



Qualitativ hochwertige Metadaten pflegen und verarbeiten

- Hinweise für Metadatenerfasser und Katalogbetreiber -

erarbeitet durch die Koordinierungsstelle GDI-NI beim Landesamt für
Geoinformation und Landentwicklung Niedersachsen (LGLN)



Niedersachsen

Inhaltsverzeichnis

1. Einleitung	3
2. Allgemeine Hinweise zum Umgang mit Metadaten.....	3
2.1 Metadaten als aussagekräftige Beschreibung der Geodatenressource	3
2.2 Datenbeschreibung – Suchdienst – Such- und Präsentationsoberfläche	4
3. Das inhaltliche Erstellen von Datenbeschreibungen	5
3.1 Sprachliche Gestaltung	6
3.2 Inhaltliche Gestaltung	8
3.3 Überprüfen der angegebenen Kommunikationswege.....	9
4. Die Präsentation von Metadaten	11
4.1 Anforderungen von Datenanbieter und Suchendem.....	11
4.2 Möglichkeiten einer Präsentation von Metadaten	13
4.3 Der Datenanbieter als Ansprechpartner	14
4.4 Zielgruppengerechte Präsentation	15
4.5 Sprachneutrale Code-Wörter, Code-Listen und sonstige Konventionen	18
4.6 Verwenden semantischer Thesauri	19
5. Hinweise für Katalogbetreiber	20
5.1 Kataloge mit direkt erfassten Metadaten (Originäre Kataloge).....	20
5.2 Kataloge mit geharvesteten Metadaten (Harvesting-Kataloge).....	23
5.3 Kataloge mit Metadaten gemischter Herkunft (Gemischte Kataloge).....	25
6. Die Bereitstellung von Metadaten im Sinne von INSPIRE	26
0. Anhang – Begriffe im Zusammenhang mit Metadaten	27
CSW – Catalogue Service Web (Katalogschnittstelle).....	27
Harvesting	27
Geodatenressourcen	27
Suchende	28
Such- und Präsentationsoberfläche.....	28

1. Einleitung

Dieser Leitfaden beantwortet die Frage, wie inhaltlich qualitativ hochwertige Datenbeschreibungen (Metadaten) erstellt werden. Qualitativ hochwertig sind die Datenbeschreibungen dann, wenn sie vom potenziellen Nutzer, also dem Suchenden, gefunden und verstanden werden. Im Anhang dieses Leitfadens finden Sie die wichtigsten Begriffe im Zusammenhang mit Metadaten kurz erläutert. Da wir die Begriffe gleich von vorn herein als „bekannt“ voraussetzen, werfen Sie ruhig als erstes einen Blick in diesen Anhang, sollten Sie mit dem Thema noch nicht allzu vertraut sein.

Der Leitfaden hat als übergreifende Zielgruppe diejenigen im Blick, die sich mit Metadaten beschäftigen. Hierunter fallen die Erfasser, die z. B. inhaltlich die Texte für die Datenbeschreibungen entwerfen. Ferner enthält er Hinweise für Katalogbetreiber und die Anbieter von Such- und Präsentationsoberflächen für Metadaten.

Nur durch das perfekte Zusammenspiel dieser Gruppen können qualitativ hochwertige Metadaten dazu beitragen, Geodaten im Rahmen der Geodateninfrastruktur interoperabel bereit zu stellen.

2. Allgemeine Hinweise zum Umgang mit Metadaten

Metadaten sind Datenbeschreibungen für Geodatenressourcen. Dies beinhaltet die Beschreibung von Geodaten, Diensten wie WMS, WFS oder CSW und ggf. auch komplette Fachinformationssysteme. Ziel der Erstellung von Metadaten ist es, einem zukünftigen Nutzer so viele Informationen an die Hand zu geben, dass er nach eigenem Ermessen fachlich beurteilen kann, ob er die beschriebene Geodatenressource zur Erledigung seiner Aufgabe einsetzen kann.

2.1 Metadaten als aussagekräftige Beschreibung der Geodatenressource

Nur aussagekräftige Beschreibungen der Geodatenressourcen führen – neben den technischen Voraussetzungen, die erfüllt sein müssen – zu interoperabel nutzbaren Geodaten innerhalb der Geodateninfrastruktur.

Im Wesentlichen werden durch die inhaltlich hochwertigen Datenbeschreibungen der Zeitaufwand für die Recherche minimiert und die Einschätzung der Nutzbarkeit von Geodaten für ein Projekt gesteigert. Ob Geodaten zu einem bestimmten Thema bereits vorliegen oder selbst erstellt werden müssen, kann nun in kürzester Zeit beantwortet werden. Zu beachten ist ebenso der persönliche Vorteil für denjenigen, der die Recherche durchführt. Häufig wiederkehrende Fragen wie die nach den vorhandenen Formaten, dem Datenmodell, den Nutzungsbedingungen oder anfallenden Kosten können direkt und ohne weitere persönliche Nachfragen im Rahmen der selbstverantwortlichen Recherche beantwortet werden. Dies schont Zeit und Nerven – sowohl auf Seiten der Datenanbieter als auch auf Seiten der Suchenden.

Eine qualitativ hochwertige Beschreibung der vorhandenen Geodatenressourcen birgt dabei die folgenden Vorteile, die je nach Art und Zugänglichkeit der Geodaten mehr oder minder zutreffen:

- Geodatenressourcen werden katalogisiert gelistet und verschlagwortet
- Geodatenressourcen können ad hoc gefunden werden
- die Entstehung der Geodatenressourcen kann sowohl intern als auch extern betrachtet nachvollzogen werden
- Testdaten stehen einem potenziellen Nutzer ohne Wartezeit zur Verfügung
- Geodaten können in ihrem inhaltlichen Umfeld angeschaut werden
- alle Informationen zu den Daten sind direkt bekannt
- alle Ansprechpartner sind direkt bekannt
- eine Recherche nach Geodaten kann jederzeit, von jedem Ort und mit geringstem Aufwand erfolgen

2.2 Datenbeschreibung – Suchdienst – Such- und Präsentationsoberfläche

Metadaten werden über so genannte Suchdienste (CSW) bekannt gemacht. Diese Suchdienste können von jedermann standardisiert angefragt werden. Üblicherweise wird für die standardisierte Anfrage eine Webapplikation erstellt, die für den Nutzerkreis „einfach“ zu bedienen ist. Hinter diese Oberfläche legt man die Anfragen an den Suchdienst. Dabei kann mit der Suchoberfläche sowohl auf nur einen Suchdienst z. B. den eines einzelnen Anbieters (z. B. Kreisfreie Stadt Braunschweig) oder aber auf mehrere Suchdienste vieler Anbieter zugegriffen werden.

Suchdienste stehen stets kostenfrei und ohne Einschränkungen für den Zugriff durch jedermann zur Verfügung. Das bedeutet, dass die URL der CSW-Schnittstelle des Suchdienstes im Internet eingesehen und genutzt werden kann. Übrigens ist auch der CSW eine Geodatenressource und sollte daher mit einem aussagekräftigen Metadatensatz beschrieben sein.

Die Suchoberflächen, mit Hilfe derer die Nutzer Geodatenressourcen ausfindig machen, orientieren sich an ihrer jeweiligen Zielgruppe. So könnte es sein, dass Sie intern eine komplexe Suchoberfläche anbieten, mit denen die Mitarbeiter nach den Geodatenressourcen im eigenen Hause recherchieren können. Für die Bürgerinformation bieten Sie hingegen eine gesonderte Suchoberfläche mit eventuell reduziertem Inhalt als Bestandteil Ihres Webauftritts an. Beide Suchoberflächen greifen im Hintergrund auf den gleichen Suchdienst zu, präsentieren jedoch die Ergebnisse nach der Zielgruppe differenziert.

Oft ist es so, dass zusätzlich zu der Recherche nach geführten Geodaten auch noch ein Zugriff auf Geodatenbeschreibungen bestehen muss, die zwar nicht zu ihren eigenen zählen, die jedoch häufig von ihnen verwendet werden. Dies könnten beispielsweise Daten aus dem Bereich Geowissenschaften und Geologie sein. Ihre Suchoberfläche wird daher nicht nur auf ihren eigenen Suchdienst zugreifen, sondern zusätzlich noch den Suchdienst ansprechen, der Informationen zu den Daten aus den Geowissenschaften und der Geologie enthält.

Ein Suchportal, das den „Bürger in Deutschland“ informieren soll, wird auf „alle“ Metadaten in Deutschland zugreifen. Das bedeutet, dass über eine gemeinsame Oberfläche Geodatenbeschreibungen aus den

unterschiedlichsten Regionen Deutschlands aber auch aus den unterschiedlichsten Themenbereichen präsentiert werden.

Die einmal von Ihnen erstellte Beschreibung muss sich damit nicht nur gegen weitere von Ihnen erstellte Metadaten „behaupten“, sondern sie steht fortan in „Konkurrenz“ mit Datenbeschreibungen von Bundesbehörden oder Gemeinden aus z. B. Niederbayern. Jemand, der nach Geodaten fahndet, wird daher bei der Eingabe des Begriffs „Straße“ nicht nur den Datensatz mit Ihren beispielsweise Gemeindewegen finden, sondern eben auch die Straßen eines Landesamtes, der Bundesbehörde und einer Gemeinde aus z. B. Niederbayern. - Dies müssen Sie bedenken, wenn Sie Ihre eigenen Geodatenressourcen beschreiben.

Ein Suchender, der nicht diejenigen Daten finden kann, die er benötigt, wird von einer weiteren Nutzung der Suchoberfläche absehen. Er unterscheidet nicht zwischen Präsentationsoberfläche und Datenbeschreibung. Für ihn ist beides „eins“. Aus diesem Grunde kann auch eine Präsentationsoberfläche nur so gut sein, wie die Daten desjenigen Katalogs, auf den die Suchapplikation gefiltert oder ungefiltert zugreift.

3. Das inhaltliche Erstellen von Datenbeschreibungen

Unsere Empfehlung für das inhaltliche Erstellen von aussagekräftigen Datenbeschreibungen zielt sowohl auf die sprachliche und inhaltliche Gestaltung der Metadaten ab als auch auf die Handhabung der zu erwartenden Reaktionen auf diesen Inhalt durch potenzielle Datennutzer. Im Idealfall führt bereits der Prozess der Entstehung von Metadaten zu einer Verbesserung des Kommunikationsflusses und der Arbeitsabläufe in Ihrem Hause.

Bei all Ihren Bemühungen gilt stets der Grundsatz: Beschreiben Sie ausschließlich **eigene** Geodatenressourcen!

