

# Konzepte zur Datenhaltung für Webseiten in einem Web-Content-Management-System

Web-Content-Management-Systeme (WCMS) dienen dazu, komplexe Websites zu verwalten und den Autoren einzelner Webseiten möglichst viele administrative Aufgaben abzunehmen. Hierzu gehört auch, daß das WCMS den Content, sprich den textuellen Inhalt einer Seite, in einer bestimmten, vordefinierten Art und Weise speichert.

Man unterscheidet hierbei mehrere unterschiedliche Konzepte, sowie ihre Mischformen. Populär sind dabei besonders die Konzepte des Publishing-/Staging-Server und des dynamischen Publishing.

## 1. Publishing-/Staging-Server

Beim Konzept des Publishing-/Staging-Servers benutzt man zwei verschiedene Bereiche zur Ablage und Bearbeitung der Seiten. Der Publishing-Server enthält das eigentliche WCMS und beantwortet die Anfragen der Autoren, Redakteure und Administratoren. Dieser Server befindet sich meist im Intranet und ist für die Öffentlichkeit nicht erreichbar. Alle Texte die von Autoren erstellt und bearbeitet wurden, werden auf diesem Server gespeichert.

Sollen die Inhalte veröffentlicht werden, wird durch das WCMS ein Prozeß gestartet, welches aus allen

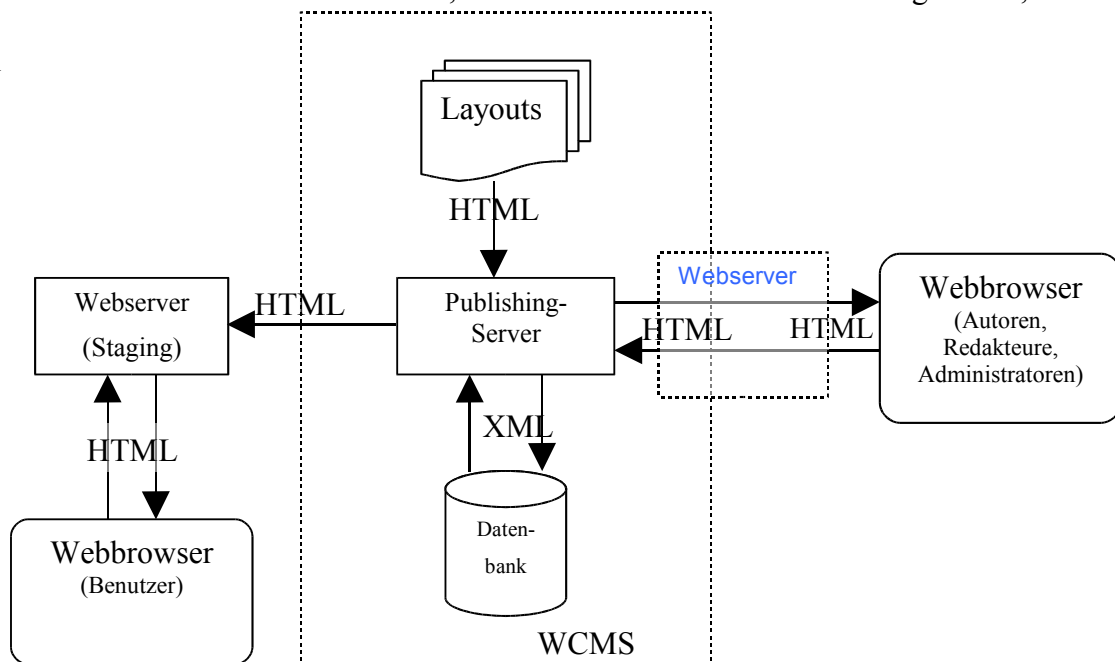


Abb. 1. Publishing-/Staging-Server

vorhandenen Texten HTML-Dateien erstellt. Diese HTML-Dateien werden dann auf dem Staging-Server gelegt und so für die Öffentlichkeit abrufbar gemacht.

Die folgende Abbildung verdeutlicht dieses Konzept: Autoren, Redakteure und Administratoren greifen zur Bearbeitung der Texte auf den Publishing-Server zu. Dies geschieht in der Abbildung je nach Wahl durch ein server- oder clientseitiges Interface. (Gegebenenfalls muß ein Webserver die Requests zwischen Autoren und Publishing-Server entgegennehmen und weiterleiten.)

