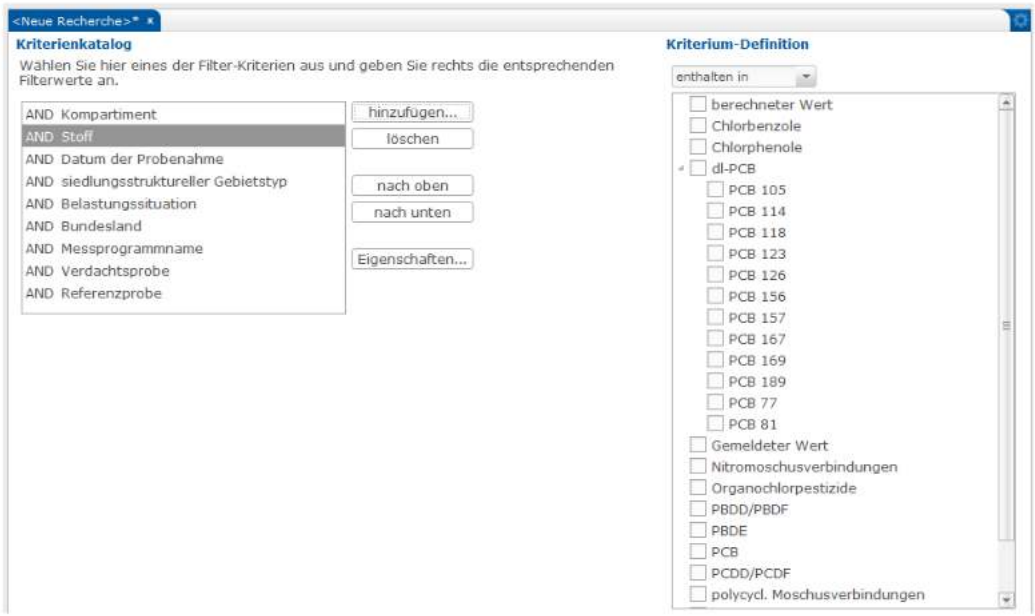


Abbildung 4: Auswahl und Definition der Filterkriterien

4.2.3 Ausführung von Recherchen für „externe Anwender“ (Stufe 2)

In die Definition von Recherchen kann die gesamte Fachkenntnis der Experten einfließen. Dabei kann für beliebige Filterkriterien festgelegt werden, ob sie bei der Ausführung der Recherche vom Nutzer parametrisiert werden können; zum Beispiel Zeitraum und Ortsbezug.

Nutzern können vordefinierte Recherchen zur Verfügung gestellt werden, die nur noch ausgeführt werden müssen, um eine aktuelle Datenauswertung zu erhalten. Der Standardnutzer muss sich lediglich aus dem Fundus der durch die Experten hinterlegten Recherchen, die auf sein spezifisches Informationsbedürfnis zugeschnittene Recherche auswählen und ausführen. Auf diesem Weg wird für die Standard-

nutzer eine sehr einfache Bedienung gewährleistet, ohne dass der fachliche Gehalt der mit Hilfe der Rechercheanwendung POP-Dioxin ermittelten Informationen eingeschränkt wird.

5 Technische Realisierung

Die „Rechercheanwendung POP-Dioxin“ wird als Web-Anwendung in einer Drei-Schichten Architektur mit den Schichten

„Datenhaltung“,

„Systemmodule“, die über einen Application-Server zur Verfügung gestellt werden und

Nutzeroberfläche im „Web-Client“.

umgesetzt.

5.1 Software-Module

Die Realisierung der Systemmodule erfolgt auf Basis verschiedener Werkzeuge und Software-Pakete, bevorzugt aus dem Eclipse™-Umfeld[12].

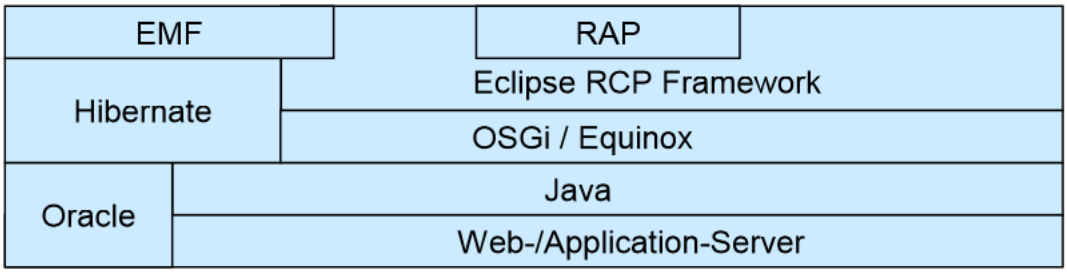


Abbildung 5: Software-Basis

Hervorzuheben ist vor allem das Software-Paket Eclipse-RAP, mit dessen Hilfe die grafische Nutzeroberfläche für das Web realisiert wird, so dass eine Web-Anwendung entsteht, die nahezu alle Vorteile einer „Rich-Client“-Anwendung besitzt.

Eclipse RAP baut auf der Eclipse-RCP-Plattform auf und gewährleistet durch die durchgehende Fokussierung auf das einer Rich-Client-Anwendung gleichende Anwendungsverhalten ein Maximum an Benutzerergonomie. Aus diesem Grunde hat man den ganzen Komfort einer Desktop-Anwendung im Browser zur Verfügung.

[12] Bei Eclipse handelt es sich um eine Open Source Community, deren Projekte die Schaffung einer offenen Software-Entwicklungsplattform zum Ziel haben.

Bäume, verschachtelte Menüs, sensitive Bereiche, auf- und zuklappbare Abschnitte, dynamische Größenänderung von Teilmasken durch den Nutzer, komplexe Auswahl-Popups mit Filterfeldern und Tabellen, Drag-and-Drop usw. sind kein Problem. Selbst das Einbetten einer anderen Web-Anwendung in einen Teil eines Eingabeformulars ist möglich.

5.2 OSGi-Module

Aufbauend auf der oben beschriebenen Software-Basis werden OSGi-Module entwickelt, die die benötigte Fachlogik bereitstellen. OSGi-Module lassen sich ohne Änderung an der Implementierung in Dienste und Anwendungen integrieren. Die Vorteile für den Einsatz von OSGi-Komponenten innerhalb der POP-Dioxin-Rechercheanwendung sind vor allem:

Unterstützung des modularen Aufbaus der Anwendung

Flexible Kombination von Komponenten für unterschiedliche Anwendungsfälle und Zielgruppen

Schnittstellenunabhängigkeit (die Module der POP-Dioxin-Rechercheanwendung können dadurch perspektivisch z.B. auch über einen Web-Service oder eine Rich-Client-Anwendung verwendet werden).

Die OSGi-Module werden für die Rechercheanwendung POP-Dioxin zu einer aus den drei Schichten „Datenhaltung“, „Logik“ und „GUI“[13] bestehenden Web-Anwendung zusammengefügt:

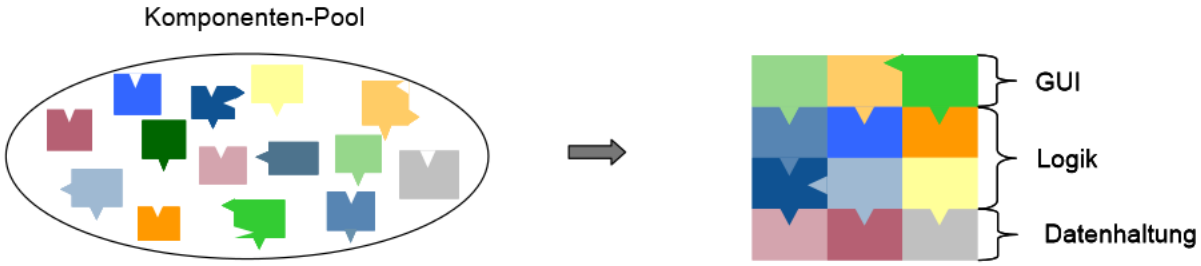


Abbildung 6: Integration der OSGi-Module in einer Drei-Schichten-Architektur

[13] GUI: Graphical User Interface, grafische Benutzeroberfläche

5.3 Werkzeuge für Statistik, Reports und Visualisierung

Für statistische Auswertungen wird das frei verfügbare Open Source Statistik-Paket „R“, eingesetzt. Es stehen hier alle für die beteiligten Behörden erforderlichen statistischen Auswerteroutinen zur Verfügung. Die Integration von R erfolgt so, dass die Statistik-Experten im UBA die Funktionalität der POP-DioxinDB durch eigene Skripte erweitern können. Insbesondere kann so der bereits vorhandene Erfahrungsschatz im UBA zu R in die neue Anwendung transferiert werden.

Die für die Präsentation der Ergebnisse der statistischen Analyse benötigten Diagramme werden über das Open-Source-Reporting-System „BIRT“ für Web-Anwendungen (Eclipse Foundation) bereitgestellt. Eclipse BIRT ist ein äußerst leistungsfähiges Reportingsystem, welches auch anspruchsvoll gestaltete Berichte in diversen Formaten mit verschiedenen Elementen generieren kann. Als Teil der Eclipse-Technologie ergibt sich eine nahtlose Integration mit den anderen Komponenten des Systems.

6 Ausblick

Perspektiven auf der Grundlage der eingesetzten Technologie Eclipse RAP

Der modulare Aufbau der Systeme und die Aktualisierung im laufenden Betrieb machen einen sanften und schrittweisen Aufbau komplexer Systeme möglich. Perspektivisch wird das UBA die Rechercheanwendung um neue Teile erweitern wollen. Mit der gewählten Technologie ist dies hervorragend möglich. Darüber hinaus können auf dieser technischen Grundlage auch mehrere Anwendungen in einem einheitlichen Erscheinungsbild zusammengeführt, modular konfiguriert und Zielgruppen-spezifisch bereitgestellt werden.

Die Schaffung und Nutzung von Dienst-Schnittstellen und die Integration von Anwendungen in eine Service-orientierte Architektur sind integrale Bestandteile der Eclipse-Technologie. Damit können die Fachfunktionen der POP-DioxinDB, wie die Recherche, der Export oder die Diagrammerzeugung perspektivisch auch extern oder von anderen UBA-Systemen genutzt werden.

Eclipse RAP wird nicht nur von allen gängigen normalen Browsern, sondern auch auf mobilen Endgeräten unterstützt. Die komfortablen RAP-Widgets sind also auch z.B. über das iPad bequem bedienbar. Außerdem ist die native Unterstützung von Touch-

Displays mit entsprechend spezialisierten Widgets bereits geplant. Um die Fülle verschiedener Endgeräte effizient unterstützen zu können, ist der Einsatz eines Frameworks wie Eclipse RAP im Prinzip unumgänglich.

Auf der Basis von Eclipse lassen sich Bestandteile der POP-DioxinDB Webanwendung z.B. als standardkonforme Java-Portlets bereitstellen, welche in ein entsprechendes zukünftiges Portal integrierbar sind. Umgekehrt wäre die POP-DioxinDB wiederum in der Lage, Bestandteile anderer Webanwendungen zu integrieren.

Perspektiven auf der Grundlage der Methode „Recherchekomponente“
Das methodische und technische Vorgehen, die Recherchekomponente konfigurierbar zu gestalten gestattet es den Experten

- flexibel und bedarfsorientiert eigene Recherchen, statistische Auswertungen und Ergebnisdarstellungen ohne Kenntnisse des technischen Datenmodells durchzuführen und
- externen Nutzern parametrisierbare Suchen in der Datenbank und weiter verarbeitbare Daten verfügbar zu machen.

Dieses Vorgehen kann grundsätzlich auch auf andere Datenbanken im Umweltbundesamt angewendet werden und dort den Experten Datenbankrecherchen erleichtern. Gleichzeitig unterstützt es die Experten bei der „Veröffentlichung“ von Daten im Rahmen einer Open Data Strategie des Hauses Umweltbundesamt.

7 Literaturverzeichnis

[Nielsen, 2000]

Nielsen, J.: Systematischer Ansatz des Internetdesigns (Erfolg des Einfachen), Markt und Technik, 2000.

[Rudlof, 2006]

Rudlof, C.: Usability Engineering (Handbuch Software-Ergonomie), UKPT, 2006

[Ballschmitter, 1980]

Ballschmitter, K.; Zell, M.: Analysis of Polychlorinated Biphenyls (PCB) by Glass Capillary Gas Chromatography (Zeitschrift für Analytische Chemie), Fresenius, 1980

[Filzmoser, 2005]

Filzmoser, P.; Garrett, R. G.; Reimann, C.: Multivariate outlier detection in exploration geochemistry (Computional Geosciences), 2005

[Knetsch, 2012]

Knetsch, G.: Integratives Auswertungskonzept von Umweltdaten mit dem Ziel der Ableitung einer Bilanz für polychlorierte Biphenyle in der Umwelt, In: Jochen Wittmann, Bernd Page (Hrsg.) Simulation in Umwelt- und Geowissenschaften, Reihe: Umweltinformatik, S. 19-30. 2012.

Einrichtung und Betrieb eines Metadatenportals für Bodenmessdaten

Jeannette Mathews, Umweltbundesamt Dessau-Roßlau, jeannette.mathews@uba.de
Thomas Bandholtz, innoQ Deutschland GmbH, thomas.bandholtz@innoq.com

Abstract

Soil monitoring programmes in Germany are widely spread by location, scope of observation, funding and temporal extent. Citizens as well as professionals have difficulties picking up the “right” data for their needs, many times not even knowing that such data exists. The Federal Environment Agency in Germany, “Umweltbundesamt (UBA)” has commissioned an R&D project to establish and operate a Metadata Portal for Soil Monitoring on a national level. It will make structured descriptions of all the distributed programmes accessible on the Web, and searchable by various criteria.

Embedded in a set of well-balanced open standards and specifications (such as OGC, W3C, INSPIRE, GS-SOIL or the national CSW profile) the portal will be built in close cooperation with the data providers and the owners of standards or re-used software components. Finally it will become a versatile information source either by itself or as a service to be integrated by INSPIRE, PortalU, and Linked Environment Data.

Einleitung

Bodendaten sind unerlässlich wenn es um Aussagen über den aktuellen Bodenzustand und seine langfristigen Veränderungen geht. Bedarf besteht sowohl an regionalisierten und über die Tiefe differenzierten Daten zu Bodeneigenschaften, Daten zur Landnutzung und Vegetation als auch an Beobachtungsdaten zu Bodenfeuchte und -temperatur, zum Grundwasserflurabstand sowie zu chemischen Bodeneigenschaften. Die Bewältigung der in den nächsten Jahren vor uns liegenden fachlichen

Herausforderungen wie der vermehrte Pflanzenanbau zur Energiegewinnung und ein demgegenüber ansteigender Nahrungsmittelbedarf (Stichwort: Flächenkonkurrenz), die zunehmende Bodendegradation/Kontamination, der Klimawandel sowie Änderungen im Boden- und Landmanagement verstärken den Ruf nach belastbaren Bodeninformationen mit zuverlässigen Zeitreihen [UBA, 2012].

