

Universeller Gateway-Ansatz auf der Basis semantisch angereicherter Web Services im Rahmen heterogener eGovernment-Anwendungen

Stefan Audersch
Zentrum für Graphische Datenverarbeitung e.V., Rostock
Joachim-Jungius-Str. 11, 18059 Rostock
stefan.audersch@rostock.zgdv.de

Abstract: Die Integration bzw. Föderation von heterogenen Informationssystemen stellt eine große Herausforderung dar. Es gibt verschiedene Ansätze, welche versuchen, die Heterogenität einzelner Systeme mittels automatischer oder semiautomatischer Verfahren aufzuheben.

Dieser Beitrag stellt einen Ansatz in Form eines universellen Gateways vor, der es ermöglicht, heterogene Systeme miteinander zu kombinieren. Betrachtete Schwerpunkte sind dabei sowohl die syntaktische als auch die semantische Integrität der Daten. Im vorgestellten Ansatz werden Web Service Technologien und semantischen Konzepten des Semantic Web miteinander verknüpft. Diese sind zusätzlich um Sichtenkonzepte erweitert, um so einen dynamischen Zugriff auf Dienste zu ermöglichen. Der um eine semantische Ebene erweiterte Web Service bietet durch die maschinenverständliche Semantik der integrierten externen Dienste eine Voraussetzung für die Dienstkomposition.

Die Ansätze werden am Beispiel des Melderegisterwesens aus dem Bereich eGovernment vorgestellt und wurden weitgehend prototypisch in eFormsConnect umgesetzt.

1 Einleitung

In der Vergangenheit und auch heute noch bestimmen anstelle einer homogenen Informationslandschaft verschiedenste Inselösungen das Bild der Datenhaltung. Aus Gründen wie beispielsweise wachsender Globalisierung der Wirtschaft und der damit verbundenen übergreifenden Kommunikation von IT-Systemen oder des Zusammenschlusses verschiedener IT-Landschaften entsteht die Notwendigkeit eines homogenen Datenzugangs zu unterschiedlichen heterogenen Informationssystemen. Dazu zählt nicht nur die technische Zusammenführung, sondern auch die inhaltliche Integration der Daten.

Da es oft nicht möglich ist, alle Daten in ein einheitliches System zu integrieren, konzentriert sich die Entwicklung auf Konzepte und Systeme, die auf Basis existierender Altsysteme eine föderative Integration von Informationen ermöglichen.

Web Services gewinnen in diesem Zusammenhang immer mehr an Bedeutung. Sie stellen eine flexible, service-orientierte Architektur dar, mit deren Hilfe die Grundlage für eine lose Anbindung bzw. Integration von Anwendungen und Informationssystemen gegeben ist. Um eine inhaltliche Äquivalenz zwischen den verschiedenen Daten zu schaffen, ist es notwendig, das Konzept der Web Services um eine zusätzliche semantische Ebene zu erweitern. Metadatengestützte Standards dienen der semantischen Beschreibung von Informationsquellen. Sie bilden eine mögliche Grundlage zur Konzeption eines universellen Ansatzes zur Schaffung intelligenter Web Services, die semantisch gesteuert die Integrations- und Transaktionsprozesse im Sinne dynamischer Services realisieren.

Im Rahmen des Projektes Bürger-Kontakt-System Mecklenburg-Vorpommern wird eine eGovernment Plattform entwickelt, die relevante Dienstangebote der Verwaltung für Bürger und Wirtschaft abbildet und zur Nutzung online zur Verfügung stellt. Ziel ist dabei unter anderem die Anbindung dieser Dienstangebote an die zumeist sehr heterogenen Fachverfahren im Backoffice-Bereich der Verwaltungen.

Ziel dieser Arbeit ist die Entwicklung eines Ansatzes zur Kopplung von eFormsDirect [AC+03], ein eGovernment-Framework für intelligente Formulare, mit verschiedenen existierenden heterogenen

Fachverfahren. Die hier vorgestellte Gateway-Lösung stellt gleichzeitig auch einen universellen Ansatz für andere Anwendungsfelder (z.B. eLearning oder Knowledge Management) dar.

2 Problemstellung und Anforderungen

eGovernment-Technologien bieten der Verwaltung neue Optionen zur Steigerung der Effizienz, zur Reduzierung von Kosten und damit zur Entlastung der Haushalte. Durch Automatisierung von Prozessen können beispielsweise Dienstleistungen schneller und unter geringerem Aufwand verfügbar gemacht werden. Deren Umsetzung ist Ziel des Projektes BKS M-V. Das im Rahmen dieses Projektes vom ZGDV e.V. Rostock entwickelte eFormsDirect ermöglicht dabei die einfache, flexible Bereitstellung sowie Nutzung von Online-Formularen.

