

Zwischenbericht zum Großen Beleg

Dokumentenverwaltung von ProjectWeb unter Berücksichtigung des Knowledge Engineerings

von

Stefan Fromm

TU Dresden

Fakultät Informatik

Institut für Software- und Multimediatechnik

Lehrstuhl Softwaretechnologie

Betreuerin: Anne Thomas

1 *Einführung in ProjectWeb*

2 Konzepte aus Wissensmanagement und Bibliothekswesen

3 Anforderungen an die Dokumentenverwaltung

4 Aktueller Stand

5 Ausblick

- ProjectWeb = Plattform für verteilte Arbeitsumgebungen von Forschungsgruppen
- Integration verschiedener Dienste in einheitlicher Oberfläche
- Internet-Anwendung → räumlich verteiltes Arbeiten
- Koordination und Kommunikation zwischen Mitarbeitern
- Mehrsprachigkeit
- Instanzenerzeugung für konkrete Arbeitsumgebungen
- Privater und öffentlicher Bereich
- Modularer Aufbau
- Kern und optionale Module

- im Rahmen des Komplexpraktikums 2001 entstanden
- Prototypische Umsetzung einer Dokumentenverwaltung
- Entwurfsfehler im modularen Konzept
- Speziell Dokumentenverwaltung:
 - keine Referenzen und Vorlagen
 - keine Koordination beim Editieren von Dokumenten
 - einfache Suchmechanismen, ...
- Verbesserungen
 - Vielseitig konfigurierbare GUI
 - Modulkonfiguration mit XML
 - Portierung auf kommerziell verwendbaren freien J2EE-Server
 - weitere Hilfsmodule

- **Instanzenmanager**: Laden und Initialisieren der Module
- **Konfigurationsmanager**: liefert Konfigurationsdaten des Projekts und der Module
- **Nutzer- und Rollenverwaltung**: Verwalten von Nutzern und ihren Rechten
- **Projektstrukturverwaltung**: hierarchisch gegliederte Projekteinheiten
- **Sichtbarkeitenverwaltung**: Zugriffsbeschränkungen für Objekte anhand der Projektstruktur
- **Dateiverwaltung**: Upload / Download von Dateien, Kompression
- **Mailverwaltung**: Senden von Mails, Verwalten von Mailinglisten
- **Nutzerprofilverwaltung**: nutzerspezifische Einstellungen ablegen

1 Einführung in ProjectWeb

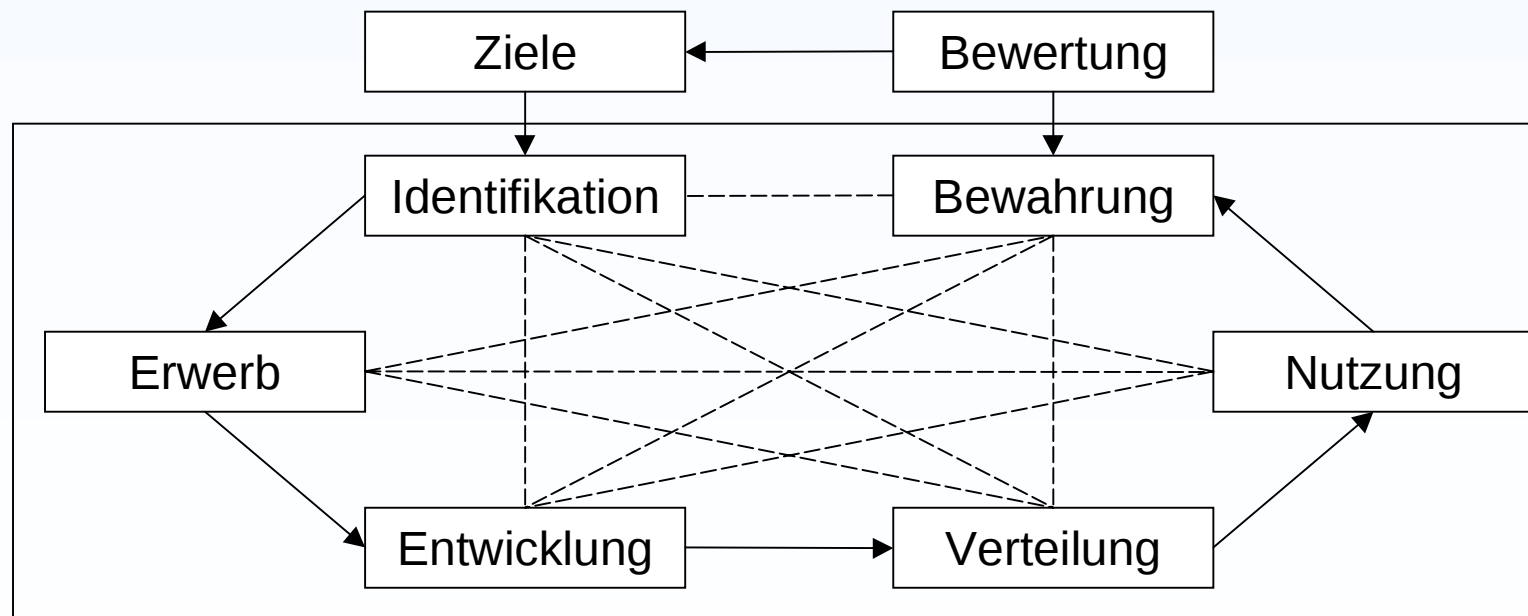
2 Konzepte aus Wissensmanagement und Bibliothekswesen

3 Anforderungen an die Dokumentenverwaltung

4 Aktueller Stand

5 Ausblick

- Wissen beeinflusst Wertschöpfung von Unternehmen
- Wissen in Dokumenten enthalten → muss verwaltet werden
- Wissensmanagement bietet dafür nötige Grundlagen: Organisationsmodelle, Nutzen der IT, Beherrschen von Komplexität, ...
- Bausteine des Wissensmanagements:



Quelle: Heidi Heilmann: Wissensmanagement – ein neues Paradigma?

