Recherchebericht

- METLAB News Clipping -

Inhaltsverzeichnis

[Allgemeines 3](#_Toc509108079)

[Einsatzumfeld 3](#_Toc509108080)

[Grundsätzliche Bedienlogik 3](#_Toc509108081)

[Mögliche Leitungsmerkmale 3](#_Toc509108082)

[Terminologie 4](#_Toc509108083)

[Übersicht über themenrelevante Applikationen 5](#_Toc509108084)

[PMG Presse-Monitor 5](#_Toc509108085)

[Argus Data Insights 5](#_Toc509108086)

[Universal Information Services - Press Clippings 5](#_Toc509108087)

[Übersicht über IT-Spezifika 6](#_Toc509108088)

[Grundlegende Architektur 6](#_Toc509108089)

[Verfügbare Technologien 6](#_Toc509108090)

[PHP 6](#_Toc509108091)

[JavaScript 6](#_Toc509108092)

[ASP.NET 6](#_Toc509108093)

[Java (EE) Webapplikationen 7](#_Toc509108094)

[RDBMS 7](#_Toc509108095)

[Dokumentenorientierte Datenbank 7](#_Toc509108096)

[Volltextsuche 7](#_Toc509108097)

[Kompetenzen innerhalb der Projektgruppe 7](#_Toc509108098)

[Beschreibung einer konkreten am Markt vorhandenen Applikation 8](#_Toc509108099)

[Leistungsmerkmale 8](#_Toc509108100)

# Allgemeines

Der Kunde fordert einen Webservice, der nach benutzerdefinierten Vorgaben Presse-Clippings erstellt. Der Kunde betreibt den Service und bietet ihn seinen Klienten zur Benutzung an. Bei den Klienten des Kunden handelt es sich um Unternehmen oder andere Organisationen, die den Dienst wiederum ausgewählten Mitarbeitern zur Verfügung stellen.

## Einsatzumfeld

Eingesetzt wird die Software in verschiedenartigen Organisationen. Dazu gehören Unternehmen unterschiedlicher Größe, Vereine und Nichtregierungsorganisationen. Die Software läuft als Webservice auf einem Server, den der Kunde selbst betreibt oder mietet. Bedient wird die Software über einen beliebigen Browser, ohne die Installation eines Plugins.

## Grundsätzliche Bedienlogik

Der Kunde verwaltet seine Klienten und die Klienten verwalten die Endnutzer in der Software. Mitarbeiter eines Klienten können sich über das Web-Interface der Software registrieren und werden dann vom Klienten verifiziert. Danach kann sich der Mitarbeiter anmelden und ein Benutzerprofil erstellen, das er auf Wunsch mit anderen Benutzern teilen kann. Basierend auf dem Benutzerprofil wird in regelmäßigen zeitlichen Abständen ein Presse-Clipping erstellt und dem Benutzer zugeschickt. Das Aktualisierungsintervall gibt der Benutzer vor. Die vollständigen Presse-Clippings werden archiviert und können vom Benutzer jederzeit betrachtet werden.

## Mögliche Leitungsmerkmale

1. Anzahl der Klienten: Die Mindestanforderung sieht 3 Klienten vor.
2. Gesamtzahl der Endnutzer / Anzahl der Benutzer pro Klient: die Mindestanforderung sieht 5 Benutzer pro Klient vor.
3. Anzahl und Vielfalt der verfügbaren Quellen: Die Mindestanforderung sieht insgesamt 6 verschiedene Quellen vor: 2 Webseiten von Radio- oder Fernsehsendern, zwei Onlineangebote von Zeitungsverlagen, 2 soziale Medien.
4. Zusätzliche Informationen zu einzelnen Artikeln: Die Mindestanforderung sieht keine zusätzlichen Informationen zu Artikeln vor. Denkbar ist z.B. die Anzahl der Aufrufe eines Artikels.
5. Sicherheitsstandards: Die Mindestanforderung sieht keine konkreten Sicherheitsstandards vor.