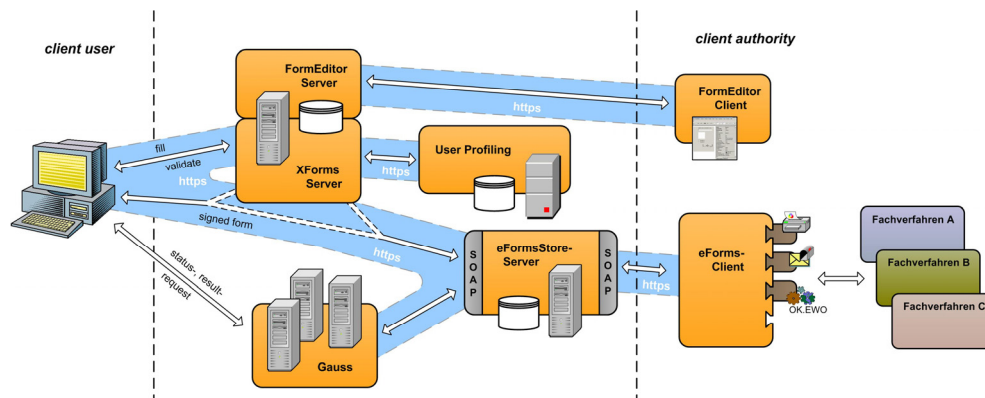


Abbildung 1: eFormsDirect Architektur

Die Kommunikation innerhalb des Frameworks basiert auf OSCI¹ (Online Service Computer Interface), ein speziell für den Bereich eGovernment entwickelter Protokollstandard (OSCI-Transport 1.2). Neben einer Spezifikation zur Datenübertragung sind in OSCI die Strukturierung der Nachrichteninhalte (OSCI-B z.B. XMeld 1.1 für den Bereich Meldewesen) definiert.

Ein echter Mehrwert und damit ein Beitrag zur Effizienzsteigerung und Kostensenkung entstehen jedoch erst durch eine Zusammenarbeit des Formularsystems mit anderen Systemen der Behörde, anderen Behörden oder auch der Privatwirtschaft. Problematisch ist dabei die Vielzahl verschiedener Systeme für die unterschiedlichsten Fachverfahren, die sich sowohl syntaktisch, logisch als auch strukturell heterogen gestalten.

Ziel ist es nun, die im Verwaltungsbereich existierenden Systeme mit heterogenen Schnittstellen innerhalb einer komplexen Gesamtlösung zu integrieren. Zudem soll die Anwendung von Kontexten möglich sein, die die Nutzung integrierter Fachverfahren zum Beispiel hinsichtlich Komplexität oder Nutzerrechte vereinfachen. Ein gewählter Kontext spiegelt dabei eine bestimmte Sicht auf die Datenstrukturen wieder.

Aus diesen Überlegungen heraus ergibt sich die Idee eines universellen Gateways zur Kommunikation und zum Austausch von Daten mit unterschiedlichen Fachverfahren, wobei eine enge Kopplung des Web Service Interfaces mit RDF Content Repository Ansätzen verfolgt wird.

3 Realisierungsaspekte

Als Ansatz für die Gateway-Lösung bietet es sich an, die verschiedenen Aspekte der Heterogenität in einzelnen Stufen zu betrachten. In dem folgendem Abschnitt wird zunächst auf die syntaktische Heterogenität eingegangen.

3.1 Syntaktische Heterogenität

Hinsichtlich der Integration von Diensten und der Überwindung syntaktischer Heterogenität sind in verschiedenen Projekten Lösungsansätze erarbeitet worden.

¹ OSCI ist ein MEDIA@Komm-Ergebnis. Der Standardisierungs-Prozess wird derzeit durch die OSCI-Leitstelle Bremen im Auftrag der KoopA-ADV (Kooperationsausschuss ADV Bund/Länder/kommunaler Bereich) fortgeführt.

Im Projekt *Service Globe* [KS+03] besteht die Idee darin, verschiedene Dienste dynamisch auswählen zu können, wobei zwischen internen (von Service Globe bereitgestellt) und externen (integrierte) Diensten unterschieden wird. Eine Verknüpfung externer Dienste erfolgt mittels Adaptoren, die jeweils für den Dienst entwickelt werden.

IBM ermöglicht mit dem *Web Services Invocation Framework* [KK+03] die Unterstützung verschiedener Service Bindings und Typsysteme. So können neben SOAP beispielsweise auch Java-Klassen oder Stored Procedures durch das Framework integriert werden. Zur Integration werden die Dienste mit WSDL entsprechend beschrieben.

