

Florian Thiery B.Sc. (900841)

**Semantic Web und Linked Data:
Generierung von Interoperabilität
in archäologischen Fachdaten
am Beispiel römischer Töpferstempel**

Masterarbeit

zur Erlangung des akademischen Grades Master of Science im
Studiengang Geoinformatik und Vermessung

Fachhochschule Mainz

Fachbereich Technik

Lehrinheit Geoinformatik und Vermessung

Betreuer: Prof. Dr. phil. Kai-Christian Bruhn

Bearbeitungszeitraum: 13. Juni 2013 – 11. Dezember 2013

Standnummer: KM0042

Mainz im Dezember 2013



Der Text dieser Arbeit liegt als  Lizenz vor. Lizenzen zitierter Quellen können von dieser Lizenz differieren. Hierbei sind die jeweiligen Rechte zu beachten. Die Rechte der genutzten Daten liegen beim RGZM, bzw. deren Erzeugern. Innerhalb dieser Arbeit wurde lediglich die Nutzung dieser Daten zur exemplarischen Darstellung eingeräumt. Die entwickelte Software (eigener JAVA Code) steht unter GPL-Lizenz.



FH MAINZ
UNIVERSITY OF
APPLIED SCIENCES



Institut für raumbezogene
Informations- und Messtechnik
Fachhochschule Mainz

R | G | Z | M

**Semantic Web und Linked Data:
Generierung von Interoperabilität
in archäologischen Fachdaten
am Beispiel römischer
Töpferstempel**

Florian Thiery B.Sc.

Kurzfassung

Gegenstand dieser hier vorgestellten Masterarbeit ist Verwendung aktueller Technologien interoperabler Datenhaltung, insbesondere das Konzept der Linked Open Data (LOD) und der semantischen Modellierung, zur Verdeutlichung ihres Potentials in archäologischen Informationen am Beispiel von Terra Sigillata-Fundorten, -Töpfern und -Keramikfragmenten. Die Arbeit zeigt eine Migration von Daten sowie die Möglichkeiten und die Problematik der Modellierung der Attribute und Beziehungen mit Hilfe bestehender LOD-Konzepte und kontrollierter Vokabularien, sowie eigene Ansätze zur Lösung. Diese Daten werden mittels ReST-Schnittstelle zur Verfügung gestellt. Ein Schwerpunkt wird auf die Verlinkung zu anderen bereits bestehenden Projekten gelegt, wodurch eine Vielzahl weiterer archäologischer und historischer Informationen z.B. über das Pelagios Projekt eingebunden werden. Zudem wird das Potential der Verlinkung und Abfrage von heterogenen Informationen zwischen Töpfern, Fragmenten und Orten deren relativ chronologische Beziehungen über LOD mit einer webbasierten Schnittstelle aufgezeigt.

Schlagwörter: Semantic Web, Linked Data, ReST, Chronologie, Terra Sigillata

Abstract

The subject matter of this master thesis is using current technologies in interoperable data management, in particular the illustration of the potential of Linked Open Data (LOD) and semantic modelling in archaeological information, as used on samian ware places and their corresponding potters and ceramic fragments. The thesis demonstrates a migration of data as well as possibilities and problems of modelling attributes and relationships using existing LOD concepts and controlled vocabularies as well as novel self-developed approaches to the solution. These data are provided by a ReST interface. One focus is linking to other existing projects, creating associations to other archaeological and historical information, for example the Pelagios project. Moreover, a web-based interface shows the potential of linking and retrieval of heterogeneous information among pottery, fragments and places and their relative chronological relationships via LOD.

keywords: semantic web, linked data, REST, chronology, samian ware

Inhaltsverzeichnis

Kurzfassung.....	4
Abstract.....	4
Erklärung	5
Inhaltsverzeichnis	6
Abbildungsverzeichnis	9
Tabellenverzeichnis	12
Abkürzungsverzeichnis	13
Vorwort.....	14
1 Motivation und Aufgabenstellung	15
2 Technologie und Standards	19
2.1 Standardisierung im Bereich der Geoinformation	19
2.2 Datenbanken.....	22
2.3 Geosever	25
2.4 ReSTful Webservices.....	25
2.5 Semantic Web und Linked Data	29
2.5.1 Semantic Web.....	29
2.5.2 Linked (Open) Data	33
2.5.3 Ressource Description Framework (RDF)	39
2.5.4 SPARQL.....	47
2.5.5 Triple Stores	51
2.5.6 Web Ontology Language (OWL).....	53
3 Interoperabilität und Semantik in archäologischen Fachdaten	56
3.1 XML- und JSON-Auszeichnungen.....	56
3.2 Linked Data	59
3.2.1 Technologien	59
3.2.2 Vokabularien.....	66
3.2.3 Linked-Data-Konzepte mit Schwerpunkt Archäologie	69
3.2.4 Projekte mit Schwerpunkt Archäologie	71
4 Client-Server Architektur.....	74

