CR-Formular

|  |  |  |
| --- | --- | --- |
| **Pos.** | **Name** | **Inhalt** |
| 1 | Titel | Qualitätssicherungswerkzeug für die GDI-DE |
| 2 | lfd. Nr. und/ oder Az. | 01/2021 |
| 3 | Bedarfsträger mit Kontaktadresse | Koordinierungsstelle GDI-DE (nach Umfrage zu Bedarfen der Kontaktstellen GDI-DE bezüglich der Unterstützung bei der Qualitätssicherung)  Koordinierungsstelle GDI-DE  Bundesamt für Kartographie und Geodäsie  Richard-Strauss-Allee 11  60598 Frankfurt am Main  mail@gdi-de.org |
| 4 | Eingangsdatum der  Bedarfsanforderung | 15.10.2020 |
| 5 | Betroffene Komponente | GDI-DE Registry Zusatzmodul „INSPIRE-Monitoring-Client“ |
| 6 | Zuordnung zu NGIS-Zielen,  ggf. Erläuterung | Ziel 2.3 „Gemeinsam abgestimmte Standards für Geoinformationen geben einen fach- und organisationsübergreifenden Qualitätsanspruch vor, der nach definierten Qualitätssicherungsmechanismen überprüft wird. Nutzer können sich über die Datenqualität informieren.“ |
| 7 | Priorität (gering/mittel/hoch), ggf. Erläuterung | Hoch, da sich neue Anforderungen an die Qualitätssicherung ergeben haben, der „INSPIRE-Monitoring-Client“ nicht mehr dem aktuellen Verfahren des INSPIRE-Monitoring der Europäischen Kommission entspricht und der „INSPIRE-Monitoring-Client“ zur Zeit nur auf einem System betrieben werden kann, das bald keine Sicherheitsupdates mehr erhält. |
| 8 | Erwartetes Ergebnis | * Bereitstellung eines Werkzeugs zur Erfüllung der Anforderungen, die u.a. bei der Umfrage zum Bedarf der Kontaktstellen GDI-DE bezüglich der Unterstützung bei der Qualitätssicherung ermittelt wurden (Anforderungsdetails: siehe 11) * Abschaltung des Zusatzmoduls „INSPIRE-Monitoring-Client“ der GDI-DE Registry (bei Entscheidung für Umsetzungsalternative a) |
| 9 | Geschätzte Umsetzungsdauer, Meilensteine | Umsetzungsalternative a)   1. Projektinitialisierung (1 Monat) 2. Anforderungsanalyse und Erstellung Lastenheft (1 Monat) 3. Ausschreibung (3 Monate) 4. Implementierung (3 Monate) 5. Test- und Abnahmephase (2 Monate)   **Gesamtdauer**: 10 Monate  Umsetzungsalternative b)   1. Erstellung Lastenheft (1 Monat) 2. Ausschreibung (3 Monate) 3. Anbindung an die aktuelle Version der GDI-DE Testsuite implementieren (1 Monat) 4. Verbesserung der Performance der aktuellen Implementierung des INSPIRE-Monitoring-Client (3 Monate) 5. Implementierung der für die neuen Anforderungen erforderlichen Funktionen (4 Monate) 6. Test- und Abnahmephase (2 Monate)   **Gesamtdauer**: 14 Monate |
| 10 | Geschätzter Aufwand | **Zeitlicher Aufwand:** siehe 9.  **Finanzieller Aufwand:**   * Umsetzungsalternative a:   + Intern: 50.000€   + Extern: 80.000€ * Umsetzungsalternative b:   + Intern: 30.000€   + Extern: 210.000€   Zusätzlich fallen nach der Inbetriebnahme der neuen oder weiterentwickelten Software fortlaufende Kosten für Wartungsaufwände in Höhe von jeweils ca. 30.000€ jährlich (Variante a und b) sowie Kosten für den Betrieb der erforderlichen IT-Infrastruktur (Server, Festspeicher, Netzwerktechnik) und Personalkosten an. Diese Kosten sind stark abhängig vom Umfang der Nutzung und den Einsatzfeldern sowie der Architektur der angestrebten Lösung. Die Kosten können erst auf Grundlage der Anforderungsanalyse (s.o.) genauer beziffert werden. |
