

Redundante Metadaten in geocat.ch

Bericht

Es handelt sich hier um einen Bericht, der dazu dienen soll, **das Problem darzustellen**, und Inputs zu liefern, um, darauf aufbauend, gemeinsame Entscheidungen zu treffen. **Keinesfalls sollen Lösungen aufgezwungen werden.**

2006 **swisstopo**
Bundesamt für Landestopografie
Office fédéral de topographie
Ufficio federale di topografia
Uffizi federal da topografia
Federal Office of Topography

Redaktion: Annina Hirschi Wyss
Postfach, CH-3084 Wabern
Tel: +41 31 963 24 23
Fax: +41 31 963 24 59
E-mail: annina.hirschi-wyss@swisstopo.ch

Inhalt

1	Einleitung	5
1.1	Kontext	5
1.2	Die Metadaten-Werkzeuge	5
1.2.1	Die Norm (Das GM03-Modell)	5
1.2.2	Die Applikation (geocat.ch)	6
1.3	Problematik der redundanten Metadaten	7
2	Wiederverkauf und Zugehörigkeit	9
2.1	Definition	9
2.2	Analyse	10
2.3	Lösungen	10
2.3.1	Erklärungen zum Schema	11
2.3.2	Rekapitulation „Kriterien“	14
2.3.3	Rekapitulation „Vor- und Nachteile“	15
3	Hierarchisierung der Datensätze	17
3.1	Definition	17
3.2	Analyse	17
3.3	Lösungen	18
3.3.1	Beispiel der Bibliothekskataloge	18
3.3.2	Anpassung der Bibliotheks-Methode an geocat.ch	19
4	Beziehung geocat.ch / envirocat	21
4.1	Definition	21
4.2	Analyse	21
4.3	Lösungen	21
4.3.1	Gleiches Modell, gleiche Applikation	21
4.3.2	Gleiches Modell, andere Applikation	21
4.3.3	Anderes Modell, andere Applikation (Ist-Zustand)	22
5	Weitere Diskussionspunkte	25
6	Zum Schluss	26
7	Anhang	27
7.1	Glossar	27
7.2	Bibliografie	28
7.3	Illustrationsverzeichnis	28
7.4	Antworten der befragten Partner	I-XI

1 Einleitung

1.1 Kontext

Da die Menge der verfügbaren Geodaten immer grösser wird, ist es auch immer wichtiger diese zu dokumentieren, so dass man sie wiederfinden, identifizieren und brauchen kann.

Somit sind [Metadaten](#) (MD) ein wichtiger Bestandteil der [NGDI](#). [KOGIS](#) wurde beauftragt, die verschiedenen Geodaten-Anbieter dafür zu sensibilisieren, dass die MD ein wesentliches Element der Geodaten darstellen.

Mit der Bereitstellung des kostenlosen Werkzeuges [geocat.ch](#), werden die Dienststellen angespornt, ihre Metadaten zu erfassen. Um potentielle Partner nicht schon im Vorhinein abzuschrecken, müssen einzig die Bedingungen des Modells [GM03](#) eingehalten werden, auf zusätzliche Regeln und Vorschriften wurde bislang verzichtet.

1.2 Die Metadaten-Werkzeuge

1.2.1 DIE NORM (DAS GM03-MODELL)

Jede Datenbank hat ihr zugehöriges Modell. Je präziser es ist, desto einfacher der Datenaustausch. Die definierten und genormten Felder führen den Nutzer und garantieren eine gewisse Homogenität zwischen den verschiedenen erfassten Objekten.

Auf der internationalen Ebene hat ISO ein Metadaten-Modell erarbeitet (die Norm ISO 19115), welches eine sehr umfangreiche Beschreibung der Geodaten erlaubt.

GM03, welches durch [KOGIS](#) (in Zusammenarbeit mit verschiedenen Fachpersonen) entworfen wurde, ist eine kompatible Erweiterung der ISO-Norm. Sie ist der schweizerischen Umgebung und deren Besonderheiten (unter anderem der Mehrsprachigkeit) angepasst. Die zwei Profile der ISO-Norm, „Core“ und „Comprehensive“, wurden in GM03 erhalten, um mehr oder weniger detaillierte Beschreibungen zu ermöglichen.

GM03, die in [UML](#) und in [INTERLIS 2](#) beschriebene Schweizer Norm (SN 612050), benutzt als Transferformat den Standard XML/INTERLIS.

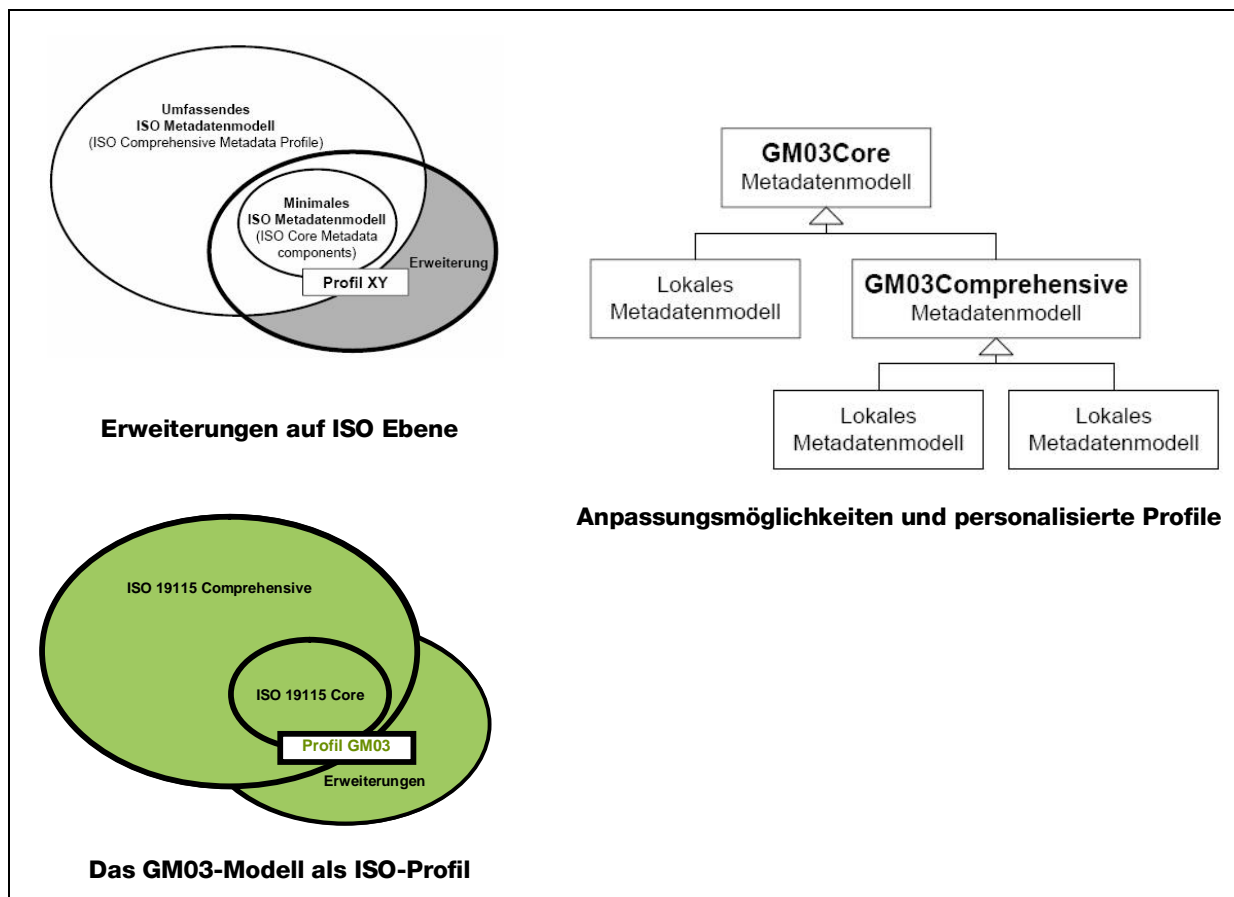


Fig. 1: Die verschiedenen Metadaten-Modelle

1.2.2 DIE APPLIKATION (GEOCAT.CH)

Die Anwendung geocat.ch basiert auf dem Modell GM03. Sie ist in der plattformunabhängigen JAVA-Sprache geschrieben und setzt sich aus dem Suchportal (Discovery service) und der Erfassungssapplikation (Metadata Management) zusammen. Eigene Profile sowie Templates können festgelegt werden, um die Katalogisierung der Daten zu vereinfachen.

Der Discovery service erlaubt einerseits in der **zentralen Datenbank** zu suchen und andererseits mit Hilfe eines Gateways **externe Kataloge** abzufragen. Drei Partnertypen stehen zur Auswahl. Die Partner A erfassen direkt in der zentralen Metadatenbank (in einem eigenen „Bereich“), die Partner B benutzen die Import/Export-Funktion, während die C-Partner unabhängig vom System sind und über den Gateway mit der Suchapplikation kommunizieren.

Die Erfassungssapplikation ist seit September 2004 in Betrieb und das Suchportal wurde im Januar 2005 aufgeschaltet.

Das folgende Schema zeigt die verschiedenen Zugriffs- und Teilnahmeoptionen von geocat.ch

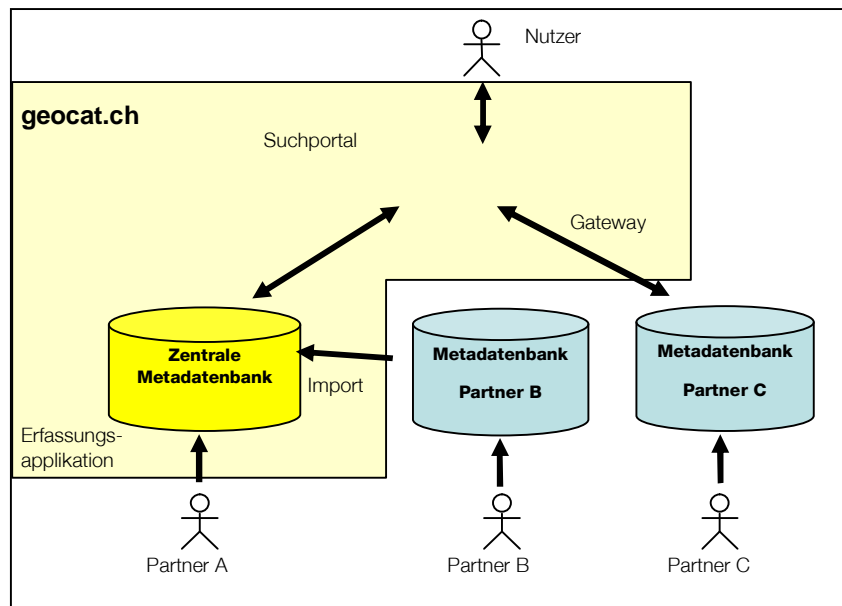


Fig. 2: Organisation von geocat.ch

1.3 Problematik der redundanten Metadaten

Nach zweijährigem Betrieb von geocat.ch ist es notwendig geworden, dessen Funktionen zu optimieren. Eine der Verbesserungen besteht darin, das Problem der redundanten Metadaten anzugehen und zu lösen.

Drei verschiedene Grundprobleme konnten ausgemacht und analysiert werden:

- 1) Als erstes die Redundanz, die auf die **Verteilung und die Zugehörigkeit von Geodaten** zurückzuführen ist. Sie entsteht sobald ein Amt erworbene Daten ein zweites Mal beschreibt, obschon der Datenherr sie bereits in geocat.ch erfasst hat. ([Kapitel 2](#))
- 2) Das folgende Kapitel ist der Problematik der **Hierarchisierungsfunktion** der Daten gewidmet, welche im Moment zu komplex ist, damit sie auch wirklich gebraucht werden kann, hat jedoch das Potential, mögliche Redundanzen verringern zu können. (Man beschreibt diejenigen Elemente nur ein Mal, die in einer Datenserie vorkommen). ([Kapitel 3](#))
- 3) Und drittens werden wir die **Beziehung zwischen envirocat und geocat.ch** analysieren, welche einige Zweideutigkeiten aufweist, die geklärt werden müssen. ([Kapitel 4](#))

Für jedes der Kapitel wird eine Definition, eine Analyse und Lösungswege vorgeschlagen. Die Argumente der Partner werden diskutiert und die Gründe erläutert, warum dieselben Datensätze in geocat.ch mehrere Male erscheinen.

2 Wiederverkauf und Zugehörigkeit

2.1 Definition

Die Unterschiede der Distribution der Daten zwischen den **Datenlieferanten des Bundes** und den **anderen Partnern** von geocat.ch (Kantone, Gemeinden, private Organisationen...) führen gelegentlich zu Redundanzen im Metadatenkatalog. Es handelt sich dabei vor allem um Daten, welche von den kantonalen Verwaltungen gebraucht werden und aus der Produktion von Bundesämtern stammen ([swisstopo](#), [BAFU](#) usw.).

Wir finden in geocat.ch also **mehrere Metadaten-Einträge, welche denselben Datensatz beschreiben** und sich nur durch kleine Variationen unterscheiden. Nehmen wir « SWISSIMAGE » als Beispiel. Es handelt sich dabei um ein digitales, flächendeckendes Orthofotomosaik der Schweiz. Dieses Produkt wird von swisstopo hergestellt und verkauft. Wie jedermann können auch die geocat.ch-Partner – zum Beispiel die Kantone – Ausschnitte aus diesem Mosaik erwerben und sie anschliessend zu bestimmten Konditionen, in der Regel mit Lizenzverträgen, an ihre Ämter und/oder Gemeinden weitergeben.

swisstopo erstellt nun Metadaten für sein Produkt. Die Kantone können diese Daten exportieren, wieder in geocat.ch importieren und sie an ihre Bedürfnisse anpassen (Beschreibung der Lizenz, geographische Ausdehnung, zusätzliche Attribute...) ([Fig. 3](#)). Theoretisch könnte man also **27 Metadatensätze für praktisch den gleichen Geodatensatz** finden.