Im UBA-Forschungsprojekt „Anwendung von Bodendaten in der Klimaforschung (BOKLIM)“ erfolgte erstmals eine bundesweite und messnetzübergreifende Betrachtung und Analyse der laufenden Aktivitäten zur Bodenzustandserhebung und zum Bodenmonitoring [Carolin Kaufmann-Boll, 2011].

Die Ergebnisse der Datenanalyse lassen folgende Rückschlüsse zu:

- Es gibt eine große Vielfalt an bundesweit laufenden Erhebungs- und Messaktivitäten für Bodendaten in Deutschland.(siehe Abbildung 1)
- In der Regel liefern die vorliegenden Datensammlungen keine Übersicht über verfügbare Messdaten.
- Metadaten stehen kaum zur Verfügung und wenn, dann sind sie nicht formalisiert vergleichbar.
- Die Informationen über Bodendaten sind uneinheitlich und sie liegen verteilt an vielen Stellen vor.

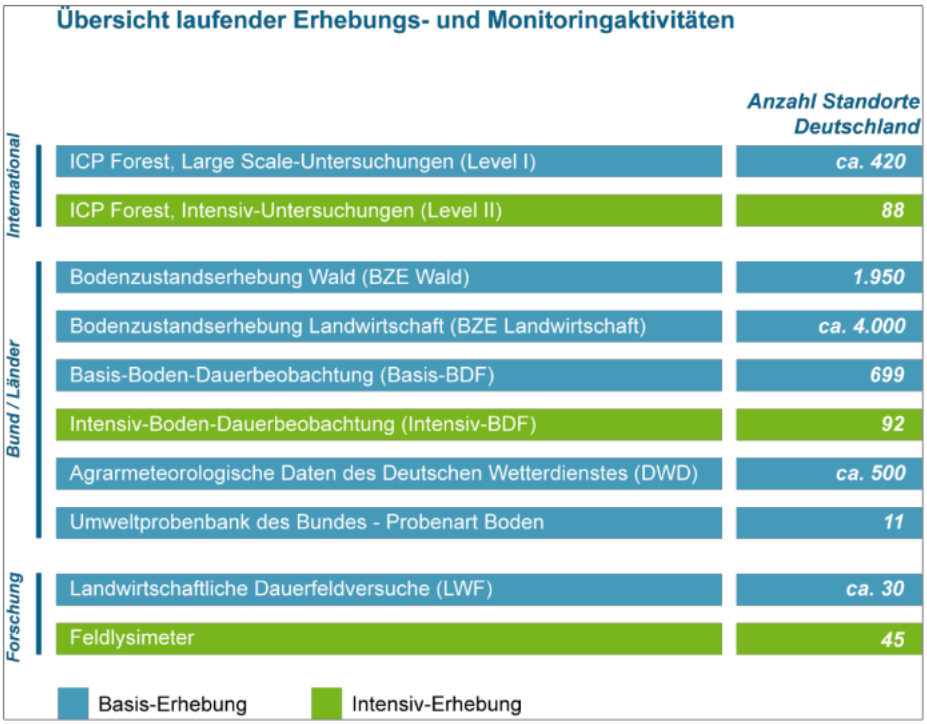


Abbildung 1: Überblick über die laufenden Erhebungs- und Monitoringaktivitäten zu Bodendaten in Deutschland, Quelle: Carolin Kaufmann-Boll, 2011

Bisher gibt es keine zentrale Auskunftsinstanz zur Recherche der vorhandenen Bodenmessdaten. Eine beispielhaft durchgeführte Recherche im Umweltportal Deutschlands (www.portalu.de) für das in der Umwelt flächenhaft verbreitete Schwermetall Cadmium (Abbildung 2) bestätigt diese Aussage.

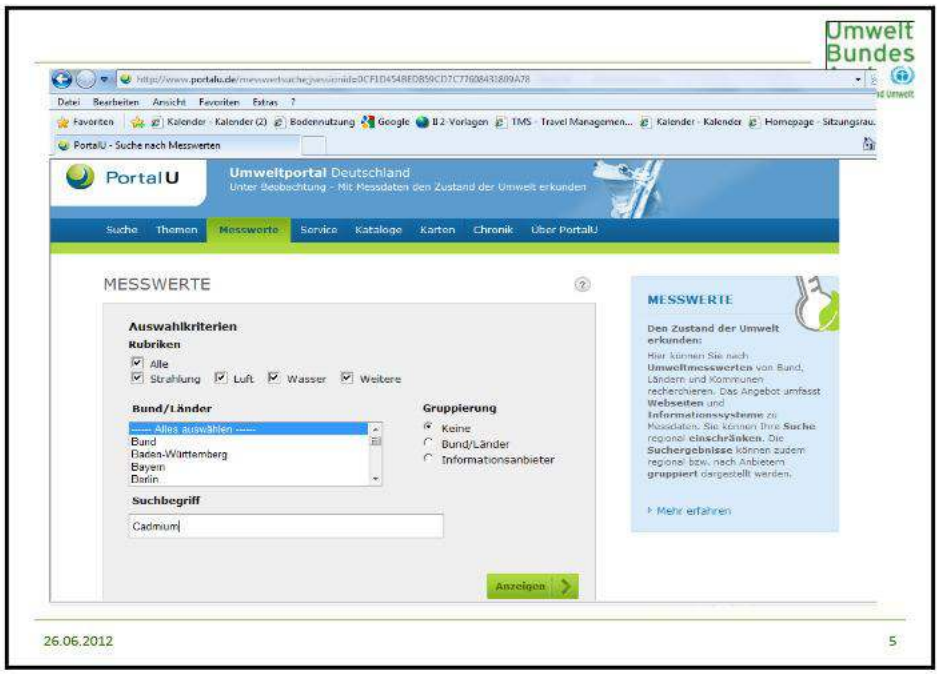


Abbildung 2: Recherche in PortalU nach Bodenmessdaten für Cadmium

Sucht man in PortalU nach Bodenmessdaten für Cadmium so erhält man gegenwärtig keine Hinweise auf Mess- und Erhebungsaktivitäten zum Boden (Abbildung 3).

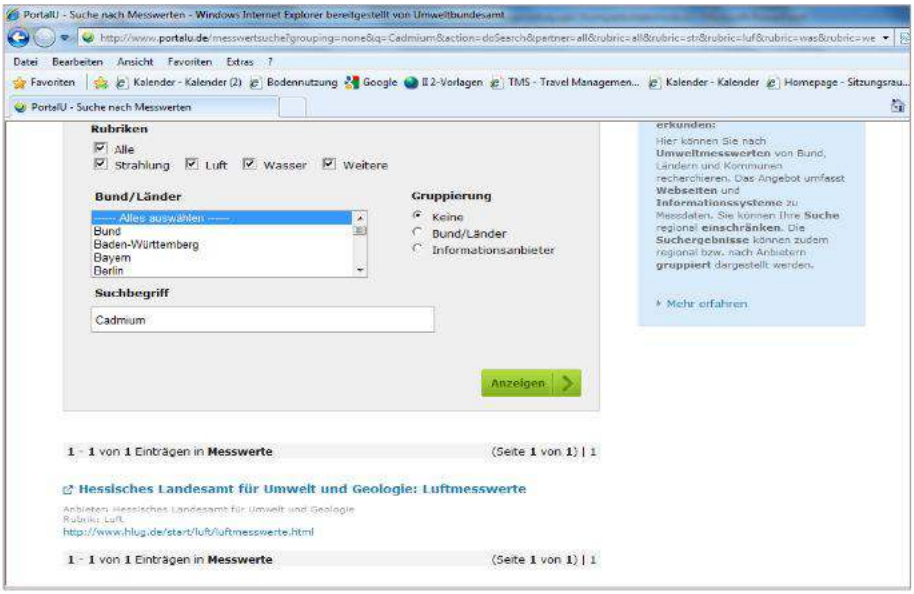


Abbildung 3: Ergebnis einer Recherche nach Bodenmessdaten für Cadmium

Cadmium wurde zum Beispiel im Rahmen der bundesweiten Bodenzustandserhebung im Wald (BZE Wald) untersucht. Darüber hinaus erfolgte in einem in Deutschland von 1990 bis 2005 im fünf-jährigem Abstand durchgeführten Moos-Monitoringprogramm die Erhebung von Cadmium-Daten [UBA, 2012].

Um über diese und viele weitere Mess- und Erhebungsaktivitäten für Böden zu informieren und die daraus gewonnenen Aussagen einem größeren Nutzerkreis bekannt zu machen, hat das UBA Anfang 2012 ein Vorhaben zum Aufbau eines Metadatenportals beauftragt.

1 Vorhaben zum Aufbau eines Metadatenportals

1.1 Projektziele

Nutzerfreundliche Bereitstellung von Metainformationen über Daten aus Bodenmonitoring und Bodenzustandserhebung. Hierzu wird eine webbasierte Informationsplattform Bodendaten neu konzipiert und umgesetzt.

Gezielte Information und Öffentlichkeitsarbeit über die bundesweiten Aktivitäten zum Bodenmonitoring und zur Bodenzustandserhebung und dem daraus resultierendem Datenangebot. Zielgruppe sind interessierte Bürgerinnen und Bürger, die wissenschaftliche Fachwelt sowie potenzielle Datennutzer.

Fragen wie z.B.: „Welche Bodenmessdaten gibt es?“, „Wer hat sie erhoben?“, „Sind die Daten frei verfügbar?“, „Wie ist die räumliche Verteilung der Untersuchungsstandorte?“, „Wer ist der zuständige Ansprechpartner?“ sollen mit dem Metadatenportal beantwortet werden können. Ein spezifisch zu entwickelnder Rechercheclient soll mögliche Suchparameter und -kombinationen sinnvoll und komfortabel bereitstellen, so dass zielgerichtet recherchiert werden kann.

Wichtig ist weiterhin ein Verweis auf bereits verfügbare Datenquellen, vorhandenen Webseiten und Publikationen.

Bei der Umsetzung dieser Ziele sind die Nutzung bereits vorliegender Softwarelösungen und die Konformität zu den laufenden Aktivitäten (INSPIRE/ GS Soil/ GDI-DE) ein wichtiger Grundsatz.

1.2 Datengrundlagen

Das Vorhaben baut auf folgenden Daten, Aktivitäten und Programmen auf:

In PortalU bereits vorliegende Metainformationen über Programme, Aktivitäten und Ergebnisse aus Bodenmonitoring und Bodenzustandserhebung.

Die seit April 2012 vorliegende UBA-Broschüre „Mess- und Erhebungsaktivitäten für Böden in Deutschland“. Die in der Broschüre dokumentierten Steckbriefe stellen eine weitere wichtige Grundlage für Metadaten über Bodenmessdaten in Deutschland dar.



Abbildung 4: Broschüre „Bodendaten in Deutschland“, veröffentlicht im UBA im April 2012, abrufbar über den Link: http://www.umweltbundesamt.de/uba-info-medien/4291.html

Sie enthalten wichtige Informationen zum Erhebungszeitraum, zur Anzahl und zur räumlichen Ausdehnung der untersuchten Standorte, zu den Untersuchungsparametern, zur Untersuchungshäufigkeit, zu den Probenahmemethoden und zur datenerhebende Institution. (siehe Abbildung 4)

Ein Metadatenprofil für Bodenmessdaten, welches im Ergebnis des BOKLIM-Projekts vorliegt (Abbildung 5).

Es umfasst insgesamt 41 Metadatenelemente, von denen 24 als verpflichtend eingestuft wurden [Carolin Kaufmann-Boll, 2011].

Anlage 4: Vorschlag BOKLIM-Metadatenprofil 30.06.2011

M	Mandatory (verpflichtend)
C	Conditional (bedingt verpflichtend)
O	Optional
X	Metadatenelement ist vorhanden (nicht differenziert als M, C, O)
(X)	Metadatenelement ist in Diskussion (nicht differenziert als M, C, O)

BOKLIM-Metadatenprofil für Datensammlungen/Datenbanken aus Bodenmonitoring und Bodenzustandserhebung							ISO/INSPIRE		GS Soil		PortalU	
Nr.	Boklim Vorschlag (Name gemäß deutsche Übersetzung ISO bzw. INSPIRE)	Bemerkungen, Beispiele	BOKLIM	ISO 19115	INSPIRE CORE	GS Soil	Portal U	ISO 19115 No.	INSPIRE No.	GS Soil Metadata Regulation Section	GS Soil Metadata Element	InGrid-Tabelle	InGrid-Feld
1) Informationen zum Metadatensatz													
1	Metadatensatzidentifikator	eindeutiger Identifikator für diesen Metadatensatz als Freitext. Beispiele: BDF-BRD, BDF-NI, BDF-064	O	O			O	2				t01_object	obj_uuid
2	Metadaten-Sprache	zur Dokumentation der Metadaten verwendete Sprache gemäß INSPIRE und ISO 639-2/B. Beispiele: ger, eng	M	C	M	M	M	3	10.3	10.3	Metadata language	t01_object	metadata_language_key
3	Metadaten-Zeichensatz	exakte Bezeichnung des Zeichencodestandards, der im Metadatensatz verwendet wird; Auswahlliste. Beispiel: UTF8	C	C		C	C	4		4.1.3	Character Encoding	t01_object	metadata_character_set
4	Metadaten-Verantwortliche Stelle	Beispiele: Umweltbundesamt, Landesamt für Bergbau, Energie und Geologie Niedersachsen	M	M	M	M	M	8	10.1	10.1	Metadata point of contact	t02_address	responsible_uuid
5	Datum der Erstellung des Metadatensatzes	Datum gemäß ISO 19103. Beispiel: 2009-08-31	M	M	M	M	M	9	10.2	10.2	Metadata date	t02_address	create_time
6	Bezeichnung des Metadatenstandards	Bezeichnung des verwendeten Metadatenstandards (einschl. Profilname). Beispiele: ISO 19115 (mit Metadatenprofil für Bodendaten UBA Deutschland, noch festzulegen)	O	O			O	10				t01_object	metadata_standard_name
7	Version des Metadatenstandards	Version des verwendeten Metadatenstandards bzw. des genutzten Profils. Beispiel: ISO 19115 2003	O	O			O	11				t01_object	metadata_standard_version
2) Datensatzbeschreibung													
8	Name der Ressource	Beispiele: Boden-Dauerbeobachtung in Deutschland, Boden-Dauerbeobachtung in Niedersachsen, Boden-Dauerbeobachtungsfläche Nr. 064	M	M	M	M	M	360	1.1	1.1	Resource title	t01_object	obj_name

Abbildung 5: Vorschlag für ein Metadatenprofil

2 Konformität zu den aktuell vorliegenden fachlichen, technischen und rechtlichen Standards (INSPIRE, ISO..)

Für die Erarbeitung und Konzeption des Metadatenprofils erfolgte eine detaillierte Auseinandersetzung mit den aus der INSPIRE-Richtlinie, dem Geodatenzugangsgesetz sowie dem Umweltinformationsgesetz resultierenden rechtlichen Anforderungen, technischen Standards und Normen. Darüber hinaus wurden die fach-

lichen Anforderungen der Projekte GS Soil und BOKLIM analysiert sowie die in PortalU enthaltenen Datenelemente. Diese Arbeiten stellen sicher, dass zum Projektbeginn eine Konformität zu den aktuell vorliegenden fachlichen, technischen und rechtlichen Standards gewährleistet ist.

