



Vorstellung der Projektaufgaben

Prof. Dr.-Ing. Robert Tolksdorf & Markus Luczak-Rösch
Freie Universität Berlin
Institut für Informatik
Netzbasierte Informationssysteme
tolk@ag-nbi.de | markus.luczak-roesch@fu-berlin.de

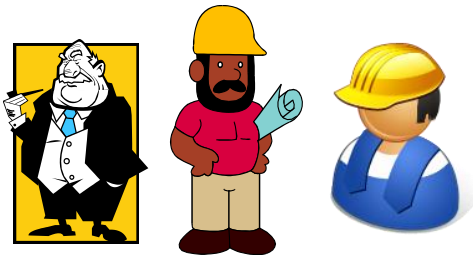


Organisatorisches → Rollen

1. Auftraggeber und Berater → Dozenten der AG NBI



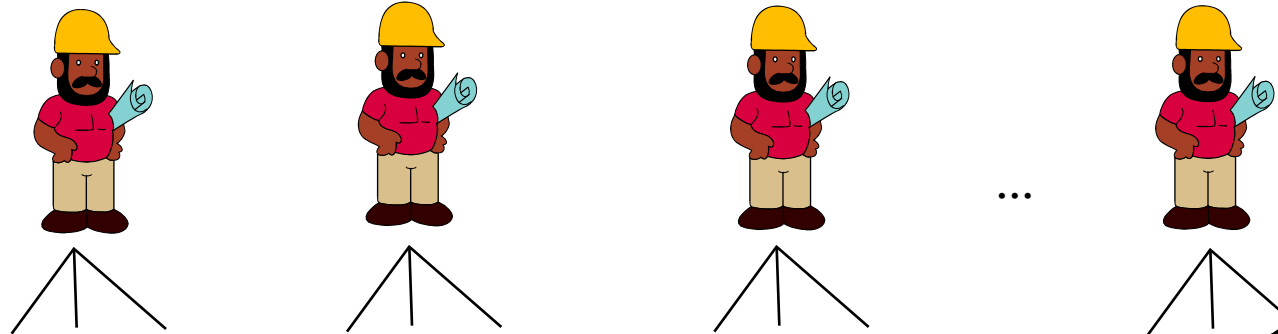
2. Projektmitarbeiter = Auftragnehmer →
Kundenprojektteilnehmer (Studenten)



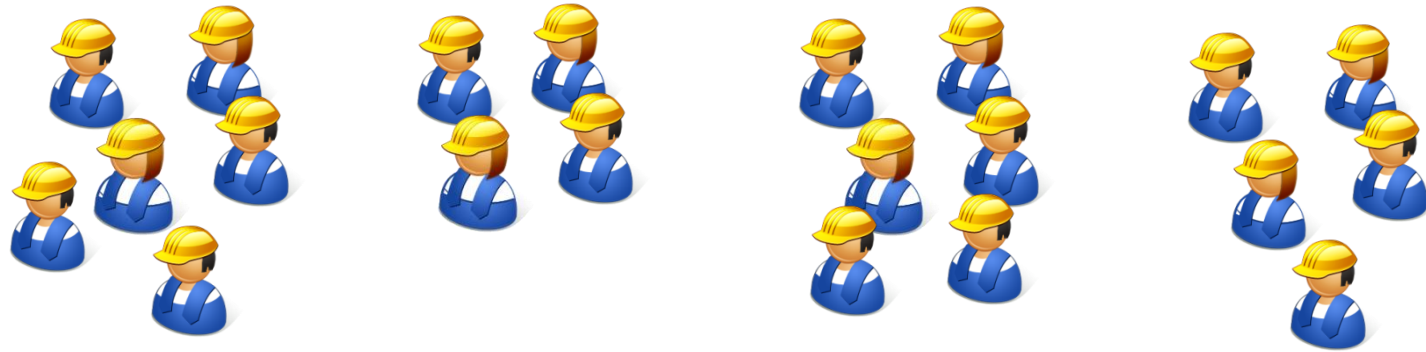
1 Projektkoordinator



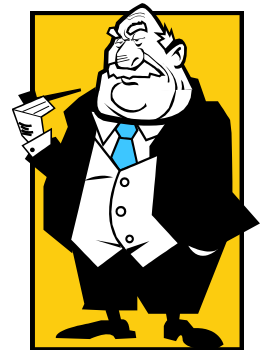
n Teilprojektleiter



Projektmitarbeiter



- **Pflichten**
 - verantwortlich/Ansprechpartner für Gesamtprojekt
 - Verantwortlich für das Projektteam
 - **verantwortlich für Zwischen-/Endpräsentationen**
- **Rechte**
 - **darf Veranstalter bei Problemen konsultieren**



