

# VISIT – TECHNISCHE DOKUMENTATION

Tobias Baumgärtner<sup>1</sup> & Emanuel Berndl<sup>2</sup> &  
Robert Kathrein<sup>3</sup> & Kris Raich<sup>3</sup> &  
Florian Schlenker<sup>4</sup>

30. April 2019



## INHALTSVERZEICHNIS

1	Einleitung	3
2	Kufstein - Backend und Fernrohr	4
3	Tobi - App Framework	5
4	Flo - Digitalisierung und Kompression	6
5	ViSIT Metadaten und die Semantische Datenbank	7
5.1	Theoretische Grundlagen für die Semantische Datenbank . .	7
5.2	Technische Details zur Semantischen Datenbank . . . . .	10
5.3	WissKI - Wissenschaftliche KommunikationsInfrastruktur . .	10
5.4	Technischer Zugang zu den Metadaten - die ViSIT REST API	10
5.5	Wichtige Technische Charakteristika der Entwicklung und den Betrieb der Semantischen Datenbank . . . . .	10
5.6	Zusatzfeatures - Erweiterung der Bedienbarkeit der Semantischen Datenbank . . . . .	10
6	Schluss	11

## ABBILDUNGSVERZEICHNIS

Abbildung 1	Informationen aus obigen Aussagen, kombiniert als Graph. . . . .	8
Abbildung 2	Grundlegende eigene Wissensbasis (oben), erweitert um zwei externe Wissensbasen (unten). . . . .	9

## TABELLENVERZEICHNIS

---

<sup>1</sup> TODO, Universität Passau

<sup>2</sup> Lehrstuhl für verteilte Informationssysteme und Data Science, Universität Passau

<sup>3</sup> TODO, FH Kufstein

<sup>4</sup> TODO, Universität Passau

## 1 EINLEITUNG

blub hi!

## 2 KUFSTEIN – BACKEND UND FERNROHR

TODO rename label and heading

[Bar+16]

### 3 TOBI – APP FRAMEWORK

TODO rename label and heading  
[Bur+16]

## 4 FLO – DIGITALISIERUNG UND KOMPRESSION

TODO rename label and heading

[Bir+16]

## 5 VISIT METADATEN UND DIE SEMANTISCHE DATENBANK

Brainstorm, things to write about:

- theoretischer background: rdf daten, CIDOC, Vismo
- datenbank: infrastruktur (hosting, allgemeiner zugriff von aussen), drupal, wisski (allgemein), grundfunktionalität
- wisski: rdf daten, pfade, konfiguration
- REST API: allgemeine beschreibung
- zusatzfeatures: copy and paste, excel import

### 5.1 Theoretische Grundlagen für die Semantische Datenbank

Dieses Unterkapitel gibt Einblicke in Teilbereiche des Semantic Webs, um eine theoretische Grundlage für die folgenden technischen Entwicklungen zu geben. Nachdem diese erläutert wurden, wird ebenfalls auf eine spezielle Ausprägung eines Metadatenmodells eingegangen, welches die Struktur für die im ViSIT Projekt verwendeten Metadaten vorgibt: das ViSIT Model **VisMo**.

Die hier angeführten Ausführungen beschränken sich jedoch nur auf jeweilige Grundlagen der Themenkomplexe, welche an manchen Stellen um weiterführende Informationen erweitert werden, wenn dies für den weiteren Verlauf von Nöten ist. Dennoch, falls angestrebt, verweisen wir für ein tieferes Verständnis auf weitere Fachliteratur, wie z.B. [Hit+07].

**SEMANTIC WEB UND RDF DATEN** Das Semantic Web ist eine Art Erweiterung zum eigentlichen World Wide Web, wie wir es aktuell kennen. Dieses ist primär für Menschen ausgelegt, die durch Homepages browsen und dabei entsprechende Informationen durch betrachten und lesen der Homepages erlangen. Diese Informationen sind dadurch jedoch nur für Menschen vorhanden, Maschinen oder Computer können auf die Informationen nicht zugreifen, um mit den entsprechenden Daten arbeiten zu können. Genau hier setzt das Semantic Web an, welches Standardisierungen, Regeln und Prozesse vorgibt, um Homepages und Applikationen so anzupassen, dass eben genau eine (semi-) automatische Informationsverarbeitung für Maschinen möglich wird.