2.1 Langfristige Konformität zu INSPIRE

Im Schlepptau von INSPIRE entstehen in den nächsten Jahren INSPIRE-konforme Portale und Fachschemata, die sich auch für klimarelevante Bodendaten eignen. Die Datendienste selbst werden vermutlich verteilt vorliegen. Es ist davon auszugehen, dass einige Institutionen selbst Infrastrukturen aufbauen, welche die entsprechenden Informationen als Dienste (z.B. Web Feature Services oder Sensor Observation Services) bereitstellen. Unter Umständen sind diese Dienste auch in dezentrale Portale eingebettet, die über spezifische Abfrageformulare verfügen und somit stärker auf die fachlichen Bedürfnisse zugeschnitten sind, als die zentralen Geoportale zur Umsetzung von INSPIRE [Carolin Kaufmann-Boll, 2011].

Vor diesem Hintergrund ist die Beobachtung der sich entwickelnden parallel laufenden Infrastrukturprojekte (v.a. PortalU, GDI-DE, GS SOIL) und Standardisierungsgremien (v.a. INSPIRE, ISO, OGC) ein wichtiger Projektbestandteil. Bei Bedarf sind entsprechende Anpassungen und Validierungen vorgesehen.

3 Wesentliche Arbeitsschritte im Projekt

Nach aktueller Planung (Stand: Ende Juni 2012) umfasst das Vorhaben 12 Arbeitspakete, von denen die ersten acht noch 2012 abgeschlossen werden, die weiteren vier bis Mitte 2013:

1. Abstimmung des Metadatenkatalogs: Im Rahmen des BOKLIM-Projektes wurde ein BOKLIM-Metadatenprofil erstellt (siehe Kapitel 2). Dieses bildet die Grundlage für die Arbeiten im Projekt.
2. Anpassung/Erweiterung der vorliegenden PortalU - Software: Der vorliegende InGrid®Editor wird für das unter 1. vereinbarte Metadatenprofil eingerichtet. Aller Voraussicht nach ist dies durch Konfiguration umsetzbar.

3. Metadatenerfassung: Die Steckbriefdaten aus der UBA-Broschüre „Mess- und Erhebungsaktivitäten für Böden in Deutschland“ [UBA, 2012] werden mit dem in 2. konfigurierten Editor erfasst.
4. Anschluss des Metadatenkatalogs an PortalU + Qualitätssicherung: Damit werden die Metadaten nicht nur über PortalU recherchierbar gemacht, sondern gleichzeitig auch über die CSW Schnittstelle von PortalU an die GDI-DE abgegeben.
5. Auswertung, Sichtung der in PortalU vorhandenen Metadaten zur Beschreibung der Aktivitäten in Bodenmonitoring und Bodenzustandserhebung: Eventuell erforderliche Anpassungen der Detaildarstellung werden in Abstimmung mit der Koordinierungsstelle PortalU erfolgen, da diese von vielen verschiedenen Datenanbietern verwendet wird.
6. Konzeption des Rechercheclients: Der Rechercheclient soll vielfältige Suchparameter und -kombinationen sinnvoll und komfortabel bereitstellen. Vorbild ist dabei die öffentliche Webanwendung der Umweltprobenbank[14].
7. Fachliches Feinkonzept für das Webportal: das Webportal ergänzt die recherchierbaren Metadaten mit weiterführenden Informationen, Kartendarstellungen und aktuellen Meldungen.
8. DV-technisches Feinkonzept: basiert auf der Einhaltung offener Standards für Webanwendungen sowie der INSPIRE-konformen CSW-Schnittstelle und den zeitlich parallel in UFOPLAN 3712 12 100 „Linked Environment Data“ entwickelten Spezifikationen.
9. Entwicklung und Abstimmung des Oberflächenprototyps: Die Frontend-Entwicklung wird eng mit der Realisierung des Webportals einschließlich des Rechercheclients verknüpft, um doppelt anfallende Arbeiten zu vermeiden. Der Oberflächenprototyp geht nahtlos in die softwaretechnische Realisierung über.
10. Softwaretechnische Realisierung des Webportals einschließlich des Rechercheclients: Die bisherigen Prototypen haben bereits Teile der Realisierung vorweggenommen. Sie werden nun nach den letzten Erkenntnissen refrakturiert und systematisch erweitert.

[14] www.umweltprobenbank.de

11. Datenvernetzung mit den Daten aus PortalU oder Schnittstellenanbindung: Voraussichtlich wird hier die von InGrid®Catalog angebotene OpenSearch-Schnittstelle verwendet.
12. Einbindung des Metadatenportals in PortalU: Die Seiten des WebPortals können grundsätzlich ohne besondere Anpassungen in den WebCrawler von PortalU eingebunden werden.

4 Zusammenarbeit und Datenvernetzung mit PortalU

Das UBA ist Partner des Umweltportals Deutschland (PortalU) (http://www.portalu.de). PortalU ist ein gemeinsames Internetportal des Bundes und der Länder und der zentrale Knoten für Umweltmetadaten in Deutschland. Es spielt eine entscheidende Rolle im Aufbau der Geodateninfrastruktur Deutschland (GDI-DE) und im Rahmen des europäischen INSPIRE-Prozesses. Das Portal bietet einen zentralen Zugriff auf die Internetseiten, Datenkataloge und Datenbankeinträge von öffentlichen Institutionen und Organisationen im Bund und in den Ländern. PortalU wurde mit der Software InGrid® (Information Grid) realisiert.

Die im Projekt erforderliche Metadatenerfassung soll mit dem InGrid®Editor erfolgen. Er dient der Erfassung und Pflege von Metadaten und ermöglicht eine getrennte Verwaltung von Adressen und Objekten. Metadaten können hierarchisch strukturiert werden. Der InGrid®Editor bietet darüber hinaus die Möglichkeit, einzeilige Textfelder mit oder ohne Auswahlliste den Standardfeldern hinzuzufügen.

Der für das Metadatenportal zu entwickelnde Rechercheclient greift auf die in PortalU vorhandenen Metadaten über die Bodenmessdaten zu. Die Anbindung/ der Datenzugriff soll nach Möglichkeit über die OpenSearch Schnittstelle von PortalU erfolgen. Über diese XML-basierte Schnittstelle kann ein InGrid®-Portal über einfache http-GET-Requests abgefragt werden. Die Suchergebnisse werden in einem standardisierten XML-Format ausgegeben. Die Ergebnisse einer solchen Anfrage sind in einem eigenen Portal mit individuellem Layout darstellbar.

InGrid kann mehrere Metadatenkataloge verwalten. Über die OGC-Katalogdienst-Funktion des Erntens („Harvesting“) ist es möglich, die jeweils anderen

standardkonformen Metadaten-Katalogdienste zu integrieren. Im Vordergrund steht dabei PortalU als zentrales Instrument für die Metadatendokumentation der deutschen Umweltbehörden, denn von ihnen stammen die meisten relevanten Datensammlungen. Mit Blick auf den INSPIRE-Grundsatz, dass die Geodaten nur einmal erfasst werden sollen, ist die zentrale Datenhaltung der Metadaten in PortalU vorgesehen. Es besteht das Ziel, dass die Webseiten des zu entwickelnden Metadatenportals im PortalU recherchierbar sind.

Konkrete Vorschläge zur Umsetzung werden im Projektverlauf erarbeitet.

5 Zusammenspiel der Komponenten

Das Bodenportal nutzt die InGrid®Catalog Komponente von PortalU und integriert sie in eine eigenständige Webanwendung.

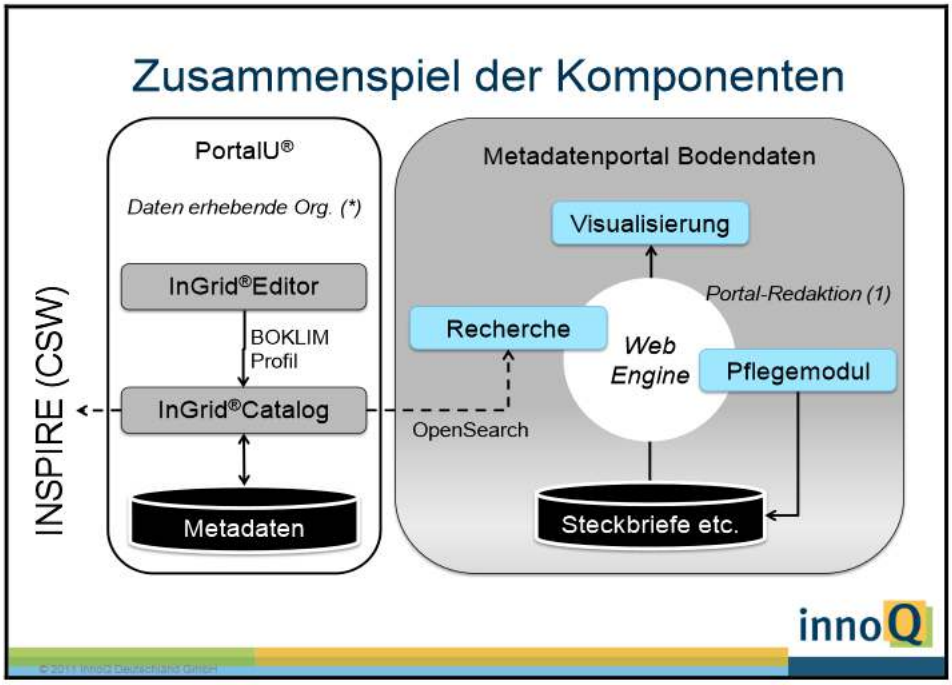


Abbildung 6: Zusammenspiel der Komponenten im UBA-Vorhaben

Bezogen auf die wesentlichen Benutzergruppen funktioniert dies wie folgt:

Die Daten erhebenden Organisationen editieren ihre Katalogdaten im InGrid®Editor und stellen der Portal-Redaktion weitergehende Steckbrief-Informationen zur Verfügung. InGrid®Catalog bietet sowohl eine INSPIRE-konforme CSW-Schnittstelle als auch eine OpenSearch-Schnittstelle, die vom Recherche-Client des Metadatenportals

benutzt wird. Dort pflegt die Portal-Redaktion weiterführende Informationen (z.B. Steckbriefe) ein, die mit den Recherche-Ergebnissen verknüpft werden.

Öffentliche Anwender können sowohl über PortalU in den Katalogdaten recherchieren als auch im Metadatenportal navigieren und mit erweiterter Funktionalität recherchieren.

Im Bild nicht dargestellt ist die Linked Open Data Repräsentation der Daten, die in Zusammenarbeit mit dem UFOPLAN Projekt 3712 12 100 „Linked Environment Data“ realisiert werden wird.

Das Vorhaben läuft 17 Monate. Im Sommer 2013 werden die Ergebnisse präsentiert und im Internet zugänglich sein.

6 Literaturverzeichnis

[UBA, 2012]

Bodendaten in Deutschland - Übersicht über die wichtigsten Mess- und Erhebungs-aktivitäten für Böden; Broschüre; 2012; Umweltbundesamt.

Link: http://www.umweltbundesamt.de/uba-info-medien/4291.html

[Kaufmann-Boll, et al., 2011]

Kaufmann-Boll, Carolin; Kappler, Wolfgang; Dr. Lazar, Silvia; Dr. Meiners, Georg; Tischler, Bettina; Dr. Baritz, Rainer; Dr. Düwel, Olaf; Dr. Hoffmann, Rainer; Dr. Utermann, Jens; Prof. Dr. Makeschin, Franz; Dr. Abiy, Mengistu; Prof. Dr. agr. Rinklebe, Jörg; Prüß, Andrea; Schilli, Carsten; Dr. Beylich, Anneke; Graefe, Ulfert: Anwendung von Bodendaten in der Klimaforschung; Texte Nr. 65/2011; 2011 Umweltbundesamt.

Link: http://www.umweltbundesamt.de/uba-info-medien/mysql_medien.php

Neue Wege in der Umweltberichterstattung - Das IT-Invest Projekt XUBetrieb des Umweltbundesamtes

Falk Hilliges, Umweltbundesamt Dessau-Roßlau, Falk.Hilliges@uba.de

Abstract

As part of the federal government's IT Investment Program, the German Federal Environment Agency (Umweltbundesamt, UBA) has developed XML repository "XUBetrieb" for collecting and reporting "company master data and environmental reporting data" from plant operators. XUBetrieb is a freely available XML structure, built on the experience gained in the development of the PRTR (Pollutant Release and Transfer Register), and it is part of the federal XÖV (XML in public administration) project. XUBetrieb is established as a free XML structure to reduce the effort required of companies to meet their environmental reporting obligations. The Federal Office for Information Technology (BIT) of the Federal Office of Administration, acting on behalf of the IT Planning Council, has successfully tested XUBetrieb (version 1.0.0) and found it compliant to the standard XÖV. XUBetrieb will be integrated into the XÖV Repository, in a platform-independent way and as an open development, so that the standards can be refined and updated.

Zusammenfassung

Das Umweltbundesamt hat im Rahmen des IT-Investitionsprogramms der Bundesregierung ein XML Repository "Betriebliche Stamm- und Berichtsdaten", kurz "XUBetrieb", entwickelt. XUBetrieb ist eine frei verfügbare XML Struktur betrieblicher Stamm- und Berichtsdaten. Es baut auf den guten Erfahrungen bei der Entwicklung des PRTR auf und ist Teil des XÖV des Bundes. Die Bundesstelle für Informationstechnik (BIT) des Bundesverwaltungsamts hat im Auftrag des IT-Planungsrats XUBetrieb (Version 1.0.0) erfolgreich geprüft. Damit ist der Standard XÖV-konform. Ziel ist, XUBetrieb zu etablieren, um die Aufwände bei der Erfüllung von Umwelt-

berichtspflichten zu verringern. Daher wurden wichtige Datenmodelle im Behörden- und Unternehmensumfeld als Grundlage für die Analyse der Struktur betrieblicher Stammdaten verwendet. Die Implementierung in das XÖV Repository ist plattformunabhängig und eine offene Entwicklung, sodass der Standard weiterentwickelt und gepflegt werden kann.