Als Hersteller von Geodaten sind Sie der Fachmann für Ihre eigenen Geodaten. Sie haben es daher auch weitestgehend selbst in der Hand, wie oft am Tag bei Ihnen nach den lieferbaren Formaten, dem Datenmodell oder dem Projektionssystem, in dem die Daten vorliegen, nachgefragt wird. Verständliche Metadaten reduzieren diesen Aufwand für Nachfragen und Auskünfte auf ein Minimum, denn Ihre Kunden werden sich nun nur noch dann bei Ihnen melden, wenn sie etwas „wirklich“ nicht verstehen. Gleichzeitig kann Ihnen jede persönliche Nachfrage, die an Sie gerichtet wird, einen Hinweis darauf geben, dass Ihre Datenbeschreibungen möglicherweise weiter verbessert werden können. Aus diesem Grund ist eine Datenbeschreibung niemals abgeschlossen, sondern unterliegt im Idealfall einem kontinuierlichen Review-Prozess.

3.1 Sprachliche Gestaltung

Im Folgenden geben wir Ihnen einige Tipps zur sprachlichen Gestaltung von Metadaten, die hoffentlich dazu führen, dass Sie den oder die Suchende von vorn herein auf „die richtige Fährte“ führen.

Schreiben Sie Fließtext

In der Geodatenbeschreibung darf ausschließlich reiner Text verwendet werden. Hierzu gehören natürlich auch Zahlen. Mit Hilfe von html gestaltete Überschriften oder Tabellen werden wenn überhaupt, dann nur in Ihrem eigenen System korrekt wieder gegeben. Von daher dürfen Elemente von Auszeichnungssprachen nicht in Ihrer Beschreibung verwendet werden. Erwartet wird reiner Fließtext ohne jegliche Formatierungen.

Schreiben Sie orthographisch korrekt

Ihre Beschreibung der Geodaten ist das Aushängeschild für Ihr Haus und Ihre Daten. Finden sich hier schon banale Fehler, wird dies ein schlechtes Licht auch auf die angebotenen Geodaten werfen. Daher ist das Verwenden einer durchgängig richtigen Rechtschreibung, Grammatik und Interpunktion in der Datenbeschreibung unerlässlich.

Vermeiden Sie Abkürzungen

Eine Geodatenressource mit dem Titel „UKN_WA_NDS“ mag für den Fachmann im Hause durchaus eingängig und effizient formuliert sein. Alle übrigen jedoch, werden zwangsläufig wenig damit anfangen können und somit eine „Übersichtskarte Waldflächen in Niedersachsen“ als Titel vorziehen.

Sobald Sie in Ihrem eigenen Hause über Geodaten sprechen, werden bestimmte „Fachbegriffe“ fallen. Unter Ihren Mitarbeitern besteht Klarheit über die Definition dieser Fachbegriffe. Sobald Sie „BoGIS“ sagen, weiß „jeder“, dass es sich dabei um Ihr Geoinformationssystem Boden handelt, indem Sie auch festhalten, welche Altlastenflächen in Ihrer Stadt bereits untersucht wurden. Für den Metadatensatz müssen Sie jedoch dies alles ganz genau beschreiben. Stellen Sie sich daher immer wieder die Frage: Kann das jemand verstehen, der noch nie vorher mit uns in Kontakt getreten ist?

Verwenden Sie ausgeschriebene Begriffe

Denken Sie daran, dass eine Abkürzung bei Ihnen auf die eine Art und Weise, bei einer anderen Institution jedoch in einem vollkommen anderen Zusammenhang verwendet werden kann. Verwenden Sie nur ausgeschriebene Begriffe und setzen Sie bei Ihnen gängige Abkürzungen in Klammern dahinter; z. B.: „Geoinformationssystem Boden (BoGIS)“ oder „Geoinformationssystem Böden und Altlasten (BoGIS)“. Besonders häufig werden Abkürzungen auch für Organisationseinheiten in einem Hause verwendet. Schreiben Sie die Organisationseinheiten, die Sie benennen möchten, dennoch immer aus und nennen Sie die Kurzbezeichnung allenfalls in Klammern hinter dem vollständigen Begriff.

Verwenden Sie gängige Begriffe

Geodaten beschreiben zum Teil hoch komplexe Sachverhalte, die nur mit entsprechendem Fachwissen fachlich korrekt und vollständig beschrieben werden können. So können beispielsweise Probebohrungen durchgeführt werden, deren Ergebnisse „sondiert“ werden. Eine andere Institution führt ähnliche Probebohrungen durch. Dort nennt man den gesamten Vorgang jedoch schlichtweg „Beprobung“, in einem anderen Falle spricht man von „Analyse“ oder „Testergebnis“.

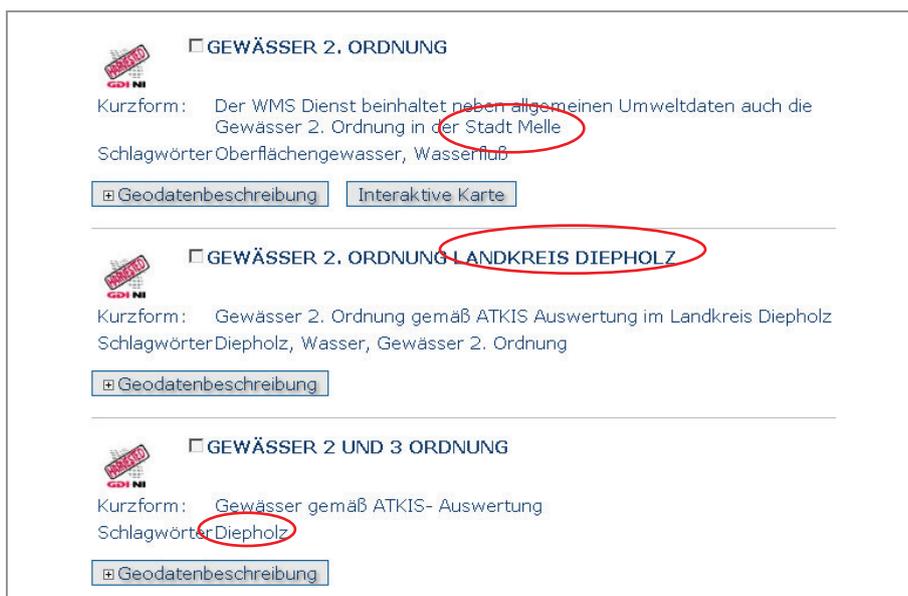
Achten Sie darauf, in Ihrer Beschreibung solche Begriffe zu verwenden, die „üblicherweise“ verwendet werden. Wenn Sie sich nicht sicher sind, benutzen Sie in Frage kommende Begriffe parallel. Denken Sie stets daran, dass Ihre Beschreibung von einem Menschen verstanden werden muss, der weder aus Ihrem Hause stammt, noch die gleiche Ausbildung hat, wie Ihre Mitarbeiter.

Folgen Sie einer Systematik

Ihre Metadaten sollten sich immer an der Handhabung der Daten in Ihrem Hause orientieren. Wenn Sie also verschiedene Geodatenressourcen beschreiben, die in einem thematischen oder technischen Zusammenhang stehen, so ergibt sich dieser thematische oder technische Zusammenhang bereits im ähnlich aufgebauten Titel für die jeweils zusammengehörigen Metadatensätze: „Übersichtskarte Wald im Landkreis“, „Übersichtskarte Boden im Landkreis“, „Übersichtskarte Gewässer im Landkreis“

Machen Sie Ihre Datenbeschreibung „spezifisch“

Überlegen Sie, ob man Ihre Geodatenbeschreibung „auf Anhieb“ von einer Geodatenbeschreibung einer anderen Behörde oder der Geodatenbeschreibung eines Nachbarkreises unterscheiden kann. Gerade bei Verwaltungen mit einem begrenzten räumlichen Wirkungskreis ist es wichtig, im Titel der Geodatenressource auf diese räumliche Begrenzung hinzuweisen. Angemessen ist auch ein zusätzlicher deutlicher Hinweis im ersten Satz der Beschreibung der Geodatenressource. Ebenso können Vorschaubilder dazu beitragen, Metadaten verständlicher zu gestalten.



GEWÄSSER 2. ORDNUNG
Kurzform: Der WMS Dienst beinhaltet neben allgemeinen Umweltdaten auch die Gewässer 2. Ordnung in der Stadt Melle
Schlagwörter Oberflächengewasser, Wasserfuß
[Geodatenbeschreibung](#) [Interaktive Karte](#)

GEWÄSSER 2. ORDNUNG LANDKREIS DIEPHOLZ
Kurzform: Gewässer 2. Ordnung gemäß ATKIS Auswertung im Landkreis Diepholz
Schlagwörter Diepholz, Wasser, Gewässer 2. Ordnung
[Geodatenbeschreibung](#)

GEWÄSSER 2 UND 3 ORDNUNG
Kurzform: Gewässer gemäß ATKIS- Auswertung
Schlagwörter Diepholz
[Geodatenbeschreibung](#)



Abb.: Wie schnell finden Sie heraus, wo es Daten gibt?



Abb.: Vorschaubilder für das direkte Verständnis

3.2 Inhaltliche Gestaltung

Überlegen Sie bei der inhaltlichen Gestaltung Ihrer Metadaten, was für den Suchenden eine hilfreiche Information ist. Nutzen Sie die Möglichkeit, auf Testdatensätze zu verlinken oder aber setzen Sie über die Geodatenbeschreibung einen Link auf ein Fachinformationssystem, in dem die Geodaten eingebunden sind. Der Suchende ist so in der Lage, die beschriebenen Daten in ihrem originären fachlichen Bezugsrahmen kennen zu lernen. Gehen Sie davon aus, dass es Suchende mit unterschiedlichem Anspruch gibt. Suchende können sowohl professionelle GIS-Nutzer als auch beispielsweise Schüler unterschiedlichen Alters sein. Die Bandbreite der Suchenden ist sehr weit zu sehen. Jeder Gruppe von Suchenden wird über ein unterschiedliches Suchvokabular verfügen, dem Sie im Idealfall durch Ihre Geodatenbeschreibung gerecht werden müssen.

Oft sind Hyperlinks in einem Metadatensatz eine der besten Möglichkeiten, die Geodatenressource gegenüber einem potenziellen Nutzer vollständig zu erläutern. Es bleibt dem Suchenden überlassen, ob er sich eine solche zusätzliche Information anschaut, indem er einen gesetzten Link durch Anklicken nutzt.

Links sind immer dann ratsam, wenn es sich um Informationen handelt, die für das pure Auffinden dieses Metadatensatzes eine untergeordnete oder keine Rolle spielen. Hierzu zählen damit z. B. Links auf Nutzungsbedingungen, die wahrscheinlich ohnehin für 99% aller Daten aus Ihrem Hause identisch sind. Oft bietet sich auch ein Link auf das verwendete Datenmodell an, welches z. B. Vektordaten und die verwendeten Attribute beschreibt. Hilfreich sind in jedem Falle auch Links auf Legenden, sofern es sich bei der Ressource um Geodaten handelt, die eine Legende benötigen, wie dies z. B. bei Rasterkarten der Fall ist.