Im hier vorgestellten Lösungsansatz werden für die Überwindung der syntaktischen Heterogenität Treiber verwendet. Ein Treiber erlaubt die Kommunikation mit Diensten, die über eine definierte Schnittstelle angesprochen werden bzw. ein spezielles Protokoll verwenden. So können beispielsweise Treiber für SOAP-, native Java- oder auch JDBC-Schnittstellen umgesetzt werden. Der eigentliche Dienst kann mit Hilfe von WSDL beschrieben werden. In der hier vorgestellten Lösung wird XML als einheitliches Eingabe- als auch Ausgabeformat verwendet. Die Umsetzung der XML-Dokumentstrukturen auf die entsprechende Schnittstelle ist Aufgabe des Treibers.

3.2 Logische und strukturelle Heterogenität

Mit XML liegt eine Beschreibungssprache für semistrukturierte Daten vor, die es erlaubt, mittels DTD oder XML Schemas die Syntax eines Dokumentes auszugeben. Allerdings können gleiche Sachverhalte (wie in Abbildung 2 dargestellt) auf unterschiedliche Art und Weise beschrieben werden.

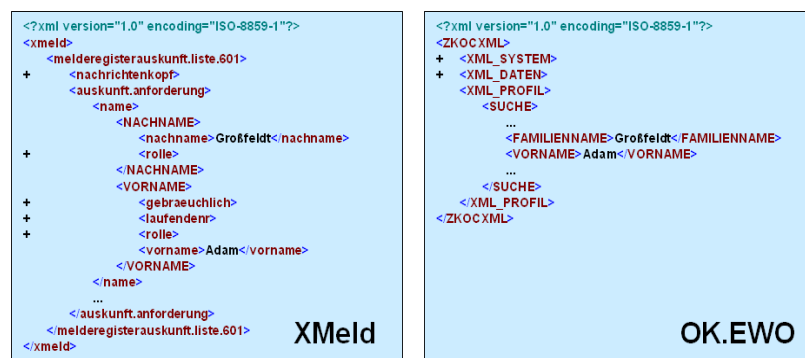


Abbildung 2: Anfrage einer Melderegisterauskunft mit XMeld und OK.EWO

Um Homogenität zwischen den verschiedenen Strukturen mit gleicher Aussage zu erzielen, müssen die Daten einer Struktur *A* semantisch korrekt in eine Struktur *B* umgesetzt werden. Zur Überführung der Datenstrukturen werden diese transformiert, welches mit Hilfe von XSL-Stylesheets realisiert werden kann. Voraussetzung für den automatischen Transformationsprozess ist eine maschinenverständliche Semantik der verschiedenen Datenstrukturen. Eine Ontologie bildet dabei die Basis für eine semantische Beschreibung. Die Abbildung von Semantik auf die Datenstrukturen lässt sich mit Mappingvorschriften auf die Ontologie erreichen.

Als Beschreibungssprache für die Ontologie wurde RDF sowie DAML+OIL gewählt. Da die Ontologie nicht allumfassend sein kann, ist eine Voraussetzung für die Definition der Ontologie eine abgeschlossene Domäne. Die prototypische Umsetzung dieser Arbeit erfolgt am Beispiel Melderegisterwesen. Die entwickelte Ontologie bildet hierzu alle Objekte und Beziehungen in diesem Anwendungsfeld ab und orientiert sich an dem Informationsmodell von XMeld (Datenaustauschformat für das Meldewesen; OSCI). Auf der Grundlage der Ontologie, die hier eine Taxonomie bildet, können nun zwischen dieser und den Schemen (Schema des Eingabe- sowie des Ausgabedokumentes) Abbildungen beschrieben werden. Dieses erfolgt mit RDF, wobei verschiedene Typen für das Mapping existieren.

Mit Hilfe der Informationen und Metainformationen aus dem Ausgangsdokument, den Dienstbeschreibungen, den Schemen, der Ontologie und dem Mapping lässt sich nun ein XSL-Stylesheet erzeugen, mit dessen Hilfe die Transformation vom Ausgangs- ins Ergebnisdokument erfolgen kann. Die Schritte zur Erstellung des Stylesheets sind in Abbildung 3 skizziert.