Auch an der „alten Siegfriedkarte“ lässt sich diese Problematik gut erkennen:

http://gbs02.geotaskserver.com/swisstopo/getPresentationRequestResult.do?fileidentif=PK_SFK25_50&profile=GM03Comprehensive&locale=de&format=html (swisstopo)

http://gbs02.geotaskserver.com/sysif/getPresentationRequestResult.do?fileidentif=SYSIF_swisstopo_022&profile=GM03Comprehensive&locale=fr&format=html (Fribourg)

Ein weiteres Beispiel ist das Terrainmodell VECTOR25. Die Kantone, welche diese Daten von swisstopo erworben hatten, beschreiben die Datensätze (normalerweise auf ihr Gebiet beschränkt), indem sie entweder jeden einzelnen (von swisstopo erfassten) thematischen Layer¹ duplizieren, oder aber nur den allgemeinen Eintrag.

geocat.ch
geographic catalogue

Recherche simple Recherche avancée Résultat

Liste des résultats
Termes de recherche (recherche libre): swissimage

Nombre de résultats: 6 (1 - 5) Serveurs consultés >

◀ Orthophoto SWISSIMAGE 2001 (ACV)
Mosaïque d'orthophotos numériques couleur de swisstopo (Office fédéral de topographie). Informations complémentaires auprès de swisstopo : <http://www.swissimage.ch>
Contact: David Giorgis, Canton de Vaud - Service de l'information sur le territoire Pierre Vuagniaux, Canton de Vaud - Service de l'information sur le territoire Jérôme Lancia, Canton de Vaud - Service de l'information sur le territoire
30/09/04
Source: [ASIT-VL](#) [charger uniquement les résultats de ce serveur]

◀ SWISSIMAGE
La mosaïque d'orthophoto SWISSIMAGE est un assemblage d'images aériennes numériques en couleur. Une orthophoto est une image aérienne dont les influences
Contact: Grandjean, Centre de Compétence SIT, CCSIT
01/01/01 - de - 2
Source: [Canton de Fribourg](#) [charger uniquement les résultats de ce serveur]

◀ SWISSIMAGE
La mosaïque d'orthophoto SWISSIMAGE est un assemblage d'images aériennes numériques en couleur. Une orthophoto est une image aérienne dont les influences
Contact: Marketing und Verkauf Geodaten, Office fédéral de topographie
01/01/01 - de - 2
Source: [swisstopo](#) [charger uniquement les résultats de ce serveur]

Page suivante >>

> Imprimer liste
Préciser la recherche Nouvelle recherche

Fribourg und Waadt haben beschlossen, den Eintrag von swisstopo in ihren Katalog zu kopieren und anzupassen.

Dies ist der ursprüngliche Datensatz von swisstopo

Fig. 3: Beispiel redundanter Metadaten von « SWISSIMAGE »

¹ Jeder der neun thematischen Layer von VECTOR25 (Gewässernetz, Bodenbedeckung, Strassennetz...) ist in geocat.ch einzeln beschrieben. Dazu gibts noch einen Gesamteintrag.

2.2 Analyse

Für die Nutzer und Produzenten stellen diese Redundanzen ein Problem dar. Der/die „Suchende“ muss **mehrere Seiten durchblättern**, jedes Mal den Eintrag im Detail anzeigen, um zur gesuchten Information zu gelangen und es ist schwierig zu erkennen, welches genau der gesuchte Datensatz ist. Der [Datenherr](#) fragt sich, warum andere Partner sein Produkt (das er vertreibt) beschreiben und warum sein „Original“ nicht an erster Stelle erscheint.

Die folgenden **zwei Extremszenarien** reflektieren den Konflikt zwischen dem Schweizer Föderalismus und dem Willen eines zentralisierten, möglichst einheitlichen Datenzugangs.

- 1) Man nimmt an, dass es sich bei veränderten Daten oder **Auszügen eines Datensatzes** um „neue“ **Daten** handelt, die ihre eigenen Metadaten haben müssen. Man akzeptiert also die Redundanzen und es liegt an der Applikation oder am Datenmodell die Suchergebnisse so zu ordnen, dass es für den Nutzer sofort klar ist welches das „Original“ und welches die „Kopien“ sind.
- 2) **Redundanzen sind nicht erlaubt**. Ein Datensatz hat nur **einen einzigen Metadatensatz**, welcher durch den Datenherrn beschrieben ist, auch wenn er zerlegt, transformiert oder von Dritten gekauft werden kann. Die Anwendung muss über eine Dublettenkontrolle verfügen, welches alle [Instanzen](#) abdeckt.

Um die Ansichten und Überlegungen der Partner zu hören, wurden Treffen organisiert und ein Fragebogen wurde einigen von ihnen gestellt².

Wie erwartet, waren **nicht alle Partner der gleichen Meinung**. Einige beschreiben nur die Daten, welche sie auch produzieren. Sie sind überzeugt davon, dass es nicht ihre Aufgabe ist, sich um gekaufte oder erhaltene Daten zu kümmern. Andere Partner sind jedoch darauf bedacht, ihren Nutzern/Partnern ein vollständiges Inventar der Daten zur Verfügung zu stellen, die in ihrer Verwaltung vorhanden sind.

2.3 Lösungen

Das folgende Schema stellt verschiedene Lösungsansätze dar.

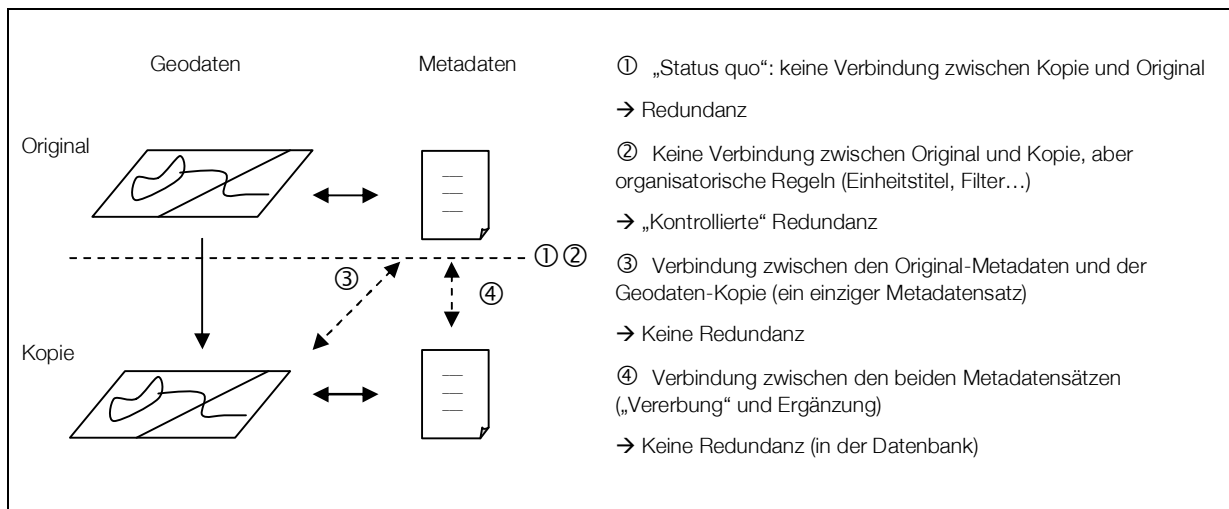


Fig. 4: Schema der Beziehungen zwischen Geodaten und Metadaten

² Die detaillierten Antworten befinden sich im Anhang S. I-XI

2.3.1 ERKLÄRUNGEN ZUM SCHEMA

2.3.1.1 ① „Status quo“

Es gibt keine Verbindung zwischen Original und Kopie. Unstrukturiertheit und Redundanzen sind die Folge.

Der Vorteil davon ist, dass **jeder so vorgehen kann wie er will**, überwiegend jedoch der Nachteil, dass es **komplizierter und unüberschaubarer** wird, je mehr Partner und Metadaten dazukommen.

2.3.1.2 ② Organisatorische Regeln

Wie bei ① gibt es keine Verbindung zwischen Original und Kopie. Jedoch werden Regeln zur Vereinheitlichung bzw. Filter definiert → „Kontrollierte“ Redundanz

- a) Es gibt Kopien des Metadatensatzes, die jedoch mit einem „standardisierten“ Titel (z.B.: Original: „SWISSIMAGE“ und die Kopien: „SWISSIMAGE (FR)“ „SWISSIMAGE (VD)“) versehen werden. So werden die **Dubletten** einigermassen **kontrolliert**.

Dies bedingt einen Export nach jedem Update der Original-Metadaten und einen Reimport in den jeweiligen Katalog mit der Anpassung der Kopie³. Die Lizenz- und Zugangsrechte müssten klar beschrieben, und der Datenherr unmissverständlich erwähnt werden.

Ein (z.B. kantonaler) Identifikator soll im Titelfeld (für A und B Partner) eingefügt werden, bei C Partnern könnte dies auch automatisch im Gateway geschehen. (Jeder C Partner kann Default-Elemente in den Gateways definieren, beispielsweise die Bounding Box des Kantons Waadt für alle Metadaten die aus dem „ASIT-VD“-Portal stammen).

Diese Lösung hat den Vorteil, **relativ einfach realisierbar** zu sein und erlaubt eine gewisse Ordnung herzustellen, während man allenfalls nach einer nachhaltigeren Lösung sucht.

- b) Eine andere Möglichkeit wäre, das Problem auf der Ebene der Rollenverteilung (Kontaktinformationen) zu lösen.

Dazu müssten klare Regeln festgesetzt werden. Man könnte sich folgendes Szenario vorstellen:

Z.B.: SWISSIMAGE wird von swisstopo beschrieben. Die Kontaktperson/-Organisation für die **Daten** (≠ Metadatenkontakt) bekommt sowohl die Rolle „Eigentümer“ und „Datenerzeuger“.

In der Kopie eines anderen Partners ist er derjenige mit der Rolle „Zuständigkeit“ oder „Anbieter“, und swisstopo ist immer noch Kontaktorganisation aber nur noch mit der Rolle „Datenerzeuger“.

Das System spuckt nach einer Suche über alle Instanzen nur jene Einträge aus, deren Kontaktperson/-Organisation beide Rollen (Eigentümer und Datenerzeuger) zugeteilt worden sind (Filter über das Feld „Rolle“). Jedoch werden ALLE Datensätze angezeigt nach einer Suche in einem oder mehreren spezifisch ausgewählten Katalogen.

Mit dieser Lösung können die Partner die Vollständigkeit ihres Katalogs bewahren, ohne den Nutzer zu „verwirren“, der eine Suche über alle Instanzen macht.

So **steht es jedem frei** was er in seinem Katalog beschreiben will, er muss jedoch **sehr rigoros** sein in der Rollenverteilung der Kontaktinformationen. Auch muss man bemerken, dass ein solcher Filter **sehr teuer** werden kann.

³ Da INTERLIS dies erlaubt, könnte eine « inkrementelle » Aktualisierung der Metadaten erwogen werden. So würden nur diejenigen Elemente ersetzt, die geändert haben, eine erneute Anpassung würde sich erübrigen.

2.3.1.3 ③ Keine Kopie der Metadaten

Es gibt eine Verbindung zwischen den Original-Metadaten und der Kopie der Geodaten (**einen einzigen Metadatensatz an einem einzigen Ort**) → Keine Redundanz

In diesem Fall ist man sehr strikt was die Datenherrschaft anbelangt. Nur der Datenherr, und nicht diejenigen die die Daten nur erworben haben, hat das Recht Metadaten zu erfassen und publizieren.

Damit die Partner dennoch die Möglichkeit haben, ihre Kunden über die bei ihnen verfügbaren Daten zu informieren, wäre eine Lösung, dass swisstopo (oder jeder andere Datenherr) in seinen Metadaten beschreibt (z.B. in der Klasse „Verteiler“), wem er seine Daten verkauft oder zur Verfügung gestellt hat, mit den jeweiligen Kontaktinfos.

Dies bedeutet zwar einen Mehraufwand für die Datenlieferanten, garantiert aber die Redundanzfreiheit. Es gibt keine Informationen zum selben Datensatz in mehreren Instanzen.

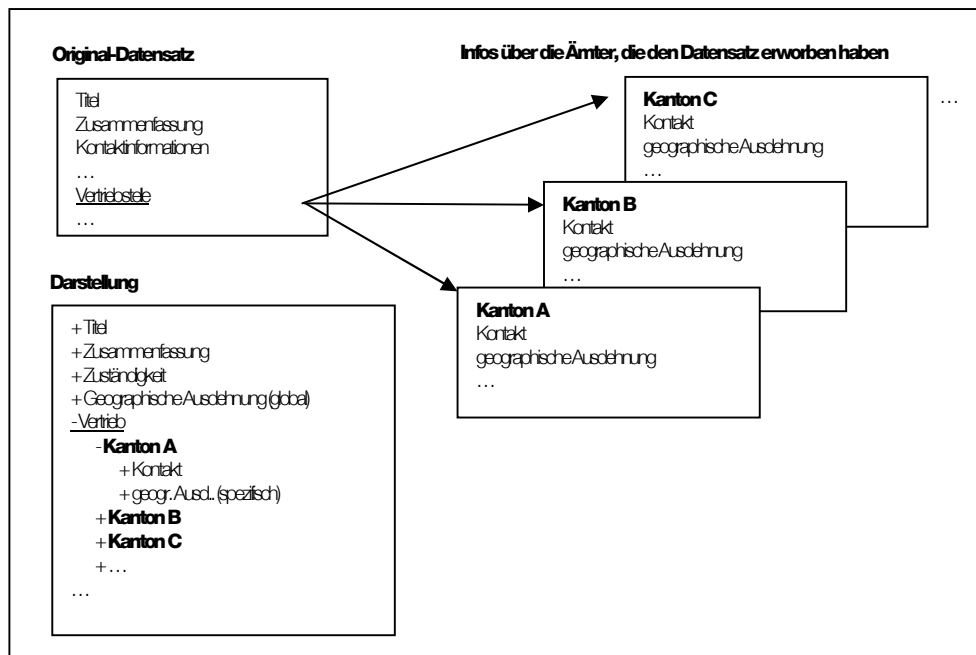


Fig. 5: Original-Metadatensatz mit „Exemplaren“

Während es vom Modell her bereits möglich ist, bei der Verteilung Kontaktinformationen anzugeben (CI_ResponsibleParty), müsste es jedoch angepasst werden, um z.B. dem Verteiler eine bestimmte geographische Ausdehnung zu zuordnen (EX_Extent).

