

FREIE UND OPEN SOURCE SOFTWARE **EINE CHANCE FÜR DIE WASSERWIRTSCHAFT**

Arnulf Christl, WhereGroup Bonn

Arnulf Christl ist Entwicklungskoordinator für räumliche Datenverarbeitung und Projektleiter der Software Mapbender. Er berät bei Konzeption, Aufbau und Umsetzung von Geodateninfrastrukturen und -Portalen. 1998 gründete er die Firma CCGIS in Bonn (seit 2007 WhereGroup <http://www.wherogroup.com>), deren Geschäftsmodell seit 2002 der Aufbau von Geodateninfrastrukturen unter Verwendung von Open Source Software Komponenten ist. Er ist Mitbegründer der Open Source Geospatial Foundation (<http://www.osgeo.org>).

Einführung

Der Einsatz von Freier und Open Source Software ist selbstverständlich nicht nur auf die Wasserwirtschaft beschränkt, doch gerade hier ergeben sich durch den Einsatz vielfältige Chancen. Der folgende Beitrag zeigt am Beispiel Fragestellungen aus der Wasserwirtschaft, die eine Bearbeitung geographischer Daten erfordert, wie durch eine sinnvolle Kombination von Freier Software und offenen Standards kostengünstig optimierte, ausbaufähige, vernetzte Systeme implementiert werden können. Als ein Beispiel sollen dabei folgende Punkte Berücksichtigung finden:

- Die Landeswassergesetze und Verordnungen schreiben Übersichtspläne zu den Charakteristika der Einzugsgebiete vor, die eine ökonomische Analyse der Gewässernutzung enthalten.
- Diese Daten werden durch eine Liste der Schutzgebiete (FFH - Natura 2000, Vogelschutzgebiete, Trinkwasserschutzgebiete, etc.) etc. ergänzt.
- Einzugsgebiete (Grundwasser-Reservoir, etc.) leiten sich häufig nicht aus Verwaltungs- oder hoheitlichen Grenzen ab, sondern folgen naturräumlichen Gegebenheiten.
- In zunehmendem Maße werden Bürgerrechte zur Informationsfreiheit festgeschrieben und auch eingefordert. Diese müssen auch technisch ermöglicht werden.

Die Fragestellungen, damit verbundene Datenbestände und vor allem die heterogenen Anforderungen der Akteure (z.B. Datenanbieter unterschiedlicher Nationalitäten) und Nutzer (Verwaltungsmitarbeiter aber auch Bürger) lassen sich optimal mit Open Source Entwicklungsmethodik und dem kollaborativen Ansatz Freier Software umsetzen.

Open Source, Freie Software und Offene Standards

Bevor ein Lösungsansatz für diese Problemstellungen entwickelt wird, erfolgt eine Erläuterung der Begriffe Open Source, Freie Software und offene Standards.

Gerade im Bereich geographischer Software (GIS) und noch stärker bei Geodateninfrastrukturen (GDI) haben sich Open Source Entwicklungsmodelle und Freie Software Lizenzmodelle bereits auf breiter Front durchgesetzt. Ein wichtiger Erfolgsfaktor ist dabei die konsequente Anwendung internationaler und offener Industriestandards.

Open Source

Der Quelltext einer Software wird im englischen Sprachgebrauch als "Source Code" bezeichnet. Er enthält alle Funktionen und Eigenschaften einer Software, die in einer Programmiersprache kodiert sind. Dieser Text ist für Menschen lesbar, aber nicht direkt von Maschinen ausführbar. In einem weiteren Schritt wird der Quelltext in eine maschinenlesbare Form überführt (kompiliert). Das Ergebnis sind binäre und damit vom Menschen nicht mehr direkt lesbare und damit unveränderbare Dateien, die aber von Maschinen ausgeführt werden können. Diese Binärdateien bilden das eigentliche Softwareprogramm, das von Endanwendern genutzt wird.

