

Matthias Plennert

Anwendungsreif? Nutzung und Potenzial von digitalen Geodaten für Stadtforschung und Raumbewachung am Fallbeispiel OpenStreetMap¹

*In diesem Beitrag wird OpenStreetMap (OSM) als ein Beispiel für digitale Geodaten vorgestellt. OSM ist ein VGI-Projekt, das nicht mit amtlichen Organisationen und Strukturen zu vergleichen ist. Das zeigt sich schon beim Zugang: OSM steht unter der Open Database License und ist sowohl dem Open Data, als auch den Open Source Bewegungen zuzuschreiben. Die Daten selbst werden von der sehr heterogenen **community** freiwillig erhoben. Durch die offene Datenhaltung bietet sich OSM als Ausgangspunkt für Nutzungsszenarien in der Stadtforschung und Raumbewachung an. In Rahmen des Beitrags soll erörtert werden, welche Besonderheiten bei der Nutzung von OSM-Geodaten zu beachten sind und welche Chancen diese bieten können. Insgesamt haben die OSM-Geodaten bereits heute, vor allem in großen Städten, einen hohen Detailgrad erreicht. Zuverlässige Aussagen, die mehrere Jahre in die Vergangenheit reichen, sind allerdings bisher nur in Ausnahmen möglich. Wenn die OSM-Beitragenden in den kommenden Jahren weiterhin so aktiv bleiben, sind in Zukunft detaillierte raum-zeitliche Analysen auf Basis von OSM-Daten möglich.*

Matthias Plennert

M. A., seit 2014 Wissenschaftlicher Mitarbeiter am Lehrstuhl für Kulturgeographie des Instituts für Geographie, FAU Erlangen-Nürnberg

✉ Matthias.plennert@fau.de

Schlüsselwörter:

Geodaten – OpenStreetMap

Das OpenStreetMap-Projekt

Das OpenStreetMap-Projekt (OSM) ist eines der am häufigsten genannten Beispiele für nutzergenerierte Geodaten. Dabei nimmt OSM eine Sonderstellung ein: das Projekt startete 2004 mit dem Ziel, eine offen zugängliche Alternative für den Ordnance Survey, das britische Vermessungsamt, zu erstellen. OSM gehört heute zu den gut organisierten und strukturierten Projekten, welche das Ziel der Datenerhebung und -bereitstellung verfolgen. Durch die offene Datenhaltung bietet sich OSM als Ausgangspunkt für Nutzungsszenarien in der Stadtforschung und Raumbewachung an. In diesem Beitrag soll erörtert werden, welche Besonderheiten bei der Nutzung von OSM-Geodaten zu beachten sind und welche Chancen diese bieten können.

Nutzergenerierte Daten, sogenannter User Generated Content (UGC), können auf verschiedene Weisen kategorisiert werden. Dabei spielt für UGC in erster Linie das Kriterium der Datenerstellung eine Rolle: die Daten werden von Nutzern und Nutzerinnen im Internet erhoben. Andere Kriterien wie die Freiwilligkeit oder die Intention der Datenerhebung sind dafür zunächst irrelevant. Dasselbe gilt auch für geographic User Generated Content (gUGC) im geographischen Kontext. Volunteered Geographic Information (VGI), ein Begriff der ursprünglich durch Goodchild (2007) geprägt wurde, bildet eine Teilmenge von gUGC. Nach längerer Diskussion in Fachkreisen hat sich inzwischen ein ausdifferenziertes Verständnis von VGI gebildet. Sowohl gUGC als auch VGI beschreiben nutzergenerierte Geodaten. Im Unterschied zu gUGC sind VGI jedoch freiwillig und bewusst gesammelt worden, um sie anschließend offen zur Verfügung zu stellen (Elwood et al. 2012: 575).

Mit dem erklärten Ziel, digitale Geodaten zu erheben und diese offen zur Verfügung zu stellen, ist OSM ein VGI-Projekt. Bei der Geodatenerhebung, dem sogenannten *mapping*, gibt es in OSM keine räumlichen oder thematischen Grenzen. OSM hat sich so zum erfolgreichsten Anbieter von VGI entwickelt. Dabei ist OSM in erster Linie als eine *wiki-community* zu verstehen: es besteht aus einer Gemeinschaft aus Beitragenden, die das gemeinsame Ziel verfolgen, ihre erstellten Inhalte zu teilen. Sie arbeiten freiwillig aus einer intrinsischen Motivation heraus. Die Beitragenden teilen das gemeinsame ideelle Ziel, Geodaten frei verfügbar zu machen.

Seit 2004, dem Gründungsjahr von OSM, hat die *community* eine umfangreiche technische Infrastruktur entwickelt. Neben den Editoren oder Kartendarstellungen stellt dabei vor allem die Datenbank das Herzstück des Projekts dar. Hier werden die gesammelten Geodaten gespeichert. Die Datenbank dient als Basis für alle Anwendungen von OSM-Daten. Die bekannteste Anwendung ist dabei wahrscheinlich die Kartendarstellung auf der Website www.osm.org. Sucht man nach OSM, so stößt man in der Regel zuerst auf diese Karte. Dabei handelt es sich beim Standard Mapnik-Style (wie auf www.osm.org zu sehen) zwar um die bekannteste Visualisierung von OSM-Geodaten, diese ist jedoch nur eine unter vielen. So existieren Visualisierungen, die bestimmte Aspekte hervorheben, wie die Fahrradkarte (www.opencyclemap.org), die Rollstuhlkarte (www.wheelmap.org), oder die Wintersportkarten (www.openpiste.org). Ferner finden OSM-Geodaten in Anwendungen Einsatz, wie bspw. dem Routingdienst Graphhopper (www.graphhopper.com).

Neben der Datenbank und darauf basierenden Anwendungen, gibt es Schnittstellen und Editoren, wie den JOSM-Editor, ID-Editor, Potlatch-Editor oder der Overpass-API, die das Abrufen und Bearbeiten der Geodaten in OSM strukturieren. Diese sind unterschiedlich komplex aufgebaut um den Nutzern und Nutzerinnen mit ihrer unterschiedlichen technischen und fachlichen Expertise gerecht zu werden. Dadurch werden sowohl einfachere Vorgänge als auch komplexe Zugriffe ermöglicht. Alle OSM-Geodaten verfügen über eine Versionierung, die es möglich macht, vergangene Inhalte abzurufen. Das bedeutet, dass man sehr einfach rekonstruieren kann, welche Inhalte zu einem bestimmten Zeitpunkt in der Vergangenheit bereits in die Datenbank eingetragen waren. Somit lässt sich der zeitliche Verlauf der Eintragungen rekonstruieren. Daneben verfügen die Inhalte des OSM-Projekts über keine abgeschlossene Struktur oder Klassifizierung. Die Geodaten werden durch eine offene Attributierung, auch *Social Tagging* genannt, mit Semantik versehen. Das *Social Tagging* arbeitet dabei mit einem Schlüssel-Wert Paar (Key-Value pair). Ein *Tag*, bestehend aus Key und Value, beschreibt genau eine Eigenschaft eines Objekts. Ein Objekt kann mit mehreren Tags versehen werden. Ein Punkt, repräsentiert durch eine Koordinate, wird somit erst durch das *Tag amenity = cafe* zu einem Café. Da dieses *Tag* bereits als offizielles Attribut zur Benennung von Cafés in die Dokumentation aufgenommen wurde, wird es systematisch verwendet, um auf die Geodaten „Cafés“ zuzugreifen.

