

# **Learning Object Metadata Generation: Von expertenerstellten zu benutzererstellten Metadaten für Lernmaterialien**

Masterarbeit  
an der  
Technischen Universität Graz

vorgelegt von

**Philip Kohler**

Institut für Wissensmanagement (IWM)  
Technische Universität Graz  
A-8010 Graz

August 2012

© Copyright 2012, Philip Kohler

Begutachterin: Univ.-Prof. Dipl.-Inf. Dr. Stefanie Lindstaedt

Betreuerin: Dr. Elisabeth Lex



# EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am .....

.....

(Unterschrift)

## Kurzfassung

Das Annotieren von Lernressourcen und das Übersetzen und Korrigieren von bestehenden Metadatenbeschreibungen für Lernressourcen stellen einen zeitaufwendigen Prozess dar. Grundsätzlich werden Lernressourcen und die dazugehörigen Metadatenbeschreibungen manuell von Experten erstellt. Der Benutzer oder der Lernende, der die Lernressource verwendet, hat häufig keine Möglichkeiten, bestehende Lernressourcen oder deren Metadatenbeschreibungen zu korrigieren. Ferner besteht für den Benutzer eines Online-Repositories, das Lernressourcen veröffentlicht, oftmals keine Möglichkeit, neue interessante Lernressourcen hinzuzufügen und somit anderen Benutzern zur Verfügung zu stellen.

Aus diesem Grund wurden im Rahmen dieser Arbeit User-Generated-Content Services und Tools entwickelt, die es den Benutzern eines Online-Repositories ermöglichen, bestehende Metadatenätze zu korrigieren, zu übersetzen und neue Lernressourcen zu annotieren. Die User-Generated-Content Services unterstützen dabei den Benutzer bei der Bearbeitung von Metadaten durch die Metadatenextraktion von neuen Lernressourcen, durch Autovervollständigung von Eingaben und durch maschinelle Übersetzung von Metadatenelementen.

Diese User-Generated-Content Services und Tools wurden in einer Komponente zusammengefasst, die als Blackbox funktioniert und einfach in Web-Portale integriert werden kann. Der Zugriff auf die Komponente basiert auf gut definierten Schnittstellen. Die User-Generated-Content Komponente verwendet eine adaptierte Version des IEEE LOM Standards als Metadatenformat für die Beschreibung von Lernressourcen. Weiters wurde der OAI-PMH Standard umgesetzt, sodass die bearbeiteten Metadaten abrufbar sind.

Diese User-Generated-Content Komponente wurde im Rahmen des Organic.Lingua Projekts entwickelt. Der erste Prototyp der User-Generated-Content Komponente wurde den Organic.Lingua Projektpartnern für eine Begutachtung vorgelegt. Dieser Prototyp wurde von diesen sehr positiv aufgenommen. Im Bereich der Usability gab es detaillierte Verbesserungsvorschläge, welche bei der Umsetzung des zweiten Prototypen beachtet werden sollen.

## **Abstract**

The annotation of learning resources as well as translating and correcting of existing learning resource metadata is a time-consuming process. Basically, learning resources and the corresponding metadata are created manually by experts. The user who uses the learning resources often has no possibility to correct existing learning resources or their metadata descriptions. Furthermore, there is often no way for online repository users to add new interesting learning resources and consequently share these learning resources with the community.

To address this issue user-generated content services and tools have been developed in this work, which enable users of an online repository to correct and translate existing metadata records and to annotate new learning resources. To support users, a set of services has been implemented to automatically generate metadata from learning resources, to auto-suggest metadata values for metadata fields and to automatically translate metadata fields.

These user-generated content services and tools have been combined into a single component that works as a black box and can be easily integrated into any web portals. Access to the component is based on well-defined interfaces. The user-generated content component uses an adapted version of the IEEE LOM standard as a metadata format for describing learning resources. Furthermore, the OAI-PMH standard has been implemented, which allows access to the edited metadata records.

The user-generated content component was developed as part of the Organic.Lingua project. The first prototype was submitted to the project partners for review, who responded very positively. Concerning usability, there were detailed suggestions for improvements that should be considered during the implementation of the second prototype.

## **Danksagung**

Ich möchte mich beim Know-Center bedanken, durch welches mir ermöglicht wurde, meine Masterarbeit in diesem Rahmen zu verfassen.

Mein Dank gilt meiner Betreuerin Dr. Barbara Kump, welche mir durch ihre positive Einstellung bereits während des gesamten Organic.Lingua Projekts viele Ideen und Möglichkeiten eröffnet hat, an mein Thema heranzugehen. Außerdem danke ich meiner nachfolgenden Betreuerin Dr. Elisabeth Lex, welche sich gerne bereit erklärt hat, meine Masterarbeit in weiterer Folge zu betreuen und stets ein offenes Ohr für meine Fragen hatte.

Des Weiteren möchte ich Univ.-Prof. Dipl.-Inf. Dr. Stefanie Lindstaedt für ihre Betreuung danken. Mein Dank gilt außerdem allen Organic.Lingua Projektpartnern und Andi Rexha für die Implementierung der Metadatenextraktions-Services.

Schlussendlich möchte ich meinen Eltern danken, die immer an mich geglaubt haben und meiner Freundin, welche mich während meines Studiums und der Masterarbeit immer unterstützt hat und mir viel Geduld und Verständnis entgegengebracht hat.

# Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Problemhintergrund.....	1
1.1.1	Das Organic.Lingua Projekt.....	1
1.1.2	Organic.Lingua Projektpartner.....	1
1.1.3	Das Organic.Edunet Web-Portal.....	3
1.1.4	Aktueller Integrationsprozess von Lernmaterialien .....	4
1.2	Ziel der Arbeit.....	7
1.3	Aufbau der Arbeit.....	9
1.4	Definitionen.....	10
2	Forschungsgrundlage.....	11
2.1	User-Generated-Content .....	11
2.1.1	Nutzergetriebene Inhaltserstellung versus industrielle Inhaltserstellung.....	14
2.1.2	Arten von User-Generated-Content und deren Einsatzgebiete.....	17
2.1.3	Motivation der Benutzer .....	20
2.1.4	Verbreitung von User-Generated-Content in der Europäischen Union . .....	24
2.2	Limitierungen von User-Generated-Content .....	26
2.3	Metadaten.....	27
2.3.1	Wichtigkeit von Metadaten.....	30
2.3.2	Metadatenerstellung .....	32
2.4	Information Retrieval Ansätze für die Metadatengewinnung .....	38
2.4.1	Öffentliche Information Retrieval Services.....	42
2.5	Qualität von Metadaten .....	43
2.6	Bereitstellen von Metadaten .....	45
2.7	Zusammenfassung .....	46
3	Die User-Generated-Content Komponente.....	51

3.1	Einleitung.....	51
3.2	Ähnliche Systeme und Projekte.....	54
3.3	Beschreibung der User-Generated-Content Szenarien.....	55
3.3.1	Szenario 1: Einen bestehenden Metadatensatz bearbeiten.....	55
3.3.2	Szenario 2: Einen bestehenden Metadatensatz übersetzen.....	57
3.3.3	Szenario 3: Eine neue Lernressource hinzufügen.....	59
3.4	Lebenszyklus der Lernressourcen-Metadatenbeschreibung.....	62
3.5	Das Organic.Edunet IEEE LOM Application Profile.....	65
3.5.1	Referenz auf den originalen Metadatensatz.....	68
3.5.2	Bearbeiter des Metadatensatzes.....	70
3.6	Nutzungsrechte der Lernressource.....	70
3.7	User-Generated-Content Services.....	75
3.7.1	Übersetzungs-Service.....	75
3.7.2	Metadatenextraktions-Service.....	76
3.7.3	Autosuggestions-Service.....	77
3.7.4	Rechtschreibkorrektur-Service.....	78
4	Technische Aspekte der Implementierung.....	79
4.1	Verwendete Technologien.....	79
4.2	Verwendete Frameworks und Bibliotheken.....	81
4.3	Cross-Domain Kommunikation.....	84
4.4	Integration der User-Generated-Content Komponente.....	88
4.4.1	Integration des Metadaten-Bearbeitungs-Widgets.....	88
4.4.2	Integration des Metadaten-Übersetzungs-Widgets.....	90
4.4.3	Integration des "Suggest new Learning Resource" Widgets.....	92
4.5	Bereitstellen der User-Generated-Metadaten.....	93
5	Reflexion und Ausblick.....	97
5.1	Zusammenfassung.....	97
5.2	Auswirkung auf die Organic.Lingua Vision.....	99

5.3	Feedback der Organic.Lingua Projektpartner.....	101
5.4	Ausblick für die Zukunft .....	102

## Abbildungsverzeichnis

Abbildung 1-1:	Integrationsprozess von Lernmaterialien (Quelle Deliverable 6.1).....	5
Abbildung 2-1:	Limitierte Feedback-Möglichkeiten bei der industriellen Inhaltserstellung (Bruns, 2008, S. 11) .....	15
Abbildung 2-2:	Die Rolle des „Producers“ (Bruns, 2007).....	16
Abbildung 2-3:	Arten von Lernressourcen im Organic.Edunet Web-Portal .....	20
Abbildung 2-4:	Motivationsniveau der Wikipedia-Autoren in Korrelation mit der Beitragshöhe. Die Standardabweichung in runden Klammern. Der Pearson-Korrelationskoeffizient in eckigen Klammern. (Nov, 2007).....	23
Abbildung 2-5:	Prozentueller Anteil der Internetbenutzer im Alter von 16 – 24 Jahren, die User-Generated-Content erzeugen, im Jahr 2005 (OECD, 2007, S. 20).....	25
Abbildung 2-6:	Prozentueller Anteil der Internetbenutzer im Alter von 16 – 74 Jahren, die User-Generated-Content erzeugen, im Jahr 2006 (OECD, 2007, S. 21).....	25
Abbildung 2-7:	Aufbau eines typischen Dublin Core Satzes (Baker, 2000) .....	28
Abbildung 3-1:	User-Generated-Content Szenarien .....	52
Abbildung 3-2:	Screenshot: Auswahl der User-Generated-Content Operation. ....	55
Abbildung 3-3:	Screenshot: „Correct Metadata Record“ Widget .....	56
Abbildung 3-4:	Screenshot: Auswahl der User-Generated-Content Operation. ....	57
Abbildung 3-5:	Screenshot: Auswahl der Ziel-Sprache .....	57
Abbildung 3-6:	Screenshot: „Translate Metadata Record“ Widget .....	58



Abbildung 3-7: Screenshot: „Suggest Learning Resource“ Button zur Lesezeichen-Symbolleiste hinzufügen .....	59
Abbildung 3-8: Screenshot: „Suggest NEW Learning Resource“ Widget .....	61
Abbildung 3-9: Publikationsprozess .....	62
Abbildung 3-10: Lebenszyklus der Metadaten .....	64
Abbildung 3-11: Screenshot: Nutzungsrechte „Custom License“ .....	72
Abbildung 3-12: Screenshot: Nutzungsrechte „Creative Commons License“ ...	72
Abbildung 4-1: Test-Webseite.....	80

## **Tabellenverzeichnis**

Tabelle 3-1: Organic.Edunet IEEE LOM Application Profile.....	68
Tabelle 3-2: Rolle des Metadaten-Contributors.....	70
Tabelle 5-1: Einfluss der User-Generated-Content Komponente auf die Ziele des Organic.Lingua Projektes.....	101

## **Codeausschnittverzeichnis**

Codeausschnitt 3-1: IEEE LOM Relation: Verweis auf originalen Metadatensatz bei der Bearbeitung eines existierenden Metadatensatzes.....	69
Codeausschnitt 3-2: IEEE LOM Relation: Verweis auf originalen Metadatensatz bei der Übersetzung eines existierenden Metadatensatzes.....	69
Codeausschnitt 4-1: Cross-Domain Kommunikation mit Hilfe des <script>-Tags .....	86
Codeausschnitt 4-2: Integrationscode für das Metadaten-Bearbeitungs-Widget .....	90
Codeausschnitt 4-3: Integrationscode für das Metadaten-Übersetzungs-Widget .....	91

Codeausschnitt 4-4: Integrationscode für das „Suggest new learning resource“ Widget.....	93
Codeausschnitt 4-5: Beispiel OAI-PMH Response von der User-Generated- Content Komponente.....	95

# 1 Einleitung

Dieses Kapitel beschreibt die Ausgangssituation, indem das Organic.Lingua Projekt und das Organic.Edunet Web-Portal beschrieben werden. Anschließend werden die Motivation, die Forschungsfrage und die Ziele dieser Masterarbeit dargestellt. Abschließend wird der Aufbau der Arbeit skizziert und wichtige Begriffe der Arbeit werden definiert.

## 1.1 Problemhintergrund

In diesem Abschnitt wird die Ausgangssituation dieser Masterarbeit beschrieben. Zuerst werden das Organic.Lingua Projekt und die Eckdaten des Organic.Edunet Web-Portals beschrieben, in Folge wird der aktuelle Integrationsprozess von Lerninhalten in das Organic.Edunet Web-Portal aufgezeigt.

### 1.1.1 Das Organic.Lingua Projekt

Die vorliegende Arbeit wurde im Zuge des Organic.Lingua ([www.organic-lingua.eu](http://www.organic-lingua.eu)) Projektes erstellt. Das Ziel des Projektes besteht darin, mehrsprachige Services und Tools zu entwickeln, die die Wiederverwendung, die Erschließung, die Erweiterung und die Auffindung digitaler Lernressourcen zum Thema biologische Landwirtschaft und Agrarökologie erleichtern.<sup>1</sup>

### 1.1.2 Organic.Lingua Projektpartner

Im Folgenden werden die für diese Arbeit relevanten Organic.Lingua Projektpartner beschrieben.

---

<sup>1</sup> Organic.Edunet bH. (2011). Projekte - Organic.Lingua. Abgerufen am 23. Juli 2012 von Know-Center.at: <http://know-center.tugraz.at/projekte>

### **Universität Alcalá**

Die Universität von Alcalá wurde 1499 gegründet und gliedert sich in 15 Fakultäten, die mehr als 20 000 Studenten betreuen (Organic.Lingua, 2010).

An der Universität von Alcalá arbeiten rund 60 Vollzeit-Angestellte an der Fakultät für Informatik. Die Forschungsschwerpunkte des Fachbereiches Informatik bestehen in E-Learning, interaktiven Systemen, Knowledge-Engineering und Fuzzy-Logik. (Organic.Lingua, 2010)

Die Universität Alcalá koordiniert das Organic.Lingua Projekt und ist Work-Package-Leader des Work-Package 5. Das Ziel des Work-Package 5 liegt darin, die bestehenden Online-Dienste des Organic.Edunet Web-Portals zu erweitern und zu reengineeren. Weiters stellt die Universität Alcalá die Infrastruktur für das Organic.Edunet Web-Portal zur Verfügung. (Organic.Lingua, 2010)

### **Xerox Research Centre Europe**

Das Xerox Research Centre Europe forscht in den Themengebieten Natural-Language-Processing, maschinellem Lernen, XML, Computer-Vision, Mathematik, Data-Mining und Ethnografie (Organic.Lingua, 2010).

Im Organic.Lingua Projekt bringt Xerox seine Expertise im Bereich maschineller Übersetzung ein und arbeitet an der Entwicklung des Multilingual-Technology-Frameworks mit (Organic.Lingua, 2010).

### **CELI**

CELI ist ein privates Unternehmen, das 1999 gegründet wurde. Das Hauptziel von CELI besteht darin, Forschungsergebnisse aus den Bereichen Human-Language- Technology, künstlicher Intelligenz und Web-Interaktion in Produkte und Lösungen für Unternehmen einzubauen. (Organic.Lingua, 2010)

Im Organic.Lingua Projekt ist CELI für die Multilingualität und das Cross-Language Information-Retrieval zuständig. Weiters entwickelt CELI auch das Multilingual-Technology-Framework für das Organic.Lingua Projekt mit (Organic.Lingua, 2010).

### 1.1.3 Das Organic.Edunet Web-Portal

Das Organic.Lingua Projekt (siehe Abschnitt 1.1.1) baut auf dem Organic.Edunet Web-Portal auf. Dieses Web-Portal enthält pädagogische Lernmaterialien über das Thema ökologische Landwirtschaft und Agrarökologie. Derzeit bietet Organic.Edunet Zugang zu mehr als 7.000 Lernressourcen aus Sammlungen von elf internationalen Institutionen.<sup>2</sup>

Das Organic.Edunet Web-Portal stellt eine beliebte Lernumgebung dar und wird von vielen Personen weltweit benutzt. Das Web-Portal weist 92.021 Unique-Visitors auf. Von diesen Unique-Visitors besuchen zirka 200 Benutzer täglich das Web-Portal. Die User-Profile der Besucher des Organic.Edunet Web-Portals umfassen die Gruppen Forscher, Lehrer und Studenten. Diese Besucher kommen aus 180 verschiedenen Ländern. (Wotquenne, Le Hénaff, & Protonotarios, 2012)

Im Rahmen von Organic.Edunet entstand ein Netzwerk von Repositories, die Lerninhalte zum Thema biologischer Landwirtschaft und Agrarökologie zur Verfügung stellen. Das Organic.Edunet ruft von allen registrierten Repositories die Metadaten ab und stellt nach einem Validierungsschritt die Lernressourcen über das Web-Portal zur Verfügung. Diesen Abrufprozess von Metadaten nennt man „Harvesting“. Dieser Harvesting-Prozess basiert auf dem OAI-PMH Standard. Dieser OAI-PMH Standard wird in Abschnitt 2.6 genauer beschrieben.

Einen wichtigen Punkt stellt die Mehrsprachigkeit des Organic.Edunet Web-Portals dar. Derzeit ist die Benutzeroberfläche in 16 Sprachen übersetzt worden und durch den Organic.Edunet Benutzer abrufbar. Der Inhalt des Web-Portals ist in bis zu elf Sprachen abrufbar. Weiters werden die Lernressourcen durch Metadaten beschrieben. Die Metadatenbeschreibungen der Lernressourcen stehen in bis zu neun Sprachen zu Verfügung. (Wotquenne, Le Hénaff, & Protonotarios, 2012)

Die Übersetzung der Lernressourcen und der dazugehörigen Metadaten wurde manuell durchgeführt. Diese manuelle Übersetzung war ein zeitaufwendiger

---

<sup>2</sup> Know-Center GmbH. (2011). Projekte - Organic.Lingua. Abgerufen am 23. Juli 2012 von Know-Center.at: <http://know-center.tugraz.at/projekte>

Prozess. Für die Übersetzung der Metadaten in eine zusätzliche Sprache musste ein Domänen-Experte aus dem jeweiligen Land der Zielsprache kontaktiert werden, der darüber hinaus der Ausgangssprache mächtig war. Bevor der Domänen-Experte den Metadatensatz übersetzen konnte, mussten ihm das Schema des Metadatensatzes und das Tool für die Metadatenübersetzung erklärt werden. Erst dann konnte er den Metadatensatz in die Zielsprache übersetzen. Dieser Übersetzungsprozess konnte für eine kleine Anzahl von Metadatensätzen bis zu vier Wochen dauern. Die zeitaufwendigste Aktivität bei diesem Prozess waren das erste Gespräch, sowie die Schulung für die Verwendung des Tools der Metadatenübersetzung. (Wotquenne, Le Hénaff, & Protonotarios, 2012)

Der folgende Abschnitt behandelt den Integrationsprozess von Lernressourcen in das Organic.Edunet Web-Portal.

### **1.1.4 Aktueller Integrationsprozess von Lernmaterialien**

Derzeit werden alle Lernressourcenbeschreibungen bzw. deren Metadaten von den Repositories angeboten und durch das Organic.Edunet geharvestet. Die Metadaten, die die Lernressourcen beschreiben, basieren auf dem IEEE LOM Standard. Dieser IEEE LOM Standard wird in Abschnitt 2.3 genauer erklärt. Derzeit sind Repositories von elf Institutionen an das Organic.Edunet Web-Portal angebunden.<sup>3</sup>

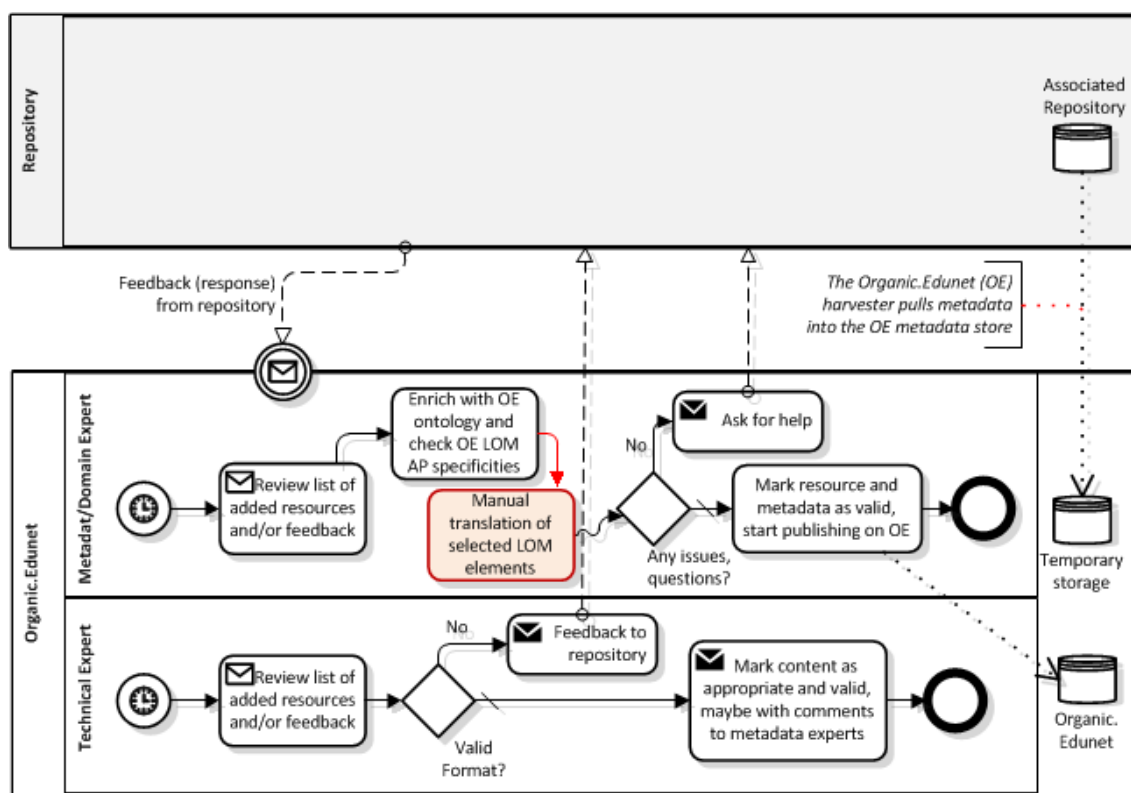
Nachdem die Metadatenbeschreibungen für die Lernressourcen durch das Organic.Edunet Web-Portal geharvestet wurden, werden die Metadaten validiert und mit weiteren Metadaten angereichert. Das folgende BPMN-Diagramm (siehe Abbildung 1-1, Quelle Deliverable 6.1), zeigt die systematische Anreicherung und Validierung der Metadaten. Unter anderem werden die Metadaten mit Konzepten der Organic.Edunet Ontologie und einer manuellen Übersetzung angereichert. In diesen Validierungs- und Anreicherungsprozess

---

<sup>3</sup> Organic.Edunet. (2012). Sammlungen. Abgerufen am 23. Juli 2012 von Organic.Edunet: [http://portal.organic-edunet.eu/index.php?option=com\\_knowmore&view=instcollections&Itemid=96](http://portal.organic-edunet.eu/index.php?option=com_knowmore&view=instcollections&Itemid=96)

sind zwei Akteure involviert. Der erste Akteur nimmt die Rolle des Metadaten- und Domain-Experten ein. Dieser reichert den Metadatensatz mit weiteren Informationen an und validiert diesen. Der zweite Akteur nimmt die Rolle des technischen Experten ein. Der technische Experte überprüft die Einhaltung des Organic.Edunet IEEE LOM Application Profiles und überprüft weitere technische Details des Metadatensatzes (Quelle Organic.Lingua Deliverable 6.1).

Anschließend an den Validierungs- und Anreicherungsprozess der Metadaten, werden diese im Organic.Edunet Web-Portal veröffentlicht (Quelle Organic.Lingua Deliverable 6.1).



**Abbildung 1-1: Integrationsprozess von Lernmaterialien (Quelle Deliverable 6.1)**

Der derzeitige Veröffentlichungsprozess von Lernmaterialien umfasst drei Integrationsmöglichkeiten (Quelle Organic.Lingua Deliverable 6.1).

- **Integrationsmöglichkeit 1: Content-Provider mit Content-Management-Systemen (CMS), die OAI-PMH unterstützen.**

Die digitale Sammlung von Lernressourcen wird in einem Content-Management-System verwaltet. Außerdem unterstützt das System das

OAI-PMH Protokoll für die Bereitstellung der Lernressourcenbeschreibungen. Bevor die Lernressourcen bzw. deren Metadatenbeschreibung vom Content-Provider veröffentlicht werden können und somit für das Organic.Edunet Web-Portal für den Abruf zur Verfügung stehen, muss ein Mapping-Prozess durchgeführt werden. In diesem Mapping-Prozess wird eine Zuordnung zwischen dem lokalen Metadaten-Schema und dem Organic.Edunet IEEE LOM Application Profile durchgeführt. Bevor die Lernressourcen vom Content-Provider im Organic.Edunet Web-Portal veröffentlicht werden, wird eine Qualitätssicherung der neuen Metadatenätze durchgeführt.

- **Integrationsmöglichkeit 2: Content-Provider mit Content-Management-Systemen, die OAI-PMH nicht unterstützen.**

Die digitale Sammlung von Lernressourcen wird in einem Content-Management-System verwaltet. Die Lernmaterialien sind online abrufbar und mit Metadaten annotiert. Das Problem besteht darin, dass der Content-Provider nicht das OAI-Protokoll unterstützt. In diesem Fall bietet AK ein Agriculture Repository Tool (AgLR) an. Diese AgLR Anwendung ermöglicht das Mapping zwischen dem lokalen Metadaten-Schema des Content-Providers und dem Organic.Edunet IEEE LOM Application Profile. Außerdem veröffentlicht das AgLR Tool die Metadaten für den Content-Provider. Die Veröffentlichung durch das AgLR Tool basiert auf dem OAI-PMH Standard.

- **Integrationsmöglichkeit 3: Content-Provider ohne Hosting-Option**

Der Content-Provider verfügt über eine Sammlung von Lernmaterialien, die nicht online zur Verfügung stehen (z. B. sind Konferenz-Papers und Präsentationen nur auf einer CD-ROM oder der Festplatte verfügbar). Der Content-Provider muss die Lernmaterialien mit Hilfe des AgLR Tools beschreiben und erstellt somit einen IEEE LOM Organic.Edunet Application Profile konformen Metadatenatz.

Der derzeitige Integrationsprozess neuer Lernressourcen lässt erkennen, dass dieser Prozess sehr statisch ist und dass nur Lernressourcen von wenigen Institutionen eingebunden werden. Dieser Umstand motiviert dazu, den



Integrationsprozess zu vereinfachen und die Barriere für das Hinzufügen neuer Lernressourcen zu verkleinern.

## 1.2 Ziel der Arbeit

### **Einschränkungen des Organic.Edunet Web-Portals**

Im Organic.Edunet existieren ein Lebenszyklus und ein Integrationsprozess (siehe Abschnitt 1.1.4) für Lernressourcen. Es werden jedoch nur Lernressourcen und die dazugehörigen Metadaten von internationalen Institutionen (wie z. B. Universitäten), die von Experten erfasst wurden, vom Organic.Edunet Web-Portal gesammelt. Die gesammelten Metadaten und somit der Verweis auf die Lernressourcen, werden, bevor sie im Organic.Edunet Web-Portal angezeigt werden, von Experten validiert und mit weiteren Metadaten angereichert.

Einen sehr wichtigen Punkt im Organic.Edunet Web-Portal stellt die Mehrsprachigkeit der Lernressourcen und deren Metadaten dar. Die Übersetzung wurde manuell und zeitaufwendig durch Domänen-Experten erstellt. Ein weiteres Problem bei der Übersetzung bestand darin, dass der Domänen-Experte sowohl der Ausgangsprache als auch der Zielsprache der zugrunde liegenden Lernressource bzw. deren Annotierung mächtig sein musste.

Das Grundproblem des bestehenden Organic.Edunet Web-Portals liegt darin, dass der Lernende bzw. der Organic.Edunet Benutzer die Lerninhalte und die dazugehörigen Metadaten nicht korrigieren, erweitern und übersetzen darf. Außerdem besteht für die Organic.Edunet Benutzer keine Möglichkeit, neue, interessante Lernressourcen zum Organic.Edunet Web-Portal hinzuzufügen.