Eine dieser Standardisierungen ist das Resource Description Framework **RDF** [MMo4], welches der de-facto Standard im Semantic Web ist, um Metadaten zu beschreiben. Daten in RDF werden als Graph modelliert und persistiert, welcher aus Knoten und Kanten besteht. Dabei entsteht eine Wissensbasis gefüllt an Informationen. Die Knoten sind hierbei die "Akteure", also diejenigen Entitäten, Sachen, Objekte, Dinge etc., ausgehend vom jeweiligen Anwendungsfall, auf die sich die im Graphen enthaltenen Informationen beziehen (diese Dinge werden im Folgenden weiterhin als "Metadatenentität" bezeichnet). Die Kanten im Graphen beschreiben Beziehungen zwischen den gegebenen Knoten und Eigenschaften der Knoten. Weiterhin sind die Knoten und Kanten durch das Grundprinzip eines **Statements** verbunden, welches eine Kapselung einer elementaren Aussage darstellt. Das Statement ist, ähnlich dem deutschen Satzbau, immer bestehend aus drei Teilen:

**SUBJEKT** Die Metadatenentität repräsentiert als ein Knoten im Graphen, von der die Aussage - und damit das Prädikat - des Statements ausgeht.

**PRÄDIKAT** Die Semantik oder die Bedeutung der Aussage.

**OBJEKT** Zweierlei Konzepte können das Objekt des Statements bilden: ein weiterer Knoten im Graphen, um das Ziel der Aussage und damit des Prädikats, um eine Relation zwischen zwei Metadatenentitäten/Knoten darzustellen, oder ein fester Wert, um eine Eigenschaft einer Metadatenentitäten/eines Knotens zu charakterisieren.

Zur Verständlichkeit für die Thematik der Aussagen und Statements im Semantic Web Kontext, soll hier ein kurzes, erfundenes Beispiel erläutert werden. Folgende Aussagen bilden die Wissensbasis:

- Peter ist vom Beruf Baumeister.
- Peter ist 40 Jahre alt.
- Peter war am Bau des Steinschlusses beteiligt.
- Das Steinschloss besteht aus Stein.
- Das Steinschloss ist 10 Jahre alt.

Wie oben beschrieben, bestehen die Aussagen jeweils aus Subjekt, Prädikat und Objekt. Als Subjekte agieren die beiden Metadatenentitäten "Peter" und das "Steinschloss", während die Objekte der Aussagen der Beruf "Baumeister", das Material "Stein", zwei "Altersangaben", sowie das "Steinschloss" selbst sind. Semantisch sind die Subjekte und Objekte über die Beziehungen bzw. Eigenschaften einer "Berufszuordnung", zwei "Alterszuordnungen", einer "Materialzuweisung" sowie der "Erbauung" eines Objekts verbunden.

Diese Aussagen können nun in einen Graphen zusammengefasst werden, dessen high-level Illustration in [Abbildung 1](#) zu sehen ist.

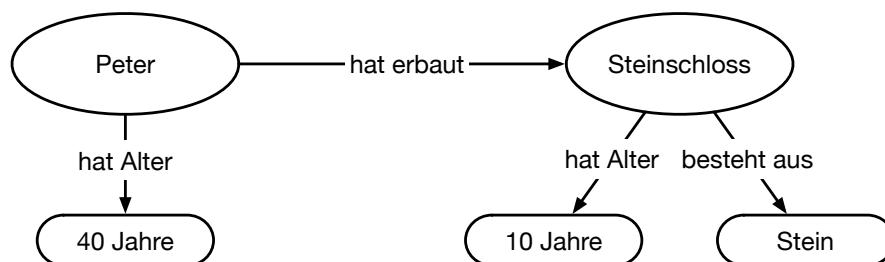


Abbildung 1: Informationen aus obigen Aussagen, kombiniert als Graph.