OpenStreetMap – Open Data

Initiativen mit dem Ziel, Daten frei verfügbar zu machen, erleben u. a. mit der Etablierung des Web 2.0 einen Aufschwung. Diese Initiativen werden häufig den „Open“ Bewegungen zugeordnet, wie bspw. Open Source, Open Content, Open Access oder Open Data, wie beim Beispiel OpenStreetMap. Die Website OpenDefinition.org (Open Definition o. J.) definiert in diesem Kontext in zahlreichen Sprachen den „Open“-Begriff:

„Wissen ist offen, wenn jedeR darauf frei zugreifen, es nutzen, verändern und teilen kann – eingeschränkt höchstens durch Maßnahmen, die Ursprung und Offenheit des Wissens bewahren.“

Aus den *Open Movements* sind verschiedene Lizenzen hervorgegangen, welche versuchen, genau diese Idee umzusetzen. Bei den Lizenzen handelt es sich in der Regel um technische Spezifizierungen, die die Offenheit gewährleisten. Das OpenStreetMap-Projekt verwendet die Open Database License (ODbL). Diese Lizenz wurde von der Open Knowledge Foundation entwickelt und findet seit 2012 Verwendung in OSM-Daten. Diese Lizenz ist in erster Linie eine technische Spezifizierung für die Umsetzung von offenen Datenbanken. Daher muss die Original-Lizenz zunächst von dem rein technischen Format in eine menschenlesbare Form umgewandelt werden. Die Übersetzung hat der Open Data Commons Beirat (Open Data Commons o. J.) sinngemäß vorgenommen:

„You are free:

To Share: To copy, distribute and use the database.

To Create: To produce works from the database.

To Adapt: To modify, transform and build upon the database.

As long as you:

Attribute: You must attribute any public use of the database, or works produced from the database, in the manner specified in the ODbL. For any use or redistribution of the database, or works produced from it, you must make clear to others the license of the database and keep intact any notices on the original database.

Share-Alike: If you publicly use any adapted version of this database, or works produced from an adapted database, you must also offer that adapted database under the ODbL.

Keep open: If you redistribute the database, or an adapted version of it, then you may use technological measures that restrict the work (such as DRM) as long as you also redistribute a version without such measures.“

Mit der ODbL wird ein umfassender Zugang zu den Daten des OSM-Projekts gewährleistet. OpenStreetMap hat damit ganz wesentlich von Web 2.0 Technologien profitiert und ist eines der erfolgreichsten Projekte der Open Movements, bzw. Open Data Projekte. Der einfache Zugang zu OSM-Daten steht zudem im starken Kontrast zu proprietären Geodaten, welche häufig nur unter Einsatz von großen finanziellen Ressourcen verwendet werden können (siehe dazu Artikel von Tim Elrick in diesem Heft). Gleichzeitig bietet die umfangreiche Datenbank großes Potenzial für Anwendungen mit Geo-Bezug.

Qualität digitaler Geodaten

Die Datenqualität von digitalen Geodaten spielt bei jeder Art der Nutzung eine zentrale Rolle. Um diese beurteilen zu können werden häufig Kriterien verwendet, welche sich direkt auf die erhobenen Daten beziehen. Dazu zählen bspw. Integrität, Präzision, oder Aktualität (DGIQ o. J.). Behörden betreiben viel Aufwand, um diesen Kriterien bei ihrer eigenen Datenerhebung und -verarbeitung gerecht zu werden: es werden elaborierte Klassifizierungssysteme entwickelt, um inhaltlich

überschneidungsfreie Daten zu erzeugen, komplexe Kontrollmechanismen angewendet, um die Integrität der Daten zu wahren oder teure Messgeräte verwendet, um ausreichend präzise Messungen vornehmen zu können. Neben den Kriterien, die sich direkt auf die Daten beziehen, gibt es Gesichtspunkte, die sich auf den Kontext der Datenerhebung beziehen. Dazu zählen bspw. Informationen zu den Methoden und ihrer Zuverlässigkeit. Der große monetäre und organisatorische Aufwand für die Erhebung von Geodaten ist grundsätzlich wichtig und richtig. Die so erhobenen Geodaten sind belastbar und dienen als Grundlage für weitreichende Entscheidungen. Betrachtet man nutzergenerierte Daten mit erstgenannten Kriterien, so schneiden diese häufig nicht gut ab. Dabei ist OSM keine Ausnahme. Zwar belegen einige Studien eine höhere Datendichte von OSM-Geodaten gegenüber einigen Anbietern aus der Navigationsbranche (Ludwig et al. 2010; Neis u. Zipf 2012; Roick et al. 2011). Eine hohe Datendichte ist allerdings nicht ausreichend um Daten eine hohe Qualität zuzuweisen, zumal Datendichte kein Kriterium der DGIQ ist. Trotzdem darf man deshalb OSM als hochwertige Geodatenquelle nicht verwerfen.

OSM ist ein außerordentlich heterogenes Projekt, was sich in den Daten und der Beteiligung deutlich zeigt (Wood 2014). Dieser Umstand muss bei der Bestimmung der Datenqualität berücksichtigt werden. Somit ist es nicht ausreichend, nur die erhobenen Daten zu betrachten, man muss vielmehr das gesamte Projekt und damit den Kontext der Datenerhebung verstehen, um es einschätzen zu können. Erst wenn man Transparenz bezüglich der Datenerhebung geschaffen hat, kann man mit Kriterien arbeiten, welche sich direkt auf die Daten beziehen.

