

Offene Geodaten – mehr Transparenz durch ein Portal der Open Data Portale

Matthias Hinz, Ralf Bill

Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät
Professur für Geodäsie und Geoinformatik
matthias.hinz@uni-rostock.de; ralf.bill@uni-rostock.de

Abstract. In den Bereichen der öffentlichen Verwaltung, Wirtschaft und Wissenschaft werden Angebote von offenen Daten und Services bereits seit Jahren gefördert und erweitert. Aufgrund technischer und organisatorischer Hemmnisse ist es für Anwender aus dem Geodatenbereich jedoch schwierig, sich umfassend über die für ein Vorhaben relevanten Datenquellen zu informieren. Im Rahmen des Projektes OpenGeoEdu wurde nun ein Einstiegspunkt zu Geodaten- und Open Data Portalen erstellt, welche über ein Web-Portal abgerufen, gesucht und visualisiert werden können. Das Portal gibt einen Überblick, wo offene Daten angeboten werden und kann somit für die systematische Datenrecherche verwendet werden.

1 Einleitung

Gegenwärtig gibt es eine Vielzahl vorhandener Daten, sei dies aus dem Umfeld der offenen Daten, die i. d. R. kostenfrei nutzbar sind, oder dem Bereich der amtlichen oder kommerziellen Daten, deren Nutzung i. d. R. nicht kostenfrei ist bzw. teilweise Einschränkungen in der Verwendung unterliegen. Alle diese Daten sind sowohl für den Bürger, die Verwaltung und Wirtschaft als auch die Wissenschaft von hohem Interesse und Nutzwert, unterliegen jedoch unterschiedlichsten Nutzungsbedingungen, Kosten- und Geschäftsmodellen. Während einzelne Datenquellen oftmals gut erschließbar sind und sich Nutzungs- und Geschäftsmodelle für Wirtschaft, Verwaltung und Wissenschaft unter klaren Rahmenbedingungen entwickeln lassen, stehen der umfassenden Nutzung und intelligenten Vernetzung dieser Daten größere technische und organisatorische Hemmnisse im Wege.

Um Nutzern einen Zugang zu bekannten Open Data Quellen und Geodateninfrastrukturen (GDI) zu ermöglichen, wird im Projekt OpenGeoEdu (Offene Daten für Lehre und Forschung in raumbezogenen Studiengängen) ein Verzeichnis von relevanten Websites und Plattformen mit Bezug auf den deutschsprachigen Raum entwickelt. Insbesondere für den Bereich Open Government Data (OGD) gibt es bereits Zusammenstellungen nationaler und internationaler Datenangebote aus verschiedenen Quellen und Verwaltungsebenen zu umfassenden Katalogen, jedoch erschöpfen sich damit nicht die Möglichkeiten der Datenrecherche. Ein Großteil der Kommunen und Regionalverbände betreiben eigene Daten- und Geoportale, die nicht immer in die nationalen Kataloge eingebunden sind und in jedem Fall bei der Suche nach offenen Daten einbezogen werden sollten. Auch außerhalb der Open Data- und Transparenzportale stellen Webseiten von Statistik- und Umweltbehörden sowie Geoportale öffentliche Daten und Services bereit. Diese unterliegen zwar nicht immer einer Open Data Lizenz, können oftmals aber unter Beachtung der jeweiligen Lizenz-Bestimmungen kostenlos genutzt werden (Zscheile 2017). Das Konzept von Open (Geo-) Data umfasst nicht nur offene Verwaltungsdaten, sondern auch Unternehmens- und Forschungsdaten sowie Daten, die von Vereinen und Bürgerinitiativen (im Sinne von Citizen Science) offen verfügbar gemacht werden.

Ziel dieser Arbeit ist es, die genannten Datenangebote in ihrer Gesamtheit systematisch zu erfassen, zu beschreiben und darzustellen, um Anwendern auf der Suche nach offenen Geodaten einen zentralen Einstiegspunkt und Orientierung zu bieten. Das hier vorgestellte Portal der Open Data Portale (<https://portal.opengeoedu.de>) bietet umfangreiche Such- und Visualisierungsfunktionen, umfassende Metadaten und Erweiterungsfunktionen, mithilfe derer es zum One-Stop-Portal für offene Geodaten in Deutschland, Österreich und der Schweiz avancieren soll.