### **Forschungsfrage der Arbeit**

Aus diesen Grundproblemen leitet sich folgende Forschungsfrage für die vorliegende Masterarbeit ab:

*Wie kann der Organic.Edunet Benutzer aktiv neue Lernressourcen zum Organic.Edunet Web-Portal hinzufügen und bestehende Annotierungen von*

*Lernressourcen verbessern, mit zusätzlichen Informationen anreichern und übersetzen?*

**Ziel der Arbeit**

Aus der expliziten Formulierung der Forschungsfrage der vorliegenden Masterarbeit lässt sich folgende implizite Formulierung der Forschungsfrage für diese Arbeit ableiten:

*Das Ziel der Masterarbeit liegt darin, dem Organic.Edunet Benutzer Möglichkeiten zu bieten, neue Lernressourcen zum Organic.Edunet Web-Portal hinzufügen zu können und bestehende Annotierungen von Lernressourcen verbessern, mit zusätzlichen Informationen anreichern und übersetzen zu können.*

Um dieses Ziel zu erreichen, wird der bestehende Integrationsprozess von Lernressourcen im Organic.Edunet Web-Portal (siehe Abschnitt 1.1.4) erweitert und Tools und Services werden entwickelt, die dem Organic.Edunet Benutzer das Annotieren von Lernressourcen ermöglichen. Ferner wird bei der Umsetzung darauf geachtet, dass die Tools für die Annotierung von Lernressourcen einfach in beliebige Systeme integriert werden können und dass Schnittstellen und Daten auf Standards beruhen oder gut definiert sind.

## **1.3 Aufbau der Arbeit**

Kapitel 1 befasst sich mit dem Problemhintergrund und dem Ziel der vorliegenden Arbeit. Weiters werden im Kapitel 1 die Definitionen für Begriffe erklärt.

In Kapitel 2 wird auf die Forschungsgrundlagen eingegangen. Zuerst wird der Begriff User-Generated-Content beschrieben, danach Limitierungen von User-Generated-Content. Anschließend behandelt das Kapitel 2 die Metadaten-Standards für Lernressourcen, die Wichtigkeit von Metadaten und die Metadatenerstellung. Weiters befasst sich das Kapitel 2 mit Information Retrieval Technologien, der Qualität von Metadaten und der Bereitstellung von Metadaten.

Das Kapitel 3 beschreibt die implementierte User-Generated-Content Komponente. Es werden ähnliche Systeme analysiert, User-Generated-Content Szenarien beschrieben und der Lebenszyklus der Lernressourcen beleuchtet. Ferner stellt das Kapitel 3 das Organic.Edunet IEEE LOM Application Profile vor und beschreibt die User-Generated-Content Services.

In Kapitel 4 werden die technischen Aspekte der Implementierung und die Integration der User-Generated-Content Komponente diskutiert.

Abgeschlossen wird diese Arbeit mit Kapitel 5, in dem die Arbeit zusammengefasst wird und die Auswirkungen dieser Masterarbeit auf das Organic.Lingua Projekt analysiert werden. Außerdem wird das Feedback des ersten Prototypen der User-Generated-Content Komponente der Organic.Lingua Projektpartner beschrieben und ein Ausblick für die Zukunft gegeben.

## 1.4 Definitionen

**CSS:** Cascading Style Sheets (CSS) ist eine Stylesheet-Sprache, mit der das Layout eines HTML-Dokumentes beschrieben wird.

**DOM:** Document Object Model (DOM) ermöglicht den einfachen Zugriff auf HTML- und XML-Dokumente.

**Harvesting:** Unter Harvesting wird das Sammeln (Ernten) von Ressourcen bzw. die Metadatenbeschreibung dieser Ressourcen verstanden.

**HTML:** HyperText Markup Language (HTML) ist die Markup-Sprache für Webseiten.

**IEEE LOM:** Institute of Electrical and Electronic Engineers Learning Object Metadata. IEEE LOM stellt einen international anerkannten Standard für die Beschreibung von Lernobjekten dar.

**LOM:** Learning Object Metadata (siehe IEEE LOM).

**OA:** Organic Agriculture. Bezeichnet die ökologische und/oder biologische Landwirtschaft.

**OAI-PMH:** Open Archives Initiative Protocol für Metadata Harvesting. Das Protokoll wurde von Open Archives Initiative entwickelt. Es dient zum Sammeln (Harvesting) von Metadaten eines Archivs.

**UGC:** User-Generated-Content.

**UI:** User Interface.

**POS:** part-of-speech. Ein Part-of-Speech-Tagger bestimmt die Wortart bzw. Wortklasse eines Wortes innerhalb eines Textes (lexikalische Kategorie).

**Repository:** Ein Repository ist eine Datenbank bzw. ein Verzeichnis und dient zur Speicherung digitaler Ressourcen und deren Metadaten.

**XML:** eXtensilbe Markup Language.

## 2 Forschungsgrundlage

Dieses Kapitel befasst sich mit der Forschungsgrundlage der zu implementierenden User-Generated-Content Komponente für das Organic.Edunet Web-Portal. Der Fokus dieses Kapitels bezieht sich auf die Themenbereiche User-Generated-Content, Metadaten, Information-Retrieval Technologien, um Metadaten aus Lernressourcen zu extrahieren, und das Bereitstellen von Metadaten.

### 2.1 User-Generated-Content

Wie in Abschnitt 1.2 erläutert, soll der Benutzer des Organic.Edunet Web-Portals nicht mehr nur reiner Konsument von Lerninhalten sein, sondern aktiv bei der Gestaltung der Lerninhalte mitwirken können. Aus diesem Grund wird in diesem Abschnitt der Begriff User-Generated-Content erläutert. Zudem werden die verschiedensten Formen von User-Generated-Content und die Plattformen, auf denen User-Generated-Content angeboten wird, erklärt. Zum Schluss wird auf Motivation der Benutzer und die Verbreitung von User-Generated-Content in der Europäischen Union eingegangen.

Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD, 2007) beschreibt in der Studie „Participative Web: User Generated Content“, dass das „partizipative Web“ auf zunehmend intelligenten Web-Services basiert, die den Benutzer befähigen, Inhalte kollaborativ zu erstellen, zu bewerten und zu teilen. Außerdem ist das Internet ein wichtiger Bestandteil im Leben der Menschen, die sich durch Beitragen von User-Generated-Content, verwirklichen (OECD, 2007).

Folglich beschreibt der Begriff User-Generated-Content verschiedene Arten von Medien-Inhalten, die von Endbenutzern erstellt wurden und öffentlich im Internet zugänglich sind.

Die Nutzung und Verbreitung von Web 2.0 Anwendungen, wie z. B. Wikis, Weblogs, sozialen Netzwerken usw., ist in den letzten Jahren stetig gestiegen (Sigala, 2007). Unter dem Begriff Web 2.0 werden interaktive und kollaborative

Anwendungen im Web zusammengefasst (O'REILLY, 2007). Die zunehmende Verbreitung von Web 2.0 Anwendungen bringt enorme Auswirkungen und Veränderungen mit sich, wie Internetbenutzer Informationen suchen, finden, entwickeln und konsumieren (Sigala, 2007). Des Weiteren hat sich durch die kollaborative Nutzung von Web 2.0 Anwendungen die Entwicklung von neuem Wissen geändert (Sigala, 2007).

User-Generated-Content ist ein wesentlicher Bestandteil des Web 2.0 und wird weiterhin eine wesentliche Rolle im Web spielen (Rennie, 2007). Wikipedia, YouTube, MySpace, Mendeley, Flickr und Yahoo! Answers sind bekannte Web-Plattformen, die dem Endbenutzer erlauben, verschiedenen Arten von User-Generated-Content beizutragen und zu teilen. Diese Web-Plattformen, die den Benutzer befähigen, User-Generated-Content zu erzeugen, zu konsumieren und zu teilen, werden auch „Social Media“ genannt. Kaplan und Haenlein (2010) definieren „Social Media“ als eine Gruppe von internetbasierten Anwendungen, die auf den ideologischen und technologischen Grundlagen des Web 2.0 aufbauen und das Erzeugen und Austauschen von User-Generated-Content erlauben. Nach Schürig (2010) wird der Begriff Web 2.0 immer mehr durch den Begriff Social Media ersetzt.

Inhalte im Internet müssen nach der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD, 2007) drei Kriterien erfüllen, um als User-Generated-Content bezeichnet zu werden. Erstens muss der Inhalt entweder auf öffentlich zugänglichen Webseiten oder in einem sozialen Netzwerk publiziert werden. Dieser Inhalt muss einer ausgewählten Gruppe von Personen zugänglich sein. Zweitens setzt die Erstellung des Inhalts oder die Anpassung einer bestehenden Ressource ein gewisses Maß an kreativer Arbeit durch den Benutzer voraus. Drittens soll der vom Benutzer erstellte Inhalt außerhalb professioneller Routinen erstellt werden, das heißt der Inhalt soll keinem kommerziellen Zweck dienen und der Benutzer soll nicht mit der Absicht eines Gewinnes oder einer Vergütung den Inhalt erstellen.

Das Internet ermöglichte mehreren Personen, Informationen zu teilen und kollaborativ zu erstellen, weil die Barriere für die Inhaltserstellung geringer wurde (Rennie, 2007). Laut Kaplan und Haenlein (2010) schaffte es das Web 2.0 und somit auch der User-Generated-Content durch drei Einflussfaktoren

zum Durchbruch. Diese drei Einflussfaktoren werden in den nächsten drei Absätzen genauer erklärt.

Der erste Einflussfaktor bestand in den technischen Veränderungen, wie z. B. der Erhöhung der Verfügbarkeit von Breitbandinternet und der verbesserten Hardwarekapazität (Kaplan & Haenlein, 2010). Sigala (2007) ergänzt, dass der Internetbenutzer keine bzw. nur mehr minimale technische Fähigkeiten haben muss, um im Internet Inhalte zu publizieren. Aus diesem Grund kann sich der Internetbenutzer auf die Kerntätigkeiten fokussieren und die Inhaltspublizierung wird nicht durch technische Barrieren erschwert. Diese Kerntätigkeiten des Endbenutzers umfassen die Erstellung von neuem Inhalt und kollaborative Aufgaben.

Der zweite Einflussfaktor nach Kaplan und Haenlein (2010) ist der wirtschaftliche Treiber, wie z. B. die erhöhte Verfügbarkeit von Werkzeugen (Plattformen) für die Erstellung von User-Generated-Content.

Der dritte Einflussfaktor, der User-Generated-Content zum Durchbruch verhalf, liegt in sozialen Faktoren, wie z. B. der Entstehung einer neuen Generation von Internetbenutzern, den sogenannten „Digital Natives“ (Prensky, 2001) bzw. „Screenagers“ (Rushkoff, 1999). „Digital Natives“ sind nach Prensky (2001) Personen, die ihr ganzes Leben lang von digitalen Technologien wie Computer, Videospiele, Handys, Video-Kameras, MP3-Player usw. umgeben sind und schon in jungen Jahren begonnen haben, damit zu arbeiten. Prensky (2001) führt weiter aus, dass sich die Denkmuster und wahrscheinlich die Gehirnstrukturen von „Digital Natives“ verändert haben. Rushkoff (1999) beschreibt Kinder, die in die elektronische Welt der Computer und TV-Monitore geboren wurden, als „Screenagers“. (Kaplan & Haenlein, 2010)

Nach Rennie (2007) ist der phänomenale Aufstieg von User-Generated-Content auf das Ergebnis eines alternativen Systems für die Produktion von Inhalten zurückzuführen. Dieses alternative System zur Inhaltserstellung besitzt keine Beschränkungen, wie z. B. physische Buchkapitel. Außerdem kann eine größere Anzahl von Personen Inhalte beitragen. Dieses kollaborative Beitragen von Inhalten durch Einzelpersonen war in der Zeit von Massenmedien, Zeitungen und Rundfunk nur schwer oder gar nicht möglich. (Rennie, 2007)

Schmidt (2009, S. 9) führt ergänzend zum Durchbruch von User-Generated-Content an, dass das Internet in den letzten Jahren immer stärker mit dem täglichen Lebensablauf der Gesellschaft verknüpft wurde. Menschen und Internet sind in den verschiedensten Situationen miteinander sehr vernetzt. Das Internet wird für persönliche, berufliche und schulische Zwecke genutzt (Schmidt, 2009, S. 9).

Bruns (2008) beschreibt in „Blogs, Wikipedia, Second Life, and Beyond: From Production to Prousage“ den Prozess der Inhaltserstellung durch den End-Benutzer. Die Erstellung von Inhalten durch den End-Benutzer ist ein iterativer und evolutionärer Prozess. In diesem Inhaltserstellungsprozess wirkt eine große Anzahl von Personen (Community) mit. Diese Personen machen eine Reihe von in der Regel sehr kleinen und inkrementellen Änderungen am Inhalt. Durch diesen inkrementellen Prozess wird die Qualität des Inhalts mit jedem kleinen Iterationsschritt verbessert. Trotz des schrittweisen Hinzufügens von Wissen zu einer Knowledge-Base kann dieser inkrementelle Inhaltserstellungsprozess durch den End-Benutzer die Geschwindigkeit von klassischen, industriellen Inhaltserstellungsprozessen übertreffen. (Bruns, 2008, S. 1)

Dieser kollaborative, iterative Inhaltserstellungsprozess nach Bruns (2008) wird im folgenden Abschnitt mit dem industriellen Inhaltserstellungsprozess verglichen.

### **2.1.1 Nutzergetriebene Inhaltserstellung versus industrielle Inhaltserstellung**

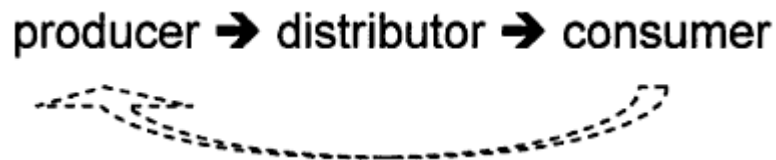
In diesem Abschnitt wird der industriell generierte Inhalt dem nutzergenerierten Inhalt gegenübergestellt und ihr Prozess bis zum fertigen „Produkt“ beschrieben. Dieser Vergleich stützt sich auf die Arbeit von Bruns (2008) mit dem Titel „Blogs, Wikipedia, Second Life, and Beyond: From Production to Prousage“.

Die industrielle Methode der Inhaltserstellung erstellt ein Produkt. Ein Produkt ist nur eine mögliche Zusammensetzung von Informationen, Wissen und kreativer Arbeit. Diese Zusammensetzung muss aber nicht zum am besten geeigneten Produkt führen. Bruns (2008, S. 2) sieht in der Bezeichnung



„Produkt“ dahingehend ein Problem, dass ein Produkt zwangsläufig ein spezifisches Ergebnis eines Prozesses impliziert. End-Benutzer können diese Produkte kaufen und verwenden. Es besteht aber keine Möglichkeit für den Endbenutzer, das Produkt zu erweitern oder bei der Gestaltung mitzuwirken. (Bruns, 2008, S. 2)

Bei traditioneller Inhaltserstellung haben Verlage Massenmedien, wie z. B. Print-Medien, produziert und vermarktet, welche durch den Kunden nur konsumiert werden konnten, ohne aktiv bei der Gestaltung des Inhalts beizutragen. Außerdem war es für den Konsumenten schwer, aktiv mit dem Ersteller des Inhaltes zu kommunizieren. Dieser Sachverhalt ist in Abbildung 2-1 dargestellt.

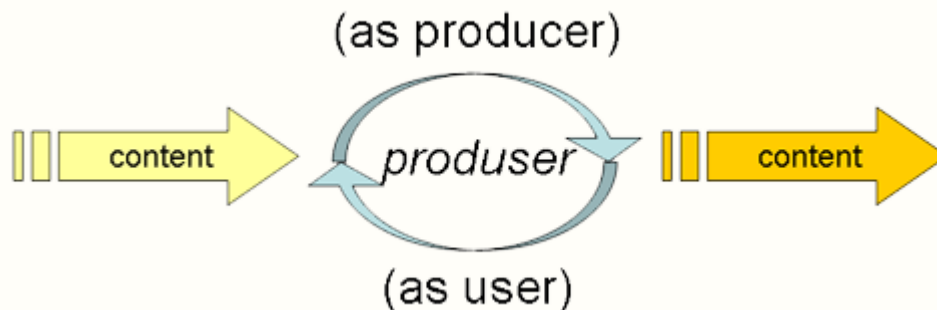


**Abbildung 2-1: Limitierte Feedback-Möglichkeiten bei der industriellen Inhaltserstellung (Bruns, 2008, S. 11)**

Für den Prozess der Inhaltserstellung durch Endbenutzer definiert Bruns (2008) einen neuen Begriff namens „Produusage“. Bruns (2008, S. 2) definiert den Begriff „Produusage“ als kollaborative, laufende Erstellung und Erweiterung von einem existierendem Inhalt und Wissen in einer Community. Das Ziel bei „Produusage“ liegt darin, die Qualität des Inhaltes zu verbessern. Benutzer sind zugleich Produzenten und Konsumenten von einer gemeinsamen Wissensbasis. Bruns (2008) definiert diese hybride, neue Rolle des End-Benutzers als „Produser“. Durch das Social Web verwischen die Grenzen zwischen der Rolle des Konsumenten und Produzenten, zwischen Lesen und Schreiben, zwischen öffentlich und privat und zwischen Schreiben und Programmierung (Sigala, 2007).

In der Abbildung 2-2 wird die Rolle des „Produsers“ im kollaborativen Inhaltserstellungsprozess dargestellt. Die Wertschöpfungskette des „Produsers“ (siehe Abbildung 2-2) unterscheidet sich deutlich von der Wertschöpfungskette der traditionellen Inhaltserstellung (siehe Abbildung 2-1). Durch die neue hybride Rolle „Produser“ verschwinden die Rollen der traditionellen Inhaltserstellung, wie

z. B. Produzent, Händler oder Konsument. Außerdem gibt es kein lineares Vorgehen bei der Inhaltserstellung, vielmehr ist es eine endlose Kette von inkrementellen Benutzeraktivitäten. Diese Benutzer nehmen zwei Rollen ein - die Rolle des Produzenten und des Konsumenten. Der Benutzer in der Rolle des Produzenten erweitert und verbessert schrittweise den Inhalt. Der Produzent erzeugt immer wieder eine neue Version des Inhaltes. Die Version des Inhaltes ist aber nur vorübergehend und ist bereit für die weitere Entwicklung. Folglich endet der Prozess der Inhaltsentwicklung nie und es existiert nie ein fertiges Produkt, sondern es entstehen nur temporäre Produkte. Der Benutzer in der Rolle des Konsumenten nutzt vorhandene Ressourcen. (Bruns, 2008, S. 20f)



**Abbildung 2-2: Die Rolle des „Producers“ (Bruns, 2007)**

Das Werkzeug für die kollaborative Inhaltserstellung stellt Social Software in ihren verschiedensten Ausprägungen dar. Der Einsatz von Social Software alleine kann aber nicht garantieren, dass Inhalte kollaborativ erstellt und geteilt werden (Bruns, 2008, S. 2). Jedoch kann eine optimale Umgebung geschaffen werden, die den End-Benutzer animiert, Inhalte beizusteuern. Die Motivationsfaktoren für Personen, um Inhalte ohne kommerziellen Nutzen einzubringen, werden im Abschnitt 2.1.3 genauer erläutert.

### **Bedeutung für die vorliegende Arbeit**

Ein Ziel dieser Arbeit liegt darin, den „industriellen“ Inhaltserstellungsprozess im Organic.Edunet Web-Portal um einen benutzergetriebenen Inhaltserstellungsprozess zu erweitern (siehe Abschnitt 1.2).

## 2.1.2 Arten von User-Generated-Content und deren Einsatzgebiete

In diesem Abschnitt werden die verschiedenen Arten von User-Generated-Content erklärt.

Die Art der Web-Plattform und der User-Generated-Content Typ sind sehr stark miteinander verknüpft. Zum Beispiel werden Posts in Blogs verfasst und Videos werden auf Online-Video-Sharing-Plattformen verbreitet. Außerdem können auf User-Generated-Content-Plattformen, wie z. B. auf Social-Networking-Webseiten, mehrere Arten von User-Generated-Content erzeugt werden, wie z. B. Musik, Videos, Kommentare usw.

Es gibt unter anderem folgende Arten von User-Generated-Content (OECD, 2007, S. 34):

- **Texte**

Es können verschiedenste Arten von Texten neu erstellt oder bestehende Text erweitert werden. Beispiele für soziale Plattformen, in denen verschiedenste Texte verfasst werden können, sind Quizilla<sup>4</sup> und Google-Docs<sup>5</sup>.

- **Bilder**

Digitale Fotos und Bilder können vom Benutzer erstellt, modifiziert und auf eine Onlineplattform hochgeladen werden. Auf den Onlineplattformen können diese Bilder geteilt, bewertet und online modifiziert werden. Benutzererstellte Bilder werden auf Webseiten wie Flickr<sup>6</sup> und Picasa<sup>7</sup> hochgeladen und mit anderen Personen geteilt.

---

<sup>4</sup> <http://quizilla.teennick.com/>

<sup>5</sup> <http://docs.google.com/>

<sup>6</sup> <http://www.flickr.com/>

<sup>7</sup> <http://picasa.google.com/>

- **Musik und Audio**

Eigene Musik und Audioaufzeichnungen können durch verschiedenste Plattformen publiziert, geteilt, kommentiert und bearbeitet werden. Myspace<sup>8</sup> ermöglicht dem Benutzer, Musikdateien hochzuladen und somit mit der Community zu teilen.

- **Videos**

Heutzutage besitzen schon sehr viele Personen ein digitales Gerät (z. B. Mobiltelefone und Digitalkameras), um Videos aufzunehmen. Diese Videos werden von Personen weiterverarbeitet und können auf Videoportalen (z. B. YouTube<sup>9</sup>) hochgeladen werden und somit mit anderen Personen geteilt werden.

- **Bürger-Journalismus**

Normale Bürger übernehmen die journalistische Berichterstattung über aktuelle Ereignisse und stellen sie online zur Verfügung. Beispiele für bürgerlichen Journalismus sind „Ohmynews International“<sup>10</sup> und Wikinews<sup>11</sup>.

- **Lehr- und Lerninhalte**

Inhalte, die im Kontext von Schulen, Universitäten oder mit dem Ziel der Weiterbildung erstellt wurden, sind öffentlich zugänglich. Beispiele für öffentlich zugängliche pädagogische Inhalte sind MIT OpenCourseWare<sup>12</sup>, Wikibooks<sup>13</sup> und Mathe-Online<sup>14</sup>.

---

<sup>8</sup> <http://www.myspace.com/>

<sup>9</sup> <http://www.youtube.com/>

<sup>10</sup> <http://international.ohmynews.com/>

<sup>11</sup> <http://www.wikinews.org/>

<sup>12</sup> <http://ocw.mit.edu>

<sup>13</sup> <http://www.wikibooks.org/>

<sup>14</sup> <http://www.mathe-online.at/>

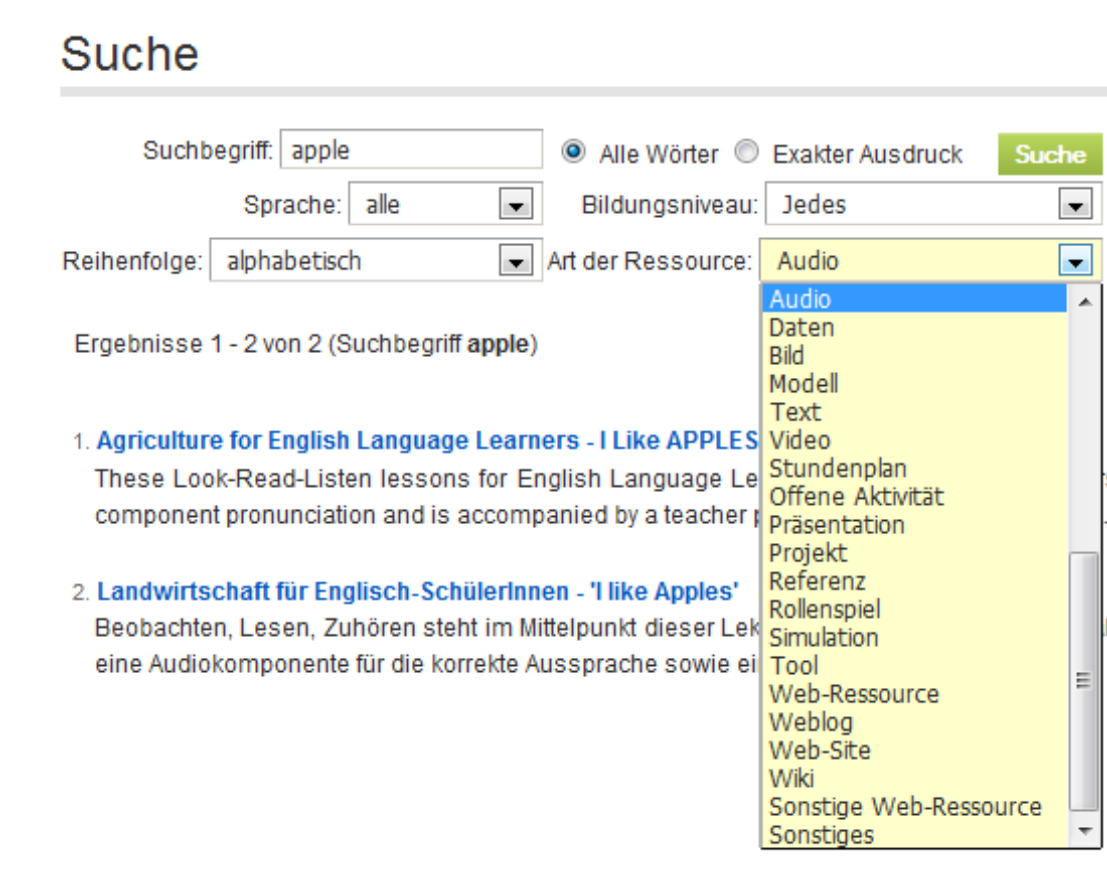
- **Virtueller Inhalt**

Ein Inhalt wird im Kontext von virtuellen Welten erzeugt und modifiziert. Ein Beispiel für eine virtuelle Welt stellt die Plattform „Second Life“<sup>15</sup> dar.

Einen guten Überblick über die verschiedenen Formen von User-Generated-Content und die User-Generated-Content Plattformen bietet die Arbeit von Kolbitsch und Maurer (2006) „The Transformation of the Web: How Emerging Communities Shape the Information we Consume“. Dabei gibt die Arbeit (Kolbitsch & Maurer, 2006) einen Überblick über folgende Services: Blogs, Wikis, soziale Netzwerke, Files-Sharing-Tools und Podcasting.

### **Bedeutung für die vorliegende Arbeit**

Das Organic.Edunet Web-Portal bietet eine große Anzahl verschiedener Arten von Lernressourcen an, die von Experten erstellt werden (siehe Abbildung 2-3).



<sup>15</sup> <http://secondlife.com/>

### **Abbildung 2-3: Arten von Lernressourcen im Organic.Edunet Web-Portal**

Diese verschiedenen Formen von User-Generated-Content, die in diesem Abschnitt beschrieben werden, sollen in Zukunft durch den Organic.Edunet Benutzer dem Organic.Edunet Web-Portal als Lernressource hinzugefügt werden können.

Das Hinzufügen von benutzergenerierten Lernressourcen soll durch die User-Generated-Content Komponente realisiert werden. Die User-Generated-Content Komponente wird in Kapitel 3 genauer beschrieben. Dabei dient das Organic.Edunet Web-Portal nicht als Speicherort der benutzergenerierten Lernressourcen, sondern der Benutzer kann mit Hilfe der User-Generated-Content Komponente Lernressourcen annotieren und somit der Organic.Edunet Community zur Verfügung stellen, wobei die Annotierung der Lernressource auch wieder eine Art von User-Generated-Content darstellt.

### **2.1.3 Motivation der Benutzer**

In diesem Abschnitt werden Motivationsfaktoren angeführt, die den End-Benutzer motivieren, ohne kommerziellen Nutzen User-Generated-Content auf verschiedensten Plattformen beizutragen.

Die Motivation der Inhaltsersteller scheint nach Nov (2007) für die Aufrechterhaltung von Wikipedia und andere kollaborative User-Generated-Content Web-Plattformen kritisch zu sein. Der Grund dafür liegt darin, dass der Inhalt aus eigenem Antrieb von Personen erstellt wird, die ohne monetäre Belohnung ihre Zeit und ihr Talent anbieten und in weiterer Folge einsetzen. Aufgrund dieser fehlenden monetären Belohnung heißt es aber nicht, dass sich ehrenamtlich tätige Personen keine Gegenleistung erwarten. Diese nicht monetären Gegenleistungen werden von einer großen Menge wissenschaftlicher Arbeiten unter dem Begriff Motivationsfaktoren von ehrenamtlichen Personen analysiert. In den folgenden Absätzen werden exemplarisch einige Motivationsfaktoren von verschiedenen wissenschaftlichen Arbeiten aufgelistet.

