

Kurzbeschreibung DIWA

Bezeichnung: DIWA (Dokumentenverwaltung in Web-Archiven der Umweltdienststellen Baden-Württemberg)

Funktion: Archivierung von Dokumenten in Web-Archiven; Einstellen von Dokumenten in die Web-Archive durch einen großen Autorenkreis; Recherche nach Dokumenten über Volltext- und Metadaten-Suche (Dateiformate MS Office, HTML, PDF, plain text, Binärdateien) innerhalb des Intranets

Beschreibung: 1. Technische Voraussetzungen

DIWA nutzt den MS Index Server (MS IS) als Suchmaschine und den MS Internet Information Server (MS IIS) als Web Server. MS IIS und MS IS arbeiten auf Windows NT 4.0 Servern. Für die Indizierung von PDF-Dokumenten wird der Adobe-PDF-Filter für den MS IS benötigt. Für die Überprüfung der Zugriffsrechte auf entfernten Servern aus DIWA heraus ist die Installation von Perl notwendig. Der Umweltthesaurus wird unter MS Access betrieben.

Der Klient benötigt einen Java 1.1-fähigen Web-Browser (Internet Explorer ab Version 4.01, Netscape Navigator ab Version 4.08).

2. Nutzung MS IIS / MS IS in DIWA:

Zur Nutzung von MS IIS / MS IS werden von DIWA eine Autoren- und eine Recherchekomponente zur Verfügung gestellt.

Im MS IIS wird eine virtuelle Ablagestruktur für ein Web-Archiv angelegt. Diese virtuellen Verzeichnisse werden mit den entsprechenden physischen (gemounteten) Verzeichnissen verbunden. Der MS IS indiziert diese Verzeichnisse. Um die virtuelle Ablagestruktur für die Autorenkomponente von DIWA und den DIWA Explorer bereitstellen zu können, wurde ein spezielles Administratorskript für den MS IIS entwickelt.

Im MS IS werden neben dem Volltext sowohl Standardeigenschaften der MS Office-Dokumente (Autor, Titel etc.) und HTML-Dateien (Meta-Tags author, keywords etc.), als auch benutzerdefinierte Eigenschaften indiziert (MS Office-Dokumente: Eigenschaften → Anpassen; HTML: selbstdefinierte Meta-Tags). Diese Funktionalität des MS IS wird in DIWA dahingehend genutzt, dass die Metadaten in den Dokumenten und nicht von ihnen getrennt verwaltet werden.

3. Einstellen von Dokumenten

Die Autorenkomponente von DIWA erlaubt das Einstellen von Dokumenten in ein Web-Archiv, welches zuvor über eine virtuelle Ablagestruktur im MS IIS eingerichtet wurde. Dabei ordnet der Autor sein mit Metadaten versehenes Dokument einem Thema der virtuellen Ablagestruktur zu. Beim Einstellen werden die Schreibrechte in diesem Verzeichnis auf Betriebssystemebene überprüft.

4. Suche:

Die Suche in DIWA erfolgt über verschiedene Suchformulare oder den navigatorischen Zugang. Bei den Suchformularen wird zwischen einer einfachen Volltextsuche und der erweiterten Suche unterschieden. Letztere enthält zusätzlich zur Volltextsuche die Recherche über die Metadaten der einzelnen Dokumentenbestände. Der navigatorische Zugang (DIWA Explorer) erlaubt die Suche nach Dokumenten über die virtuelle Ablagestruktur in Form eines Themenbaums.

Die Suchformulare in DIWA sind auf die Funktionalität des MS IS zugeschnitten. Die Abfragesprache umfasst die im MS IS angebotenen Funktionalitäten. Die Inhalte der einzelnen Suchfelder eines erweiterten Suchformulars werden UND-verknüpft.

Für eine Suche kann der Suchbereich auf einzelne Dokumentenbestände/Themen der virtuellen Ablagestruktur im MS IIS über den DIWA-Explorer eingeschränkt werden.

Die Anwender erhalten stets alle Ergebnisse für ihre Suchanfrage. Erst beim Öffnen/Herunterladen einer Datei werden die Zugriffsrechte auf Betriebssystemebene überprüft.

Technisch wird für eine Suche eine spezielle Anfragedatei (IDQ) für den MS IS in DIWA dynamisch generiert. Die in der Oberfläche eingegebenen Suchbegriffe werden in diese Datei integriert, die schließlich an den MS IS weitergeleitet wird. Die Ergebnisformatierung erfolgt wiederum über spezielle Dateien (HTX). Diese erlauben eine einheitliche Visualisierung der Ergebnisse. Über Variablen in den speziell für DIWA entwickelten HTX-Dateien werden die von den Anwendern gewünschten Ergebnisformatierungen realisiert.

Anwendung im
UIS B-W: DIWA wird im UVM und in der LfU eingesetzt.

Versionen: DIWA befindet sich in der Testphase

Lizenzen: DIWA ist Eigentum des Landes Baden-Württemberg, vertreten durch das Ministerium für Umwelt und Verkehr. DIWA kann nach besonderer Vereinbarung und Regelung des Entgelts zur Nutzung außerhalb der Zuständigkeitsbereiche des Eigentümers weitergegeben werden.

Informationen: R.Mayer-Föll, K.Tochtermann (Hrsg.): Dokumentenverwaltung in Web-Archiven der Umweltdienststellen Baden-Württemberg – DIWA 1999, Ministerium für Umwelt und Verkehr Baden-Württemberg, 2000.

<http://www.uvm.baden-wuerttemberg.de/lfu/uis/globus/globus6/08-faw-diwa/gl0608.html>