Der Begriff "quelloffene Software" oder Open Source wird für Software verwendet, deren Quelltext oder Programmcode offengelegt ist. Open Source ist damit zunächst eine Entwicklungsmethodologie. Die Offenlegung des Quelltextes als "Open Source" alleine ist für die meisten Endanwender jedoch weitgehend irrelevant, da hierdurch noch nicht definiert ist, in welcher Form die Software von wem genutzt werden kann.

Wie der Quellcode oder die daraus hergestellten Binärdateien genutzt, verändert, kopiert und erweitert werden dürfen wird in Lizenzmodellen geregelt. Die Open Source Initiative [4] (<http://www.opensource.org>) erweitert den Begriff der Quelloffenheit deshalb um das Recht der Allgemeinheit den Quellcode nicht nur einsehen zu dürfen, sondern auch Änderungen vorzunehmen und gestattet, für die Endanwender das gewichtigere Argument, dass die aus diesen Quellen kompilierte Software so oft und zu welchem Zweck auch immer eingesetzt und genutzt werden kann. Die Rechte und Freiheiten der Anwender und Entwickler einer Software werden in den entsprechenden Lizenzmodellen festgeschrieben.

Die Vorteile einer Open Source Entwicklungsmethodologie werden sehr treffend in Eric S. Raymonds Aufsatz »The Cathedral and the Bazaar« (<http://www.catb.org/~esr/writings/homesteading/cathedral-bazaar/>) beschrieben.

Freie Software

Der Begriff "Free Software" wurde bereits 1984 durch Richard Stallman im GNU-Projekt geprägt (GNU ist das rekursive Akronym von "GNU is not Unix", <http://www.gnu.org>). Die Open Source Entwicklungsmethodologie ist eine der Voraussetzungen für Freie Software. Sie betont zusätzlich den gesellschaftlichen und kulturellen Wert von Software und ist Grundlage der weit verbreiteten Software Lizenz GNU GPL (GNU General Public License), durch die auch der Betriebssystem-Kernel Linux geschützt wird.

Freie Software bedeutet nicht primär, dass die Software kostenfrei genutzt werden kann (das ist ein schöner Nebeneffekt), sondern dass die Freiheiten der öffentlichen Wissensgemeinschaft dauerhaft gewahrt werden. Um dieses Ziel zu erreichen wird ein rechtlich sicheres Konstrukt benötigt, in dem diese Rechte und Freiheiten geregelt werden. Diese Art von Rechten werden üblicherweise in Form von Lizenzen geregelt und unter dem Begriff Copyright zusammengefasst. In Anlehnung dazu wurde das Copyleft erfunden, das auf den Seiten der Free Software Foundation (<http://www.fsf.org>) wie folgt beschrieben wird:

»Um ein Programm unter das Copyleft zu stellen, stellen wir es zuerst unter das Copyright; dann fügen wir als Rechtsmittel Vertriebsbedingungen hinzu, die jedermann das Recht geben, den Code des Programms oder jedes davon abgeleiteten Programms zu nutzen, zu ändern oder weiter zu verteilen, aber nur, wenn die Vertriebsbedingungen unverändert bleiben. So werden der Code und die gewährten Freiheiten rechtlich untrennbar.«

Open Source ist dabei eine Voraussetzung für Freier Software, da nur die Offenlegung des Quelltextes gewährleistet, dass Anwender die Software selbst verändern oder durch Dritte verändern lassen können. Open Source in Verbindung mit dem rechtlichen Vehikel Copyleft ermöglicht die umfassende Definition von Freier Software wie sie in der GNU GPL Lizenz festgeschrieben ist.

Die GNU General Public License

Die Präambel der GNU GPL (<http://www.gnu.org/licenses/gpl.html>) beginnt mit der Feststellung: »Die meisten Softwarelizenzen sind dafür gemacht, Ihnen die Freiheit wegzunehmen, Software mit anderen zu teilen und sie zu verändern. Im Gegensatz dazu hat die GNU General Public License die Absicht, Ihre Freiheit abzusichern, Freie Software zu teilen und sie zu verändern – um sicherzustellen, dass die Software für alle ihre Nutzer frei ist.«