2 Angebotsübersichten und Kataloge zu offenen Daten

Auf internationaler Ebene gibt es bereits vergleichbare Übersichten. So erstellte das Unternehmen OpenDataSoft eine erweiterbare Web-Karte von über 2.600 Open Data Portalen weltweit (<https://opendatainception.io/>), welche zum Teil auch in das Verzeichnis von OpenGeoEdu eingeflossen sind. Die Non-profit-Organisation CTIC stellt mit dem Public Dataset Catalogs Facets Browser eine ähnliche Plattform bereit, welche auf den Prinzipien von Linked Open Data (LOD) und Semantic-Web-Technologien aufbaut (CTIC-CT) (<http://datos.fundacionctic.org/sandbox/catalog/faceted/>). Diese Projekte sind

wesentlich ambitionierter hinsichtlich des Umfangs der zu erfassenden Daten, jedoch bedürfte es einer weltweiten, aktiven Nutzergemeinschaft oder intensivem administrativen Support, um diese durchgängig vollständig, detailliert und aktuell zu halten.

Länderübergreifende und nationale Portale für Open Government Data (OGD) bzw. offene Verwaltungsdaten werden seit mehreren Jahren staatlich unterstützt. So finanziert die Europäische Union seit 2015 das Europäische Datenportal (<https://europeandataportal.eu>) (EPSI platform 2016), welches einen gesammelten Katalog aus gegenwärtig mehr als 70 Datenkatalogen aus dem öffentlichen Sektor und dem Geodaten-Bereich der EU umfasst, darunter 6 nationale Portale aus Deutschland (GovData, GDI-DE), Österreich (data.gv.at, Open Data Portal Austria, INSPIRE Portal Austria) und der Schweiz (opendata.swiss). Diese nationalen Portale fassen wiederum regionale und kommunale Datenangebote zusammen. Neben kategorie- und stichwortbasierter Suche unterstützen sowohl das Europäische Datenportal als auch GovData (<https://www.govdata.de>), das Datenportal für Deutschland, eine kartenbasierte Suche, bei der Datensätze anhand einer Rechteckauswahl (Bounding Box) gefunden werden können, sofern sie mit entsprechenden Metadaten versehen sind.

Ein offenes Datenportal für Deutschland ging bereits 2010 aus einer Bürgerinitiative hervor (<https://offenedaten.de>). Aus einer ähnlichen Motivation heraus wurde das bis Ende 2014 aktive Projekt Open Data Monitor ins Leben gerufen, welches erstmalig offene Datensätze deutschlandweit und umfassend kartierte (Kubicek und Lippa 2015a, 2015b). Die daraus resultierende interaktive Open-Data-Landkarte (<http://open-data-map.de>) wurde abschließend in das Portal offenedaten.de überführt. Datensätze auf der Web-Karte sind mit Markern und mit thematischen Schlagwörtern versehen. Im Unterschied zum staatlichen Datenportal GovData, welches lediglich Datensätze führt, die von Einrichtungen des Bundes, der Bundesländer und Kommunen registriert wurden, basiert der Datenkatalog der Open-Data-Landkarte auf proaktiver Suche. Es wurden Datenkataloge ausgewertet, Suchmaschinen verwendet und einschlägige Webseiten mittels eines selbstentwickelten Web-Crawlers nach Datensätzen durchsucht. Mit diesem Vorgehen konnten viele bis dahin nicht umfassend katalogisierte Datensätze, insbesondere von Kommunen, registriert werden, allerdings war die Datenerhebung mit erheblichem manuellen Vor- und Nachbereitungsaufwand verbunden.

3 Ein Portal der Open Data Portale

Das hier vorgestellte Portal unterscheidet sich von den anderen Einstiegsseiten dadurch, dass es nicht nur auf Open Data Portale beschränkt ist, sondern Geodateninfrastrukturen (GDI) und andere Datenquellen einbezieht, bei denen es neben nutzungsbeschränkten und kommerziellen Produkten eine Tendenz zu offenen Daten oder kostenfrei nutzbaren Angeboten gibt. Tabelle 1 listet die Kategorien auf, nach welchen die Datenportale klassifiziert werden. Kernbestandteil des Portals der Portale sind die beiden in Abbildung 1 dargestellten Karten- und Tabellen-Ansichten. Datenportale sind auf der Karte durch Marker repräsentiert, deren Form und Farbe sich gemäß der Tabelle nach Portal-Typ und Reichweite richten. Eine textuelle Beschreibung sowie ein Link zur Webseite kann entweder durch Pop-up Fenster direkt in die Karte eingeblendet oder in der Tabelle gesucht werden.