- **Aufgaben**
 - Anforderungsanalyse
 - Planung, Leitung & Durchführung
- **Rollen (Auszug)**
 - Projektleiter & Teilprojektleiter
 - Software-Architekt
 - Entwickler
 - Tester
 - Autor der Dokumentation



- Jeder soll im Rahmen des Projekts mit XML-Technologien in Berührung kommen!
 - nicht nur leiten
 - nicht nur dokumentieren
 - ...
- Wir händigen (per Mail an Koordinatoren) eine Vorlage für Wochenberichte aus, die dokumentiert, was die einzelnen Personen getan haben
 - Projektkoordinator hat das Recht hier zu bewerten, ob die Angaben stimmen
 - Angaben mit Zeitaufwand (Stundenzettel)



Aufgabe 1

- **Ermöglichen von Xquery Anfragen für Datensätze aus allen (oder vielen) der registrierten OAI-Repositories und Visualisierung der Suchergebnisse unter anderem als navigierbares SVG**
- Was braucht man dazu?
 - Viele Daten (XML...)
 - Schnittstelle zum Anfrage stellen (Webservice/Applikation)
 - Mechanismus zum Duplikate filtern (?)
 - ...

- OAI- **Open Archives Initiative** stellt Liste von OAI-Repositories bereit
- Repositories zugreifbar via HTTP-Webservice
- Webservices liefern OAI-PMH 2.0 (ein OAI-Format-Standard) konformes XML

- **<http://www.openarchives.org/Register/ListFriends>**
- **Liste in XML-Format mit allen aktuell registrierten Daten-Providern**
- lesbare Fassung:
 - <http://www.openarchives.org/Register/BrowseSites>
- Services werden über URL mit Verb-Parameter (Aktionsaufruf) angesteuert, z.B.
 - Identify (Meta-Infos), ListRecords, ListSets, ListIdentifiers, etc.

← → ↻ www2.informatik.uni-stuttgart.de/cgi-bin/OAI/OAI.pl

OAI 2.0 Request Results

[Identify](#) | [ListRecords](#) | [ListSets](#) | [ListMetadataFormats](#) | [ListIdentifiers](#)

You are viewing an HTML version of the XML OAI response. To see the underlying XML use your web browsers view source option. More information about this XSLT is at the [bottom of the page](#).

Datestamp of response 2011-05-06T07:31:07Z

Request URL <http://www2.informatik.uni-stuttgart.de/cgi-bin/OAI/OAI.pl>

OAI Error(s)

The request could not be completed due to the following error or errors.

Error Code badVerb

No verb was specified.

[Identify](#) | [ListRecords](#) | [ListSets](#) | [ListMetadataFormats](#) | [ListIdentifiers](#)

← → ↻ www2.informatik.uni-stuttgart.de/cgi-bin/OAI/OAI.pl?verb=Identify ☆ » 🔍

OAI 2.0 Request Results

[Identify](#) | [ListRecords](#) | [ListSets](#) | [ListMetadataFormats](#) | [ListIdentifiers](#)

You are viewing an HTML version of the XML OAI response. To see the underlying XML use your web browsers view source option. More information about this XSLT is at the [bottom of the page](#).

Datestamp of response	2011-05-06T07:36:31Z
Request URL	http://www2.informatik.uni-stuttgart.de/cgi-bin/OAI/OAI.pl

Request was of type Identify.