Datenqualität in OSM: Community verstehen

Möchte man OSM verstehen, so ist es nicht ausreichend, sich die formalen Strukturen anzusehen, zumal diese kaum vorhanden sind. Es existieren zahlreiche informelle Strukturen und Hierarchien, welche das Projekt regeln. So hat OSM im Jahr 2015 mehr als 2 Millionen Mitglieder. Obwohl nur wenige formalisierte Strukturen vorhanden sind, die Entscheidungen oder Konflikte regulieren, ist OSM weiterhin erfolgreich: die Mitgliederzahlen nehmen weiter zu, es werden immer mehr Geodaten erhoben und Entscheidungen sowie Konflikte werden zur Zufriedenheit des Großteils der *community* gelöst. Um also Transparenz bezüglich des Kontexts der Datenerhebung herzustellen, ist es wichtig zu verstehen, welche Aspekte, formelle und informelle Strukturen, Abläufe und Prozesse bei der Organisation des Projekts und insbesondere bei der Entscheidungsfindung eine Rolle spielen.

Das OSM Projekt wird formal durch die OSM Foundation vertreten, welche 480 Mitglieder hat (Glazze u. Perkins 2015: 153 f.). Es ist jedem möglich, gegen einen Beitrag der Foundation beizutreten. Die Foundation selbst wählt einen Vorstand aus sechs Mitgliedern. Neben dem Vorstand existieren acht Arbeitsgruppen, die sich den Themen Kommunikation, Daten, Lizenz, Ortsgruppen, Entwicklung, der „State of the Map“ Konferenz und Strategie widmen. Diese formale hierarchische Struktur spiegelt jedoch nicht die tatsächlichen Machtverhältnisse in

der *community* wider. Die Arbeitsgruppen und der Vorstand treten eher moderierend auf. Wenn es zu Konflikten oder Problemen in der *community* kommt, fordert in der Regel eine Arbeitsgruppe eine Lösung der direkt betroffenen (lokalen) Beitragenden. Bisher sind keine Fälle bekannt, in denen eine Arbeitsgruppe direkt und kompromisslos in einen Konflikt eingegriffen und diesen durch ein „Machtwort“ beendet hätte. Die formale Struktur und die mit ihr verbundene Macht wird mit dem Begriff der Bürokratie zusammengefasst (Glazze u. Perkins 2015: 153 f.; Ramm 2013).

Die Zurückhaltung der OSM Foundation und damit die geringe Bedeutung der Bürokratie in OSM spiegelt die *community*-Priorität des Projekts wider. Alle formalen Entscheidungen sollen, zumindest theoretisch, demokratisch und unter Beteiligung der gesamten *community* gefällt werden. Unter formale Entscheidungen fallen bspw. der Lizenzwechsel 2012, neue *Tagging-Schemata* oder Bestimmungen für die Datenerhebung. Jedes OSM-Mitglied ist stimmberechtigt und kann sich an den teilweisen regen Diskussionen beteiligen. Dass Abstimmungen jedoch nur theoretisch demokratisch sind, hat verschiedene Gründe, wobei das Long-Tail Phänomen wohl der wichtigste ist. Das Long-Tail Phänomen beschreibt, im Zusammenhang mit OSM, die ungleiche Beteiligung der Nutzer und Nutzerinnen am Projekt. Einfach ausgedrückt tragen sehr wenige Nutzer sehr viel und sehr viele Nutzer sehr wenig zu OSM bei. Dieses Phänomen zieht sich durch alle Bereiche des Projekts: von der Datenerhebung über die Dokumentation bis hin zur Entwicklung neuer Software (Wood 2014).

Das Long-Tail Phänomen spiegelt sich jedoch nicht nur in der ungleichen Verteilung der Beiträge wider, sie beeinflusst auch die *community*-Struktur in OSM. Nutzer bzw. Nutzerinnen, die besonders viel beitragen, erfahren eine höhere Wertschätzung. Gleichzeitig sind es vor allem diese Beitragenden, die sich an Diskussionen und Abstimmungen beteiligen. Dadurch ergibt sich eine informelle Hierarchie, die besonders aktive Beitragende an die Spitze setzt. Diese Art der *community*-Struktur wird von Glazze und Perkins (2015: 153) mit dem Begriff Meritokratie beschrieben.

Da innerhalb des OSM-Projekts neben der Erstellung von Daten (bspw. Dokumentation im OSM Wiki oder Erhebung von Geodaten) auch die dafür notwendige Software entwickelt wird, sind informatische, bzw. technische Kenntnisse von besonderer Bedeutung. Auch diese Wertschätzung findet in Form einer Technokratie Einfluss in die informelle Hierarchie. Technokratie bedeutet in OSM, dass Nutzer bzw. Nutzerinnen mit besonderen technischen Kenntnissen (z. B. Softwareentwicklung) ein höheres Gewicht bei der Meinungsbildung haben (Glazze u. Perkins 2015: 153).

Somit ergibt sich ein Konglomerat aus Bürokratie, Meritokratie und Technokratie, die die de facto Entscheidungsfindung prägen. Neben den genannten Kategorien der Entscheidungsfindung gibt es noch andere Prinzipien, anhand derer Entscheidungen getroffen werden.

Da OSM über keine umfassenden, festgelegten Strukturen und Abläufe verfügt, gibt es zahlreiche Bereiche, die keine formalisierten Entscheidungsprozesse vorgeben. In diesen Bereichen werden in der Regel Fakten geschaffen: kommt ein Beitragender in eine bisher noch nicht geregelte Situation, entscheidet er sich für eine Lösung. Wenn die Mehrheit der Bei-

tragenden seinem Beispiel folgt, ergibt sich automatisch eine Entscheidung und eine Struktur etabliert sich. Dieses Prinzip wird auch als Do-ocracy bezeichnet (Community Wiki o. J.). Dabei handelt es sich um Entscheidungen, die ohne direkte Abstimmung, sondern vielmehr durch die Ausführung getroffen werden. Grundsätzlich ist es auch bei Do-ocracy möglich, seine Meinung zu äußern und über Entscheidungen zu diskutieren. In der Regel sind von dieser Art der Entscheidungsfindung weniger grundlegende Bereiche im OSM-Projekt betroffen, wie beispielsweise Dokumentationseinträge im OSM Wiki.

Vor allem bei der Entwicklung von Software-Komponenten spielte von Anfang an das *community-building* eine wichtige Rolle: das gesamte Projekt sollte so gestaltet werden, dass es einfach zu verstehen und zu bedienen ist. Klassische informatische Kriterien wie Performance oder Datenintegrität spielten daher eine eher zweitrangige Rolle. Diese *community-Orientierung* spiegelt sich in dem Prinzip wieder, das mit „Look for the simplest useful thing“ umschrieben wird (Glasze u. Plennert 2015). Es werden also einfach Lösungen gesucht, die einen schnellen und problemlosen Einstieg in das Projekt gewährleisten.