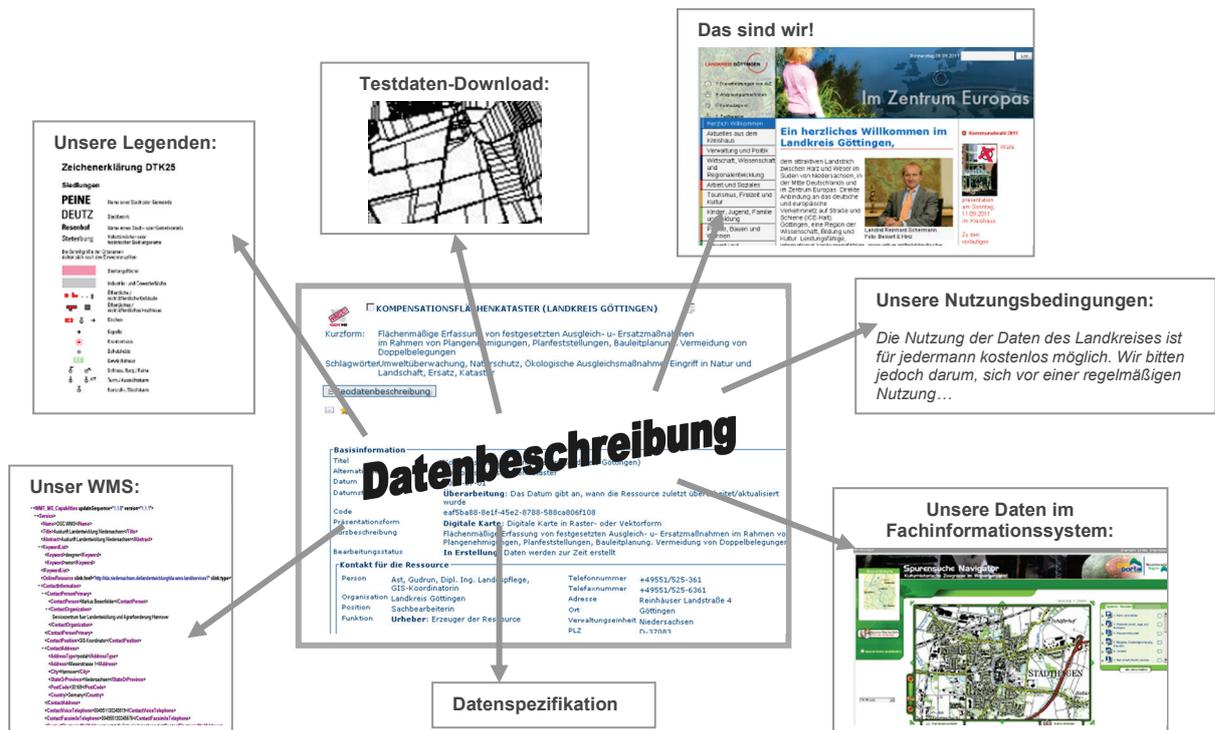


Abb.: Metadaten als Ausgangspunkt zur umfassenden Beschreibung einer Geodatenressource

Beim inhaltlichen Gestalten eines Metadatensatzes geht es darum, die Balance zu wahren. Einerseits müssen Sie so viele Informationen zu der Geodatenressource anbieten, dass sie ausreichen, um die Geodatenressource möglichst schnell zu finden und genau zu beschreiben. Andererseits sollten es tatsächlich nur so viele Informationen sein, dass ein Nutzer sich „ein Bild“ von der Geodatenressource machen kann. Sofern ihn die Geodatenressource interessiert, wird er gerne bereit sein, weitere Informationen direkt z. B. in Ihrem Internetauftritt abzurufen, sofern Sie ihn mit den entsprechenden Links versorgen.

Wenn Sie die Möglichkeit haben, kontrollieren Sie Ihre Metadatensätze nicht nur in Ihrer eigenen Suchoberfläche, sondern vor allem in fremden Suchoberflächen. Ist die Darstellung auch dort so, wie Sie sich das wünschen? Sind Sie in der Lage binnen kürzester Zeit Ihren eigenen Metadatensatz aus einer Vielzahl von Datensätzen heraus zu suchen?

3.3 Überprüfen der angegebenen Kommunikationswege

Neben einer umfassenden sprachlich-inhaltlichen Darstellung der Geodatenressource besteht ein hochwertiger Metadatensatz durch weitere Eigenschaften, die jedoch nur einen kleinen inhaltlichen Teil von ihm ausmachen und der in der Regel deswegen auch „schnell“ erledigt sind. Dahinter stehen umfangreiche und vorausschauende Überlegungen in Bezug auf die zu planenden innerbetrieblichen Abläufe. Was nämlich passiert in dem Moment, wo ein potenzieller Nutzer den Wert Ihrer Geodatenressource für sich erkennt und sich bei Ihnen aufgrund der Datenbeschreibung z. B. telefonisch meldet?

Um Ihnen Anhaltspunkte für eine Qualitätskontrolle im Hinblick auf einen reibungslosen innerbetrieblichen Ablauf zu gewährleisten, bieten wir Ihnen im Folgenden einen Ansatz, das Thema in Ihrem Hause aufzugreifen.

Im Metadatensatz gibt es die folgenden Kontakte, die Auskunft geben können:

- Kontakt für die Geodatenressource
- Kontakt für den Vertrieb der Geodatenressource
- Kontakt für den Inhalt des Metadatensatzes

Überprüfen Sie, ob die angegebenen Kontaktpersonen oder Kontaktteams tatsächlich zu den jeweiligen Themen Auskunft geben können:

- Weiß die Kontaktperson für die Geodatenressource in ausreichendem Maße über diese Geodaten Bescheid und ist sie in der Lage, dieses Wissen an einen Anfragenden verständlich weiter zu geben oder an den Vertrieb durchzustellen, sofern es um die Abgabe der Daten geht?
- Ist der Vertrieb dazu in der Lage Informationen zu den Kosten, dem Zugang und zur Nutzung herauszugeben oder bei Bedarf an den Kontakt für die Geodatenressource oder den Ansprechpartner für den Inhalt des Metadatensatzes weiterleiten?
- Kann die Kontaktperson für den Metadatensatz z. B. ad hoc einen Schreibfehler oder anderen offensichtlichen Fehler im Metadatensatz berichtigen? Verfügt die Kontaktperson für Metadaten über ausreichende Informationen zum technischen Betrieb der CSW-Schnittstelle und ggf. zum Harvesting?

Alle Ansprechpartner müssen voneinander Kenntnis besitzen. Jeder agiert in einem zuvor abgesprochenen Rahmen und leitet bei Bedarf an den jeweiligen anderen Ansprechpartner weiter. Jeder kann gleichermaßen einen notwendigen Revisionsprozess, den der Metadatensatz nach festem Muster durchlaufen muss, anstoßen.

Spielen Sie den „Notfall“ durch! – Was passiert, wenn sich ein Interessent...

- an die Kontaktperson für die Geodatenressource wendet?
- an den Vertrieb für die Geodatenressource wendet?
- an die Kontaktperson für den Inhalt des Metadatensatzes wendet?
- meldet, der etwas zum CSW wissen will, in dem er Ihren Metadatensatz gefunden hat?

Prüfen Sie zusätzlich, wie die Kontaktpersonen bei der Nutzung unterschiedlicher Kommunikationswege durch einen Suchenden reagieren.

Machen Sie sich selbst Zielvorgaben um einen angemessenen Qualitätsstandard zu etablieren:

- Wie lange darf es dauern, bis jemand eine Antwort erhält, der Ihnen eine Frage zu dieser Geodatenressource stellt?

Nehmen Sie sich dabei selbst als Maßstab. Überlegen Sie, wie lange Sie selbst auf eine Antwort zu warten bereit wären. Legen Sie im Zweifelsfall fest, nur bestimmte Kommunikationsmedien anzubieten, wenn sich andere als schwierig erweisen.

Der Metadatensatz wirkt wie Ihre Visitenkarte. Er und das Umgehen mit den in der Geodatenbeschreibung angebotenen Kontaktmöglichkeiten entscheiden darüber, wie Ihre Institution von Außenstehenden wahrgenommen wird. Durch ein festgeschriebenes Qualitätsmanagement vereinfacht und verbessert sich nicht nur das Zusammenspiel aller Beteiligten in ihrem Hause, sondern tatsächlich auch die Qualität Ihrer Produkte.

4. Die Präsentation von Metadaten

Im Geodatenbereich durchsuchen spezielle Metadatensuchen CSW-Kataloge. Sie haben eine für die jeweilige Zielgruppe geeignete Oberfläche, mit der die Suche durchgeführt und die gefundenen Ergebnisse präsentiert werden. Dabei besteht die Schwierigkeit vor allem darin, dass die Inhalte von Metadaten, die in xml vorliegen, keine eigene Darstellung, sondern lediglich eine vorgefertigte Struktur besitzen. Zwar liegt in der Regel eine Präsentation vor, die von dem originären Metainformationssystem gebildet wird. Sie ist jedoch nicht Bestandteil der Metadaten selbst.

Fachdatenhalter unterschiedlicher Disziplinen legen Wert auf spezielle Informationen, die Ihre Daten kennzeichnen. Dies führt dazu, dass neben den Pflichtfeldern verschiedene zusätzliche Elemente der ISO durch die Metadatenkataloge bedient werden. Nicht sämtliche Elemente sind für alle späteren Nutzer der Metadaten tatsächlich verständlich. Durch die gemeinsame Präsentation der Ergebnisse aus unterschiedlichen Metadatenkatalogen kommt somit diese weitere Erschwernis hinzu.

Als Betreiber einer Suchoberfläche für Metadaten arbeiten Sie im Spannungsfeld zwischen den Anforderungen des Datenanbieters und des Suchenden. Seien Sie in erster Linie bestrebt, eine „nützliche“ Suchoberfläche bereit zu stellen. Welche Form der Metadatenpräsentation sie wählen, ist von Ihrer Zielgruppe abhängig, die sie informieren möchten. Optimal ist eine Präsentationsoberfläche, die den Inhalt der Metadaten vollständig bedarfsgerecht interpretiert und damit für die Zielgruppe verständlich darstellt.

4.1 Anforderungen von Datenanbieter und Suchendem

Als Anbieter einer Präsentationsoberfläche werden Sie versuchen, dem Anliegen von Datenanbieter und Suchendem gerecht zu werden. Der Katalogbetreiber spielt bei der Überlegung eine untergeordnete Rolle, da

er durch den Betrieb des Kataloges lediglich die gewünschte Homogenität und die technische Verfügbarkeit der Metadaten sicherstellt.

Im Folgenden werden die unterschiedlichen Anforderungen an Geodatenbeschreibungen skizziert.

