**Projektangebot**

**1. Zielbestimmung**

Die Internetseite der Personendatenbank des Projekts "Wissenschaftsbeziehungen im 19. Jahrhundert zwischen Deutschland und Russland auf den Gebieten Chemie, Pharmazie und Medizin" der Sächsischen Akademie der Wissenschaften zu Leipzig soll in eine moderne kollaborative semantischen Webapplikation überführt werden. Dazu müssen bestehende Daten in eine semantische Datenbank überführt werden, umfangreiche Möglichkeiten geschaffen werden die Daten in einem Backend zu editieren und neue Daten einzufügen. Des weiteren soll ein schlichtes Frontend aufgebaut werden, in dem die Daten übersichtlich dargestellt werden. Dieses soll auch Suchmethoden integrieren, die angemessene Möglichkeiten bietet domänenspezifische Fragestellungen zu beantworten.

Die Personendatenbank enthält bio- und bibliographischen Daten über Wissenschaftler, die sowohl in Russland als auch in Deutschland tätig waren.

**2. Produkteinsatz**

Die Web Applikation wird auf der einen Seite dazu genutzt bestehende Daten zu bearbeiten und zu erweitern sowie neue Daten hinzuzufügen. Es sollen hierbei bio- und bibliographische Daten, die bei der Arbeit der Wissenschaftlerinnen, die am Projekt beteiligt sind, durch das einfügen in das Backend semantisch aufgewertet werden(Zwei Sätze). Insbesondere muss die zu modellierende Datenstruktur in der Lage sein alle Daten aufzunehmen, bzw. leicht zu erweitern sein um die Datenstruktur mit im Forschungsprozess erwartbaren neu auftauchende Strukturen erweitern zu können.

Die Arbeit der MedizingeschichtlerInnen an diesem Projekt ist langfristig ausgelegt. Das bedeutet, dass die Plattform eine stabile Basis bilden muss und die Persistenz der Daten in angemessenen Rahmen sichergestellt werden muss. Auch muss die Plattform einfache Erweiterungsmöglichkeiten bieten um auf technischer Seite leicht anpassbar zu sein und damit der Funktionsumfang leicht auszubauen ist und neuen Zielstellungen angepasst werden kann.

Auf der anderen Seite soll die Webapplikation dem Web User die Möglichkeit geben einfach im Datenbestand zu recherchieren und Suchmöglichkeiten zu nutzen, die dabei helfen fachspezifische Fragestellungen zu klären.

Weiterhin soll es möglich sein den Datenbestand in das Semantic Web einzupflegen. Das bedeutet, dass die Daten nicht nur visuell, sondern auch semantisch im Web zugänglich gemacht werden sollen. Damit soll es Projekten der Fachdomäne ermöglicht werden leicht auf die Daten zuzugreifen und die Daten mit eigenen Beständen abzugleichen, bzw. mit Daten anderer Projekte zu verknüpfen.

**4. Funktionalität und Arbeitspakete**

**4.1 Funktionalitäten**

**4.2 Arbeitspakete:**

**Vorprojekt 10%**

**Im Rahme des Vorprojekts sollen die Grundfunktionen des Softwarepaketes implementiert werden. Zielsetzung ist es ein OntoWiki System aufzusetzen und einen Versuchsdatensatz mit geringen Umfang (ca. 10 Datensätze) Einzufügen um die Funktionsweise von OntoWiki zu testen und dem Auftraggeber Grundsätzliche Funktionsweisen demonstrieren zu können.**

**Auf der anderen Seite soll ein einfaches Frontend die Grundfunktionalität der Darstellung der Daten für den Webuser zeigen und einfache Suchfunktionen demonstriert werden können.**

**Insbesondere ist dies die Umsetzung folgender Funktionalitäten in einer Testumgebung.**

**Datentransformation/ Haltung 40%**

**Ein Schwerpunkt der Arbeit wird der Umgang mit den Daten darstellen, da die Datengrundlage der Schlüssel für das Funktionieren des Softwarepakets darstellt. Je mehr Zeit in diese Grundlagenarbeit verwendet wird, desto größer sind die Möglichkeiten später das System zu erweitern.**

**Teilaspekt sind die Modellierung des Datenstruktur als Ontologie in enger Zusammenarbeit mit dem Auftraggeber. Hierbei soll durch Vokabular-Alignement eine größtmögliche Kompatibiltät mit bestehenden Semantischen Web Ressourcen ermöglicht werden.**

**Das Hauptaugenmerk muss hier auf Personen, Institutionen, Zeitpunkt, berufliche Laufbahn sowie Fachgebiete gesetzt werden, da diese die Hauptgrundlage für die Recherchemöglichkeiten liefern.**