Ein weiteres Prinzip bezieht sich direkt auf die Geodatenerhebung in OSM. Um Geodaten nach dem *Tagging-Schema* richtig einzutragen, gibt es verschiedene Dokumentationen, an denen man sich orientieren kann. Ist ein Objekt nicht eindeutig zuordenbar oder es besteht Uneinigkeit bei der Einteilung, gilt das Ground-Truth-Prinzip. Dieses Prinzip besagt, dass die Daten in der Geodatenbank so eingetragen werden, wie sie vor Ort erkennbar sind. Baustellen werden also auch als Baustellen eingetragen und nicht bereits als fertiges Gebäude. Das dieses Prinzip nicht immer eindeutig ist, haben bereits Glasze und Elrick (2015) in ihrem Vortrag „Maps & Mosques: (k)eine kleine Signatur und die Transformation von Geoinformation und Kartographie im digitalen Zeitalter“ offengelegt. Darin beschreiben sie die Kartierung von religiösen Andachtsstätten in OSM. Sie berichten von einer Diskussion unter OSM-Beitragenden, welche sich uneins über die Kartierung einer Hinterhof-Moschee waren. Da sie vor Ort nur als Kulturverein zu erkennen war, so argumentierten die Mapper, darf sie auch nur als Kulturverein kartiert werden. An diesem Beispiel erkennt man, wann das Ground-Truth-Prinzip an seine Grenzen stößt.

Die Entscheidungsfindung in OSM erscheint zunächst sehr komplex und undurchsichtig. Jedoch ist der Prozess der Entscheidungsfindung meist relativ kurz: in der Regel dauert der gesamte Zeitraum vom Anstoßen einer Diskussion bis zur Abstimmung nicht länger als einige Monate (wie man aus der Mailing List entnehmen kann, hat bspw. die Diskussion und Etablierung über den Key „product“ weniger als einen Monat gedauert). Hinzu kommt, dass das Tagging-Schema flexibel ist und somit offen für neue Beitragende mit neuen Ideen. Zudem ist die gesamte Kommunikation und Dokumentation frei einsehbar, was den Prozess transparent macht. Das führt zu einer großen Flexibilität des Tagging-Schemas: es kann schnell neuen Anforderungen angepasst werden.

Insgesamt darf OSM nicht als *eine* Organisation missverstanden werden. OSM als *community* ist vielmehr die Summe seiner zahlreichen verschiedenen Mitglieder und Mitgliederinnen. Diese verfügen über mehrere Wege, Meinungen zu bilden und Entscheidungen zu treffen: Meritokratie, Bürokratie und Tech-

nokratie. Abstimmung erfolgen demokratisch-formalisiert oder als informelle Do-ocracy. Dabei orientieren sich die Mitglieder an Prinzipien wie Ground-Truth oder „Look for the simplest useful thing“, die jedoch nicht formal festgelegt sind. Ist man sich diesen verschiedenen Formen der Entscheidungsfindung und -bildung bewusst, wird es möglich, die Qualität von Daten anhand dessen Entscheidungskontexts einzuschätzen.

OSM-Datenqualität am Beispiel von Cafés

Bisher wurden die Mechanismen der Entscheidungsfindung, welche die Struktur und Methodik des OSM-Projekts bestimmen, erläutert. Nun soll mit diesem Verständnis gezeigt werden, wie genau eine Einschätzung der Datenqualität aussehen könnte. Das soll am Beispiel von erhobenen Cafés geschehen. Möchte man Cafés aus der OSM-Datenbank auswerten und zuvor die Güte dieser Daten klären, gibt es verschiedene Möglichkeiten, diese festzustellen. Zunächst kann man recherchieren, ob es eine „offizielle“ Struktur für das Eintragen von Cafés gibt. Diese werden im OSM Wiki unter der Subseite „Map Features“ gelistet. Im besten Fall existiert hier eine detaillierte Vorgabe mit genauer Bezeichnung, einer Beschreibung des Objekts, dem Kartensymbol sowie einem Beispielfoto des Objekts. Eine Auflistung im OSM Wiki bedeutet, dass über die Empfehlung diskutiert und mehrheitlich dafür gestimmt wurde. Im Fall des Café wurde Ende 2007 über ein Tagging-Schema diskutiert und abgestimmt. Mit 18:0 Stimmen wurde das Schema *amenity = cafe* angenommen. In dem Beispiel lassen sich bereits indirekt Einflüsse des Long Tail-Phänomens erkennen: obwohl zu diesem Zeitpunkt bereits ca. 10.000 Beitragende bei OSM registriert waren, haben sich nur wenige, besonders aktive und sozial vernetzte Beitragende an der Abstimmung beteiligt. Diese werden auch als *heavyweight-user* bezeichnet (Budhathoki u. Haythornthwaite 2013: 2 ff.). Durch ihre Beteiligung an der Abstimmung ist die Wahrscheinlichkeit sehr hoch, dass das neue Schema schnell in der Datenbank Anwendung findet. Nach der Abstimmung wird das neue Schema in die Dokumentation aufgenommen und als Grundlage für Eintragungen und Korrekturen verwendet. Jedoch kann man aus einer Eintragung in der Dokumentation keine sicheren Aussagen über die tatsächliche Verwendung in der Datenbank treffen. Dafür ist die Website taginfo.openstreetmap.org hilfreich. Diese stellt die Häufigkeit aller verwendeten Eintragungen dar, unabhängig davon, ob sie bereits dokumentiert wurden oder nicht. Somit ist es möglich zu erkennen, welche verschiedenen Eintragungsarten in Zusammenhang mit Cafés erkennen: *building = cafe*, *amenity = internet_cafe* oder *shop = cafe*. Die mit großen Abstand am häufigsten verwendete Art ist jedoch dieselbe, die auch in der Dokumentation empfohlen wird: *amenity = cafe*. An dem Beispiel der Eintragungen von Cafés in OSM kann man erkennen, dass es in OSM relativ früh eine „offizielle“ Empfehlung für das Eintragen gab. Mit der Website *Taginfo* lässt sich zudem belegen, dass das festgelegte *Tagging-Schema*, bis auf einige Ausnahmen, beim Großteil der Eintragungen umgesetzt wurde. Es existieren also wenige Einträge, die dem Tagging-Schema widersprechen. Die „offizielle“ Dokumentation schafft außerdem die Voraussetzung für das Redigieren

von Inhalten: Beitragende können sich bei der Korrektur von Inhalten auf das OSM Wiki und der darin enthaltenen Dokumentation berufen.

Mit dem beschriebenen Beispiel lassen sich Aussagen über die räumlich-globale Gesamtheit der eingetragenen Objekte machen. TagInfo sagt dabei nichts über die räumliche Heterogenität der Daten aus. Um bei dem Beispiel der Cafés zu bleiben: die Dokumentation und die zahlreichen „richtigen“ Eintragungen in der Datenbank geben erste Hinweise auf die Güte der Daten. Inwiefern Cafés in einem bestimmten Raum wie Berlin, München oder Köln nach demselben Schema tatsächlich eingetragen wurden, kann man so jedoch nicht klären. Nun soll gezeigt werden, wie man sich dann der regionalen Güte der Daten anhand von lokalen Beitragenden nähern kann.