Typische Sichtweise eines Datenanbieters:

- Metadaten sind mein Aushängeschild
- Ich liefere eine fachlich korrekte Beschreibung
- Ich bilde die Daten technisch korrekt ab (z. B. Metadatenprofil des Bundeslandes, INSPIRE-Metadatenprofil, etc.)
- Inhaltlich ist die Beschreibung sowohl für meine Mitarbeiter als auch für Außenstehende geeignet
- Ich will, dass meine Daten gefunden werden
- Immer wenn eine Nachfrage zu einem Metadatensatz kommt, durchläuft dieser einer Revision, um ihn inhaltlich zu verbessern

Zwischen der Sichtweise eines Datenanbieters und eines Suchenden gibt es einen gravierenden Unterschied. Während Datenanbieter immer fachlich und technisch versierte Experten für **ihre eigenen** Geodaten sind, ist die Menge der Suchenden eine weit gefächerte Gruppe von Personen aus unterschiedlichsten Erfahrungsbereichen.

Suchende können sehr versiert sein, weil sie vielleicht selbst Daten halten. Häufig sind sie jedoch fachfremd und besitzen kein tiefer greifendes Verständnis dessen, was überhaupt „Geodaten“ sind. Viele Suchende haben lediglich eine erste Idee für Ihr Projekt im Kopf, für das Sie ein festes Budget besitzen, welches Sie einsetzen können. In der Regel können Sie auch ihre eigenen Fähigkeiten zumindest grob einschätzen, viele unterschätzen diese Fähigkeiten jedoch, da sie „Fachbegriffe“ als fremdartig empfinden.

Typische Sichtweise eines Suchenden:

- Ich möchte schnell die gesuchten Daten für mein Projekt finden
- Ich erwarte Hinweise auf artverwandte Daten, die mich ebenfalls in meinem Projekt weiterbringen könnten
- Das ist für mich die einfachste Möglichkeit „alles“ über die gewünschten Geodaten zu erfahren, ohne dass ich lange herumtelefonieren oder gar eine Firma beauftragen muss, die womöglich die Recherche für mich erledigt
- Ich kann Testdaten direkt herunterladen und prüfen, ob die Daten meinem Kriterienkatalog standhalten
- Ich kann die Bedingungen für einen Zugriff und den Kostenaufwand abschließend beurteilen

Bezüglich Ihrer Ziele ähneln sich Datenanbieter und Suchender, den beide versprechen sich Vorteile aus der Recherchemöglichkeit. Dem Minimieren des Aufwands für die Datenabgabe und dem Gewinnen neuer Kunden steht eine Zeit- und Kostenersparnis durch die Vereinfachung der Recherche auf Seite der Suchenden gegenüber.

4.2 Möglichkeiten einer Präsentation von Metadaten

Bei der Präsentation von Metadaten bieten sich unterschiedliche Möglichkeiten für die Umsetzung an.

Metadaten können rein **schematisch** präsentiert werden. Hierbei orientiert sich die Darstellung nah am Aufbau, also der xml-Struktur des Metadatensatzes. Der Vorteil ist, dass sich so sämtliche Elemente der ISO gleichermaßen abbilden lassen. Dem Nutzer der Daten werden alle vorhandenen Informationen offen angeboten. Die Hilfen, diese Informationen jedoch auch inhaltlich vollständig zu verstehen, sind in diesem Falle zwangsläufig aufgrund der Fülle an möglichen ISO-Elementen eher gering.

Eine **inhaltlich strukturierte** Darstellung wirkt in der Regel auf den ersten Blick aufgrund ihrer erhöhten Übersichtlichkeit anwenderfreundlicher. Die Information wird auf ausgewählte Metadatenelemente reduziert und inhaltlich gebündelt. Nicht dargestellte Elemente dienen vor allem der internen Verwaltung von Metadaten. Sie werden dem Anwender wissentlich vorenthalten, da sie für die angestrebte Zielgruppe keinen Wert besitzen.

Ergänzend zu diesen Formen der rein schematischen und rein inhaltlich strukturierten Darstellung von Metadaten wird häufig auch ein **direkter Link auf den xml-Metadatensatz** selbst angeboten. Dieses Angebot kann jedoch nur von Metadatenexperten korrekt genutzt werden.

Kegelrobbezählung im Nationalpark Niedersächsisches Wattenmeer Winter/Frühjahr 2010/2011 (UIG) Grey seals in the Wadden Sea National Park of Lower Saxony

Um ein genaueres Bild über den Bestand vor Ort zu bekommen, führt die Nationalparkverwaltung im Winter und Frühjahr spezielle Erfassungsfüge durch. Im Winter wird der Nachwuchs geboren und im Frühjahr schließt sich der jährliche Fellwechsel der Kegelrobben an - beide Lebensabschnitte bieten eine gute Gelegenheit, die Tiere an Land zu zählen.

Thematik: biota, oceans, environment
Schlüsselwörter: Kegelrobbe, Kegelrobben, Meeressäuger, Halichoerus grypus, Grey seal, Nationalpark Niedersächsisches Wattenmeer, Wadden Sea National Park of Lower Saxony, Nordsee, Niedersächsisches Küstenmeer, Biologie, Fauna, inspireidentifiziert, INSPIRE, opendata, Environmental monitoring facilities, Umweltüberwachung
Sprache: ger
Durchschnittliche Bewertung: ★★★★★
Anbieter: Nationalparkverwaltung Niedersächsisches Wattenmeer

Details	Kontakt	Beschränkungen	Kommentare
Status			
Datum der Erstellung:	2012-11-05		
Datum der Veröffentlichung:	2012-11-05		

Abb.: Inhaltlich strukturierter Metadatensatz

Weiterhin besteht die nutzerfreundliche Möglichkeit, **nur ausgewählte Elemente** aus Metadaten dazu zu verwenden, z.B. Listen für vorhandene Geodaten auf Webseiten aktuell anzubieten. Einzelne Elemente aus Metadaten werden also dazu benutzt, erläuternde Webseiten automatisiert mit aktuellem Inhalt zu befüllen.

Im Hintergrund arbeitet ein Metadatenkatalog (CSW). So können „Metadaten“ auch mit den üblichen Suchmaschinen schnell auffindbar gemacht werden. Dem Nutzer bleibt an dieser Stelle verborgen, dass überhaupt eine Metadatenutzung im eigentlichen Sinne erfolgt. Somit ist diese Lösung in der Regel die anwenderfreundlichste Methode der Präsentation von Metadaten.

4.3 Der Datenanbieter als Ansprechpartner

Innerhalb der Metadaten ist ein besonderes Augenmerk auf die Ansprechpartner zu richten. Gerade durch die gezielte Benennung von Verantwortlichkeiten für die Geodatenressourcen, den Vertrieb und die Metadaten wird die Kommunikationsaufnahme zwischen Datenanbieter und Suchendem deutlich verbessert. Wird in der Präsentation von Metadaten die getrennte Darstellung der Ansprechpartner realisiert, kann sich der Suchende direkt an den passenden Ansprechpartner wenden, was eine deutliche Zeitersparnis im Rechercheprozess bedeutet.

Derzeit weisen die meisten Metadaten für diese drei Kontaktangaben eine identische Angabe auf. Sobald sich jedoch die vorgesehene Nutzung von Metadaten in der Praxis durchsetzt und die damit verbundenen Möglichkeiten voll ausgeschöpft werden, ist damit zu rechnen, dass die Datenanbieter hier genauere Unterscheidungen werden treffen müssen, um eine effiziente Abwicklung innerhalb der GDI-Prozesse zu gewährleisten.

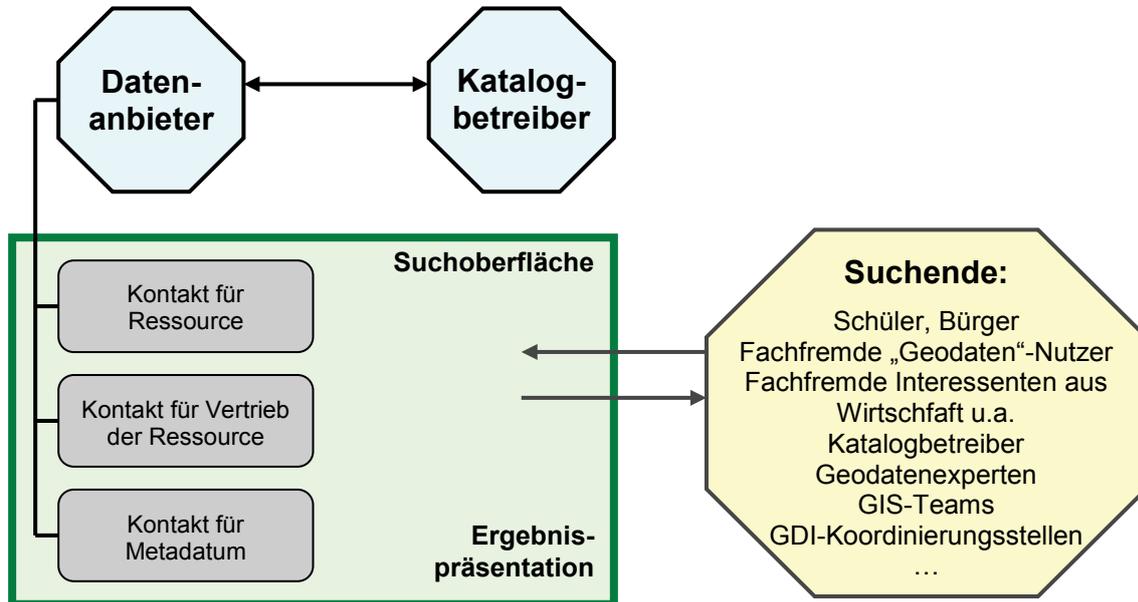


Abb.: Die Rollen „Katalogbetreiber“, „Datenanbieter“ und „Suchende“ in Bezug auf die Entwicklung einer Suchoberfläche für Metadaten

4.4 Zielgruppengerechte Präsentation

Um eine zielgruppengerechte Präsentation zu erreichen müssen die Ergebnisse der Suche in einer angemessenen Reihenfolge angeboten werden als auch die Aussagekraft der Metadaten selbst gewährleistet sein.

Bei der Präsentation des Suchergebnisses werden die gefundenen Metadaten in eine Reihenfolge gebracht, die dem Nutzer die „nützlichsten“ Daten an erster und die weniger „nützlichen“ an letzter Stelle präsentieren. Die Bildung einer Reihenfolge nach Relevanz bezeichnet man auch als „Ranking“.

Die einfachste Form eines Rankings erfolgt durch das alleinige Auszählen bestimmter Wörter in einer Beschreibung. Wird das Suchwort besonders häufig in der Datenbeschreibung gefunden, so wird der Metadatensatz weiter vorne in der Liste geführt. Es bestehen jedoch weitere Kriterien ein Ranking durchzuführen, wie z. B. die Zugehörigkeit von Daten einer bestimmten Verwaltungsebene zuerst zu präsentieren oder Daten eines bestimmten Fachgebietes den Vorzug zu geben. Auch Kombinationsmöglichkeiten oder Vorabgruppierungen wie die Unterscheidung zwischen gefundenen Geodaten und Geodatendiensten bieten sich an.