**Optional soll versucht werden einen möglichst großen Teil der weiteren Daten semantisch zu spezifizieren.**

**Dann muss eine Schnittstelle aufgebaut werden, die es ermöglicht die Altdaten (vorhanden in HTML-Struktur) zu importieren und semantisch aufzuwerten.**

**Backendendwicklung 30%**

**Die Entwicklung des Backends umfasst das Aufsetzen des Frameworks, Anpassung und Implementierung von Widgets zur Dateneingabe und Darstellung der Daten im Backend. Die Spezifizierung und Realisierung von Schnittstellen. Und implementierung von Serverfunktionalitäten.**

**Weiterhin ist hier die Implementierung von optionalen in den Funktionalitäten formulierten Umsetzungen von Anforderungen zu nennen.**

**Frontendentwicklung 30%**

**In diesem Arbeitspaket soll die Darstellung für den web user realisiert werden. Hier liegt das Hauptaugenmerk auf die schlichte, Übersichtliche Darstellung der Daten und die Implementierung der spezifizierten Suchmöglichkeiten sowie die Implimentierung optionaler Funktionalitäten für das Frontend wie graphische Darstellungen.**

**Einrichten der Software auf dem Server der SAW und Übergabe 10%**

**Der letzte Teil der Arbeit wird die Portierung des Softwaresystems auf die Server der SAW und die evtl Schulung des Personals.**

**5. Qualitätssicherung**

**Um nach der Fertigstellung des Produkts schnell gewährleisten zu können, dass es produktiv eingesetzt werden kann ist es nötig das Hauptaugenmerk auf Funktionalität, Zuverlässigkeit und Benutzbarkeit zu legen.**

**Es muss gewährleistet sein, dass alle recherchierten Daten des Projekts leicht integriert werden können und die Benutzbarkeit darauf ausgerichtet ist, Nutzern mit normalen Computerkenntnissen einen leichten Einstieg in die Arbeit zu ermöglichen.**

**Da das Projekt langfristig ausgelegt ist, muss der Zuverlässigkeit ein hoher Stellenwert eingeräumt werden. In dem Zusammenhang ist auf eine ausgereifte Testphase vor dem produktiven Einsatz unabdingbar.**

**Insbesondere ist eine Rückkopplung zum Auftraggeber und eine in einem iterativen Prozess stattfindende Anpassung an die Bedürfnisse des Auftraggebers unabdingbar um einen schnellen Übergang in den produktiven Einsatz der Applikation zu ermöglichen. Die Umsetzung des Projekt Catalogus Professorum Lipsiensis hat gezeigt, das der Umfang dieser Arbeit nicht zu unterschätzen ist (Vgl. Riechert et al. 2010). Dieser Punkt wiegt schwerer als die Effizienz, da ohne Zuverlässigkeit das Produkt nicht einsatzfähig ist.**

**Die Wartbarkeit muss angemessen im normalen Bereich möglich sein. Hierbei sind keine Besonderen Anforderungen zu beachten.**

**Die Übertragbarkeit muss möglich sein, da aber das Produkt eine Grundlage der Arbeit der WissenschaftlerInnen bieten soll ist nicht damit zu rechnen, dass das Produkt schnell durch ein anderes ersetzt wird. Zusätzlich ist durch die geplante Architektur ohne viel Aufwand möglich vorhandene Schnittstellen für die Übertragung auf Folgeprojekte zu gewährleisten.**

**Die Produktqualität muss dem Projekt angemessen sein und den produktiven Einsatz ermöglichen.**

**Quellen:**

**Riechert, Thomas/Morgenstern, Ulf/Auer, Sören/Tramp, Sebastian/Martin, Michael: Knowledge Engineering for Historians on the Example of the *Catalogus Professorum Lipsiensis.* In:** P.F. Patel-Schneider et al.(Eds.): ISWC 2010, Part II, LNCS 6497, pp. 225–240, 2010.

**Übersicht der Qualitätsanforderugen**

**Sehr gut Gut Normal Weniger Relevant**

**Produktqualität X**

**Funktionalität X**

**Zuverlässigkeit X**

**Benutzbarkeit X**

**Effizienz X**

**Wartbarkeit X**

**Übertragbarkeit X**

**6. Glossar**

**Frontend**

**Das Frontend ist die Schnittstelle des Systems zu dem web user. Hier werden die Daten für die Öffentlichkeit visuell dargestellt**

**SPARQL Endpoint**

**Durch diese Schnittstelle wird es externen Projekten ermöglicht dierekt auf die Daten zuzugreifen.**

**Backend**

**Das Backend ist der Kern des Systems, dass die Datenverwaltung und Datenhaltung realisiert. Dem Administrator und den Content Editoren werden hier alle sie betreffenden Funktionalitäten zur Verfügung gestellt.**

**Altdaten**

**Das sind alle Daten, die bisher erfasst wurden und die in Form der bestehenden HTML-Daten zur Verfügung gestellt werden.**