Um sicherzustellen, dass die Software auch in Zukunft frei bleibt, unterliegen die Freiheiten Bedingungen, »die es jedem verbieten, dem Nutzer diese Freiheiten zu verweigern oder ihn auffordern, auf sie zu verzichten«. Das Hinzufügen weiterer Restriktionen wird untersagt. Die Freiheiten und ihre Bedingungen umfassen im Einzelnen:

- Die Freiheit, das Programm für jeden Zweck auszuführen (Ziff. 0)
- die Freiheit, den Quellcode des Programms wörtlich zu kopieren und zu verbreiten, sofern der Copyright-Vermerk und die Lizenz mit kopiert und verbreitet wird. Die Erhebung einer Gebühr für die physikalische Übertragung einer Kopie und für andere Dienstleistungen, wie eine Gewährleistung, wird ausdrücklich erlaubt (Ziff. 1),
- die Freiheit, das Programm zu verändern und diese veränderte Version zu kopieren und zu verbreiten, sofern das abgeleitete Werk Angaben über die Änderung enthält und gebührenfrei und unter denselben Lizenzbedingungen veröffentlicht wird, wie das ursprüngliche Programm. Von der Bedingung ausgenommen sind Teile des veränderten Programms, die unabhängige Werke darstellen und separat verbreitet werden (Ziff. 2)
- die Freiheit, das Programm oder abgeleitete Versionen in Objektcode oder ausführbarer Form zu kopieren und zu verbreiten, sofern der dazugehörige maschinenlesbare Quellcode oder ein schriftliches, mindestens drei Jahre gültiges Angebot, diesen Quellcode auf Anfrage bereitzustellen, beigefügt ist (Ziff. 3).

Neben der GNU GPL gibt es eine Vielzahl weiterer Open Source und Free Software Lizenzen, die sich zum Teil erheblich in der Definition der Freiheiten und Rechte der Anwender unterscheiden. Bevor Sie eine Lizenzbedingungen akzeptieren, sollten Sie deshalb immer prüfen, um welche Form der Lizenzierung es sich genau handelt (http://www.fsf.org/licensing/licenses/index_html). Das gilt natürlich für jegliche Software, egal ob Frei oder proprietär.

Eine umfassende Beschreibung der Unterschiede zwischen Freier Software Entwicklung und proprietären Geschäftsmodellen wurde im Auftrag der Bundeszentrale für Politische Bildung von Volker Grassmuck in deutscher Sprache zusammengestellt und kann kostenfrei aus dem Internet heruntergeladen werden (<http://freie-software.bpb.de/>).

Offene Standards

Offene Standards bedienen sich eines an die Open Source Entwicklungsmethodologie angelehnten Konzepts, um konsensgetrieben Festlegungen über Formate und Schnittstellen zu definieren. Die Offenheit bezieht sich dabei hauptsächlich auf die öffentliche Verfüg- und Nutzbarkeit dieser Standards. Bitte beachten Sie aber, dass es keinen inhärenten Zusammenhang zwischen Open Source und offenen Standards gibt. Allerdings sind Freie Software Lizenz- und Open Source Entwicklungsmodelle besonders geeignet, um Software mit offenen Standards zu implementieren.

Im GIS-Umfeld haben sich die Standards des OGC (Open Geospatial Consortium, <http://www.opengeospatial.org/>) durchgesetzt, die von einem freiwilligen Zusammenschluss von privatwirtschaftlichen, öffentlichen und wissenschaftlichen Organisationen in enger Abstimmung mit der ISO (International Organisation for Standardisation, <http://www.iso.org/>) entwickelt werden.

Quasi-Standards

Offene Standards sind von sogenannten Quasi-Standards zu unterscheiden, wie sie z.B. durch die Marktbeherrschung eines einzelnen Herstellers entstehen können. Zu solchen Quasi-Standards

zählen z.B. das Dokumentformat PDF der Firma Adobe (<http://www.adobe.com/>) oder das Textformat DOC der Firma Microsoft (<http://www.microsoft.com/>). Auch wenn diese Formate eine hohe Marktdurchdringung erzielt haben, handelt es sich dabei keineswegs um offene Standards.