Um die Aussagekraft von einzelnen Metadaten zu erhöhen werden die einzelnen Elemente auf der Suchoberfläche mit Titeln oder Beschreibungen versehen, die für die jeweilige Zielgruppe verständlich und nachvollziehbar sind.

Folgende Punkte beeinflussen typischerweise eine Applikation für die Metadatensuche:

- ist auf eine bestimmte Zielgruppe gerichtet
- zeigt nur die für diese Zielgruppe angemessenen Inhalte
- verarbeitet fremde Inhalte
- Inhalte werden so präsentiert, dass sie für die Zielgruppe verständlich sind
- wird an ein bestehendes Design angepasst
- die Handhabung orientiert sich an bereits vorhandenen Applikationen in diesem Webauftritt

Der Nutzer soll sich „heimisch“ fühlen, den Auftritt als nützlich bewerten und zu einem späteren Zeitpunkt erneut nutzen. Beachten Sie bei Ihrer gewählten Darstellung auch die Intention des Datenanbieters, die mit der Erstellung von Metadaten verbunden ist. Er möchte die Nachfragen zu den angebotenen Geodatenressourcen auf ein Minimum reduzieren.

In den folgenden Abbildungen sehen Sie einen Metadatensatz in seinem originären Katalogsystem und in zwei Katalogsystemen, die den gleichen Metadatensatz nach dem Harvesting anzeigen.



Basisinformation

Titel	Flächennutzungsplan der Stadt Helmstedt
Datum	2014-02-24
Datumstyp	Publikation: Das Datum gibt an, wann die Ressource publiziert wurde
Code	e07251ac-5c82-40bd-bef1-42b172c09e77
Präsentationsform	Digitale Karte: Digitale Karte in Raster- oder Vektorform
Kurzbeschreibung	<p>Der Flächennutzungsplan ist ein vorbereitender Bauleitplan gemäß BauGB. Aus dem Flächennutzungsplan werden die Bebauungspläne entwickelt.</p> <p>Die Aufstellung des Flächennutzungsplanes ergibt sich aus §§ 5-7 BauGB.</p> <p>Dargestellt wird die aktuelle Leseversion.</p>
Zweck	Information
Bearbeitungsstatus	Kontinuierliche Aktualisierung: Daten werden fortlaufend aktualisiert

Kontakt für die Ressource

Abb.: Auszug der Ansicht der Original-Datenbeschreibung im originären Erfassungssystem (Metadatenerfassung im Geodatenportal GDI-NI)

Flächennutzungsplan der Stadt Helmstedt

2014-02-24 Stadt Helmstedt, Fachbereich Planen und Bauen ★ ★ ★ ★ ★ (0) [kommentieren](#)

Der Flächennutzungsplan ist ein vorbereitender Bauleitplan gemäß BauGB. Aus dem Flächennutzungsplan werden die Bebauungspläne entwickelt. Die Aufstellung des Flächennutzungsplanes ergibt sich aus §§ 5-7 BauGB. Dargestellt wird die aktuelle Leseversion.

INSPIRE

[Gebiet](#) € [Bestellen](#) [Download](#) [Drucken](#) [XML](#) [KML](#)
▼ [Weniger](#)

Flächennutzungsplan der Stadt Helmstedt

Der Flächennutzungsplan ist ein vorbereitender Bauleitplan gemäß BauGB. Aus dem Flächennutzungsplan werden die Bebauungspläne entwickelt. Die Aufstellung des Flächennutzungsplanes ergibt sich aus §§ 5-7 BauGB. Dargestellt wird die aktuelle Leseversion.

Thematik: planningCadastre

Schlüsselwörter: Bauleitplanung, Flächennutzung, Flächennutzungsplan, Flächennutzungsplanung, Bodennutzung, inspireidentifiziert, F-Plan, FNP, vorbereitender Bauleitplan

Sprache: ger

Durchschnittliche Bewertung: ★★★★★

Anbieter: Stadt

Details

Kontakt

Beschränkungen

Kommentare

Abb.: Auszug der Ansicht der geharvesteten Daten (Geoportal.de)

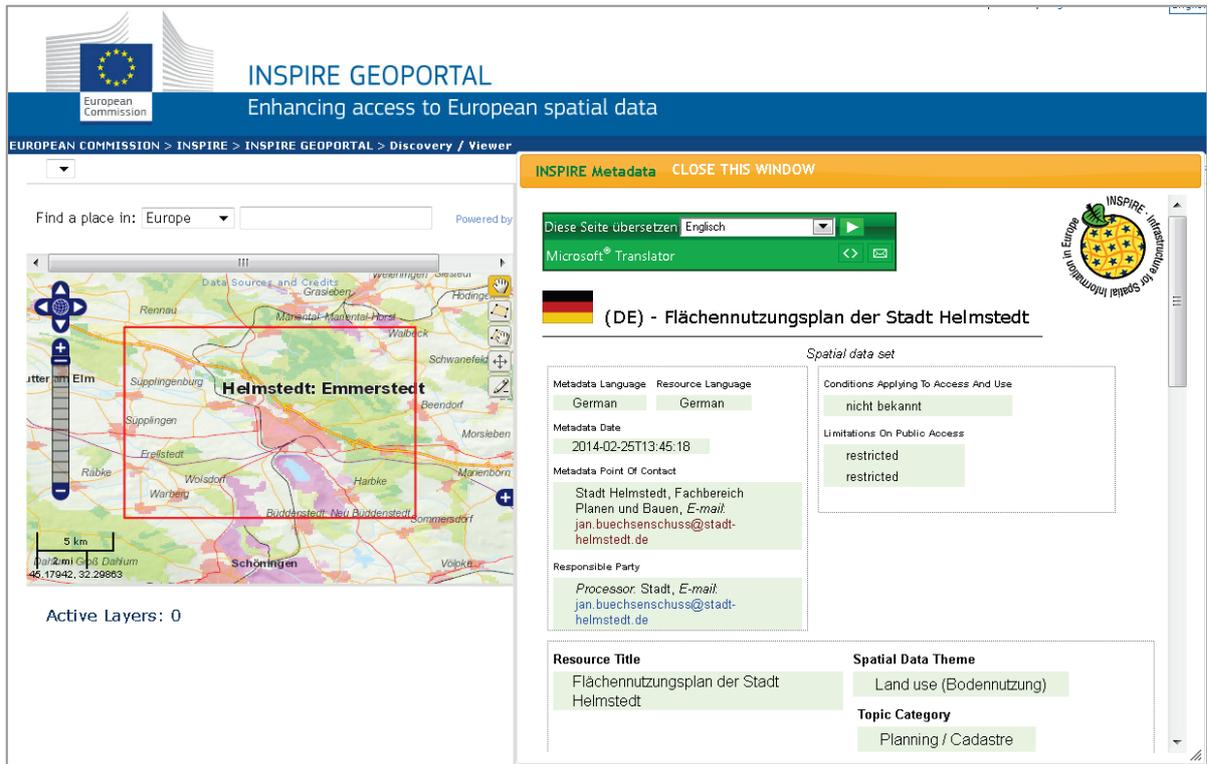


Abb.: Auszug der Ansicht der geharvesteten Daten (INSPIRE-Portal)

Den Nutzern Ihrer Präsentationsoberfläche dürfen, trotz der unter Umständen auf nur bestimmte Elemente beschränkten Ansicht, keine Informationen vorenthalten werden. Eine Lösung dieses Paradoxon kann mit einem Link zum originären Metadatensatz aufgelöst werden. Damit besteht für den Suchenden die Möglichkeit, den Metadatensatz in derjenigen Oberfläche anzuschauen, für die er ursprünglich angefertigt wurde. In dieser Oberfläche sind dann auch diejenigen Metadatenelemente zu sehen, die aus Verständnisgründen in Ihrer Oberfläche möglicherweise bewusst weggelassen wurden.

Das Metadaten-Element 397 enthält im Idealfall einen **Link auf die Originalpräsentation des Metadatensatzes im Internet**, über das ein Suchender per Klick zu der originären Darstellung der Datenbeschreibung gelangen kann.

Hier ein Auszug aus der xml-Ansicht eines Metadatensatzes für das Element 397

MD_Metadata.contact.CI_ResponsibleParty.contactInfo.CI_Contact.onlineResource.CI_onlineResource.linkage
[1].[8].[374].[378].[387].[390].[396].[397]:

```
<gmd:MD_Metadata>
  <gmd:contact>
    <gmd:CI_ResponsibleParty>
      <gmd:contactInfo>
        <gmd:onlineResource>
          <gmd:CI_OnlineResource>
            <gmd:linkage>
              <gmd_URL>http://link.auf-originalpräsentation.de/</gmd_URL>
            </gmd:linkage>
          </gmd:CI_OnlineResource>
        </gmd:onlineResource>
      </gmd:contactInfo>
    </gmd:CI_ResponsibleParty>
  </gmd:contact>
</gmd:MD_Metadata>
```

Denkbar sind ebenso die folgenden drei Varianten:

1. Link auf den GetCapabilities-Request zum originären CSW
2. Link zur Internetoberfläche des originären Erfassungskataloges
3. Hinterlegen eines GetRecordByld-Request auf den Metadatensatz

Das Hinterlegen eines GetCapabilities-Requests auf Ihren Katalog oder aber auf die Weboberfläche, die Ihren Katalog im Internet abbildet sind für eine weiterführende Information weniger gut geeignet, da durch beide Varianten der Metadatensatz nicht direkt von einem Suchenden in seinem originären Katalog gefunden werden kann. Beim Hinterlegen des GetRecordByld Request wird die xml Ausprägung des Metadatensatzes aufgerufen. Diese ist für Metadaten-Laien in der Regel unverständlich.

Im Idealfall kann dem Nutzer die Suchoberfläche alle vier Varianten entsprechend präsentieren und eine Hilfestellung mit auf den Weg geben, wie die hinterlegte Angabe vom Suchenden zu nutzen ist, um den Metadatensatz in seinem originären Katalog anschauen zu können.

4.5 Sprachneutrale Code-Wörter, Code-Listen und sonstige Konventionen

Üblicherweise werden Metadaten in der GDI-DE in der deutschen Sprache abgefasst. Innerhalb der ISO-Metadaten wird jedoch in etlichen Fällen auf sprachneutrale Code-Wörter oder Code-Listen sowie weitere spezielle Begriffen und Strukturen zurückgegriffen. Die Angaben in einem Metadatensatz beziehen sich nicht ausschließlich auf die beschriebene Geodatenressource, sondern z. B. ebenso auf den Metadatensatz und dessen eigener Ausgestaltung.

Auf der einen Seite handelt es sich somit um in Metadaten abgelegte Informationen, die sich nicht mit der Beschreibung der gesuchten Geodatenressource beschäftigen. Auf der anderen Seite handelt es sich gegebenenfalls um erklärungsbedürftige Inhalte, die dem zukünftigen Nutzerkreis mit Bedacht und angemessen zu präsentieren sind. Die Bedeutung all dieser Besonderheiten erschließt sich demjenigen, der auf der Suche nach Geodaten und Diensten ist, in der Regel nicht von allein. Eine Unterstützung durch eine angemessene Präsentation ist damit erforderlich.