Tabelle 1: Kategorien und Symbolik des Portals der Portale. Jedes Datenangebot wird gemäß seiner Art und Reichweite bzw. Verwaltungsebene klassifiziert.

Portal-Typ		Reichweite	
	Open Data Portal		international
	GDI/ Geoportale		national
	Statistikportal		regional
	Umweltdatenportal		kommunal
	Forschungsdatenportal		
	Citizen Science Projekt		

Kartenmarker sind i. d. R. am Ort des Herausgebers eines Datenangebotes platziert. So verdeutlichen dessen Position und Reichweite den räumlichen Bezug eines Portals, was ermöglicht, Verzeichniseinträge gemäß einem Untersuchungsgebiet zu erfassen. Dafür können auch Filter angewendet werden, welche zugleich auf der Karte und der Tabelle wirksam sind. Dazu gehören Auswahl-Boxen, mittels derer man Datenportale gemäß den Kategorien aus- und einblenden kann. Mithilfe einer Rechteck-Auswahl auf der Karte (Bounding Box) können Einträge in der Tabelle für ein bestimmtes Gebiet gefiltert werden. Zum Beispiel können die Marker und Tabelleneinträge mithilfe der Checkbox-Filter auf regionale Geodateninfrastrukturen und Geoportale eingegrenzt und mittels

der Rechteck-Auswahl auf das Gebiet um Nordrhein-Westfalen beschränkt werden. In der Tabelle ausgewählte Einträge werden außerdem zugleich in der Karte hervorgehoben.

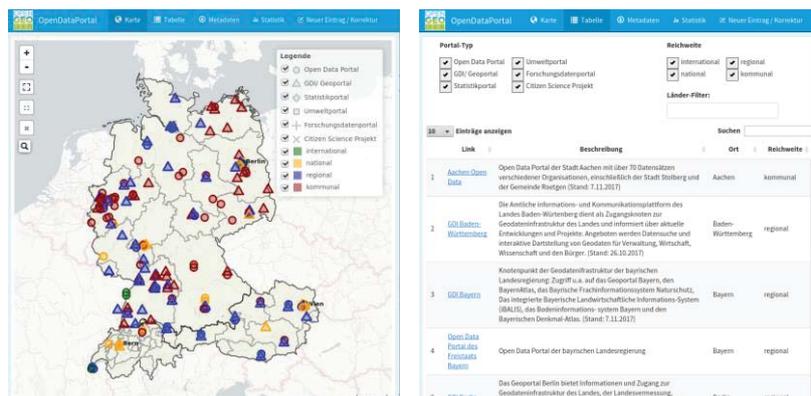


Abbildung 1: Interaktive Web-Karte und Tabellenansicht

Das Verzeichnis der Datenportale besteht gegenwärtig aus 250 Einträgen, zusammengesetzt aus Titel, URL und einer kurzen textuellen Beschreibung, sowie Portal-Typ, Reichweite, Position und Name des Bezugsortes (z. B. Land, Region oder Stadt). Ein Großteil dieser Daten basiert auf den Recherchen des OpenGeoEdu-Teams. Ein Web-Formular ermöglicht aber auch Nutzern des Portals, neue Einträge zu verfassen oder die Änderung vorhandener zu beantragen. Da OpenGeoEdu an verschiedenen deutschsprachigen Hochschulen beworben wird, ergibt sich das Potenzial einer wachsenden Nutzergemeinschaft, die das Verzeichnis von Portalen auch in Zukunft aktualisiert und ergänzt.

Das Verzeichnis der Datenportale kann von dem Portal als Tabelle oder in gängigen Geodaten-Formaten wie GeoJSON, GeoPackage oder ESRI Shapefile heruntergeladen werden, sodass dieses selbst als Open Data verfügbar ist. Auch das Portal ist quelloffen und besteht, wie im folgenden Abschnitt beschrieben, ausschließlich aus Open Source-Komponenten.

4 Technischer Aufbau des Portals

Das Portal der Datenportale ist eine Client/Server-Anwendung, deren Architektur in Abbildung 2 dargestellt wird. Die interaktive Kartendarstellung basiert auf der JavaScript Bibliothek Leaflet (Agafonkin 2018). Serverlogik, Datenverarbeitung und clientseitige Darstellung wurden mit der freien Programmiersprache und Softwareumgebung R (R Core Team 2017) umgesetzt, welche sich mit einer Vielzahl von Erweiterungspaketen zu einer universellen Sprache entwickelt, so auch im Bereich der Geodatenanalyse und Geoinformation (Bivand et al. 2013).