**LINKED OPEN DATA GEDANKE** Ein weiterer Eckpfeiler des Semantic Web ist ein weiteres Konzept, das unter dem Namen **Linked Open Data - LOD** bekannt ist. Oft wird dieser Name ebenfalls für das Semantic Web selbst benutzt, die punktgenauen Definitionen überschneiden und ergänzen sich.

Einfach übersetzt zielt LOD auf öffentlich zugängliche Daten ab, die untereinander vernetzt und verlinkt sind. Somit soll es möglich sein, verteilte Datenbanken mit ihren eigenen entsprechenden Wissensbasen, miteinander zu verbinden, um so jedem Beteiligten mehr Informationen zur Verfügung

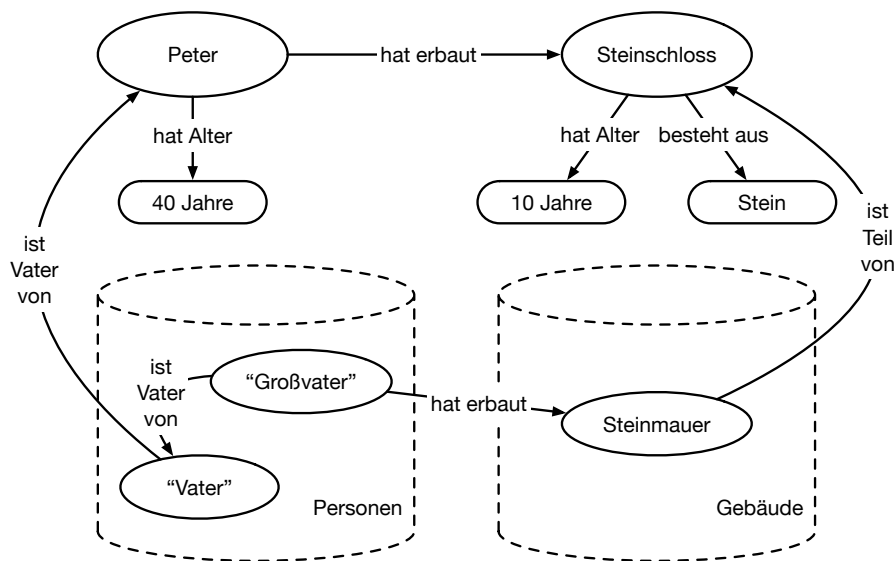


zu stellen, da durch die Verlinkung einzelner Graphen ein großer Gesamtgraph entsteht. Auf diese Weise macht es Sinn, dass jede Wissensbasis ihren eigenen spezialisierten Kontext besitzt. Sollte eine Wissensbasis weitere Informationen aus einem anderen Kontext benötigen, müssen diese Daten nicht auf eigene Hand erforscht und aufbereitet werden, da eine LOD Verbindung zu einer anderen Wissensbasis hergestellt werden kann. Zur weiteren Veranschaulichung dieser Thematik und dessen Vorteile, zeigt der folgende Paragraph zwei Anwendungsfälle im geschichtswissenschaftlichen Kontext.

#### ZWEI ANWENDUNGSFÄLLE FÜR RDF IM GESCHICHTSWISSENSCHAFTLICHEN KONTEXT Verbund verschiedener Kontexte

##### Erschliessen bisher unbekannten Verbindungen

Ein zweiter großer Vorteil davon, geschichtswissenschaftliche Daten in Form von Metadaten und RDF zu persistieren, ist das mögliche Erschließen von vorher nicht bekannten oder erforschten Zusammenhänge der persistierten Objekte. Dazu folgendes (frei erfundenes) Beispiel: Ausgehend von der eigenen Wissensbasis, die die Daten aus [Abbildung 1](#) enthält, sollen nun zwei weitere Wissensbasen angekoppelt werden, welche auf der einen Seite weitere Informationen über Personen und vor allem deren familiärer Beziehungen beinhaltet, und auf der anderen Seite eine Wissensbasis, die mehr Informationen über Gebäude und deren Geschichte beinhaltet. Dies ist in [Abbildung 2](#) visualisiert.