Durch die Verwendung dieser Formate unterwirft sich der Anwender den Vorgaben des jeweiligen Besitzers (z.B. ein Softwarehersteller). Das ist vor allem dann problematisch, wenn durch die Speicherung in einem bestimmten Format Inhalte nur noch unter Verwendung einer bestimmten Software eingesehen oder verwandelt werden können. Durch eine Änderung dieser Software oder den Lizenzbedingungen kann der Besitzer der Copyright-Rechte dann sogar maßgeblichen Einfluss auf weitere die Verwendbarkeit der Inhalte nehmen.

Diese Einsicht hat sich herumgesprochen und inzwischen hat auch Microsoft reagiert und wird in einer der nächsten Versionen der Office-Software auf ein Standard XML Format anbieten (<http://www.microsoft.com/presspass/features/2005/jun05/06-01XMLFileFormat.mspx>).

Aufbau von Geodateninfrastrukturen für die Wasserwirtschaft

Nach dieser kurzen Einführung in Open Source Entwicklungsmodelle und Freie Software Lizenzmodelle kommen wir zurück zu der Eingangs gestellten Ausgangssituation und beleuchten die Möglichkeiten, die eine GDI unter Verwendung Freier Software Komponenten bietet.

Übersichtspläne zu den Charakteristika der Einzugsgebiete

Zur Erstellung und Analyse der Übersichtspläne wird ein Geographisches Informationssystem (GIS) benötigt, welches heute zweckmäßigerweise als interoperable Dienstarchitektur mit geographischer Funktionalität ausgeprägt sein sollte.

Für die Digitalisierung neuer Geometrien über eine GDI wird ein sogenannter Fat Client (früher Desktop-GIS) eingesetzt. Die Software sollte in der Lage sein, Kartenwerke als Digitalisierungsgrundlage dynamisch über ein Netzwerk als Dienst einzubinden (siehe OGC WMS weiter unten).

Die neu erstellten Geometrien (die Basis der Übersichtspläne) werden heute üblicherweise in einer zentralen Datenbank gespeichert, um eine möglichst hohe Flexibilität bei der Weiterverarbeitung zu ermöglichen. Die Datenbank muss über räumliche Datentypen und Operationen verfügen. Die neu erstellten Objekte werden über die ebenfalls standardisierte Zwischenschicht OGC WFS-T in die Datenbank zurückgeschrieben.

Damit der Ersteller der Daten deren Qualität und Richtigkeit gewährleisten kann, behält er zweckmäßigerweise die Hoheit über diesen Teil des Systems, indem er es selbst aufbaut.

Durch die Verwendung der zentralen Datenbank wird die Sicherung und Pflege der Daten vereinfacht. Es können mehrere Arbeitsplätze gleichzeitig darauf zugreifen, sowie auch weitere Dienste, die eine Beauskunftung der Daten (siehe unten) ermöglichen.

Ergänzung durch Schutzgebiete (FFH - Natura 2000, Vogelschutzgebiete, Trinkwasserschutzgebiete, etc.)

Die ergänzenden Daten werden nicht selbst erstellt, sondern von Dritten bezogen. Die Daten werden heute allerdings nicht mehr kopiert, sondern dynamisch über ein Netzwerk direkt vom Anbieter eingebunden. Der Zugriff erfolgt über die Standardschnittstelle OGC WMS, die auch vom Bundesamt für Kartographie und Geodäsie (BKG) empfohlen wird. Das BKG (<http://www.bkg.bund.de/>) betreibt eine Geodateninfrastruktur und bietet über das Geodatenzentrum (<http://www.geodatenzentrum.de/>) selbst Kartenwerke über diese Schnittstelle an.

Das bedeutet, dass Geodaten nicht mehr vollständig physikalisch kopiert werden müssen, sondern dynamisch über ein Netzwerk immer nur der Teil in dem Maßstab angefordert wird, der gerade benötigt wird. Hochaufgelöste Luftbilder belegen beispielsweise enorm viel Festplatten-Speicher und stellen ausserdem hohe Hardwareanforderungen an den Digitalisierungsarbeitsplatz. Bei jeder Aktualisierung der Daten wäre es zusätzlich erforderlich den Datenbestand erneut zu kopieren. Beim Zugriff auf eine Geodateninfrastruktur verbleiben die Daten auf dem Server des Anbieters, der sich um ausreichenden Speicherplatz, Ausfallsicherheit, regelmäßige Aktualisierung und Pflege kümmert.