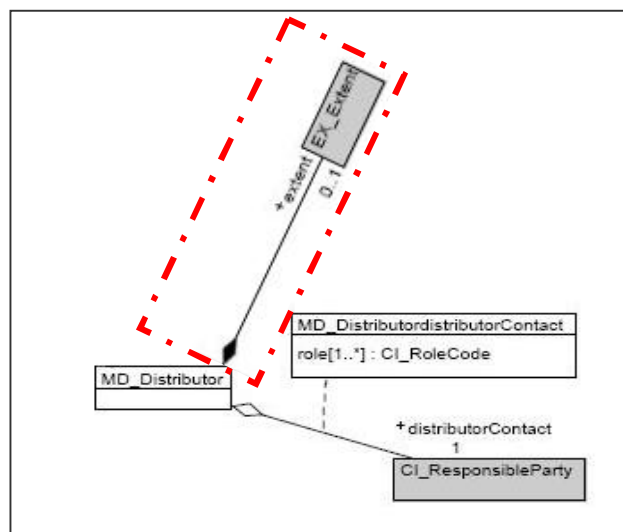


Fig. 6: Die Klasse EX_Extent wird an den Verteiler (MD_Distributor) gehängt

2.3.1.4 ④ Zwei sich ergänzende Metadatensätze

Es gibt eine Verbindung zwischen den beiden Metadatensätzen. Der „Kinddatensatz“ **erbt** Elemente der „Mutter“ und wird zusätzlich **ergänzt**. → Keine Redundanz

Diese Lösung ist ziemlich komplex und sollte weiter überdacht werden. Hier daher nur ein paar Ansätze.

Ein neuer Metadatensatz, der nur die ergänzenden Informationen enthält (über die geographische Ausdehnung, die Kontaktinfos...) bezieht sich direkt (via Unique Identifier) auf die Original-Metadaten. Die gemeinsamen Informationen befinden sich nur im Original und somit werden auch die Updates nur an einem Ort gemacht.

Die Darstellung wird übers [XSLT](#) geregelt.

Im ersten Fall (A), werden die Attribute des „Mutterdatensatzes“ (Objekt M) übernommen und zusätzliche Elemente im „Kind“ (Objekt K) angefügt.

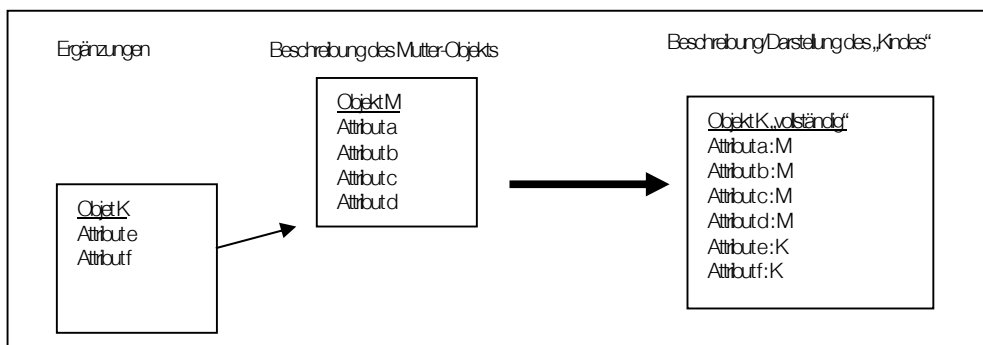


Fig. 7: Darstellung Fall A

Im zweiten Fall (B), sind Attribute im Kind anders beschrieben als im Mutter-Objekt. Die Änderungen im Kind sind stärker als die Informationen im Mutter-Objekt (Der „Titel“ im untenstehenden Beispiel).

☑ = Definiertes Element, ☒ = Nicht definiertes Element.

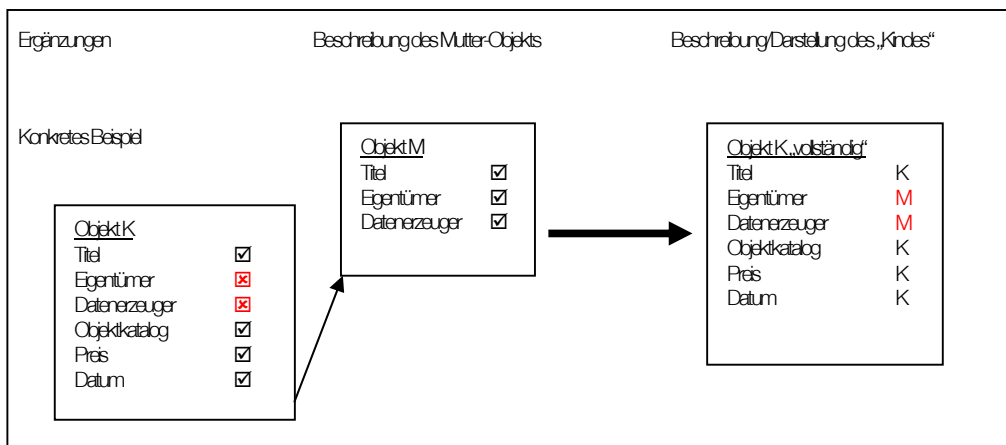


Fig. 8: Darstellung Fall B

Obschon es sich hier um eine saubere Lösung handelt, wo jeder verantwortlich ist für die Informationen die er liefert, muss man doch anmerken, dass das Redundanz-Problem nur auf Datenbank-Niveau gelöst ist. Es gilt also noch mehrere Fragen zu beantworten:

- In den Suchresultaten **erscheinen immer noch alle Datensätze**. Man muss also noch definieren in welcher Reihenfolge sie erscheinen, oder dass nur der Mutter-Datensatz angezeigt wird, und man von ihm aus zu den daran hängenden gelangt. Dafür muss die Applikation angepasst werden⁴.
- Die technische Machbarkeit abklären, da es sich um Einträge in **verschiedenen Instanzen** handelt.
- Ein **Unique Identifier** über alle Instanzen hinweg muss möglich sein.
- Alle Partner müssen den **parametrisierten URL-Aufruf** für komplette Metadatenätze gewährleisten (dies ist für Partner, die die geocat.ch-Erfassungssapplikation benutzen bereits der Fall).

Aus all diesen Gründen, ist dieser Lösungsweg wahrscheinlich mit grossem **technischen und finanziellen Aufwand** verbunden.

2.3.2 REKAPITULATION „KRITERIEN“

Lösung	Redundanz-Freiheit	Darstellung der Suchresultate	Langzeit-Lösung	Einfache Realisierung	Tiefe Kosten	Vereinfachung der Erfassung	Freie Wahl in der Erfassung
① Status quo	--	--	--	++	++	++	++
②a Einheitstitel	-	+	-	++	++	+	-
②b Filter	--	++	-/+	--	--	--	++
③ Keine Kopie der MD	++	++	+	-	+	--	-
④ Sich ergänzende Metadaten	++	--	-/+	--	--	--	++

⁴ In der jetzigen Applikation ist es nur möglich vom Kind-Datensatz zur Mutter zu gelangen aber nicht umgekehrt. Im GM03 ist die Beziehung jedoch beidseitig modelliert.

2.3.3 REKAPITULATION „VOR- UND NACHTEILE“

Lösung	Vorteile	Nachteile
① Status quo	<ul style="list-style-type: none"> – Man braucht nichts zu ändern – Nichts zu investieren – Jeder kann machen was er will 	<ul style="list-style-type: none"> – Keine nachhaltige Lösung, es wird immer komplizierter und unübersichtlicher je mehr Daten dazukommen
②a) Einheitstitel	<ul style="list-style-type: none"> – Einfach realisierbare Lösung – Wenig Ressourcen nötig – Bessere Übersicht der Suchresultate – Kann mit anderen Lösungen kombiniert werden 	<ul style="list-style-type: none"> – Zwischenlösung, eliminiert nicht wirklich die Redundanzen
②b) Filter	<ul style="list-style-type: none"> – Übersichtlichkeit/Klarheit in der Suche – Jeder kann es machen so machen, wie er will 	<ul style="list-style-type: none"> – Grosse technische/finanzielle Investition – Ev. Anpassung des Modells nötig (neues Filter-Feld, falls es nicht möglich ist über das Rollenfeld zu filtern)
③ Keine Kopie der MD	<ul style="list-style-type: none"> – Keine Redundanz – Klarheit für den Nutzer – Eindeutigkeit/Einmaligkeit der Information 	<ul style="list-style-type: none"> – Mehrarbeit für den Datenherrn – Datenschutzfragen – Wiederverkäufer? „Normalkunde“? → Verschiedenes Vorgehen?
④ Sich ergänzende Metadaten	<ul style="list-style-type: none"> – Keine Redundanz auf Datenniveau – Jeder ist für seine Daten verantwortlich – Diese Methode kann allenfalls bei der Hierarchisierung angewandt werden (s. Kap. 3.3.2.2) 	<ul style="list-style-type: none"> – Der Nutzer sieht noch immer alle Einträge durcheinander, falls nicht ein zusätzlicher Filter (oder so) existiert. – Grosse technische Investitionen (Unique Identifier über alle Instanzen hinweg) – Ev. Anpassung des Modells nötig

Um eine gewisse Ordnung ins System zu bringen, ist es wichtig zu einem Konsens zu finden. Wenn die Regeln einmal definiert sind, müssen sie von allen angewandt werden, alle sollten in gleicher Weise vorgehen. Ist dies nicht der Fall, verliert der geocat.ch-Nutzer leicht den Überblick, und weiss nicht an wen er sich wenden muss, um die gewünschten Daten oder Informationen zu erhalten.

3 Hierarchisierung der Datensätze

3.1 Definition

Die Hierarchisierungs-Funktion erlaubt einen Zusammenhang zwischen Datensätzen herzustellen (VECTOR25 und seine verschiedenen Layer, die Landeskarten mit ihren Blättern, die Gewässerschutzkarte mit ihren Komponenten...) und hilft so, Redundanzen zu vermeiden. Im übergeordneten Datensatz erfasst man die Attribute, die alle Daten betreffen (z.B. die geographische Ausdehnung, Kontaktinformationen...) und in den untergeordneten Datensätzen nur die speziellen Elemente (eigener Titel, Zusammenfassung, Aktualisierungsdatum...).

Im Moment - wir werden später die Gründe dafür erörtern - wird die Funktion von den Partnern nicht genutzt. Die hierarchisierten, bzw. hierarchisierbaren Geodaten werden **sehr uneinheitlich beschrieben**:

VECTOR25 zum Beispiel hat Metadaten, die das gesamte Produkt beschreiben und zudem je einen Metadatensatz pro Themen-Layer. Die Landeskarten sind pro Massstab einmal beschrieben (inklusive Informationen zu den Zusammensetzungen) und für die Pixelkarten gibt es einen einzigen Eintrag für alle Massstäbe von 1:25'000 bis 1:1 Mio.

3.2 Analyse

Weshalb also, verzichten die Ämter darauf ihre Daten wirklich zu hierarchisieren? Die Funktion ist zwar im GM03-Modell vorgesehen, jedoch sehr kompliziert gelöst in der Applikation. Sie verwirrt mehr, als dass Sie hilft die Datensätze zu vereinfachen. Im Moment ist es zudem auch nicht möglich, vom übergeordneten Datensatz auf die „Kinder“ zuzugreifen, die Beziehung existiert nur in der anderen Richtung.

Um trotzdem einen Zusammenhang zwischen mehreren Datensätzen herzustellen, wird die „Hierarchie“ z. B. in der Zusammenfassung erwähnt.

Betrachten wir noch einmal das Beispiel VECTOR25: Die Metadaten zu den Themen (Strassennetz, Eisenbahnnetz, Gewässernetz...) sind nicht mit dem allgemeinen Eintrag verhängt, jedoch werden in jedem Datensatz alle Themenlayer in der Zusammenfassung, bzw. in den Zusatzinformationen erwähnt (→ Redundanzen!). Wenn man sich nun entscheidet, einem Layer einen anderen Namen zu geben oder einen neuen hinzuzufügen, muss man dies in jedem einzelnen Eintrag ändern. Dies bedeutet eine Vermehrung von Fehlerquellen.

http://gbs02.geotaskserver.com/swisstopo/getPresentationRequestResult.do?fileIdentifier=V25_all&profile=GM03Comprehensive&locale=de&format=html → VECTOR25 „Gesamteintrag“

http://gbs02.geotaskserver.com/swisstopo/getPresentationRequestResult.do?fileIdentifier=V25_str&profile=GM03Comprehensive&locale=de&format=html → VECTOR25 „Strassennetz“

Abgesehen von der komplexen Technik, spielen auch die Darstellung und die Lesbarkeit der Daten eine Rolle. Mit der jetzigen „Lösung“ verliert man leicht den Überblick, wenn man die übergeordneten Metadaten aus den untergeordneten heraus anzeigen möchte. Es gibt keine visuelle Unterscheidung zwischen den beiden Datensätzen.

Damit die Partner von der Hierarchie-Funktion profitieren können, muss die **Erfassungsprozedur vereinfacht**, und die **Darstellung übersichtlicher** gemacht werden.

3.3 Lösungen

3.3.1 BEISPIEL DER BIBLIOTHEKSKATALOGE

In Bibliothekskatalogen stösst man oft auf **mehrbändige Werke**, beispielsweise Enzyklopädien, das Gesamtwerk eines Autors...). Eine Verbindung via „Unique Identifier“ (für den Nutzer sichtbar als Link auf den Titel des übergeordneten bzw. des untergeordneten Datensatzes) erlaubt zwischen den Katalogeinträgen hin und her zu wechseln.