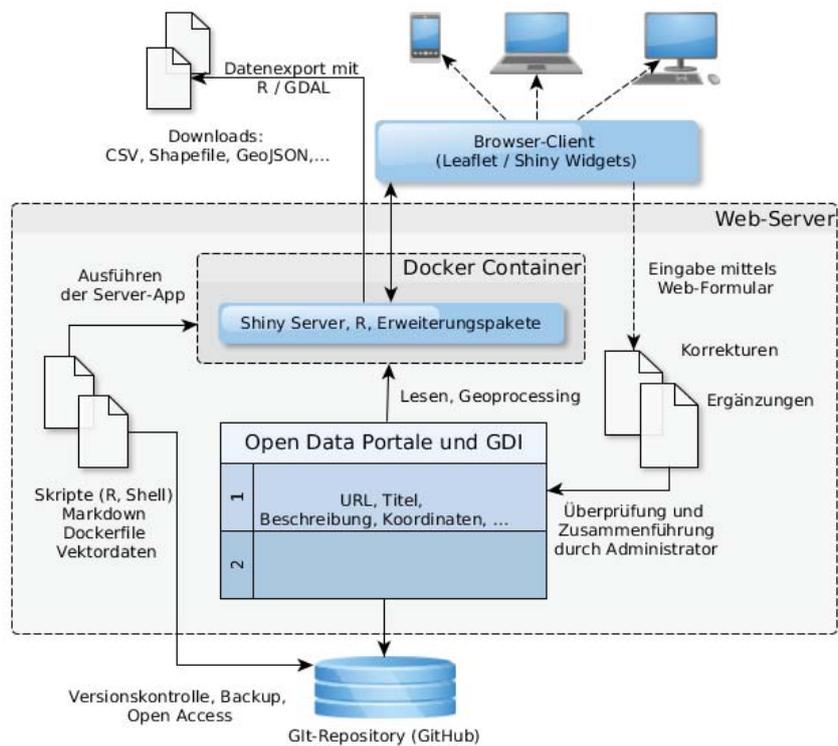


Abbildung 2: Software-Architektur des Portals mit R und Leaflet als Hauptbestandteilen

Mithilfe eines R-Markdown-Dokuments ist es möglich, mit wenig Aufwand ein webbasiertes und responsives Dashboard (Flexdashboard) zu entwickeln, welches dem Grundlayout des Portals entspricht (Borges und Allaire 2017). R-Markdown-Dokumente kombinieren R-Code, Markup-Syntax, CSS und JavaScript und können in ein browser-kompatibles HTML-Dokument übersetzt werden. Shiny (Chang et al. 2017) bietet eine Laufzeit-Umgebung, in der solche R-basierten Web-Anwendungen gehostet werden können und serverseitig R-Code ausgeführt werden kann, was das Realisieren von dynamischen und interaktiven Web-site-Elementen (Shiny-Widgets) erlaubt, sowie die serverseitige Verarbeitung von Benutzereingaben.

Das R-Erweiterungspaket Leaflet ermöglicht nicht nur das Erstellen der Web-Karte auf Basis der gleichnamigen JavaScript-Bibliothek, sondern auch das Erweitern der Funktionalität in Kombination mit anderen R-Paketen. Die oben erwähnte Interaktion zwischen Checkbox-Filtern, Karte und Tabelle sowie die Auswahlwerkzeuge sind mit dem Paket Crosstalk (Cheng 2017) umgesetzt. Die Shiny-Laufzeitumgebung ermöglicht das Ein- und Ausblenden von Kartenelementen z. B. bei verschiedenen Zoomstufen oder Nutzereingaben. Sie gewährleistet auch das persistente Speichern von Nutzereingaben auf dem Server (Editorfunktion) sowie das dynamische Nachladen von Informationen. Mithilfe der Geospatial Data Abstraction Bibliothek (GDAL) (GDAL Development Team 2017) und dem zugehörigen R-Paket rgdal (Bivand et al. 2017) können die Portal-Daten in verschiedene Geodatenformate exportiert werden.

Ein Git-Repository ermöglicht externe Sicherungen und Versionskontrolle des Quellcodes und der Rohdaten in Tabellenform. Über das Hosting-Portal GitHub sind diese Ressourcen im Sinne von Open Access öffentlich verfügbar (<https://github.com/opengeoedu>).