## Terminologie

* Kunde: erwirbt und betreibt die Software, um sie als Dienstleistung anzubieten
* Klient (des Kunden): Unternehmen oder andere Organisation, die die Software verwendet
* Benutzer / Endnutzer: Mitarbeiter eines Klienten
* Presse-Clipping: Sammlung von vollständigen Artikeln oder Artikelzusammenfassungen aus bestimmten Quellen und zu bestimmten Themen
* Benutzerprofil: Jeder Endnutzer legt ein Profil an, in dem er Quellen auswählt und relevante Suchbegriffe angibt. Ein solches Profil kann vom Benutzer mit anderen Benutzern geteilt werden.
* Framework: Infrastruktur und Sammlung vorgefertigter Bausteine, die einen Rahmen für die Entwicklung einer Software zur Verfügung stellen
* Webanwendung: Software, die zum größten Teil auf einem Server ausgeführt wird; generiert typischerweise dynamische Inhalte, die auf dem Client in einem Webbrowser dargestellt werden
* XML: **E**xtensible **M**arkup **L**anguage, eine Auszeichnungssprache zur Darstellung hierarchisch strukturierter Daten im Textfomat
* RSS: **R**ich **S**ite **S**ummary bzw. **R**eally **S**imple **S**yndication, eine Erweiterung von XML zur Darstellung von Inhalten einer Webseite (typischerweise Nachrichten) in einer einheitlichen, strukturierten Form

# Übersicht über themenrelevante Applikationen

## PMG Presse-Monitor

* umfangreichste deutschsprachige, tagesaktuelle Pressedatenbank[[1]](#footnote-1)
* Pressespiegel im PDF-, XML- oder HTML-Format online oder per Mail
* Medientrendanalyse
* Webservice, zusätzliche lokal installierte Software verfügbar
* Wird täglich aktualisiert
* Keine Einbindung von sozialen Medien
* Zielt vorrangig auf den deutschsprachigen Raum
* Richtet sich an Unternehmen, Vereine, öffentliche Institutionen und PR-Agenturen

Quelle: <https://www.pressemonitor.de/leistungen/pressespiegel/>

## Argus Data Insights

* Nationale und international Medienbeobachtung
* Tagesaktuell
* Bezieht auch soziale Medien ein und analysiert diese auf Trends
* Bietet Marktvorhersagen
* Onlineportal, auf Tablets und Smartphone optimiert
* Nutzer- und Rechteverwaltung
* Medienanalyse mit Handlungsempfehlungen
* Clippings können kommentiert und geteilt werden

Quelle: <https://www.argusdatainsights.de/de/produkte-und-dienstleistungen/medienspiegel/>

## Universal Information Services - Press Clippings

* Hochgradig individualisierte Lösungen, die speziell für Kunden erstellt werden
* Lokale und nationale (USA) Print-Medien, internationale Online-Medien
* Umfangreiche Beratung
* Interaktives Portal, in dem Clippings gesucht, sortiert, archiviert und mit anderen Mitarbeitern innerhalt einer Organisation geteilt werden können

Quelle: <https://universal-info.com/press-clippings/>

# Übersicht über IT-Spezifika

## Grundlegende Architektur

Bei der Software handelt es sich um einen Webservice, der sich grundlegend in zwei Bestandteile gliedert:

1. Das Front-End, das im Browser des Benutzers ausgeführt wird
2. Das Back-End, das auf einem Anwendungsserver ausgeführt wird

Dabei zwängt sich die Implementierung des *Model-View-Controller* (MVC) Entwurfsmuster auf, um eine saubere funktionale Trennung zwischen Front-End und Back-End zu gewährleisten.

Das Front-End stellt dabei die *View* dar, die ausschließlich die Visualisierung und Benutzerinteraktion enthält.

Das *Model* ist Bestandteil des Back-Ends und verwaltet sämtlich Daten der Anwendung.