Map-Guarding: Territoriale Administration von Geodaten

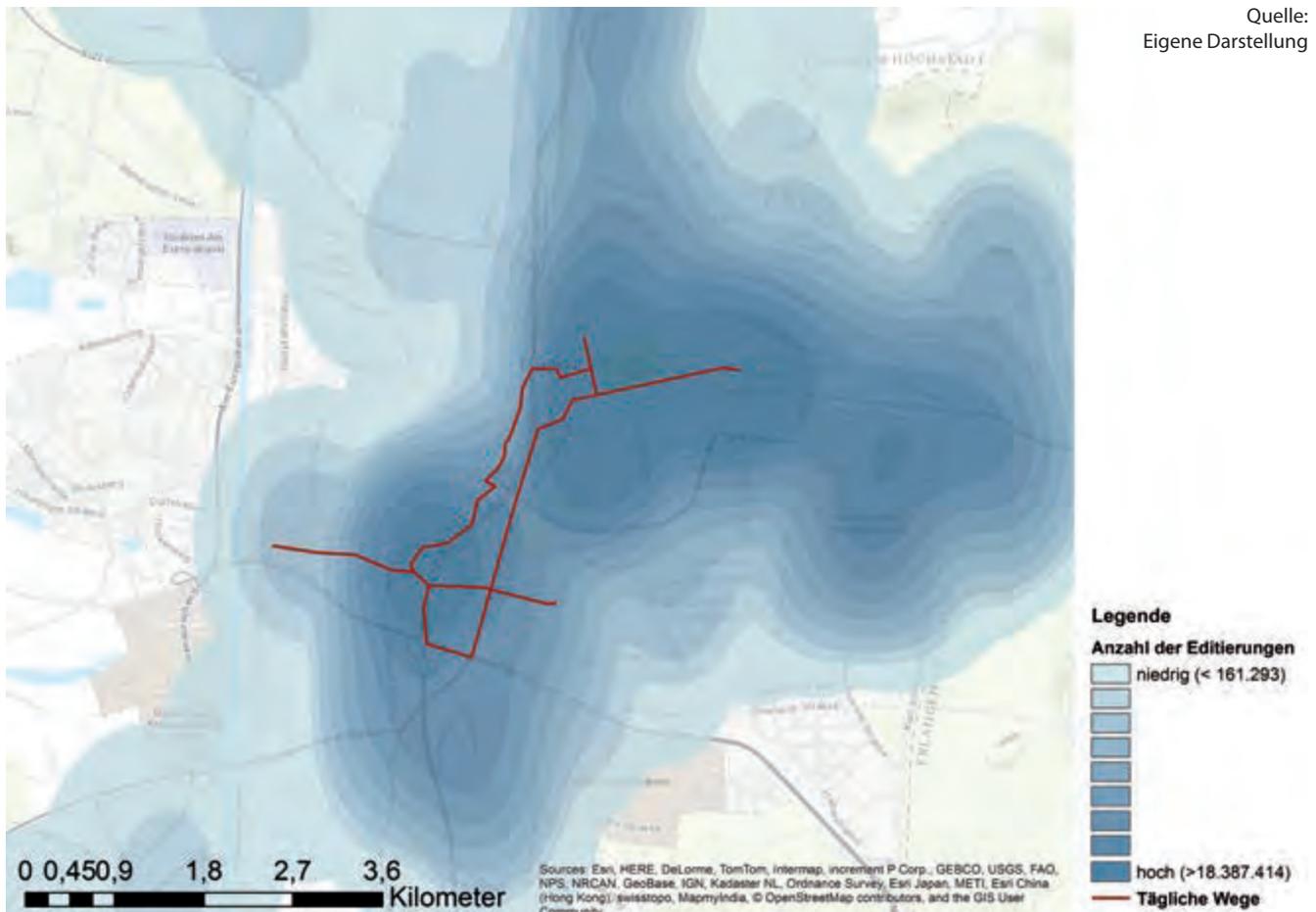
Die *community* in OSM ist sehr heterogen aufgebaut, wie mit dem Long-Tail Phänomen bereits beschrieben wurde. Es gibt nicht nur eine ungleiche Beteiligung unter den Mitgliedern, diese sind auch räumlich heterogen verteilt. Betrachtet man die Verteilung der Editierungen der Mitglieder (Abb. 1), ergibt sich folgendes Bild: die Geodaten, welche nach den jeweiligen Nutzer bzw. Nutzerinnen eingefärbt wurden, erzeugen einen Flickenteppich.

Abbildung 1: OSM-Geodaten in Norddeutschland, eingefärbt nach dem jeweiligen Beitragenden der Geodaten



Quelle:
Screenshot <http://bl.ocks.org/lxbarth/raw/6545162/#8/53.046/10.239>

Abbildung 2: Editierungen und tägliche Wege eines Beitragenden im Raum Erlangen



Quelle:
Eigene Darstellung

Abbildung 3: Gastronomische Eintragungen und deren Dokumentation im OSW Wiki

Schlüssel	Wert	Element	Kommentar	Darstellung	Foto
Verpflegung, Einkauf					
amenity	bar		Bar, Nachtclub . Es werden hauptsächlich alkoholische Getränke serviert. Siehe auch Beschreibung von amenity:bar und amenity:pub zur Unterscheidung von Bar und Pub (Wikipedia)		
amenity	bbq		Grillplatz . Kann mit amenity:bar (woodgas electric) kombiniert werden.		
amenity	biogarten		Biogarten		
amenity	cafe		Café, Eiscafé, Biers, Teeladen, Kaffeeladen , letztere beide stehen allerdings in Konkurrenz zu shop:cafe und shop:ice_cream		
amenity	drinking_water		Eine öffentlich zugängliche Trinkwasserstelle		
amenity	fast_food		Schnell-Restaurant, Imbiss . Die Art kann mit amenity:fast_food angegeben werden als Wert z.B. burger für ein Burgerrestaurant oder kebab für einen Dönerimbiss		