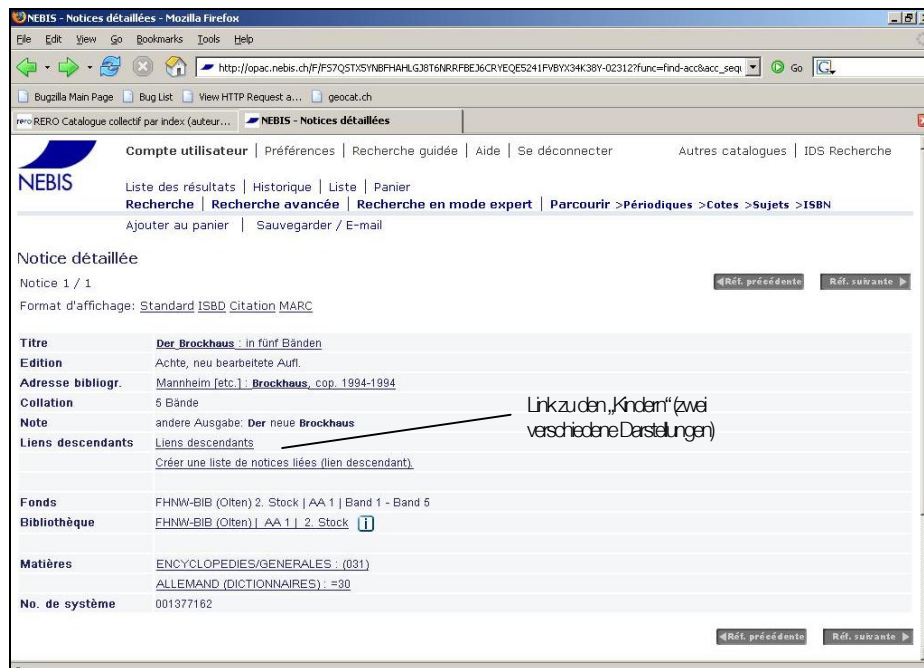


Fig. 9: Ein übergeordneter Datensatz

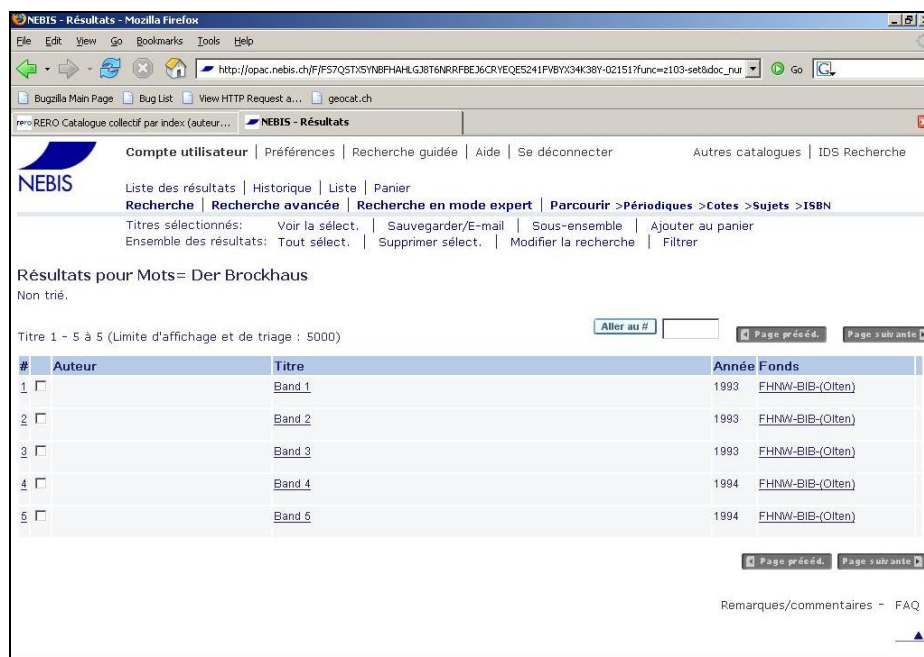


Fig. 10: Liste der „Kinder“

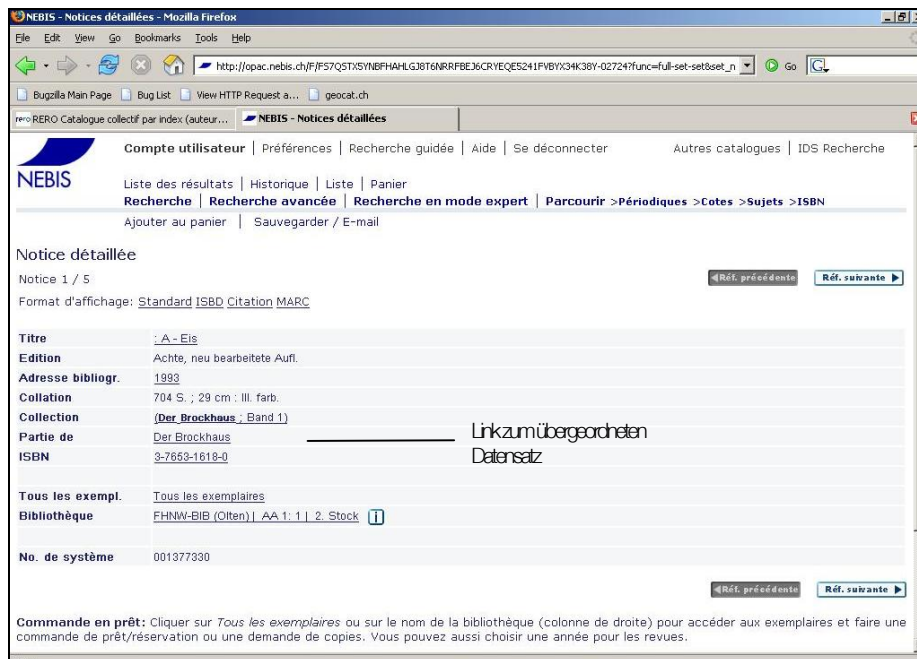


Fig. 11: Vollanzeige eines untergeordneten Datensatzes

In einem solchen System erkennt man sofort, ob ein Werk aus mehreren Bänden zusammengesetzt ist. Falls man auf den Datensatz eines Einzelbandes trifft, erkennt man sogleich den Titel des Gesamtwerks. Wenn man weitere Details will, klickt man auf den Link, der zur Vollanzeige in einer neuen Ansicht führt.

3.3.2 ANPASSUNG DER BIBLIOTHEKS-METHODE AN GEOCAT.CH

Angenommen, man lässt sich von den Bibliothekskatalogen inspirieren für eine Lösung in geocat.ch, muss man dennoch beachten, dass Geo-Metadaten einige Informationen mehr enthalten als eine einfache bibliographische Notiz. Es gibt also einige Herausforderungen zu bewältigen.

3.3.2.1 Erfassungshilfe

Man muss die Erfassung vereinfachen und den Anwender leiten. Wenn ein Kind-Datensatz erzeugt wird, möchte man wissen welche Felder im Mutter-Datensatz bereits erfasst sind. Sobald also der „Unique Identifier“ (der „Mutter“) ausgefüllt wird (z.B. aus einer Auswahlliste), werden die Felder automatisch in einer anderen Farbe übernommen. So braucht man nur noch die Felder für das „Kind“ zu korrigieren und ergänzen.

3.3.2.2 Darstellung

XSLT

geocat.ch bietet die Möglichkeit, die Visualisierung der Hierarchiestufen in Stylesheets zu definieren. Die [XSLT](#) können den verschiedenen Datentypen und Sprachen angepasst werden und damit Zweideutigkeiten vermindern. Jeder Partner soll die Möglichkeit haben, das zur Verfügung gestellte XSLT anzupassen. Es dürfte also möglich sein das Problem der Darstellung zu lösen, ohne die GM03-Modellstruktur anzupassen. Da wir uns in diesem Fall innerhalb *einer* Instanz befinden, gibt es auch kein Problem mit der eindeutigen Identifikation (Unique identifier, siehe auch [Kapitel 2.3.1.4](#)).

Mutter-Kind

Das Darstellen der über- bzw. untergeordneten Datensätze in verschiedenen Ansichten oder Fenstern (inspiriert am Beispiel der Bibliothekskataloge) ist unseres Erachtens eine gute Lösung. Wenn der gesamte Mutter-Datensatz im Kind eingeschachtelt ist (wie es zur Zeit der Fall ist), ist es sehr schwierig sich zu orientieren; man weiss schnell nicht mehr in welchem Datensatz man sich befindet.

4 Beziehung geocat.ch / envirocat

4.1 Definition

geocat.ch ist ein Portal welches den Zugriff zu komplett autonomen Katalogen ermöglicht. Die Verantwortung für die Aktualität, die Exaktheit und die Pflege der Daten liegt bei den Partnern.

Während die Partner A und B zumindest in GM03Core (mit oder ohne Erweiterungen) erfassen, steht es den C-Partnern frei, ein eigenes Metadatensystem zu wählen, auch sind sie nicht verpflichtet GM03 zu brauchen⁵. Wir haben es also mit verschiedensten Strukturen und Arbeitsweisen zu tun.

Das typische Beispiel (und im Moment das einzige) für die Beziehung zwischen geocat.ch und anderen thematischen Katalogen ist [envirocat](#). Hierbei handelt es sich um eine zentral(istisch)e Datenbank, worin verschiedene Ämter, Kantone, Städte und andere Partner Umweltmetadaten erfassen. Seit 2005 ist envirocat C-Partner von geocat.ch. Das heisst, Umweltmetadaten mit Raumbezug, die in der Datenbank des [BAFU](#) erfasst worden sind, sind via geocat.ch auffindbar.

Nun entsteht folgendes Problem: Man trifft auf Datensätze (z.B. „Cadastre du bruit routier“) die vom einen Partner in envirocat, vom andern jedoch in geocat.ch erfasst worden sind. Es gibt sogar Partner, die sich nicht entscheiden wollen und die Daten in beiden Katalogen eintragen. Dies erschwert, ganz abgesehen vom zweimaligen Erscheinen in den geocat.ch-Suchresultaten, die Pflege und die Kohärenz der Metadaten.

Darüber hinaus muss man unterstreichen, dass envirocat auf einem anderen Datenmodell als GM03 basiert. Dies verunmöglicht einen richtigen Datenaustausch und die Kompatibilität der beiden Kataloge. So sind nun die Metadaten zum „cadastre du bruit des stands de tir“ vom Kanton Genf (envirocat) nicht vergleichbar mit denjenigen des Kantons Fribourg (geocat.ch), weder auf Präzisions- noch auf Strukturniveau.

4.2 Analyse

Da die Schnittstelle zwischen den Katalogen (noch?) nicht in beiden Richtungen besteht (man findet nur die envirocat-Daten in geocat.ch, aber nicht umgekehrt) und nicht über alle Attribute funktioniert, ist die Homogenität der Suchresultate nicht gewährleistet.

Welches sind „Geoinformationen“ mit Umweltthematik und welches sind georeferenzierte „Umweltdaten“? Diese Frage ist zwar der Grund der Problematik, kann jedoch nur subjektiv beantwortet werden. Man muss also nach einer Lösung suchen, die diesen Konflikt gar nicht erst aufkommen lässt. Man sollte die Partner frei entscheiden lassen, wo sie welche Daten erfassen wollen, ohne damit Redundanzen zu schaffen.

4.3 Lösungen

4.3.1 GLEICHES MODELL, GLEICHE APPLIKATION

Die beste Lösung wäre sicher eine Fusion der beiden Kataloge. Es gäbe dann nur eine Datenbank und daher nur ein Datenmodell. Dies bedeutet am Anfang zwar einen grossen Aufwand beider Seiten, würde aber längerfristig helfen, einiges an Ressourcen zu sparen und Redundanzen vermeiden.

4.3.2 GLEICHES MODELL, ANDERE APPLIKATION

Wenn sich die erste Lösung als unmöglich erweist, aus welchen Gründen auch immer, könnten sich die beiden Parteien zumindest auf das gleiche Datenmodell einigen. Dies würde erlauben, die Qualität der Daten zu vergleichen und die Nachhaltigkeit garantieren. Wird diese Lösung gewählt, sollte darauf geachtet werden, dass der Datenaustausch ohne Schwierigkeiten möglich ist, und eine Suche in geocat.ch Daten aus envirocat liefert und umgekehrt.

⁵ Man muss nur ein Mapping einzelner Elemente machen und den Gateway konfigurieren, um die Kommunikation der beiden Kataloge herzustellen.

4.3.3 ANDERES MODELL, ANDERE APPLIKATION (IST-ZUSTAND)

4.3.3.1 Vollständiges Mapping und Import-Export in envirocat

Um die aktuelle Situation zu verbessern (ohne Modell- bzw. Applikationswechsel), müssten die beiden Kataloge durchwegs kompatibel sein. So sollte envirocat das Mapping seines Modells mit GM03 zumindest über **alle obligatorischen Elemente** der Norm, fertig stellen. Der Export/Import dieser Elemente müsste möglich sein, um einen Datensatz von der einen zur anderen Applikation zu übertragen. So wäre zumindest auch eine gewisse Uniformität gewährleistet für diejenigen Partner, die nicht darauf verzichten wollen, ihre Metadaten in beiden Applikationen zu erfassen⁶.

4.3.3.2 Vorschlag von St. Gallen

Zwei Ämter der St. Galler Administration haben ein Schema erarbeitet zur Verbesserung der Situation. Der Vorschlag lautet, die Metadaten zu „Produkten“ (a) in envirocat zu erfassen, während die dazugehörigen Daten (b) in geocat.ch zu finden sind (z.B. die „Eignungskarte für Erdwärmesonden“ gehört in envirocat und alle GIS-Layer, die diese Karte bilden, sind in geocat.ch erfasst). Um die GIS-Daten auch über envirocat zu finden, wurde ein neues Feld definiert um direkt via URL auf die Datensätze in geocat.sg zugreifen zu können.

Das Schema (siehe Fig. 12) wurde in einer Sitzung mit St. Gallen, dem BAFU und KOGIS diskutiert.