| 11 | Beschreibung einschl.  Begründung | Der CR beruht auf den gewünschten Anforderungen, die u.a. bei der Umfrage zum Bedarf der Kontaktstellen GDI-DE erfasst und im GDI-DE Wiki zusammengetragen wurden:  <https://wiki.gdi-de.org/x/Q4AqLQ>  Eine Erläuterung der Kst. GDI-DE zur Bedarfserhebung ist als Anlage beigefügt.  **1. Anforderungen Auswertung INSPIRE-Monitoring:**  Prioritäten (gemäß Umfrage)   * M - MUST HAVE (unbedingt erforderlich) * S - SHOULD HAVE (sollte umgesetzt werden, ist aber nicht zwingend erforderlich) * C - COULD HAVE (kann umgesetzt werden, Nice-to-have)  1. Metadaten über Katalogschnittstelle importieren (M)    * vorkonfigurierte Katalogschnittstelle aus Liste auswählen oder manuelle Eingabe der URL des Dienstes    * Filter Encoding eingeben (zur selektiven Abfrage der Katalogschnittstelle, um nur Metadatensätze zu importieren, die das Keyword "inspireidentifiziert" beinhalten    * **Erwartetes Ergebnis:** Die INSPIRE-identifizierten Metadatensätze sind importiert und können für das INSPIRE-Monitoring ausgewertet werden.    * **Alternativ** kann auch der komplette Geodatenkatalog.de in regelmäßigen Abständen importiert und ausgewertet werden. In einer Ergebnisansicht müssen dann Möglichkeiten zur Filterung nach Bundesland, Daten des Bundes, Kommune und geodatenhaltender Stelle vorhanden sein. 2. Ergebnisse des Metadaten-Imports auswerten (M)    * Liste der importierten Metadaten zu Datensätzen (FID, Titel, Organisation, E-Mail-Adresse, Regionalschlüssel, INSPIRE-Themen, Konformität, FID gekoppelter Darstellungsdienst, FID gekoppelter Downloaddienst) = "Rohdaten", auf deren Basis die Berechnung der Monitoring-Indikatoren erfolgt    * Liste der importierten Metadaten zu Diensten (FID, Titel, Organisation, E-Mail-Adresse, Diensttyp, Konformität) = "Rohdaten", auf deren Basis die Berechnung der Monitoring-Indikatoren erfolgt    * Änderungen gegenüber dem vorherigen Import auflisten und menschenlesbar darstellen (neue, fehlende und geänderte Metadatensätze)    * Ergebnisse des Metadaten-Imports sind in Excel (XLSX, CSV) exportierbar    * online Übersicht der Ergebnisse. Zudem können die Ergebnisse in weitere Webseiten (z.B. durch HTM-Export oder IFrames) integriert werden    * **Erwartetes Ergebnis:** Die Änderungen gegenüber dem letzten Import sind für den Nutzer erkennbar und auswertbar. 3. Berechnung der Indikatorwerte für das INSPIRE-Monitoring gemäß der von der EU verwendeten Methodik für eine bestimmte Auswahl an Ressourcen (Ressourcen des Bundes, Ressourcen eines Landes, Ressourcen einer Kommune, Ressourcen einer geodatenhaltenden Stelle) (M)    * Ressourcen auswählen, für die die Indikatorwerte berechnet werden sollen (siehe Anforderung #1)    * Indikatoren für Auswahl berechnen auf Basis der von der EU verwendeten Methodik    * Konformitätsindikatoren Metadaten auf Basis der Durchführung entsprechender Konformitätstests berechnen    * **Erwartetes Ergebnis:** Die Indikatoren des INSPIRE-Monitoring können für eine bestimmte Auswahl an Ressourcen berechnet werden. 