

Bachelor-Thesis 2011

OpenAddresses goes OpenStreetMap



Autor: Manuel Bröchin

Examinator: Prof. Hans-Jörg Stark

Experte: Jean-Marc Buttlinger

OpenAddresses goes OpenStreetMap

Diese Arbeit konzipiert und implementiert einen Synchronisationsmechanismus zwischen den beiden Projekten OpenAddresses (OA) und OpenStreetMap (OSM), welche auf dem Konzept von Volunteered Geographic Information basieren. Die Gebäude-Adressen der beiden Datenquellen sollen synchronisiert werden. Herausfordernd dabei sind die unterschiedlichen Datenstrukturen in den beiden Projekten.

Schlagworte: Adressen, OpenAddresses, OpenStreetMap, Osmosis, Synchronisation, PostgreSQL

1. Bestehende Projekte von Freiwilligen

1.1. OpenAddresses

OA ist eine Onlineplattform, welche Adressen georeferenziert erfasst. Die Adressen werden als Punkte mit den dazugehörigen Adressattributen gespeichert. Ein Adressformular und eine Online-Karte unterstützen den Benutzer. Adressdaten von OA können kostenlos von Interessierten bezogen und verwendet werden. Im Rahmen dieser Arbeit wurde dieses laufende Projekt mit OSM synchronisiert, d.h. Anpassungen an der Datenbank wurden nur auf Seiten OA vollzogen.

1.2. OpenStreetMap

OSM ist eine freie Geodatenbank der Welt. Aus den Daten werden beliebige Karten generiert oder sie dienen als Grundlagendaten für weitere Anwendungen. OSM besitzt mit über 400'000 registrierten Personen eine grosse weltweite Community, welche Daten sammelt und verarbeitet. Die metrischen Grundelemente der Datenbank sind Punkte, Linienzüge und Relationen. Beliebige Eigenschaften können den Grundelementen mittels Schlüssel und Wert (XML-Tags), zugewiesen werden – so auch Adressattribute.

2. Synchronisation

2.1. Datenströme

OSM bietet einen Datendownload an, welcher die Änderungen einer Minute, Stunde oder eines Tages bezüglich des gesamten Datenbestandes enthält. Der gesamte Datenbestand bzw. die entsprechenden Änderungsdaten sind mittels Osmosis, einer Kommandozeilen-Java-Applikation, in eine PostgreSQL Datenbank eingetragen worden. Darauf extrahieren eigens geschriebene Datenbanktrigger die Adressattribute mit dazugehöriger Geometrie und übergeben sie zur Synchronisierung und Speicherung an OA.

Löst ein Benutzer über die OA-Webanwendung eine Datenbankänderung aus, wird diese unmittelbar ausgelesen und über die Programmierschnittstelle (API) von OSM in deren Datenbestand eingepflegt. Allfällige auftretende Konflikte werden mit entsprechenden Kontrollmechanismen behandelt.

2.2. Kontrollmechanismen

Als Kern für den Kontrollmechanismus wurde die Versionierung der Daten von OSM verwendet. Das heisst, jedes Grundelement enthält eine Versionsnummer. Ändert sich etwas am Grundelement, wird dessen Versionsnummer um eins erhöht. Somit kann über diese

Versionsnummer bei einer Synchronisation festgestellt werden, ob der OSM- oder der OA-Datenbestand jünger ist und an welcher Stelle eine Nachführung notwendig ist.

2.3. Datenaufbereitung

Bevor die Initial-Synchronisation angestossen werden kann, muss geklärt sein, wie mit den schon vorhandenen Daten umgegangen wird. Unter dem Aspekt den OSM-Datenbestand möglichst unberührt zu lassen, wird erst nach einer Benutzeraktion ein OA-Datensatz in OSM eingetragen. Dies impliziert, dass der Adress-Datenbestand bei OA liegen muss. Um Adressen auf vorhandene OSM-Gebäude referenzieren zu können, muss vorab für alle Gebäude je ein Punkt in OA angelegt werden. Ein Nachteil davon ist eine allfällige redundante Adresshaltung, da z.B. eine OSM-Adresse mit unvollständigen Adressangaben in OA nicht automatisch gelöscht oder ergänzt werden darf, auch wenn sie in OA vollständig erfasst ist.

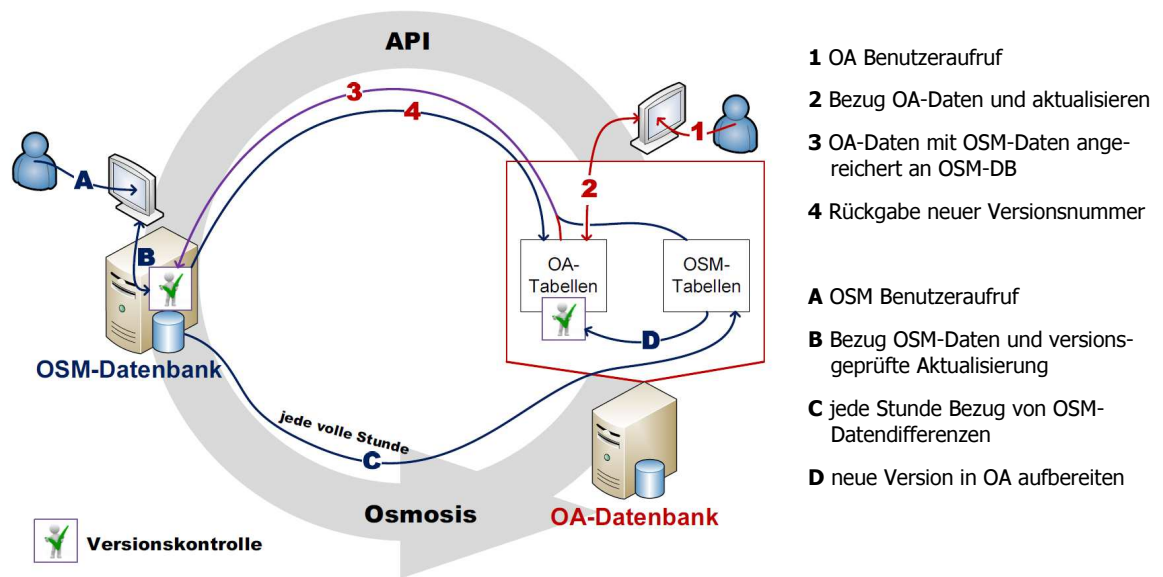


Abb. 1 Darstellung der Datenströme inkl. Kontrollmechanismen

3. Fazit

Es ist gelungen, knapp 41 Millionen Gebäude aus der OSM Datenbank nach OA als Punkte zu transferieren. Dank des implementierten Synchronisationskonzepts, erhält das OSM-Projekt mit OA ein neues Werkzeug, um Adressen einfach zu erfassen und zu bearbeiten: OA trägt nun automatisch die Anwender generierten Adressendaten in OSM ein. Dies funktioniert sowohl für neu erhobene als auch bearbeitete Adressen, die bereits existieren. Ausserdem visualisiert OA neu die Vollständigkeit der Adressangaben mittels Farbcodierung direkt in der Karte und bietet somit einen schnelle Übersicht über die Qualität des Adressdatenbestands. Des Weiteren konnte auch die regelmässige stündliche Synchronisation von OSM nach OA erfolgreich implementiert werden.

Autor:	Manuel Bröchin	manuel.broechin@students.fhnw.ch
Examinator:	Prof. Hans-Jörg Stark	hansjoerg.stark@fhnw.ch
Experte:	Jean-Marc Buttlinger	jean-marc.buttlinger@bl.ch