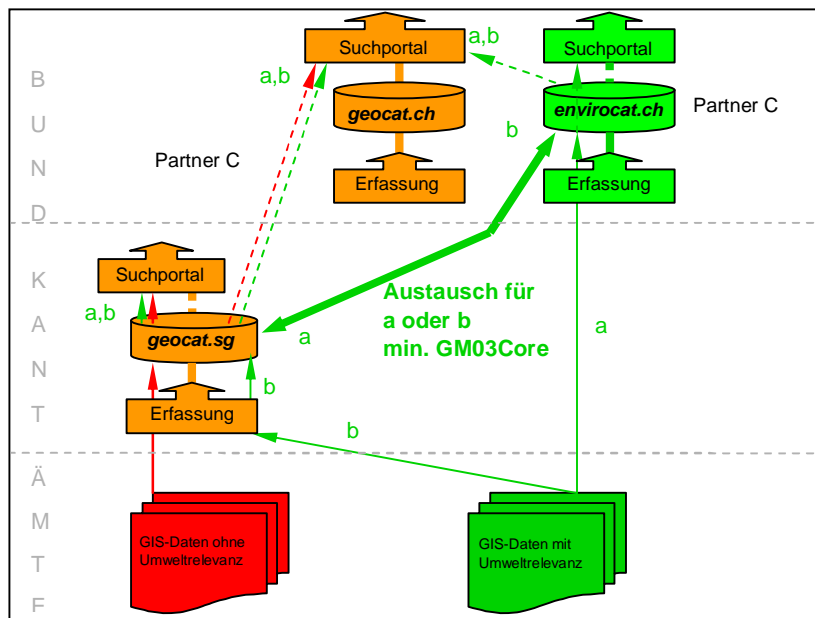


Fig. 12: Schema der Verbindungen geocat / envirocat für St. Gallen

Offene Fragen:

- Da es (noch) keine Schnittstelle zwischen geocat.sg und envirocat (hier fettgedruckt) gibt, funktioniert dieser Austausch nicht.
- Gibt es eine klare Definition für den Unterschied zwischen einem „Produkt“ und „Daten“?
→ Kann man diese Unterscheidung auf alle Geoinformationen mit Umweltrelevanz anwenden?
- Die „Produkte“ sind nicht GM03-konform (im Vergleich zu den Daten)
- Wie könnte man diese Lösung auf andere Kataloge (u.a. C-Partner) anwenden?

⁶ Bei einer Recherche in geocat.ch (und in envirocat, falls die Schnittstelle in beide Richtungen realisiert wird) sollten diese „Dubletten“ jedoch herausgefiltert werden.

4.3.3.3 Rekapitulation

Bezeichnung	Redundanz-Freiheit	Benutzerhilfe, klare Darstellung der Suchresultate	Langzeidlösung	Einfache Realisierung	Tiefe Kosten	Freie Wahl in der Erfassung
4.3.1 Gleiche Applikation	++	++	++	+/-	--/++ kurz- /langfristig	erübrigt sich
4.3.2 Gleiches Modell	+	+	+	+/-	--	-/+
4.3.3.1 Mapping/Export-Import	-/+	+	-/+	+	-	++
4.3.3.2 St.Galler Lösung	++	+	-/+	-/+ (Schnittstellen- Installation)	--	- strikte Erfassungsregeln

Es wäre wünschenswert, eine Arbeitsgruppe (mit verschiedenen Parteien) zu bilden, um allenfalls auf den St. Galler Erfahrungen aufzubauen und weitere Lösungswege zu suchen.

5 Weitere Diskussionspunkte

Während der Erstellung dieses Berichts sind wir uns weiterer Probleme bewusst geworden, die jedoch nur im weitesten Sinn etwas mit Redundanz zu tun haben. Dennoch wollten wir sie nicht vollständig ausklammern. Hier also ein kurzer Überblick.

Kohärenz

- Angenommen man autorisiert die Partner Kopien zu erstellen. Werden Regeln erstellt oder **duldet man folgende Situation** (am Beispiel der **Pixelkarten**)?
 - o swisstopo hat beschlossen alle Massstäbe (25, 50, 100, 200, 500, 1000) in einem Metadatensatz zusammenzufassen. Zwei Kantone haben sich entschieden die Daten nochmals zu beschreiben, jedoch je einen Metadatensatz pro Massstab. Es wurde auch ein anderer Titel gewählt.
 - o **Alte Versionen** wurden von einem Kanton erfasst, mit dem Link zu swisstopo. Diese sind dort jedoch so nicht mehr verfügbar. Wie löst man dieses Problem? Einfach den Link entfernen, oder braucht es eine weitere Erklärung?
- Werden in Zukunft **C-Partner** akzeptiert, die **nicht** mindestens das **GM03core** respektieren? (Wie es zur Zeit mit envirocat der Fall ist)
- Ist es die einzige Aufgabe von geocat.ch den Zugang zu einzelnen Katalogen zu gewähren (C-Partner wären bevorzugt, die Verantwortung liegt einzig bei den Teilnehmern) oder sollten auch **inhaltliche Kontrollen** durchgeführt werden? Sollte man **verbindliche Regeln** aufstellen um eine gewisse Homogenität/Kohärenz zu garantieren?

Datenherrschaft

- Bis jetzt haben wir nur von Redundanzen zwischen den geocat.ch-Partnern gesprochen. Allerdings existieren auch **Daten von externen Produzenten**. So findet man in geocat.ch z.B. zahlreiche Datensätze von „GeoPost“, (mit den jeweiligen geographischen Ausdehnungen) jedoch keinen der das Produkt als ganzes beschreibt.
- Manchmal ist es **nicht klar, wer die Datenherrschaft besitzt** oder sie wird von mehreren Ämtern/Organisationen geteilt. Beispielsweise die Bundesinventare. Im Moment kümmert sich das BFS/GEOSTAT um den Vertrieb, aber die Datenherrschaft liegt beim BAFU⁷, allerdings sind es aber die Kantone die die Daten liefern.

⁷ Auf der BFS-Website sind alle Daten beschrieben die GEOSTAT vertreibt:

http://www.bfs.admin.ch/bfs/portal/de/index/dienstleistungen/servicestelle_geostat/datenbeschreibung.html

6 Zum Schluss

geocat.ch hilft, die Geodaten zu verwalten und zu beschreiben und auch wiederzufinden. Wie wir in diesem Bericht gesehen haben, liegt die Ursache der meisten Redundanzprobleme in der dezentralen Verwaltungsstruktur und in den unterschiedlichen Ansichten der einzelnen Partner. Ein dezentrales System hat natürlich seine Vorteile und wir Schweizer hängen sehr daran. Allerdings darf man in einem solchen System nicht vergessen miteinander zu kommunizieren, um nicht völlig in entgegen gesetzte Richtungen zu rennen. Der Dialog sollte sowohl zwischen KOGIS und den Partnern bestehen, als auch bei den Partnern untereinander.

Kurze Rückschau

Zur Lösung des ersten Redundanzproblems, **(Wieder-)Verkaufs und Zugehörigkeit**, wurden vier Vorschläge präsentiert. In einer ersten Phase könnte man also die Titel der verschiedenen Metadatenätze harmonisieren. Wenn sich alle ein bisschen anstrengen, kann ein grosser Schritt in Richtung Homogenität gemacht werden. Weiter, unter der Bedingung, dass die Bundesämter einverstanden wären in ihren Datensätzen die verschiedenen Kunden zu erwähnen, wäre die [3. Lösung](#) durchaus denkbar.

Wichtig ist jedoch, dass die Kantone die erwähnten Punkte diskutieren und zu einem Konsens finden. Eine gemeinsame Strategie könnte dann durch Repräsentanten der [KKGEO](#) und der Bundesverwaltung erarbeitet werden.

Was die **Hierarchisierung** anbelangt, heisst es die Applikation zu verbessern, das XSLT und wenn nötig das Modell anzupassen. Dabei könnte man sich an den Bibliothekskatalogen orientieren. Zudem würden **klare Direktiven bezüglich Erfassung und Pflege** der Metadaten, einerseits neuen Partnern den Einstieg erleichtern und andererseits den Fortgeschritteneren in Zweifelsfällen helfen.

Die Diskussion zwischen **geocat.ch und envirocat** sollte wieder aufgenommen werden um zu einer Lösung zu finden. Es wäre nicht nur für KOGIS die beste Lösung sich auf ein einziges Datenmodell zu einigen, welches die Dauerhaftigkeit der Informationen und internationale Kompatibilität garantiert. Für die Nutzer wäre es ebenfalls von Vorteil, mit nur einer Metadatenapplikation zu arbeiten und sich nicht die Frage stellen zu müssen, welche Daten nun in welchen Katalog gehören.

Pflege und Unterhalt

Was sicher nicht vergessen werden darf ist, dass ein Katalog dynamisch ist und in ständigem Wandel. Die Daten verändern sich, oder man entdeckt (Flüchtigkeits-)Fehler; somit braucht es kontinuierliche Anpassungen. Einen Metadatenkatalog zu pflegen erfordert einen nicht zu unterschätzenden zeitlichen und finanziellen Aufwand. Anstatt Datensätze mit einer grossen Anzahl Attributen zu beschreiben, sollte man sich vorerst auf kleine Datensätze beschränken (dafür vielleicht eine grössere Menge?). Je mehr Attribute ein Datensatz hat, desto mehr Zeit braucht es für die Pflege. Es ist also besser, die Daten im Nachhinein zu ergänzen, wenn man sicher ist, dass man die nötigen Ressourcen zur Verfügung hat.

Schliesslich wurde auch seitens der Partner der Wunsch geäussert **besser über die Entwicklungen von geocat.ch informiert** zu sein. KOGIS kann sich daher vorstellen, einen regelmässigen Newsletter zu lancieren.

7 Anhang

7.1 Glossar

BAFU: Bundesamt für Umwelt. <http://www.bafu.admin.ch>

Datenherr: Bezeichnung für ein Amt oder eine Instanz welche die Autorität über bestimmte Daten besitzt. Der Datenherr ist verantwortlich für die Pflege und die Aktualität der Daten.

e-geo.ch: Das Impulsprogramm e-geo.ch im Rahmen der nationalen Geodaten-Infrastruktur NGDI ist eine Initiative von KOGIS. Das Ziel von e-geo.ch ist, den Zugang zu Geodaten zu garantieren und deren Austausch zu fördern. Weiter steht der Aufbau von Netzwerken zwischen den verschiedenen Datenproduzenten und – Nutzern im Vordergrund. Das Steuerungsorgan setzt sich zusammen aus dem Bund (KOGIS), den Kantonen (KKGEO), den Gemeinden und Städten und der Schweizerischen Organisation für Geo-Information (SOGI). http://www.e-geo.ch/home_d.htm

envirocat: Der Schweizerische Umweltdatenkatalog. Entwickelt vom BAFU in Zusammenarbeit mit UNEP/GRID. C-Partner von geocat.ch. <http://www.envirocat.ch>

Geobasisdatenkatalog des Bundesrechts: Dieser Katalog beinhaltet sämtliche Geodaten von nationaler Bedeutung, die eine Gesetzesgrundlage haben. Die Geobasisdaten, sowie die Ausarbeitung und der Unterhalt des Katalogs sind im neuen Geoinformationsgesetz verankert.

geocat.ch: Die schweizerische Applikation zur Erfassung, Suche und Lokalisierung von Geo(meta)daten. Sie basiert auf der Norm GM03. http://www.geocat.ch/home_d.htm

GM03: Das Schweizer Metadatenmodell GM03 ist eine Schweizer Norm geworden: SN 612050. GM03 ist ein Profil der internationalen Metadatennorm ISO19115. Mehrere Erweiterungen wurden gemacht, um die verschiedenen Anforderungen der schweizerischen Partner zu erfüllen. Es gibt zwei Profile: Das Kernmodell (GM03core) und das erweiterte Modell (GM03Comprehensive). http://www.geocat.ch/GM03_d.htm

Instanz: Der Katalog eines Partners zugänglich über geocat.ch. Er kann vom [Typ](#) A, B oder C sein.

INTERLIS: INTERLIS ist ein Beschreibungs- und Transfermechanismus, der, speziell um die Bedürfnisse von Geodaten-Modellierung und –Integration, bei den aktuellen und zukünftigen Geoinformations-Systemen, ganz zu erfüllen, entwickelt wurde. INTERLIS ist software- und systemunabhängig und bestimmt ein standardisiertes Transfer-Format (ITF für die Version 1 und XML für die Version 2). Seit 2003 ist INTERLIS 2 Schweizer Norm (SN 612031). <http://www.interlis.ch>

KKGEO: Die Konferenz der Kantonalen Geodaten-Koordinationsstellen und GIS-Fachstellen vereinigt die kantonalen Geodaten-Koordinationsstellen und GIS-Fachstellen mit dem Ziel, eine gemeinsame Interessenvertretung der Kantone zu gewährleisten. Die gemeinsame Behandlung fach- und verwaltungstechnischer Fragen, die sich im Bereich Geoinformation in einem Kanton stellen, steht dabei im Vordergrund. <http://www.kkgeo.ch/>

KOGIS: Das interdepartementale Koordinationsorgan für Geoinformation und geografische Informationssysteme des Bundes. <http://www.kogis.ch>

Metadaten: „Daten über Daten“. (Geo-)Metadaten beschreiben Geodaten durch die Angaben von Autor, Quelle, Zusammenfassung, Themen, Datenzugang... Durch Metadaten erhält man auch wesentliche Informationen zur Datenqualität.