Auf dem Publishing-  
 verwaltet und  
 gewünschter  
 werden dann daraus  
 erstellt und auf den  
 geschoben.  
 Die HTML-Dateien  
 sind dabei statisch  
 eingebundenen  
 JavaScript, Flash oder  
 Die Bearbeitung von  
 dem Publishing-  
 eine HTML-Datei auf  
 geändert, wird diese  
 Veröffentlichung  
 Server überschrieben.  
 organisatorisches  
 bietet dieses  
 der Redundanz: Fällt  
 kann trotzdem noch  
 Server gearbeitet  
 Publishing-Server aus,  
 können immer noch die  
 Webseiten vom Staging-  
 Server gelesen  
 werden.

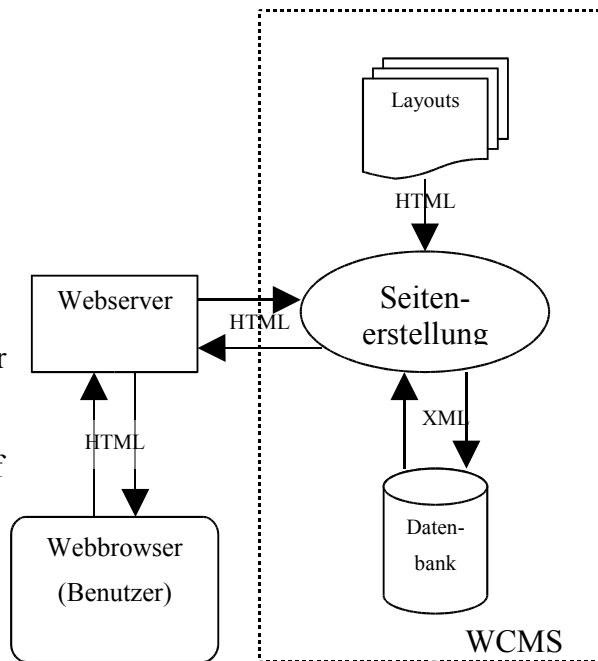


Abb. 2. Dynamisches Publishing

Server werden alle Texte  
 bearbeitet. Bei  
 Veröffentlichung  
 die HTML-Dateien  
 Staging-Server  
 auf dem Staging-Server  
 (sieht man von  
 Webtechniken, wie  
 ähnlichem ab).  
 Texten kann nur auf  
 Server erfolgen. Wird  
 dem Staging-Server  
 bei der nächsten  
 durch den Publishing-  
 Zwar bringt dies ein  
 Problem mit sich, jedoch  
 Vorgehen den Vorteil  
 der Staging-Server aus,  
 auf dem Publishing-  
 werden, bzw. fällt der

Die Vor- und Nachteile für die Verwendung von statischen Seiten über den Staging-Server lassen sich wie folgt zusammenfassen:

- + Gute Performance
- + Unkompliziertes Backup
- + Statische Seiten werden von Suchmaschinen erfasst
- Nicht geeignet für häufige Aktualisierungen
- /+ Seitenrecherche für Benutzer durch eigene, aber ggf. spezialisierte, Suchmaschine nötig.
- Layoutänderungen für mehrere Dateien aufwendig

Eingesetzt wird dieses Konzept zum Beispiel in dem VIP Content Manager von Gauss Interprise oder von SchemaText.

## 2. Dynamisches Publishing

Beim dynamischen Publishing werden angeforderte Seiten direkt aus der Datenbank publiziert. Abbildung 2 schematisiert dieses Konzept: Beim Aufruf einer Webseite über den Webbrowser des Benutzers gibt der Webserver den Request weiter an das WCMS, welches die Seiteninhalte aus einer Datenbank anfordert, diese mit dem Layout verknüpft und dann die HTML-Ausgabe an den Webserver und damit den Benutzer zurückgibt.

Die verwendete Datenbank des WCMS' kann in diesem Verfahren auch die eines anderen Herstellers, wie Oracle, Informix oder sogar mySQL sein.

Da die eigentlichen HTML-Seiten erst bei der Anfrage durch einen Benutzer erstellt werden, spricht man hier von dynamischen Seiten. Man erkennt diese leicht daran, daß die URL zu diesen Seiten nicht fest ist, sondern meist eine kryptische Gestalt besitzt und Parameter betreffend des Aufrufes enthält. Ein Verzeichnis mit statischen HTML-Dateien existiert nicht mehr.