Repository Name	Universität Stuttgart, Fakultät 5, Germany, Computer Science Archive
Base URL	http://www2.informatik.uni-stuttgart.de/cgi-bin/OAI/OAI.pl
Protocol Version	2.0
Earliest Datestamp	2000-11-10T10:11:00Z
Deleted Record Policy	transient
Granularity	YYYY-MM-DDThh:mm:ssZ
Admin Email	reportadm@informatik.uni-stuttgart.de

OAI-Identifier

Scheme	
Repository Identifier	informatik.uni-stuttgart.de
Delimiter	:
Sample OAI Identifier	oai:informatik.uni-stuttgart.de:INPROC-2006-49

[Identify](#) | [ListRecords](#) | [ListSets](#) | [ListMetadataFormats](#) | [ListIdentifiers](#)

- **Lösungsansätze**

- **Variante A) Nicht erlaubt!**
 - HTTP-Anfragen an Repositories stellen wenn Suchanfrage eingeht und Ergebnisse zusammenstellen

- **Variante B)**
 - Daten einsammeln
 - XML-Datenbank aufbauen
 - XML-Datenbank anfragen
 - Ergebnisse Visualisieren

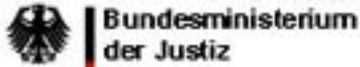
- Seitenweise (10.000 Treffer?) Ausgabe der Daten nach Suchanfrage
- nicht nur Datensätze, auch alle zugehörigen Quellen (repository-URLs)
- ggf. Datum der Aufnahme in die Datenbank (Aktualität der Daten...)
- navigierbares SVG

- Ausgabe z.B. in OAI-PMH XML (mit Quellen-Ergänzungen) transformiert durch **OAI2 to HTML XSLT Style Sheet**
 - <http://www.eprints.org/software/xslt.php>



Aufgabe 2

- **Gesetztestexte sowie deren Verweise crawlen und in einer XML-Datenbank speichern, Daten und navigierbares SVG über einen Web Service ausgeben (Daten speichern in einem eigenen XML-Format, Ausgeben zusätzlich in GraphML und SVG)**
- Datenquelle: www.gesetze-im-internet.de
 - iso-8859-1 codiertes XHTML
 - leider nicht immer valide aber grundsätzlich sauberer (X)HTML-Code



Startseite

Gesetze / Verordnungen

Aktualitätendienst

Titelsuche

Volltextsuche

Translations

Hinweise

Impressum

Tastenkombinationen

Verwaltungsvorschriften im Internet

Gesetz über die Durchführung von Maßnahmen des Arbeitsschutzes zur Verbesserung der Sicherheit und des Gesundheitsschutzes der Beschäftigten bei der Arbeit

zur Gesamtausgabe der Norm im Format: [HTML](#) [PDF](#)

Erster Abschnitt

Allgemeine Vorschriften

- [§ 1 Zielsetzung und Anwendungsbereich](#)
- [§ 2 Begriffsbestimmungen](#)

Zweiter Abschnitt

Pflichten des Arbeitgebers

- [§ 3 Grundpflichten des Arbeitgebers](#)
- [§ 4 Allgemeine Grundsätze](#)
- [§ 5 Beurteilung der Arbeitsbedingungen](#)
- [§ 6 Dokumentation](#)
- [§ 7 Übertragung von Aufgaben](#)
- [§ 8 Zusammenarbeit mehrerer Arbeitgeber](#)
- [§ 9 Besondere Gefahren](#)
- [§ 10 Erste Hilfe und sonstige Notfallmaßnahmen](#)

- **§ 2 Begriffsbestimmungen**
- (1) Maßnahmen des Arbeitsschutzes im Sinne dieses Gesetzes sind Maßnahmen zur Verhütung von Unfällen bei der Arbeit und arbeitsbedingten Gesundheitsgefahren einschließlich Maßnahmen der menschengerechten Gestaltung der Arbeit.
- (2) Beschäftigte im Sinne dieses Gesetzes sind:
 - ...
 - 3. arbeitnehmerähnliche Personen im *Sinne des § 5 Abs. 1 des Arbeitsgerichtsgesetzes*, ausgenommen die in Heimarbeit Beschäftigten und die ihnen Gleichgestellten
 - ...