Nach der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD, 2007, S. 10) gibt es sechs Faktoren, die den Benutzer motivieren, neue Inhalte im Internet zu publizieren und bestehende Inhalte zu erweitern:

- Verbindung mit gleichartigen Personen
- Erlangen eines gewissen Maßes an Ruhm
- Erlangen von Berühmtheit
- Berühmtheit
- Prestige
- der Wunsch der Benutzer, sich selbst auszudrücken

Nov (2007) untersucht in seiner Studie „What motivates wikipedians?“, warum Menschen Inhalte zu Wikipedia beitragen. Er führte eine empirische Studie, die auf die sechs Motivationsfaktoren von Clary et al. (1998) aufbaut, durch. Die sechs Motivationsfaktoren von Clary et al. (1998, zitiert nach (Matthias, 2010, S. 71) und (Nov, 2007)) sind:

- **Value**

Engagement als Ausdruck von persönlichen Werten. Der Wille, eigenen altruistischen und humanitären Werten Ausdruck zu verleihen.

- **Career**

Freiwilliges Engagement kann die Möglichkeit bieten, berufliche Vorteile zu erlangen, wie z. B. Vorbereitung für eine neue Karriere oder die Aufrechterhaltung von berufsrelevanten Fähigkeiten. Am Beispiel Wikipedia können Mitwirkende anhand der geschriebenen Artikel zukünftigen Arbeitgebern ihr Wissen und ihre schriftliche Ausdrucksfähigkeit signalisieren.

- **Social**

Durch freiwilliges Engagement haben die mitwirkenden Personen die Möglichkeit, mit ihren Freunden zusammen zu arbeiten. Die Aktivität des Freiwilligen kann von der Community auch positiv bewertet werden.

- **Understanding**

Durch das freiwillige Mitwirken haben die Personen die Möglichkeit, neue Dinge zu lernen und ihr bestehendes Wissen, Fertigkeiten und Fähigkeiten werden durch die freiwillige praktische Übung erweitert. Dadurch versteht der Freiwillige die Thematik besser, bei der er kollaborativ mitwirkt und auch die Menschen, mit denen er kollaborativ tätig ist.

- **Protective**

Dieser Motivationsfaktor beinhaltet die Überwindung der eigenen Probleme, von Einsamkeit oder Schuldgefühlen. Beispielsweise schützt der Motivationsfaktor „Protective“ das Ego der mitwirkenden Personen, indem die Schuldgefühle reduziert werden, wenn der Mitwirkende glücklicher ist als die Anderen. Außerdem verletzt das direkte Ansprechen der Probleme das Ego der mitwirkenden Personen nicht folgenscher.

- **Enhancement**

Der Motivationsfaktor „Enhancement“ steht in Beziehung mit dem Motivationsfaktor „Protective“. Der Unterschied der zwei Faktoren liegt darin, dass „Enhancement“ die positiven Faktoren stärkt, um das Ego aufzubauen, wohingegen „Protective“ negative Einflussfaktoren für das Ego eliminiert. Kurz gefasst geht es in der Kategorie „Enhancement“ darum, dass sich Mitwirkende besser und wichtiger fühlen.

Nov (2007) ergänzte diese sechs Motivationsfaktoren von Clary (1998) um die Faktoren „Fun“ und „Ideology“, da diese Motivationsfaktoren sehr oft in den empirischen Studien über die Motivation der Open-Source-Programmierung vorkommen, wie zum Beispiel in den Studien von Stewart und Gosain (2006) und Hars und Ou (2001). Das Ergebnis der empirischen Studie von Nov (2007) wird in der Abbildung 2-4 ersichtlich. Durchschnittlich investierten die Wikipedia-Autoren 8,27 Stunden pro Woche, um Wikipedia-Artikel anzupassen oder neue zu erstellen. Die Top-Motivationsfaktoren waren Spaß („Fun“) und Ideologie („Ideology“), während soziale, schützende und karriere-technische



Motivationsfaktoren keine starke Motivation für die Mitwirkung bei Wikipedia darstellten.

Motivation	Mean
Fun	<b>6.10</b> (1.15) [0.322***]
Ideology	<b>5.59</b> (1.71) [0.110]
Values	<b>3.96</b> (1.55) [0.175**]
Understanding	<b>3.92</b> (1.48) [0.296***]
Enhancement	<b>2.97</b> (1.39) [0.313***]
Protective	<b>1.97</b> (1.05) [0.306***]
Career	<b>1.67</b> (0.94) [0.185**]
Social	<b>1.51</b> (0.92) [0.027]

\*significant at 0.05 level  
 \*\*significant at 0.01 level  
 \*\*\*significant at 0.001 level

**Abbildung 2-4: Motivationsniveau der Wikipedia-Autoren in Korrelation mit der Beitragshöhe. Die Standardabweichung in runden Klammern. Der Pearson-Korrelationskoeffizient in eckigen Klammern. (Nov, 2007)**

Yang und Lai (2010) analysierten ebenfalls die Motivationsfaktoren der Wikipedia-Autoren, um zu verstehen, warum Menschen unentgeltlich ihr Wissen auf Wikipedia teilen. Das Ergebnis der Studie von Yang und Lai (2010) zeigte, dass das interne Selbstverständnis den größten Motivationsfaktor der Wikipedia-Autoren darstellt, um Wissen zu teilen. Somit deckt sich die Studie von Yang und Lai (2010) mit der Studie von Nov (2007), bei der die Ideologie ebenfalls zu den Top-Motivationsfaktoren gehörte.

### **Bedeutung für die vorliegende Arbeit**

Die oben angeführten Studien lassen den Schluss zu, dass Ideologie und Spaß die größten Motivationsfaktoren sein werden, damit Organic.Edunet Benutzer ihr Wissen und ihre Zeit in die Organic.Edunet Web-Plattform investieren.

Folglich ist es nicht notwendig, besondere Anreize für Organic.Edunet Benutzer anzubieten, sodass sie aktiv bei der Gestaltung des Lernressourcenangebotes mitwirken.

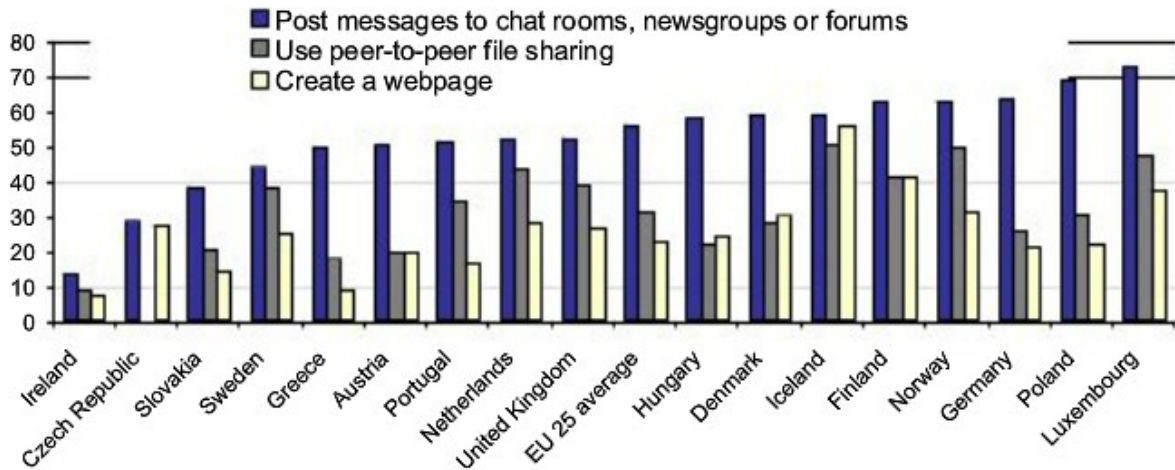
Im nächsten Abschnitt wird auf die Verbreitung von User-Generated-Content eingegangen. Diese Verbreitung lässt Rückschlüsse auf die Beteiligung der Organic.Edunet Benutzer ziehen, die benutzererstellte Metadaten beitragen.

### **2.1.4 Verbreitung von User-Generated-Content in der Europäischen Union**

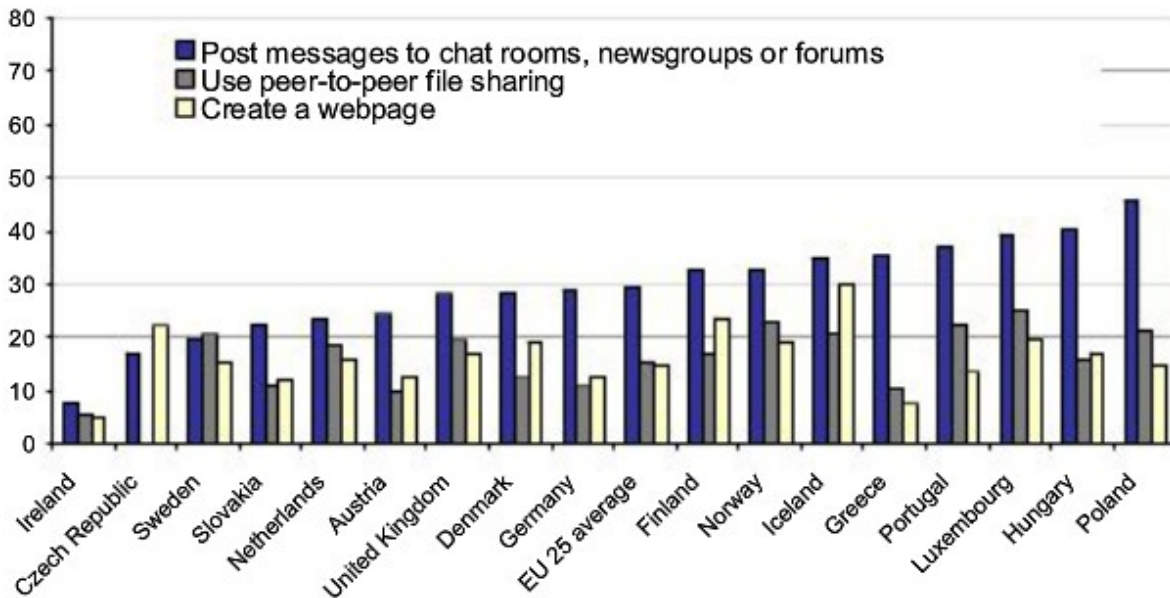
In diesem Abschnitt wird näher auf die Verbreitung von User-Generated-Content in der Europäischen Union eingegangen.

Inhaltserstellung im Internet ist eine sehr beliebte Aktivität bei jungen Altersgruppen. Die Abbildung 2-5 und Abbildung 2-6 zeigen, dass das Schreiben von Beiträgen in Chatrooms, Newsgroups oder Foren, Verwendung von Peer-to-Peer-Filesharing und das Erstellen von Webseiten unter Internetbenutzern bereits weit verbreitet und beliebt ist. Dabei ist jedoch darauf zu achten, dass Filesharing nicht unbedingt User-Generated-Content sein muss. In Ländern wie Finnland, Norwegen, Island, Portugal, Luxemburg, Ungarn und Polen, waren im Jahr 2005 rund ein Drittel aller Internet-Nutzer im Alter von 16 bis 74 Jahren mit dem Teilen von Inhalten im Internet beschäftigt (siehe Abbildung 2-6). (OECD, 2007)

Jüngere Altersgruppen sind aktivere Content-Ersteller im Internet. In Ländern wie Ungarn, Dänemark, Island, Finnland, Norwegen, Deutschland, Polen und Luxemburg (in aufsteigender Reihenfolge), haben im Jahr 2005 zwischen 60 und 70 Prozent der Internetnutzer im Alter von 16 bis 24 Jahren Beiträge in Chatrooms, Newsgroups oder Foren geschrieben (siehe Abbildung 2-5). Ein Drittel aller Internetbenutzer hat im Durchschnitt im Jahr 2005 in der Altersgruppe von 16 bis 24 Jahren eine Webseite erstellt (siehe Abbildung 2-5). (OECD, 2007)



**Abbildung 2-5: Prozentueller Anteil der Internetbenutzer im Alter von 16 – 24 Jahren, die User-Generated-Content erzeugen, im Jahr 2005 (OECD, 2007, S. 20)**



**Abbildung 2-6: Prozentueller Anteil der Internetbenutzer im Alter von 16 – 74 Jahren, die User-Generated-Content erzeugen, im Jahr 2006 (OECD, 2007, S. 21)**

Ein weiteres Indiz für die weite Verbreitung von User-Generated-Content im Internet zeigt die Studie der „Kansas State University“ (2008). Diese Studie bringt vor, dass mehr als 80 % der hochgeladenen Videos benutzererstellte Videos sind und keinem kommerziellen Zweck dienen.

### **Bedeutung für die vorliegende Arbeit**

Aus diesen Statistiken über die Verbreitung von User-Generated-Content lässt sich das große Potenzial von User-Generated-Metadaten im Organic.Edunet Web-Portal erschließen. Angenommen, die 60 Prozent (OECD, 2007) der Internetbenutzer zwischen 16 und 24 Jahren, die Beiträge in Chatrooms, Newsgroups oder Foren schreiben, erklären sich auch bereit, Annotierungen von Lernressourcen zu erweitern, zu verbessern und zu übersetzen. Diese Annahme lässt den Schluss zu, dass 200 Unique-Organic.Edunet-Benutzer pro Tag (siehe Abschnitt 1.1.3) 120 Metadaten von Lernressourcen hinzufügen, verbessern oder übersetzen. Wird weiters angenommen, dass diese 120 Organic.Edunet Benutzer pro Tag genau einen Metadatenatz modifizieren oder eine neue Lernressource annotieren, so geht hervor, dass pro Jahr 43.800 Lernressourcenannotierungen verbessert, erweitert, übersetzt oder hinzugefügt werden. Bei derzeit 7.000 abrufbaren Lernressourcen im Organic.Edunet Web-Portal (siehe Abschnitt 1.1.3) führt der zukünftige Einsatz der User-Generated-Content Komponente (siehe Kapitel 3) zu einer Vervielfachung der annotierten Lernressourcen im Web-Portal.

## **2.2 Limitierungen von User-Generated-Content**

Grundsätzlich liegt das Problem von User-Generated-Content darin, dass benutzererstellte Inhalte unstrukturiert und nicht annotiert sind. Folglich sind benutzererstellte Inhalte schwer auffindbar und bleiben zum Teil ungenutzt. Weiters werden Lernressourcen für gewöhnlich für kleine Gruppen erstellt und diese Gruppen tauschen die Lernressourcen untereinander nicht aus. Die Probleme der nicht genutzten Lernressourcen und der nicht gegebenen Interoperabilität zwischen zwei Gruppen, die Lernressourcen erstellen, kann mit dem Einsatz von standardisierten Metadaten gelöst werden. Der folgende Abschnitt 2.3 befasst sich aus diesem Grund mit der Wichtigkeit von Metadaten, Metadatenstandards und der Metadatenerstellung.

## 2.3 Metadaten

Bargmeyer und Gillmann (2000) definieren Metadaten als Daten über Daten. Obwohl diese Definition kurz und leicht zu merken ist, stellt diese Definition eine nicht sehr präzise dar. Metadaten sind Daten, die verwendet werden, um andere Daten zu beschreiben. Jedoch ist es nicht einfach, Metadaten zu identifizieren, weil auf den ersten Blick nicht gesagt werden kann, ob es sich um Daten oder Metadaten handelt. (Bargmeyer & Gillman, 2000)

Nach Bargmeyer und Gillman (2000) wird von den nationalen und internationalen Standard-Communities viel Arbeit investiert, um einen Konsens über Metadaten in verschiedenen Kontexten zu erreichen und in weiterer Folge Standardisierungen von Metadaten zu definieren.

Es existieren zwei bekannte Metadaten-Initiativen, die Standards für Metadaten entwickelten. Erstens die „*Dublin Core Metadata Initiative (DCMI)*“ und zweitens die „*Institute for Electrical and Electronics Engineers (IEEE) Learning Object Metadata (LOM) Working Group*“ (Duval, Hodgins, Sutton, & Weibel, 2002). Diese zwei Metadaten-Standards „Dublin Core“ und „IEEE Standard Learning Object Metadata“, die aus diesen zwei Gruppen hervorgingen, werden in den folgenden Absätzen erklärt. Sie sind die zwei am häufigsten verwendeten Sammlungen von Konventionen zur Beschreibung von Lernressourcen.

- **Dublin Core**

Dublin Core ist eine Sammlung von Elementen, die eine Ressource beschreiben. Dublin Core ist eine Sprache, um eine Menge von Aussagen über eine bestimmte Ressource zu treffen. Das Dublin Core besteht aus 15 Kernelementen. Diese 15 Kernelemente sind dc:title, dc:creator, dc:subject, dc:description, dc:publisher, dc:contributor, dc:date, dc:type, dc:format, dc:identifier, dc:source, dc:language, dc:relation, dc:coverage und dc:rights. Der Aufbau einer Dublin Core Aussage ist in der Abbildung 2-7 dargestellt. Ein Dublin Core Satz besteht aus drei Teilen, nämlich einem Subjekt, einem Prädikat und einem Objekt. Das Subjekt ist die Ressource, die durch Dublin Core näher beschrieben werden soll. (Baker, 2000)

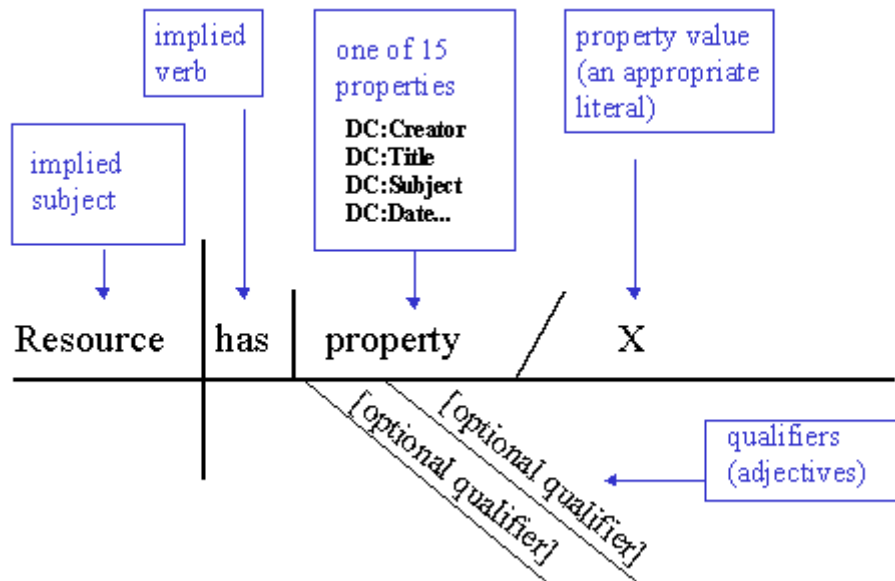


Abbildung 2-7: Aufbau eines typischen Dublin Core Satzes (Baker, 2000)

- **IEEE-Standard Learning Object Metadata**

Nach Lucke (2011) hat sich für die Beschreibung von Lernressourcen im Bereich eLearning der IEEE-Standard Learning Object Metadata (LOM) durchgesetzt. Der IEEE LOM Standard umfasst 58 verschiedene Attribute, die auf neun Hauptkategorien aufgeteilt sind. Diese neun Hauptkategorien beschreiben verschiedene Aspekte der Lernressource, wie z. B. inhaltliche, technische, didaktische und rechtliche Aspekte. Durch die Beschreibung der Lernobjekte werden diese leichter auffindbar und wiederverwendbar. Damit die Beschreibung der Lernressource keine isolierte Einheit darstellt, können auch mit Hilfe des IEEE LOM Standards Beziehungen zwischen zwei Lernobjekten annotiert werden. (Lucke, 2011)

Es existieren sechs definierte Beziehungen zwischen zwei Lernobjekten im IEEE LOM Standard:

- isPartOf, hasPart
- isVersionOf, hasVersion
- isFormatOf, hasFormat

- references, isReferencedBy
- isBasedOn, isBasisFor
- requires, isRequiredBy

In den folgenden Absätzen wird der Einsatz von IEEE LOM in verschiedenen Arbeiten gezeigt.

Meyer (2011) beschreibt in seiner Arbeit den Einsatz von IEEE LOM für die Beschreibung von digitalen Lernmaterialien und die damit verbundenen Vorteile. Der Hauptvorteil liegt nach Meyer (2011) darin, dass die Lernmaterialien durch IEEE LOM wiederverwendbar sind und dadurch Kosten gespart werden können. Weiters führt er die didaktischen Metadaten-Attribute des IEEE LOM Standards als Vorteil an, weil sie den richtigen Einsatz der Lernressource beschreiben und somit der Lernerfolg der Lernenden gesteigert wird.

Pansanato & Fortes (2005) verwenden für die Beschreibung der Lernressourcen in der kollaborativen Lernumgebung CoTeia ebenfalls eine Teilmenge des IEEE LOM Standards, um die Lernressourcen zu annotieren.

Skår Heiberg und Kongsli (2003) beschreiben in ihrer Arbeit den Einsatz von XML und IEEE LOM, um bestehende Provider von Lernressourcen in ein Such-Portal einzugliedern. Für die Realisierung entwickelten sie ein Application-Profile, das aus 21 von 90 Metadatenelementen des IEEE LOM Standards besteht.

Nach Yang, Li und LIN (2000) verbessert der Einsatz des IEEE LOM Standards die Flexibilität, die Interaktivität und die Anpassbarkeit des internetbasierten „Distance-Learning“.

Duval et al. (2002) definieren vier Prinzipien für die Metadaten-Implementierung, sowohl für Metadaten-Schemata als auch für Metadaten-Application Profiles. Diese vier Prinzipien sind „Modularity“, „Extensibility“, „Refinement“ und „Multilingualism“. Diese Prinzipien zeigen die Vorteile von Metadaten und sollen eine Orientierungshilfe für die Entwicklung von praktischen Lösungen für semantische Interoperabilität darstellen. (Duval et al., 2002)

### **Bedeutung für die vorliegende Arbeit**

Die Literatur (Skår et al., 2003; Yang et al., 2000; Pansanato & Fortes, 2005; Dahl & Vossen, 2008; Currier et al., 2004; Barton, Currier, & Hey, 2003; Ochoa & Duval, 2006; Brasher & Mcandrew, 2004; Meyer et al., 2011; Cardinaels, Meire, & Duval, 2005) belegt, dass IEEE LOM gewissermaßen den Standard darstellt, um Lernressourcen im Web zu beschreiben. Aus diesem Grund beruht die Annotierung von Lernressourcen im Organic.Edunet (siehe Abschnitt 1.1) und in weiterer Folge in der User-Generated-Content Komponente (siehe Kapitel 3) auf dem IEEE LOM Standard.

### **2.3.1 Wichtigkeit von Metadaten**

Durch das rasante Wachstum des World Wide Webs und den Online Informationssystemen stehen mehr und mehr Informationen online zur Verfügung und können vom Benutzer einfach abgerufen werden. Diese Explosion von Informationen hat zu einer Informationsüberflutung geführt. Die Menschen haben allerdings keine Zeit, um alles zu lesen, jedoch haben sie Entscheidungen zu treffen, die auf den verfügbaren Informationen basieren. Aus diesem Grund sind Metadaten, wie z. B. Keywords und Textzusammenfassungen, unverzichtbar, um in dieser Menge an Informationen die richtigen zu finden. (Mani & Maybury, 1999, S. IX)

Metadaten sind ein wichtiger Bestandteil in der Informations-Infrastructure und helfen, eine Ordnung in das Chaos des Webs zu bringen (Duval et al., 2002).

Lassila (1998) führt weiters an, dass die Suche im Web ohne semantische Annotationen bzw. Metadaten aufwendig ist. Barton et al. (2004) zeigen in ihrer Arbeit, dass standardisierte Metadaten für das einfache und schnelle Finden von relevanten Ressourcen unverzichtbar sind.

Obwohl viele Lernressourcen im Internet verfügbar sind und diese durch Suchmaschinen indexiert sind, kann die Suche im gesamten Web nach bestimmten Lernressourcen zeitaufwendig sein und möglicherweise nicht das gewünschte Ergebnis liefern. Um dies zu verhindern, ist es von Vorteil, Repositories zu verwenden. Repositories sammeln und verwalten Lernressourcen bzw. deren Metadaten. Diese Lernressourcen in Repositories



werden detailliert und konsistent mit Informationen (Metadaten) beschrieben. Dies ermöglicht dem Benutzer, eine genaue und gezielte Suche durchzuführen und in weiterer Folge relevante Lernressourcen abzurufen. (Currier et al., 2004)

Meyer, Rensing und Steinmetz (2011) halten fest, dass die Produktion von hoch qualitativen Lernressourcen arbeitsintensiv ist und folglich mit hohen Kosten verbunden ist. Daher ist es vorteilhaft, bereits existierende Lernressourcen wiederzuverwenden, um Kosten zu sparen. Meyer et al. (2011) beschreiben, dass jeder Autor daran interessiert sein sollte, Lernressourcen von anderen Autoren wiederzuverwenden. Diese Autoren können natürlich unterschiedlichen Organisationen angehören. Grundsätzlich verwenden verschiedene Organisationen unterschiedliche Konventionen, um Lernressourcen zu beschreiben, sofern sie keinen Metadatenstandard verwenden oder einen Konsens über die Konvention der Metadaten im Vorhinein definiert haben. Im Fall, dass ein Autor eine Lernressource verwenden will, die von einem Autor einer anderen Organisation, erstellt wurde, besteht das Problem der Interoperabilität. Viele Lernplattformen verwenden beliebige Formate für den Inhalt und die Metadatenbeschreibung (Currier et al., 2004). In Folge können Lernressourcen zwischen zwei verschiedenen Lernplattformen nicht ausgetauscht werden. Um das Problem der Interoperabilität zu lösen, sind standardisierte Lernressourcenbeschreibungen und Austauschmechanismen von Nöten. Standardisierte Lernressourcenbeschreibungen werden in Abschnitt 2.3 aufgelistet und ein Austauschmechanismus, wie Lernressourcen zwischen zwei Lernplattformen ausgetauscht werden können, wird in Abschnitt 2.6 erklärt.

## 2.3.2 Metadatenerstellung

Duval et al. (2002) definieren folgende praktische Aspekte für Metadaten-Erstellung und -Verwaltung im Internet:

- **Application Profiles**

Kein Metadata-Element-Set kann die funktionalen Anforderungen jeder Anwendung erfüllen. Aus diesem Grund soll ein Application Profile definiert werden. Ein Application Profile ist ein angepasstes Metadata-Element-Set für eine spezifische Anwendung in einem bestimmten Kontext, das aus mehreren Metadaten-Schemata und/oder eine Teilmenge eines Metadaten-Schemata besteht. Außerdem kann auch das Application-Profile Metadaten-Elemente beinhalten, die auf keinem Metadaten-Standard basieren.

Das eingesetzte Application-Profile in der User-Generated-Content Komponente wird in Abschnitt 3.5 genauer erklärt.

- **Syntax and Semantics**

Die Semantik und die Syntax des Metadaten-Schemas müssen definiert werden. Diese Definition ist notwendig, wenn Metadaten zwischen zwei Communities geteilt werden.

- **Association Models**

Duval et al. (2002) beschreiben drei Möglichkeiten, wie Metadaten mit der Ressource verbunden werden können. Der erste Ansatz sind „Embedded metadata“, wobei die Metadaten direkt in die Ressource eingebunden werden. Dies impliziert, dass die Metadaten zur Zeit der Ressourcen-Erstellung erzeugt werden. Die zweite Möglichkeit stellen „Associated Metadata“ dar. Hier sind die Metadaten in einer eigenen Datei gespeichert, die eng mit der Ressource gekoppelt ist. Die letzte Variante ermöglichen „Third-Party metadata“. Dabei werden die Metadaten in einem eigenem Repository gespeichert und verwaltet.

Die User-Generated-Content Komponente beruht auf der Metadaten Speicherart „Third-Party metadata“. Ein Vorteil von „Third-

Party metadata“ liegt darin, dass die User-Generated-Content Komponente flexibler ist, da der Organic.Edunet Benutzer jede Art von Web-Ressource als Lernobjekt annotieren kann. Eine weiterer positiver Aspekt besteht dadurch, dass die User-Generated-Content Komponente nicht die Verwaltung und Speicherung der Lernressource übernehmen muss. Der Benutzer muss auf bestehende Plattformen zurückgreifen, um Lernmaterialien zu veröffentlichen.

- **Identifying and Naming Metadata Elements: Tokens Versus Labels**

Jedes Element soll eindeutig in einem Set von Elementen durch einen URI identifizierbar sein.

- **Metadata Registries**

Metadaten-Registries sind die zentrale Stelle, bei der Metadaten-Schemata und Application-Profiles gespeichert und verwaltet werden.

