

Das Offline Web Archiv 200.035.200¹

Das *Offline Web Archiv* (OWA) ist ein System der *oia GmbH* zur Archivierung von Webpräsenzen.

Redundanzfreie Datenspeicherung

Mit dem Offline Web Archiv können Websites im Internet und Intranet in beliebig dichten Intervallen gespiegelt und gespeichert werden. Durch die redundanzfreie Speicherung der Dokumentdaten ist der Speicherplatzbedarf so gering wie möglich. Die Archivierung erfolgt vollständig unabhängig vom jeweiligen produktiven Content-Management-System.

Langzeitspeicherung

Das *Offline Web Archiv* unterstützt die Funktionen für die Langzeitspeicherung wie sie in ISO14721 gefordert sind. Darüber hinaus können die Daten in systemunabhängigen „Containern“ mit vollständiger XML-basierter Metadatenbeschreibung abgelegt werden (Archival Information Package – AIP). Die AIP ermöglichen den verlustfreien Austausch von Daten und Metadaten zwischen OWA Systemen. Wahlweise steht für die Datenablage auch das WARC-Format (ISO 28500) zur Verfügung.

Webweite Verfügbarkeit

Der Webclient bietet den gesamten Funktionsumfang. Durch die integrierte Softwareverteilung stehen alle Funktionen gemäß Sicherheitsrichtlinien webweit zur Verfügung.

Bestands- und Sammlungsbildung

Die Einrichtung und Verwaltung von Sammlungen und Beständen ist einfach und kann durch den Administrator leicht an die Archivorganisation und Bestandsstruktur angepasst und bedarfsgerecht erweitert werden. Dazu kann der Administrator eine hierarchische Verwaltungsstruktur mit beliebig vielen Ebenen aufbauen, in denen er Projekte zu Projektgruppen zusammenfasst und beschreibt.

Projektspezifische Spiegelungsparameter

Die Spiegelungsparameter und Spiegelungsintervalle können projektweise nach fachlichen Gesichtspunkten eingestellt werden – das System steuert und überwacht die Ausführung.

Selektive Nachbearbeitung und Online-Analyse

Mit dem „Online-Abgleich“ können die zu archivierenden Ressourcen seiten- und dateigenau ausgewählt und archiviert werden. Die integrierte Seitenanalyse erlaubt die Datei- und protokollgenaue Selektion der zu einer Seite gehörigen Ressourcen.

¹ Stand: März 2012

Archivierung interaktiver multimedialer Inhalte

Durch das speziell für OWA entwickelte Verfahren zur Analyse und Archivierung interaktiver multimedialer Inhalte („Flashes“) können jetzt auch Seiten mit „navigierbaren Flashes“ und eingebundene Videoclips archiviert werden. Die Funktionalität der „navigierbaren Flashes“ bleibt im Archiv weitgehend erhalten.

Archivierung von Intranet- und passwortgeschützten Webseiten

Die selektive Archivierung von passwortgeschützten Intranetseiten ist ebenso möglich wie die Archivierung von Internetseiten, die teilweise ohne Rootzertifikate (http) und teilweise mit Rootzertifikaten (https) veröffentlicht werden.

Einfache Navigation im Web Archiv

Unabhängig von Umfang und Anzahl der archivierten Websites und Dokumente kann der Benutzer selbstständig und einfach recherchieren, historische Websites und Dokumente lesen und drucken oder auch auf einen lokalen Datenträger kopieren. Die Navigation in den archivierten Daten ist einfach – das System unterstützt die vielfältigsten Nutzungen von der Rekonstruktion ganzer Websites zu unterschiedlichen Zeitpunkten über die gezielte Verfolgung der Veränderungen über einen frei definierbaren Zeitverlauf.

Volltextrecherche in Meta- und Primärdaten

Alle Meta- und Primärdaten sind über eine Volltextdatenbank recherchierbar. Die Treffermenge kann über Kriterien wie Bestandszugehörigkeit und Zeitpunkt oder Zeitraum eingegrenzt werden. Die Darstellung der Treffermengen erfolgt wahlweise in Linklisten oder Verlaufsstatistiken. Durch die Recherche in den Verlaufsstatistiken nach dem „Drill-Down-Prinzip“ werden auch große Datenmengen erschlossen, übersichtlich dargestellt und leicht nutzbar gemacht.

Daten- und Erhaltungsanalyse

Die integrierten Analysewerkzeuge ermöglichen die Bewertung der zu archivierenden und archivierten Webpräsenzen angefangen vom Datenumfang bis hin zu Dateitypen und Dateiversionen. Das integrierte Transformationswerkzeug bietet Möglichkeiten zur Konvertierung (Transformation) von Medien- und Dateitypen oder Dateiversionen, ohne Verlust der Originaldateien.

Benutzerverwaltung

Die integrierte Benutzerverwaltung stellt sicher, dass Projekte ausschließlich von autorisiertem Personal angelegt und die Daten vor der Archivierung nur von berechtigten Personen bearbeitet werden können.

Systemvoraussetzungen *Offline Web Archiv V 200.035.*

Stand: Januar 2012

Alle Funktionen können auf einem Rechnersystem installiert und genutzt werden. Um eine optimale Performance bei großem Datendurchsatz, insbesondere während laufender „Spiegelungen“ zu erreichen, sollten jedoch mindestens drei Rechnersysteme eingesetzt werden: ein Benutzerclient für alle Verwaltungs- und Recherchearbeiten, ein „DataAnalyser“ für die Durchführung der Spiegelungen, die Zwischenspeicherung der Spiegelungsdaten und die Aufbereitung der Daten für das Archiv sowie ein Datenbankserver.

Benutzerclient

Hardware:

Intel Pentium 4 Prozessor oder vergleichbar; 2 GB RAM Hauptspeicher empfohlen. *Software:* Windows XP (ab SP3) oder höher, Windows 7, .NET 4.0.

DataAnalyser

Hardware:

Prozessor 1x Dual-Core, empfohlen 2x Dual-Core oder vergleichbar; Hauptspeicher: mindestens 4 GB RAM, empfohlen 8 GB RAM; Plattenspeicher mindestens 250 GB.

Software:

Microsoft Server 2008 oder höher und .NET 4.0.

Datenbankserver und ServiceBroker

Hardware:

Prozessor 1x Quad-Core, empfohlen 2x Quad-Core oder vergleichbar; Hauptspeicher mindestens 8 GB RAM, empfohlen 16 GB RAM. Plattenspeicher mindestens 1000 GB, empfohlen² 2000 GB.

Software:

Microsoft Server 2008 oder höher und .NET 4.0, empfohlen Microsoft SQL-Server 2008³.

² Die Planungsgrößen für Ihr *Offline Web Archiv* richten sich danach, wie viele URLs in welcher Frequenz archiviert werden sollen. Gerne beraten wir Sie bei der Planung Ihres *Offline Web Archivs*.

³ Weitere Datenbanken auf Nachfrage.