Quelle: Screenshot http://wiki.openstreetmap.org/wiki/DE:Map_Features

Abbildung 4: Kategorie „office“ und deren dokumentierten Ausprägungen

Schlüssel	Wert	Element	Kommentar	Darstellung / Rendering	Foto
office	accountant		Buchhalter, Steuerberater, Wirtschaftsprüfer		
office	administrative		Kreis-, Gemeinde-, Verwaltungs- und Aufsichtsbehörden, die keine Bundes- oder Landesbehörden sind.		
office	advertising_agency		An advertising agency or ad agency is a service based business dedicated to creating, planning, and handling advertising (and sometimes other forms of promotion) for its clients.		
office	architect		Architekturbüro		
office	association		gemeinnütziger, bzw. nicht gewinnorientierter Verein / Vereinigung / Interessensgemeinschaft , z.B. Sportverein oder Studentenverbindung		
office	company		Sitz einer privaten Firma (allgemeiner Wert, falls kein anderer Wert passt)		
office	educational_institution		Bildungseinrichtung		
office	employment_agency		Job-Center / Arbeitsvermittlung		
office	estate_agent		Immobilienmakler / Wohnungsbaugenossenschaft		
office	forestry		Forstamt / Revierförsterei		
office	foundation		Geschäftsstelle einer Stiftung		
office	government		Büro einer Regierung / Behörde / Regierungseinrichtung		

Quelle: Screenshot http://wiki.openstreetmap.org/wiki/DE:Map_Features

Abbildung 5: Die Tags eines Restaurants

```
<node id="1364036605" lat="53.5442963" lon="9.9770079">
  <tag k="addr:city" v="Hamburg"/>
  <tag k="addr:country" v="DE"/>
  <tag k="addr:houseNumber" v="42"/>
  <tag k="addr:postcode" v="20559"/>
  <tag k="addr:street" v="Fossetten"/>
  <tag k="amenity" v="restaurant"/>
  <tag k="cuisine" v="portugiesisch"/>
  <tag k="name" v="Restaurang Sagres Plus"/>
  <tag k="note" v="Es gibt 2 Restaurants Sagres mit gleichem Inhaber."/>
  <tag k="opening_hours" v="closed"/>
  <tag k="website" v="http://www.sagres.org"/>
  <tag k="wheelchair" v="limited"/>
</node>
```

Quelle: Screenshot <https://overpass-turbo.eu/>

Dieser Flickenteppich zeigt die verschiedenen Aktionsgebiete der *heavyweight user*, die von Napolitano und Mooney (2012: 6 ff.) als *pet-locations* umschrieben wurden. *Pet-locations* sind Räume, für die sich *heavyweight user* verantwortlich fühlen. Dabei bilden sich diese Räume nicht zufällig, sie spiegeln vielmehr die lokale Expertise des *heavyweight users* wieder. Das bedeutet, dass sie die Geodaten in ihrem persönlichen räumlichen Umfeld pflegen. Sie kennen sich innerhalb der *pet-locations* aus. Häufig sind die Räume durch ihre täglichen Wege durchzogen, wie in Abbildung 2 zu erkennen ist (Plennert 2014: 18 ff.).

Die Geodaten in diesen Räumen werden besonders intensiv gepflegt. Dieses Pflegen von Daten kann auch als *map-guarding*, bzw. territoriale Administration bezeichnet werden (Plennert 2014: 18 ff.). Das Administrieren von Inhalten umfasst „das Melden und Korrigieren von Fehlern, das Aktualisieren und Pflegen von Inhalten bis hin zum Kommunizieren mit anderen *community*-Mitgliedern, um sie über Fehler aufzuklären“ (Plennert 2014: 50). *Heavyweight user*, welche territorial administrieren, können auch als *Map-guards* bezeichnet werden.

Map-guards nehmen im Bezug zur Datenqualität eine Schlüsselstellung ein. Sie können den ausführenden Arm der *community* verkörpern. Werden neue Bestimmungen bezüglich des Datenschemas beschlossen, sind es vor allem die *Map-guards*, die die Daten einpflegen und anpassen. Sie stellen *gatekeeper* dar, die versuchen, die Datenintegrität zu wahren (Plennert 2014: 49 f.). In Aktionsgebieten von *Map-guards* kann eine hohe Datenqualität erreicht werden, bezogen auf Kriterien, die sich direkt auf die Daten beziehen. Jedoch sind die Bereiche mit hoher Datenqualität, genauso wie *Map-guards*, räumlich heterogen verteilt. Gelingt es, diese zu identifizieren, ist es jedoch möglich, auf hochwertige Daten zurückgreifen zu können.

Nutzungspotenzial/Nutzungsbeispiele

Ist für OSW die Frage der Datenqualität geklärt, so ergibt sich ein großes Nutzungspotenzial. Durch den einfachen Datenzugang, der Offenheit, der Flexibilität, der hohen Datendichte und der Aktualität ergeben sich viele Szenarien, in denen sich OSW-Geodaten gewinnbringend einsetzen lassen. Im Folgenden sollen drei Nutzungsbeispiele erläutert werden.

Detaillierte Datengrundlage

Betrachtet man den Standard OSW-Kartenstil, so fällt im Vergleich zur amtlichen topographischen Karte die hohe Informationsdichte auf. Wie man in Abbildung 3 erkennen kann, hat allein die Kategorie „Gastronomie“ auf der OSW-Karte viele verschiedene Ausprägungen: Bar, BBQ, Biogarten, Café, Fast Food, Food Court und Pub. Dabei wird nur ein Bruchteil der hinterlegten Informationen im Standard-Stil visualisiert. Die Datenbank bietet ein noch ausdifferenzierteres Datenschema. Das zeigt das Beispiel der Kategorie „Office“ in Abbildung 4, welche allein 25 verschiedene Ausprägungen aufweisen kann. *Keine* davon wird auf der OSW-Karte abgebildet; sie wären jedoch für eine Analyse über entsprechende Werkzeuge/Software abrufbar.

Die Vielfalt dieser „versteckten“ Informationen wird noch deutlicher, wenn man ein spezifisches Objekt auf der Karte genauer betrachtet. Abbildung 5 zeigt ein Restaurant-Symbol auf der OSM-Karte in Hamburg. Hinter dem Symbol verbergen sich jedoch noch zahlreiche weitere Informationen: Adressdaten (Land, Stadt, Postleitzahl, Straße, Hausnummer), Küche des Restaurant (portugiesisch), Name des Restaurants, Öffnungszeiten, Website, Eignung für Rollstuhlfahrer und sogar einen Hinweis darauf, dass es ein weiteres Restaurant in Hamburg mit demselben Namen und Inhaber gibt.

Diese Beispiele zeigen, welche Informationsfülle in der Datenbank zu finden ist. Die detaillierte Datengrundlage bietet neue Möglichkeiten für Anwendungen, die auf ausdifferenzierte Kategorien zurückgreifen.