- **Completeness of Description**

Bei der Erstellung eines Metadatensatzes sollen alle Metadatenelemente ausgefüllt werden, da dadurch das Auffinden der Ressource erleichtert wird. Der Benutzer kann unterstützt werden, indem die Anwendung einige Metadatenelemente ausfüllt.

- **Mandatory Versus Optional Elements**

Bei Designen eines Metadaten-Standards für fachübergreifende Informationen wird ein hohes Maß an Flexibilität vorausgesetzt. In einer Domäne können bestimmte Metadaten-Elemente sehr wichtig sein, in einer anderen hingegen weniger sinnvoll. Aus diesem Grund sollen nur bestimmte Elemente zwingend definiert werden.

- **Subjective and Objective Metadata**

Grundsätzlich gibt es subjektive und objektive Metadatenelemente, die bei der Metadatenerstellung ausgefüllt werden müssen. Objektive Metadatenelemente sind z. B. der Autor, Erstellungsdatum und andere Attribute, die in einer objektiven Art und Weise bestimmt werden

können. Diese objektiven Metadatenelemente können in den meisten Fällen maschinell erzeugt werden. Subjektive Metadatenelemente sind Attribute, die nicht eindeutig bestimmt werden können. Beispiele für subjektive Attribute sind Keywords, Zusammenfassung des Inhalts der Ressource usw.

- **Automated Generation of Metadata**

Die Metadatenerstellung durch den Menschen ist eine sehr arbeitsintensive Tätigkeit. Aus diesem Grund soll versucht werden, Metadaten-Elemente teilweise automatisch zu generieren.

Lernressourcen-Metadaten können auf viele Arten erstellt werden. Jede dieser Metadatengenerierungsarten hat ihre Vor- und Nachteile. Vorrangig können zwei verschiedene Arten identifiziert werden: erstens ein Top-Down Ansatz, bei dem Metadaten gezielt und explizit erstellt werden und zweitens der Bottom-up Ansatz, bei dem Lernressourcen im Laufe der Zeit gesammelt wurden und diese im Nachhinein analysiert und verarbeitet werden, um relevante Metadaten zu gewinnen. (Dahl & Vossen, 2008)

Dahl und Vossen (2008) beschäftigen sich mit der Fragestellung:

*„Wer, wann und wie werden Metadatensätze für Lernressourcen erzeugt?“*

Dahl und Vossen (2008) beschreiben Ansätze für die Metadatengenerierung in der „Web 2.0 ERA“ und diskutieren automatische, halb-automatische und manuelle Methoden für die Metadatenerstellung. Außerdem beschreiben die Autoren die Metadatengenerierung der e-Learning Plattform der Uni Münster.

Brasher und McAndrew (2004) untersuchen den Erstellungsprozess von nutzergenerierten Metadaten für e-Learning-Ressourcen. Sie zeigen unter anderem Probleme im Erstellungsprozess auf, z. B. fehlende Motivation der Benutzer oder unterschiedliches Verständnis für Lernressourcen. Brasher und McAndrew (2004) identifizierten Situationen, in denen die menschliche Produktion von Metadaten unvermeidlich ist. Nötig ist die Erstellung von Metadaten durch den Benutzer für nicht textuelle Lernressourcen, wie z. B. Filme und Bilder. Der Grund liegt darin, dass aus Multimedia-Lernressourcen schwer Metadaten extrahiert werden können. Des Weiteren müssen

Metadatenelemente vom Benutzer definiert werden, die nicht maschinell von der Lernressource extrahiert werden können. Beispielsweise können pädagogische Annotierungen, wie z. B. Schwierigkeitsgrad, Altersgruppe, typische Lernzeit usw., nur durch den Benutzer realisiert werden. Aus diesem Grund führt er Motivationsfaktoren innerhalb von Communities an und wie diese genutzt werden können, um eine gute Qualität von Endbenutzer generierten Metadaten zu gewährleisten. (Brasher & McAndrew, 2004)

Bezüglich des Zeit- und Kostenaufwandes ist es für eine einzelne Person fast unmöglich, sowohl hochquantitative als auch hochqualitative Metadaten für eine Reihe von Lernobjekten zu erstellen. Ein kollaborativer Ansatz für die Metadatenerstellung ist weder definiert, noch wird er explizit von einem Tool unterstützt. (Dahl & Vossen, 2008)

Pansanato und Fortes (2005) analysierten eine kollaborative Lernplattform namens CoTeia auf die Vollständigkeit der Metadatenbeschreibung für die einzelnen Lernressourcen. Die Annotierung der Lernressourcen wurde durch den Benutzer der Lernplattform erstellt. Die Metadaten für die Beschreibung einer Lernressource umfassen die Metadatenelemente Titel, Autor und Keywords. Pansanato und Fortes (2005) kamen zu dem Ergebnis, dass 88,41 % der Metadatenfelder von 3.131 Lernressourcen nicht ausgefüllt waren.

Eine mögliche Lösung für das Problem der zeitaufwendigen Metadatenerstellung durch den Endbenutzer kann die automatische Erstellung der Metadaten darstellen. Die automatische Erzeugung von Metadaten extrahiert relevante Informationen aus Lernobjekten und den Kontext, in dem sie gespeichert und genutzt werden. Auf diese Weise muss der Benutzer sich nicht mit der Metadatenerstellung quälen, wenn er es nicht will. Nur wenn der Benutzer möchte, kann er Metadaten korrigieren, ergänzen oder löschen. Nach Cardinaels et al. (2005) allerdings wollen die meisten Benutzer keine Zeit damit verbringen, Metadaten zu bearbeiten. (Cardinaels et al., 2005)

Pansanato und Fortes (2005) beschreiben die Entwicklung und die Verwendung von Strategien für die automatische Generierung von Metadaten für Lernressourcen in einem Wiki-System. Das eingesetzte Metadaten-Profil von Pansanato und Fortes (2005) ist eine Teilmenge der Attribute des LOM Standards. Für die Metadatengenerierung verfolgen die Autoren vier Strategien.

Diese vier Strategien sind Ressource, Kontext, Template und Regeln. In den folgenden Absätzen werden die Strategien Ressource und Kontext genauer erklärt. Auf die Strategien Template und Regeln wird nicht näher eingegangen, weil sie für diese Arbeit keine Relevanz darstellen.

Die „Ressource“-Strategie basiert auf der Inhaltsanalyse der zugrunde liegenden Lernressource. Um Metadaten aus der Lernressource extrahieren zu können, werden Information-Retrieval-Technologien, wie z. B. Language-Classification, Automatic-Summarization und Keyword-Extraction, eingesetzt. (Pansanato & Fortes, 2005)

Die „Kontext“-Strategie basiert auf dem Kontext, in der die Lernressource verwendet wird. Es wird zwischen drei Kontext-Typen unterschieden, aus denen Metadaten generiert werden können. Die drei Kontexttypen sind Dateisystem, Struktur und Benutzermanagement. Diese drei Kontexte liefern Metadaten, wie z. B. Dateigröße (die aus dem Dateisystem ausgelesen wird), Autor (der vom Benutzermanagement abgefragt werden kann) und Relationship (die von der hierarchischen Struktur der Lernressource bestimmt werden kann). (Pansanato & Fortes, 2005)

Cardinaels, Meire und Duval (2005) beschreiben die Entwicklung von Web-Services für die automatische Generierung von Metadaten für Lernressourcen. Sie beschreiben vier Metadatenquellen, aus denen Metadaten extrahiert werden können.

- **Text-Analyse**

Die naheliegendste Quelle für Metadaten zu einer Ressource ist die Ressource selbst. Typische Content-Analyser sind Keyword-Extrahierer und Language-Analyzers.

- **Kontext-Analyse**

Auch der Kontext, in dem die Lernressource verwendet wird, kann als Quelle für Metadaten herangezogen werden. Cardinaels, Meire und Duval (2005) betrachten drei Kontexttypen. Erstens die Creator-Profiles, die Informationen über den Autor einer Lernressource beinhalten. Zweitens das Learning-Content-Management-System, in

dem Lernressourcen und deren Metadaten gespeichert werden. Die existierenden Metadaten im Learning-Content-Management-System können als Quelle für neue Lernressourcen verwendet werden. Diese Informationen werden typischerweise verwendet, wenn die neue Lernressource in Beziehung zu einer existierenden Lernressource steht, beispielsweise wenn die neue Lernressource eine neue Version von einer bestehenden Lernressource darstellt. Drittens das Learning-Management-System, das umfangreiche Kontextinformationen liefern kann, z. B. Kurse, in denen die Lernressource verwendet wird, die Anzahl der Downloads der Lernressource usw.

- **Verwendung des Dokuments**

Die Verwendung einer Lernressource kann ebenfalls als Quelle für Metadaten dienen. Um aus der Verwendung Metadaten zu extrahieren, müssen die Benutzerinteraktionen mit der Lernressource mitprotokolliert werden. Beispielsweise kann die Metadateninformation Autor einer Lernressource aktualisiert werden, wenn ein Benutzer eine Lernressource bearbeitet. Außerdem können durch die Verwendung der Lernressource durch die Benutzer pädagogische Metadaten, wie z. B. Interaktivitätsstufe und Altersgruppe, bestimmt werden.

- **Struktur-Analyse (Kompositum)**

Neue Lernobjekte können ein Teil von einem anderen bzw. ganzen Lernobjekt sein. In einem solchen Fall, dass ein neues Lernobjekt ein Teil eines Ganzen ist, sind die Metadaten für das Ganze eine interessante Quelle für die Metadaten der neuen Lernressource.

Während auf der einen Seite Kosten für Personal und Aufwand durch die automatische Erstellung von Metadaten reduziert werden, müssen auf der anderen Seite fehlerhafte Metadaten in Kauf genommen werden. (Dahl & Vossen, 2008)

Die automatische Erstellung von Metadaten aus textbasierten Lernressourcen ist relativ zuverlässig, hingegen können aus multimedialen Lernressourcen keine oder nur wenige Metadaten, die sich auf den Inhalt der Ressource beziehen, extrahiert werden. (Dahl & Vossen, 2008) Beispielsweise kann der

Titel einer textbasierten Lernressource aufgrund der Formatierung und der Position im Text ziemlich einfach extrahiert werden, hingegen kann aus einem Bild ein möglicher Titel nur mit großem Aufwand extrahiert werden, indem z. B. Objekterkennungsalgorithmen eingesetzt werden.

Die automatische Metadatenerzeugung stellt einen vielversprechenden Weg dar, um Metadaten zu erstellen, jedoch sollte dieser automatische Ansatz nicht ohne menschliche Interaktion eingesetzt werden. Die menschliche Interaktion bei der Metadatenerstellung ist aufgrund der Fehleranfälligkeit der automatischen Metadatengenerierung wichtig. Somit kann der Benutzer fehlerhaft generierte Metadaten-Attribute korrigieren und den Metadatensatz vervollständigen. (Dahl & Vossen, 2008)

Die Kombination der automatischen Metadatengenerierung mit anschließender manueller Vervollständigung und Korrektur der automatisch generierten Metadaten durch den Benutzer wird als halb-automatische Metadatengenerierung bezeichnet (Dahl & Vossen, 2008).

### **Bedeutung für die vorliegende Arbeit**

Die User-Generated-Content-Komponente baut auf der halb-automatischen Methode auf. Die Vorteile der halb-automatischen Generierung von Metadaten liegen darin, dass der Benutzer bei der Annotierung von Lernressourcen weniger Zeit benötigt und durch die Korrekturmöglichkeit der automatisch generierten Metadaten die Qualität der Metadaten erhöht wird (Dahl & Vossen, 2008).

## **2.4 Information Retrieval Ansätze für die Metadatengewinnung**

In den folgenden Absätzen werden Information Retrieval Ansätze vorgestellt, die bei der automatischen Ermittlung von Metadaten verwendet werden können. Konkret wird auf die Extrahierung von den Metadatenelementen Titel, Keywords und Zusammenfassung genauer eingegangen.



Bevor Metadaten aus einem Dokument extrahiert werden können, müssen die Dokumente gefiltert werden, um den bedeutungsvollen Text zu erhalten. Beispielsweise sollen in einem HTML-Dokument keine HTML-Tags indexiert werden. Aus diesem Grund wird ein Filter angewendet, der die HTML-Tags entfernt. (Belew, 2000, S. 41)

Jedoch können auch HTML-Tags sinnvoll sein, um Metadaten zu extrahieren. Der Überschriften-Tag <h1> oder der <title>-Tag im HTML-Kopf können auf den Titel des Dokumentes schließen lassen. Weiters können <meta>-Tags<sup>16</sup>, eingebettet im HTML-Kopf, eine nützliche Quelle für die automatische Generierung eines LOM sein. Diese <meta>-Tags kategorisieren und beschreiben das HTML-Dokument und umfassen typischerweise die Metadaten „description“, „keywords“, „copyright“, „date“ und „author“. Diese <meta>-Tags werden in HTML-Dokumenten eingebunden und sind nicht sichtbar im Webbrowser. Die Tags erleichtern das maschinelle Kategorisieren der Webseite und helfen dem Benutzer, zum Beispiel die Webseite mit der Suchmaschine leichter zu finden. Google.com verwendet beispielsweise das „description“ Meta-Tag, das die Webseite beschreibt, und zeigt diese kurze Beschreibung der Webseite in der Auflistung der Suchergebnisse an.<sup>17</sup> Diese <meta>- , <h1>- und <title>-Tags können viele Informationen über die Online-Lernressource liefern, ohne den Text der Lernressource analysieren zu müssen.

Sowohl für das Extrahieren von Keywords als auch bei Zusammenfassungen ist es üblich, Wörter, die im Titel eines Dokumentes vorkommen, höher zu gewichten (Raskutti, Ferrá, & Kowalczyk, 2001).

Chien (1997) beschreibt eine Keyword-Extraction-Methode basierend auf PAT-Trees. Dieser Ansatz für die Extraktion von Keywords aus chinesischen Texten wird in mehreren Information Retrieval Systemen eingesetzt.

---

<sup>16</sup> W3C. (24. Dezember 1999). The global structure of an HTML document. Abgerufen am 19. Juli 2012 von HTML 4.01 Specification: <http://www.w3.org/TR/html401/struct/global.html#h-7.4.4.2>

<sup>17</sup> Google.com. (3. September 2012). Meta-Tags. Abgerufen am 17. Juli 2012 von Support.Google.com: <http://support.google.com/webmasters/bin/answer.py?hl=de&answer=79812>

Besteht eine Basis von annotierten Lernressourcen, kann auf induktive Lernalgorithmen zurückgegriffen werden, um Texte zu kategorisieren und folglich Keywords und Zusammenfassungen von Dokumenten zu extrahieren. Dumais, Platt, Heckerman und Sahami (1998) beschreiben und vergleichen fünf induktive Lernalgorithmen für die Text-Kategorisierung:

- Find Similar
- Decision Trees
- Naive Bayes
- Bayes Nets
- Support Vector Machines (SVM)

Die Effektivität dieser fünf Ansätze wurde auf Basis ihrer Lerngeschwindigkeit, Real-Time-Klassifizierungsgeschwindigkeit und Klassifizierungsgenauigkeit bestimmt. Die Autoren Dumais et al. (1998) kommen zu dem Ergebnis, dass der Support Vector Machine (SVM) Algorithmus den besten Klassifikator darstellt, weil er sehr präzise und schnell zu trainieren ist.

Matsuo und Ishizuka (2004) beschreiben einen Keyword-Extraction Ansatz für alleinstehende Dokumente. Die Keywords werden durch Kookkurrenzen bestimmt.

In den nächsten Absätzen werden Arbeiten angeführt, die sich mit der Zusammenfassung von Texten beschäftigen.

Kolcz, Prabakarmurthi und Kalita (2001) beschreiben, dass die Qualität der Zusammenfassung eines Textes subjektiv ist. Aus Sicht der Menschen ist es wichtig, dass eine Zusammenfassung den „Kern“ eines Dokuments enthält (Kolcz, Prabakarmurthi, & Kalita, 2001).

Grundsätzlich wird zwischen zwei Zusammenfassungsmethoden unterschieden. Erstens die „Extractive-Summarization“ Methode, bei der wichtige Sätze vom Lernobjekt bestimmt werden. Anschließend werden die wichtigsten Sätze von der Lernressource kopiert und die Gesamtheit der kopierten Sätze bildet die Zusammenfassung. Zweitens die „Abstractive-Summarization“ Methode, die auf

einer tieferen Analyse des Textes basiert und den Text grammatikalisch zusammenfasst. (Lin, 2009)

Die ersten Ansätze, um wichtige Sätze aus einem Text zu extrahieren, beschrieben Luhn (1958), Baxendale (1958) und Edmundson (1969).

Luhn (1958) beschäftigt sich in seiner Arbeit mit den statistischen Informationen der Wort-Häufigkeit und Verteilung. Diese statistischen Informationen werden herangezogen, um die Wichtigkeit von Wörtern und in weiterer Folge von Sätzen in einem Text zu bestimmen. Die Sätze mit der höchsten Signifikanz bilden die Zusammenfassung. (Luhn, 1958)

Baxendale (1958) beschäftigt sich mit der Position von Sätzen und Wörtern im Text, um aufgrund der Position die Signifikanz des Satzes zu bestimmen. Baxendale (1958) untersuchte 200 Absätze und kam zu dem Ergebnis, dass in 85 % der Absätze der erste Satz einen Topic-Sentence darstellt und somit wichtige Informationen für die Zusammenfassung enthält.

Edmundson (1969) beschäftigt sich mit Key-Phrases und ergänzte die zwei oben genannten Methoden „Extrahierung von Sätzen auf Basis der am häufigsten vorkommenden Wörter“ (Luhn, 1958) und „Extrahierung von Sätzen auf Basis deren Position im Text“ (Baxendale, 1958) um weitere zwei Methoden für die Extrahierung von signifikanten Sätzen. Die erste Methode basiert auf dem Vorhandensein von Cue-Words, wie z. B. „significant“ und „hardly“. Zweitens bestimmt Edmundson (1969) signifikante Sätze aufgrund des Aufbaus eines Dokumentes, beispielsweise stellen Wörter im Titel und Überschriften eine wichtige Rolle für die Zusammenfassung dar.

Das & Martins (2007) erhoben in ihrer Arbeit automatische „Text Summarization“ Algorithmen. Sie erarbeiteten in ihrer Publikation mehrere Ansätze für automatische Textzusammenfassungen, sowohl im Bereich „Single-Document-Summarization“ als auch „Multiple-Document-Summarization“. Die Autoren beschreiben ältere Ansätze, wie z. B. „Naive-Bayes“, „Decision Tree“ und „Hidden Markov Model“, um wichtige Sätze zu extrahieren. Sie führen aber auch neuere Sprachanalysetechniken an, die mit Hilfe von Heuristiken die Zusammenfassung von Dokumenten realisieren.

## 2.4.1 Öffentliche Information Retrieval Services

Es gibt einige öffentlich zugängliche Information Retrieval Services, die genutzt werden können, um Metadaten aus Online-Lernressourcen zu extrahieren. Die folgende Auflistung gibt einen Auszug der freien Metadatenextraktions-Web-Services:

- **Open Xerox API<sup>18</sup>**

Open Xerox bietet einige Web-Services an, die zur Extrahierung von Metadaten genutzt werden können. Diese Services sind frei zugänglich, nachdem ein Entwicklerkonto angelegt wurde. Unter anderem bietet die Open Xerox API einen Language-Identifizierer<sup>19</sup> und einen Incremental-Parser<sup>20</sup> an. Dieser inkrementelle Parser führt eine linguistische Analyse des Textes durch<sup>21</sup>.

- **Yahoo API<sup>22</sup>**

Die Yahoo API bietet Services für Term-Extraction<sup>23</sup> und Content-Analysis<sup>24</sup> an, die genutzt werden können, um Konzepte und Keywords

---

<sup>18</sup> <http://open.xerox.com>

<sup>19</sup> XEROX CORPORATION. (2012). Language Identifizierer. Abgerufen am 12. Mai 2012 von Open.Xerox.com: <http://open.xerox.com/Services/LanguageIdentifizierer>

<sup>20</sup> XEROX CORPORATION. (2012). Xerox Incremental Parser. Abgerufen am 12. Mai 2012 von Open.Xerox.com: <http://open.xerox.com/Services/XIPParser>

<sup>21</sup> XEROX CORPORATION. (2012). Xerox Incremental Parser. Abgerufen am 12. Mai 2012 von Open.Xerox.com: <http://open.xerox.com/Services/XIPParser>

<sup>22</sup> <http://developer.yahoo.com/>

<sup>23</sup> Yahoo! Inc. (2012). Term Extraction Documentation for Yahoo! Search. Abgerufen am 11. Mai 2012 von Yahoo! Developer Network: [http://developer.yahoo.com/search/content/V1/termExtraction.html#access\\_via\\_yql](http://developer.yahoo.com/search/content/V1/termExtraction.html#access_via_yql)

<sup>24</sup> Yahoo! Inc. (2012). Content Analysis Documentation for Yahoo! Search. Abgerufen am 11. Mai 2012 von Yahoo! Developer Network: <http://developer.yahoo.com/search/content/V2/contentAnalysis.html>

aus Lernressourcen zu extrahieren. Die Anzahl der Aufrufe der Yahoo-Services ist mit 5.000 Abfragen pro IP-Adresse und Tag für die nicht kommerzielle Nutzung beschränkt.

- **Bing API<sup>25</sup>**

Die Bing Search API kann dazu genutzt werden, um den Titel und eine kurze Beschreibung der Online-Lernressource zu extrahieren. Die Bing API ist für die nicht kommerzielle Nutzung auf sieben Abfragen pro IP und Tag beschränkt.<sup>26</sup>

## 2.5 Qualität von Metadaten

Bruce und Hillmann (2004) definieren sieben Kriterien, die Metadaten mit hoher Qualität erfüllen müssen. Die Qualitätsmerkmale nach Bruce und Hillmann (2004) umfassen Vollständigkeit, Richtigkeit, Herkunftsangabe (Autor), Erfüllung der Erwartung einer Community, logische Konsistenz und Kohärenz, Aktualität und Zugänglichkeit.

Ochoa und Duval (2006) erstellten Qualitätsmetriken für Metadaten, die Lernobjekte beschreiben. Konkret haben sie Metriken für den IEEE LOM Standard umgesetzt. Mit Hilfe der Qualitätsmetriken können Metadaten maschinell in Repositories auf ihre Qualität überprüft werden. Die Qualitätskennzahlen nehmen auf die oben genannten Qualitätsmerkmale von Bruce und Hillmann (2004) Bezug.

Um die Qualität der Metadaten zu steigern, zeigen die Autoren Barton et al. (2003) unter anderem folgende Maßnahmen auf:

- Erkläre den Autoren, warum Metadaten wichtig sind.
- Bereitstellung von mehr Dokumentation über den Metadatenerstellungsprozess.

---

<sup>25</sup> <http://www.bing.com/developers/>

<sup>26</sup> Bing. (2012). API Basics. Abgerufen am 12. Mai 2012 von Bing-Developers: <http://www.bing.com/developers/s/APIBasics.html>

- Einsatz von Metadaten-Validatoren, die die Metadaten verbessern.

Guy & Powell (2004) definieren eine Reihe von Verfahren und Maßnahmen, durch die die Qualität der Metadaten gesteigert werden kann. Der wichtigste Punkt für hochqualitative Metadaten besteht darin, dass am Beginn die relevanten Metadaten bestimmt werden. Um die relevanten Metadaten ermitteln zu können, müssen die funktionalen Anforderungen an die Anwendung bestimmt werden und auf Basis dieser Anforderungen können die relevanten Metadaten bzw. das Application-Profil abgeleitet werden. Außerdem sollen die möglichen Werte und das Format der einzelnen Metadatenattribute definiert werden. Weiters kann die Verbesserung der Usability des Metadaten-Erstellungstools dazu beitragen, die Qualität der Metadaten zu erhöhen. Zusätzlich soll ein geeigneter Prozess für die Qualitätskontrolle der Metadaten umgesetzt werden. Ergänzend kann die Qualität durch Usability-Tests getestet werden. (Guy & Powell, 2004)

Die Qualität der Metadaten für Lernobjekte ist wichtig, um die sinnvolle Verwendung und Interoperabilität zu gewährleisten (Barton, Currier, & Hey, 2003; Hughes, 2004; Currier, Barton, O'Beirne, & Ryan, 2004; Guy & Powell, 2004; Beall, 2005; Ochoa & Duval, 2006).

Fehler in Metadatenansätzen können in vielen Formen auftreten (Beal, 2005; Barton et al, 2003). Barton et al (2003) beschreiben einige Fehlerquellen bei der Metadatenerstellung, die die Qualität der Metadaten mindern. Diese Fehlerquellen sind Rechtschreibung, verschiedene Schreibweisen von Personen, mehr als ein möglicher Titel, Keywords, Datumsformate usw. Das Problem liegt darin, dass Fehler den Zugriff auf verfügbare Lernmaterialien verhindern. Beispielsweise kann eine Lernressource über Apfelbäume unauffindbar sein, wenn die Lernressource mit dem Keyword „Aplebäume“ annotiert wurde und ein Benutzer über die Keyword-Suche nach „Apfelbäume“ sucht. Fehler in Metadatenansätzen sind nach Beall (2005) schwerwiegend, wenn Metadaten als Ersatz für die Objekte dienen und keine Volltextsuche verfügbar ist.

Auch Barton et al. (2003) zeigen, dass durch Metadaten in schlechter Qualität die Lernressourcen für den Benutzer unauffindbar werden und somit die Lernressourcen unbenutzt bleiben. Folglich kamen Barton et al. (2003) zur Erkenntnis, dass die Qualität der Metadaten sehr wichtig ist. Sie führen

folgende zwei Beispiele an, in denen eine hohe Qualität der Metadaten unabdingbar ist:

- für das Auffinden von Metadaten und dadurch auch das Finden von Lernressourcen
- für die Wiederverwendung von relevanten Materialien

Cardinaels, Meire und Duval (2005) sehen mehrere Gründe, warum Personen keine Metadaten für Lernressourcen anlegen oder Metadaten in mangelhafter Qualität erstellen. Der Hauptgrund nach Cardinaels et al. (2005) liegt darin, dass aktuelle Tools für die Erstellung von Metadaten nicht benutzerfreundlich sind. Die meisten Metadatenerstellungs-Tools repräsentieren nur den Metadatenstandard. Der Benutzer muss eine große Anzahl an Feldern ausfüllen. Jedoch wurden die Metadatenstandards nicht dazu definiert, um direkt für den Endbenutzer sichtbar zu sein. Eine direkte Darstellung der Standards macht es sehr schwierig und zeitaufwendig, Metadaten für eine Lernressource zu erstellen. (Cardinaels et al., 2005)

### **Bedeutung für die vorliegende Arbeit**

Dieser Abschnitt macht die Wichtigkeit der Qualität der Metadaten deutlich. Es zeigt sich, dass Lernressourcen, die Metainformationen in schlechter Qualität aufweisen, mit nicht annotierten Lernressourcen gleichzusetzen sind sofern nur eine Metadaten-Suche existiert. In beiden Fällen sind die Lernressourcen über ihre Metadaten nicht auffindbar und somit werden sie nicht so häufig wiederverwendet. Somit spielt die Qualität der zukünftig benutzergenerierten Metadaten im Organic.Edunet Web-Portal eine wichtige Rolle.

## **2.6 Bereitstellen von Metadaten**

Grundsätzlich werden Lernressourcen in einer Gruppe erstellt und verwendet. Diese Lernressourcen werden allerdings nur in seltenen Fällen geteilt und anderen Gruppen oder Personen zur Verfügung gestellt. Der erste Schritt, um Lernressourcen für die Nutzung durch andere Personen zugänglich zu machen, liegt darin, dass Metadaten für die Lernressourcen angelegt werden. Das Thema Metadaten und Metadatenerzeugung wurde im Abschnitt 2.3 ausführlich

behandelt. Nachdem die Metadaten für die Lernressourcen erstellt wurden, können die Metadaten für andere Gruppen und Personen bereitgestellt werden.

Durch den Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) Standard können Lernressourcen bzw. deren Metadaten veröffentlicht und geharvestet werden (Lagoze, Sompel, Nelson, & Warner, 2002). Der OAI-PMH Standard (Lagoze, Sompel, Nelson, & Warner, 2002) unterscheidet zwischen zwei Klassen von Teilnehmern:

- **Data-Provider (Repository)**

Der Data-Provider veröffentlicht die Metadaten mit Hilfe des OAI-PMH Standards. Er verwaltet das Metadaten-Repository und kann sechs definierte OAI-PMH Anfragen verarbeiten.

- **Service-Provider (Harvester)**

Der Service Provider ruft die zur Verfügung gestellten Metadaten ab und nimmt sie als Grundlage für weitere Services.

Sompel et al. (2004) beschreiben das Bereitstellen und das Abrufen von Ressourcen mittels des OAI-PMH Standards und beschreiben den üblichen Ansatz von der Bereitstellung von Metadaten und deren Abruf.