4. Die Ergebnisse des INSPIRE-Monitoring (Liste mit Ressourcen und Indikatoren) für alle Ressourcen aus Deutschland oder für eine bestimmte Auswahl an Ressourcen sind für die weitere Verwendung/ Auswertung in verschiedene Formate exportierbar: (S)    * Excel (XLSX)    * In einem Format zur Integration in Webseiten (z.B. HTML)    * **Erwartetes Ergebnis:** Die Ergebnisse des INSPIRE-Monitoring (Liste mit Ressourcen und Indikatoren) sind in verschiedene Formate exportierbar. 5. Für die Fehleranalyse werden von der EU (JRC) Fehlerlisten mit den Ergebnissen der Validierung der Metadatensätze herausgegeben ("Summary of metadata failing validation"). Zusätzliche Anforderungen: (M/C)    * Fehlerlisten nach Bund/Land filtern (für die Weitergabe an die Kontaktstellen GDI-DE) - M    * Fehlerlisten nach geodatenhaltender Stelle filtern (für die Weitergabe an die geodatenhaltenden Stellen) - M    * Fehlerlisten selbst erzeugen (auf Basis der Ergebnisse aus Anforderung #1) – C    * **Erwartetes Ergebnis:** Die Fehlerlisten der EU mit den Ergebnissen Validierung der Metadatensätze sind nach Bund/Land bzw. geodatenhaltender Stelle filterbar. 6. Analyse der Monitoring-Ergebnisse: automatische Ableitung von Diagrammen, Grafiken (z.B. Karten) und Tabellen für alle erfassten Monitoring-Jahre (M)    * Zeitpunkt/Zeitraum auswählen, auf die sich die Analyse beziehen soll (nach Jahr(en) filtern, historische Entwicklung)    * Ressourcen auswählen, auf die sich die Analysen beziehen sollen (nach Ressourcen filtern, z.B. Bund, Land, Kommune, geodatenhaltende Stelle)    * Folgende Analysen müssen mindestens umgesetzt werden (wenn automatisiert auf Basis der Metadateninhalte umsetzbar):      + Anzahl der Datensätze      + Anzahl der Anhang I-Datensätze, Anhang II-Datensätze, Anhang III-Datensätze      + Anzahl der Datensätze differenziert nach Themen      + Anzahl der Datensätze differenziert nach Sektor (Verwaltung, Wirtschaft, Wissenschaft)      + Anzahl der Datensätze differenziert nach Verwaltungsebene (Bund, Land, Kommune)      + Anzahl der Open-Data-Datensätze      + Anzahl der Dienste      + Anzahl der Dienste differenziert nach Diensttyp      + Anzahl der geodatenhaltenden Stellen      + Anzahl der Kommunen      + Anzahl der geodatenhaltenden Stellen aus Wirtschaft / Wissenschaft      + Indikatoren-Werte    * Nutzer kann weitere Analysen konfigurieren und anzeigen lassen    * die Diagramme, Grafiken und Tabellen sind exportierbar (PDF und PNG bei Grafiken/Diagrammen, PDF und XLSX bei Tabellen)    * **Erwartetes Ergebnis:** Es sind konfigurierbare und filterbare Analysen der Monitoring-Ergebnisse in Form von Diagrammen, Grafiken und Tabellen verfügbar. 7. Auswertung der Flächendeckung (C)    * auf Basis der in den INSPIRE-Anwendungs-Schemata definierten "Feature Types"    * auf Basis der "Betroffenheitsmatrix" (Berücksichtigung von Vereinbarungen zur Bereitstellung und Berücksichtigung der Einschätzung der Betroffenheit)    * visualisiert in Form von Karten    * **Erwartetes Ergebnis:** Es sind Karten zur Flächendeckung der INSPIRE-Feature Types verfügbar.   **2. Anforderungen Qualitätssicherung allgemein:**   1. Auflistung der Zeitpunkte des Harvestings in den Geodatenkatalog.de + letztes erfolgreiches Harvesting, Harvesting-Intervall, Anzahl geharvesteter Daten, Anzahl Veränderung im Vergleich zum letzten Harvesting (Hinzufügungen, Löschungen, Aktualisierungen) 2. Prüfung der Metadaten im Geodatenkatalog.de mit der GDI-DE Testsuite 3. Auflistung valider und invalider Daten 4. Darstellung der Anzahl valider und invalider Daten beim letzten Harvesting (aufgeschlüsselt nach geharvesteten Katalogen, Datenbereitstellern) 5. Ableitung von Indikatoren zur Bewertung der Konformität, Qualität und Vollständigkeit der Metadaten im Geodatenkatalog.de bzw. in der GDI-DE.   Die Anforderungen sind zur fortlaufenden Kontrolle und Validierung des Anschlusses von externen Katalogen (Metadatenkataloge der Länder, des Bundes und der Wissenschaft und Forschung) an den Geodatenkatalog.de immens wichtig. Daher sollten diese, wenn möglich, alle umgesetzt werden.  **3. Benutzerverwaltung/Nutzerkonten**  Zur Anzeige, Konfiguration und Speicherung von spezifischen Auswertungen sind Nutzerkonten für Katalogbetreiber, geodatenhaltende Stellen und weitere Akteure (z.B. Kst. GDI-DE, Kontaktstellen, AK der GDI-DE) bereitzustellen.  **Umsetzungsalternativen:**   1. Neu- bzw. Weiterentwicklung eines Werkzeugs, das die gesamten Anforderungen abdeckt auf Basis des Qualitätssicherungswerkzeugs, das zurzeit vom Kompetenzzentrum Geodateninfrastruktur Baden-Württemberg entwickelt wird. Das Ergebnis ist eine von der GDI-DE Registry und anderen Komponenten losgelöste Anwendung. Abschaltung des Zusatzmoduls „INSPIRE-Monitoring-Client“ der GDI-DE Registry. Das Kompetenzzentrum Geodateninfrastruktur Baden-Württemberg hat angeboten, seine bisherigen Entwicklungen zur Verfügung zu stellen und das Qualitätssicherungswerkzeug in einem gemeinsamen Projekt zwischen BW, Betrieb GDI-DE und Kst. GDI-DE weiterzuentwickeln. Zudem soll das Projekt weiteren Partnern aus der GDI-DE offenstehen. 2. Erweiterung, Anpassung, Korrektur und Optimierung des Zusatzmoduls „INSPIRE-Monitoring-Client“ der GDI-DE Registry.   Zur Begründung der Abschaltung des Zusatzmoduls „INSPIRE-Monitoring-Client“ der GDI-DE Registry und den Vorteilen einer nicht in die GDI-DE Registry integrierten Anwendung zur Erfüllung der o.g. Anforderungen siehe den Block zur Umsetzungsalternative b in Abschnitt 13.  Der Aufwand zur Abschaltung des Zusatzmoduls „INSPIRE-Monitoring-Client“ der GDI-DE Registry ist voraussichtlich gering, da die übrigen Funktionen der GDI-DE Registry ohne den „INSPIRE-Monitoring-Client“ unverändert weiter genutzt werden können. Es gibt nur eine Abhängigkeit vom „INSPIRE-Monitoring-Client“ zur GDI-DE Registry, nicht auch in die andere Richtung.  **Umsetzungsalternativen im Detail:**   1. Bei der Wahl dieser Umsetzungsalternative ist eine Anwendung zu implementieren, die die o.g. Anforderungen bestmöglich erfüllt. Als Basis dieser Anwendung soll das Qualitätssicherungswerkzeug dienen, das zurzeit vom Kompetenzzentrum Geodateninfrastruktur Baden-Württemberg entwickelt wird und bereits viele der gewünschten Anforderungen abdeckt. Um Nutzern die vom „INSPIRE-Monitoring-Client“ zur Analyse großer Datenmengen benötigte Verarbeitungszeit zu ersparen und Auswertungsergebnisse möglichst schnell zu liefern, sollen sämtliche Daten aus dem Geodatenkatalog.de kontinuierlich automatisch analysiert werden (Anforderung 1.1). Die aktuellen Analyseergebnisse sowie historische Analysedaten (Anforderung 1.6) werden Nutzern in einer Webanwendung auf Basis von Kibana bereitgestellt. In der im Rahmen dieses CRs beschriebenen, mit der Koordinierungsstelle GDI-DE abgestimmten, ersten Entwicklungsstufe des Werkzeugs soll sich die Datenanalyse auf den Geodatenkatalog.de beschränken. Durch die Möglichkeit, die Rohdaten der Analyse in Kibana zu filtern und zu segmentieren, werden die Anforderungen 1.1, 1.5 und 1.6 erfüllt. Ergänzend dazu können in GeoNetwork virtuelle CSW Entry Points (Katalog-IDs) verwendet werden, um den Katalog zu segmentieren. Um nutzerspezifische Analysen (Anforderungen 1.1, 1.5, und 1.6) anzuzeigen, wird ein Nutzerkontenmanagement benötigt. Hierzu kann auf die in Kibana integrierte Nutzerverwaltung zurückgegriffen werden. Zur Analyse der aus dem Geodatenkatalog.de bezogenen Daten soll u.a. eine lokale Instanz des INSPIRE Validators bzw. der GDI-DE Testsuite verwendet werden (Anforderungen 1.2 und 1.3). Die Anbindung an den INSPIRE Validator des JRC ist im Qualitätssicherungswerkzeug aus Baden-Württemberg bereits vorhanden. Die Anbindung an eine lokale Instanz des INSPIRE Validators bzw. die GDI-DE Testsuite ist im Rahmen dieses CRs zu implementieren, um nicht von der Verfügbarkeit und mäßigen Geschwindigkeit der INSPIRE Validator Instanz des JRC abhängig zu sein. Ebenfalls ist zur Sicherstellung einer hohen Performance und Zuverlässigkeit die Anbindung an eine lokale Kopie der INSPIRE Registry umzusetzen. Der Export von Analysen, Diagrammen und Grafiken kann vorwiegend über Kibana erfolgen (Anforderung 1.4). In der Feinkonzeption zur Entwicklung der Software soll geprüft werden, welche Anpassungen am Qualitätssicherungswerkzeug aus Baden-Württemberg zur konkreten Erfüllung der o.g. Anforderungen noch erforderlich sind. Die Umsetzbarkeit der optionalen Anforderung 1.7 (Auswertung der Flächendeckung) ist nicht gesichert aber auch nicht zwingend erforderlich. Das Qualitätssicherungswerkzeug soll zentral vom Betrieb GDI-DE bereitgestellt werden. Länderspezifische Auswertungen sind – falls gewünscht – auf Basis der o.g. Möglichkeiten zur Datenfilterung und Nutzerverwaltung möglich. 2. Soll der „INSPIRE-Monitoring-Client“ weiter verwendet und ausgebaut werden, ist eine Anbindung an die aktuelle GDI-DE Testsuite zu implementieren, da der „INSPIRE-Monitoring-Client“ zur Zeit nur mit der veralteten Version der GDI-DE Testsuite zusammenarbeiten kann. Parallel dazu müssen die Funktionen zur Erfüllung der neuen Anforderungen implementiert werden. Zudem muss die Verarbeitungsgeschwindigkeit des „INSPIRE-Monitoring-Client“ deutlich erhöht werden, um bei der gestiegenen Anzahl auszuführender Prüfungen eine annehmbare Verarbeitungsgeschwindigkeit für alle Nutzer zu erreichen. Bei allen Änderungen ist die Kompatibilität des „INSPIRE-Monitoring-Client“ mit dem übrigen Quellcode der GDI-DE Registry sicherzustellen, um ihren Betrieb nicht zu gefährden. |