Leerstandsmonitoring

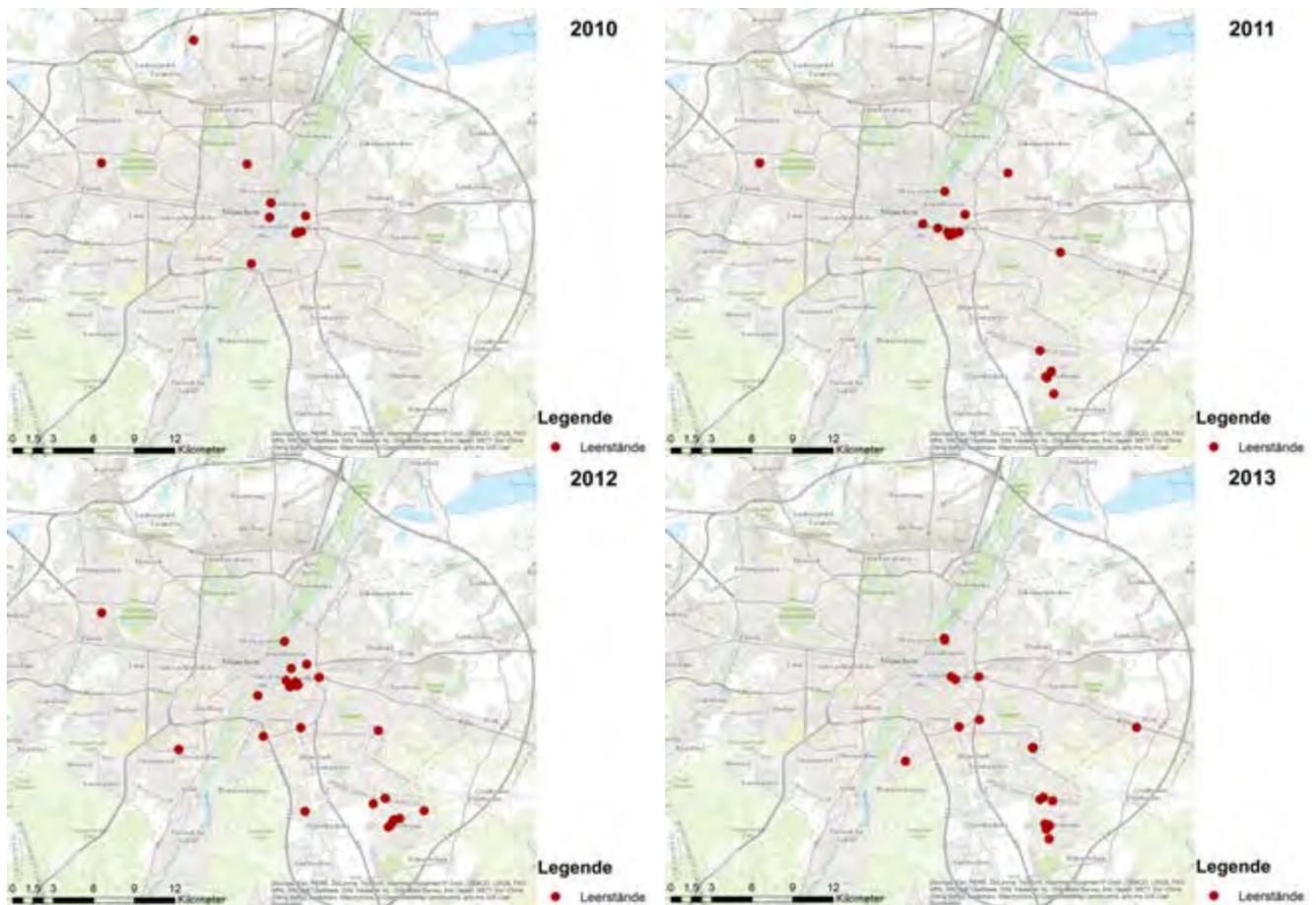
Die detaillierte Datengrundlage kann Informationen bieten, welche von Behörden nicht erhoben werden. Der Leerstand von Gebäuden ist eine Kategorie, welche in OSM über ein dokumentiertes Datenschema verfügt und die gleichzeitig nicht auf der OSM-Karte angezeigt wird. Zugleich werden Leerstände von Behörden nicht systematisch erhoben. Somit ergibt sich das Potenzial eines Leerstandsmonitoring auf Basis von OSM-Daten.

Dies wurde hier am Beispiel von München im Zeitraum von 2010 bis 2013 erstellt. Abbildung 6 zeigt die so entstandenen Karten. Die Karten der Jahre 2010 und 2011 zeigen noch sehr wenige Daten, weshalb man hier davon ausgehen kann, dass der Datenbestand noch erhoben wurde. Ab dem Jahr 2012 hingegen lassen sich erste räumliche Konzentrationen erkennen. Durch die zeitliche Auflösung kann man gleichzeitig die dynamischen Veränderungen nachverfolgen. Dabei ist zu beachten, dass das Beispiel des Monitoring ohne die Berücksichtigung auf *Map-guards* und deren Gewährleistung von Datenintegrität und Aktualität gezeigt wurde. Grundsätzlich kann man jedoch erkennen, dass die Verteilung von Leerständen und damit auch anderer OSM-Daten über einen Zeitraum hinweg verfolgt werden kann.

Landnutzungskarte

In einer Arbeit von Vaz und Arsanjani (2015) wurde gezeigt, wie man auf Basis von OSM-Daten Landnutzungskarten erstellen kann. Wie man in Abbildung 7 erkennen kann, verfügt die so entstandene Karte für den Raum Toronto über mehr Details als offizielle, von Behörden herausgegebene Karten. Gleichzeitig stellt eine auf diese Weise erstellte Landnutzungskarte eine kosteneffiziente Lösung dar. Zudem ist sie, aufgrund der dynamischen OSM-Daten, zeitlich skalierbar.

Abbildung 6: Zeitliche und räumliche Verteilung von Leerständen in München, basierend auf OSM-Geodaten von 2010 bis 2013



Eigene Darstellung

Landnutzungskarten aus OSM-Daten können also eine Alternative zu behördlichen Landnutzungskarten darstellen. Durch das Ground-Truth-Prinzip erlauben diese Karten zudem einen Einblick, wie die Lage vor Ort tatsächlich ist, nicht wie sie laut Bebauungsplan oder Flächennutzungsplan sein müsste.