**Abbildung 2:** Grundlegende eigene Wissensbasis (oben), erweitert um zwei externe Wissensbasen (unten).

In dem Beispiel beinhaltet die eigene Wissensbasis Informationen über "Peter" und das "Steinschloss". Durch die beiden hinzugenommenen Wissensbasen wird Peter aus dem Anwendungsbeispiel mit seinem "Vater", und dieser wiederum mit seinem "Großvater" verbunden (die Namen sind hier zur Einfachheit ersetzt). Zudem wird die "Steinmauer" als ein Teil des Steinschlusses deklariert. Die beiden neuen Wissensbasen enthalten darüber hinaus bereits implizit eine eigene Verbindung, die semantisch besagt, dass der "Großvater" am Bau der Steinmauer beteiligt ist.

Dadurch erweitern die beiden externen Wissensbasen die eigenen Informationen durch die neu erstellten Relationen. Darüber hinaus jedoch lässt sich so ebenfalls eine neue Erkenntnis in den Daten schliessen: sowohl "Peter" als auch dessen "Großvater" sind direkt oder indirekt am Bau des "Steinschlusses" beteiligt.

DAS CIDOC CRM blub

DAS VISIT MODEL - VISMO blub

## 5.2 Technische Details zur Semantischen Datenbank

infrastruktur, hosting, allgemeiner zugriff

zertifizierung, sicherheit

drupal

wisski (allgemein)

rdf4j

grundfunktionalität

## 5.3 WissKI - Wissenschaftliche Kommunikationsinfrastruktur

aufhänger: wissenschaftlicher zugang, zugang für die geisteswissenschaftler und forser

connection zu den rdf daten

pfade

konfiguration

## 5.4 Technischer Zugang zu den Metadaten - die ViSIT REST API

aufhänger: technischer Zugang im gegensatz zum allgemeinen zugang  
allgemeine beschreibung

## 5.5 Wichtige Technische Charakteristika der Entwicklung und den Betrieb der Semantischen Datenbank

alle sachen wie scripte und details für die lauffähige API und DB erklären

## 5.6 Zusatzfeatures - Erweiterung der Bedienbarkeit der Semantischen Datenbank

copy and paste feature

excel importer

## 6 SCHLUSS

blub

## REFERENCES

- [Bar+16] Emanuel Barndl u. a. "Anno4j - Idiomatic Access to the W3C Web Annotation Data Model". In: *The Semantic Web - ESWC 2016 Satellite Events, Heraklion, Crete, Greece, May 29 - June 2, 2016, Revised Selected Papers*. 2016, S. 257–270. DOI: [10.1007/978-3-319-47602-5\\_43](https://doi.org/10.1007/978-3-319-47602-5_43). URL: [https://doi.org/10.1007/978-3-319-47602-5\\_43](https://doi.org/10.1007/978-3-319-47602-5_43).
- [Bir+16] Emanuel Birndl u. a. "Anno4j - Idiomatic Access to the W3C Web Annotation Data Model". In: *The Semantic Web - ESWC 2016 Satellite Events, Heraklion, Crete, Greece, May 29 - June 2, 2016, Revised Selected Papers*. 2016, S. 257–270. DOI: [10.1007/978-3-319-47602-5\\_43](https://doi.org/10.1007/978-3-319-47602-5_43). URL: [https://doi.org/10.1007/978-3-319-47602-5\\_43](https://doi.org/10.1007/978-3-319-47602-5_43).
- [Bur+16] Emanuel Burndl u. a. "Anno4j - Idiomatic Access to the W3C Web Annotation Data Model". In: *The Semantic Web - ESWC 2016 Satellite Events, Heraklion, Crete, Greece, May 29 - June 2, 2016, Revised Selected Papers*. 2016, S. 257–270. DOI: [10.1007/978-3-319-47602-5\\_43](https://doi.org/10.1007/978-3-319-47602-5_43). URL: [https://doi.org/10.1007/978-3-319-47602-5\\_43](https://doi.org/10.1007/978-3-319-47602-5_43).
- [Hit+07] Pascal Hitzler u. a. *Semantic Web: Grundlagen*. Springer-Verlag, 2007.
- [MM04] Frank Manola und Eric Miller. *RDF Primer*. W3C Recommendation. W3C, Feb. 2004. URL: <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>.