- **Gesetzestexte bestehen aus:**
- Name des Gesetzes/der Norm
- Abschnitte
 - Paragraphen
 - Absätze
 - Sätze in Absätzen
 - Aufzählungen
- Referenzen auf andere Elemente von Gesetzestexte (z.B. Paragraphen)

- Auswerten der HTML-Ausgabe
- Konvertierung der Daten in ein eigenes XML-Format (Repräsentation der Verweise)
- XSLT Transformation in das GraphML Format
- „Textanalyse“ (z.B: Reguläre Ausdrücke) zur Auffindung von Querverweisen („siehe §5 ArbSchG“, „ im Sinne des §5 Abs. 1 des Arbeitsgerichtsgesetzes“)
- Ausgabe via Webservice als
 - GraphML und navigierbares SVG
 - Demoseite, die das SVG nutzt

- **Technologien für beide Aufgaben?**
- XML
- XSLT
- Xquery
- HTTP
- HTML / CSS
- SVG



Ablauf, Abgabe und Scheinkriterien

24.05.11			Tutorium A (Übungsblatt 1)
25.05.11			Tutorium B
31.05.11			Tutorium A (Übungsblatt 2)
01.06.11		Projektarbeit & Tutorium	Tutorium B
07.06.11			Tutorium A (Übungsblatt 3)
08.06.11	Praktischer Block (Übungsphase)		Tutorium B
14.06.11			
15.06.11		Präsentation	Meilensteinpräsentation Projektphase - Thema 1 & 2 (MS 1)
21.06.11			Tutorium A (Übungsblatt 4)
22.06.11		Projektarbeit & Tutorium	Tutorium B
28.06.11			Tutorium A (Übungsblatt 5)
29.06.11			Tutorium B

- **Tutorium: 45 Minuten am Beginn des Termins**
 - Di. & Mi. danach: 45 Minuten Technologiecoaching durch Tutor (Fragen Sie ihn)
- **immer mittwochs ab 15:05: 5 Minuten Stand-Up für jede Projektgruppe mit Dozenten (Projektkoordinator berichtet)**
 - Projektplanung im ersten Stand-Up präsentieren (Arbeitspakete, Gantt-Chart...)

24.05.11			Tutorium A (Übungsblatt 1)
25.05.11			Tutorium B
31.05.11			Tutorium A (Übungsblatt 2)
01.06.11		Projektarbeit & Tutorium	Tutorium B
07.06.11			Tutorium A (Übungsblatt 3)
08.06.11	Praktischer Block (Übungsphase)		Tutorium B
14.06.11			
15.06.11		Präsentation	Meilensteinpräsentation Projektphase - Thema 1 & 2 (MS 1)
21.06.11			Tutorium A (Übungsblatt 4)
22.06.11		Projektarbeit & Tutorium	Tutorium B
28.06.11			Tutorium A (Übungsblatt 5)
29.06.11			Tutorium B

- Meilensteinpräsentation am 15.06. durch Projektkoordinator
 - 10 min. pro Team
 - Was wollten Sie erreichen?
 - Was wurde erreicht?
 - Was wird noch erreicht?

- Demo als Teil der Abschlusspräsentation
- Code und Doku auf Datenträger bis Fr. 08.07.
(Codefreeze 06.07.!!!)
- Doku in angemessener Form und Länge einfach gedruckt bis Fr. 08.07.

- Das Projekt wird im Rahmen der regelmäßigen und aktiven Teilnahme gewertet
 - von Aktivität hängt das Projektergebnis ab (schwaches oder unvollständiges Ergebnis stellt die Aktivität in Frage)

- Abgabe des Stundenzettels als fortlaufend geführtes Dokument durch Projektkoordinator an Veranstalter
 - Projektkoordinator bekommt Wochenberichte mit Zeitangabe per Mail von Team
 - bewertet die Stichhaltigkeit der Angabe durch Kommentar
 - schickt jeweils montags bis 12 Uhr das für die Vorwoche aktualisierte Dokument an Veranstalter
- Veranstalter schickt die Vorlage in dieser Woche an die Koordinatoren

Was nun?

- Entscheidung für eine Aufgabe
- Gruppenliste ausfüllen
- Sternchen bei Koordinator und zurück an Veranstalter

- Koordinator muss
 - schnell Infrastruktur aufbauen (Trac?, Protokollvorlage,...)
 - Projekt planen (im Dialog mit seinem Team)