

agroXML Inhaltslisten – Konzeption und Inhalte

Daniel Martini, Jürgen Frisch, Martin Kunisch

Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL)
Bartningstraße 49
64289 Darmstadt
d.martini@ktbl.de
j.frisch@ktbl.de
m.kunisch@ktbl.de

Abstract: Um mißverständnisfreien Datenaustausch zwischen Maschinen zu gewährleisten, ist es nötig, Zusammenstellungen von eindeutigen Bezeichnern für bestimmte Objekte bereitzustellen. Im Rahmen der Dokumentation im Agrarbereich müssen beispielsweise Pflanzenschutzmittel oder Sorten eindeutig identifizierbar sein. agroXML stellt hierfür Inhaltslisten bereit. Diese sind als XML-Dictionaries umgesetzt. Die Inhalte selbst stammen in aller Regel aus national und international anerkannten Quellen. Ergänzt werden diese durch einen Satz von Metadaten, der jede Inhaltsliste begleitet. Ein Webservice ermöglicht es Farmmanagement-Informationssystemen, die jeweils aktuellen Inhaltslisten herunterzuladen.

1 Einleitung

Eine Kommunikation zwischen zwei Partnern wird erst möglich, wenn diese sich auf ein gemeinsames Verständnis von Begriffen zur Beschreibung von Objekten und deren Eigenschaften geeinigt haben. agroXML hat zum Ziel, diese Einigung mit Hilfe einer standardisierten Datenaustauschsprache für den Agrarsektor voranzutreiben. Ein XML Schema gibt dabei die Struktur, vergleichbar einer Grammatik, vor. Dabei werden Elemente definiert, die bestimmte Inhalte transportieren können [KBR07]. Oft kommt es vor, dass diese Inhalte selbst Bezeichner für Objekte sind, wie beispielsweise ein eingesetztes Pflanzenschutzmittel oder eine angebaute Sorte. Um eine maschinelle Auswertung zu ermöglichen, müssen in diesem Zusammenhang Schreib- oder Übertragungsfehler vermieden werden, das heißt, es muss eine Einigung über die zu verwendenden Bezeichner erzielt werden.

2 Problemstellung

Ein Landwirt möchte zu einem Dienstleister Daten senden, auf deren Basis eine automatisierte Auswertung durchgeführt wird. Die Daten hat er vorher in seine Schlagkartei oder in ein elektronisches Formular eingegeben. Dabei hat er als Kulturart die Kartoffel angegeben, als Sorte „Leila“. Richtig wäre für die Sorte im Zusammenhang mit der Kartoffel die Schreibweise „Leyla“. Allerdings gibt es bei der Kulturart Besenheide eine Sorte mit dem Namen „Leila“. Allein durch diesen trivialen Tippfehler wird eine maschinelle Auswertung ohne Eingriff des Menschen unmöglich. Die Entscheidung, ob die Fehleingabe bei der Sorte oder bei der Kulturart erfolgte, ist ohne aufwändige Plausibilisierungsalgorithmen für den Rechner allein nicht möglich.

Eine Möglichkeit, solche Schwierigkeiten zu vermeiden, ist es, strukturierte Listen mit eindeutigen Bezeichnern verknüpft mit den Namen von Objekten und gegebenenfalls einer näheren Erläuterung bereitzustellen. Diese Listen können zum einen dazu dienen, dem Nutzer eine Auswahlfunktion anzubieten. Andererseits müssen beim Datenaustausch dann auch lediglich die eindeutigen Bezeichner übertragen werden, damit der Empfänger einer Nachricht versteht, welches Objekt gemeint ist.

3 Anforderungen

Sowohl inhaltlich als auch technisch werden an die Umsetzung der Listen gewisse Anforderungen gestellt. So sollten die Einträge national und gegebenenfalls auch international anerkannten Standards und Begriffssammlungen folgen. Die Listen sollen außerdem in einem vom agroXML-Schema unabhängigen Veröffentlichungszyklus aktualisiert und erweitert werden können. Wenn beispielsweise neue Sorten zugelassen werden, so muss es möglich sein, diese neuen Einträge in agroXML-Instanzen einbinden zu können, ohne dass sich am agroXML-Schema selbst etwas ändert. Diese Anforderung lässt sich nur erfüllen, wenn die Listen getrennt vom Schema gehalten und gepflegt werden. Verschiedene Listen enthalten Einträge zur Befüllung unterschiedlicher agroXML-Elemente. Wünschenswert ist ein generisches Konstrukt, das erlaubt, Listen trotz verschiedener Inhalte über eine standardisierte Schnittstelle einlesen und aktualisieren zu können.