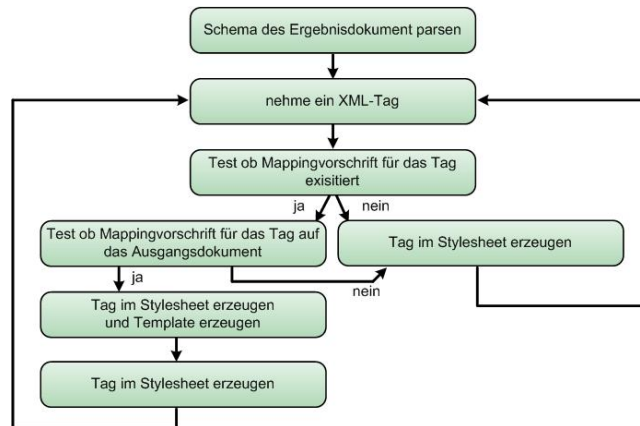


Abbildung 3: Erzeugung des XSL-Stylesheets

3.3 Sichten

Sichten können als allgemeines Abstraktions- und Transformationskonzept dienen und dabei Funktionen wie Datenreduktion oder die Bereitstellung von Nutzerfunktionalitäten abbilden. Eine Sicht, bestehend aus einer Menge von Selektionen und Projektionen, kann über die Ontologie definiert und so sowohl auf ein Eingangs- und/oder Ausgabedokument bei der Transformation angewendet werden. Durch dieses Sichtenkonzept ist es beispielsweise möglich, die Anmeldung in verschiedenen Ländern miteinander zu realisieren, indem entsprechende Sichten für die jeweiligen Nationalitäten bestimmte Werte vordefinieren bzw. ausblenden. Das Sichtenkonzept erlaubt ebenfalls die Kombination von Sichten. So lässt sich zum Beispiel eine Sicht für Mecklenburg-Vorpommern und eine speziellere Sicht für ein Amt definieren, die beide in den Transformationsprozess einfließen. Ebenso lassen sich die Sichten bei verschiedenen Fachverfahren anwenden. Eine Sicht eines bestimmten Amtes kann somit sowohl bei der Anmeldung als auch bei einer Melderegisterauskunft eingesetzt werden.

Durch den Einsatz von Sichten ergibt sich die Möglichkeit, Anfragen an das Gateway zu stellen, deren Anfragedokument nicht vollständig der Definition entspricht, in Kombination mit den verwendeten Sichten jedoch korrekt ist.

Die Definition der Sichten erfolgt analog zu der Definition des Mapping mit RDF, wobei die Daten im Metadatenrepository des Systems abgelegt werden.

3.4 Prototypische Implementierung

Im Rahmen der Arbeit wurde ein Prototyp *eFormsConnect* entwickelt, der die oben genannten Konzepte entsprechend umsetzt. Den zentralen Kern der Architektur bildet der Transformationsprozess, indem auf Basis des eingehenden Dokumentes und der verschiedenen Metainformationen das XSL-Stylesheet erzeugt und das Dokument transformiert wird. Die notwendigen Metainformationen wie Dienstbeschreibungen, Ontologie, Mapping und Sichten werden im Metadaten-Repository verwaltet. Die Dienstbeschreibung erfolgt in WSDL, wobei zur Verarbeitung der Dokumente WSDL4J und zum Parsen der XML-Schemas Castor eingesetzt wird. Alle anderen Metainformationen werden in RDF beschrieben und im RDF-Repository [Au02] verwaltet. Grundlage hierfür ist die RDF-Datenbank Sesame mit einer RQL-Schnittstelle als Anfragesprache.

Der Driver Pool bindet verschiedene Treiber (Connectoren) für die unterschiedlichsten Schnittstellenimplementierungen ein, die zur Kommunikation mit den verschiedenen externen Diensten eingesetzt werden. Hierzu wurden Treiber mit SOAP-, RMI-, nativer Java- oder auch JDBC-Schnittstelle umgesetzt. Beim Start des Gateways werden diese registriert. Die Informationen zur Initialisierung eines jeweiligen Treibers für eine Kommunikation mit einem externen Dienst ergeben sich aus der WSDL-Beschreibung des Dienstes, welche im Metadaten-Repository hinterlegt ist.

Das Gateway selbst präsentiert sich als Web Service und stellt eine Schnittstelle auf der Basis von SOAP bereit.