1 Hintergrund

Die Bereitstellung von Umweltinformationen, hat in den vergangenen Dekaden zunehmend an öffentlicher Bedeutung gewonnen und wurde durch die Etablierung einer Vielzahl von Umweltberichtspflichten vorangetrieben. Diese Umweltberichtspflichten sind in der Regel national, europäisch oder international rechtlich verankert. An den Berichtsprozessen ist nicht selten eine Vielzahl institutionell oft sehr divers organisierter Ebenen involviert. Das können zum Beispiel Betriebe, Landes- und Bundesbehörden oder auch europäische Institutionen sein. Die Erfahrung in der Umweltberichterstattung zeigt, dass häufig gemeinsame Standards für die Datenbereitstellung und -übermittlung fehlen. Ursachen dafür sind beispielsweise die Heterogenität der verwendeten Datenformate, redundante Datenhaltung, häufige Medienbrüche durch fehlende inhaltliche und technische Stringenz sowie fehleranfällige Individualschnittstellen. Diese Voraussetzungen führen oft zu erhöhtem Zusatzaufwand für die Datenaufbereitung und fehleranfälligen ad-hoc Prozessen. Schlechte Datenqualität und fehlende Transparenz sind dann das Ergebnis individualisierten Vorgehens. Für das Funktionieren einer Prozesskette zwischen verschiedenen Berichtsebenen innerhalb der Umweltberichterstattung sind einheitliche Standards unabdingbar.

Hier setzt das Projekt XUBetrieb des Umweltbundesamtes an. Bei diesem Projekt sollte, basierend auf den Erfahrungen, die beim PRTR[15] gemacht wurden, eine frei verfügbare XML Struktur betrieblicher Stamm- und Berichtsdaten erstellt werden, um die Aufwände bei der Erfüllung von Umweltberichtspflichten deutlich zu verringern. Beschlossen wurde das Vorhaben im Rahmen des IT-Investitionsprogramms als Teil des Konjunkturpakets II (02/2009) der Bundesregierung mit einem Gesamtvolumen von 500 Mio. Euro. Das Projekt hatte eine Laufzeit von April 2011 bis November

[15] PRTR (Pollutant Release and Transfer Register, Schadstofffreisetzungs- und Verbringungsregister, http://prtr.bund.de)

2012. Projektpartner waren Industriebetriebe, Verbände, Kommunal-, Landes- und Bundesbehörden sowie weitere eGovernment-Institutionen.

2 Projektziele und Aufgaben

Ziel des Vorhabens war die Entwicklung einer frei verfügbaren XML Struktur für betriebliche Stamm- und Berichtsdaten, insbesondere aus den Bereichen Genehmigung, Vollzug und Berichterstattung. Diese XML Struktur sollte als XÖV[16] Standard für den Umweltbereich etabliert werden. XÖV ist Teil der eGovernment Strategie der Bundesregierung. XÖV beinhaltet Fachstandards für den elektronischen Datenaustausch in der öffentlichen Verwaltung auf Basis von XML. Ziel von XÖV sind standardisierte, durchgängig elektronische Geschäftsprozesse innerhalb der Verwaltung. Die XÖV Standards gewährleisten die Sicherstellung der Wiederverwendbarkeit und die Förderung der Interoperabilität. Weiterhin wird der Aufwand zur Schnittstellenmodellierung durch die bestehenden Standards deutlich reduziert. Abgelegt sind die Standards (Datenmodelle, Schemata, Dokumentationen) in einem XRepository[17], das vom Bundesland Bremen geführt wird.

3 Projektverlauf und Ergebnisse

Das Projekt gliederte sich in der Umsetzung in drei Teilprojekte auf. Projektsteuerung und Kommunikation, Bestandsaufnahme und Kernkomponentenidentifizierung. Es wurden zu Beginn des Projektes zwei Kick-Off Workshops mit Industrie- und Behördenvertretern durchgeführt. Dabei gab es große Einigkeit bei den Problemerfahrungen und daraus resultierte ein breites Interesse an Mitarbeit und Unterstützung. Darauf folgte eine Fragebogenaktion, mit der die Berichtspflichten mit den höchsten Aufwänden herausgearbeitet werden sollten. Weiterhin sollten Erleichterungs- und Verbesserungspotenziale in den einzelnen Berichtsprozessen aufgezeigt werden. Im Ergebnis wurden die Berichtspflichten mit den größten Arbeitsaufwänden festgestellt:

[16] XÖV XML in der öffentlichen Verwaltung
[17] https://www.xrepository.deutschland-online.de/

E-PRTR-VO (VO 166/2006 EU)

11. BImSchV (Emissionserklärung)

91/271/EWG (Kommunalabwasser)

13. BImSchV (Großfeuerungsanlagen)

WHG § 61 (Kläranlagen Eigenüberwachung)

2008/1/EG IVU, zukünftig IED (Industrial Emissions Directive)

Daraus resultierend ist auf der Grundlage des XÖV Leitfadens ein UML-Modell und das zugehörige XML Schema für betriebliche Stamm- und Berichtsdaten entwickelt und XÖV-zertifiziert worden. Die Nutzung für einzelne Fachanwendungen kann auf der Basis eines Web-Service erfolgen. Die XUBetrieb Modellkomponenten dienen der Zusammenstellung von Datenmodellen für den Aufbau von Nachrichten und Datenbanken auf einer standardisierten Basis. Sie bilden eine Bibliothek, aus der sich jeder die passenden Bausteine für seine Aufgabe wählen kann. Die Modellkomponenten sind frei verfügbar. Die offene Lizenz sorgt dafür, dass dies auch zukünftig gewährleistet ist.

Die XUBetrieb Modellkomponenten sind von der Bundesstelle für Informationstechnik (BIT) des Bundesverwaltungsamts im Auftrag des IT-Planungsrats erfolgreich geprüft worden. Somit ist XUBetrieb (Version 1.0.0) zertifizierter XÖV-Standard und auch im XRepository hinterlegt. Sie können die Modellkomponenten als XML-Schema oder UML-Modell auf der Projekthomepage herunterladen. Um sowohl die Tauglichkeit des Ansatzes als auch die Tauglichkeit der Modellkomponenten für die Entwicklung von Datenmodellen für Schnittstellen zu zeigen und gleichzeitig eine Dokumentation für die Nutzung zu erstellen, wurde aus den XUBetrieb Modellkomponenten ein zweites UML-Modell XUKommunalabwasser abgeleitet und für dieses eine Referenzimplementation unter Verwendung von RESTful Webservices durchgeführt. XUKommunalabwasser ist ebenfalls zertifizierter XÖV-Standard. Die XUKommunalabwasser Referenzimplementation soll zukünftig als Alternative zu den bestehenden Schnittstellen für den Austausch von Berichtsdaten der Kommunalabwasserrichtlinie zwischen Bundesländern und der zentralen Webanwendung e-Kommunalabwasser des Umweltbundesamts dienen. Durch eine allgemein verfügbare, moderne XML

Struktur wird die Grundlage zur Prozesslevelanhebung geschaffen, sodass Behörden und Unternehmen weniger Ressourcen durch Bürokratiekosten binden.

4 Fazit

Die XÖV standardisierten XUBetrieb Modellkomponenten können von neu zu entwickelnden Schnittstellen genutzt werden, um die Struktur der ausgetauschten Nachrichten festzulegen. Die Verwendung der Komponenten erzeugt an keiner Stelle mehr Aufwand, sondern reduziert den Aufwand bei der Definition der Schnittstellen. Je mehr Anwendungen von den XUBetrieb Modellkomponenten Gebrauch machen, desto wahrscheinlicher wird es, dass es auf einer Seite der Kommunikation bereits eine XUBetrieb kompatible Schnittstelle gibt. Wenn ein oder beide Partner eines Datenaustauschs (Nachrichtenaustauschs) bereits die Modellkomponenten einsetzen, reduziert sich der Programmieraufwand erheblich. Am Ende haben alle Kommunikationspartner gleiche Schnittstellenstrukturen. Dabei wird keinem Partner vorgeschrieben, welche Infrastruktur zu verwenden ist. Wenn sich Anwender (Unternehmen, untere Behörden, Landesbehörden oder Bundesbehörden etc.) zusammenschließen und eine Anwendung mit der XUBetrieb Schnittstellenstruktur entwickeln, dann werden die Daten sehr leicht austauschbar sein.

Bausteine zu einer Technologie für Participatory Sensing Anwendungen

Andreas Abecker, disy Informationssysteme GmbH
abecker@disy.net
Valentin Zacharias, Julio de Melo Borges, FZI Forschungszentrum Informatik
zach@fzi.de, borges@fzi.de

Abstract

We present the idea of Participatory Sensing which is a Crowdsourcing approach where people use mobile sensors and their mobile phones or tablet PCs for data transfer in order to collect data for some dedicated purpose. After a comprehensive Participatory Sensing definition and a review of applications for nature conservation and environmental protection, we sketch elements of a generic software framework and the current implementation status of the PartSense project.

Zusammenfassung

Wir stellen die Idee des Participatory Sensing vor, bei dem im Sinne des Crowdsourcing Menschen mit mobiler Sensorik und Datenübertragung per Mobiltelefon bzw. Tablet-PC zu einem übergeordneten Zweck Daten sammeln. Nach einer umfassenden Definition des Participatory Sensing und einer Darstellung von Anwendungen im Bereich Natur- und Umweltschutz, skizzieren wir Elemente einer generischen Software-Lösung und den aktuellen Implementierungsstand im Projekt PartSense.

1 Einleitung

Der Ansatz des Participatory Sensing (PS, auch: Human Sensing, Urban Sensing, Opportunistic Sensing) kombiniert (1) die technischen Möglichkeiten moderner mobiler Endgeräte (Smartphones und Media Tablets) mit ihrer GPS-Lokalisierung und ihrer vielfältigen eingebauten Sensorik mit (2) der Idee des Crowdsourcing für das Delegieren einer Daten-Sammel-Aufgabe an eine große Anzahl von Menschen

(Mitarbeiter, registrierte Freiwillige oder auch die allgemeine Öffentlichkeit) (Goldman et al, 2009).

In den vergangenen fast 10 Jahren gab es bereits eine große Zahl hoch interessanter Prototypen und Forschungsanwendungen. Zieht man aber die gewaltigen aktuellen Wachstumsraten bei Smartphones, Media Tablets wie auch sozialen Medien und Online Communities in Betracht, liegt es nahe, die größten Chancen für PS-Anwendungen sind erst noch in der Zukunft zu erwarten. Auch sind viele ältere akademische PS-Beispiele heute technologisch nicht mehr auf der Höhe der Zeit. Sie waren von Anfang an zeitlich begrenzte Experimente und fanden keine ausreichende Unterstützergruppe. Sie basierten auf aufgabenspezifischen ad-hoc Implementierungen oder waren vom Anwendungsfall her auf US-amerikanische Rahmenbedingungen abgestellt.

All dies hat die Idee motiviert, sich in einem neuen Forschungsprojekt mit einer aktuellen, wissenschaftlich-technisch fundierten Herangehensweise an PS-Anwendungen zu beschäftigen und die Grundlagen für die ingenieurgemäße Erstellung von PS-Anwendungen zu konsolidieren. Dies sollte auch die Untersuchung professioneller bzw. kommerzieller Anwendungen umfassen, während sehr viele der akademischen Beispiele offene Anwendungen für die allgemeine Öffentlichkeit darstellen und sich häufig auf Altruismus als einzige Motivation verlassen haben. Die vorgenannten Ziele werden im BMBF Projekt PartSense (Laufzeit 01/2012 – 12/2013 verfolgt). In diesem Beitrag präsentieren wir erste Projektergebnisse sowie eine Sammlung bekannter Anwendungsfälle aus dem Bereich Umwelt- und Naturschutz.

2 Eine Arbeitsdefinition von Participatory Sensing

Unsere Arbeitsdefinition versucht, das breite Spektrum von durch den Participatory-Sensing Begriff erfassten Ansätzen möglichst gut zu überdecken:

Beim Participatory Sensing sammelt (a) eine Gruppe von Personen (b) im Rahmen einer Datensammlungskampagne (c) raumbezogene Daten, unter Verwendung (d) marktüblicher Internet-fähiger mobiler Endgeräte und deren (e) eingebauter Sensor- und Lokalisierungsmechanismen. Die Personen, die Daten sammeln, senden diese mithilfe ihrer Mobilgeräte an einen zentralen Server, der die Daten für einen definierten Zweck weiter verarbeitet. Art und Format der zu sammelnden Daten

sowie weitere Eigenschaften der Sammelkampagne werden im Szenario vorher festgelegt. In komplexen Anwendungsfällen kann der Server aktiv mögliche Datensammler ansprechen bzw. deren Aktivitäten steuern.

Konkrete PS-Projekte variieren entlang der Ausprägungen der Merkmale (a) – (e):

a. Die **Datensammler** können dem System bekannte Personen sein (z.B. *registrierte Nutzer*) oder sich *anonym* an der Kampagne beteiligen (durch Download einer Smartphone-App). Sie können in einer *formalen Beziehung* zum Kampagnen-Organisator stehen (z.B. Angestellte, bekannte Mitglieder einer NGO oder registrierte Mitglieder eines Freiwilligen-Pools), aber sie *müssen nicht* (in vielen Fällen stammen die Datensammler aus der „allgemeinen Öffentlichkeit").

b. Eine **Daten-Sammel-Kampagne** findet i.A. in einem *räumlich abgegrenzten Gebiet* statt. Zeitlich gesehen, kann eine Kampagne eine *feste Dauer* besitzen, *unbegrenzt* laufen oder die Dauer kann *implizit gegeben* sein (z.B. wenn so lange Daten gesammelt werden, bis eine gewisse Menge erreicht ist).

c. Der **Raumbezug** der Daten ist in den meisten Anwendungen der Bezug auf eine einfache *Punktkoordinate. D*iese wird von der GPS-Lokalisierungsfunktion des mobilen Geräts geliefert, kann aber typischerweise manuell berichtigt werden, wenn nötig. Wenige Anwendungen erfassen auch *Liniengeometrien* wie GPS Tracks, häufig auch mit Höheninformationen, *z.B.* eine Mountain-Bike- oder Wanderstrecke oder eine Skipiste. Im Allgemeinen kann man sich auch auf ein *Polygon* beziehen.

d. Wir betrachten keinerlei aufgabenspezifische Spezial-Hardware, sondern Internet- und GPS-fähige **Standard-Geräte** wie Smartphones und Media Tablets (Apple iPad, Samsung Galaxy Tab, etc.) – hier trifft sich der PS-Ansatz mit der "bring your own device" Philosophie.

e. Bei der **Sensortechnology** verlässt sich PS weitgehend auf die *eingebauten Mechanismen der Mobilgeräte*: Tastatur, Audio-, Video- und Foto-Aufnahmen, GPS-Funktionen für die Lokalisierung, Drei-Achsen-Beschleunigungssensor, Gyroskop, Umgebungslichtsensor, Kompass, Feuchtigkeitssensor. In der nahen Zukunft werden wahrscheinlich weitere Sensoren hinzukommen, wie z.B. ein Barometer. In manchen Fällen kann es sinnvoll sein, auch weitere tragbare/

mobile Sensoren zu verwenden, die das Mobilgerät für die Datenübertragung ansprechen können, z.B. zur Überwachung von Vitalsignalen. Ein Beispiel wären die in Bekleidung integrierbaren Sensoren für die Herzfrequenz, ECG, Körpertemperatur, Hautwiderstand u.a.m.. Solche Sensoren verbinden sich für die Datenübermittlung per Bluetooth mit dem Mobilgerät.