Die Vor- und Nachteile des dynamischen Publishing sind:

- + Leichte Aktualisierungen
- + Einfache Verwaltung der Seiten
- + Flexible Nutzung der Datenbanken von Drittherstellern
- + Seitenrecherche eingebaut
- Für ausreichend Performance sind hohe Rechner-Ressourcen nötig
- Nachteile bei der Erfassung durch Suchmaschinen: Viele Suchmaschinen erfassen keine dynamischen Seiten, so daß diese bei einer Suche nicht gefunden werden.

Dieses Konzept wird unter anderem von SIXCMS verwendet. Der Einsatz kann dabei zum Beispiel auf der Site des Internet-Magazins „Internet World“ betrachtet werden.

### 3. Mischformen

Eine Erweiterung des dynamischen Publishing stellt das **Caching** dar. Hier wird dem Fakt, das einige Seiten stärker frequentiert werden als andere, Rechnung getragen. Um lange Ladezeiten für die jeweilige Neuerstellung von Seiten aus der Datenbank zu vermeiden, werden solche Seiten, die häufig angefordert werden entweder statisch an festen Positionen abgelegt, oder aber in ihrer fertigen Form als HTML-Seite in einen Cache-Bereich gespeichert. Der Cache ist dabei entweder ein Speicherbereich auf dem Webserver oder aber ein temporärer Bereich auf der Festplatte des Servers.

Die Frage, welche Seiten gecached werden, kann automatisch über die Anzahl der Zugriffe, als auch über die Administration durch Redakteure beantwortet werden.

Um Aktualität zu garantieren muß der Cache automatisch aktualisiert werden, wenn im WCMS sich Inhalte der betreffenden Seiten geändert haben.

Aufgrund seiner Struktur ist das Caching eine Mischung der Konzepte des dynamischen Publishing und des Publishing-/Staging-Servers.

### 4. PSE-Konzept

Ein weiteres Konzept, welches im Januar durch die Diplomarbeit des Autors erstmals vorgestellt wurde, ist das **Publishing-, Staging- und Extract-Konzept**, kurz PSE-Konzept.

Das Konzept baut auf dem des Publishing-/Staging-Servers auf und erstellt auf dem Webserver statische Webseiten. Neu ist jedoch die Art der Speicherung und die Verarbeitung und Verwaltung neuer Seiten.

Abbildung 3 schematisiert das Konzept. In dem Konzept sind die HTML- und Objektdateien, welche von WCMS bearbeitet wurden, für den Webserver voll les- **und schreibbar**. Dies bedeutet nichts anders, als daß das WCMS die vollständige Kontrolle über alle Inhalte aufgibt und es erlaubt, das andere Applikationen darauf Zugriff nehmen dürfen. Einzige Prämisse dabei ist, daß die Dateien statisch auf dem Webserver liegen.

Für den Benutzer, der eine Seite abrufen, verhält es sich daher wie ein normales statisches System.

Ebenso verhält es sich, wenn ein vorgegebenes Layout verwendet wird. Die Inhalte der Datei werden gespeichert. Die Inhalte der Datei (Bilder, Grafiken, etc.) werden in dieser HTML-Datei gespeichert. Einzige Ausnahme: Meta-Angaben wie „Titel“, „Description“, „Keywords“, „Version“, usw. werden in eine Meta-Datenbank geschrieben.