| 12 | WiBe | Als Anlage beigefügt |
| 13 | Umsetzungsempfehlung | Empfohlen wird **Umsetzungsalternative a**, da so sichergestellt werden kann, dass alle wichtigen Anforderungen vollständig erfüllt werden und bereits existierende Entwicklungen in der GDI-DE nachgenutzt werden können. In einem gemeinsamen Projekt mit der GDI-BW wird das Qualitätssicherungswerkzeug aus BW weiterentwickelt, um die Anforderungen und Bedürfnisse der GDI-DE vollständig erfüllen zu können. Hierzu ist zum Projektstart zu analysieren welcher Anpassungs- und Erweiterungsbedarf besteht. Hiernach kann ein detaillierter Projektplan mit einzelnen Schritten und Zeitplan erarbeitet werden. Angestrebt wird eine sukzessive Weiterentwicklung des QS-Werkzeugs um Funktionalität, sodass schon relativ früh erste Lösungen zur Verfügung gestellt werden können. Der Betrieb des QS-Werkzeugs wird zentral vom Betrieb GDI-DE übernommen. Hierzu muss der Leistungskatalog der NTK der GDI-DE in Absprache mit dem AK Architektur entsprechend angepasst werden. Eine Beteiligung weiterer GDI-DE Partner im Rahmen der Projektorganisation ist vorgesehen.  Die Weiterverwendung des „INSPIRE-Monitoring-Client“ (**Alternative b**), der für das INSPIRE Monitoring bis zum Berichtsjahr 2018 genutzt wurde, lässt einen hohen Anpassungsaufwand der Software erwarten. Dabei ist die Verzahnung mit der GDI-DE Registry hinderlich und bietet Potenzial für eine hohe Fehlerrate sowohl im „INSPIRE-Monitoring-Client“ als auch in der GDI-DE Registry selbst. Da eine zuverlässige Verfügbarkeit der GDI-DE Registry essentiell für z.B. die Daten-Dienste-Kopplung ist, wird davon abgeraten, sie durch Werkzeuge, die nichts mit ihren Kernfunktionen zu tun haben, komplexer als nötig zu machen.  Der „INSPIRE-Monitoring-Client“ stellt hohe Anforderungen an die verwendete Hardware, die bei zusätzlichen Prüfungen innerhalb der Anwendung zur Erfüllung der gewünschten Anforderungen noch steigen werden und damit ggf. die bereitgestellten Kapazitäten überfordern. Eine weitere vertikale Skalierung (mehr Prozessoren und Arbeitsspeicher) ist mit der vorhandenen IT-Infrastruktur nicht ohne weiteres möglich. Eine horizontale Skalierung (zusätzliche Server) wird von der Anwendung aktuell nicht unterstützt.  Ein weiterer Hinderungsgrund für die Weiterentwicklung des „INSPIRE-Monitoring-Client“ ist die Anbindung an die GDI-DE Testsuite. Der „INSPIRE-Monitoring-Client“ kann aktuell nur mit der bereits stillgelegten veralteten Version der GDI-DE Testsuite zusammenarbeiten, um manche Prüfungen auszuführen. Bevor der „INSPIRE-Monitoring-Client“ weiter verwendet werden kann, müsste daher zunächst eine Anbindung an die aktuelle Version der GDI-DE Testsuite implementiert werden. |
| 14 | Anlagen | * Erläuterung der Kst. GDI-DE zur Bedarfserhebung vom 01.02.2021 * WiBe: Entwicklung einer Lösung zur Auswertung des INSPIRE-Monitoring vom 29.03.2021 |