Es gibt noch weitere Unterscheidungsdimensionen konkreter PS-Anwendungen:

Leute können an der Datensammlung *anonym* oder *authentifiziert* teilnehmen.

Die Übernahme einer bestimmten Sammelaufgabe kann *auf eigene Initiative* des Nutzers erfolgen oder aufgrund einer *aktiven Kontaktaufnahme durch den Server*. Vom konkreten Anwendungsfall abhängig, kann eine solche Anfrage durch den Server einen unverbindlichen oder einen verbindlichen Charakter besitzen. Wenn der Server potenzielle Sammler aktiv kontaktiert, kann die Adressatenauswahl rollen-, benutzerprofil- oder lokations-basiert geschehen – typischerweise werden verschiedene Aspekte kombiniert.

Aufgaben zur Datensammlung können die *Ersterfassung* eines Datenpunkts betreffen, seine *Aktualisierung* oder *Verifikation*.

Participatory Sensing kann eine Reihe von Vorteilen bieten. In dem Fall, dass ein PS-Ansatz in einen bereits existierenden Geschäftsprozess eingebaut wird, der zurzeit noch ohne jede mobile Software-Unterstützung abläuft, können natürlich alle Vorteile einer mobilen Software-Lösung zum Tragen kommen. Dies betrifft die Eliminierung von Medienbrüchen, Zeitersparnis und die Ausnutzung ohnehin gegebener Mobilität von Mitarbeitern. Spezifischer für PS im engeren Sinn ist aber die Ausnutzung des Crowdsourcing-Gedankens, also das Verteilen einer Aufgabe an *viele* Akteure mit sehr kleinen und einfachen Teilaufgaben; was auch insbesondere die Möglichkeit eröffnet, *neue* Personenkreise einzubinden, gelegentliche Mitarbeiter, Freiwillige, Ehrenamtliche oder sogar die allgemeine Öffentlichkeit.

Was Anwendungsfälle und –domänen des PS angeht, gibt es in der Fachliteratur eine große Bandbreite von Pilotprojekten; die Mehrzahl davon lässt sich in eine der folgenden Gruppen einordnen (1) Bürger-Feedback an die öffentliche Verwaltung (Mängelmelder-Szenarien); (2) Naturschutz und Umwelt-Monitoring; (3) Preisvergleiche; (4) Sport- und Gesundheits-Monitoring; (5) Verkehrs- und Mobilitäts-Monitoring.

Es gibt weitere Anwendungsszenarien, wo PS-Ansätze offensichtlich nützlich sein könnten, zum Beispiel: (i) Überwachung und Wartung großer, verteilter Anlagen und Infrastrukturen (wie große Fabrikgelände, eine Wohnsiedlung oder ein Nahverkehrsschienennetz mit den dazugehörigen Anlagen); oder (ii) Naturkatastrophen (wenn z.B. Einwohner eines Krisengebiets, Rettungs- und Einsatzkräfte verteilt in Echtzeit Informationen über die Lage sammeln können bzw. auch gleichzeitig Aufgaben kontinuierlich an Hilfs- und Rettungskräfte verteilt werden müssen). Andere Anwendungsideen haben sich aus Gesprächen mit Kunden ergeben, z.B. in der Forstverwaltung, bei der Dokumentation von Altlasten oder bei der Straßenbauverwaltung.

Da Umwelt- und Naturschutz eine verhältnismäßig große Rolle bei den PS-Anwendungen in der Fachliteratur spielt, skizzieren wir im folgenden Kapitel eine Reihe von Anwendungsbeispielen aus diesem Bereich.

3 Participatory Sensing für Umweltanwendungen

Die wichtigsten umwelt-orientierten PS-Anwendungsfelder:

Artenzählungen: Die *Artenfinder* App (http://artenfinder.de/) ermöglicht bspw., Sichtungen seltener Tiere oder Pflanzen durch ein Foto plus Zusatzinformationen zu dokumentieren. Es laufen zurzeit Pilotprojekte in Rheinland-Pfalz und in Nordrhein-Westfalen.

- Interessante Aspekte: Artenfinder bietet einen *Kompass* und eine *Entfernungsangabe,* um weiter entfernte Sichtungen besser lokalisieren zu können. Außerdem werden *Hintergrundinformationen* angeboten, die bei der Bestimmung der vorliegenden Spezies helfen sollen. Die Pilotprojekte laufen in enger Zusammenarbeit mit Naturschutzverbänden und mit Fachleuten im Back-Office, die die eingehenden Meldungen bewerten und die Ergebnisse an die Naturschutzbehörden weiterleiten.
- Ähnlich: What's invasive (http://whatsinvasive.com/) – hier werden Sichtungen von invasiven Spezies gesammelt, zurzeit in ca. 60 Nationalparks, überwiegend in Nordamerika.
- Schon vor dem Siegeszug von Mobiltelefonen wurden ähnliche Ideen mithilfe von Web-Eingabemasken realisiert, bspw. bei *webfauna.ch* (Sichtungen von

Amphibien, Reptilien und Säugetieren in der Schweiz). *naturgucker.de / enjoynature.net* ist eine Community-Plattform mit vielen nützlichen Funktionen für Kommunikation und Datenverwaltung, mit der Privatpersonen (typischerweise Mitglieder von Naturschutz- oder ornithologischen Vereinen) ihre Naturbeobachtungen (Flora, Fauna) speichern und inspizieren können. Sichtungen können entweder *einem Punkt oder einem Gebiet* zugeordnet werden. Sichtungen bedrohter Arten können von anderen nicht im Detail angesehen werden. Das Informationssystem "Vogelz hlung in Deutschland" ist ein Web-basiertes System, die Zuweisung konkreter Zählaufgaben in bestimmten Gebieten wird zentral organisiert, vom Dachverband der ornithologischen Verbände (DDA).

Überwachen einer bestimmten Gegend: Die *SenseTheBeach* App ist dafür gedacht, verschiedenste Arten von Beobachtungen zu sammeln, die für den *guten Zustand einer bestimmten geographischen Region* relevant sind, also z.B. im Fall eines bestimmten Strandes Informationen zu: (i) invasiven Arten (Algenblüte, Seegras), (ii) Fauna (Vögel, Quallen, tote Tiere), und (iii) vom Menschen verursachten Einflüssen (überfüllter Strand, verbotene Fahrzeuge, Müllablagerungen, Abwassereinleitungen) (http://sensethebeach.appspot.com/).

Überwachen der Luftqualität: in den Jahren 2005-2010 wurden in den USA viele Experimente zum Monitoring der Luftqualität mit PS durchgeführt, was natürlich besonders nützlich ist, wenn nur eine geringe Dichte staatlicher Messstationen gegeben ist. Über die Möglichkeiten der öffentlichen Verwaltung hinausgehend, hat PS auch die Chance, *Indoor-Messungen* zu sammeln bzw. kann auch dort funktionieren, wo überhaupt keine staatlichen Messungen durchgeführt werden (Paulos et al, 2008).

- Es ist auch zu erwähnen, dass Luftqualitätsüberwachung unsere obige PS-Definition im engeren Sinn ein wenig "dehnt", weil es bisher noch keine in Mobiltelefonen eingebauten Luftqualitätsmessgeräte gibt; stattdessen nutzt man leichte mobile Sensoren (die man z.B. in ein Arm- oder Stirnband einbauen könnte oder die ein Fahrradkurier an seinem Fahrradrahmen anbringt), die ihre Daten per Bluetooth an das Mobiltelefon zur Weiterleitung übertragen.

- Interessant: Luftqualitätsdaten können auch an den Nutzer zurück gegeben werden, der bspw. bei Warnung vor hohen Ozon- oder Verschmutzungswerten in einem bestimmten Stadtviertel oder entlang einer belebten Verkehrsstraße seine Route ändern könnte.

Lärmkartierung: Viele Projekte betrachten die mobile Sammlung von Umgebungslärmwerten in Städten (What's noisy (http://whatsnoisy.appspot.com/); NoiseSpy (Kanjo, 2009); Ear-Phone (Rana, 2010); NoiseMap (Schweizer et al, 2011); Noisetube (http://noisetube.net/)). Solche Ansätze können die etablierten Methoden (reine Simulationen auf der Basis von Modellen, fest installierte Messnetze oder regelmäßige Mess-Initiativen) mit ihren bekannten Problemen (hohe Kosten, zu grobe Messnetze, zu seltene Aktualisierungen) ergänzen, verbessern oder gar (teilweise) ersetzen. Wissenschaftlich-technisch bietet die Lärmkartierung einige Herausforderungen, wie bspw. die Signalverarbeitung (inkl. Sensorkalibrierung) auf dem Mobilgerät oder auch die vollständige Kartenberechnung auf „dünner" und ungleichmäßiger Datenbasis (Rana, 2010).

Beobachten von Reaktionen der Flora auf das Klima: Das Projekt *BudBurst* (http://neoninc.org/budburst/) ist eine sog. "Bürgerwissenschaftler"-Kampagne („Citizen Science"), um mit der Hilfe von Freiwilligen zu studieren, wie Ereignisse im Lebenszyklus von Pflanzen (Beginn der Blüte u.ä.) zeitlich mit Umweltveränderungen zusammen hängen. Man verspricht sich Erkenntnisse zur Frage, wie Bäume, Sträucher und Gräser auf den globalen Klimawechsel reagieren.

- Budburst macht die Unterscheidung zwischen *regelmäßigen Beobachtungen* bei täglich besuchten Pflanzen und *Einzelbeobachtungen* von Pflanzen, bei denen man nur unregelmäßig vorbei kommt.
- Außerdem wird ein sog. "Serious Game" ("*Floracaching*") angewandt, um spielerisch Anreize zu setzen und bspw. den Fokus auf bestimmte Arten zu schieben (http://networkednaturalist.org/floracaching/).

Dies sind soweit die bekanntesten PS-Anwendungen im Umwelt- und Naturschutz. Im weiteren Umfeld ist auch die *FracTracker* Initiative zu sehen, die Daten und Informationen über die Effekte der unkonventionellen Erdgasförderung („Fracking") in den USA sammelt, insbesondere auch zu potenziellen Umweltschäden. Allerdings wer-

den zur Datenerfassung offenbar bisher noch keine Mobilgeräte verwendet, sondern eine WebGUI, so dass die Initiative eher in der Tradition des „Public Participation GIS“ zu sehen ist. Dies lässt sich verallgemeinern zur Idee, PS beim Finden und Dokumentieren von Umweltschäden (wie Grubensenkungen, wilde Müllkippen, Tierkadaver, kranke Bäume, Gewässerverschmutzungen etc.) zu verwenden. Diese Anwendungsfälle haben wir allerdings bisher in der Literatur noch nicht gefunden.

Eine weitere Verallgemeinerung führt zur Idee, PS beim Finden und Dokumentieren von Schäden oder Wartungsnotwendigkeiten an verteilten Infrastrukturen zu nutzen – für kommerzielle Anwendungen kann das beliebige Infrastrukturen betreffen (Fabrikgelände, Schienennetz), im engeren Umweltbereich eher Dämme oder Deiche, Wander-, Reit- oder Mountainbike-Wege, usw. Dies führt auch zur Idee, PS bei Naturkatastrophen zu nutzen (in der Prävention oder im Schadensfall), wie Hochwasser/Sturmfluten, Waldbrände, Erdbeben, usw. In solchen Fällen können betroffene Bürger oder frewillige Helfer die offiziellen Hilfskräfte durch Sammlung aktueller Situationsberichte dabei unterstützen, Rettungsmaßnahmen zu planen und priorisieren.

Ein weiteres typisches PS-Anwendungsszenario mit *indirekten* Auswirkungen auf die Umwelt betrifft Verkehr und Personentransport. Viele Projekte befassen sich mit der Status- und Problemerfassung auf Straßen und im öffentlichen Nahverkehr, mit der Mustererkennung im persönlichen Pendel- und Mobilitätsverhalten, der effizienten Nutzung freier Parkplätze, Verspätungen, Staus, usw. Diese Projekte können zum Einen stadt- und raumplanerische Aktivitäten (oder auch die Planungen von Mobilitätsanbietern) informieren, zum Anderen aber auch in Echtzeit persönliche Transportentscheidungen unterstützen – so dass es insgesamt zu Ressourceneinsparungen kommt. In einem allgemeinen Umweltbegriff gibt es natürlich auch direkte Verbindungen zum Bereich Gesundheit, der ebenfalls in verschiedenen PS-Anwendungen adressiert wird. Ein prominentes Beispiel ist *MobAsthma* (Kanjo et al, 2009), das es erlaubt, den persönlichen Asthma-Status mit der erfahrenen Luftverschmutzung zu korrelieren. Ganz allgemein, stellen Allergien und Atemwegsprobleme ein interessantes Feld für Participatory Sensing dar (vgl. auch *PEIR – Personal Environmental Impact Report* (Mun et al, 2009)).

4 Eine generische Software-Architektur für Participatory Sensing

Aus einer Anforderungsanalyse und bereits existierenden PS-Architekturen wurde im PartSense-Projekt eine initiale funktionale Software-Architektur abgeleitet (Abbildung 1), deren Moduln wir skizzieren:

Szenario- oder Kampagnen-Definition: Um eine neue Datensammel-Aktion aufzusetzen, werden die wesentlichen Parameter spezifiziert, z.B. welche Art von Daten mit welcher Sensorik in welchem Gebiet und Zeitraum zu sammeln sind.

Teilnehmerverwaltung: In PS-Szenarien mit registrierten Nutzern kann die Benutzerverwaltung unterschiedlich komplex ausgeprägt sein, beginnend mit dem Speichern einfacher technischer Nutzerdaten (vorhandene Hardware und Software) bis hin zu typischen Bewegungsmustern oder Informationen über die Motivationsstruktur. Auch Wissen über frühere Datensammel-Aktivitäten kann hilfreich sein.