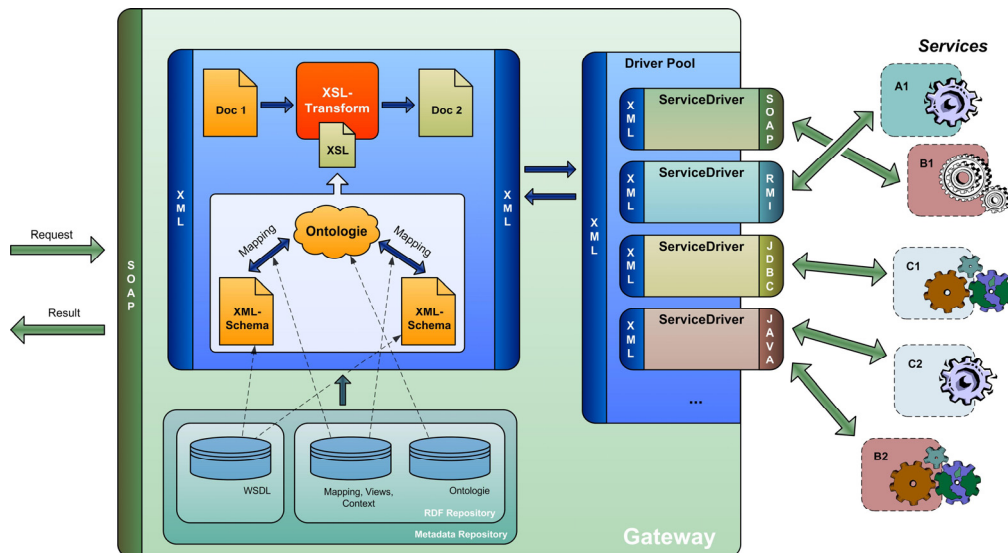


Abbildung 4: eFormsConnect Architektur

4 Zusammenfassung und Ausblick

Der hier vorgestellte Beitrag stellt einen Ansatz in Form eines universellen Gateways vor, der es ermöglicht, heterogene Systeme miteinander zu kombinieren. Betrachtete Schwerpunkte sind dabei sowohl die syntaktische als auch die semantische Integrität der Daten. Im vorgestellten Ansatz wurden Web Service Technologien und semantischen Konzepten des Semantic Web miteinander verknüpft. Diese wurden zusätzlich um Sichtenkonzepte erweitert, um so einen dynamischen Zugriff auf Dienste zu ermöglichen.

Die vorgestellten Ansätze wurden am Beispiel des Melderegisterwesens aus dem Bereich eGovernment vorgestellt. In den derzeit laufenden Forschungs- und Entwicklungsarbeiten erfolgt die Abbildung des beschriebenen Lösungsansatzes auf weitere Anwendungsbereiche. Im Rahmen des WIESEL-Projektes [RuFl04] erfolgt so z.B. die flexible und offene Lernmanagementsystem-Föderation und Integration von Wissensmanagement-Systemen auf der Basis semantisch angereicherter Web Services. Dabei soll die Generierung und Erzeugung multimedialer Lernmodule sowie die Unterstützung mobiler Lernszenarien kontextabhängig gesteuert werden. Darüber hinaus wird derzeit im Rahmen eines Biotechnologie-Anwendungsszenarios [AC+03b] die Nutzung des obigen Ansatzes zur inhaltsorientierten Fusion molekularbiologischer Analyse-Daten genutzt.

5 Literaturverzeichnis

- [AC+03] S. Audersch, T. Courvoisier, G. Flach: eFormsDirect – XML-basiertes eGovernment-Framework für Intelligente Formulare auf der Basis von XForms. Proc. XMIDX 2003, Berlin, 2003.
- [AC+03b] S. Audersch, T. Courvoisier, G. Flach: KnowledgeDirect – Wissensmanagement im Unternehmensnetzwerk BioCon Valley, 4. Oldenburger Fachtagung Wissensmanagement, 2003
- [Au02] S. Audersch: Universelle Metadatenverwaltung zur Steuerung von Transformations- und Daten-Analyse-Prozessen. 14. GI-Workshop Grundlagen von Datenbanken, Darß, 2002.
- [KK+03] D. König, M. Kloppmann, F. Leymann, G. Pfau, D. Roller: Web Services Invocation Framework: A Step towards Virtualization Components. In (R. Eckstein, R. Tolksdorf Hrsg.) Proceedings zum Workshop XMIDX2003. Köllen Druck+Verlag GmbH, Bonn, 2003: S. 33-44.
- [KS+03] M. Keidl, S. Seltzsam, C. König, A. Kemper: Kontext-basierte Personalisierung von Web Services. In (G. Weikum, H. Schöning, E. Rahm Hrsg.) Tagungsband der 10. BTW-Konferenz. LNI, Leipzig, 2003; S. 344-363.
- [RuFl04] M. Rust, G. Flach: Integration von Wissensmanagement und eLearning im Rahmen der WIESEL-Frameworkarchitektur. 16. GI-Workshop Grundlagen von Datenbanken, Monheim, 2004.