Häufig finden sich so genannte sprachneutrale Code-Wörter aus relevanten EU-Verordnungen und auch Einträge aus ISO Code-Listen in den Metadaten. Code-Wörter und Code-Listen tragen dazu bei, Metadaten über sprachliche Grenzen hinaus auffindbar zu machen und Daten und Dienste in inhaltliche Klassen zu unterteilen. Die Klassen können genutzt werden, um Metadaten in bestimmte Gruppen zu unterteilen. Eine solche Gruppe bildet z. B. die Gruppe der deutschsprachigen Metadaten. Denkbar ist jedoch auch eine thematische Klassifikation der beschriebenen Geodatenressourcen.

Zu unterschiedlichen Zwecken erfolgt die Verwendung bestimmter innerhalb der GDI-DE vereinbarter Schlagwörter. Sie schaffen die Grundlage für eine einwandfreie technische Zuordnung der Metadaten zu bestimmten Gruppen, die deutschlandweit geregelt wird. Um Metadaten beispielsweise in das GovData-Portal transferieren zu können, wurden konkrete Vereinbarungen innerhalb der GDI-DE getroffen, die Nutzungsbedingungen in einer bestimmten, für Menschen jedoch schwer lesbaren Form, innerhalb der xml-Struktur abzubilden.

Diese Vielfalt an zunächst nicht lesbaren Informationen gilt es auch für einen Laien in Sachen Metadaten in verständlicher Form zu interpretieren und zu präsentieren. Aus diesem Grunde ist es erforderlich, sich mit der Architektur der GDI-DE, hier mit dem Dokument „Konventionen zu Metadaten der Geodateninfrastruktur Deutschland (GDI-DE)“ eingehend zu beschäftigen. In diesem Dokument sind die konkreten Anforderungen an Metadaten formuliert. Hinzu kommen die inhaltlichen Umsetzungen aus den einschlägigen ISO-Standards und EU-Verordnungen zu Metadaten.

4.6 Verwenden semantischer Thesauri

Um Fachfremden die eigenen Geodatenressourcen näher zu bringen, kann durch die Anbindung semantischer Thesauri das in den Metadaten vorhandene Fachvokabular vereinheitlicht und von einem größeren Personenkreis genutzt und erschlossen werden.

Mit Hilfe des Resource Description Framework (RDF) werden logische Aussagen in Bezug auf die Verknüpfung und Strukturierung von Begriffen getroffen. Semantische Thesauri verbinden auf diese Art und Weise bei Bedarf unterschiedliche Begriffsgruppen so, dass z.B. bei Eingabe eines übergeordneten Begriffs wie „Süßmittel“ ebenso die Ergebnisse für die Suchbegriffe „Honig“ oder „Stevia“ gefunden werden.

Die Möglichkeit der Verschlagwortung über Begriffe des **General Multilingual Environmental Thesaurus (GEMET)** unterstützt sowohl eine semantische Suche als auch eine automatisierte Mehrsprachigkeit von Metadaten im Sinne von INSPIRE. Zusätzlich können mehrere semantische Fachthesauri zu Rate gezogen werden, um logische Verbindungen zwischen Inhalten unterschiedlicher Fachdisziplinen herzustellen. Der maßgebliche deutschsprachige Fachthesaurus im Umweltbereich, der **Umwelt-Thesaurus** des Umweltbundesamtes (UMTHES) bietet ebenfalls seine Inhalte über RDF an. Denkbar sind weiterhin räumliche Bezüge, die mit Hilfe von RDF ermittelt werden können. Ein Beispiel ist der Getty Thesaurus of Geographic Names[®], eine hierarchische Verknüpfung von weltweiten Orten.

In der Praxis bestehen aktuell jedoch erhebliche technische Probleme zwischen den verschiedenen zu bedienenden Schnittstellen und Standards, um eine derartige semantische Verbindung herzustellen. Einer

vorausschauenden Gestaltung von Metadateninhalten und deren Präsentationsoberflächen spricht dies jedoch keineswegs entgegen.

5. Hinweise für Katalogbetreiber

Welche Hinweise für Katalogbetreiber wichtig sind, ist im Wesentlichen davon abhängig, ob Ihr Katalog ausschließlich mit originären Metadaten ihrer eigenen Institution arbeitet oder ob Sie einen Katalog betreiben, der zwar ausschließlich mit originären Metadaten, jedoch möglicherweise mit Metadaten unterschiedlicher Institutionen bestückt ist.

Sofern Sie zusätzlich die Beschreibung fremder Geodatenressourcen in Ihrem Katalog vorrätig halten möchten, importieren Sie in regelmäßigen Abständen die Beschreibungen anderer Anbieter. Diesen Vorgang nennt man „Harvesting“. Die Praxis zeigt, dass die Grenzen zwischen Katalogen mit originären Metadaten und Katalogen mit geharvesteten Metadaten fließend sind. Sollten Sie einen Katalog betreiben, der ausschließlich mit geharvesteten Metadaten verschiedener Institutionen befüllt ist, ergeben sich Anforderungen, die wir Ihnen ebenfalls in diesem Kapitel kurz vorstellen.

Grundsätzlich gilt, dass ein originärer Katalog immer über eine CSW-Schnittstelle verfügen sollte, über die er Daten abgeben kann. Harvesting-Kataloge werden zumeist ausschließlich für bestimmte Aufgaben zur Bündelung von Informationen betrieben und müssen CSW-Schnittstellen somit vor allem lesen können. Eine Abgabe von Metadaten per CSW ist für Harvesting-Kataloge nicht zwingend erforderlich.

5.1 Kataloge mit direkt erfassten Metadaten (Originäre Kataloge)

Als Katalogbetreiber gewährleisten Sie den reibungslosen Betrieb des CSW. Für den Inhalt der Metadaten ist vorrangig ist der Erfasser von Metadaten (Ansprechpartner für den Metadatenkontakt) verantwortlich. Nichts desto trotz tragen Sie als Katalogbetreiber wesentlich dazu bei, einen gleichbleibenden inhaltlich und technischen Qualitätsstandard aller Metadaten im Katalog zu erreichen.

Wenn in einem Katalog originäre Metadaten geführt werden, so bedeutet das, dass dies die Originalapplikation ist, in der die Metadaten gehalten werden. Es ist möglich, dass die Erfassungsoberfläche eine andere Software ist als diejenige, mit der die Metadaten als CSW abgegeben werden. Mit diesem Kapitel dürfen Sie sich grundsätzlich angesprochen fühlen, wenn Sie eine Anwendung betreiben, die Metadaten originär hält, denn daraus ergeben sich für Sie eine Reihe von Verantwortlichkeiten, derer Sie sich bewusst sein müssen.

In der Regel werden Metadaten in einem solchen originären System gepflegt, weil z. B. die technische Anbindung an die vorliegenden Geodatenressourcen besonders komfortabel gestaltet ist. Beispielsweise können bestimmte Bestandteile der Metadaten aus einer vorhandenen Datenbank oder einem bestehenden Archiv abgeleitet werden und die Grundlage für die zu erstellenden Metadaten bieten, die von einem Erfasser

in einem weiteren Arbeitsprozess entsprechend ergänzt werden. Ebenso ist es möglich, Metadaten komplett neu anzulegen oder z.B. analoge Datenbeschreibungen in eine einheitliche digitale Form zu bringen.

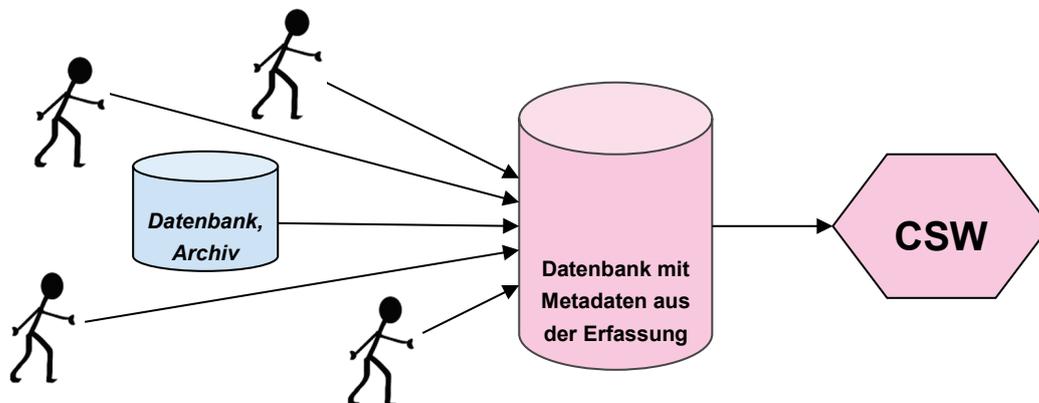


Abb.: Katalog mit direkt erfassten oder erstellten Metadaten (Originärer Katalog)

Als Katalogbetreiber gewährleisten Sie gegenüber Ihren Metadaten erfassenden Institutionen, dass die Metadaten den geltenden Profilen entsprechen. Aus diesem Grunde sollten Sie auch Ihre Ansprechpartner für die Erfassung der Metadaten kennen und sie z. B. über einen E-Mail Verteiler informieren können, sofern bestimmte Sachverhalte dies erfordern. Beispielsweise könnte es möglich sein, „ab sofort“ nur noch bestimmte Schlagwörter zu verwenden, um Datensätze z. B. für die im Aufbau befindliche Nationale Geodatenbasis oder aber auch für INSPIRE automatisiert zu melden.

Deutschland- und europaweit werden die technischen Metadatenprofile weiterentwickelt. Wir bitten Sie als Katalogbetreiber daher, in Kontakt mit uns zu treten, damit wir Sie über Neuerungen in Bezug auf die Profile oder die deutschlandweiten Festsetzungen angemessen informieren können. Des Weiteren ist es für die Bereitstellung von Metadaten auf der europäischen Ebene unerlässlich, dass Sie uns die URL Ihrer CSW-Schnittstelle mitteilen. Die Inhalte Ihres Kataloges werden über die Software des Geodatenportals Niedersachsen damit ein Bestandteil der GDI-DE.

Legen Sie für Ihren Katalog fest, in wie weit technisch bedingte Änderungen am Metadatensatz dazu führen, dass dieser ein aktualisiertes Metadaten-Datum erhält. Bedenken Sie, dass ein Metadatensatz nur dann erneut von fremden Applikationen geharvestet wird, sofern das Metadatendatum aktualisiert wurde. Bedenken Sie aber auch, dass die Metadatenerfasser das Datum der letzten Änderung des Metadatensatzes ebenso benutzen und ein automatisch von Ihnen aktualisiertes Metadatum zu Missverständnissen führen könnte.