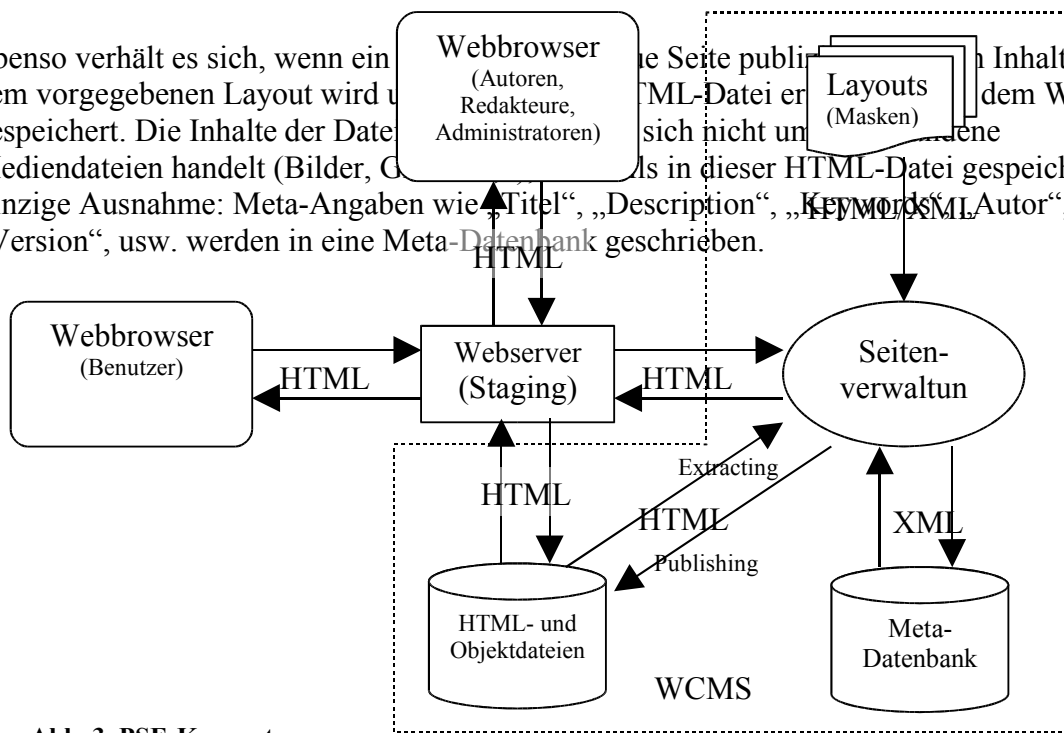


Abb. 3. PSE-Konzept

Soll nun eine HTML-Datei verändert werden, wird das WCMS sie vom Webserver laden. Danach wird das Layout, mit dessen Hilfe die Datei erstellt wurde, genutzt um alle Layout-Angaben in der HTML-Datei zu löschen. Es handelt sich also um die Inversion des Prozesses der Seitenerstellung aus der Addition von Inhalt und Layout: Die HTML-Seite, abzüglich dem HTML-Layout ergibt den Inhalt.

Nach diesem Schritt bleiben nur noch die Inhalte übrig und können somit direkt geändert werden.

Die Meta-Datenbank ist nötig um Daten, die nicht in der HTML-Datei gespeichert werden, zu sichern. Darunter auch Angaben zum verwendeten Layout, Kontroll- und Zugriffsdaten. Die in der Abbildung zusätzlich angegebenen Layouts können ebenfalls als HTML-Dateien abgespeichert und somit verwaltet und geändert werden. Basiert das Konzept auf serverseitige Interfaces ist es möglich eine Selbstadministration zu erreichen, indem alle Layout- und Kontrollausgaben zur Verwaltung durch Redakteure und Autoren ebenfalls Teil der änderbaren Seiten sind.

Vor- und Nachteile dieses Konzeptes:

- + Gute Performance
- + Unkomplizierte Backups
- + Einfache Aktualisierungen
- + Erlaubt die gleichzeitige Verwendung anderer WCM-Lösungen, die nach dem Pushing-/Staging-Konzept arbeiten
- + Erlaubt Autoren mittels eigener Editoren direkt in HTML-Dateien zu arbeiten
- + Import bestehender Seiten einfach
- Benötigt Abbildung des File-Systems innerhalb der Meta-Datenbank
- Wenig Praxiserfahrungen, da neu
- /+ Seitenrecherche für Benutzer durch eigene, aber ggf. spezialisierte, Suchmaschine nötig.
- Layoutänderungen für mehrere Dateien aufwendig

Das Konzept wurde im OpenSource-WCMS Altjira implementiert, welches auf Sourceforge gefunden werden kann (<http://altjira.sourceforge.net>)

### **Anmerkung**

Dieser Artikel ist ein überarbeiteter Auszug aus der Diplomarbeit „Konzeption und Realisierung eines Web-Content-Management-Systems“ von Wolfgang Wiese. Diese Arbeit kann über die Diplomarbeitenbörse [Diplom.de](http://Diplom.de) käuflich erworben werden.