- Unternehmenskultur wichtig für Erfolg von Wissensmanagement
- Dokumentenverwaltung unterstützt:
 - Wissensverteilung (Pull- und Push-Modell)
 - Wissensbewertung
 - Wissensbewahrung
 - Wissensnutzung
- Integration der Dokumentenverwaltung in übergeordnetes Informations- und Assistenzsystem
- Beherrschung von Komplexität
 - Metadaten zu Dokumenten
 - Konzept der strukturierten Vernetzung

- In Bibliothek: Verwaltung von Bücherbeständen in Regalen
- Suche von Büchern durch ein Informationssystem anhand von Metadaten
- Dokumentreferenzen verweisen auf Bücher etc.

- Thesaurus definiert Beziehungen zwischen Begriffen (Synonym, Oberbegriffe, verwandte Begriffe, Antonyme)
- Thesaurus verwendbar für:
 - Beheben von Terminologieproblemen
 - Ähnlichkeitsdefinition zwischen Begriffen
- in Dokumentenverwaltung verwendet für Suche nach Schlagwörtern

1 Einführung in ProjectWeb

2 Konzepte aus Wissensmanagement und Bibliothekswesen

3 *Anforderungen an die Dokumentenverwaltung*

4 Aktueller Stand

5 Ausblick

- Verwaltet werden: Dokumente, Referenzen, Vorlagen, Beziehungen zwischen Objekten (Literaturquellen, ...)
- Strukturierung von Objekten
- Versionierung von Dokumenten
- Suche vs. Blättern
- Bewertung und Löschen
- Interessenprofile
- projektinterne und öffentliche Dokumente
- Zugriff projektintern anhand definierbarer Sichtbarkeiten
- umfangreiche Schnittstellen

- Anforderungen:
 - fehlertolerante Suche nach Dokumenten und Referenzen
 - einfache und erweiterte Suche
 - Verwenden eines Thesaurus für Schlagworte
 - Sortieren der Trefferliste nach Relevanz
 - keine Volltextsuche im Inhalt von Dokumenten
 - keine fortgeschrittene Textanalyse (phonetische Analyse, Grundformenzerlegung, Vergleichswortverzeichnis, ...)
- Benötigt wird: Ähnlichkeit von Objekten zu einer Suchanfrage
- Berechnung der Ähnlichkeit: eigener Algorithmus
- Mustersuche in Texten mit Leichtvariante von *agrep*
- Umsetzung erfordert Einschränkungen, weil hohe Netzlast bei großen Datenmengen

- Suchanfrage = Anzahl von Kriterien (Auswahllisten, Muster, Datumsrelationen)
- Vorverarbeitung bei Kriterien mit Thesaurus → Einbeziehung der zu den eingegebenen Suchwörtern in Beziehung stehenden Wörter
- Festlegen der jeweiligen Begriffsrelevanz der endgültigen Suchwörter
- Vergabe der Punkte pro Treffer t für ein Suchwort w_i bei Mustern
$$p(t, w_i) = \text{brel}(w_i) * [1 + f_{\max} - \text{fehler}(t)] \quad (f_{\max} - \text{max. Fehleranzahl})$$
- bei Auswahllisten und Datumsrelationen je einen Punkt pro Übereinstimmung
- Gewichtete Summe der Punkte für ein Kriterium: Gewichte α , β , γ je nach Typ des Suchworts (Muss, Kann, Darf-Nicht):
$$P_i = \alpha * \sum p(t, w_i) + \beta * \sum p(t, w_i) - \gamma * \sum p(t, w_i)$$
- **Ähnlichkeit** = Mittelwert aller Quotienten $P_i / \text{mögliche Punkte } P_{\max}$

1 Einführung in ProjectWeb

2 Konzepte aus Wissensmanagement und Bibliothekswesen

3 Anforderungen an die Dokumentenverwaltung

4 Aktueller Stand

5 Ausblick

- Einarbeitung in Richtlinien des Wissensmanagements und Bibliothekswesens geschehen
- Anforderungsheft erarbeitet
- Teilsysteme identifiziert und analysiert
 - Dokumentenverwaltung
 - Thesaurus
 - Textfunktionen für Suche, Behandlung von Falscheingaben, ...
 - Generator z.B. für Referenznummern von Dokumenten
- Algorithmus für Relevanzberechnung erstellt
- Spezifikation der Modulschnittstelle und des Datenmodells für Dokumentenverwaltung erledigt

- Statistiken zur Dokumenten- und Referenzenbewertung folgt
- Spezifikation der restlichen Module (Schnittstellen und Datenmodell)
- Betrachten von nebenläufigen, ähnlichen Produktentwicklungen
- Beginn der Implementation ab August bis vsl. Ende Oktober