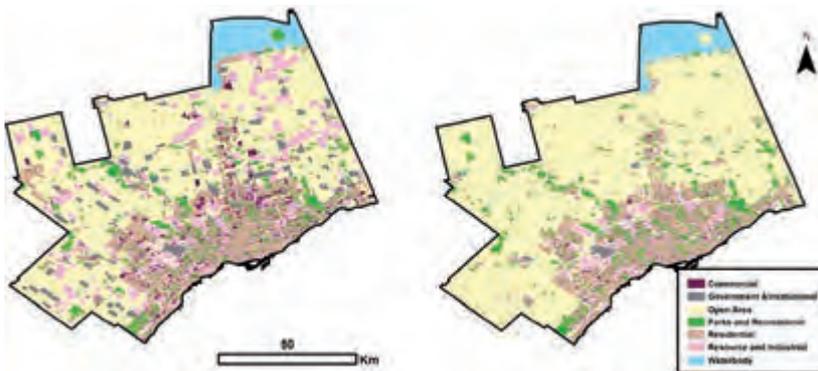
Fazit

In diesem Artikel wurde OpenStreetMap als ein Beispiel für digitale Geodaten vorgestellt. OSM ist ein VGI-Projekt, das nicht mit amtlichen Organisationen und Strukturen zu vergleichen ist. Das zeigt sich schon beim Zugang: OSM steht unter der Open Database License und ist sowohl dem Open Data, als auch den Open Source Bewegungen zuzuschreiben. Die Daten selbst werden von der sehr heterogenen *community* freiwillig erhoben. Das gesamte Projekt ist *community*-orien-

tiert aufgebaut. Das heißt, alle Aspekte werden in erster Linie *community*-freundlich ausgerichtet. Erst danach kommt z. B. Benutzerfreundlichkeit für Endnutzer außerhalb der *community*. Das zeigt sich auch bei den Entscheidungsprozessen: zwar wirken diese auf den ersten Blick komplex und undurchsichtig. Sie sind jedoch grundsätzlich demokratisch und transparent aufgebaut. Gleichzeitig erlauben die Strukturen eine hohe Offenheit und Flexibilität, wodurch OSM sich neuen Herausforderungen und Ideen schnell anpassen kann. Auch die Datenqualität ist stark *community*-abhängig. Sind lokal besonders aktive Beitragende vorhanden, sogenannte *Map-guards*, welche die Daten administrieren, kann von einer hohen Aktualität, Datendichte und -integrität ausgegangen werden. Die Identifizierung und Lokalisierung von *Map-guards* ist dabei eine wichtige Herausforderung für die Nutzung von OSM-Daten. Dies muss das Ziel für zukünftige Forschungsarbeiten sein.

Insgesamt haben die OSM-Geodaten bereits heute, vor allem in großen Städten, einen hohen Detailgrad erreicht. Zuverlässige Aussagen, die mehrere Jahre in die Vergangenheit reichen, sind allerdings bisher nur in Ausnahmen möglich. Wenn die OSM-Beitragenden in den kommenden Jahren weiterhin so aktiv bleiben, sind in Zukunft detaillierte raum-zeitliche Analysen auf Basis von OSM-Daten möglich.

Abbildung 7: Landnutzungskarte von Toronto. Basierend auf OSM-Geodaten (links) und amtlicher Erhebung (rechts)



Quelle: Vaz, Arsanjani 2015: 8

1 Dieser Beitrag beruht auf einem Vortrag des Autors im Rahmen der Sitzung „Jenseits von Registern und Umfragen: Nutzung und Potenzial von neuen Datenquellen für Planung und Politik“ aus dem Programm des VDSt zur Statistischen Woche 2015 in Hamburg.

Literatur

Budhathoki, Nama Raj; Haythornthwaite, C. (2013): Motivation for Open Collaboration Crowd and Community Models and the Case of OpenStreetMap. In: *ABS* 57 (5), S. 548–575. Community Wiki: DoOcracy. Online verfügbar unter: <https://communitywiki.org/de/DoOcracy>.
 Elrick, Tim; Glasze, Georg (2015): Maps & Mosques: (k)eine kleine Signatur und die Transformation von Geoinformation und Kartographie im digitalen Zeitalter. Neue Kultur-geographie. Bamberg, 2015.
 Elwood, Sarah; Goodchild, Michael F.; Sui, Daniel Z. (2012): Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice 102 (3), S. 571–590. DOI: 10.1080/00045608.2011.595657.
 Glasze, Georg; Perkins, Chris (2015): Social and Political Dimensions of the OpenStreetMap Project: Towards a Critical Geographical Research Agenda. In: Jamal Jokar Arsanjani, Alexander Zipf, Peter Mooney und Marco Helbich (Hg.): *OpenStreetMap in GIScience. Experiences, Research, and Applications*. Cham: Springer International Publishing (Lecture notes in geoinformation and cartography), S. 143–166.

Glasze, Georg; Plennert, Matthias (2015): Die Rolle von software für digitale Geoinformation: das OpenStreetMap-Geodatenmodell. Deutscher Kongress für Geographie. Deutsche Gesellschaft für Geographie. Berlin, 2015.
 Goodchild, Michael F. (2007): Citizens as sensors: the world of volunteered geography. In: *GeoJournal* 69 (4), S. 211–221.
 Ludwig, Ina; Voß, Angi; Krause-Traudes, Maike (2010): Ein Vergleich der Straßennetze von OpenStreetMap und NAVTEQ. In: Josef Strobl (Hg.): *Angewandte Geoinformatik 2010*. Beiträge zum 22. AGIT-Symposium Salzburg. Berlin, Offenbach: Wichmann.
 Mooney, Peter; Napolitano, Maurizio (2012): MVP OSM: A Tool to identify Areas of High Quality Contributor Activity in OpenStreetMap. In: *The SoC Bulletin*.
 Neis, Pascal; Zipf, Alexander (2012): Analyzing the Contributor Activity of a Volunteered Geographic Information Project – The Case of OpenStreetMap. In: *IJGI* 1 (2), S. 146–165. DOI: 10.3390/ijgi1020146.
 Open Data Commons: ODC Open Database License (ODbL) Summary. Online verfügbar

unter <http://opendatacommons.org/licenses/odbl/summary/>
 Open Definition: The Open Definition. Online verfügbar unter: <http://opendefinition.org/>
 Plennert, Matthias (2014): Vom Map-Gardening zum Map-Guarding. Territoriale Administration in OpenStreetMap.
 Ramm, Frederik (2013): Wer ist der Boss bei OpenStreetMap? Online verfügbar unter: http://wiki.openstreetmap.org/wiki/FOSSGIS_2013/Videomitschnitte.
 Roick, Oliver; Neis, Pascal; Zipf, Alexander (2011): Volunteered Geographic Information – Datenqualität und Nutzungspotentiale am Beispiel von OpenStreetMap. Deutsche Gesellschaft für Kartographie. Königslutter, 2011.
 Vaz, Eric; Jokar Arsanjani, Jamal (2015): Crowdsourced mapping of land use in urban dense environments. An assessment of Toronto. In: *The Canadian Geographer/Le Géographe canadien* 59 (2), S. 246–255. DOI: 10.1111/cag.12170.
 Wood, Harry (2014): The Long Tail of OpenStreetMap. Online verfügbar unter: <http://harrywood.co.uk/blog/2014/11/17/the-long-tail-of-openstreetmap/>