4 Umsetzung

Die Umsetzung der Listen erfolgte als XML-Dictionaries. Im Gegensatz zur Einschränkung von Elementinhalten mit Hilfe einer so genannten enumeration [BM04] erlauben diese die oben genannte, vom Schema unabhängige Erweiterung. Außerdem ist bei Bedarf die Einbindung spezieller, angepasster Listen möglich. Jedem Eintrag einer Inhaltsliste wird in einem XML-Dictionary eine ID (über das Attribut `agrodictionary:id` zum Element `<Definition>`), ein Name (Element `<Name>`) und eine Beschreibung (Element `<Description>`) zugeordnet:

```

<DictionaryEntry>
  <Definition agrodict:id="S12">
    <Description>
      Kornfraktionen (Masse-%): Ton: 5-8; Schluff: 10-25;
      Sand: 67-85
    </Description>
    <Name>schwach lehmiger Sand</Name>
  </Definition>
</DictionaryEntry>

```

Das `id`-Attribut ist hierbei als `xsd:id` vereinbart, so dass lediglich bestimmte Zeichenklassen erlaubt sind. Außerdem werden in jeder Liste Metadaten zum dictionary selbst gehalten. Diese sind im Einzelnen:

- Name: Der Name des Dictionaries selbst.
- Description: Die Beschreibung des Dictionaries. Sie enthält in aller Regel auch eine Quellenangabe.
- LastChangedAt: Der Zeitpunkt des letzten Hinzufügens eines neuen Eintrags.
- LastAccessedAt: Der Zeitpunkt der letzten Überprüfung.
- LastChangedBy: Kontaktdaten der Person, die die letzte Änderung durchgeführt hat.
- Maintainer: Kontaktdaten der Person, die für die Liste verantwortlich zeichnet.
- Version: Die Versionsnummer der Liste, eine ganze Zahl, die stetig hochgezählt wird.

In die Instanz eingebunden wird lediglich der im Attribut `id` angegebene Bezeichner. Eindeutig wird dieser Bezeichner aber erst, wenn auch angegeben wird, aus welcher Liste der Eintrag stammt. Diese Angabe wird ins `codeSpace`-Attribut des entsprechenden Elements in der Instanz eingetragen:

```

<Bodenart
codeSpace="http://www.agroxml.de/content/SoilTexture.xml"
codeListVersionID="1">S12</Bodenart>

```

Der dringende Bedarf nach Inhaltslisten bestand im Bereich der Pflanzenschutzmittel, insbesondere weil dort gesetzliche Vorgaben eine ordnungsgemäße Dokumentation verlangen. Die Liste der Pflanzenschutzmittelnamen enthält als Bezeichner die Zulassungsnummer des jeweiligen Mittels. Sie wird vom Bundesamt für Verbraucherschutz und Lebensmittelsicherheit gepflegt und für agroXML aufbereitet. Im Datenaustausch mit Abnehmern landwirtschaftlicher Produkte, die besondere Qualitätsansprüche haben, wie beispielsweise Mälzereien oder Mühlen, sind häufig auch genaue Angaben über eingesetzte Sorten notwendig. Informationen zu den Sortennamen werden vom Bundessortenamt bezogen.

Weitere Listen geben Bezeichner für Bodenarten, Maschinenarten, Düngemitteltypen und vieles mehr aus unterschiedlichsten Informationsquellen. Die derzeit verfügbaren Listen sind auf <http://www.agroxml.de/content> abrufbar.

Farmmanagement-Informationssysteme und andere Softwaresysteme, die agroXML nutzen, sollen den jeweils aktuellen Stand der Inhaltslisten zur Befüllung von Instanzen lokal vorhalten. Dabei gibt es verschiedene Caching-Strategien. agroXML macht hierzu keine Vorgaben. Als Unterstützung zur Überprüfung der Aktualität und zum Download der Listen steht ein Webservice zur Verfügung, der als Eingabeparameter eine Listenversionsnummer sowie einen Listennamen erwartet. Sollte eine aktuellere Liste als die mit der übergebenen Versionsnummer bezeichnete existieren, so wird diese als Ausgabe geliefert.

5 Ausblick

Künftig werden weitere Listen im Rahmen des IT-FoodTrace-Projektes für die Anwendung von agroXML im Bereich der Tierhaltung zusammengestellt werden. Insbesondere von Interesse für Dokumentationsbelange sind hierbei Tierarzneimittel und Futtermittel. Außerdem in Arbeit ist eine leicht bedienbare Anwendung zur Pflege der Inhaltslisten, die es ermöglichen soll, die Bereitstellung neuer Einträge den jeweiligen fachlichen Experten zu überlassen, auch wenn diese keine speziellen IT-Kenntnisse besitzen.

Danksagung

Wir danken dem Bundesamt für Verbraucherschutz und Lebensmittelsicherheit sowie dem Bundessortenamt für ihre Bereitschaft, die entsprechenden Listen zur Verfügung zu stellen. Wir bedanken uns beim BMBF für die Förderung (Förderkennzeichen 0330664 und 0330761) im Rahmen der Verbundprojekte preagroII und IT FoodTrace, und dem Projektträger des BMBF und BMWi (Forschungszentrum Jülich GmbH). Außerdem danken wir der Deutschen Bundesstiftung Umwelt für die Förderung im Rahmen des abgeschlossenen Projektes „Integrierte Pflanzenproduktion – Standardisierte Datenformate zur Dokumentation landwirtschaftlicher Produktionsprozesse“.

Literaturverzeichnis

- [BM04] Biron, P. V., Malhotra, A. (Hrsg.): XML Schema Part 2: Datatypes Second Edition. The World Wide Web Consortium, 2004; <http://www.w3.org/TR/xmlschema-2/>, zitiert am 05.11.2006.
- [KBR07] Kunisch, M.; Böttinger, S.; Rodrian, H.-C.: Stand der Entwicklung von agroXML. In: Proceedings 27. GIL-Jahrestagung, Stuttgart 2007.