5 Zusammenfassung und Ausblick

Mit dem Portal der Datenportale steht Anwendern aus dem Geodatenbereich ein nützliches WebGIS-Tool für die Recherche nach offenen Daten zur Verfügung, das einen schnellen und einfachen Überblick über freie Datenangebote im deutschsprachigen Raum vermittelt und auch für Branchen aus dem Data Science Bereich interessant sein dürfte. Im Rahmen von OpenGeoEdu wird dieses beständig weiterentwickelt und zukünftig in die Lehre zum Umgang mit offenen Geodaten für raumbezogene Studiengänge eingebunden. Da die Daten und das Portal selbst als offene Ressourcen verfügbar sind, besteht auch für Dritte die

Möglichkeit der Wiederverwendung, beispielsweise, um das Verzeichnis der Portale in anderer Art zu visualisieren, mit eigenen Daten zu verknüpfen oder die Software weiterzuentwickeln. Offene Daten haben nur dann einen Mehrwert für die Gesellschaft, wenn sie nicht nur verfügbar, sondern auch auffindbar sind. Somit ist das Portal ein Beitrag zur Erschließung der bestehenden „Datenschätze“.

Danksagung

Die Verfasser danken dem Bundesministerium für Verkehr und Digitale Infrastruktur (BMVI) für die Förderung im Rahmen des mFUND-Programms (FKZ: 19F2007A).

Literaturverzeichnis

- Agafonkin, Vladimir (2018): Leaflet: An open-source JavaScript library for mobilefriendly interactive maps, zuletzt geprüft am 31.01.2018.
- Bivand, Roger S.; Keitt, Tim; Rowlingson, Barry (2017): rgdal. Bindings for the 'Geospatial' Data Abstraction Library. Online verfügbar unter <https://CRAN.R-project.org/package=rgdal>.
- Bivand, Roger S.; Pebesma, Edzer; Gómez-Rubio, Virgilio (2013): Applied Spatial Data Analysis with R. New York, NY: Springer New York.
- Borges, Barbara; Allaire, J. J. (2017): flexdashboard. R Markdown Format for Flexible Dashboards. Online verfügbar unter <http://rmarkdown.rstudio.com/flexdashboard>, zuletzt geprüft am 22.12.2017.
- Chang, Winston; Cheng, Joe; Allaire, J. J.; Xie, Yihui; McPherson, Jonathan (2017): shiny. Web Application Framework for R. Online verfügbar unter <https://CRAN.R-project.org/package=shiny>, zuletzt geprüft am 22.12.2017.
- Cheng, Joe (2017): crosstalk. Inter-Widget Interactivity for HTML Widgets. Online verfügbar unter <https://rstudio.github.io/crosstalk/>, zuletzt geprüft am 22.12.2017.
- CTIC-CT: Open Data @ CTIC. Public Dataset Catalogs Faceted Browser. Online verfügbar unter <http://datos.fundacionctic.org/sandbox/catalog/faceted/>, zuletzt geprüft am 24.11.2017.
- EPSI platform (2016): Understanding the European Data Portal. Online verfügbar unter https://www.europeandataportal.eu/sites/default/files/2016_understanding_the_european_data_portal.pdf, zuletzt geprüft am 22.12.2017.
- GDAL Development Team (2017): GDAL - Geospatial Data Abstraction Library, Version 2.2.1. Open Source Geospatial Foundation. Online verfügbar unter <http://www.gdal.org>.

- Kubicek, Herbert; Lippa, Barbara (2015a): Open Data: Appelle reichen nicht. In: *Kommune 21* (5/2015), S. 12–13. Online verfügbar unter http://www.kommune21.de/meldung_21279_Appelle+reichen+nicht.html, zuletzt geprüft am 22.12.2017.
- Kubicek, Herbert; Lippa, Barbara (2015b): Open Data: Kommunale Krux. In: *Kommune 21* (4/2015), S. 12–13. Online verfügbar unter http://www.kommune21.de/meldung_21140_Kommunale+Klux.html, zuletzt geprüft am 22.12.2017.
- R Core Team (2017): R. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Online verfügbar unter <https://www.R-project.org/>.
- Zscheile, Falk (2017): Mit Open Government Data Planen – Lizenzkonforme Nutzung von Datensätzen. In: Ralf Bill, Andreas Golnik, Marco L. Zehner, Tobias Lerche, Jörg Schröder und Sarah Seip (Hg.): *GeoForum MV 2017 – Mit Geoinformationen planen!* Berlin: Gito mbH Verlag, S. 113–121.