- Ein OAI-PMH Harvester (Service-Provider) ruft Dublin-Core-Datensätze aus einem OAI-PMH-Repository (Data-Provider) ab.
- Der Harvester analysiert jeden Dublin-Core-Datensatz und extrahiert die dc-identifier Information, um den Speicherort der beschriebenen Ressource zu bestimmen.
- Ein separater Prozess, der nicht Teil des OAI-PMH Standard ist, sammelt die Ressource von ihrem Speicherort.

## 2.7 Zusammenfassung

Web 2.0 bzw. Social Media ermöglicht dem Benutzer, aktiv an der Gestaltung des Inhaltes im Web beizutragen. Dieser vom Benutzer erstellte Inhalt wird als



User-Generated-Content bezeichnet. Der nutzergetriebene Inhaltserstellungsprozess unterscheidet sich von der industriellen Inhaltserstellung deutlich. Die größten Unterschiede bestehen darin, dass bei der nutzergetriebenen Inhaltserstellung der Benutzer sowohl Konsument als auch Produzent von Inhalten ist und durch das kollaborative und iterative Verbessern und Erweitern des Inhalts im Web 2.0 nie ein fertiges Produkt existiert.

Dieser Übergang von der industriellen Inhaltserstellung hin zur nutzergetriebenen Inhaltserstellung soll im Organic.Edunet Web-Portal durch diese Arbeit realisiert werden.

Es gibt verschiedene Arten von User-Generated-Content, wie z. B. Texte, Bilder, Musik, Audio usw. Das Organic.Edunet Web-Portal bietet eine große Anzahl von verschiedenen Arten von Lernressourcen an, die von Experten erstellt wurden. Diese verschiedenen Arten von Lerninhalten sollen in Zukunft durch den Organic.Edunet Benutzer annotiert werden können.

Die größten Motivationsfaktoren für Benutzer, um User-Generated-Content zu erstellen, stellen Ideologie und Spaß dar. Dies war das Ergebnis von zwei Studien von Yang und Lai (2010) und Nov (2007), die analysierten, warum Menschen auf Wikipedia Wissen teilen. Daraus lässt sich schließen, dass diese zwei Motivationsfaktoren auch den Organic.Edunet Benutzer motivieren werden, in Zukunft unentgeltlich Metadaten zu bearbeiten.

Die Statistik der OECD (2007) über die Verbreitung von User-Generated-Content in der Europäischen Union lässt den Schluss zu, dass rund 60 % der Organic.Edunet Benutzer zukünftig aktiv an der Gestaltung des Lerninhaltes im Organic.Edunet Web-Portal beitragen werden. Folglich kümmern sich eine große Anzahl von Organic.Edunet Benutzern um die Aktualität der Metadaten und erweitern das Angebot an Lernressourcen.

Grundsätzlich liegt User-Generated-Content unstrukturiert auf verschiedenen Plattformen. Aus diesem Grund sind benutzererstellte Lernressourcen schwer auffindbar und bleiben zum Teil ungenutzt. Der Einsatz von Metadaten, um Lernressourcen zu beschreiben, löst das Problem der Unstrukturiertheit. Weiters sollen die eingesetzten Metadaten auf einem Standard beruhen, sodass keine Interoperabilitätsprobleme beim Austausch von Lernressourcen zwischen

zwei Organisationen entstehen. Der Standard für die Beschreibung von Lernressourcen ist der IEEE LOM Standard. Folglich werden die Lernressourcen im Organic.Edunet Web-Portal und bei der praktischen Umsetzung dieser Arbeit mit Hilfe des IEEE LOM Standards beschrieben.

Bei dem Einsatz von Metadaten in einer Anwendung soll ein Application-Profile definiert werden, das der funktionalen Anforderung der Anwendung entspricht. Weiters sollen "mandatory" und optionale Metadatenelemente definiert werden.

Es wird zwischen drei verschiedenen Metadatenerstellungsmethoden unterschieden. Erstens die nutzergenerierten Metadaten, zweitens die automatisch generierten Metadaten, bei welchen mit Hilfe von Information Retrieval Technologien maschinell Metadaten aus der Lernressource extrahiert werden, drittens eine Kombination aus dem nutzergenerierten Metadatenerstellungsansatz und der maschinellen Metadatenextraktion. Bei dieser Methode, die halb-automatische Metadatengenerierung genannt wird, werden zuerst die Metadaten maschinell aus der Lernressource extrahiert und anschließend hat der Benutzer die Möglichkeit, die Metadaten für die Lernressource zu vervollständigen und zu korrigieren. Die User-Generated-Content Komponente baut auf der halb-automatischen Methode der Metadatengenerierung auf, weil dadurch der Organic.Edunet Benutzer weniger Zeit benötigt, um Metadaten zu erstellen und er die automatisch generierten Metadaten ergänzen kann.

Um automatisch Metadaten aus Lernressourcen extrahieren zu können, werden Information Retrieval Algorithmen eingesetzt, die z.B. Titel, Keywords und eine Zusammenfassung aus der Lernressource extrahieren. Dafür stehen öffentliche Information Retrieval Services zur Verfügung, die frei bzw. begrenzt kostenlos genutzt werden können.

Einen wichtigen Punkt bei der Metadatengenerierung stellt die Qualität der generierten Metadaten dar. Qualitätsmerkmale für Metadaten sind unter anderem Vollständigkeit, Richtigkeit, Herkunftsangabe, Aktualität und die Erwartung der Community (Ochoa & Duval, 2006). Das Problem von Metadaten in schlechter Qualität liegt darin, dass die beschriebenen Lernressourcen über die Metadaten nicht auffindbar sind.

Der OAI-PMH Standard ermöglicht die standardisierte Veröffentlichung und das Abrufen von Lernressourcen bzw. deren Metadaten. Dieser Standard ermöglicht die einfache Bereitstellung der benutzergenerierten Metadaten, die von Service-Providern wiederverwendet werden können.



## **3 Die User-Generated-Content Komponente**

Im Kapitel 2 wurde ausführlich die Forschungsfrage (siehe Abschnitt 1.2) theoretisch erarbeitet. Dieses Kapitel befasst sich mit der praktischen Umsetzung der Forschungsfrage.

Zuerst werden die User-Generated-Content Szenarien und der erweiterte Lebenszyklus von Lernressourcen für das Organic.Edunet Web-Portal erläutert. Anschließend werden ähnliche Systeme aufgezeigt. Danach beschäftigt sich das Kapitel mit dem Application-Profil, den Nutzungsrechten von neuen Lernressourcen und den Services der User-Generated-Content Komponente. Zum Schluss werden die Integration der User-Generated-Content Komponente in das Organic.Edunet Web-Portal und der Abruf der benutzererstellten Metadaten beschrieben.

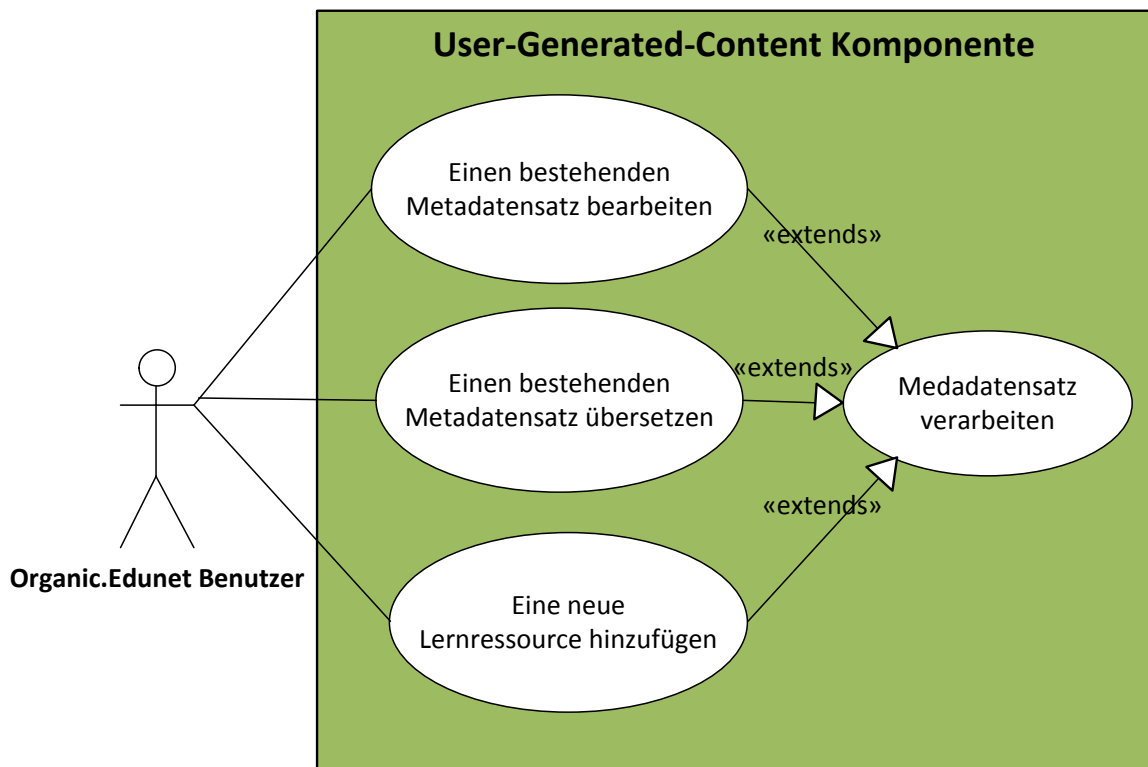
### **3.1 Einleitung**

In den vergangenen Jahren hat ein Paradigmenwechsel hinsichtlich Web Interaktion von Benutzern stattgefunden. Benutzer sind inzwischen Konsumenten und Produzenten von Inhalten. Dieser Paradigmenwechsel ist im Abschnitt 2.1 ausführlich erklärt. Der Paradigmenwechsel zum User-Generated-Content wurde im bestehenden Organic.Edunet (siehe Abschnitt 1.1) noch nicht berücksichtigt. Die abrufbaren Lernressourcen im Organic.Edunet sind von Experten erstellt und annotiert worden. Der Organic.Edunet Benutzer nimmt die Rolle des Konsumenten ein. Aus diesem Grund wurde die in dieser Arbeit beschriebene User-Generated-Content Komponente entwickelt. Die User-Generated-Content Komponente ermöglicht dem Organic.Edunet Benutzer, aktiv an der Gestaltung der Lerninhalte beizutragen.

Im Kontext des Organic.Lingua Projektes umfasst der Begriff User-Generated-Content (UGC) das Hinzufügen, Modifizieren und Übersetzen von kompatiblen IEEE LOM Datensätzen durch den Endbenutzer. Diese Operationen werden durch die User-Generated-Content-Komponente, die aus drei Widgets besteht, unterstützt. Jedes User-Generated-Content-Widget unterstützt eine Operation.

Zusätzlich umfasst die UGC-Komponente Services, die die Kommunikation zwischen den User-Generated-Content Widgets und der Organic.Edunet Plattform bewerkstelligt. Die User-Generated-Content-Komponente speichert alle vom Benutzer generierten Metadaten in einem neuen Organic.Edunet Provider-Repository. Folglich befindet sich die User-Generated-Content-Komponente außerhalb des Organic.Edunet Systems. Die User-Generated-Content-Komponente stellt die benutzererstellten Metadaten mit Hilfe des OAI-PMH Standards zur Verfügung.

Die User-Generated-Content Komponente kann in drei Endbenutzer-Szenarien (siehe Abbildung 3-1) unterteilt werden.



**Abbildung 3-1: User-Generated-Content Szenarien**

Diese drei User-Generated-Content Szenarien sind:

- **einen bestehenden Metadatenatz bearbeiten.**

Szenario 1: Der Organic.Edunet Benutzer bearbeitet einen vorhandenen Metadatenatz:

Der Benutzer bemerkt eine falsche oder unvollständige Beschreibung einer Lernressource. Mit einem Klick auf einen "Complete/Correct this record"-Link im Organic.Edunet Web-Portal öffnet sich das User-Generated-Content Widget und ermöglicht dem Benutzer, den Metadatenatz zu vervollständigen und/oder zu korrigieren.

- **einen bestehenden Metadatenatz übersetzen.**

Szenario 2: Der Organic.Edunet Benutzer übersetzt einen bestehenden Metadatenatz:

Der Organic.Edunet Benutzer findet eine Beschreibung eines Lernobjektes in einer Sprache, die er nicht versteht. Die User-Generated-Content Komponente ermöglicht ihm, eine automatische Übersetzung in seine Muttersprache anzufordern, indem er auf einen „Translate this record“-Link im Organic.Edunet Web-Portal klickt. Die Übersetzung wird dem Benutzer im User-Generated-Content Widget angezeigt. Diese Übersetzung kann der Organic.Edunet Benutzer bei Bedarf korrigieren und anschließend im User-Generated-Content Repository speichern.

- **eine neue Lernressource hinzufügen.**

Szenario 3: Der Organic.Edunet Benutzer fügt eine Metadatenbeschreibung für eine neue Lernressource zum Organic.Edunet hinzu:

Der Organic.Edunet Benutzer surft im Internet und trifft auf eine interessante Online-Ressource, die noch nicht im Organic.Edunet Web-Portal enthalten ist und würde diese gerne mit der Organic.Edunet Community teilen. Ein leicht zugänglicher Button, der im Browser integriert ist, ermöglicht dem Organic.Edunet Benutzer, das User-Generated-Content Widget zu öffnen. Wenn der Benutzer das Widget öffnet, werden relevante Attribute für die Beschreibung der Lernressource

automatisch mittels Information-Retrieval Services (siehe Abschnitt 3.7.2) ausgefüllt.

Der Benutzer kann die Metadaten vervollständigen und die vorausgefüllten Felder korrigieren. Findet der Organic.Edunet Benutzer die neue Lernressource vollständig beschrieben, kann er den Metadatensatz, der die neue Lernressource beschreibt, im User-Generated-Content Repository speichern.

Die drei Szenarien basieren auf der gleichen Grundfunktionalität und unterscheiden sich nur minimal.

## 3.2 Ähnliche Systeme und Projekte

Merlot<sup>27</sup> und OER-C<sup>28</sup> sind Lern-Repositories, die bereits Tools für das Sammeln von Online-Lernressourcen durch den End-Benutzer zur Verfügung stellen. Diese bestehenden Tools zeigen Möglichkeiten auf, wie Lernressourcen-Annotierungen realisiert werden können. Das Problem dieser zur Verfügung gestellten Tools der Lern-Repositories zur Lernressourcen Annotierung, liegt in folgenden Punkten:

- Die Tools funktionieren nicht als Blackbox-Systeme.
- Das Extrahieren von Metadaten wurde nicht bzw. unvollständig umgesetzt.
- Bei der Erfassung der Metadaten wird keine automatische Rechtschreibkorrektur durchgeführt.
- Das Übersetzen von Metadaten wird nicht ermöglicht.
- Die Tools bieten keine mehrsprachigen Benutzeroberflächen an.

---

<sup>27</sup> <http://www.merlot.org/>

<sup>28</sup> <http://www.oercommons.org/>



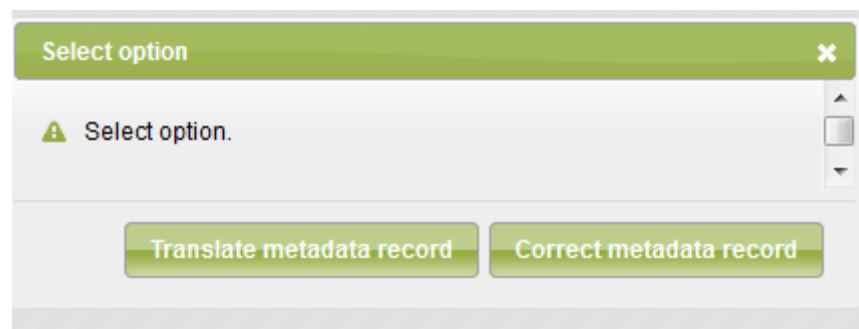
## 3.3 Beschreibung der User-Generated-Content Szenarien

Dieser Abschnitt setzt sich mit den drei User-Generated-Content Szenarien auseinander. Es werden die Szenarien genauer beschrieben, indem die zusammengehörenden Transaktionen, die von dem Organic.Edunet Benutzer im Dialog mit der User-Generated-Content-Komponente ausgeführt werden, bis der bearbeitete Metadatenatz gespeichert ist.

### 3.3.1 Szenario 1: Einen bestehenden Metadatenatz bearbeiten

Folgende Auflistung beschreibt die Prozessschritte, bis der Organic.Edunet Benutzer einen unvollständigen oder inkorrekten Metadatenatz korrigiert bzw. vervollständigt hat.

1. Der Organic.Edunet Benutzer lokalisiert einen inkorrekten Metadatenatz im Organic.Edunet Web-Portal und klickt auf diesen. Der Dialog (siehe Abbildung 3-2) öffnet sich und fragt den Organic.Edunet Benutzer, ob er den aktuellen Metadatenatz korrigieren oder übersetzen will. Der Organic.Edunet Benutzer klickt auf den Button „Correct Metadata Record“.



**Abbildung 3-2: Screenshot: Auswahl der User-Generated-Content Operation**

2. Das „Correct Metadata Record“-Widget (siehe Abbildung 3-3) öffnet sich, die einzelnen Metadatenelemente sind im Widget vorausgefüllt.

The screenshot shows a web interface for editing a metadata record. At the top, there is a green header bar with the text 'Correct metadata record' and a close button (X). Below the header, there are three tabs: 'General', 'Author', and 'Rights', with 'General' selected. The form contains several sections: 'TITLE' with a text input field containing '2001 Consultation Of Organic Farming Research Priorities In The U'; 'DESCRIPTION' with a text area containing 'The following reports a consultation of organic farming research priorities in the UK in 2001. The information was collated by Organic Centre Wales, Institute of Rural Studies, University of Wales Aberystwyth on behalf of MAFF organic'; 'LANGUAGE' with a dropdown menu set to 'English'; and 'KEYWORD +' with a list of five keywords: 'organic farming', 'organic research', 'agricultural practice', 'Technology transfer', and 'farming research'. Each keyword has a small green 'X' button to its right. At the bottom right of the form, there are two green buttons: 'Save' and 'Cancel'.

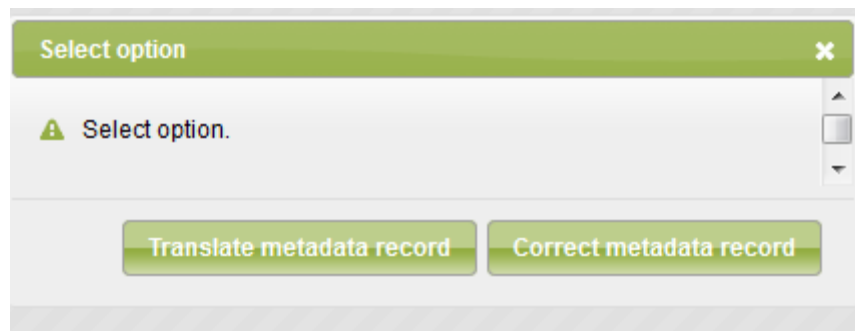
**Abbildung 3-3: Screenshot: „Correct Metadata Record“ Widget**

3. Der Organic.Edunet Benutzer korrigiert oder vervollständigt mit Hilfe des User-Generated-Content Widgets den Metadatensatz.
4. Hat der Organic.Edunet Benutzer die Korrektur des Metadatensatzes abgeschlossen, klickt er auf den „Save“-Button. Durch das Klicken auf den „Save“-Button wird dieser im User-Generated-Content Repository gespeichert.

### 3.3.2 Szenario 2: Einen bestehenden Metadatensatz übersetzen

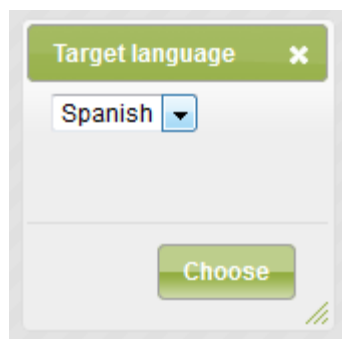
Folgende Auflistung beschreibt die einzelnen Dialog-Schritte, bis der Organic.Edunet Benutzer einen Metadatensatz übersetzt hat.

1. Ein spanischer Organic.Edunet Benutzer findet im Organic.Edunet Web-Portal einen Metadatensatz, der noch nicht in der spanischen Sprache vorliegt. Aus diesem Grund klickt er auf den Titel der Lernressource. Der Dialog (siehe Abbildung 3-4) öffnet sich und fragt den Organic.Edunet Benutzer, ob er den aktuellen Metadatensatz korrigieren oder übersetzen will. Der Organic.Edunet Benutzer klickt auf den Button „Translate Metadata Record“.




**Abbildung 3-4: Screenshot: Auswahl der User-Generated-Content Operation**

2. Ein Dialog für die Auswahl der Zielsprache öffnet sich. Der Organic.Edunet Benutzer wählt die Sprache „Spanish“ aus und klickt auf den Button „Choose“.



**Abbildung 3-5: Screenshot: Auswahl der Ziel-Sprache**

3. Das „Translate Metadata Record“-Widget (siehe Abbildung 3-6) öffnet sich, die einzelnen Metadatenelemente, die durch die User-Generated-Content Komponente automatisch in die Zielsprache übersetzt wurden, sind im Widget vorausgefüllt.



**Abbildung 3-6: Screenshot: „Translate Metadata Record“ Widget**

4. Der Organic.Edunet Benutzer korrigiert und vervollständigt den automatisch übersetzten Metadatensatz.
5. Hat der Organic.Edunet Benutzer die Übersetzung des Metadatensatzes abgeschlossen, klickt er auf den „Save“-Button. Durch das Klicken auf den

„Save“-Button wird dieser im User-Generated-Content Repository gespeichert.

### 3.3.3 Szenario 3: Eine neue Lernressource hinzufügen

Dieser Abschnitt erläutert die Prozessschritte für das Hinzufügen einer neuen Lernressource zum Organic.Edunet Web-Portal anhand des User-Generated-Content Szenarios „Eine neue Lernressource hinzufügen“.

Um das User-Generated-Content Widget für das Vorschlagen von neuen Lernressourcen aufrufen zu können, muss der Organic.Edunet Benutzer den „Suggest Learning Resource“ Button zur Lesezeichen-Symbolleiste hinzufügen. Dies kann der Organic.Edunet Benutzer durch Drag-and-Drop realisieren (siehe Abbildung 3-7).

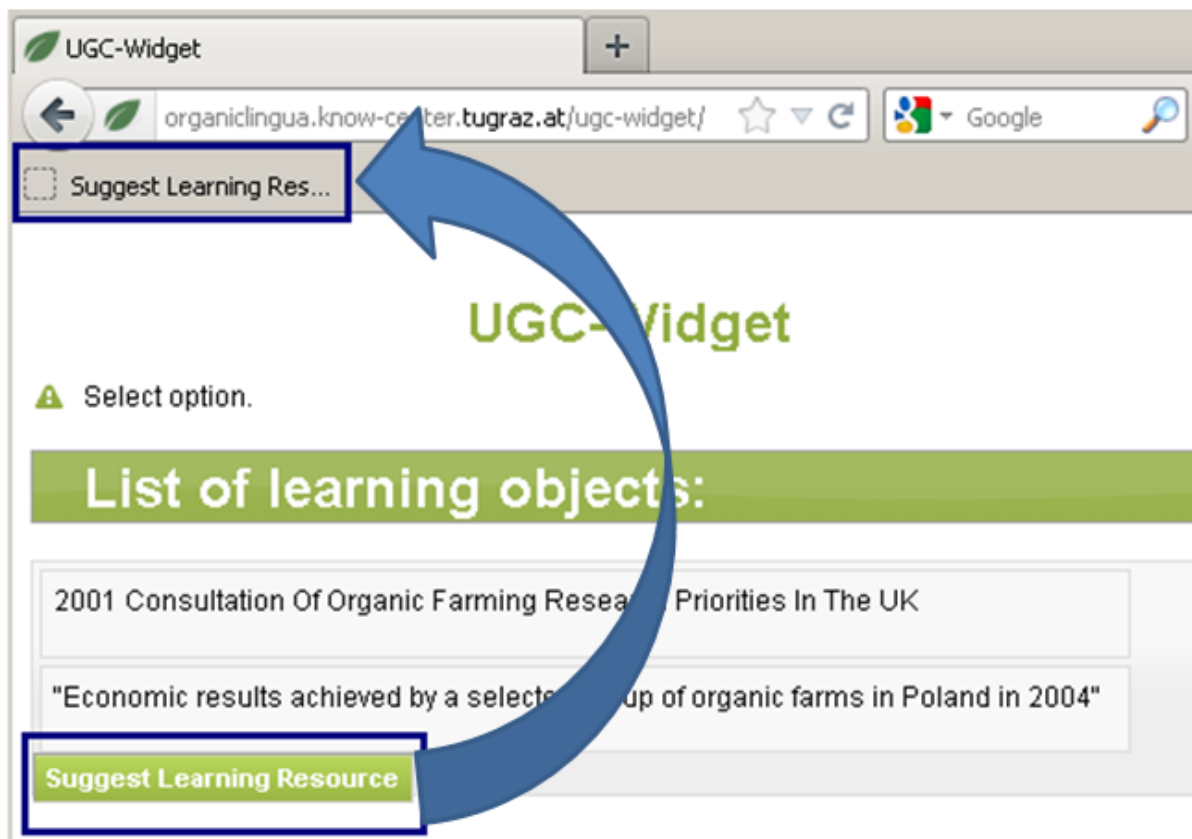
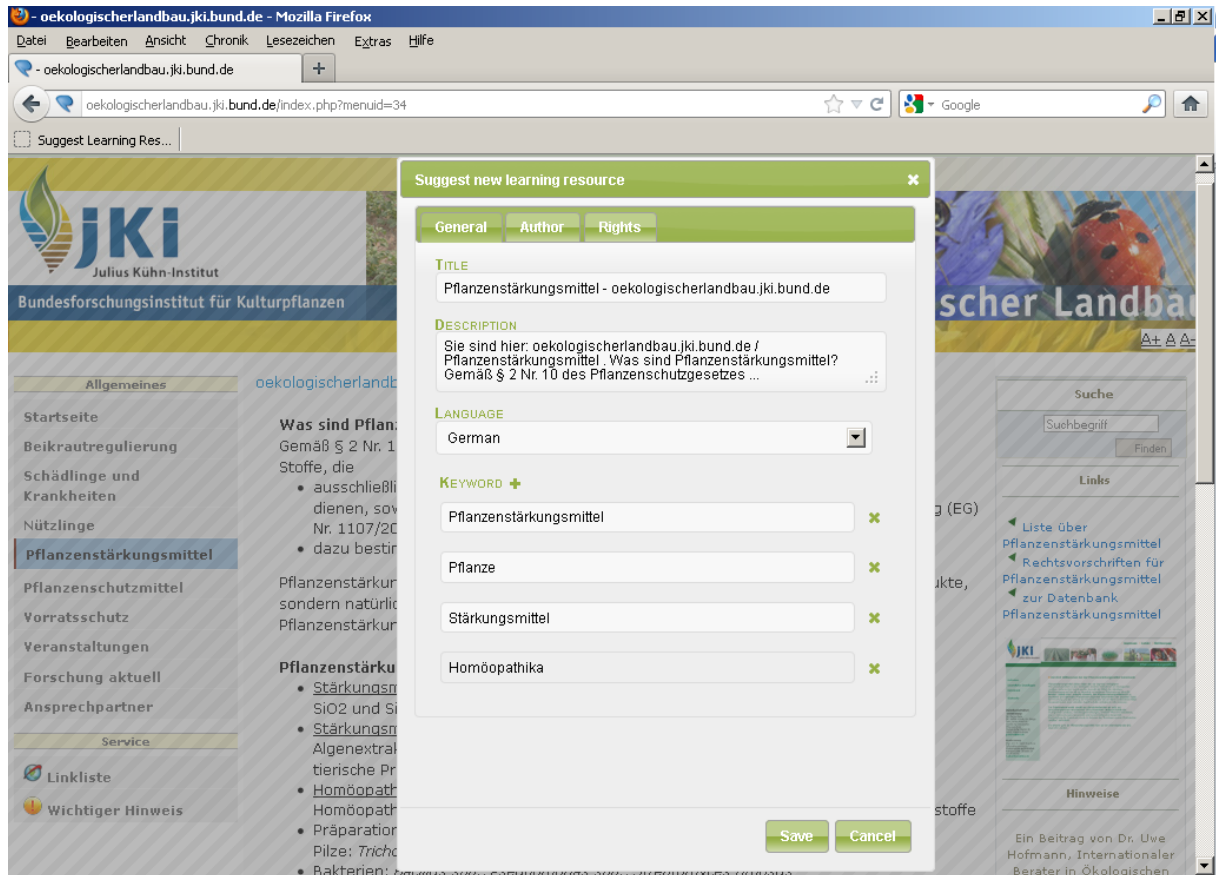


Abbildung 3-7: Screenshot: „Suggest Learning Resource“ Button zur Lesezeichen-Symbolleiste hinzufügen

Folgende Auflistung beschreibt die einzelnen Dialog-Schritte, bis der Organic.Edunet Benutzer eine neue Lernressource vollständig annotiert hat:

1. Ein Organic.Edunet Benutzer findet im Web eine interessante Lernressource zum Thema biologische Landwirtschaft, die noch nicht im Organic.Edunet Repository erfasst wurde. Der Organic.Edunet Benutzer klickt auf sein Lesezeichen „Suggest Learning Resource“ in der Lesezeichen-Symbolleiste (siehe Abbildung 3-7).
2. Die User-Generated-Content Komponente extrahiert von der aktuell geöffneten Webseite mit Hilfe von Information-Retrieval-Services die Metadaten der neuen Lernressource. Für eine genauere Beschreibung der Metadatenextraktion siehe Abschnitt 3.7.2. Da das Extrahieren je nach Lernressource einige Zeit dauern kann, wird dem Benutzer ein Ladesymbol angezeigt, somit ist er immer über den aktuellen Fortschritt informiert.
3. Nachdem der Metadatenextraktionsprozess abgeschlossen ist, öffnet sich das „Suggest New Learning Resource“-Widget (siehe Abbildung 3-8). Das Widget überdeckt die aktuelle Lernressource und ist in der Webseite integriert. Die einzelnen Metadatenelemente, die durch die User-Generated-Content Komponente automatisch extrahiert wurden, sind im Widget vorausgefüllt.