Grenzüberschreitende Zusammenarbeit

Einzugsgebiete, Flußläufe, Grundwassergebiete, etc. leiten sich häufig nicht aus Verwaltungs- oder hoheitlichen Grenzen ab, sondern folgen naturräumlichen Gegebenheiten. Gerade in der Wasserwirtschaft erfordert das oft eine grenzüberschreitende Zusammenarbeit. Diese wird erheblich vereinfacht, wenn sich beide Seiten an die gleichen Standards halten. Hierbei ist zu beachten, dass es sich nicht nur um technische, sondern auch um inhaltliche Standards handeln kann. Die X-Border Initiative (<http://www.x-border-gdi.org/>) berücksichtigt diese Punkte.

Hier können die Vorteile der Open Source Entwicklungsmethodik auch ausserhalb der reinen Softwareentwicklung zum Einsatz kommen, indem die beteiligten Partner gemeinsame Methoden entwickeln, wie die Daten abgebildet werden. Die technische Standardisierung ist bereits so weit gediehen, dass eine Kooperation hier keine größeren Probleme mehr verursacht. Die inhaltliche Zusammenarbeit ist wesentlich schwieriger, da sie oft mit unterschiedlichen Sprachen erfolgt und zusätzlich kulturelle und legislative Unterschiede überbrücken muss.

Hier bietet sich der Einsatz kollaborativer Wissensmanagementsysteme an, wie sie z.B. durch Wiki Software ermöglicht werden (<http://de.wikipedia.org/wiki/Wissensmanagement>).

Durch die vollständige, technische Interoperabilität wird praktisch nebenbei eine erhebliche Qualitätssicherung im System verankert, da fehlerhafte Eingaben vor allem in Grenzbereichen schnell offensichtlich werden.

Bürgerinformation über das Internet

Nachdem die Übersichtspläne erfolgreich erstellt wurden, sollen sie auch veröffentlicht werden, um einerseits der Informationspflicht dem Bürger gegenüber gerecht zu werden und andererseits externen Dritten einen geschützten Zugriff auf ausgewählte Dateninhalte zu ermöglichen.

Die Veröffentlichung erfolgt unter Verwendung der gleichen Komponenten, die bereits bei der Erstellung zum Einsatz kamen. Ergänzend wird eine internetfähige Anwendungsoberfläche erstellt, die auch dem technisch nicht versierten Bürger ermöglicht die Kartenwerke einzusehen.

Um den Anforderungen des Datenschutz gerecht zu werden, wird der Zugriff auf schützenswerte Daten durch eine Berechtigungsstruktur geregelt und verschlüsselte Netzwerkverbindungen eingesetzt. Damit wird es möglich, sowohl lesend als auch schreibend, transparent nachvollziehbar auf den zentralen Datenbestand zuzugreifen.

Die Verwendung standardkonformer Textverarbeitung rundet den Einsatz von Open Source Komponenten bis zum Büro-Endarbeitsplatz ab.

Vorteile durch Freie Software

Das beschriebene Szenario lässt sich mit verschiedenen Architekturen umsetzen. Dabei ist es von technischer Seite her zunächst weitgehend unerheblich, ob Freie oder proprietäre Lizenzmodelle zum Einsatz kommen. Die Technologieentscheidung kann also – die Einhaltung offener Standards vorausgesetzt – zunächst unabhängig von der Wahl des Lizenzmodells erfolgen.

Allerdings bieten Open Source Entwicklungsmodelle die Möglichkeit direkt Einfluss auf die Produktentwicklung zu nehmen. Vor allem bei speziellen Fragestellungen ist es deshalb möglich, kostengünstig und zeitnah Erweiterungen der Software selbst vorzunehmen oder durch Dritte vornehmen zu lassen. Die Finanzierung kann auch durch die Bildung kleinerer Interessengruppen ermöglicht werden, die ein gemeinschaftliches Ziel verfolgen.