Sofern Sie technisch bedingte inhaltliche Änderungen an den Metadatensätzen vornehmen, informieren Sie Ihre Metadatenerfasser über diese Änderungen. Im Folgenden geben wir ein Beispiel für eine technisch bedingte inhaltliche Änderung an den Metadatensätzen eines Kataloges:

Versehentlich wurde eine ISO codeList für einen angegebenen Wert beim Anlegen eines neuen Metadatensatzes durch die verwendete Software nicht gesetzt, sondern die entsprechende Angabe blieb leer:

```
...  
<gmd:characterSet>  
  <gmd:MD_CharacterSetCode codeList="" codeListValue="utf8"/>  
</gmd:characterSet>  
...
```

Über die gesamte Datenbank musste dieser Parameter nun mit dem zutreffenden Wert für die codeList versehen werden. Es wurde daraufhin ein SQL ausgeführt, welches in allen Metadatensätzen den Fehler berichtigte.

```
...  
<gmd:characterSet>  
  <gmd:MD_CharacterSetCode codeList="MD_CharacterSetCode"  
    codeListValue="utf8"/>  
</gmd:characterSet>  
...
```

Der Fehler im Erfassungsprogramm wurde zwischenzeitlich behoben, so dass neu erstellte Datensätze den Fehler nicht mehr aufwiesen und die codeList wie vorgesehen befüllt wurde.

Inhaltliche Änderungen an einem Metadatensatz dürfen grundsätzlich nicht durch den Katalogbetreiber erfolgen. Insbesondere bei ungeübten Erfassern kann jedoch die direkte Unterstützung bei der Erfassung eines Metadatums durch den Katalogbetreiber in Einzelfällen zwingend notwendig werden. Sie darf jedoch nur in direkter Absprache mit dem jeweiligen Erfasser des Metadatensatzes zur z. B. Veranschaulichung einer bestehenden Eingabemöglichkeit und auf dessen Wunsch erfolgen. Ein besser geeignetes Mittel, als Katalogbetreiber inhaltlich notwendige Änderungen durchzusetzen ist es, den Metadatensatz vorübergehend aus dem Katalogangebot zu entfernen und dem Erfasser den Änderungswunsch aufzuzeigen. Nach der Überprüfung des Metadatums durch den Erfassern gibt dieser den Metadatensatz wieder zur Veröffentlichung über die Katalogschnittstelle frei.

Erarbeiten Sie Regeln, mit denen Sie die Qualität und die Aktualität der Metadaten aus Ihrem Katalog gewährleisten. Dies kann sich auf eine inhaltliche Prüfung der hinterlegten Daten beziehen, kann aber auch auf bestimmte technische Vorgaben abzielen. Zu einer gleichbleibend hohen Qualität trägt eine Erfassungsmaske bei, die per Listenauswahl und Plausibilitätsprüfung die fehlerfreie Erfassung unterstützt. Differenzieren Sie, welche Qualitätsanforderungen durch die Metadatenerfasser und welche durch sie als Katalogbetreiber herbeigeführt werden.

Als Katalogbetreiber stehen Sie in zuverlässigem Kontakt mit Ihren Erfassern und uns als Koordinierungsstelle GDI-NI. Sie klären ebenfalls Probleme des Harvestings mit den Stellen, die Ihren Katalog harvesten möchten. Geben Sie außerdem die Nutzungsbedingungen bekannt, die bei der Nutzung Ihrer CSW-Schnittstelle einzuhalten sind.

Im Rahmen des Harvestings durch Dritte werden Sie häufig ein Unterscheidungsmerkmal zwischen den Metadaten Ihres Katalogs und den Metadaten aus anderen Katalogen benennen müssen. Hierfür gibt es im Metadatensatz den so genannten „Metadatenkontakt“. In diesem Kontakt kann eine URL auf das Erfassungssystem dieses Metadatensatzes hinterlegt werden. Achten Sie als Katalogbetreiber darauf, dass an dieser Stelle in jedem Metadatensatz die URL „von diesem Metadatensatz selbst“ angegeben wird. Es handelt sich hierbei um das Metadatenelement 367 MD_Metadata.contact.CI_ResponsibleParty.contactInfo.CI_Contact.onlineResource.CI_onlineResource.linkage [1].[8].[374].[378].[387].[390].[396].[397].

Durch einen Klick auf den dort hinterlegten Link sollte der Metadatensatz in derjenigen Applikation dargestellt werden, in der oder für die er erfasst wurde. Diesen Link könnten Sie beispielsweise durch einen automatisierten Lauf auf die Datenbank standardmäßig setzen.

Geben Sie im GetCapabilities-Dokument Ihres CSW bzw. zusätzlich in der für diesen CSW zur Verfügung stehenden Oberfläche Auskunft zu den folgenden Sachverhalten:

- Wer sind Sie?
- Welche Institutionen erfassen in diesem Katalog Metadaten?
- Durchlaufen die Metadaten bei oder nach der Erfassung eine Qualitätskontrolle?
- Geben Sie die URL für diesen Katalog an (CSW, bei Bedarf Oberfläche).

Wie aber erfahren nun andere von der Existenz Ihres CSW? – Damit wären wir wieder beim Anfang, denn als Betreiber einer solchen Fachapplikation sind Sie gehalten, einen inhaltlich aussagekräftigen Metadatensatz für Ihren Katalog anzulegen.

5.2 Kataloge mit geharvesteten Metadaten (Harvesting-Kataloge)

Ein Katalog mit geharvesteten Metadaten enthält in seiner reinsten Form ausschließlich Metadaten fremder CSW-Schnittstellen. Das Einrichten eines Harvesting-Katalogs ist immer dann sinnvoll, wenn Datenbeschreibungen aus sehr vielen unterschiedlichen Quellen zusammengetragen werden müssen, um z. B. eine entsprechende Suchoberfläche für eine bestimmte Zielgruppe bereit zu stellen. Ziel ist es, ausschließlich solche Metadaten im Katalog zusammen zu tragen, die dieser Zielgruppe „nützlich“ sind. Die Schwierigkeit liegt klar in der einwandfreien Abschätzung dieser „Nützlichkeit“, denn wer außer den Suchenden will entscheiden, was „nützlich“ und was vielleicht „überflüssig“ ist?

Als Betreiber eines Katalogs mit geharvesteten Metadaten tragen Sie die Verantwortung dafür, dass die geharvesteten Metadaten dem potenziellen Nutzer in der Originalfassung zur Verfügung gestellt werden.

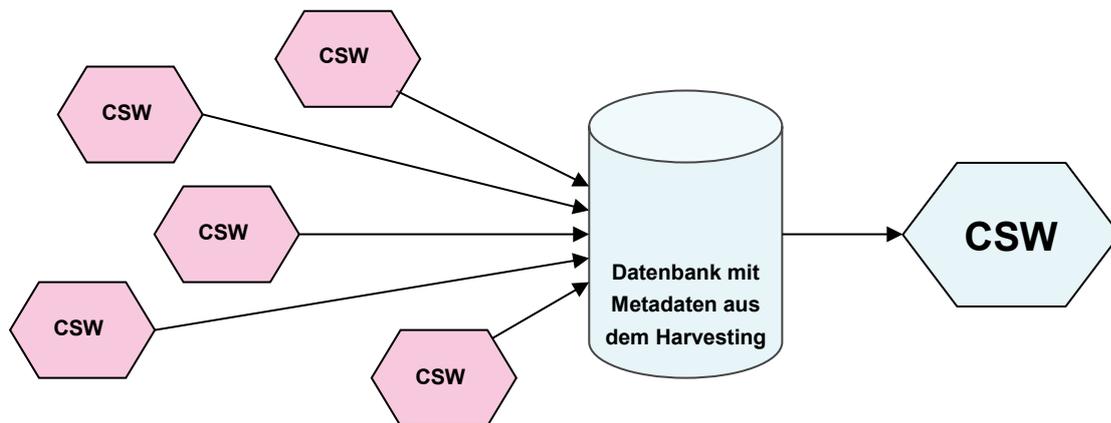


Abb.: Katalog mit geharvesteten Metadaten (Harvesting-Katalog)

Geben Sie im GetCapabilities-Dokument Ihres CSW bzw. zusätzlich in der für diesen CSW zur Verfügung stehenden Oberfläche Auskunft zu den folgenden Sachverhalten:

- Wer sind Sie?
- Wer ist die Zielgruppe dieses Katalogs?
- Welche CSWs harvesten Sie mit dieser Applikation?
- In welchem Turnus werden die CSWs geharvestet?
- Gegebenenfalls Link auf die Suchoberflächen der geharvesteten CSWs mit originären Metadaten setzen
- Geben Sie die URL für diesen Harvesting-Katalog an (Oberfläche, bei Bedarf CSW).
- Beschreiben Sie die Art und Weise des Harvestings einzelner Kataloge, sofern Sie dies als notwendig erachten (Wird gefiltert? Erfolgt das Harvesting eines Gesamtkatalogs? Werden UUIDs explizit ausgeklammert?)

Wenn Sie das Harvesting durchführen, ist es erforderlich, eine feste Reihenfolge der zu harvestenden Kataloge vorab festzulegen. Grundsätzlich sollten Sie immer zuerst Kataloge harvesten, die ausschließlich originäre Metadaten enthalten. Danach folgen Kataloge mit sowohl originären als auch geharvesteten Daten. Kataloge, die ausschließlich geharvestete Daten erhalten, sollten nicht erneut geharvestet werden. Verwenden Sie stattdessen die URLs derjenigen CSW, die hier Daten zuliefern.

Teilen Sie den Katalogbetreibern der durch Sie geharvesteten CSW mit, dass Sie beabsichtigen, die Daten zu harvesten. Besprechen Sie mit dem Katalogbetreiber jeden CSWs, in welchem Turnus Sie seine Daten harvesten möchten und bestimmen Sie gemeinsam bei Bedarf zusätzlich einen Termin bzw. legen Sie die Uhrzeit fest, an dem Sie beabsichtigen, das Harvesting durchzuführen.

Stellen Sie sicher, dass das Harvesting einzelner Kataloge durch Sie technisch korrekt erfolgt und ausschließlich Metadaten geharvestet werden, die aus Sicht des Betreibers des originären Katalogs auch geharvestet werden sollen. Fragen Sie beim Betreiber des originären Metadatenkataloges nach, wie viele Metadatensätze vorhanden und für das Harvesting durch Sie freigegeben sind. Überprüfen Sie

stichprobenhaft zusätzlich die inhaltliche Qualität der geharvesteten Metadaten. Machen Sie den Betreiber des originären Kataloges auf vorhandene Defizite aufmerksam und bitten Sie um kurzfristige Abhilfe.

Ziehen Sie immer auch in Erwägung, bestimmte Metadaten, die nicht Ihren Qualitätsanforderungen genügen, durch den Einsatz von Filtern nicht in den von Ihnen betriebenen Katalog aufzunehmen. Denkbar sind hier entweder inhaltliche Filter oder aber auch das Anlegen einer Blacklist mit UUIDs der Metadatenatz-Identifikatoren. Durch den Einsatz der Blacklist können Metadaten gekennzeichnet werden, die entweder nicht geharvestet werden oder die nach einem erfolgten Harvesting automatisiert gelöscht werden.