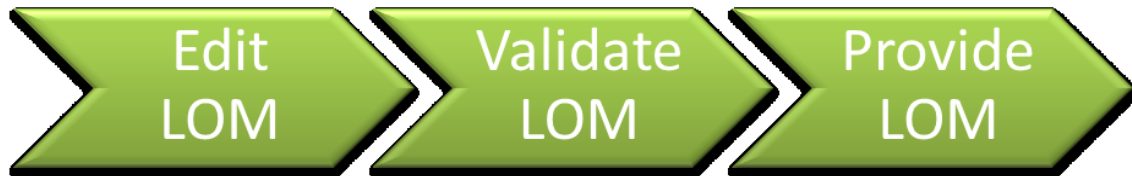


**Abbildung 3-8: Screenshot: „Suggest NEW Learning Resource“ Widget**

4. Der Organic.Edunet Benutzer korrigiert und vervollständigt die automatisch extrahierten Metadatenelemente.
5. Hat der Organic.Edunet Benutzer die Annotierung der neuen Lernressource abgeschlossen, klickt er auf den „Save“-Button. Durch Klicken auf den „Save“-Button wird die Annotierung der Lernressource im User-Generated-Content Repository gespeichert.

## 3.4 Lebenszyklus der Lernressourcen-Metadatenbeschreibung

In diesem Abschnitt wird der Lebenszyklus einer neuen bzw. bearbeiteten Metadatenbeschreibung genauer erklärt.



**Abbildung 3-9: Publikationsprozess**

Der Veröffentlichungsprozess einer benutzererstellten Metadatenbeschreibung besteht aus drei Schritten (siehe Abbildung 3-9):

- **Ein Metadatensatz wird bearbeitet, übersetzt oder neu erstellt**

Der Organic.Edunet Benutzer bearbeitet bzw. übersetzt einen bestehenden Metadatensatz oder erstellt einen neuen Metadatensatz für eine neue Lernressource. Diese vom Organic.Edunet Benutzer bearbeiteten oder hinzugefügten Metadatensätze werden im User-Generated-Content Repository gespeichert.

- **Validierung der endbenutzererstellten Metadatensätze**

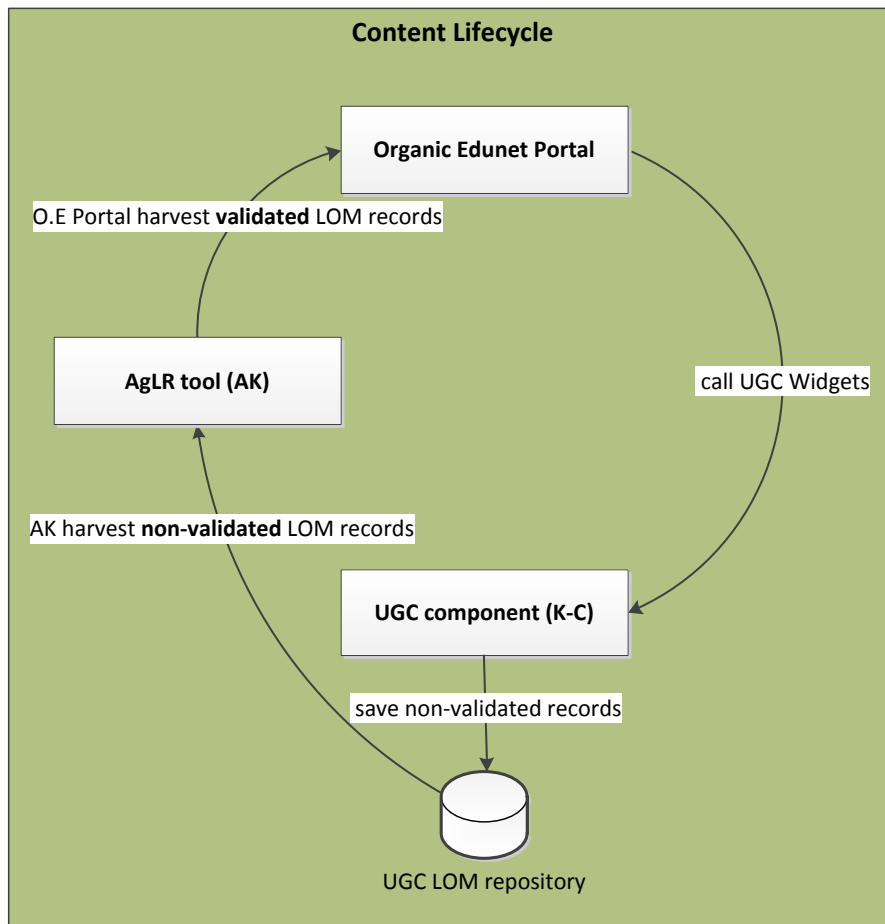
Das AgLR Tool ruft die nicht-validierten Metadatensätze, die im User-Generated-Content Repository gespeichert sind, ab. Das AgLR Tool ist eine Anwendung für Experten, die es den Experten ermöglicht, Metadatensätze zu validieren und gegebenenfalls zu korrigieren und zu vervollständigen. Für das Bereitstellen der nicht-validierten Metadaten durch die User-Generated-Content Komponente und das Abrufen der Metadatensätze wird auf den OAI-PMH Standard zurückgegriffen. Dieser OAI-PMH Standard wird im Abschnitt 2.6 genauer erklärt und die konkrete Umsetzung durch die User-Generated-Content Komponente wird im Abschnitt 4.5 skizziert. Anschließend überprüfen Experten die



von Organic.Edunet Benutzern bearbeiteten bzw. erstellten Metadaten mit Hilfe des AgLR Tools. Die Experten haben drei Optionen, wie sie mit den nicht-validierten Metadatenätzen verfahren. Erstens können sie die Metadatenätze ohne Anpassung für das Organic.Edunet Web-Portal veröffentlichen. Zweitens können sie, bevor die Metadatenätze für das Organic.Edunet Web-Portal veröffentlicht werden, diese korrigieren und vervollständigen. Drittens können die Experten den veränderten bzw. neuen Metadatenatz vom User-Generated-Content Repository ablehnen und somit nicht für das Organic.Edunet Web-Portal veröffentlichen.

- **Bereitstellung der nutzererstellten Metadatenätze für das Organic.Edunet Web-Portal:**

Nachdem die Experten mit Hilfe des AgLR Tools die benutzererstellten Metadaten validiert und mit weiteren Metadaten angereichert haben, werden die Metadatenätze mit Hilfe des OAI-PMH Standards für das Organic.Edunet Web-Portal veröffentlicht. Das Organic.Edunet Web-Portal ruft die Metadatenätze ab und somit werden die Änderungen durch die Organic.Edunet Benutzer wieder für die gesamte Organic.Edunet Community sichtbar.



**Abbildung 3-10: Lebenszyklus der Metadaten**

Der gesamte Lernressourcen-Lebenszyklus ist in der Abbildung 3-10 dargestellt und wird in den folgenden Absätzen genauer erklärt.

Der Organic.Edunet Benutzer kann bestehende Metadatensätze bearbeiten und übersetzen. Außerdem kann der Organic.Edunet Benutzer neue Lernressourcen vorschlagen, indem er einen neuen Metadatensatz für diese anlegt. Eine neue Lernressource wird der Benutzer vorschlagen, wenn er sie für das Organic.Edunet Web-Portal für relevant hält und mit der Organic.Edunet Community teilen will.

Um einen neuen Metadatensatz anzulegen oder einen bestehenden Metadatensatz zu bearbeiten, öffnet der Organic.Edunet Benutzer das entsprechende User-Generated-Content Widget. Diese Widgets sind im Organic.Edunet Web-Portal integriert. Die Integration der User-Generated-Content Widgets wird in Abschnitt 4.4 ausführlich erklärt.

Nachdem der Organic.Edunet Benutzer den Metadatensatz mit Hilfe des User-Generated-Content Widgets bearbeitet hat, wird der neue Metadatensatz im User-Generated-Content Repository gespeichert. In dem Fall, dass der Organic.Edunet Benutzer einen existierenden Metadatensatz im Organic.Edunet Web-Portal bearbeitet oder übersetzt, wird im User-Generated-Content Repository eine Kopie des originalen Metadatensatzes gespeichert. Dieser neue Metadatensatz referenziert auf den ursprünglichen Metadatensatz. Die Beziehung zwischen dem neuen Metadatensatz und dem originalen Metadatensatz wird im Abschnitt 3.5.1 ausgeführt.

Die vom Organic.Edunet Benutzer erstellten Metadatensätze werden als nicht-validierte Metadatensätze im User-Generated-Content Repository gespeichert. Diese nicht-validierten Metadatensätze werden mit Hilfe des OAI-PMH Standards veröffentlicht und die AgLR Anwendung ruft die veröffentlichten Metadatensätze ab und speichert sie im AgLR Repository. Das Interface zwischen der User-Generated-Content Komponente und der AgLR Anwendung wird im Abschnitt 4.5 anhand einer Beispiel-Anfrage und -Antwort erklärt.

Anschließend werden die nicht-validierten, endbenutzererstellten Metadaten, die im AgLR Repository gespeichert sind, durch Experten mit weiteren Informationen, wie z. B. Ontologie-Konzepten, angereichert und validiert.

Zum Schluss werden die validierten Metadatensätze durch das AgLR Tool veröffentlicht. Dieser Veröffentlichungsprozess basiert wieder auf dem OAI-PMH Standard. Das Organic.Edunet Web-Portal ruft in bestimmten Intervallen diese veröffentlichten Metadaten ab und somit werden die Veränderungen durch den Organic.Edunet Benutzer im Organic.Edunet Web-Portal sichtbar.

### **3.5 Das Organic.Edunet IEEE LOM Application Profile**

Im Organic.Lingua Projekt wird eine adaptierte Version des IEEE LOM Standards verwendet, um eine Lernressource zu beschreiben. Aufgrund der Adaptierung des IEEE LOM Standards werden die Wiederverwendung und der Austausch von Lernressourcen in verteilten Systemen, wie es das Organic.Edunet darstellt, erleichtert. Die folgende Tabelle 3-1 gibt einen

Überblick über das aktuelle Organic.Edunet IEEE LOM Application Profile und die Verwendung der einzelnen LOM Elemente in den drei User-Generated-Content Szenarien (siehe Abschnitt 3.1 )

Die zweite Spalte „Mandatory“ der Tabelle 3-1 zeigt die verpflichtenden Metadaten-Elemente des Organic.Edunet IEEE LOM Application Profiles, die im Organic.Lingua Deliverable 3.1.2.1 (Sanchez, Stoitsis , Lezcano, Manouselis, & Dimitropoulos, 2011) beschrieben und definiert sind. Die „Mandatory“ Metadatenelemente, die mit einem „X“ gekennzeichnet sind, werden vollständig von allen drei oben genannten User-Generated-Content Szenarien unterstützt. Die „Mandatory“ Metadatenelemente, die mit „(X)“ gekennzeichnet sind, werden nur teilweise oder gar nicht von den User-Generated-Content Szenarien unterstützt. Das Metadatenelement „3.2. contribute“ wird zum Beispiel beim Übersetzen eines bestehenden Metadatensatzes nicht berücksichtigt, weil die Übersetzung dieses Elements nicht sinnvoll ist.

Außerdem wurde das existierende Organic.Edunet LOM Application Profile (Sanchez, Stoitsis , Lezcano, Manouselis, & Dimitropoulos, 2011) um das Metadatenelement „3.5 Relation“ erweitert. Diese Erweiterung ist im Abschnitt 3.5.1 genauer erklärt.

	<b>Mandatory</b>	<b>Revise LOM</b>	<b>Translate LOM</b>	<b>Suggest Learning Resource</b>
<b>1 GENERAL</b>	<b>x</b>			
1.1 Identifier	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
1.1.1 Catalog	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
1.1.2 Entry	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
1.2 Title	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
1.3 Language	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
1.4 Description	<b>x</b>	<b>x</b>	<b>x</b>	<b>x</b>
1.5 Keyword		<b>x</b>	<b>x</b>	<b>x</b>
1.6 Coverage				
1.7 Structure				
1.8 Aggregation Level				
<b>2 LIFE CYCLE</b>				
2.1 Version				
2.2 Status				

2.3 Contribute		x		x
2.3.1 Role	(x)	x		x
2.3.2 Entity	(x)	x		x
2.3.3 Date	(x)	x		x
<b>3 META-METADATA</b>				
3.1 Identifier		x	x	x
3.1.1 Catalog	x	x	x	x
3.1.2 Entry	x	x	x	x
3.2 Contribute		x	x	x
3.2.1 Role	x	x	x	x
3.2.2 Entity	x	x	x	x
3.2.3 Date	x	x	x	x
3.3 Metadata Schema				
3.4 Language				
3.5 Relation				
3.5.1 Kind		x	x	
3.5.2 Resource		x	x	
3.5.2.1 Identifier		x	x	
3.5.2.1.1 Catalog		x	x	
3.5.2.1.2 Entry		x	x	
<b>4 TECHNICAL</b>				
4.1 Format	(x)	x		x
4.2 Size				
4.3 Location	(x)	x		x
4.4 Requirement				
4.5 Installation Remarks				
4.6 Other Platform Requirements				
4.7 Duration				
<b>5. EDUCATIONAL</b>				
5.1 Interactivity Type				
5.2 Learning Resource Type				
5.3 Interactivity Level				
5.4 Semantic Density				
5.5 Intended End User Role				
5.6 Context				
5.7 Typical Age Range				
5.8 Difficulty				
5.9 Typical Learning Time				
5.10 Description				
5.11 Language				
<b>6 RIGHTS</b>	(x)	x		x
6.1 Cost	(x)	x		x
6.2 Copyright and Other Restrictions	(x)	x		x

6.3 Description		<b>x</b>		<b>x</b>
<b>7 RELATION</b>				
7.1 Kind				
7.2 Resource				
7.2.1 Identifier				
7.2.1.1 Catalog	<b>(x)</b>			
7.2.1.2 Entry	<b>(x)</b>			
7.2.2 Description	<b>(x)</b>			
<b>8 ANNOTATION</b>				
8.1 Entity				
8.2 Date				
8.3 Description				
<b>9 CLASSIFICATION</b>				
9.1 Purpose				
9.2 Taxon Path				
9.3 Description				
9.4 Keyword				

**Tabelle 3-1: Organic.Edunet IEEE LOM Application Profile**

In den folgenden Abschnitten werden die Metadatenelemente „3.5 Relation“ (siehe Abschnitt 3.5.2) und „3.2 Contribute“ (siehe Abschnitt 3.5.13.5.2) behandelt.

### 3.5.1 Referenz auf den originalen Metadatensatz

Die User-Generated-Content Komponente ermöglicht dem Organic.Edunet Benutzer, bestehende Metadatensätze vom Organic.Edunet Web-Portal zu übersetzen und zu überarbeiten. In diesen zwei genannten Modifizierungsszenarien wird eine Kopie vom originalen Metadatensatz im User-Generated-Content Repository gespeichert. Aus diesem Grund wurde das Organic.Edunet IEEE LOM Application Profile überarbeitet, um auf den originalen, ursprünglichen Metadatensatz zu referenzieren. Diese Referenzierung wurde durch eine neue Organic.Edunet IEEE LOM meta-metadata Kategorie mit der Bezeichnung „3.5 Relation“ realisiert. Die neue Kategorie ist in der Tabelle 3-1 mit der Farbe Grün geschrieben. Die Metadatenelemente der Kategorie „3.5 Relation“ werden automatisch bei folgenden Szenarien gesetzt:

- einen bestehenden Metadatensatz bearbeiten
- einen bestehenden Metadatensatz übersetzen

Die neue Kategorie „3.5 Relation“ besteht aus dem Metadatenelement „3.5.1 Kind“ und der Metadatenkategorie „3.5.2 Resource“. Das Metadatenelement „3.5.1 Kind“ beschreibt die Beziehung zwischen dem aktuellen und dem originalen Metadatensatz. Der originale Metadatensatz wird durch die Metadatenkategorie „3.5.2 Resource“ definiert. Das Metadatenelement „3.5.1 Kind“ kann zwei Werte annehmen. Der erste Wert ist „isrevisedversionof“, dieser definiert, dass der aktuelle Metadatensatz eine überarbeitete Version des originalen Metadatensatzes ist, der durch die Kategorie „3.5.2 Resource“ definiert ist. Der zweite Wert ist „istranslatedversionof“. Diese Beziehung beschreibt, dass der aktuelle Metadatensatz eine Übersetzung vom originalen Metadatensatz darstellt. Folgende zwei Codeausschnitte (siehe Codeausschnitt 3-1 und Codeausschnitt 3-2) zeigen Beispiele für Organic.Edunet IEEE LOM Relation Kategorien:

```
-<relation>
  <kind>isrevisedversionof</kind>
  -<resource>
    -<identifier>
      <catalog>URI</catalog>
      <entry>http://orgprints.org/11021/</entry>
    </identifier>
  </resource>
</relation>
```

**Codeausschnitt 3-1: IEEE LOM Relation: Verweis auf originalen Metadatensatz bei der Bearbeitung eines existierenden Metadatensatzes**

```
-<relation>
  <kind>istranslatedversionof</kind>
  -<resource>
    -<identifier>
      <catalog>URI</catalog>
      <entry>http://orgprints.org/7102/</entry>
    </identifier>
  </resource>
</relation>
```

**Codeausschnitt 3-2: IEEE LOM Relation: Verweis auf originalen Metadatensatz bei der Übersetzung eines existierenden Metadatensatzes**

### 3.5.2 Bearbeiter des Metadatensatzes

Der Verfasser und der Bearbeiter von Metadatensätzen werden in der Kategorie „3.2 Contribute“ vom Organic.Edunet IEEE LOM Application Profile gespeichert. Diese Information über den Organic.Edunet Benutzer, der am Metadaten-Lebenszyklus beteiligt ist, wird automatisch im Metadatensatz gespeichert. Um dies bewerkstelligen zu können, wird das existierende Organic.Edunet IEEE LOM Application Profile erweitert, indem der Wertebereich des Metadatenelements „3.2.1 Role“ erweitert wird. Dem Wertebereich des Metadatenelements „3.2.1 Role“ werden die Werte „enricher“ und „translator“ hinzugefügt. Somit wird ermöglicht, den Bearbeiter bzw. Übersetzer eines existierenden Metadatensatzes zu speichern.

Die folgende Tabelle 3-2 zeigt, welche Rolle der Metadaten-Contributor in den verschiedenen Use-Cases hat:

Use case	Wert des Metadatenelements '3.2.1 Role'
Einen bestehenden Metadatensatz überarbeiten	enricher
Einen bestehenden Metadatensatz übersetzen	translator
Eine neue Lernressource hinzufügen	creator

**Tabelle 3-2: Rolle des Metadaten-Contributors**

### 3.6 Nutzungsrechte der Lernressource

In diesem Abschnitt wird die Problematik des Nutzungsrechtes von Lernressourcen beschrieben. Außerdem wird die „Creative Commons“ Lizenz genauer beschrieben, die vollständig in die User-Generated-Content Komponente implementiert wurde.

Der Organic.Edunet Benutzer kann mit Hilfe der User-Generated-Content Komponente neue Lernressourcen hinzufügen. Bei Hinzufügen neuer



Lernressourcen muss der Organic.Edunet Benutzer die Nutzungsrechte für das Lernobjekt definieren.

Der IEEE LOM Standard (IEEE Learning Technology Standards Committee, 2002) sieht für die Speicherung der Nutzungsrechte eine eigene Kategorie vor. Diese Metadaten Kategorie namens „6 Rights“ beschreibt die Rechte am geistigen Eigentum und Nutzungsbedingungen des Lernobjekts. Die Metadaten Kategorie „6 Rights“ besteht aus folgenden drei Metadatenelementen (IEEE Learning Technology Standards Committee, 2002):

- **6.1 Cost**

Der Wertebereich dieses Metadatenelements umfasst „yes“ und „no“ und beschreibt, ob die Verwendung des Lernobjekts eine Zahlung erfordert.

- **6.2 Copyright and Other Restriction**

Der Wertebereich dieses Metadatenelements umfasst „yes“ und „no“ und beschreibt das Urheberrecht oder andere Einschränkungen für die Nutzung des Lernobjekts.

- **6.3 Description**

Dieses Metadatenelement beschreibt textuell die Bedingungen für die Nutzung des Lernobjekts, sofern ein Urheberrecht oder andere Einschränkungen für die Nutzung des Lernobjekts bestehen.

The screenshot shows a dialog box titled "Correct metadata record" with a close button (X) in the top right corner. Below the title bar are three tabs: "General", "Author", and "Rights". The "Rights" tab is active. Under the "Rights" tab, there are three sections: "COST" with a dropdown menu set to "no"; "COPYRIGHT RESTRICTIONS" with a dropdown menu set to "yes"; and "CREATIVE COMMONS LICENSE (VERSION 3.0) Info" with a radio button that is not selected. Below this is "CUSTOM LICENSE" with a radio button that is selected. Under "CUSTOM LICENSE" is a "DESCRIPTION" text area with a small grid icon in the bottom right corner.

**Abbildung 3-11: Screenshot: Nutzungsrechte „Custom License“**

The screenshot shows the same "Correct metadata record" dialog box. The "Rights" tab is still active. The "COST" dropdown is "no" and "COPYRIGHT RESTRICTIONS" is "yes". Under "CREATIVE COMMONS LICENSE (VERSION 3.0) Info", the radio button is now selected. Below it are two dropdown menus: "Allow modifications?" set to "No" and "Allow commercial uses?" set to "No". Below these is "CUSTOM LICENSE" with a radio button that is not selected.

**Abbildung 3-12: Screenshot: Nutzungsrechte „Creative Commons License“**

Der Organic.Edunet Benutzer hat bei der Annotierung von neuen Lernressourcen zwei Möglichkeiten, um die Nutzungsrechte zu definieren. Erstens kann der Organic.Edunet Benutzer eine benutzerdefinierte Urheberrechtslizenz beschreiben (siehe Abbildung 3-11). Diese Beschreibung wird im Metadatenelement „6.3 Description“ gespeichert. Zweitens kann er im

User-Generated-Content Widget die entsprechende Creative Commons Lizenz<sup>29</sup> auswählen (siehe Abbildung 3-12) und mit Hilfe dieser, die Nutzungsrechte für die neue Lernressource definieren.

Creative Commons ist eine gemeinnützige Organisation, die Lizenzen für den Austausch und die Verwendung von Wissen zur Verfügung stellt. Die Creative Commons Lizenzen sind kostenlos und bieten einen einfachen und standardisierten Weg, nach benutzerdefinierten Bedingungen „kreative Arbeit“ zu veröffentlichen und zu teilen.<sup>30</sup>

Die User-Generated-Content Komponente unterstützt die Creative Commons Lizenz in der Version 3.0 um Vervielfältigungs- und Verbreitungsrechte für die neuen Lernressourcen zu beschreiben. Der Organic.Edunet Benutzer kann bei Hinzufügen neuer Lernressourcen eine aus sechs möglichen Creative Commons Lizenzen wählen, um das Urheberrecht festzulegen.<sup>31</sup>

Creative Commons (2012c) definiert folgende sechs Lizenzen:

- **Namensnennung (CC BY)**

Diese Lizenz erlaubt anderen Personen, den Lerninhalt zu verbreiten, zu bearbeiten, zu verbessern und zu erweitern. Außerdem ist eine kommerzielle Nutzung der Lernressource erlaubt, solange der originale Urheber angeführt wird.

- **Namensnennung und Weitergabe unter gleichen Bedingungen (CC BY-SA)**

Diese Lizenz erlaubt anderen Personen, den Lerninhalt zu verbreiten, zu bearbeiten, zu verbessern und zu erweitern, solange auf den Urheber der originalen Lernressource verwiesen wird. Außerdem ist

---

<sup>29</sup> Creative Commons. (2012). Abgerufen am 2. Juli 2012 von Creative Commons : <http://creativecommons.org/>

<sup>30</sup> Creative Commons. (2012). About. Abgerufen am 2. Juli 2012 von Creative Commons: <http://creativecommons.org/about>

<sup>31</sup> Creative Commons. (2012). About The Licenses. Abgerufen am 2. Juli 2012 von Creative Commons: <http://creativecommons.org/licenses/>

die kommerzielle Nutzung erlaubt und modifizierte Lerninhalte, die auf den originalen Lerninhalt aufbauen, müssen unter der gleichen Lizenz veröffentlicht werden.

- **Namensnennung und keine Bearbeitung (CC BY-ND)**

Diese Lizenz erlaubt anderen Personen, die Lernressource weiterzuverbreiten. Der Lerninhalt muss aber vollständig weiterverbreitet werden und darf nicht modifiziert werden. Weiters ist die kommerzielle Nutzung erlaubt, solange der originale Autor des Lerninhaltes genannt wird.

- **Namensnennung und nicht kommerzielle Nutzung (CC BY-NC)**

Diese Lizenz erlaubt anderen Personen, den Lerninhalt zu verbreiten, zu bearbeiten, zu verbessern und zu erweitern, solange auf den Urheber der originalen Lernressource verwiesen wird. Außerdem ist die kommerzielle Nutzung nicht erlaubt.

- **Namensnennung, nicht kommerzielle Nutzung und Weitergabe unter gleichen Bedingungen (CC BY-NC-SA)**

Diese Lizenz erlaubt anderen Personen, den Lerninhalt zu verbreiten, zu bearbeiten, zu verbessern und zu erweitern, solange der Autor des originalen Lerninhaltes angeführt wird. Außerdem ist die kommerzielle Nutzung nicht erlaubt und modifizierte Lerninhalte, die auf den originalen Lerninhalt aufbauen, müssen unter der gleichen Lizenz veröffentlicht werden.

- **Namensnennung, nicht kommerzielle Nutzung und keine Bearbeitung (CC BY-NC-ND)**

Diese Lizenz erlaubt anderen Personen, die Lernressource weiter zu verbreiten, solange der Autor der Lernressource angeführt wird. Der Lerninhalt darf aber nicht verändert und nicht kommerziell genutzt werden. Ergo ist diese Lizenz die strengste bzw. restriktivste der sechs Creative Common Lizenzen.

Wenn keine der sechs angeführten Creative Common Lizenzen passend für die neue Lernressource ist, kann der Organic.Edunet Benutzer eine benutzerdefinierte Lizenz erstellen (siehe Abbildung 3-11).

## **3.7 User-Generated-Content Services**

Ein Ziel der User-Generated-Content Komponente besteht darin, dass sie eine sehr intuitive und einfache Benutzeroberfläche bereitstellt, die es ermöglicht, mehrsprachige Lernressourcen zu annotieren. Das Annotieren von Lernressourcen umfasst das Korrigieren, Erweitern und Übersetzen von existierenden Metadatenätzen und die Möglichkeit, dem User-Generated-Content Repository neue Lernressourcen hinzuzufügen. Die User-Generated-Content-Komponente ermöglicht die automatische Übersetzung von vorhandenen Metadatenätzen im Organic.Edunet Web-Portal. Außerdem wird der Benutzer beim Hinzufügen einer Lernressource durch eine automatische Metadatenextraktion unterstützt. Darüber hinaus werden dem Benutzer bei der Eingabe von Keywords bereits existierende Keywords im Organic.Edunet Repository vorgeschlagen. Um falschen Schreibweisen von Wörtern, wie beispielsweise Tippfehlern, entgegen zu wirken, wird die Eingabe durch den Organic.Edunet Benutzer einer automatischen Rechtschreibkorrektur unterzogen.