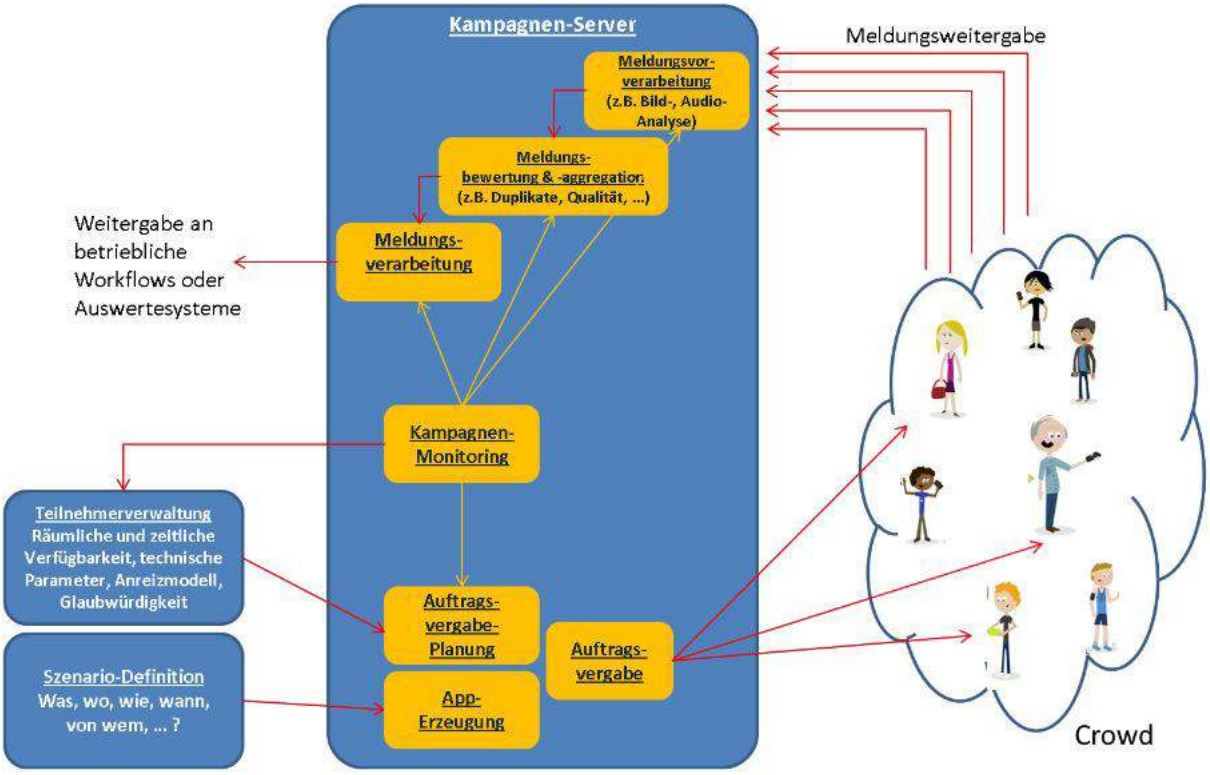


Abbildung 1: Generische Architektur für Participatory Sensing

App-Erzeugung: Aus den Spezifikationen der Kampagnen-Definition sollte die Datensammel-App für die verbreiteten mobilen Betriebssysteme automatisch erzeugt werden.

Auftragsvergabe und Auftragsbvergabe-Planung: Wenn Datensammel-Aufgaben *aktiv* an mögliche Sammler verteilt werden, muss eine voraussichtlich hinreichend

große Anzahl von Adressaten identifiziert und automatisch angesprochen werden. Die Auswahl kann die verschiedensten Kriterien berücksichtigen und z.B. auch erst dann erfolgen, wenn ein Nutzer sich einem interessanten Ort nähert.

Datensammlung und Meldungsweitergabe: Die Datensammlung mit der verteilten App erfolgt entweder mit *expliziter manueller Interaktion* oder *automatisch* durch das Mobilgerät, wenn die App aktiv ist. Meldungen werden normalerweise *sofort* nach Erfassung an den Server geleitet, es gibt aber auch Szenarien mit verzögerter Übermittlung. Bei verzögerter Übermittlung ist auch ein Datentransfer über Medien wie USB-Sticks möglich. Vor dem großen Erfolg internetfähiger mobiler Endgeräte auf dem Massenmarkt war die manuelle Dateneingabe zuhause über eine Web-GUI üblich. Die Datenerfassung selber geschieht durch die eingebaute Sensorik bzw. Eingabemechanismen des Geräts, evtl. ergänzt um preisgünstige und leichtgewichtige mobile Sensoren, die über das Mobiltelefon ihre Daten weitergeben können.

Pipeline zur Meldungsverarbeitung: Abb. 1 zeigt exemplarisch drei Elemente der Pipeline zur Meldungsverarbeitung. Diese Pipeline muss zwar für jeden Anwendungsfall spezifisch konfiguriert werden, kann aber in vielen Fällen auch sehr einfach gehalten werden. Die dargestellten Schritte: (1.) abhängig von der Art der gesammelten Daten kann eine *Vorverarbeitung* erforderlich sein, z.B. zur Analyse eines gesendeten Fotos. (2.) Weitere Schritte zur *Meldungsbewertung und -analyse* können folgen, z.B. zur Duplikaterkennung oder zur *Qualitätsbewertung* (Erkennung von „Spam“ oder Ausreißer-Werten). (3.) Die Meldungen können dann an nachgelagerte *Back-end Systeme* weiter gegeben werden oder mit den eingebauten Mecnhanismen der PS-Plattform verwaltet (Speicherung, Inspektion, Bearbeitung, Visualisierung, Auswertung von Meldungen).

Kampagnen-Monitoring: Der Server kann eine laufende Datensammel-Aktion mit Hinblick auf vorgegebene Kriterien (Menge und Qualität von Meldungen) überwachen und bei Bedarf steuernd eingreifen (z.B. durch Aktivieren zusätzlicher Datensammler, Veränderung der Anreizmechanismen, etc).

5 Implementierung und Projektstatus

Technisch gesehen, startet das PartSense Projekt mit zwei bereits existierenden Entwicklungssträngen, die schrittweise miteinander verwoben werden sollen.

Der erste dieser Startpunkte ist das KA-Feedback System, das seit Frühling 2012 in Karlsruhe öffentlich im experimentellen Operativbetrieb ist. Mit KA-Feedback können Bürger über ihr iPhone zeit- und ortskodierte Text- und Foto-Informationen über eine Reihe vordefinierter Mängel (Gefahrenstellen wie Straßenschäden, überfüllte Müllbehälter, Verkehrsverstöße, Ampelprobleme, defekte Straßenbeleuchtung) im Karlsruher Stadtbild an die Stadtverwaltung weiter geben. KA-Feedback basiert auf einer generativen Software, die – für die spezifischen Parameter einer bestimmten Mängelmeldungsanwendung – eine native iOS App erzeugt, die der Bürger herunter laden kann (Borges et al., 2012). Ähnliche Anwendungen für andere Städte können leicht erzeugt werden.

Der KA-Feedback Server bietet zwei Web-GUIs an: (i) für die allgemeine Öffentlichkeit, in der die gesammelten Meldungen zusammen mit den Antworten und dem Bearbeitungsstatus der Stadtverwaltung als Listen- und als Kartenansicht gezeigt werden; (ii) für den Systemadministrator und die Sachbearbeiter der Stadtverwaltung. Eingehende Meldungen werden auch direkt per e-Mail an den zuständigen Sachbearbeiter weiter geleitet (Abecker et al., 2012; Borges et al, 2012).

KA-Feedback ist (momentan noch) eine iOS-spezifische, sehr einfache Umsetzung der Architekur aus Kapitel 4, speziell ausgelegt für den Anwendungsfall der Mängelmeldung. Es besitzt keine Benutzerverwaltung, keine Aufgabenverteilungsplanung (weil jeder Bürger anonym und ohne Registrierung die App herunter laden kann) und kein ausgeprägtes Kampagnen-Monitoring. KA-Feedback bietet eine anwendungsfall-spezifische GUI für die Kampagnen-Definition, auf deren Basis die iPhone-App automatisch erzeugt wird. Die Meldungsverarbeitung ist sehr einfach: Speichern von Meldungen, Anzeige in den zwei GUIs, Weiterleitung per e-Mail an die zuständigen Sachbearbeiter, einfache Auswertungen für den Kampagnen-Administrator.

Der zweite Startpunkt ist die im Mai 2012 veröffentlichte mobile GIS Software “Cadenza Mobile” der disy GmbH (Hofmann et al., 2012). Sie erlaubt die einfache Erfassung, Nutzung und Änderung von GIS-Daten auf Mobilgeräten (typischerweise auf sog. Media Tablets). Durch die Integration von “Cadenza Mobile” im PartSense-Projekt soll der Rahmen möglicher PS-Anwendungen von einfachen Datensammelaufgaben für die allgemeine Öffentlichkeit auf mehr professionelle Fachanwendungen für Firmen oder die öffentliche Verwaltung ausgedehnt werden. “Cadenza

Mobile" unterstützt Eingabe und Editieren komplexer Fachanwendungsobjekte, die mit Linien- oder Polygon-Geometrien verbunden sind (anstelle rein punkthaften Ortsbezuges) und in der disy-Software "Cadenza Professional" zum räumlichen Datenmanagement und –reporting gespeichert sind. "Cadenza Mobile" ermöglicht (1) Export von Teilen der Cadenza Sach- und GIS-Daten und Übertragung auf ein mobiles Endgerät; (2) Verarbeitung dieser Daten im Offline-Modus auf dem Mobilgerät; (3) Zurückspeichern der Daten ins Desktop-System. Die Möglichkeit zur Offline-Bearbeitung ist wichtig, weil wir in vielen Outdoor-Szenarien (Forst- oder Landwirtschaft, Regionalplanung, Flurneuordnung, Straßenbau, …) nicht davon ausgehen können, im Gelände jederzeit und überall einen hoch performanten drahtlosen Internet-Zugang vorzufinden. Im Gegensatz zur einfachen Nutzung eines WebGIS im mobilen Browser bietet "Cadenza Mobile" exakt dieselbe Kartendarstellung von Fachkarten im mobilen und im Desktop-Arbeitsmodus und es nutzt alle gewohnten Funktionen der Multitouch Gestensteuerung mobiler Endgeräte. Abbildung 2 stellt in einem Screenshot Cadenza Mobile auf Media Tablet dar.



Abbildung 2: Screenshot Cadenza Mobile auf Media Tablet

Bisher implementiert “Cadenza Mobile” noch keine PS-spezifischen Funktionalitäten, sondern hat eher die Grundlagen dafür geschaffen, PS-Ideen auf schwergewichtigere GIS-Fachanwendungen zu übertragen. Außerdem wurden technische Fragestellungen untersucht, die für eine sinnvolle vollständige Umsetzung der PartSense-Ziele wichtig sind, wie insbesondere die plattform-übergreifende Entwicklung. Der Kern der “Cadenza Mobile” App-Entwickung ist plattform-neutral und mit HTML5 / Javascript sowie dem Phonegap / Apache Cordova Framework umgesetzt worden. Phonegap erlaubt die Erzeugung plattform-spezifischer Apps (zurzeit arbeiten wir mit Apple-iOS and Google-Android). Funktionen, die nicht einfach mit HTML5 umgesetzt werden können, kann man als nativen Code für die jeweiligen Systeme entwickeln und (wieder plattform-neutral) dem HTML5-Kern über den Phonegap Abstraktionslayer zur Verfügung stellen. Dieser Mechanismus ermöglicht eine ökonomische App-Entwicklung für unterschiedliche mobile Betriebssysteme. Weitere Technologien, die in der Entwicklung von “Cadenza

Mobile" erfolgreich eingesetzt wurden, sind das Senchatouch GUI-Framework für mobile Web-Anwendungen und das Web-Mapping Framework Open Layers.

6 Zusammenfassung

Wir haben das viel versprechende Thema Participatory Sensing vorgestellt, mit einem ersten umfassenden Definitionsversuch, der Vorstellung einer einfachen generischen Architektur für die Umsetzung und einer Sammlung von Anwendungsmöglichkeiten im Umwelt- und Naturschutz. Unsere Implementierungsarbeiten kombinieren das KA-Feedback System, eine einfache, aber funktionsfähige Mängelmelder-Lösung, mit Arbeiten aus disy "Cadenza Mobile", welches PS-Szenarien auf komplexere Fachanwendungen mit komplizierterem Raumbezug ausweiten soll. Dabei liegt der aktuelle Arbeitsschwerpunkt auf Offline-Nutzung und Experimentieren mit Entwicklungsstacks für die Crossplatform-Entwicklung.

Spätere Forschungsarbeiten (insbesondere auch experimenteller Natur mit Pilotanwendungen) sollen die vermutlich wichtigsten Themen für erfolgreichere PS-Anwendungen zumindest teilweise adressieren: (i) Datenschutz und Datensicherheit beim Umgang mit persönlichen Daten (z.B. bei der Auswertung von Bewegungsdaten; vgl. (Christin et al., 2011)); (ii) Qualität und Quantität der gesammelten Daten – was eng verwandt ist mit der Frage nach Anreizsystemen zur Sicherstellung der Nutzerbeteiligung.

Danksagung: Die vorgestellten Arbeiten werden im Projekt "PartSense – Participatory Sensing für Firmen und die öffentliche Verwaltung" vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen des Programms "KMU Innovativ" gefördert (FKZ 01IS11029A).

7 Literaturverzeichnis

[Abecker, et al, 2012]

Abecker, A., Braun, S., Valikov, A., Zacharias, V.: Towards a Technology for Participatory Sensing Applications. In: P. & M. Cunningham (Hrsg.): eChallenges 2012. In Vorbereitung.

[Borges, et al, 2012]

Borges, J. de Melo, Zacharias, V., Plessing, N.: PartSense - A Participatory Sensing Platform and its Instantiation KA-Feedback. In: M.-O. Löwner et al. (Hrsg.): GEO-INFORMATIK 2012 – Mobilität und Umwelt, pp. 277-284, Shaker Verlag, Aachen, 2012.

[Christin, et al, 2011]

Christin, D., Reinhardt, A., Kanhere, S.S., Hollick, M.: A Survey on Privacy in Mobile Participatory Sensing Applications, Journal of Systems and Software 84(11):1928-1946, Elsevier, 2011.

[Goldman, et al, 2009]

Goldman, J., Shilton, K., Burke, J., Estrin, D., Hansen, M., Ramanthan, N., Reddy, S., Samanta, V., Srivastava, M., West, R.: Participatory Sensing: A Citizen-Powered Approach to Illuminating The Patterns That Shape Our World, Woodrow Wilson Center for International Scholars, May 2009.

[Hofmann, et al, 2012]

Hofmann, C., Briesen, M., Schmidt, K.: Cadenza Mobile – Geo- und Fachdaten mobil nutzen. In: K. Weissenbach et al. (Hrsg.): F+E-Vorhaben MAF-UIS - Phase I 2011/12, KIT Scientific Reports No. 7616, Karlsruhe, 2012.

[Kanjo, 2009]

Kanjo, E., Bacon, J., Roberts, D., Landshoff, P.: MobSens - Making Smart Phones Smarter, IEEE Pervasive Computing 8(4):50-57, 2009.

[Mun, et al., 2009]

Mun, M., Reddy, S., Shilton, K., Yau, N., Burke, J., Estrin, D., Hansen, M., Howard, E., West, R., Boda, P.: PEIR, The Personal Environmental Impact Report, as a Platform for Participatory Sensing Systems Research. In: MobiSys '09, S. 55-68, ACM, 2009.

[Paulos, 2008]

Paulos, E., Honicky, R.J., Hooker, B.: Citizen Science: Enabling Participatory Urbanism. In: M. Foth (Hrsg.): Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City, IGI Global, Hershey, 2008.

[Rana, et al, 2010]

Rana, R.K., Chou, C.T., Kanhere, S.S., Bulusu, N., Hu, W.: Ear-phone: An End-to-End Participatory Urban Noise Mapping System,. In 9th ACM/IEEE Int. Conference on Information Processing in Sensor Networks (IPSN '10), S. 105-116, ACM, 2010.