Der *Controller* gehört ebenfalls zum Back-End. Er dient als Bindeglied zwischen *View* und *Model*. Benutzeraktionen von der *View* an den *Controller* gesendet, der sie auswertet und darauf reagiert.

Das MVC-Muster ermöglicht eine klare Trennung der Zuständigkeiten und eine schlanke Schnittstelle zwischen Front-End und Back-End. Dadurch können einzelne Komponenten der Software ohne großen Aufwand ausgetauscht werden. So kann zum Beispiel ein Datenbankmanagementsystem, in dem Benutzerdaten verwaltet werden, durch ein anderes ersetzt werden, ohne dass dabei Änderungen am User-Interface nötig werden.

Die eigentliche Geschäftslogik der Software kann wahlweise im *Model* oder im *Controller* implementiert werden, sollte jedoch nicht zwischen diesen beiden Schichten aufgeteilt werden.

## Verfügbare Technologien

### PHP

Die am häufigste serverseitig verwendete Programmiersprache ist PHP. Der PHP-Interpreter wird auf einem Webserver, meist Apache, ausgeführt und generiert HTML-Code, der in einem Webbrowser dargestellt wird. Ein PHP-Skript kann auch direkt in statischen HTML-Code eingebettet werden.

### JavaScript

Das Front-End einer Webapplikation wird häufig in JavaScript programmiert, oft in Verbindung mit einem clientseitigen Framework wie AngularJS.

### ASP.NET

*Active Server Pages .NET* ist ein Framework für .NET basierte Webanwendungen, die unter anderem in C# programmiert werden können. ASP.NET bietet diverse Funktionalitäten an, von GUI-Elementen bis hin zur Datenbankanbindung.

ASP.NET MVC bietet eine Vorlage für die Implementierung des oben genannten MVC-Entwurfmusters. Mit ASP.NET Core steht auch ein separates, quelloffenes Framework mit offizieller Unterstützung für Linux zur Verfügung.

### Java (EE) Webapplikationen

Die Java Platform, Enterprise Edition (JavaEE) spezifiziert eine Softwarearchitektur für Webanwendungen. Java EE setzt einen Anwendungsserver wie WildFly oder TomEE voraus.

Mit sogenannten *Servlets* können dynamische Webanwendungen erstellt werden. Dabei bildet ein Servlet selbst ein *Model* und einen *Controller* ab, die *View* kann z.B. als *Facelet* realisiert werden, das XHTML-Seiten erzeugt. Mit einem zusätzlichen Framework wie Spring MVC kann das MVC-Muster noch strikter implementiert werden. Für die Entwicklung von Java Webanwendungen steht eine Vielzahl unterschiedlichster Frameworks zur Verfügung. Die große Auswahl an verfügbaren Frameworks und Werkzeugen ermöglicht eine hohe Flexibilität bei der Entwicklung einer Java Webanwendung.

### RDBMS

Für die Verwaltung von Klienten und Endnutzern empfiehlt sich ein relationales Datenbanksystem (**R**elational **D**atabase **M**anagement **S**ystem), weil hier der Umfang und die Komplexität einzelner Datensätze vergleichsweise gering ausfallen. Die Beziehungen zwischen Klienten und ihren Mitarbeitern lassen sich leicht in tabellarischer Form darstellen. Um die Brücke zwischen einer objektorientierten Programmiersprache wie Java und einem RDBMS zu schlagen, kann ein *ORM*-System (beispielsweise Hibernate ORM) verwendet werden. Ein solches System zur Objektrelationale Abbildung soll den sogenannten *Object-relational impedance mismatch* abmildern.

### Dokumentenorientierte Datenbank

Zur Verwaltung von Artikeln und daraus zusammengesetzten Clippings ist ein RDBMS wenig geeignet, nicht zuletzt, weil RDBMS Dokumente nur sehr ineffizient indizieren können. Stattdessen kommt hier ein dokumentenorientiertes Datenbanksystem zum Einsatz.