Diese Services, Rechtschreibkorrektur, automatische Übersetzung, Auto-Suggestion und Metadatenextraktion, erhöhen die Qualität der Metadaten (siehe Abschnitt 2.5).

Die User-Generated-Content Services werden in den folgenden Abschnitten genauer erklärt.

### **3.7.1 Übersetzungs-Service**

Die Übersetzungsservices werden aufgerufen, wenn der Organic.Edunet Benutzer eine automatische Übersetzung eines existierenden Metadatenatzes im Organic.Edunet Web-Portal anfordert. Das Übersetzungsservice übersetzt das Organic.Edunet IEEE LOM Application Profile. Das Übersetzungsservice

kombiniert die Machine-Translation-Services und dictionary-based Übersetzungsservices der Organic.Lingua Projektpartner Xerox und Celi. Die Übersetzungsservices der Projektpartner sind im Organic.Edunet Deliverable 4.1 (Nikoulina, et al., 2011) genauer beschrieben. (Kohler & Lex, 2012)

Die folgende Auflistung zeigt die IEEE LOM Elemente auf, die durch die Hilfe der Übersetzungsservices der Organic.Lingua Projektpartner übersetzt werden (Kohler & Lex, 2012):

- General.Title (Xerox)
- General.Description (Xerox)
- General.Coverage (Celi)
- Educational.TypicalAgeRange (Celi)
- General.Keyword (Celi)

### **3.7.2 Metadatenextraktions-Service**

Die Metadatenextraktionsservices unterstützen den Benutzer bei der Erstellung von Metadatenätzen für eine neue Lernressource, die der Organic.Edunet Benutzer zum User-Generated-Content Repository hinzufügen will. Diese Services ermöglichen dem Benutzer, schneller eine neue Lernressource zu annotieren, weil folgende Metadatenelemente durch die Metadatenextraktionsservices automatisch extrahiert werden:

- General.Title
- General.Description
- General.Keywords
- General.Language

Die automatisch extrahierten Metadatenelemente werden beim Öffnen des User-Generated-Content Widgets vorausgefüllt. Somit muss der Organic.Edunet Benutzer die einzelnen Metadatenelemente nur gegebenenfalls korrigieren bzw. anpassen. Dieser kombinierte Ansatz von automatischer Metadatenextraktion

und anschließender Benutzer-Interaktion wird als halb-automatische Metadatengenerierung bezeichnet (Dahl & Vossen, 2008).

Die Sprache der neuen Lernressource wird mit Hilfe des Language-Identifizier-Services von Celi bestimmt. Dieses Language-Identifizier-Service von Celi ist im Organic.Lingua Deliverable 4.1 (Nikoulina, et al., 2011) genauer beschrieben.

Die Beschreibung der Lernressource wird automatisch generiert, indem die Key-Sentences vom Inhalt der neuen Lernressource extrahiert werden.

Bevor die Metadaten von der neuen Lernressource extrahiert werden können, wird mit Hilfe einer Format-Converters-Library der Inhalt der Webseite, die zum User-Generated-Content Repository hinzugefügt werden soll, extrahiert. Anschließend werden vom extrahierten Text die Wortarten der einzelnen Wörter innerhalb des Textes bestimmt. Für diese Bestimmung wird der CELI-POS-Tagger verwendet, der im Organic.Lingua Deliverable D4.1 (Nikoulina, et al., 2011) genauer beschrieben ist.

Zum Schluss werden die Beschreibung und die Schlüsselwörter vom Text extrahiert. Diese Extraktion basiert auf dem TextRank- Algorithmus von Rada Mihalcea und Paul Tarau (2004).

### **3.7.3 Autosuggestions-Service**

Bei dem Bearbeiten und Hinzufügen von neuen Metadatenätzen wird der Organic.Edunet Benutzer durch ein Autovervollständigungs-Service bei der Eingabe von Keywords durch die User-Generated-Content-Komponente unterstützt. Die Vorschläge für Keywords für den Metadatenatz basieren auf ähnlichen Metadatenätzen, die im Organic.Edunet Repository gespeichert sind. Das Service ruft die gespeicherten Metadatenätze vom Organic.Edunet Repository durch die OAI-PMH Schnittstelle ab. Anschließend werden die gesammelten Keywords indexiert. Das daraus resultierende Wörterbuch wird wöchentlich aktualisiert. Für das Autosuggestions-Service wird das Apache Lucene-Framework<sup>32</sup> verwendet.

---

<sup>32</sup> <http://lucene.apache.org/>

### **3.7.4 Rechtschreibkorrektur-Service**

Das Rechtschreibkorrektur-Service unterstützt den Organic.Edunet Benutzer bei der Bearbeitung, Übersetzung und dem Hinzufügen von neuen Metadatenätzen in das User-Generated-Content Repository. Bei der Eingabe von einem Keyword durch den Organic.Edunet Benutzer wird automatisch die Rechtschreibung der Keywords geprüft. Das Rechtschreibkorrektur-Service liefert einen Rechtschreibkorrektur-Vorschlag, wenn das Wort im erstellten Index vorhanden ist. Um den Index für die Keywords zu erstellen, werden alle existierenden Metadatenätze im Organic.Edunet Repository wöchentlich abgerufen und die Keywords in den existierenden Metadatenätzen werden dem Wörterbuch hinzugefügt. Auch dieses Information-Retrieval-Service basiert auf dem Apache Lucene-Framework<sup>33</sup>.

---

<sup>33</sup> <http://lucene.apache.org/>



## 4 Technische Aspekte der Implementierung

In diesem Kapitel wird auf technische Aspekte der Implementierung und der Integration der User-Generated-Content Komponente eingegangen. Zuerst werden die eingesetzten Technologien und die verwendeten Frameworks erklärt. Anschließend wird auf die Problematik der Cross-Domain Kommunikation eingegangen. Zum Schluss wird die Integration der User-Generated-Content Komponente erklärt.

### 4.1 Verwendete Technologien

Dieser Abschnitt „Verwendete Technologien“ stellt die verwendeten Programmiersprachen, Markup Languages und Server vor, die für die Implementierung der User-Generated-Content Komponente notwendig waren.

Als Web-Server dient ein Apache-Server in der Version 2.2.16. Die Implementierung der Business-Logic ist in der Programmiersprache PHP realisiert. Die aktuelle Version der User-Generated-Content Komponente verwendet PHP in der Version 5.3.3.7.

Die User-Generated-Content Widgets basieren auf den Technologien JavaScript, HTML und CSS. Daher sind die Widgets plattformunabhängig und die Benutzer müssen keine gesonderte Software, Plug-Ins, etc. installieren. Außerdem wurden bei der Entwicklung der Widgets die verschiedenen Webbrowser berücksichtigt, sodass die User-Generated-Content Widgets unter folgenden aktuellen Web-Browsern funktionieren:

- Mozilla Firefox
- Opera
- Safari
- Google Chrome
- Internet Explorer

Eine prototypische Entwicklung der User-Generated-Content Widgets kann unter folgender URL aufgerufen werden.

**URL des User-Generated-Content Widgets:**

<http://organiclingua.know-center.tugraz.at/ugc-widget/>

Um die Widgets zu testen und zu präsentieren, wurde eine Webseite entwickelt, die es ermöglicht, die drei User-Generated-Widgets aufzurufen. Diese Test-Webseite wird später durch das Organic.Edunet Web-Portal ersetzt und die User-Generated-Content Widgets werden direkt vom Organic.Edunet Web-Portal aufgerufen. Die folgende Abbildung 4-1 zeigt die Benutzeroberfläche der Test-Webseite.



**Abbildung 4-1: Test-Webseite**

## 4.2 Verwendete Frameworks und Bibliotheken

Dieser Abschnitt gibt einen Überblick über die verwendeten JavaScript-Frameworks, die in der User-Generated-Content Komponente eingesetzt wurden.

Für die Umsetzung der User-Generated-Content Komponente wurden folgende JavaScript Frameworks verwendet:

- **jQuery 1.6.2**

Die jQuery ist eine umfangreiche JavaScript-Klassenbibliothek. Die Klassenbibliothek kann unter den Lizenzbedingungen der MIT-Lizenz oder „General Public Licence“ (GPL) verwendet werden.<sup>34</sup>

In der User-Generated-Content Komponente wurden folgende Funktionen, die die jQuery Bibliothek zur Verfügung stellt, eingesetzt:

- Elementselektion in DOM

In Anlehnung an CSS unterstützt jQuery die Elementselektion in XML- und HTML-Dokumenten.<sup>35</sup>

Die Elementselektionsfunktion wurde verwendet, um bestimmte HTML-Elemente zu selektieren. Diese selektierten Dokumente wurden in weiterer Folge manipuliert, ausgelesen oder mit Events verknüpft.

- Events

Mit Hilfe der Event-Funktionalität können einzelnen HTML-Elementen Events zugewiesen werden. Der Benutzer löst die Events aus, wenn er mit dem HTML-Dokument interagiert.

---

<sup>34</sup> jQuery Foundation. (2012). License. Abgerufen am 19. Juni 2012 von jQuery: <http://jquery.org/license/>

<sup>35</sup> jQuery Foundation. (2012c). Selectors. Abgerufen am 19. Juni 2012 von jQuery API: <http://api.jquery.com/category/selectors/>

Außerdem kann das Verhalten nach dem Auslösen des Events definiert werden.<sup>36</sup>

- Manipulation von Inhalten und Attributen von HTML-Elementen

JQuery stellt eine breite Funktionalität für die DOM-Manipulation zur Verfügung. Es können Attribute und der Style von HTML-Elementen geändert werden. Außerdem können sie sich selbst verändern, indem sie sich kopieren, löschen und einfügen.<sup>37</sup>

- AJAX

Die jQuery-Bibliothek verfügt über AJAX-Funktionen, die es ermöglichen, Daten vom Server zu laden, ohne dass der Benutzer den Browser aktualisieren muss.<sup>38</sup>

In der User-Generated-Content Komponente wurde die AJAX-Funktionalität für die Kommunikation zwischen dem Frontend, das den drei Widgets entspricht, und der Anwendungslogik, die serverseitig implementiert wurde, verwendet.

- **jQuery UI 1.8.16**

jQuery UI bietet eine umfassende Sammlung von Plug-ins für die Benutzeroberfläche, Widgets und visuellen Effekte an.<sup>39</sup>

Grundsätzlich ist jQuery UI eine Widget- und Interaktions-Bibliothek, die auf die jQuery JavaScript Bibliothek aufbaut. jQuery UI ermöglicht dem Entwickler, hoch-interaktive Web-Anwendungen zu erstellen.<sup>40</sup>

---

<sup>36</sup> jQuery Foundation. (2012). Events. Abgerufen am 19. Juni 2012 von jQuery API: <http://api.jquery.com/category/events/>

<sup>37</sup> jQuery Foundation. (2012). Manipulation. Abgerufen am 19. Juni 2012 von jQuery API: <http://api.jquery.com/category/manipulation/>

<sup>38</sup> jQuery Foundation. (2012). Ajax. Abgerufen am 19. Juni 2012 von jQuery API: <http://api.jquery.com/category/ajax/>

<sup>39</sup> jQuery Foundation und das jQuery UI TEAM. (2012). Demos & Documentation. Abgerufen am 19. Juni 2012 von jQuery User Interface: <http://jqueryui.com/demos/>

Die JavaScript-Bibliothek jQuery UI ist kompatibel mit dem Internet Explorer 6.0+, Firefox 3+, Safari 3.1+, Opera 9.6+ und Google Chrome.<sup>41</sup>

Für die Umsetzung wurden unter anderem vier jQuery UI Features verwendet. Das erste jQuery UI Feature, das eingesetzt wurde, ist das jQuery UI Dialog Plug-In. Außerdem wurde das jQuery UI Tab Plug-In verwendet, um die Form, die die Modifikation von Metadatensätzen erlaubt, in mehrere Abschnitte aufzuteilen. Durch diese Aufteilung der Form wird Platz gespart und die User-Generated-Content Widgets sind auf den ersten Blick nicht zu überladen. Darüber hinaus wird für die User-Generated-Content Widgets das jQuery UI Autocomplete Plug-In verwendet. Diese Autovervollständigungsfunktion wird für Keyword-Felder verwendet.

- **CVI-Busy-Lib 1.3**

Die CVI-Busy Bibliothek ermöglicht das Hinzufügen und Entfernen von verschiedenen Ladeelementen in Webseiten. Diese Ladeelemente können jedem beliebigen HTML-Element oder dem gesamten Fensterrahmen (viewport) zugewiesen werden. Die Bibliothek ist sehr einfach anzusprechen und kann beliebig konfiguriert werden. Die CVI-Busy-Bibliothek 1.3 funktioniert in allen gängigen Browsern – Mozilla Firefox 1.5+, Opera 9+, Safari und Internet Explorer IE6+. (Effenberger, 2012)

Für die Nutzung der Bibliothek stehen zwei Lizenzen zur Verfügung. Erstens die „Netzgesated NON-commercial Software Licence“, die die Nutzung für nicht-kommerzielle Websites und private Websites erlaubt. Zweitens ist auch eine kommerzielle Software-Lizenz auf Anfrage erhältlich. (Effenberger, 2012)

---

<sup>40</sup> jQuery Foundation und das jQuery UI Team. (2012b). What is jQuery UI? . Abgerufen am 19. Juni 2012 von jQuery User Interface: [http://jqueryui.com/docs/Getting\\_Started](http://jqueryui.com/docs/Getting_Started)

<sup>41</sup> jQuery Foundation und das jQuery UI TEAM. (2012a). Demos & Documentation. Abgerufen am 19. Juni 2012 von jQuery User Interface: <http://jqueryui.com/demos/>

In der User-Generated-Content Komponente wurde die Bibliothek CVI Busy verwendet, um dem Benutzer bei rechenintensiven Aktionen eine Rückmeldung zu geben. Rechenintensive Aktionen in der User-Generated-Content Komponente sind das Extrahieren der Metadaten von neuen Lernressourcen und das Übersetzen eines bestehenden Metadatensatzes. Bei diesen zwei Aktionen wird die Webseite mit einem Ladesymbol überblendet.

### 4.3 Cross-Domain Kommunikation

In diesem Abschnitt werden die Problematik und die Gefahren der Cross-Domain Kommunikation erklärt und in Folge, warum diese Art der Kommunikation von allen Browsern geblockt wird. Anschließend wird auf die speziellen Anforderungen an die User-Generated-Content Komponente seitens der Organic.Lingua Projektpartner eingegangen. Um diese Anforderungen zu realisieren, bedarf es der Cross-Domain Kommunikation. Abschließend wird in diesem Abschnitt erklärt, wie die Cross-Domain Kommunikation trotz Browserblockierung realisiert wurde.

Die Cross-Domain Kommunikation kann wie folgt beschrieben werden. Jedes Dokument (HTML-Dokument, JavaScript-Datei, usw.), das in einem Web-Browser geladen wird, hat eine Domain-Eigenschaft. Die Domain ist eine IP-Adresse oder ein Hostname. Folglich wird unter dem Begriff „Cross-Domain Kommunikation“ die Kommunikation zwischen zwei Dokumenten mit unterschiedlicher Domain verstanden.

Aktuelle Browser-Sicherheitskonzepte wurden nicht darauf ausgelegt, Cross-Domain Kommunikationen zu unterstützen (De Keukelaere, Bholá, Steiner, Chari, & Yoshihama, 2008). Diese Browser-Sicherheitskonzepte schreiben vor, dass Inhalte unterschiedlicher Herkunft nicht miteinander interagieren können, während Inhalte von der gleichen Domain ohne Beschränkungen kommunizieren können (De Keukelaere et al., 2008). Konkret wird dieses Sicherheitskonzept, das JavaScript nur auf Objekte aus derselben Domain zuzugreifen erlaubt, Same-Origin-Policy (SOP) bezeichnet. Aufgrund dieses Sicherheitskonzeptes muss für die Integration der User-Generated-Content

Komponente in das Organic.Edunet Web-Portal, ein Workaround realisiert werden, weil die Interaktion zwischen der User-Generated-Content Komponente und dem Organic.Edunet Web-Portal erwünscht ist.

De Keukelaere et al. (2008) beschreiben zwei Möglichkeiten, die es ermöglichen, die Beschränkungen der Same-Origin-Policy zu umgehen: erstens den Einsatz einer Web-Proxy-Server Anwendung, welche den Inhalt von verschiedenen Domains holt und der Web-Client-Anwendung zur Verfügung steht. Zweitens kann die Beschränkung umgangen werden, indem Dateien unterschiedlicher Domains unter Verwendung des `<script>`-Tags in einer Webseite geladen werden.

Der erste Lösungsansatz von De Keukelaere et al. (2008) kann für die Integration der User-Generated-Content Komponente in das Organic.Edunet Web-Portal nicht verwendet werden. Der Grund dafür liegt darin, dass die User-Generated-Content Komponente einfach und mit wenig Aufwand in das Organic.Edunet Web-Portal integriert werden können soll. Würde der erste Lösungsvorschlag von De Keukelaere et al. (2008) verfolgt werden, hieße das, dass beim Organic.Edunet serverseitig ein Proxy-Server realisiert werden müsste, der die Kommunikation mit der User-Generated-Content Komponente realisiert. Die Umsetzung des Proxy-Servers wäre sehr aufwendig, weil das gesamte Interface der User-Generated-Content Komponente durch den Proxy-Server implementiert werden müsste. Außerdem wäre bei jeder Veränderung des Interfaces der User-Generated-Content Komponente eine Anpassung des Proxy-Servers nötig.

Der zweite Lösungsansatz von De Keukelaere et al. (2008) wurde verwendet, um im Organic.Edunet Web-Portal die User-Generated-Content Stylesheets und JavaScript-Dateien dynamisch nachzuladen. Das dynamische Nachladen von JavaScript-Dateien funktioniert, weil die Same-Origin-Policy das dynamische Nachladen von Dateien nicht verhindert. Die einzige Einschränkung der Browser Same-Origin-Policy besteht darin, dass das nachgeladene Script so vom Browser gehandhabt wird, als wäre das geladene Skript von der Domain der Webseite. Versucht die nachgeladene JavaScript-Datei ein Dokument von einer anderen Domain als der Domain der Webseite zu laden, würde es aufgrund der Same-Origin-Policy wieder nicht funktionieren. Folgender Codeausschnitt 4-1

zeigt beispielhaft die Integration des Metadatenübersetzungs-Widgets mit Hilfe des <script>-Tags.

```
function callLOMTranslateService(Name, Username, Email, LOMID, LOMLocation){  
  
    var WIDGET_HOST = "http://organiclingua.know-center.tugraz.at/";  
  
    var path_js = "/ugc-widget-server/";  
  
    var x = document.createElement("SCRIPT");  
  
    x.type = "text/javascript";  
  
    x.src = WIDGET_HOST + path_js + "loadUGC.js";  
  
    x.setAttribute("NameOfUser", Name);  
  
    x.setAttribute("Username", Username);  
  
    x.setAttribute("Email", Email);  
  
    x.setAttribute("Operation", "translate");  
  
    x.setAttribute("id", "LOMWidget");  
  
    x.setAttribute("LOMID", LOMID);  
  
    x.setAttribute("LOMLocation", LOMLocation);  
  
    document.getElementsByTagName("head")[0].appendChild(x);  
  
}
```

### **Codeausschnitt 4-1: Cross-Domain Kommunikation mit Hilfe des <script>-Tags**

Außerdem wird ein zweiter Lösungsansatz für die Realisierung der Cross-Domain Kommunikation für die Realisierung der User-Generated-Content Komponente eingesetzt, weil der erste Lösungsansatz zwar das dynamische Nachladen von domainfremden JavaScript-Dateien erlaubt, aber diese nachgeladenen JavaScript-Dateien keine domainfremden Anfragen absetzen können. Der dritte Lösungsansatz realisiert den Aufruf der User-Generated-Content Services und wird in den folgenden Absätzen genauer erklärt.



Grundsätzlich wird der Datenaustausch zwischen Browser (JavaScript) und Server mit Hilfe von Ajax realisiert (Garrett, 2005). Ajax steht für „**A**synchronous **J**ava**S**cript **a**nd **X**ML“ und ist im Grunde ein Technologiemix, der mehrere Standards kombiniert. Der Begriff Ajax fasst den Einsatz von den Technologien wie zum Beispiel HTML, CSS, DOM, JavaScript und XMLHttpRequest zusammen. Die Kerntechnologie, auf die der Ajax-Ansatz aufbaut, ist das standardisierte XMLHttpRequest Objekt. Dieses XMLHttpRequest Objekt realisiert die asynchrone Datenübertragung zwischen einem Browser und dem Server. Der große Vorteil von Ajax liegt darin, dass Daten mit Hilfe von http mit einem Server ausgetauscht werden können und die gleichzeitige Aktualisierung von Teilen einer Webseite vonstattengehen kann, ohne dass die Webseite neu geladen werden muss. (Garrett, 2005)

Aufgrund der Sicherheitsbeschränkung der Browser für den Datenaustausch über Domaingrenzen (Same-Origin-Policy) hinweg, funktioniert Ajax nicht. Mittels Ajax können grundsätzlich Daten nur von einer Domain abgerufen werden, die gleich der aktuellen Domain der Web-Seite ist. (Özses & Ergül, 2009)

Dieses Problem von Ajax kann mit Hilfe von JSONP umgangen werden. JSONP steht für „JSON with padding“. JSONP ermöglicht den Abruf von Daten von domainfremden Servern. Der domainübergreifende Datenaustausch basiert auf dem Prinzip des <script>-Tags, der keiner Cross-Domain Beschränkung unterliegt. Für jeden JSONP-Aufruf wird ein <script>-Tag in der aktuellen Webseite hinzugefügt. Der Script-Tag zeigt mit Hilfe des „src“-Attributes auf das Server-Service und teilt dem Server mit einem Parameter die Callback-Funktion mit. Der Server verpackt die Daten in den Funktionsaufruf der übergebenen Callback-Funktion und liefert diese als Rückgabe zurück. Der größte Vorteil von JSONP besteht darin, dass die Technik von jedem Browser unterstützt wird und jQuery den Einsatz von JSONP erleichtert. (MacCaw, 2011, S. 43)

Zusammenfassend wurden zwei Lösungsansätze in der User-Generated-Content Komponente eingesetzt, um die domainübergreifende Kommunikation zwischen Organic.Edunet Web-Portal und der User-Generated-Content Komponente zu bewerkstelligen. Erstens wurde das <script>-Tag eingesetzt, um JavaScript-Dateien und Stylesheets dynamisch nachzuladen. Zweitens wurde JSONP

eingesetzt, um domainfremde Daten abzurufen, im Konkreten wurde damit die Kommunikation zwischen den nachgeladenen User-Generated-Content JavaScript-Dateien und den User-Generated-Content Services realisiert.

## 4.4 Integration der User-Generated-Content Komponente

Der folgende Abschnitt beschreibt die Integration der User-Generated-Content Komponente in das Organic.Edunet Web-Portal. Spezielles Augenmerk wurde bei dem Design und der Implementierung der gesamten User-Generated-Content Komponente auf eine einfache Integration der User-Generated-Content Widgets gelegt. Dies bedeutet, dass nur geringfügige Änderungen im Organic.Edunet erforderlich sind, um die User-Generated-Content Komponente zu integrieren. Die User-Generated-Content Komponente besteht aus drei webbasierten Widgets. Für die Integration jedes einzelnen Widgets müssen zehn Zeilen JavaScript-Code im Organic.Edunet Web-Portal eingefügt werden. Diese Codezeilen fügen dynamisch neue HTML-Header-Elemente zur existierenden Webseite hinzu. Diese neu hinzugefügten HTML-Header-Elemente laden die Widgets und die Funktionalität der User-Generated-Content Komponente. In den folgenden Abschnitten wird die Integration für jedes der drei Widgets im Detail beschrieben.

### 4.4.1 Integration des Metadaten-Bearbeitungs-Widgets

Folgende Codezeilen, die im Codeausschnitt 4-2 angeführt sind, integrieren das Metadaten-Bearbeitungs-Widget, das dem Organic.Edunet Benutzer ermöglicht, bestehende Metadatensätze zu bearbeiten.

Folgende Attribute müssen beim Aufruf des User-Generated-Content Metadaten-Bearbeitungs-Widgets übergeben werden:

- **NameOfUser:** Vorname und Nachname des Organic.Edunet Benutzers.
- **Username:** Benutzername des Organic.Edunet Benutzers.

- **Email:** Die E-Mail-Adresse des Organic.Edunet Benutzers.
- **id:** Der Standard-Wert ist „LOMWidget“, damit wird der User-Generated-Content Komponente mitgeteilt, dass das Widget für die IEEE LOM Eingabe angefordert wird.
- **Operation:** Name der Funktion, die angefordert wird. In diesem Fall ist der Wert des Attributes „edit“, da ein bestehender Metadatensatz im Organic.Edunet Web-Portal bearbeitet werden soll.
- **LOMID:** Eine eindeutige Kennung des Metadatensatzes (zum Beispiel der Wert des IEEE LOM Elements „1.1.2“)
- **LOMLocation:** Der Ort bzw. URL an der die zu bearbeitende Metadatenbeschreibung im IEEE LOM Format abrufbar ist. (zum Beispiel [http://portal.organic-edunet.eu/harvested\\_files/oai\\_scam\\_http\\_confolio\\_vm\\_grnet\\_gr\\_scam\\_12\\_entry\\_181.xml](http://portal.organic-edunet.eu/harvested_files/oai_scam_http_confolio_vm_grnet_gr_scam_12_entry_181.xml)).
- **Language:** Gibt die Sprache an, in welcher der Metadatensatz vom Organic.Edunet Benutzer bearbeitet werden soll.

```
function callLOMEditService(Name, Username, Email, LOMID, LOMLocation, Language){  
  
    var WIDGET_HOST = "http://organiclingua.know-center.tugraz.at/";  
  
    var path_js = "/ugc-widget-server/";  
  
    var x = document.createElement("SCRIPT");  
  
    x.type = "text/javascript";  
  
    x.src = WIDGET_HOST + path_js + "loadUGC.js";  
  
    x.setAttribute("NameOfUser", Name);  
  
    x.setAttribute("Username", Username);  
  
    x.setAttribute("Email", Email);  
  
    x.setAttribute("Operation", "edit");  
  
}
```

```
x.setAttribute("id", "LOMWidget");  
  
x.setAttribute("LOMID", LOMID);  
  
x.setAttribute("LOMLocation", LOMLocation);  
  
x.setAttribute("Language", Language);  
  
document.getElementsByTagName("head")[0].appendChild(x);  
  
}
```

**Codeausschnitt 4-2: Integrationscode für das Metadaten-Bearbeitungs-Widget**

## 4.4.2 Integration des Metadaten-Übersetzungs-Widgets

Folgende Codezeilen, die im Codeausschnitt 4-3 angeführt sind, integrieren das Metadaten-Übersetzungs-Widget, das dem Organic.Edunet Benutzer ermöglicht, bestehende Metadatensätze zu übersetzen.

Folgende Attribute müssen beim Aufruf des User-Generated-Content Metadaten-Übersetzungs-Widgets übergeben werden:

- **NameOfUser:** Vorname und Nachname des Organic.Edunet Benutzers.
- **Username:** Benutzername des Organic.Edunet Benutzers.
- **Email:** Die E-Mail-Adresse des Organic.Edunet Benutzers.
- **id:** Der Standard-Wert ist „LOMWidget“, damit wird der User-Generated-Content Komponente mitgeteilt, dass das Widget für die IEEE LOM Eingabe angefordert wird.
- **Operation:** Name der Funktion, die angefordert wird. In diesem Fall ist der Wert des Attributes „translate“, da ein bestehender Metadatensatz im Organic.Edunet Web-Portal übersetzt werden soll.
- **LOMID:** Eine eindeutige Kennung des Metadatensatzes (zum Beispiel der Wert des IEEE LOM Elements „1.1.2“)

- **LOMLocation:** Der Ort bzw. URL an der die zu übersetzende Metadatenbeschreibung im IEEE LOM Format abrufbar ist. (zum Beispiel [http://portal.organic-edunet.eu/harvested\\_files/oai\\_scam\\_http\\_confolio\\_vm\\_grnet\\_gr\\_scam\\_12\\_entry\\_181.xml](http://portal.organic-edunet.eu/harvested_files/oai_scam_http_confolio_vm_grnet_gr_scam_12_entry_181.xml)).

```
function callLOMTranslateService(Name, Username, Email, LOMID, LOMLocation){  
  
    var WIDGET_HOST = "http://organiclingua.know-center.tugraz.at/";  
  
    var path_js = "/ugc-widget-server/";  
  
    var x = document.createElement("SCRIPT");  
  
    x.type = "text/javascript";  
  
    x.src = WIDGET_HOST + path_js + "loadUGC.js";  
  
    x.setAttribute("NameOfUser", Name);  
  
    x.setAttribute("Username", Username);  
  
    x.setAttribute("Email", Email);  
  
    x.setAttribute("Operation", "translate");  
  
    x.setAttribute("id", "LOMWidget");  
  
    x.setAttribute("LOMID", LOMID);  
  
    x.setAttribute("LOMLocation", LOMLocation);  
  
    document.getElementsByTagName("head")[0].appendChild(x);  
  
}
```

**Codeausschnitt 4-3: Integrationscode für das Metadaten-Übersetzungs-Widget**

### 4.4.3 Integration des "Suggest new Learning Resource" Widgets

Folgende Codezeilen, die im Codeausschnitt 4-4 angeführt sind, integrieren den Bookmark-Button für das User-Generated-Content Widget, das dem Organic.Edunet Benutzer ermöglicht, neue Lernressourcen zum Organic.Edunet Web-Portal hinzuzufügen. Diesen Bookmark-Button kann der Organic.Edunet Benutzer in seine Lesezeichen-Symbolleiste ziehen. Findet der Organic.Edunet Benutzer eine interessante Lernressource, während er im Internet surft, kann er mittels eines Klicks auf den „Suggest new Learning Resource“ Button, der in der Lesezeichen-Symbolleiste aufscheint, die aktuelle Webseite zum User-Generated-Content Repository hinzufügen.