[Schweizer, et al, 2011]

Schweizer, I., Bärtl, R., Schulz, A., Probst, F., Mühläuser, M.: NoiseMap - Real-time Participatory Noise Maps. In: Second International Workshop on Sensing Applications on Mobile Phones, 2011.

PortalU und Open Data: Möglichkeiten der Darstellung und Abfrage von Umweltinformationen

Stefanie Haß, Koordinierungsstelle PortalU, stefanie.hass@portalu.de
Fred Kruse, Koordinierungsstelle PortalU, fred.kruse@portalu.de

Abstract / Einleitung

Since the introduction of the Environmental Information Portal, PortalU (www.portalu.de) in 2006 the Portal publishes environmental information from public agencies. Visitors to the portal can search environmental web pages, documents, maps (WMS) and metadata, published by more the 450 providers all over Germany. Environmental data in open data portals are requested frequently and play an important role in developing applications for people generally interested in this field. PortalU sets standards in publishing and searching environmental data. Large sets of metadata are collected on the basis of ISO-conform data catalogues. The link between the metadata and the published information in the PortalU is a key element for improving the presentation and content of information in the portal. Further aspects in optimizing the information retrieval is the analysis of user search log files, which allows identifying certain search pattern. With these findings the PortalU search machine can be adapted to user requirements. The coordination centre PortalU (Koordinierungsstelle PortalU) plans to expand the portal service portfolio by testing and applying those methods.

1 Verfügbarkeit von Umweltinformationen in PortalU

PortalU bietet derzeit Zugriff auf rund 3 Mio Umweltinformationen in Form von Webseiten, Datenbankeinträgen, Dokumenten, WMS-Diensten und Metadaten, die dezentral in rund 450 öffentlichen Institutionen und Organisationen verwaltet werden. Der größte Teil der Umweltinformationen stammt aus Inhalten von Webseiten, die täglich von der PortalU eigenen Umweltsuchmaschine indexiert werden. Mittels einer

Suchmaske können diese Inhalte recherchiert und die Suchresultate nach eigenen Kriterien gefiltert werden. Seit den frühen neunziger Jahren werden Umweltdaten von öffentlichen Verwaltungen mittels Umweltdatenkatalogen (UDK) beschrieben. In PortalU kann über standardisierte Schnittstellen auf die Umweltdatenkataloge und die darin enthaltenen Metadaten zugegriffen werden. Die Erfassung der Metadaten basiert auf einem einheitlichen ISO 19115- und ISO 19119-konformen Datenmodell, welches sowohl die Beschreibung von Karten als auch Webdiensten ermöglicht. Als besonderes Charakteristikum des InGrid-Catalogs können aber auch Organisationseinheiten, Dokumente, Projekte und Datensammlungen beschrieben werden. Dadurch ist bereits heute eine Vielzahl behördlicher Umweltinformation in Form von Metadaten, Dokumenten und Webseiten über PortalU frei zugänglich. Nicht zuletzt im Zuge der Open Data Diskussionen, beginnen öffentliche Institutionen nach und nach auch Umweltdaten im Sinne von Rohdaten oder aggregierten Daten bereitzustellen. PortalU bietet mit seinem integrierten Datenkatalog (InGrid-Catalog) bereits heute die Technik, auf downloadbare Daten zu verweisen. Innerhalb der Objektklassen des InGrid-Catalogs ist es möglich, Verweise auf verfügbare Dokumente, Geo- und andere Daten zu setzen. Besucher können dadurch frei verfügbare Datensätze in PortalU herunterladen. Künftig wird dieses zusätzliche Feature in PortalU noch stärker hervorgehoben werden.

2 Abfrage und Darstellung von Umweltinformationen in PortalU

2.1 Umweltinformationen aus Nutzersicht

Aus Nutzersicht gibt es zwei Kategorien von verfügbaren Informationen in PortalU. Zum einen Umweltinformationen, die direkt zugänglich und für eigene Bedürfnisse nutzbar sind. Dazu zählen Informationen in schriftlicher und digitaler Form (wie beispielsweise relevante Webseiten oder Dokumente) ebenso wie Geo- oder Messdaten, die für den eigenen Informationsbedarf oder eigene Anwendungen direkt genutzt und gegebenfalls weiterverarbeitet werden können. Zum anderen zählen zum Informationsangebot Metadaten, die häufig keinen direkten Zugang zu den Datenquellen ermöglichen, sondern eher „Auskunftscharakter“ haben: Welche Umweltdaten sind in welcher Form, in welcher Qualität und zu welchen Konditionen in einer Behörde verfügbar? Bisher wurden verfügbare Informationen beider Katego-

rien in PortalU als Resultat von Suchanfragen in Trefferlisten dargestellt. Insbesondere die wachsende Nachfrage unterschiedlichster Zielgruppen an frei verfügbaren Daten stellt PortalU vor neue Herausforderungen in puncto Recherche und Darstellung von Umweltinformationen: Während Fachanwender / Entwickler aus Wissenschaft oder Wirtschaft spezifische Anforderungen an Daten, wie beispielsweise maschinelle Lesbarkeit, Vollständigkeit und die Verfügbarkeit von Rohdaten haben, spielen diese Kriterien für den interessierten Bürger eher eine untergeordnete Rolle. Hingegen werden aufbereitete Daten, Apps oder von Usern favorisierte Datensätze von der Öffentlichkeit häufig nachgefragt und in gängigen Open Data Portalen auch an zentraler Stelle positioniert.

2.2 Technische Umsetzung in PortalU

Die Koordinierungsstelle verfolgt auf technischer Ebene derzeit zwei Ansätze. Unter dem Stichwort Daten-Dienste-Kopplung sollen künftig in PortalU Metadaten aus Umweltdatenkatalogen auch an Web Map Services gekoppelt werden. Darüber hinaus ist geplant, Vorschaubilder (thumbnails) der Web Map Services in der Trefferliste von PortalU anzuzeigen.

Diskutiert wird derzeit auch, ob der direkte Zugriff auf Daten über Links aus den Metadaten heraus prominenter dargestellt werden kann.

Als Proof-Of-Concept gilt derzeit das Vorhaben, mobile Applikationen mit Webdiensten zu kombinieren. Ziel ist es, PortalU-Inhalte mobilen Endgeräten zugänglich zu machen.

2.3 Inhaltliche Weiterentwicklung

Welche Umweltinformationen/-daten unterschiedliche Zielgruppen wünschen, diese Frage gewinnt im Zuge der Open Government / Open Data-Diskussionen immer mehr an Bedeutung. Und auch die zunehmende „Öffnung“ behördlicher Institutionen wird in den kommenden Jahren zu einem breiten Spektrum verfügbarer Daten führen. Daraus ergeben sich auch spezifische Anforderungen an das Informationsangebot und die Informationspräsentation in PortalU. So könnten beispielsweise themenorientiertes Browsen und personalisierte Informationsangebote die bisherige suchmaschinenkonzentrierte Gestaltung von PortalU ergänzen. Die Informationspräsentation nimmt in der konzeptionellen Entwicklung von PortalU eine zentrale

Stelle ein. Eine Pilotumfrage unter den PortalU Nutzern lieferten erste wichtige Erkenntnisse, welche Angebote erwünscht sind und welche Anforderungen es an die Ergebnispräsentation gibt. Daraus resultiert, dass insbesondere die Kartenkomponente, Umweltereignisse, Publikationen und Daten nachgefragt werden. In Planung sind weitere zielgruppenorientierte Befragungen, mit dem Ziel, Umweltinformationen bedarfsgerecht bereitzustellen. Darüber hinaus fließen erste Auswertung von Suchmustern auf Basis von Log-Files in die Weiterentwicklung der Suchmaschine ein.

Die Koordinierungsstelle plant durch die genannten Beispiele der Daten-Dienste-Kopplung und Suchmaschinenoptimierung die Informationspräsentation und die Sichtbarkeit nutzerorientierter Angebote in PortalU weiter zu optimieren.

Digitale Biodiversitätsatlanten – Tools zur Unterstützung regionaler und internationaler Expertennetzwerke für die Erfassung und Bewertung biologischer Diversität

Heiko Brunken, Hochschule Bremen, heiko.brunken@hs-bremen.de
Carl-Heinz Genzel, Hochschule Bremen, calle_genzel@web.de
Heide-Rose Vatterrott, Hochschule Bremen, heide-rose.vatterrott@hs-bremen.de
Martin Winkler, Hochschule Bremen, mawinkler@stud.hs-bremen.de

Abstract / Einleitung

The paper describes a framework for the interactive implementation of Web-GIS applications by biologists and interested non-professionals. The framework is developed by biologists and computer scientists of the University of Applied Sciences Bremen in order to support for gathering, management and visualization of information about biodiversity.

Der Beitrag stellt ein Framework zur online-Unterstützung der Erfassung und Bewertung der Biodiversität durch Experten und interessierte Laien vor, das im Rahmen einer Kooperation zwischen Biologen und Informatikern der Hochschule Bremen entstanden ist.

1 Biologische Motivation

Die biologische Vielfalt nimmt weltweit und auch in Deutschland stark ab. Die derzeitige Aussterberate übertrifft die vermutete natürliche Rate um das 100 – 1000fache und ist durch menschliches Handeln bedingt. Die Völkergemeinschaft hat erkannt, dass gegen den Schwund der biologischen Vielfalt energisch gegengesteuert werden muss. Die Vereinten Nationen haben daher die internationale UN-Dekade Biologische Vielfalt von 2011 bis 2020 ausgerufen (ZUK 2012). Eine grundlegende Voraussetzung zum Schutz der biologischen Vielfalt sind genaue Informationen über die Biologie und die Verbreitung von Arten.

Bei der Erfassung und Kommunikation von faunistischen und floristischen Daten zeichnen sich in Deutschland derzeit divergierende Trends ab. Bestimmte Artengruppen und ausgewählte Gebiete werden auf Grund EU-rahmenrechtlicher Vorgaben (insbesondere Fauna-Flora-Habitat-Richtlinie, Wasserrahmenrichtlinie) von behördlicher Seite oft mit professionellem Aufwand (meist von spezialisierten Planungsbüros) erfasst, ausgewertet und über GIS-Systeme direkt an die EU gemeldet. Das wissenschaftliche Arbeiten an Hochschulen und Forschungseinrichtungen orientiert sich heute aus verschiedenen Gründen (z.B. Evaluation von Forschungsergebnissen) an hoch spezialisierten Einzelthemen, Publikationen erfolgen überwiegend in Form von isolierten Fachbeiträgen. Klassische monographische Bearbeitungen z.B. über bestimmte Arten oder Artengruppen oder qualifizierte Gebietsbeschreibungen (Faunen- und Florenwerke) treten zunehmend in den Hintergrund. Somit ist im angewandten Naturschutz paradoxerweise eine Situation entstanden, in der es, trotz der allgemeinen Erkenntnis zur Notwendigkeit der Erfassung von Biodiversitätsdaten und vermehrten behördlichen Anstrengungen, einen großen Mangel an der Erfassung und Aktualisierung von faunistischen und floristischen Verbreitungsdaten gibt. In dieser Situation kommt der verstärkten Einbindung des ehrenamtlichen Elementes (wieder) eine zunehmende Bedeutung zu. Dieses kann heute im Zeitalter des WEB 2.0 auch durch interaktive Online-Werkzeuge und Publikationsplattformen im Internet erfolgen.

2 Anforderungen an die technische Unterstützung

Nachdem im Rahmen wissenschaftlicher Arbeiten von Biologen der Hochschule Bremen umfangreiches Datenmaterial über unterschiedliche Spezies und deren Verbreitung in unterschiedlichen Systemen, z.B. [Brunken & Brunschön, 2006; Brunken et al., 2011; Moura et al., 2011] zusammen getragen worden war, hatten sich eine Reihe von Anforderungen an die Entwicklung eines Frameworks herauskristallisiert.

Das zu entwickelnde Framework sollte es einerseits gestatten, vielfältige fachwissenschaftliche Informationen in Form von Texten, Bildern, u.a. zu unterschiedlichen Spezies zu verwalten und diese für Fachwissenschaftler leicht zugänglich sowie änder- und erweiterbar anzubieten.

Andererseits sollte es möglich sein, Daten über die Verbreitung der Spezies aus unterschiedlichen Quellen (amtliche Erfassung, Auswertung wissenschaftlicher Untersuchungen, Meldungen interessierter Laien) in einem System zusammenzuführen und unter Wahrung arten- und datenschutzrechtlicher Auflagen für unterschiedliche Auswertungszwecke online zur Verfügung zu stellen. Die Integration zusätzlicher Verbreitungsinformationen sollte durch Experten der Biologie komfortabel realisierbar sein. Darüber hinaus waren intuitiv bedienbare Lösungen für raumbezogene Anfragen gewünscht, wobei Kartenmaterial unterschiedlicher Anbieter zur Visualisierung genutzt werden sollte.

Zur Unterstützung vielfältiger internationaler Forschungskontakte sollte es möglich sein, sowohl die öffentliche Präsentation der Informationen als auch die Pflege der Daten durch Fachwissenschaftler und Studierende der Biologie in geeigneten Webanwendungen mehrsprachig zu gestalten, wobei neben Deutsch und Englisch brasilianisches Portugiesisch Berücksichtigung finden sollte.

Die Lösung sollte zukunftssicher sein, aktuellen Trends der Softwareentwicklung gerecht werden und ausschließlich mit kostenlos verfügbarer Software realisiert werden. Die Portierung der vorhandenen umfangreichen Datenbestände in das neue System sollte weitgehend automatisiert erfolgen können.

3 Das entwickelte Framework

Da umfangreiche Teile des vorhandenen Datenbestandes des Fischartenatlasses für Deutschland und Österreich [Brunken et al., 2011] mit Hilfe des Content-Management-Systems (CMS) Joomla! [Joomla, 2011] verwaltet wurden und die Experten der Biologie dieses System hinsichtlich seiner Benutzungsfreundlichkeit als geeignet charakterisierten, ergab sich die Herausforderung, Joomla! um geeignete Komponenten für die Verwaltung und Präsentation von artspezifischen und Geodaten zu erweitern. Dabei sollten die speziellen Anforderungen unterschiedlicher Nutzergruppen Berücksichtigung finden.

Das CMS Joomla! bietet Unterstützung für die Realisierung von Benutzungsoberflächen für unterschiedliche Nutzergruppen in Form einer Unterscheidung von Backend und Frontend-Komponente an. Dieses Konzept wird im Rahmen des entwickelten Frameworks (Abb. 1) wie folgt genutzt:

Nutzer, die sich über das Angebot des Systems informieren möchten, haben als Gäste unbeschränkten Zugang zum Frontend.

Nutzern, die darüber hinaus mit zusätzlichen Informationen zur Weiterentwicklung des Systems beitragen möchten, wird nach Anmeldung zusätzliche Funktionalität im Frontend zur Verfügung gestellt.

Fachexperten der Biologie, die den Datenbestand pflegen und eingegangene Informationen auf ihre Qualität prüfen, erhalten administrativen Zugang zu Teilen des Front- und Backend.

