<Datenhaltung

Die Datenhaltung von OntoWiki erfolgt vollständig über den sogenannten Triplestore. In dieser Datenbank, die in Form eines MySQL oder eines Virtuoso Servers existieren kann, sind sowohl die von OntoWiki selbst verwendeten Texte und Daten gespeichert, als auch die Datenbestände, die von OntoWiki verwaltet werden. Aufgrund von Performanzbedenken entschieden wir uns für die MySQL Variante.

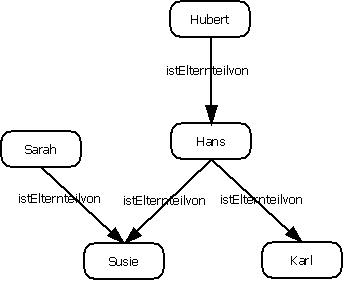
Die Speicherung der Daten in dieser Datenbank wird in der folgenden Grafik schematisch dargestellt.

Grafik 1:



Der Index ist hierbei eine fortlaufende Nummer, die als Primärschlüssel dient. Die eigentlichen Daten sind als Tripel der Form Subjekt-Prädikat-Objekt gespeichert. Das Subjekt bezeichnet hierbei stets eine Instanz einer Klasse, das Objekt eine andere Klasse oder ein Literal (String, Bool, Int, ... ) und das Prädikat die Beziehung zwischen den beiden. Mittels dieser Strukturen können sehr komplexe Beziehungen beschrieben und mithilfe von geeigneten Algorithmen weitere Informationen gewonnen werden. Es folgt der aus den Daten in Grafik 1 konstruierbare Baum.

Grafik 2:



Das Auslesen dieser Beziehungen wird, wie an anderer Stelle erwähnt, vom Erfurt Framework übernommen.

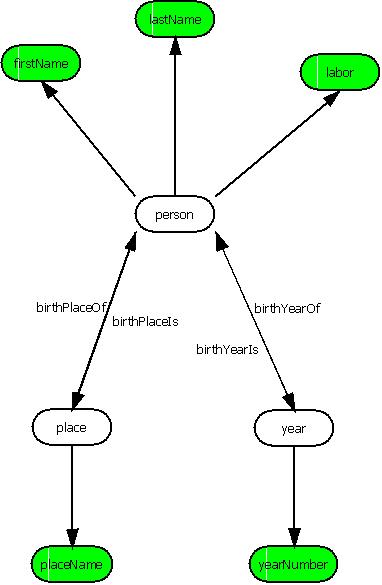
Aus der elementaren Beziehung "ist Elternteil von" kann man nunmehr weitere Informationen entnehmen, z.B. indem man abfragt welche Person Elternteil von einem Elternteil von Susie ist. Diese spezialisierte Form der Abfrage ist mittels der Abfragesprache SPARQL möglich.

Beispiel für SPARQL: (Abfrage aller Personen, die Elternteil von einem Elternteil von Susie sind)

SELECT $Großelternteil WHERE { $Großelternteil abstammung:istElternteilvon Susie}

Es ist durch Erweiterungen möglich, Datenbestände sowie Klassenmodelle aus XML Dateien wie RDF, OWL oder Turtle zu importieren oder als solche zu exportieren. Ein Klassenmodell wird dabei als Vokabular bezeichnet, und gibt einen Überblick über alle Objekte sowie über alle Beziehungen zwischen diesen (Vererbung, funktionale Relationen, ...). Für die grafische Darstellung verwenden wir UML-konforme Zeichen. Es folgt eine Grafik unseres für das Vorprojekt entwickelten Vokabulars. Es ist aus Zeit- und Aufwandsgründen natürlich erheblich einfacher als das im Hauptprojekt entwickelte, jedoch eignet es sich bereits gut, um die Funktionalität der Daten zu verdeutlichen.

Grafik 3:



Die farblosen Flächen sind Klassen von Objekten, denen über die Pfeile, die Object Properties darstellen, andere Objekte und Daten zugeordnet sind. Sie können innerhalb der semantischen Interpretation als Subjekte oder Objekte auftreten.

Die Object Properties treten an die Stelle von Prädikaten, die Instanzen von Klassen miteinander verbinden.

Die grün hinterlegten Flächen stellen Data Properties, also literale Daten, dar. Sie können nur als Objekte im Schema dienen, da sie nicht eindeutig identifizierbar sind. Dennoch sind sie die einzige Möglichkeit wirkliche Daten abzuspeichern.

Durch die inversen Object Properties birthPlaceOf und birthPlaceIs sowie birthYearOf und birthYearIs ist es möglich, ohne erheblichen Mehraufwand Orten die Personen zuzuordnen, die dort geboren wurden oder starben. Hierzu ist es nötig für Orte und Jahre eigene Klassen zu erstellen, wobei die spezifischen Informationen in placeName und yearNumber gespeichert werden.