NGDI: „Bei einer „Nationalen Geodaten-Infrastruktur“ handelt es sich um ein allgemein verfügbares System von Verfahren, institutionellen Einrichtungen, Technologien, Daten und Personen, die den gemeinsamen Austausch und die effiziente Nutzung geografischer Daten ermöglichen. Den Rahmen dazu bildet in der Schweiz das Kontaktnetzwerk e-geo.ch, über das alle Leistungsfaktoren koordiniert und gesteuert werden.“ (aus: http://www.e-geo.ch/NGDI_d.htm)

Partner-Typen: A = Direkte Erfassung in der geocat.ch-Metadatenbank, B = Erfassung in einem externen System, Datentransfer durch Export-Import im XML(INTERLIS)-Format, C = Erfassung in einem externen Katalog der durch einen Gateway mit der Suchapplikation von geocat.ch kommuniziert.

swisstopo: Bundesamt für Landestopografie. <http://www.swisstopo.ch/>

UML: „Die Unified Modeling Language ist eine von der Object Management Group (OMG) entwickelte und standardisierte Sprache für die Modellierung von Software und anderen Systemen. Im Sinne einer Sprache definiert die UML dabei Bezeichner für die meisten Begriffe, die für die Modellierung wichtig sind, und legt mögliche Beziehungen zwischen diesen Begriffen fest. Die UML definiert weiter grafische Notationen für diese Begriffe und für Modelle von statischen Strukturen und von dynamischen Abläufen, die man mit diesen Begriffen formulieren kann.“ <http://www.uml.org> (aus: <http://de.wikipedia.org/wiki/UML>)

XSLT: (eXtended Stylesheet Language Transformations). XSLT ist eine funktionale Programmiersprache zur Transformation von XML-Dokumenten. Es baut auf der logischen Baumstruktur eines XML-Dokumentes auf und erlaubt die Definition von Umwandlungsregeln. XSLT-Programme, so genannte XSLT-Stylesheets, sind dabei ebenfalls nach den Regeln des XML-Standards aufgebaut (es handelt sich also genau genommen um XML-Dokumente). Mit XSLT kann man:

- XML-Dateien in HTML (und andere Formate) verwandeln und Ihnen verschiedene Layouts geben,
- automatische Extraktionen machen um z.B. eine Zusammenfassung oder ein Inhaltsverzeichnis zu generieren,
- den Inhalt eines Dokuments in ein „handlicheres“ Format konvertieren.

(u.a. aus: <http://de.wikipedia.org/wiki/XSLT>)

7.2 Bibliografie

KOGIS (2005). GM03 – Metadatenmodell : ein Schweizer Metadatenmodell für Geodaten : Version 2.3. Wabern.

[Online] http://www.geocat.ch/docu/GM03/GM03_V23_de.pdf (letzter Zugriff 25.09.06)

KOGIS (2005). Metadata Management : Applikation zur Erzeugung und Verwaltung von Metadaten zu geografischen Daten : Version 1.5. Basel, GeoTask.

[Online] http://www.geocat.ch/docu/application/manual_geocat_MM_1.5_de.pdf (letzter Zugriff 25.09.06)

DORFSCHMID, Joseph et BRAWER, Sascha (2005). Modellieren raumbezogener Daten : eine Einführung unter Berücksichtigung von UML und INTERLIS. Wabern, COSIG.

[Online] http://www.interlis.ch/interlis2/download_d.php (letzter Zugriff 25.09.06)

FINK, Karin [et al.] (2004). „envirocat : a Swiss catalogue for sharing environmental information“. In: [GRID-Europe-Article], p. 482-491.

[Online] http://www.grid.unep.ch/product/publication/download/article_envirocat.pdf (letzter Zugriff 25.09.06)

7.3 Illustrationsverzeichnis

Fig. 1:	Die verschiedenen Metadaten-Modelle.....	6
	Aus: GM03-Metadatenmodell und einer KOGIS-Präsentation	
Fig. 2:	Organisation von geocat.ch	7
Fig. 3:	Beispiel redundanter Metadaten von « SWISSIMAGE »	9
	Adaptierter Screenshot aus geocat.ch	
Fig. 4:	Schema der Beziehungen zwischen Geodaten und Metadaten.....	10
	Nach einer Idee von A. Buogo	
Fig. 5:	Original-Metadatenatz mit „Exemplaren“.....	12
Fig. 6:	Die Klasse EX_Extent wird an den Verteiler (MD_Distributor) gehängt	12
Fig. 7:	Darstellung Fall A	13
Fig. 8:	Darstellung Fall B	13
Fig. 9:	Ein übergeordneter Datensatz.....	18
Fig. 10:	Liste der „Kinder“.....	18
Fig. 11:	Vollanzeige eines untergeordneten Datensatzes.....	19
	Die Screenshots (9-11) stammen aus dem Verbundskatalog NEBIS http://www.nebis.ch	
Fig. 12:	Schema der Verbindungen geocat / envirocat für St. Gallen	22
	Aus: "4er Arbeitsgruppe": Abgrenzung envirocat/geocat.ch. 1. Sitzung am 15. Juni 2004 in St. Gallen	

7.4 Antwoorden der befragten Partner

7.4.1 KANTON FRIBOURG (M. SYLVESTRE, TREFFEN VOM 20.06.2006)

1. Quelles sont les données que vous décrivez dans votre base de métadonnées (geocat ou application personnelle)? Celles dont vous êtes propriétaire ? Producteur ? Utilisateur ?

Toutes les données géographiques (données dont la référence spatiale est enregistrée sous forme de coordonnées, de nom de lieu, d'adresse postale ou de tout autre critère) doivent à terme figurer dans le géorépertoire. Il s'agit donc aussi bien de données vecteurs que de données sous forme de raster, tables, papier...→ Propriétaire (produits achetés), producteur et utilisateur

2. Qui saisit les MD dans votre service ?

Le Centre de compétences de SIT (CCSIT)

Mandat donné à une entreprise extérieure pour compléter

Depuis juillet 2006 3 autres services du canton gèrent leurs propres métadonnées: (1 pers. dans chacune des 3 : Forêts, Protection de la nature, Aménagement)

3. Qui les publie ?

Le CCSIT. Jusqu'à présent, tous ceux qui saisissent publient aussi, car tout le monde utilise le même compte d'utilisateur ! → Vont le changer quand ils savent comment le faire...

Rem. : Autrefois il n'y avait qu'un seul compte utilisateur car le CCSIT était le seul gestionnaire des métadonnées pour le canton de Fribourg.

Depuis juillet 2006 trois autres comptes ont été créés pour les trois services qui seront en charge de gérer leurs propres métadonnées. Ils ne peuvent par contre pas publier les métadonnées, seul le CCSIT possède ce droit.

4. Comment gérez-vous la mise à jour des métadonnées ?

Un « profil fribourg » sous forme de questionnaire est mis à disposition des producteurs afin qu'ils le remplissent. La mise à jour des métadonnées repose donc sur l'implication des différents services à transmettre les informations au CCSIT.

Pour les données externes Mise à jour des MD à la réception de la nouvelle version des géodonnées concernées

Pas de contrôle sur les 3 services qui saisissent eux-mêmes

5. Consultez-vous geocat.ch avant de saisir un nouveau jeu de métadonnées ?

Pas toujours, quand il y a une mise à jour des données (ex. Vector200) → copie de la notice, puis adaptation à Fribourg

En outre, on part du principe qu'il s'agit des données fribourgeoises, non disponibles ailleurs

6. Quel(s) est/sont votre/vos argument(s) pour décrire un jeu de données dont vous n'êtes pas producteur ?

Le but est d'avoir une base de données exhaustive des données existantes Sont donc décrites les données mises à disposition des utilisateurs par le biais de serveurs de fichiers ou SDE et celles stockées en internes dans les différents services du canton de Fribourg

Afin de faciliter la consultation des métadonnées disponibles au sein du Canton de Fribourg la liste des données est consultable sur le site internet de SYSIF. Chaque donnée renvoie, via un URL, vers sa description dans geocat.

Rem. : Pourquoi alors ne pas faire le lien direct sur les métadonnées du producteur (ex. swisstopo pour les cartes VECTOR) ? → **car dans l'outil de recherche de geocat, on ne connaîtrait pas les données disponibles à Fribourg**

7. Comment voyez-vous le problème en tant qu'utilisateur de geocat.ch à la recherche d'une information pertinente ?

Conscients du problème, mais leur priorité sont les utilisateurs de Fribourg qui doivent avoir connaissance de ce qui est à leur disposition

8. Avez-vous eu des réactions (positives ou négatives) concernant geocat de la part de vos clients ?

Très bonne initiative essentielle pour la qualité des géodonnées, mais :

Difficulté de faire intéresser les gens aux métadonnées, → cours/information en début juillet

L'outil ne fonctionne souvent pas, est très lent dans le chargement des données

En général très peu de réactions

9. Lequel des scénarios « extrêmes » préféreriez-vous : Pourquoi ?

- a) On considère qu'une donnée « transformée » ou un extrait d'un jeu de données est une « nouvelle » donnée qui doit avoir ses propres métadonnées. Il ne s'agit donc pas de redondances, et c'est à l'application ou au modèle de « mettre de l'ordre » dans les résultats de recherche pour que l'utilisateur sache que ces données ont la même origine mais des propriétés différentes.
- b) Aucune redondance n'est permise. Chaque jeu de géodonnées n'a qu'un seul jeu de métadonnées, décrit par le producteur, même s'il est découpé, transformé, acheté par des tiers. L'application doit contenir un outil de contrôle de doublons, couvrant tous les catalogues et instances. On dispose alors d'un outil de recherche performant et cohérent.

a) car correspond à la politique fribourgeoise (outil de saisie et recherche exhaustive)

10. Utilisez-vous un thésaurus ou une liste « contrôlée » pour les mots-clés ?

« Les mots-clés sont rentrés au fur et à mesure qu'ils apparaissent, lorsque que l'on en a besoin... »

**Problématique des objets réutilisables : On n'a pas moyen de trouver le jeu de métadonnées qui utilise ces objets
→ quasiment impossible de faire des corrections (mots-clés, contacts...)**

11. Propositions pour améliorer la situation, comment imagineriez-vous un scénario idéal ?

Pas de proposition concrète, mais :

- **Possibilité de pouvoir afficher toutes les métadonnées d'une instance (sans devoir entrer un autre critère de recherche)**
- **Le problème des redondances n'est pas tellement dérangeant**

12. Autres remarques

Serait bien de pouvoir accéder aux MD d'un seul gestionnaire (rem : ne savaient pas que ceci est déjà possible par le biais de comptes spécifiques)

Etre informé sur le suivi du traitement des bugs, problèmes etc.

Questions supplémentaires (e-mail du 13/07/2006) :

Est-ce que vous utilisez la fonction de la hiérarchisation dans geocat?

La classe "Identifiant du parent" n'est pas intégrée dans notre modèle fribourgeois. Nous n'en voyons pas l'utilité pour la saisie de nos métadonnées. En effet, il est rare qu'une métadonnée se rapporte à une autre métadonnée de rang supérieur. Pour donner un exemple concret, nous ne décrivons pas le Vector200 Topic par Topic, il ne nous est donc d'aucune utilité de faire référence à une donnée de rang supérieur puisque celle-ci n'existe pas. Cependant je suis tout à fait disposée à intégrer cette classe dans notre profil si cette information est nécessaire.

Comment gérez-vous la question envirocat-geocat.ch? Pourquoi le "cadastre du bruit routier" ne se trouve que dans envirocat, tandis que le "cadastre de bruit des stands de tir" se trouve dans les deux, mais avec deux titres différents?

Aucune solution n'existe pour le moment. Ces deux catalogues sont gérés par deux services différents : geocat est de notre responsabilité (SIT) et envirocat est du ressort du Service de l'environnement. Notre inventaire pour les géométadonnées s'est basé, entre autre, sur les données existantes dans envirocat mais :

- **un grand nombre de données présentes dans envirocat ne sont pas des données géographiques, celles-ci n'ont pas été reprises dans geocat**
- **un grand nombre de données présentes dans geocat ne sont pas des données environnementales, celles-ci ne figureront pas dans envirocat**
- **les modèles de description des métadonnées dans geocat et envirocat sont différents**
- **les mises à jour dans ces deux répertoires ne sont actuellement pas coordonnées puisque réalisées par deux services différents.**

Pour résumer, ces deux répertoires sont actuellement gérés de façon indépendante, nous souhaiterions pouvoir trouver une solution afin d'améliorer la synergie entre ces deux outils. Cette thématique nous préoccupe et nous sommes ouverts à des propositions et serions intéressés de savoir comment les autres cantons gèrent cette problématique.

Fribourg en tant que canton bilingue. Est-ce que vous avez prévu de décrire vos métadonnées dans les deux langues? Quelle est votre "politique" à ce sujet-là?

La description des métadonnées dans les deux langues n'est pas à l'ordre du jour. La majorité des collaborateurs de l'Etat de Fribourg étant francophone, la description des métadonnées en langue allemande n'est pas une priorité mais serait effectivement souhaitable à terme.

7.4.2 KANTON NEUENBURG (F. VOISARD, ANTWORT PER E-MAIL)

1. Quelles sont les données que vous décrivez dans votre base de métadonnées ?

Nous décrivons essentiellement les données produites et gérées par le canton. Nous ne décrivons pas de données fédérales.

2. Qui saisit les MD dans votre service ?

La saisie se fait par toute l'équipe du SITN. Une personne s'occupe de la gestion et coordination du domaine (F. Voisard). Nous avons formé une personne de la ville de La Chaux-de-fonds pour faire la saisie des données communales. Pour les autres communes, les ressources et compétences nécessaires sont difficiles à trouver.

3. Qui les publie ?

Chaque personne faisant la saisie peut le faire mais c'est le gestionnaire qui s'en occupe.

4. Comment gérez-vous la mise à jour des MD ?

Actuellement la mise à jour des métadonnées se fait quand il y a des modifications majeures d'une donnée. Il est prévu de faire un contrôle annuel.

5. Consultez-vous geocat.ch avant de saisir un nouveau jeu de métadonnées pour vérifier s'il n'est pas déjà disponible ?

Oui

6. Utilisez-vous la fonction de hiérarchisation des données ?

Non

7. Quelle(s) est/sont votre/vos motivation(s) pour décrire un jeu de données déjà décrit par quelqu'un d'autre ?

Nous essayons de ne pas le faire.

8. Quel(s) est/sont votre/vos argument(s) pour décrire un jeu de données sur laquelle vous ne détenez pas l'autorité ?

Si les personnes en charge des données n'ont pas les ressources ou compétences pour le faire (petites communes).

9. Comment gérez vous les données environnementales (envirocat vs geocat) ?

Nous recommandons l'utilisation de Geocat pour toutes les données géographiques environnementales. Pour les autres documents dans le domaine : mesures, rapports, ... que conseille la confédération ?