Joomla!-Administratoren erhalten uneingeschränkten Zugang zu Front- und Backend.

Primär bedingt durch die derzeitige Systemarchitektur von Joomla! wurde als Datenbankbetriebssystem für die Verwaltung aller Daten MySQL [MySQL, 2011] gewählt.

Als Konsequenz dieser Entscheidung war zusätzlich Funktionalität zur Realisierung geographischer Projektionen in das System zu integrieren, wobei die Wahl auf die Open-Source-Bibliothek Proj4js [Proj4js, 2011] fiel, die im Rahmen eines eigens entwickelten Plugins in Joomla! integriert wurde.

Die Kombinationen der im Rahmen des Frameworks verwalteten Verbreitungsdaten mit Kartenmaterial unterschiedlicher Anbieter erfolgt mit Hilfe der Web-GIS Bibliothek Open Layers [OL, 2011].

Vergleichsweise geringer Aufwand war bei der Realisierung mehrsprachiger Benutzungsoberflächen zu investieren, da Joomla! Standardkomponenten für die Realisierung mehrsprachiger Benutzungsoberflächen anbietet, die im Rahmen des Frameworks in vielen Fällen zum Einsatz kommen konnten.

Der Kern des hier charakterisierten Frameworks wurde im Rahmen einer Bachelor-Arbeit [Genzel, 2011] realisiert.

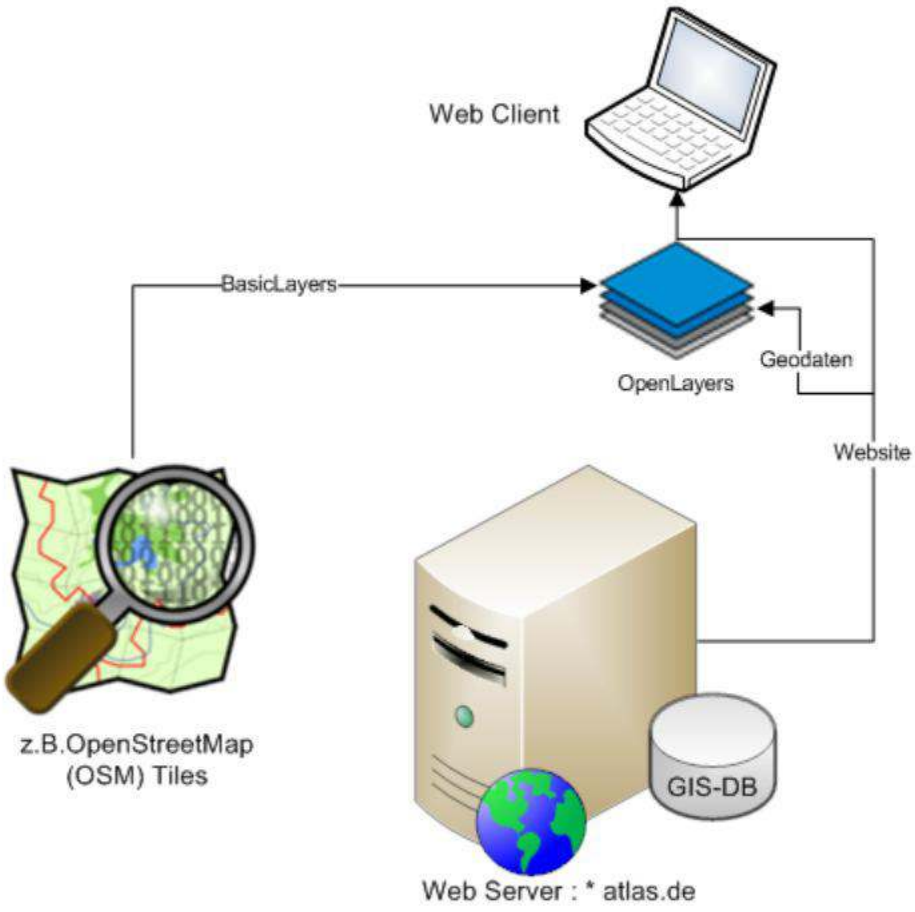


Abbildung 2: Architektur des aktuellen Frameworks für Biodiversitätsatlanten

4 Das Framework in der Praxis

Das Framework dient zur Reimplementation des an der Hochschule Bremen entwickelten „Fischartenatlas von Deutschland und Österreich“ [Brunken et al., 2011] und wird darüber hinaus für die Entwicklung von Biodiversitätsatlanten für andere Artengruppen (u.a. Säugetiere von Nordwestdeutschland, Reptilien und Amphibien von Nordostbrasilien) eingesetzt. Mit dem Atlas der Amphibien und Reptilien (Herpetofauna) von Pernambuco (pt.: *Atlas da Herpetofauna do Estado de Pernambuco*) wird im Rahmen einer Kooperation zwischen der Hochschule Bremen und der Landwirtschaftuniversität in Recife (Universidade Federal Rural de Pernambuco), Pernambuco, Nordostbrasilien, in einem gemeinsamen Forschungs- und Entwicklungsprojekt eine erste Anwendung des neuen Frameworks realisiert und

getestet (Moura et al., 2011). Abbildung 3 zeigt einen Ausschnitt aus dem Frontend des derzeit im Aufbau befindlichen Atlanten. Dargestellt ist die Artseite der Echse Schwarzweißer Teju (*Tupinambis merianae*), die neben Informationen zur Art eine Karte mit Markierungen von Fundorten enthält. Beim „Mouse-over" über einen Verbreitungspunkt werden zusätzliche Kurzinformationen zum Ort, zum Nachweisdatum, zur Autorenschaft des Nachweises sowie ein Kurzkommentar angezeigt. Abbildung 4 veranschaulicht die interaktive Eingabe von Informationen über den Nachweisort einer Art. Der genaue Fundort kann durch numerische Eingabe der Koordinaten (Formular links) oder durch Anklicken auf der Karte bestimmt werden. Nach Bestätigung durch den Nutzer werden die Koordinatenangaben zusammen mit Informationen über das Nachweisdatum, die Datenherkunft und die Datenqualität in das System übernommen.

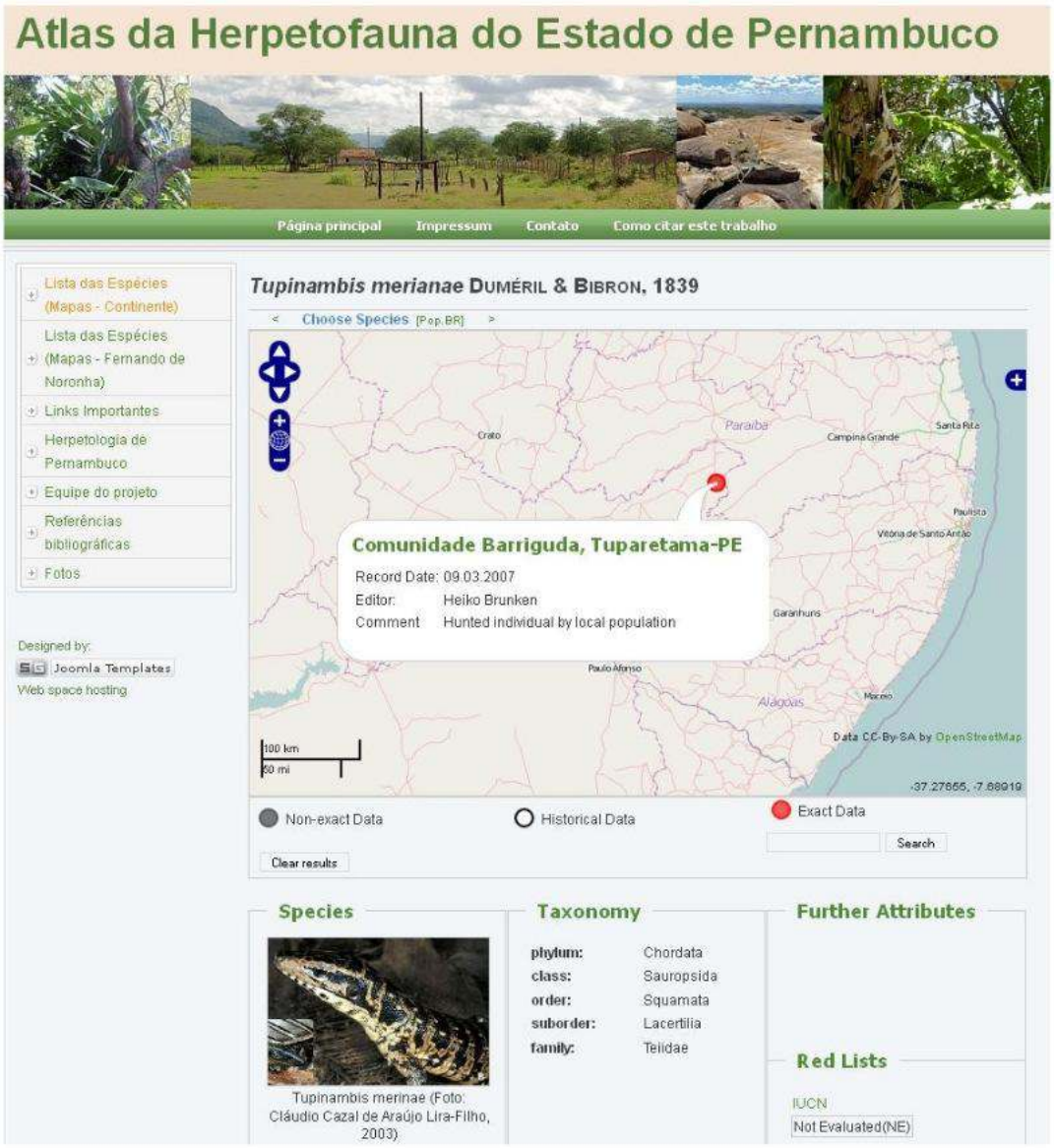


Abbildung 3: Frontend, Beispiel: Atlas da Herpetofauna do Estado de Pernambuco

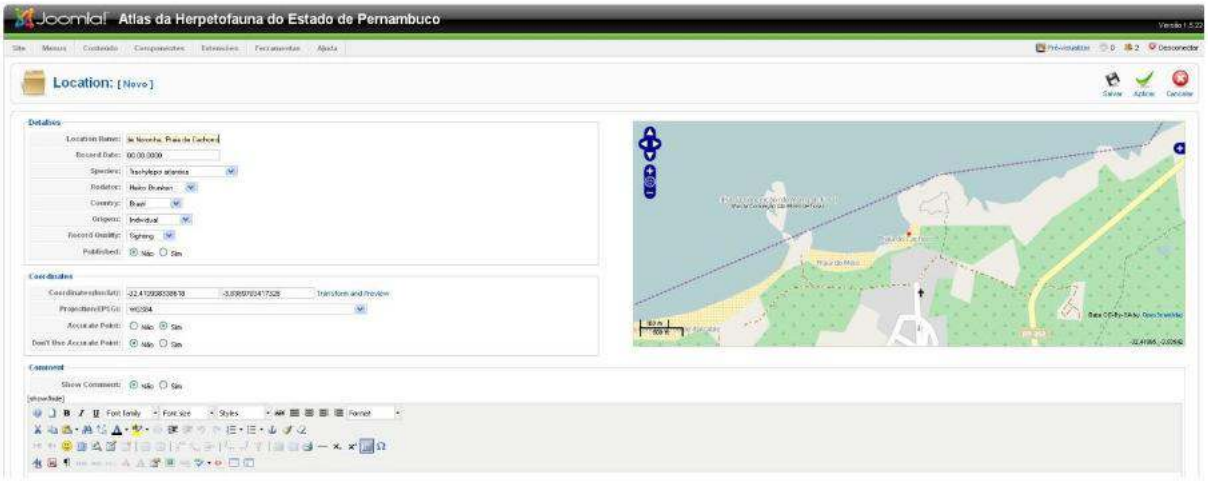


Abbildung 4: Backend, Beispiel: Atlas da Herpetofauna do Estado de Pernambuco

5 Ausblick

Um den zunehmenden Anforderungen an das Framework, speziell die Möglichkeiten von GIS-Anfragen betreffend, perspektivisch gerecht werden zu können, sollte an den Einsatz eines Web-Map-Servers (WMS) sowie eines Web-Feature-Servers (WFS) gedacht werden. Hinzu kommt die Entwicklung von Werkzeugen für die mobile Eingabe über Android-Smartphones.

Mit dem derzeit vorliegenden Framework und seinen geplanten Weiterentwicklungen steht ein technisches Hilfsmittel zur Verfügung, das geeignet ist, den Informationsaustausch zwischen Fachwissenschaftlern der Biologie sowie die Kommunikation mit interessierten Laien zu unterstützen. Es wird zukünftig den Aufbau und die Pflege von biologischen Expertennetzwerken fördern.

In Ländern, die derzeit noch über keine Systeme der behördlichen Umweltdatenverwaltung (z.B. staatliche Tierartenerfassungsprogramme) verfügen, kann das hier entwickelte, einfach einzusetzende und vollständig auf open-source-Software basierende System kurzfristig einen bedeutenden Beitrag zur Erfassung und Kommunikation von Biodiversität und damit zu deren Erhalt beitragen.

6 Literaturverzeichnis

[Brunken et al., 2006]

Brunken, H. & Brunschön, C.: Digitaler Fischartenatlas von Deutschland, eine Projektbeschreibung. - Verhandlungen der Gesellschaft für Ichthyologie 5: 27-34.

[Brunken et al., 2011]

Brunken, H., Brunschön, C., Sperling, M. & Winkler, M.: Digitaler Fischartenatlas von Deutschland und Österreich. Eine ichthyologische Informations- und Kommunikationsplattform. - Hrsg. Gesellschaft für Ichthyologie e.V. World Wide Web electronic publication. http://www.fischartenatlas.de

[Genzel, 2011]

Genzel, C.-H.: Web GIS Erweiterung des CMS Joomla! im Kontext der Biodiversität. – Bachelorthesis an der Fakultät Elektrotechnik und Informatik, Hochschule Bremen.

[Joomla, 2011]

Joomla: Home-Page des CMS Joomla http://www.joomla.de/

[Moura, et al. 2011]

Moura, G.B.M., Santos, E.M.S., Brockmann, O., Genzel, C.H., Steinhausen, K., Vatterrott, H.R., Winkler, M. & Brunken, H.: Atlas da Herpetofauna do Estado de Pernambuco. World Wide Web electronic publication. http://herpetofauna.biodiversidade-pe.com/

[MySQL, 2011]

MySQL: Home-Page des DBMS MySQL http://www.mysql.de/

[OL, 2011]

OL: Home-Page Open Layers http://openlayers.org/

[Proj4js, 2011]

Proj4js:Home-Page der Bibliothek Proj4js http://trac.osgeo.org/proj4js/

[ZUK, 2012]

ZUK: Homepage des Zentrums für Umweltkommunikation (ZUK) http://www.dbu.de/336.html