Auch für die Verwaltung der Suchprofile ist eine dokumentenorientierte Datenbank denkbar.

Bei diesen beiden Anwendungen ist ein XML-basiertes Datenbanksystem wie BaseX besonders günstig.

### Volltextsuche

Um Artikel nach bestimmten Begriffen zu dursuchen, ist eine sogenannte Volltextindexierung nötig. Eine verbreitete Programmbibliothek zur Volltextsuche ist Apache Lucene, auf dem wiederum mehrere eigenständige Frameworks basieren. Manche dokumentenorientierten Datenbanksysteme verfügen über eingebaute Funktionen zur (Volltext-) Indexierung.

## Kompetenzen innerhalb der Projektgruppe

Alle Gruppenmitglieder verfügen mindestens über Grundkenntnisse der objektorientierten Programmierung mit Java. Darüber hinaus sind alle Gruppenmitglieder mit der Entwicklung von relationalen Datenbanken vertraut. Konkrete Erfahrungen mit der Entwicklung einer vollständigen Webanwendung sind nicht vorhanden. Die einzelnen Gruppenmitglieder haben sich bereits jeweils in bestimmte Aspekte der Software eingearbeitet und erste Erfahrungen mit konkreten Technologien und Frameworks gesammelt.

Die Kompetenzen der Gruppenmitglieder sprechen für die Entwicklung einer Java Webapplikation.

# Beschreibung einer konkreten am Markt vorhandenen Applikation

Im Folgenden wird ein Konkurrenzprodukt der Echobot Media Technologies GmbH näher beschrieben. Echobot kombiniert Presse-Clippings mit Online- und Social-Media-Monitoring.

## Leistungsmerkmale

* Neben Print- und Online-Medien werden auch soziale Medien miteinbezogen (Facebook, Twitter, YouTube, Instagram, Pinterest und Google+). Dabei kann nach Themen und nach bestimmten Accounts gefiltert werden.

Die Art der verwendeten Quellen kann vom Benutzer eingeschränkt werden.

* Es werden zusätzliche Metadaten zu den Artikeln bereitgestellt, wie zum Beispiel die Zugriffszahlen und die geschätzte Reichweite.
* Machine-Learning-gestützte Algorithmen filtern kontextabhängig firmenrelevante Artikel heraus. Dies ist besonders nützlich bei Firmennamen, die häufig auch in anderen Kontexten auftauchen (wie zum Beispiel „Müller“ oder „Mars“).
* Ein Frühwarnsystem benachrichtigt den Benutzer per E-Mail, wenn sich ein starker Negativtrend abzeichnet.
* Firmeninterne Nachrichten können eingespeist werden.
* Eine zielgerichtete Wettbewerbsbeobachtung ist möglich, einschließlich einer Firmen- und Finanzanalyse.
* Die optische Gestaltung des Clippings kann individualisiert werden.
* Das Presse-Clipping kann direkt auf dem Webinterface gelesen oder in verschiedenen Formaten exportiert werden, darunter PDF, RSS und XML.

In Abbildung 1 werden folgende Use-Cases der Applikation abgebildet:

* Pressespiegel erstellen mit Artikeln, Social Media Posts, etc. die die Firma betreffen
* Aufrufen von archivierten Pressespiegeln
* Wettbewerbsvergleich mit ausgewählten Konkurrenzunternehmen
* Überwachung von Zulieferern und anderen Partnerunternehmen anhand von Mitarbeiterentlassungen, Streiks, Standortschließungen, etc.
* Spezifische Kundensuche wie z.B. "Firmen, die eine neue Lagerhalle bauen"
* Veröffentlichen des Pressespiegels auf der eigenen Website
* Kundenbewertungen anhand von Social Media Posts erhalten
* E-Mail-Benachrichtigung kann zu jeder Medienzusammenfassung eingestellt werden



Abbildung 1: Anwendungsfalldiagramm für Echobot

1. <http://www.clipping-anbieter.de/> [↑](#footnote-ref-1)