10. Comment voyez-vous le problème en tant qu'utilisateur de geocat.ch à la recherche d'une information pertinente ?

Il faut simplifier au maximum l'application (Google – Geocat) : + facile à utiliser + performante + stable. Recherche classifiée selon la pertinence (cf. Google).

11. Avez-vous eu des réactions (positives ou négatives) concernant geocat de la part de vos clients ?

Nous n'avons quasiment pas eu de réactions.

12. Lequel des scénarios « extrêmes » préféreriez-vous : Pourquoi ?

a) On considère qu'une donnée « transformée » ou un extrait d'un jeu de données est une « nouvelle » donnée qui doit avoir ses propres MD. Il ne s'agit donc pas de redondances, et c'est à l'application ou au modèle de « mettre de l'ordre » dans les résultats de recherche pour que l'utilisateur sache que ces données ont la même origine mais des propriétés différentes.

b) Aucune redondance n'est permise. Chaque jeu de géodonnées n'a qu'un seul jeu de métadonnées, décrit par le *producteur*, même s'il est découpé, transformé, acheté par des tiers. L'application doit contenir un outil de contrôle de doublons, couvrant tous les catalogues et instances. On dispose alors d'un outil de recherche performant et cohérent.

Favorables au scénario compromis : minimiser les redondances. Une documentation redondante ne doit se justifier que sur la base de critères de transformation importants : ex. simple découpage de cartes topographiques : non, interpolation de données ponctuelles et passage en mode image : oui.

13. Utilisez-vous un thésaurus ou une liste « contrôlée » pour les mots-clés ?

Non, pas pour le moment.

14. Propositions pour améliorer la situation

**Edicter des propositions générales pour éviter au maximum les redondances
Un modérateur doit analyser les redondances**

15. Autres remarques, commentaires

Où en sont les réflexions de migration de l'application actuelle sur la plate-forme améliorée pour l'Alsace ?

7.4.3 KANTON ST. GALLEN (C. BRENK KVA, S. HEGGLI AFU, RENCONTRE DU 18.07.2006)

1. Welche Daten beschreiben Sie in Ihrer Metadatenbank? Ihre eigenen? Diejenigen die Sie besitzen? Auch solche die Sie nur benutzen?

Eigene Daten (Geobasisdaten, versch abgeleitete Produkte, Orthophotos...)

Fremddaten (andere Ämter des Kantons)

2. Wer erfasst in Ihrem Betrieb die Metadaten? Wer publiziert sie?

Christian Brenk, noch keine definitive Regelung der Zuständigkeiten und Organisation

3. Wie wird die Aktualisierung der Metadaten gehandhabt? Nur falls sich die Geodaten ändern? Regelmässige, systematische Kontrollen und updates? Falls ja, wie sehen die Kontrollen aus?

Noch keine definitive Regelung

4. Konsultieren sie geocat.ch bevor Sie Metadaten erfassen?

Bislang nicht, es wurden zunächst KVA-relevante Daten erfasst

5. Wie sehen Sie das Problem der Redundanzen aus Sicht eines Benutzers auf der Suche nach einer treffenden Information?

Wenn ein Datensatz sehr häufig bei verschiedenen Stellen aufgeführt ist, kann dies verwirren, wenn man den eigentlichen Eigentümer sucht. Ansonsten kann es von Vorteil sein, Geodaten der swisstopo bei seinem Heimatkanton zu bestellen.

6. Gab es Reaktionen (positive oder negative) von Seiten Ihrer Kunden, Partnern... bezüglich geocat.ch?

Bislang fast gar keine, nur hie und da Störungsmeldungen

Vorschlag: Nutzungsstatistiken veröffentlichen, auch präzise: Wie weit gehen die Suchanfragen

7. Welches der folgenden „Extremszenarien“ bevorzugen Sie? Weshalb?

- a) Man bezeichnet veränderte Daten, Datenauszüge als neuen Datensatz der seine eigenen Metadaten erhält. Es handelt sich also nicht um Redundanzen. Es ist die Aufgabe der Applikation bzw. des Modells die Resultate/Daten so darzustellen, dass man erkennt welches das Original und welches die „Derivate“ sind.

Es sollte immer erkennbar sein, um welchen Datensatz es sich handelt. Grundsätzlich sollte jeder Datensatz einen Mindestsatz an Metadaten enthalten, um ihn identifizieren zu können. Über die Angabe der Herkunft ist dies auch bei Ausschnitten möglich.

Es muss aber auch unterschieden werden zwischen öffentlich relevanten Datensätzen und temporären Arbeitskopien o.ä. Es ist sicher nicht sinnvoll, jeden kleinen Teildatensatz im geocat zu dokumentieren. Dies ist primär für den jeweiligen Benutzer interessant.

- b) Redundanzen sind absolut nicht erlaubt. Jeder Geodatensatz hat nur einen Metadatensatz der vom Datenherrn erfasst wird, auch wenn von Dritten nur Ausschnitte unter teilweise anderen Bedingungen verwendet werden. Die Applikation muss katalogübergreifend nach Dubletten kontrolliert werden. So wird ein kohärentes und leistungsstarkes Suchwerkzeug gewährleistet.

Die Frage ist, welchen Datenbestand soll geocat.ch repräsentieren? Alle Daten der Schweiz? Oder nur übergeordnete Datensätze?

Wären evtl. Untergruppen von Metadatenbanken notwendig? Dies wäre möglich, wenn die Erfassung bei jedem Erzeuger direkt (im GIS) stattfindet. Jeder erfasst Metadaten zu jeder Bearbeitung. Welche Daten dann an ein zentrales Metadaten system weitergegeben werden, muss definiert werden.

8. Benutzen Sie einen Thesaurus oder eine kontrollierte Liste für Ihre Schlüsselwörter?

Soweit sind wir noch nicht

9. Benutzen Sie die Möglichkeit der Hierarchisierung von Datensätzen? („übergeordneter Datensatz“)

Soweit sind wir noch nicht

10. Verbesserungsvorschläge, wie stellen Sie sich zum Beispiel ein „Idealszenario“ vor?

- **Die Metadatenerfassung findet grundsätzlich auf allen Ebenen der Erzeugung und Bearbeitung von Geodaten direkt im jeweiligen GIS statt. Dabei kann das GIS unterstützend helfen. Hier kann ein sehr detailliertes Profil eingesetzt werden.**
- **Metadaten werden als untrennbarer Bestandteil eines Geodatensatzes betrachtet, nicht als zusätzlicher Aufwand.**
- **Bei der Weitergabe der Daten werden die notwendigen Metadaten mit weitergegeben. Der Attributumfang kann dabei variieren.**
- **Die Metadaten öffentlich relevanter Daten werden in einem zentralen Metadatenkatalog veröffentlicht.**

Voraussetzungen:

- **Alle eingesetzten GIS-Systeme unterstützen den Schweizer Standard und bieten Schnittstellen zu geocat**
- **Metadaten werden als integraler Bestandteil von Geodaten betrachtet**
- **Es braucht eine koordinierende Stelle, die entscheidet, welche Daten öffentlich relevant sind**
- **Es braucht evtl. abgestufte Metadatenbanken mit unterschiedlicher Ausrichtung (Erzeuger, Benutzer)**

11. Weitere Bemerkungen

- **Bis jetzt wurden nur Beispieldatensätze in geocat, bzw. envirocat eingetragen, da eine definitive Lösung des „was, wo“ noch nicht gelöst wurde. Der Vorschlag von St.Gallen, die „Karten“ (d.h. die Produkte mit den Links zu den Bestandteilen) in envirocat und die einzelnen Datensätze (GIS-Layer) in geocat einzutragen funktioniert nur teilweise, da die Karten zwar in geocat.ch erscheinen (mit Hilfe der Schnittstelle zu envirocat), auf kantonaler Ebene jedoch nicht.**
- **Um die Sache übersichtlich zu halten, ist nur eine Einzelperson für die Pflege der Umwelt-Metadaten zuständig. (Kontrolle, Aktualisierung, Erfassung...)**
- **Die „4er Arbeitsgruppe“ die 2004 gestartet wurde (BAFU, KOGIS, AfU und KVA) sollte wieder belebt werden, allenfalls (zu einem späteren Zeitpunkt) auch kantonsübergreifend erweitert werden.**
- **Eine Erfassungsempfehlung sollte zumindest für die Geobasisdaten herausgegeben werden. (Einheitstitel (jener der im Gesetzestext erscheint) möglichst in den drei offiziellen Landessprachen, Schlüsselwörter-Vorschläge, Präzisionsniveau... mit Beispielen).**
- **Vorschlag: Anstelle eines „Geobasisdaten-Ateliers“ am e-geo Forum in Liestal, könnte man einen Workshop zum Thema Metadaten einschieben, um über Grundsatzfragen zu diskutieren (welche MD für wen, Erarbeiten von Empfehlungen...)**
- **Generell wird zu wenig kommuniziert von Seiten des Bundes (News, Erweiterungen, Geodienste...). Oft wird zu sehr Wert gelegt auf die Technik, das Werkzeug, Empfehlungen werden (wenn überhaupt) erst im Nachhinein abgegeben, und die Leute die schon begonnen haben, werden dann „bestraft“ und jene die abwarten haben die besseren Karten. Dies hat zur Folge dass momentan sowohl envirocat als auch geocat der Treibstoff (d.h. die erfassten Daten) fehlt. (Bsp.: Nur 2 Kantone haben die Gewässerschutzkarte in envirocat erfasst)**
- **Redundanzenproblem auch auf Gemeindeebene: Bsp: 1 Metadatensatz für alle Zonenpläne des Kantons oder einen für jeden einzelnen Zonenplan? → tausende Zonenpläne in geocat!?!?**
- **Solange Metadaten etwas sind, das zusätzlichen Aufwand bedeutet, und nicht obligatorisch, oder vielmehr von wirtschaftlichem Interesse ist, wird es schwer sein, die Leute zur Erfassung zu bewegen. Es braucht den Druck der Kunden um die Sache zum Laufen zu bringen. Eine Qualitäts- und Fortschrittskontrolle sollte auf Bundesebene geschehen.**
- **Wünschenswert: geocat-Daten sollten via google auffindbar sein. Test: „Gewässerschutzkarte“ als google-suche → man erhält zuoberst Resultate von Kantonen die (noch) nicht in geocat erfassen... (Klar ist aber, dass google geocat nicht ersetzt!)**

7.4.4 KANTON WAADT (D. GIORGIS, ANTWORT PER E-MAIL)

1. Quelles sont les données que vous décrivez dans votre base de métadonnées?

Toutes les données utilisées, acquises (chez swisstopo ou autres sociétés privées) ou produites en interne.

2. Qui saisit les MD dans votre service ?

Une seule personne dans chaque service concerné.

3. Qui les publie ?

Une seule personne pour toute l'administration a le droit d'activer la publication.

4. Comment gérez-vous la mise à jour des métadonnées ?

Uniquement quand il y a une modification des géodonnées.

5. Consultez-vous geocat.ch avant de saisir un nouveau jeu de métadonnées ?

Non.

6. Quelle(s) est/sont votre/vos motivation(s) pour décrire un jeu de données déjà décrit par quelqu'un d'autre ?

Y inscrire nos propres conditions d'utilisation.

7. Quel(s) est/sont votre/vos argument(s) pour décrire un jeu de données dont vous n'êtes pas producteur ?

Informers nos utilisateurs.

8. Comment voyez-vous le problème en tant qu'utilisateur de geocat.ch à la recherche d'une information pertinente ?

Pas compris la question ...

9. Avez-vous eu des réactions (positives ou négatives) concernant geocat de la part de vos clients ?

Non.

10. Lequel des scénarios « extrêmes » préféreriez-vous : Pourquoi ?

a) On considère qu'une donnée « transformée » ou un extrait d'un jeu de données est une « nouvelle » donnée qui doit avoir ses propres métadonnées. Il ne s'agit donc pas de redondances, et c'est à l'application ou au modèle de « mettre de l'ordre » dans les résultats de recherche pour que l'utilisateur sache que ces données ont la même origine mais des propriétés différentes.

b) Aucune redondance n'est permise. Chaque jeu de géodonnées n'a qu'un seul jeu de métadonnées, décrit par le *producteur*, même s'il est découpé, transformé, acheté par des tiers. L'application doit contenir un outil de contrôle de doublons, couvrant tous les catalogues et instances. On dispose alors d'un outil de recherche performant et cohérent.

La solution a) est préférable, car de nombreux produits dérivés existent parfois et il est nécessaire de les décrire ou d'indiquer comment les données originales ont été "transformées" pour obtenir les produits dérivés.

11. Utilisez-vous un thésaurus ou une liste « contrôlée » pour les mots-clés ?

Pas utile

12. Propositions pour améliorer la situation, comment imagineriez-vous un scénario idéal ?

Pas de proposition

Questions supplémentaires/Précisions (e-mail du 24/07/2006)

Comment gérez-vous les redondances dans le portail ASIT-VD (pourquoi y a-t-il par ex. 2 notices pour SWISSIMAGE, 2 notices pour la CP25...)

Nous ne pouvons pas éviter ce genre de redondance car swisstopo est également membre de l'ASIT-VD. Par conséquent, chaque membre est autorisé à renseigner ses métadonnées. Nous essayons néanmoins de distinguer nos données de celles de swisstopo en insérant le label (ACV) qui signifie Administration Cantonale Vaudoise.

Pourquoi saisissez-vous 3 fois le point de contact quand il s'agit du même service qui s'occupe de diffusion, contact et gestion?

Ce sont des champs obligatoires. Par conséquent, la saisie de métadonnées ne peut pas être validée si ces champs ne sont remplis !

La personne qui publie les métadonnées, effectue-t-elle un dernier contrôle des informations? Si oui de quelle manière (ex.: vérification de doublon, orthographe...) Si non, pourquoi?

Non, nous n'effectuons aucuns contrôles. Chaque service est responsable de ses métadonnées. Un certain nombre de champs sont obligatoires et un document d'aide à la saisie avait été présenté afin d'avoir une certaine homogénéité de la saisie.