Sofern Metadatenätze eine mangelhafte Qualität aufweisen, teilen Sie dies bitte dem Katalogbetreiber des originären Katalogs mit oder wenden Sie sich direkt an den Metadatenkontakt für diesen Datensatz. Ziel ist es, langfristig die Qualität der vorhandenen Metadaten gemeinsam zu verbessern.

Erstellen Sie einen Metadatenatz für Ihre Fachapplikation und veröffentlichen Sie diesen in einem geeigneten System.

5.3 Kataloge mit Metadaten gemischter Herkunft (Gemischte Kataloge)

Das Einspeisen von Metadaten in einen gemischten Katalog kann sowohl durch das Harvesting Kataloge Dritter, als durch die direkte Datenerfassung, aber auch durch vollautomatisierte Importprozesse von automatisiert erzeugten Metadaten oder Metadaten aus Datenbanken oder Dateisystemen sein.

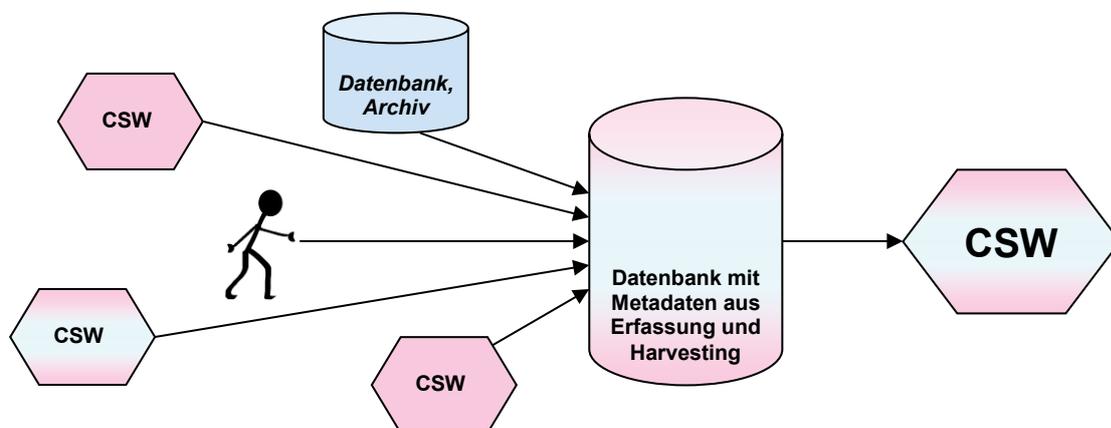


Abb.: Katalog mit sowohl direkt erfassten Metadaten als auch geharvesteten Metadaten (Gemischter Katalog)

Kataloge mit Metadaten gemischter Herkunft dienen häufig dazu, bestimmte Systeme mit Datenbeschreibungen zu bestücken. Denkbar wäre beispielsweise ein übergreifendes Shop-System, welches befüllt wird oder eine Auskunftsplattform für ein bestimmtes Fachpublikum, welches zusätzliche spezielle Anforderungen an das Auskunftssystem hat. Um den speziellen Anforderungen gerecht zu werden, werden weitere Metadaten eingespeist, die für ein Publikum außerhalb der in sich geschlossenen Institution oder des in sich geschlossenen Arbeitsprozesses unwichtig sind oder ihm ein „zu viel“ an Information böten.

Betreiben Sie einen solchen gemischten Katalog, achten Sie darauf, dass über die CSW-Schnittstelle, so diese öffentlich zur Verfügung steht, direkt erfasste Metadaten und geharvestete Metadaten aus Fremdkatalogen separat angefragt werden können. Da diese Handhabung nach dem derzeitigen allgemeinen Entwicklungsstand in der Umsetzung eher schwierig ist, empfiehlt es sich, für die direkt erfassten Metadaten eine eigene CSW-Schnittstelle anzubieten, um Harvesting-Daten Dritter und originäre Metadaten besser voneinander unterscheiden zu können.

6. Die Bereitstellung von Metadaten im Sinne von INSPIRE

Bei der Bereitstellung von Metadaten im Sinne von INSPIRE ist eine enge Zusammenarbeit zwischen dem Erfasser für Metadaten, dem Katalogbetreiber und der Koordinierungsstelle GDI-NI zwingend erforderlich. Erkundigen Sie sich bitte direkt bei uns nach den aktuellen Anforderungen, sofern Sie Datenhalter im Sinne von INSPIRE sind oder zu sein glauben.

Als einziger Katalogdienst in Deutschland muss der CSW aus dem Geodatenkatalog.de der Koordinierungsstelle GDI-DE alle Anforderungen an einen Suchdienst im Sinne von INSPIRE erfüllen.

Für die Kataloge der GDI-NI ist es ausreichend, wenn die enthaltenen Metadaten die relevanten ISO-Standards einhalten und entweder die CSW-Schnittstelle von der Koordinierungsstelle GDI-NI in gleichbleibenden Abständen geharvestet oder aber die dateibasierten Metadaten in die Geodatenuche Niedersachsen eingelesen werden können.

Über den CSW der Geodatenuche Niedersachsen gelangen sämtliche Metadaten der GDI-NI zum Geodatenkatalog.de. Metadaten, die zusätzlich mit dem Schlagwort „inspireidentifiziert“ gekennzeichnet wurden, gelangen von dort aus in das INSPIRE Portal (Discovery).

0. Anhang – Begriffe im Zusammenhang mit Metadaten

CSW – Catalogue Service Web (Katalogschnittstelle)

Bei einer CSW-Schnittstelle handelt es sich um eine standardisiert abfragbare Schnittstelle, die Datenbeschreibungen (Metadaten) abgibt. Die Anfrage wird wie auch bei anderen OGC-Diensten „Request“, die Antwort „Response“ genannt. Eine Katalogschnittstelle kann auf Anfrage benennen, welche Fähigkeiten sie besitzt (GetCapabilities), über wie viele Datenbeschreibungen sie verfügt (DescribeRecords) und es können neben weiteren Möglichkeiten gezielt bestimmte Datenbeschreibungen unter Angabe des Metadatensatz-Identifikators abgerufen werden (GetRecordById). Alles in allem ist die Kommunikation mit einer CSW-Schnittstelle um ein Vielfaches komplexer als beispielsweise die Kommunikation mit einem WMS (WebMapService), der Kartenbilder und Sachinformationen ausliefern kann. Aus diesem Grunde ist es sinnvoll, für die Kommunikation mit einem CSW eine auf die eigenen Bedürfnisse angepasste z. B. webbasierte Suchoberfläche in einem Browser zu verwenden.

Harvesting

Harvesting wird betrieben, damit Datenbeschreibungen von Geodatenressourcen in Suchoberflächen in ihrem originalen Wortlaut möglichst schnell zur Verfügung stehen.

Im Rahmen des Harvesting werden Geodatenbeschreibungen (Metadatensätze) von einem Metadatenkatalog in einen anderen gezogen. Das Original verbleibt im Original-Katalog und wird dort aktualisiert. Durch den Metadatensatz-Identifikator (UUID) und das hinterlegte Datum der letzten Änderung am Metadatensatz ist sichergestellt, dass der Datensatz nicht nur im Original-Katalog sondern auch – zeitversetzt – im harvestenden Katalog aktualisiert wird. Metadaten sind damit die einzigen Daten in einer Geodateninfrastruktur, die gezielt kopiert werden, um den Nutzern zur Verfügung zu stehen.

Durch ein Harvesting dürfen Metadaten eines fremden Katalogs niemals abgeändert werden.

Geodatenressourcen

Geodatenressourcen werden alle Daten genannt, die man innerhalb einer Geodateninfrastruktur als Daten mit räumlichem Bezug (Geodaten) sieht. Üblicherweise unterscheidet man zwischen Geodaten (z. B. Shape- oder Tiff-Dateien), Diensten (z. B. WMS, WFS, CSW) und Fachinformationssystemen (z. B. Viewer für ein spezielles Fachthema). Es können aber auch Tabellen oder Listen mit geographischem Bezug Merkmale aufweisen, die sie für die Nutzung in einer Geodateninfrastruktur qualifizieren. Um allen möglichen Datenquellen gerecht zu werden, spricht man im Fachjargon von „Geodatenressourcen“. Gleichzeitig zeigt dieser Begriff, dass Geodaten in Ihren unterschiedliche Ausprägungen durchaus als „Rohstoff“ zu sehen sind, die es gilt weiter zu nutzen, zu verarbeiten oder zu veredeln.

Suchende

Der Suchende ist der, der Geodatenressourcen ausfindig machen möchte. Hierfür verwendet er im Idealfall eine Metadatenuche, die mittels einer Suchoberfläche realisiert wurde. Der Grund für eine Suche nach Geodatenressourcen kann vollkommen unterschiedlich motiviert sein. Der Suchende kann aufgrund seines Berufs recherchieren, als Bürger oder aus einem persönlich motivierten Interesse. Ebenso unterschiedlich wie seine Motivation können seine genauen Beweggründe sein, die von einer ersten grundlegenden Information hin bis zu einer detaillierten Nachverfolgung eines sehr speziellen Sachverhaltes reichen können.

Nicht für jeden Suchenden ist jede Such- / Präsentationsoberfläche gleichermaßen geeignet. Er wird diejenige Oberfläche wählen, die ihm gestalterisch, inhaltlich und sprachlich am nächsten kommt.

Such- und Präsentationsoberfläche

Unter Suchoberfläche wird in diesem Leitfaden eine Anwendung verstanden, die die technischen Requests (Anfragen) an die CSW-Schnittstelle schickt und die Responses (Antworten) in Empfang nimmt. Einerseits kann ein Suchender über die Suchoberfläche bestimmte Dinge auswählen (z. B. Schlagwörter) oder eingeben (Suchbegriffe) und die Suche auslösen. Die Suche schickt daraufhin die technische Anfrage an den oder die angeschlossenen CSW-Schnittstellen und erhält eine Antwort. Im Idealfall sind in der Antwort diejenigen Datenbeschreibungen aufgelistet, die der Suchanfrage entsprechen. Diese werden dann dem Nutzer über eine Präsentationsoberfläche in angemessener Form angeboten. Üblich ist, dass eine solche Präsentationsoberfläche beispielsweise URLs als anklickbare Hyperlinks abbildet, weil die meisten Präsentationsoberflächen Browseranwendungen sind. Denkbar ist aber auch eine Such- und Präsentationsoberfläche, die z. B. direkt in einem Geoinformationssystem implementiert ist.

Such- und Präsentationsoberflächen richten sich stets nach den Bedürfnissen der jeweiligen Zielgruppe unter den Suchenden.

Weitere Begriffe werden im Geodatenportal Niedersachsen erklärt. Wählen Sie dort bitte den Menüpunkt „GDI-Standards / Glossar“.