Die Freiheiten, die durch die Lizenzierung als Freie Software gegeben sind gewährleisten, dass die getätigten Investitionen unbeschränkt genutzt werden können. Da die Nutzung der Software in keiner Weise verpflichtend ist, muss sie sich bewähren. Weitere Interessenten werden sich nur freiwillig für dieses System entscheiden, wenn es wirklich gut ist. Ab einer kritischen Masse an Anwendern, die bei jedem Softwareprojekt unterschiedlich ist, wird die Entwicklung zu einem Selbstläufer, da das Potential zur Weiterentwicklung mit jedem produktiven System wächst.

Softwareauswahl

Auf diese Weise ist bereits eine Vielzahl von Softwarekomponenten entstanden, die sofort und ohne irgendeine weitere Änderungen uneingeschränkt eingesetzt werden können. Zu den bekanntesten Beispielen aus der Mainstream-IT zählen mit einem weltweiten Marktanteil von über 70% der Webserver Apache http Server oder das Betriebssystem GNU/Linux. Diese hundertausendfach bewährten Softwarepakete werden durch die großen Systemhersteller (z.B. IBM, Sun oder Novell) im täglichen Einsatz betrieben. Der Apache Webserver, die Skriptsprache PHP und die Datenbank PostgreSQL sind sogar Bestandteil der meisten Linux-Distributionen und können direkt zur Installation ausgewählt werden. Sie bilden das Rückgrat der GDI, deren räumliche Datenverarbeitung (das GIS) durch folgende Komponenten implementiert werden kann:

- PostgreSQL Geo-Spracherweiterung: PostGIS (<http://postgis.refrations.net>)
- Kartenserver (OGC WMS): UMN MapServer (<http://www.umn-mapserver.de>)
- Geometrieserver (OGC WFS-T): GeoServer (<http://www.geoserver.org>)
- Internetfähiges Desktop-GIS: uDig (<http://udig.refrations.net>)
- Internet-Kartenclient, GDI Mangement-Suite: Mapbender (<http://www.mapbender.org>)

Auch GDI-Komponenten sind bereits bei einigen Linux Distributionen enthalten, z.B bei Debian (<http://www.debian.org>), das auch vom Bundesamt für Sicherheit in der Informationstechnik empfohlen wird.

Zusammenfassung und Fazit

Die weltweiten, unabhängigen Anwender- und Entwicklergemeinschaften gewährleisten die kontinuierliche, pragmatische und durch wirtschaftliche Partikularinteressen uneingeschränkte Weiterentwicklung und Pflege der Software. Der breite Einsatz dieser Software in vielen produktiven Systemen gewährleistet Investitionssicherheit. Die Geodatenverarbeitung entwickelt sich zunehmend in Richtung Mainstream-IT und wird zunehmend einfacher zu installieren und zu pflegen.

Die Einhaltung internationaler, offener Standards ist unabdingbare Voraussetzung für den Aufbau und Betrieb von Geodateninfrastrukturen. Die Fragestellung und die Natur der Daten im Bereich der Wasserwirtschaft erfordern eine hohen Vernetzung, die Interoperabilität unabdingbar machen. Das für den Aufbau und Betrieb der Systeme benötigte Know-How ist allgemeingültig und kann einfach übertragen werden, da es nicht auf eine bestimmte Software festgelegt ist.

Durch die ausschließliche Verwendung von Freien Software Lizenzen entstehen durch die erweiterte Nutzung auch einer Internet-basierten Anwendung keine zusätzlichen Lizenznutzungskosten. Falls die Nachfrage es erfordert, kann das Produktivsystem jederzeit verdoppelt, verdreifacht oder durch ein Sekundärsystem vollständig ersetzt werden. Dadurch kann mit einfachen Mitteln und geringen Kosten eine hohe Ausfallsicherheit des Produktivsystems bei gleichzeitig stetiger Weiterentwicklung erzielt werden.