Folgende Attribute müssen beim Aufruf des User-Generated-Content „Suggest new Learning Resource“ Widgets übergeben werden:

- **NameOfUser:** Vorname und Nachname des Organic.Edunet Benutzers.
- **Username:** Benutzername des Organic.Edunet Benutzers.
- **Email:** Die E-Mail-Adresse des Organic.Edunet Benutzers.
- **id:** Der Standard-Wert ist „LOMWidget“, damit wird der User-Generated-Content Komponente mitgeteilt, dass das Widget für die IEEE LOM Eingabe angefordert wird.
- **Operation:** Name der Funktion, die angefordert wird. In diesem Fall ist der Wert des Attributes „add“, da ein neuer Metadatensatz für die neue Lernressource im User-Generated-Content Repository gespeichert werden soll.
- **URL:** Gibt den Ort bzw. die Lokalisierung der Webseite an, die als neue Lernressource in das User-Generated-Content Repository aufgenommen werden soll.

```
<a type="button" id="button_add_loma" name="button_add_lom"
style="width: 120px; cursor:pointer" class="button"
href="javascript:(function() {WIDGET_HOST = 'http://organiclingua.know-
center.tugraz.at/';var path_js = '/ugc-widget-server/';try {var x =
```

```
document.createElement('SCRIPT');x.type = 'text/javascript';x.src =
WIDGET_HOST + path_js + 'loadUGC.js';x.setAttribute('NameOfUser', 'Philip
Kohler');x.setAttribute('Username', 'kohph');x.setAttribute('Email',
'pkohler@know-center.at');x.setAttribute('Operation', 'add');x.setAttribute('id',
'LOMWidget');x.setAttribute('URL',
window.location.href);document.getElementsByTagName('head')[0].appendChild(x)} catch (e) {}})();">Suggest Learning Resource</a>
```

**Codeausschnitt 4-4: Integrationscode für das „Suggest new learning resource“ Widget**

## 4.5 Bereitstellen der User-Generated-Metadata

Die vom End-Benutzer generierten Metadaten werden dem Organic.Edunet Web-Portal mittels des OAI-PMH Standards zur Verfügung gestellt. Dieser Standard basiert auf XML und REST und ermöglicht dem Organic.Edunet Web-Portal die einfache Einsammlung und Weiterverarbeitung (Validierung) der Metadaten. Eine genaue Beschreibung des OAI-PMH Standards wird in Abschnitt 2.6 gegeben.

Folgende URL zeigt eine Beispiel-Anfrage für neue User-Generated-Content Metadatenätze:

<http://organiclingua.know-center.tugraz.at/ugc-widget-server/server/oai-pmh/oai2.php?verb=ListRecords&from=2011%E2%80%9006%E2%80%9002>

Diese Anfrage kann mit einem „from“ Attribut parametrisiert werden. Dies ermöglicht dem Service-Provider (siehe Abschnitt 2.6), neue Metadatenätze vom User-Generated-Content Repository in definierten Zeitintervallen abzurufen.

Folgender Codeausschnitt 4-5 zeigt einen Beispielresponse der User-Generated-Content Komponente.

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
  http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
<responseDate>2011-11-22T15:20:30Z</responseDate>
<request verb="ListRecords"
  from="2011-11-22">
  http://url_K-C:port/harvesting/oai </request>
<ListRecords>
<record>
<header>
  <identifier>oai.k-c.12346</identifier>
  <datestamp>2011-09-25</datestamp>
</header>
<metadata>
  <lom>
  </lom>
  ....
</metadata>
<about>
</about>
</record>
<record>
```



```
<header>

  <identifier>oai.k-c.11350</identifier>

  <timestamp>2011-08-20</timestamp>

</header>

<metadata>

  <lom>

    ....

  </lom>

</metadata>

<about>

</about>

</record>

</ListRecords>
```

**Codeausschnitt 4-5: Beispiel OAI-PMH Response von der User-Generated-Content Komponente**



## 5 Reflexion und Ausblick

In diesem Kapitel wird die gesamte Arbeit zusammengefasst. Weiters wird das Feedback der Organic.Lingua Projektpartner zum ersten Prototypen der User-Generated-Content Komponente angeführt. Abgeschlossen wird das Kapitel mit einem Ausblick für die Zukunft, der eventuelle Weiterentwicklungsmaßnahmen für die User-Generated-Content Komponente aufzeigt.

### 5.1 Zusammenfassung

Im Rahmen dieser Masterarbeit wurden User-Generated-Content Services und Tools entwickelt, sodass der Organic.Edunet Benutzer nicht mehr reiner Konsument der Lerninhalte ist, sondern aktiv bei der Gestaltung der Lerninhalte mitwirken kann. Diese Services und Tools wurden in der User-Generated-Content Komponente (siehe Kapitel 3) zusammengefasst. Durch die User-Generated-Content Komponente wurde der bestehende Lebenszyklus für Lernressourcen bzw. deren Metadaten um den „End-Benutzer“ als Akteur erweitert. Durch die Erweiterung des Lebenszyklus wird dem Organic.Edunet Benutzer die Möglichkeit geboten, bestehende Metadatensätze zu erweitern, zu korrigieren und zu übersetzen und Metadatensätze für neue Lernressourcen, die noch nicht in den Organic.Edunet Repositories erfasst wurden, anzulegen.

Es wurden intuitive und einfache Widgets für das Bearbeiten und Übersetzen von bestehenden Metadatensätzen und für das Annotieren neuer Lernressourcen entwickelt.

Mit Hilfe des „Suggest-new-Learningressource“-Widgets kann der Organic.Edunet Benutzer, während er im Web surft, sehr einfach und mit geringem Aufwand neue Lernressourcen in das User-Generated-Content-Repository hinzufügen, ohne die gefundene Internetseite verlassen oder eine neue Anwendung öffnen zu müssen.

Die User-Generated-Content Komponente unterstützt den Organic.Edunet Benutzer bei der Erfassung von Metadatensätzen für neue Online-Lernressourcen, indem möglichst viele Informationen von der zugrunde

liegenden Lernressource extrahiert werden (Titel, Beschreibung, Keywords, Sprache).

Dieser halb-automatische Metadatengenerierungsansatz hat den Vorteil, dass der Benutzer mit weniger Zeitaufwand Lernressourcen annotieren kann, weil die Metadatenfelder vorausgefüllt sind und der Benutzer die automatisch generierten Metadaten nur mehr korrigieren und vervollständigen muss.

Bei der Übersetzung von bestehenden Metadatenätzen wird der Organic.Edunet Benutzer durch eine maschinelle Übersetzung, welche die Projektpartner CELI und Xerox zur Verfügung gestellt haben, unterstützt. Folglich wird durch die Benutzerinteraktion die maschinelle Übersetzung verbessert und der Benutzer kann bei der Übersetzung Zeit sparen, da er nur die maschinelle Übersetzung gegebenenfalls korrigieren muss.

Zusätzlich wird dem Organic.Edunet Benutzer die Bearbeitung von Metadatenätzen durch Autovervollständigungs- und Rechtschreibkorrektur-Services erleichtert werden. Diese Services erhöhen die Qualität der Metadaten, indem beispielsweise der Organic.Edunet Benutzer bei der Bearbeitung von Metadatenätzen auf Tipp-Fehler aufmerksam gemacht wird.

Weiters wurde darauf geachtet, dass die Oberfläche der Widgets nicht überladen ist und dem Benutzer nur relevante Metadatenelemente des IEEE LOM Standards angezeigt werden.

Die User-Generated-Content Komponente funktioniert als Blackbox und kann in jedes beliebige System zur Annotierung von Metadaten eingebunden werden. Die Metadatenätze werden in einem IEEE LOM konformen Format in einem eigenen Repository abgespeichert und können durch die bereitgestellte OAI-PMH Schnittstelle vom Service-Provider abgerufen werden. Bei der Integration der User-Generated-Content Komponente in das Organic.Edunet Web-Portal müssen nur geringfügige Änderungen am Portal vorgenommen werden.

Im folgenden Abschnitt werden die Auswirkungen der User-Generated-Content Komponente auf die Organic.Lingua Vision beschrieben.

## 5.2 Auswirkung auf die Organic.Lingua Vision

Der folgende Textausschnitt präsentiert die Vision des Organic.Lingua Projektes. Diese Vision wurde im Organic.Lingua Deliverable 2.1 „White Paper on Organic.Lingua Vision“ definiert (Dimitropoulos, Koutoumanos, & Manouselis, 2011).

### ***Organic.Lingua Vision***

*"Boost the usage, exploitation and extension of knowledge on Organic Agriculture and Agroecology by using new/ innovative ways to facilitate searching and to minimize the segmentation of knowledge due to lingual barriers through the introduction of automated multilingual services, supported by a quality-controlled collaborative and interactive approach while developing a sustainable low-cost and re-usable solution to the problem of multilinguality across Europe and the whole globe."*

In der folgenden Tabelle 5-1 wird der Einfluss der User-Generated-Content Komponente auf die wichtigsten Ziele der Organic.Lingua Vision aufgezeigt.

<b>Ziele der Organic.Lingua Vision</b>	<b>Einfluss der User-Generated-Content Komponente</b>
<b>Verbessern der Suche</b>	<p>Durch die Möglichkeit der Korrektur fehlerhafter oder unvollständiger Metadatenätze durch den Organic.Edunet Benutzer, wird die Qualität der Metadaten erhöht. Folglich wird auch die Qualität der Suchergebnisse erhöht, weil die Suche mitunter die einzelnen Metadatenelemente heranzieht.</p> <p>Durch die Möglichkeit der Übersetzung von bestehenden Metadatenätzen werden die mehrsprachigen Metadaten bereichert. Folglich wird die Qualität der mehrsprachigen Suche erhöht und die Organic.Edunet Benutzer können die Lernressourcen in ihrer Muttersprache suchen und finden.</p>
<b>Minimierung der Sprachbarrieren für den Benutzer.</b>	<p>Die User-Generated-Content Komponente ermöglicht dem Organic.Edunet Benutzer die Korrektur der automatischen Übersetzung der Metadatenätze. Weiters kann der Benutzer eine Übersetzung zu einem bestehenden Metadatenatz hinzufügen. Folglich sind die sprachlichen Barrieren reduziert, weil andere Organic.Edunet Benutzer die Metadatenätze in deren Muttersprache lesen und durchsuchen können.</p>
<b>Ausbau der Nutzung und Erweiterung der Wissensbasis über biologische Landwirtschaft und Agrarökologie.</b>	<p>Die User-Generated-Content Komponente ermöglicht dem Organic.Edunet Benutzer das Korrigieren, Vervollständigen, Übersetzen und Hinzufügen von Metadaten. Folglich werden immer neue Lernressourcen hinzugefügt und bestehende Metadaten korrigiert und übersetzt. Dies erhöht die Abdeckung von Wissen über das</p>

	Thema 'Biologische Landwirtschaft und Agrarökologie' in verschiedenen Sprachen.
<p><b>Entwicklung einer nachhaltigen, kostengünstigen und wiederverwendbaren Lösung für mehrsprachige Systeme.</b></p>	<p>Die User-Generated-Content Komponente funktioniert als Blackbox. Daher kann die Komponente in verschiedenen Systemen mit unterschiedlichen Kontexten verwendet werden. Außerdem ist die Komponente leicht in jede beliebige Web-basierte Plattform zu integrieren. Darüber hinaus ist die Komponente plattformunabhängig, weil sie auf den Technologien HTML, CSS, JavaScript und XML basiert.</p>

**Tabelle 5-1: Einfluss der User-Generated-Content Komponente auf die Ziele des Organic.Lingua Projektes**

## 5.3 Feedback der Organic.Lingua Projektpartner

Nach der Fertigstellung des ersten Prototypen wurde das Feedback der Organic.Lingua Projektpartner eingeholt. Unter anderem erfolgten Rückmeldungen von folgenden Projektpartnern:

- Universität Alcalá
- Xerox Research Centre Europe
- CELI

Das Feedback der Projektpartner viel sehr positiv aus. Sie gaben sehr detaillierte Anregungen für die Verbesserung der User-Generated-Content Komponente.

Das Ergebnis des Reviews der Organic.Lingua Projektpartner lässt erkennen, dass Verbesserungen in der Bezeichnung der Metadatenfelder vonnöten wären.

Der Projektpartner der Universität Alcalá macht deutlich, dass der Einsatz von „Tabs“ nicht sinnvoll ist, weil für den Benutzer nicht ersichtlich ist, ob er die weiteren Tabs ausfüllen muss. Aus diesem Grund sollen die Tabs „General“, „Author“ und „Rights“ vom bestehenden Widget entfernt werden und durch einen Prozess mit drei Schritten ersetzt werden. Bei jedem Schritt soll ein „Vor“ und „Zurück“ Button existieren, sodass der Benutzer zwischen den drei Bereichen „General“, „Author“ und „Rights“ wechseln kann.

Die Feedbacks der Projektpartner haben gezeigt, dass noch Verbesserungen im Bereich Benutzerfreundlichkeit nötig sind. Diese Verbesserungen werden in den zweiten Prototypen der User-Generated-Content Komponente eingebaut werden. Der folgende Abschnitt gibt einen Überblick für die Weiterentwicklung der User-Generated-Content Komponente.

### **5.4 Ausblick für die Zukunft**

Der erste Prototyp der User-Generated-Content Komponente wurde im Organic.Lingua Deliverable 4.2.3 – „Multilingual content annotation components“ – beschrieben (Kohler & Lex, 2012). Weiters wurde der erste Prototyp den Projektpartnern für eine Begutachtung vorgelegt. Dieser Prototyp wurde von diesen sehr positiv aufgenommen. Im Bereich der Usability gab es detaillierte Verbesserungsvorschläge, welche bei der Umsetzung des zweiten Prototypen beachtet werden sollen. Ferner stellen folgende Punkte die nächsten Schritte für die Weiterentwicklung der User-Generated-Content Komponente dar:

- Die Integration der User-Generated-Content Komponente in das Organic.Edunet Web-Portal.
- Entwicklung einer mehrsprachigen Benutzeroberfläche. In der Zukunft soll die Benutzeroberfläche in der Muttersprache des Benutzers sein.
- Darüber hinaus ist geplant, ein Webinterface für die User-Generated-Content Komponente zu implementieren. Dieses Webinterface ermöglicht dem Benutzer, neue Sprachen für die Benutzeroberfläche der User-Generated-Widgets hinzuzufügen und existierende automatische



Übersetzungen zu korrigieren. Der Zugriff auf dieses Webinterface ist nur berechtigten Personen erlaubt.

- Entwicklung weiterer Webservices für die Abfrage von verschiedenen Statistiken über den Benutzer und Lernressourcen. Typische Abfragen sind zum Beispiel:
  - Wie oft wurde ein bestimmter Metadatensatz bearbeitet?
  - Welche Lernressourcen hat ein bestimmter Organic.Edunet Benutzer hinzugefügt?
  - Welche Metadatensätze hat ein bestimmter Organic.Edunet Benutzer übersetzt oder bearbeitet?
- Entwicklung und Integration eines Ontologie-Lookup-Services für die User-Generated-Content Widgets. Das Ontologie-Lookup-Service soll dem Organic.Edunet Benutzer ermöglichen, bei der Erstellung eines Metadatensatzes für eine neue Lernressource die Lernressource mit einem oder mehreren Konzepten der Organic.Edunet Ontologie<sup>42</sup> zu verknüpfen. Die Organic.Edunet Ontologie wurde von Experten erstellt und bildet die Domäne „biologische Landwirtschaft und Agrarökologie“ ab. Durch die Zuordnung von Lernressourcen zu einem oder mehreren Konzepten der Organic.Edunet Ontologie wird die Lernressource durch semantische Suche im Organic.Edunet Web-Portal gefunden.
- Der Organic.Edunet Benutzer soll bei der Erstellung von neuen Metadatensätzen und bei der Bearbeitung/ Übersetzung von existierenden Metadatensätzen textuelle Hilfestellungen für das Ausfüllen der Metadatenelemente erhalten.

---

<sup>42</sup> Organic.Edunet Ontologie. Abgerufen am 17. August 2012 von Organic.Edunet: [http://portal.organic-edunet.eu/index.php?option=com\\_navigational&view=navigational&Itemid=89](http://portal.organic-edunet.eu/index.php?option=com_navigational&view=navigational&Itemid=89)



## Literaturverzeichnis

- Baker, T. (2000). A grammar of Dublin Core. D-lib magazine.
- Bargmeyer, B. E., & Gillman, D. W. (2000). Metadata standards and metadata registries: an overview.
- Barton, J., Currier, S., & Hey, J. M. N. (2003). Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice. Proceedings of the 2003 international conference on Dublin Core and metadata applications: supporting communities of discourse and practice---metadata research & applications (pp. 66-73).
- Baxendale, P. B. (1958). Machine-made index for technical literature—an experiment. IBM Journal of Research and Development.
- Beall, J. (2005). Metadata and data quality problems in the digital library. Journal of Digital Information, 6.
- Belew, R. K. (2000). Finding Out About: A Cognitive Perspective on Search Engine Technology and the WWW. New York: Cambridge University Press.
- Brasher, A., & McAndrew, P. (2004). Human-Generated Learning Object Metadata Relevance to Learning Object Metadata. On the Move to Meaningful Internet Systems 2004 (pp. 723-730). Springer.
- Bruce, T. R., & Hillmann, D. I. (2004). The continuum of metadata quality: defining, expressing, exploiting. Metadata in Practice. Retrieved from <http://www.ecommons.cornell.edu/handle/1813/7895>
- Bruns, A. (2008). Blogs, Wikipedia, Second Life, and Beyond: From Production to Producership. New York: Peter Lang.
- Bruns, A. (31. Dezember 2007). Producership: A Working Definition. Abgerufen am 13. Juli 2012 von Producership.org: <http://producership.org/node/9>
- Cardinaels, K., Meire, M., & Duval, E. (2005). Automating metadata generation: the simple indexing interface. WWW '05 Proceedings of the 14th international conference on World Wide, 548-556.

- Chien, L.-F. (1997). PAT-tree-based keyword extraction for Chinese information retrieval. *ACM SIGIR Forum* (Vol. 31, pp. 50–58). ACM.
- Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., & Stukas, A. A. (1998). Understanding and assessing the motivations of volunteers: A functional approach. *Journal of personality*.
- Currier, S., Barton, J., O’Beirne, R., & Ryan, B. (2004). Quality assurance for digital learning object repositories: issues for the metadata creation process. *AltJ*, 12(1), 5-20. Taylor Francis Ltd. Retrieved from <http://dx.doi.org/10.1080/0968776042000211494>
- Dahl, D., & Vossen, G. (2008). Learning Object Metadata Generation in the Web 2. 0 Era. *International Journal of Information and Communication Technology Education*, 4(3), 1–10. IGI Publishing, 701 E. Chocolate Ave, Suite 200, Hershey, PA, 17033-1240, USA,. Retrieved from [http://www.iadis.net/dl/final\\_uploads/200711L014.pdf](http://www.iadis.net/dl/final_uploads/200711L014.pdf)
- Das, D., & Martins, A. (2007). A Survey on Automatic Text Summarization Single-Document Summarization. *Language*, 4, 1–31. Citeseer
- De Keukelaere, F., Bholra, S., Steiner, M., Chari, S., & Yoshihama, S. (2008). Smash: secure component model for cross-domain mashups on unmodified browsers. *Proceeding of the 17th*, 535-544.
- Dimitropoulos, A., Koutoumanos, A., & Manouselis, N. (2011). *Organic.Lingua Deliverable 2.1 - White Paper on Organic.Lingua Vision*. Organic.Lingua.
- Dumais, S., Platt, J., Heckerman, D., & Sahami, M. (1998). Inductive learning algorithms and representations for text categorization. *Proceedings of the seventh international conference on Information and knowledge management - CIKM '98*, 148-155. New York, New York, USA: ACM Press. doi:10.1145/288627.288651
- Duval, E., Hodgins, W., Sutton, S., & Weibel, S. L. (2002). Metadata principles and practicalities. *D-lib Magazine*.
- Edmundson, H. P. (1969). New Methods in Automatic Extracting. *Journal of the ACM*, 16(2), 264-285. doi:10.1145/321510.321519

- Effenberger, C. (2012). CVI-Busy-Lib. Abgerufen am 19. Juni 2012 von busy.netzgesta.de: <http://www.netzgesta.de/busy/>
- Garrett, J. J. (18. Februar 2005). Ajax: A New Approach to Web Applications. Abgerufen am 16. August 2012 von adaptivepath.com: <http://www.adaptivepath.com/ideas/ajax-new-approach-web-applications>
- Guy, M., & Powell, A. (2004). Improving the quality of metadata in Eprint archives. *Ariadne*, (38). Retrieved from <http://www.ariadne.ac.uk/issue38/guy>
- Hars, A., & Ou, S. (2001). Working for free? Motivations of participating in open source projects. *System Sciences, 2001. Proceedings of the 34th Annual, 00(c)*, 1-9.
- Hughes, B. (2004). Metadata quality evaluation: Experience from the open language archives community. *Digital Libraries: International Collaboration and Cross-Fertilization*, 320-329.
- IEEE Learning Technology Standards Committee. (2002). Draft Standard for Learning Object Metadata. New York: Learning Technology Standards Committee.
- Kansas State University. (13. August 2008). Digital Ethnography @ KSU. Abgerufen am 7. Juli 2012 von ksudigg: <http://ksudigg.wetpaint.com/page/YouTube+Statistics>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1), 59-68. Elsevier.
- Kohler, P., & Lex, E. (2012). Organic.Lingua Deliverable 4.2.3 - Multilingual content annotation components. *Organic.Lingua*.
- Kolbitsch, J., & Maurer, H. (2006). The Transformation of the Web: How Emerging Communities Shape the Information we Consume. *Journal of Universal Computer Science*, 12(2), 187-213.

- Kolcz, A., Prabaharmurthi, V., & Kalita, J. (2001). Summarization as feature selection for text categorization. Proceedings of the tenth international conference on Information and knowledge management, 365-370.
- Lagoze, C., Sompel, H., Nelson, M., & Warner, S. (14. Juni 2002). The Open Archives Initiative Protocol for Metadata Harvesting. Abgerufen am 22. Juli 2012 von OpenArchives.org: <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>
- Lassila, O. (1998). Web metadata: A matter of semantics. Internet Computing, IEEE, 2(4), 30-37.
- Lin, J. (2009). Summarization. Encyclopedia of Database Systems. Springer.
- Lucke, U. (2011). Netzbaasierte Systeme in Lehre und Forschung: Innovative IT-Infrastrukturen für die Hochschule der Zukunft. Berlin: Logos Verlag Berlin GmbH.
- Luhn, H. P. (1958). The automatic creation of literature abstracts. IBM Journal of research and development.
- MacCaw, A. (2011). JavaScript Web Applications. Sebastopol(CA, Vereinigte Staaten): O'Reilly Media, Inc.
- Mani, I., & Maybury, M. T. (1999). Advances in Automatic Text Summarization. USA: MIT Press.
- Matsuo, Y., & Ishizuka, M. (2004). Keyword extraction from a single document using word co-occurrence statistical information. International Journal on Artificial Intelligence Tools, 13(1), 157-170. WORLD SCIENTIFIC PUBLISHING.
- Matthias, H. (2010). Frewilligenarbeit und Religiosität: Der Zusammenhang von religiösen Einstellungen und ehrenamtlichem Engagement. Berlin: Lit Verlag.
- Meyer, M., Rensing, C., & Steinmetz, R. (2011). Multigranularity reuse of learning resources. ACM Transactions on Multimedia Computing, Communications, and Applications, 7(1), 1-23. doi:10.1145/1870121.1870122

- Nikoulina, V., Dymetman, M., Sandor, A., Coursimault, J.-M., Sicilia, M. Á., Cano, R., et al. (2011). Organic.Lingua Deliverable 4.1 - Specification of multilingual technologies and components. Organic.Lingua.
- Nov, O. (2007). What motivates wikipedians? *Communications of the ACM*, 50(11).
- Ochoa, X., & Duval, E. (2006). Quality metrics for learning object metadata. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2006* (pp. 1004–1011).
- OECD. (12. April 2007). *Participative Web and User-Created Content: Web 2.0, Wikis and Social Networking*. Paris: OECD Publishing.
- O'REILLY, T. (2007). *What is Web 2.0: Design patterns and business models for the next generation of softwa*. Sebastopol (CA) USA: O'Reilly Media.
- Organic.Lingua. (2010). Organic.Lingua "Description of Work" - Demonstrating the potential of a multilingual Web portal for Sustainable Agricultural & Environmental Education.
- Özses, S., & Ergül, S. (24. Februar 2009). Cross-domain communications with JSONP, Part 1: Combine JSONP and jQuery to quickly build powerful mashups. Abgerufen am 3. Juli 2012 von ibm.com: <http://www.ibm.com/developerworks/library/wa-aj-jsonp1/>
- Pansanato, L. T. E., & Fortes, R. P. M. (2005). Strategies for filling out LOM metadata fields in a Web-based CSCL tool. *Web Congress, 2005. LA-WEB 2005. Third Latin American* (p. 4–pp). IEEE. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1592376](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1592376)
- Prensky, M. (2001). Digital natives, digital immigrants Part 1. *On the horizon*, 9(5), 1-6. doi:10.1108/10748120110424816
- Rada Mihalcea and Paul Tarau. TextRank: Bringing order into texts. In *Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 2004.

- Raskutti, B., Ferrá, H., & Kowalczyk, A. (2001). Second order features for maximising text classification performance. *Machine Learning: ECML 2001*, 419-430.
- Rennie, E. (2007). Community media in the prosumer era. *3C Media Journal of Community, Citizen's and Third Sector Media and Communication*, (3).
- Rushkoff, D. (1999). *Playing the Future: What We Can Learn from Digital Kids*. New York: Riverhead Trade.
- Sanchez, S., Stoitsis, G., Lezcano, L., Manouselis, N., & Dimitropoulos, A. (2011). Organic.Lingua Deliverabe 3.1.2.1 - Revised Organic.Edunet ontology and IEEE LOM application profile. *Organic.Lingua*.
- Schmidt, J. (2009). *Das neue Netz: Merkmale, Praktiken und Folgen des Web 2.0. Design* (1st ed.). Konstanz: UVK Verlagsgesellschaft mbH.
- Schürig, H. (31. März 2010). Social Media statt Web 2.0. Abgerufen am 15. August 2012 von Hennig Schürig Blog: <http://www.henningschuerig.de/blog/2010/social-media-statt-web-20/>
- Sigala, M. (2007). Integrating web 2.0 in e-learning environments: a socio-technical approach. *Int. J. Knowledge and Learning*, 3(6), 628-648.
- Skår, L. A., Heiberg, T., & Kongsli, V. (2003). Reuse learning objects through LOM and XML. Companion of the 18th annual ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications (pp. 78-79). ACM.
- Sompel, H. V. de, Nelson, M., & Lagoze, C. (2004). Resource harvesting within the OAI-PMH framework. *D-lib magazine*, 10.
- Stewart, K. J., & Gosain, S. (2006). The impact of ideology on effectiveness in open source software development teams. *Mis Quarterly*.
- Wotquenne, M., Le Hénaff, D., & Protonotarios, V. (2012). Organic.Lingua Deliverabe 6.1 - Multilingual Content Population Methodology. *Organic.Lingua*.



Yang, D., Li, H., & LIN, Z. (2000). The Applications of LOM and XML in Web-Based Distance Learning. Management.

Yang, H.-L., & Lai, C.-Y. (2010). Motivations of Wikipedia content contributors. (L. Cheng-Yu, Ed.)Computers in Human Behavior, 26(6), 1377-1383. Elsevier Ltd. doi:10.1016/j.chb.2010.04.011