4.1	Server-Technologie	74
4.2	Realisierung	76
4.3	Architektur	77
4.4	Übersicht der Applikationen	80
4.5	Beurteilung	82
5	Datenaufbereitung und Datenmigration	83
5.1	Ausgangsdaten und Datenaufbereitung	84
5.2	Datenmodell in PostGIS	87
5.3	Datenüberführung	90
5.4	Beurteilung	92
6	Datenbereitstellung	93
6.1	Die ReST-Schnittstelle	94
6.1.1	Soll-Eigenschaften der Ressourcen	96
6.1.2	HTML Repräsentationen der Ressourcen	97
6.1.3	XML Repräsentationen der Ressourcen	98
6.1.4	JSON Repräsentationen der Ressourcen	103
6.1.5	RDF Repräsentationen der Ressourcen	104
6.1.6	Pelagios Entry-Point	109
6.1.7	Bereitstellungen für den Explorer	111
6.1.8	Triple-Dateien zum Befüllen des Triplestore	112
6.1.9	Programmiertechnische Umsetzung	113
6.2	OGC-Webdienste	119
6.3	Triple Store	124
6.4	Beurteilung	127
7	Datenmapping und Datenverknüpfung	128
7.1	Pleiades und Pleiades Places	128
7.2	Pelagios und die Pelagios API	130
7.3	Pleiades Mapping	133
7.4	PelagiosConnectionAPI	139
7.5	Beurteilung	143
8	Die App - Prototypische Webanwendung	144
8.1	Die Map - Kartenansicht der Findspots	145
8.2	Der Linked Samian Ware Explorer (LSWE)	147
9	Relative Chronologie	152
9.1	Temporale Logik nach Allen und Freksa	154
9.2	Ontologie relativ chronologischer Bezüge (Prototyp)	160
9.3	Visualisierungsstrategien und Evaluation	162

9.4	Der Time Explorer	169
9.5	Zeitmodellierung in der Samian Ware.....	177
10	Fazit und Ausblick	182
	Anhang A: RDF Beispiel Cube	184
	Anhang B: Installationshinweise Server	187
	Anhang C: JAVA-Quellcode	188
	Anhang D: Kilnsites – Karten und Ausgangsdaten	189
	Anhang E: SQL-Skripte	190
	Anhang F: Eigene Ontologie.....	191
	Anhang G: SLD der Kilnsites	196
	Anhang H: Liste der Mapping Ergebnisse.....	197
	Anhang J: Quellcode der Apps	201
	Anhang K: Allen und Freksa Schlussfolgerungen	202
	Anhang L: Visualisierungsideen	205
	Anhang M: Sonstiges.....	207
	Quellenverzeichnis.....	208
	Literatur	208
	Websites und Online-Ressourcen	210
	Abbildungen.....	214

Vorwort

*“It's difficult to imagine the power that you're going to have
when so many different sorts of data are available.”*

Sir Timothy John Berners-Lee¹

Erfinder der Hypertext Markup Language und Begründer des World Wide Web

Die erste Seite dieser Masterarbeit möchte ich nutzen, um all denen zu danken, die mich bei der Anfertigung dieser Arbeit in den letzten Wochen und Monaten unterstützt haben.

Mein größter Dank gilt Herrn Prof. Dr. Bruhn, der es mir ermöglicht hat, meine Masterarbeit am Institut für Raumbezogene Informations- und Messtechnik der Fachhochschule Mainz (i3mainz) in Zusammenarbeit mit dem Römisch-Germanischen Zentralmuseum Mainz (RGZM) zu schreiben. Für die sehr gute Betreuung und tatkräftige Unterstützung zu jeder Tag- und Nachtzeit sowie das meinem Thema entgegengebrachte Interesse bedanke ich mich vielmals. Ein weiterer großer Dank ergeht ebenfalls den Mitarbeitern des RGZM Guido Heinz und Allard Mees, die mich sowohl in technischen, wie vor allem in archäologischen Fragestellungen unterstützt haben. Ein großer Dank ergeht auch an den ‚technischen Support‘: Nikolai Bock, Thomas Engel und Georg Paternoster, wie auch an sonstige Mitarbeiter diverser Institute wie Torsten Schrade, Martin Unold, Michael Haft und Tobias Kohr.

Meiner Mutter Susanne, meinem Vater Peter wie auch allen anderen Familienangehörigen kann ich für die lebenslange Unterstützung, besonders während des Studiums und die mir entgegengebrachte Liebe und Geduld nicht genug danken. Bedanken möchte ich mich zudem bei all meinen Freunden, die ihre Ansichten über die Ergebnisse meiner Arbeit mit mir geteilt und diskutiert haben.

Florian Thiery, Mainz den 11. Dezember 2013

¹ vgl. <http://www.w3.org/People/Berners-Lee>