Comment gérez-vous les liens vers l'extérieur (pas de vérification? Outil de vérification automatique? Contrôle manuel? --> ex: "Cartes pixel de l'Atlas géologique de la Suisse (ACV)": le lien vers swisstopo n'aboutit pas)

Nous avons malheureusement de la peine à suivre les divers changements informatiques (changement de serveur, nouveau lieu de stockage des fichiers, etc.). Cela se fait au coup par coup lorsque nous constatons que cela ne marche pas ou lorsque que quelqu'un nous le signal.

Informez-vous vos utilisateurs sur l'existence de geocat (si oui comment: présentations, liens vers geocat depuis vos sites Internet, de bouche à oreille..., si non pourquoi?

geocat a été présenté aux différents responsables des métadonnées de nos services. Mais il se trouve que dans notre administration les collaborateurs sont rarement amenés à travailler en dehors de notre canton et par conséquent ils ont rarement besoins de consulter des données qui seraient dans geocat et pas dans le Géocatalogue de l'ASIT-VD.

Précision de ma question sur les utilisateurs: Quelqu'un a entendu parler du produit SWISSIMAGE mais ne se souvient plus qui en est le fournisseur et ou il peut commander le produit. Il fait donc une recherche sur geocat et tombe sur plusieurs notices de la même donnée. Ce serait peut-être uniquement après un coup de fil chez vous qu'il découvre que la bonne adresse est swisstopo (s'il ne vient pas du canton de Vaud). Auriez-vous une proposition de clarifier la chose pour lui? Ou n'accordez-vous pas tant d'importance à ce sujet?

Cette personne devra consulter le contenu des métadonnées pour constater que pour la métadonnée du canton de Vaud (ASIT-VD Géocatalogue Condition de diffusion ou geocat ... ?), seules les personnes membres de l'administration ou mandataire de l'administration peuvent utiliser et disposer de ces données. L'autre solution serait de convaincre swisstopo de partager la propriété de certain produit (p.ex. la copropriété canton-swisstopo pour le produit SWISSIMAGE) ce qui permettrait à n'importe quelle personne (bureau privé, mandataire de l'administration etc.) de s'adresser au canton pour disposer de ce produit. Cela pourrait aussi éventuellement éviter les redondances de geocat !

7.4.5 KANTON ZUG (C. OBERHOLZER (GIS-FACHSTELLE), A. KONRAD (GIS-FACHSTELLE), A. FURRER (AFU), TREFFEN VOM 27.06.2006)

envirocat-Diskussion:

Wenig Ressourcen für die Erfassung, begonnen 2000

GIS-Fachstelle trug die raumbezogenen Daten in geocat ein, da sie auch bei deren Erstellung dabei war. Daten wurden danach aus envirocat gelöscht

AFU: Hauptsache Datenzugang, ob es nun in Geocat oder Envirocat sei. Grösstes Problem dabei ist, dass der externe envirocat-Kunde diese Daten nicht mehr findet, da keine Schnittstelle vorhanden ist. Gibt nun das AFU Zug diese Daten noch einmal in envirocat ein, entsteht ein erneutes Redundanzen-Problem in geocat.

Metadaten sind wichtig, aber man sollte sie auch nicht überbewerten, je kompakter desto besser.

Generell erscheint der Detaillierungsgrad dem AFU zu hoch, je mehr Daten erfasst werden, desto höher ist auch der Aktualisierungs- und Überarbeitungsaufwand.

Es gibt zwar externe Datenanfragen auf Grund von envirocat, jedoch sind die telefonischen Kontakte doch noch viel häufiger. Daher sind das à jour halten der Kontaktinfos das wichtigste.

Betonung der Kurzfristigkeit der Informatikanwendung, AFU glaubt kaum, dass in 5 Jahren noch jemand diese Applikation benutzen wird. Datensicherung sei auch nicht garantiert.

Problem auch der „Freiwilligkeit“. Es gibt anscheinend noch keine gesetzliche Grundlage bezüglich Umweltdaten-Dokumentation.

Lösungsvorschläge: Ideal: Die Schnittstelle muss auf beide Seiten funktionieren

Eventuell Vorschlag von St. Gallen was aber entweder mit einem grösseren Aufwand seitens GIS und/oder AFU verbunden ist

Am einfachsten: Der Link zu geocat.ch sollte auf „allen“ envirocat-Seiten sichtbar sein, eventuell mit dem Vermerk „für raumbezogene Daten, siehe auch geocat.ch“

Ratlosigkeit was das weitere Vorgehen beim AFU angeht. Irgendwelche Daten (beispielsweise Elektromog) beschreiben und dann den Link in envirocat zur Karte in geocat machen, damit wenigstens ein Anhaltspunkt in envirocat aufzufinden ist? Abwarten bis die Schnittstelle existiert?

geocat.ch

GIS-Fachstelle koordiniert alle Rauminfos im Kanton Zug

1) Welche Daten beschreiben Sie in Ihrer Metadatenbank? Ihre eigenen? Diejenigen die Sie besitzen? Auch solche die Sie nur benutzen?

Nur Daten die den Kanton Zug betreffen, deren Datenherr die GIS-Fachstelle ist. Der Stand ist aktuell, ziemlich alles erfasst

2)3) Wer erfasst in Ihrem Betrieb die Metadaten? Alle Mitarbeiter? Eine Einzelperson? Der Datenproduzent? Ein Team? / Wer publiziert sie?

Eine Einzelperson erfasst die Daten und publiziert sie auch

4) Wie wird die Aktualisierung der Metadaten gehandhabt? Nur falls sich die Geodaten ändern? Regelmässige, systematische Kontrollen und updates? Falls ja, wie sehen die Kontrollen aus?

Nachführung wenn's neue Datensätze gibt

5-8) Konsultieren sie geocat.ch bevor Sie Metadaten erfassen? Aus welchem Grund, bzw. welchen Gründen erfassen Sie Metadaten zu einem Geodatensatz der bereits in geocat.ch beschrieben ist? Nennen Sie ein (paar) Argument(e) weshalb Sie Geodaten eines anderen „Datenherrn“ (Datenproduzenten) beschreiben. Wie sehen Sie das Problem der Redundanzen aus Sicht eines Benutzers auf der Suche nach einer treffenden Information?

Dieses Problem gibt's nicht im Kanton Zug, (siehe punkt 1)

9) Welches der folgenden „Extremszenarien“ bevorzugen Sie? Weshalb?

- a) Man bezeichnet veränderte Daten, Datenauszüge als neuen Datensatz der seine eigenen Metadaten erhält. Es handelt sich also nicht um Redundanzen. Es ist die Aufgabe der Applikation bzw. des Modells die Resultate/Daten so darzustellen, dass man erkennt welches das Original und welches die „Derivate“ sind.
- b) Redundanzen sind absolut nicht erlaubt. Jeder Geodatensatz hat nur einen Metadatensatz der vom Datenherrn erfasst wird, auch wenn von Dritten nur Ausschnitte unter teilweise anderen Bedingungen verwendet werden. Die Applikation muss katalogübergreifend nach Dubletten kontrolliert werden. So wird ein kohärentes und leistungsstarkes Suchwerkzeug gewährleistet.

Version b) entspricht der Handhabung im Kanton

10) Benutzen Sie einen Thesaurus oder eine kontrollierte Liste für Ihre Schlüsselwörter?

Schlüsselwörter nach Gutdünken, keine Kontrolle

11) Verbesserungsvorschläge, wie stellen Sie sich zum Beispiel ein „Idealszenario“ vor?

Alle sollten so arbeiten (wie Szenario b)

12) Weitere Bemerkungen

Anfragen die auf geocat zurückzuführen sind, sind eher selten,

Gefahr dass die Leute „zu viele“ Daten sehen und dann mal alles bestellen, ohne vorher zu überlegen, was sie wirklich brauchen → Mehraufwand

Zug hat den Vorteil ein kleiner Kanton mit einer relativ einfachen Verwaltungsstruktur zu sein.

Gut gibt es geocat aber: Mehr Kantone müssten ihre Daten eintragen, und man sollte es als „Mittel zum Zweck“ sehen, „Aufwand muss mit Ertrag übereinstimmen“. Die Hauptsache ist, dass der Benutzer weiss, an wen er sich wenden soll (Meine Bemerkung: Zuerst muss er die Info finden!), denn schlussendlich braucht es sowieso den persönlichen Kontakt, um die Modalitäten zu besprechen.

7.4.6 SWISSTOPO : (WERNER BALMER, RAPHAËL BOVIER, ANTWORT PER E-MAIL)

1. Quelles sont les données que vous décrivez dans votre base de métadonnées ?
Les produits de l'office fédéral de topographie dont il est le producteur et propriétaire.
2. Qui saisit les MD dans votre service ?
Cela dépend des produits. Chaque domaine décrit ses produits et une unité est responsable de la gestion de geocat (support aux personnes qui saisissent, préparation de profils, etc.).
3. Qui les publie ?
Chaque domaine décrit les données puis l'unité responsable les publie. Cette unité ne fait qu'un contrôle sur la forme et pas sur le contenu.
4. Comment gérez-vous la mise à jour des MD ?
La mise à jour s'effectue lorsqu'il y a une modification des géodonnées. Un contrôle général annuel a également lieu.
5. Consultez-vous geocat.ch avant de saisir un nouveau jeu de métadonnées pour vérifier s'il n'est pas déjà disponible ?
Non, si ce n'est dans les métadonnées swisstopo.
6. Utilisez-vous la fonction de hiérarchisation des données ?
Non, car l'implémentation dans geocat ne nous a pas convaincu. Mais, la fonction nous intéresse.
7. Quelle(s) est/sont votre/vos motivation(s) pour décrire un jeu de données déjà décrit par quelqu'un d'autre ?
En tant que producteur et distributeur nous nous devons de décrire nos produits.
8. Quel(s) est/sont votre/vos argument(s) pour décrire un jeu de données sur laquelle vous ne détenez pas l'autorité ?
Nous ne les décrivons pas.
9. Comment gérez vous les données environnementales (envirocat vs geocat) ?
Pas de gestion
10. Comment voyez-vous le problème en tant qu'utilisateur de geocat.ch à la recherche d'une information pertinente ?
La multiplication des résultats positifs pour certaines recherches demande un travail de tri important. Cela nous fait penser aux premiers outils de recherche sur le web. Une certaine intelligence dans le tri des résultats serait un plus.
11. Avez-vous eu des réactions (positives ou négatives) concernant geocat de la part de vos clients ?
Aucune réaction par rapport à geocat en général.
12. Lequel des scénarios « extrêmes » préféreriez-vous : Pourquoi ?
 - c) On considère qu'une donnée « transformée » ou un extrait d'un jeu de données est une « nouvelle » donnée qui doit avoir ses propres MD. Il ne s'agit donc pas de redondances, et c'est à l'application ou au modèle de « mettre de l'ordre » dans les résultats de recherche pour que l'utilisateur sache que ces données ont la même origine mais des propriétés différentes.
 - d) Aucune redondance n'est permise. Chaque jeu de géodonnées n'a qu'un seul jeu de métadonnées, décrit par le producteur, même s'il est découpé, transformé, acheté par des tiers. L'application doit contenir un outil de contrôle de doublons, couvrant tous les catalogues et instances. On dispose alors d'un outil de recherche performant et cohérent.
**La solution a) est préférable. L'utilisateur doit par exemple pouvoir choisir où il souhaite s'adresser pour recevoir des données. Encore une fois, le tri des résultats de recherche doit être optimisés. Quelques propositions:
Les extraits d'un jeu de données devraient être clairement identifiables comme tel par rapport au jeu de données "mère". De même pour les dérivés.
Lors d'une recherche avec un filtre spatial, les résultats avec un taux de couverture à 100% devraient être mis en évidence par rapport aux autres.**
13. Utilisez-vous un thésaurus ou une liste « contrôlée » pour les mots-clés ?
Une liste "swisstopo" a été élaborée pour la première saisie, mais depuis elle n'est plus du tout contrôlée. Cette liste rassemblait simplement tous les mots-clés souhaités par les utilisateurs. Elle contenait encore des synonymes. Avec une saisie et des responsabilités décentralisées, il est difficile de mettre en place une liste contrôlée. Une telle liste voir même un thésaurus serait un plus indéniable mais demanderait un important travail entre les domaines pour la mettre au point et la gérer.

7.4.7 ENVIROCAT : (KARIN FINK, INFORMATIONS-TREFFEN VOM 15.05.06)

envirocat est la base de (méta)données centralisée pour les données, documents, produits etc. liés aux questions environnementales. Les différents partenaires y entrent directement leurs données. La coordination et l'administration se trouvent au sein de l'OFEV.

Volume : Env. 1800 jeux de données dont env. 260 « géo-métadonnées » reliées à geocat par les champs : Données géoréférencées : soit, Géométrie des données, système de référence ou Résolution spatiale des données, échelle. Pour les documents ils s'agit des attributs : Couverture géographique du contenu : unités administratives, coordonnées géographiques correspondant au système de référence suisse, description sous forme de texte

Des problèmes lors de la migration de l'ancien système CDS et beaucoup de données redondantes et/ou incomplètes n'ont pas permis une publication directe. Il y a d'abord un long travail de correction et finition de ces données à effectuer.

« Catalogage » : Manuel d'utilisateur, Workshops, visites chez les partenaires, certain contrôle par envirocat. Mots-clés et Thèmes connexes du Thésaurus de l'EUA → normalisation, mais tout de même un problème d'homogénéité, car le propriétaire des données peut choisir plus ou moins librement lesquelles et le degré de précision/généralisation

Différents types d'éléments → masque de saisie correspondant aux attributs nécessaires (Données, récolte de données, document, projet, site Internet, outil, base légale, personne, organisation)

Remarques : Pas la même approche que geocat. Tandis que pour geocat, la standardisation, l'échange, la réutilisation et surtout la durabilité des métadonnées sont la première préoccupation, envirocat met la priorité sur un outil accessible, axé sur besoins des utilisateurs, dont un souci principal est d'être clair et convivial à l'utilisation.