

## CyMON - SDMS

### Semantisches Dokumenten Management System



### Warum SDMS ?

Die Informationsgesellschaft basiert auf schneller Verfügbarkeit und einfacher Verwendbarkeit von großen Mengen an Daten und Informationen. Insbesondere mit der Ausbreitung der Web-Infrastruktur und der damit verbundenen Technologien spielen der effiziente Zugang zu korrekten Informationen und ihrer Anwendung eine überragende Rolle für den Erfolg oder Misserfolg der meisten Unternehmungen.

#### Informationsüberlastung

Der weitaus größte Teil heutiger Informationen wird in der Form von textbasierten Dokumenten gehalten und verarbeitet. Sie können unterschiedlichster Natur sein, z.B. liegen sie in unterschiedlichen Sprachen vor, in verschiedenen Dokument- oder Druckformaten. Der Überfluss an Information in der Informationsgesellschaft hat, neben der Grundlage für erfolgreiche Geschäfte, auch neue Herausforderungen hinsichtlich der Verarbeitung dieser Dokumente gebracht. Unter anderem bedeutet dies eine Komplexität und höhere Anforderungen an Ressourcen und Expertise zum Finden der richtigen Information von der weltumspannenden Datenbank an Dokumenten.

#### Relevanz von Information

Aus Marktanalysen lässt sich entnehmen, dass Angestellte einer Firma typischerweise ca. 80% (je nach Art seiner Beschäftigung) seiner Arbeitszeit mit der Suche nach relevanten und nützlichen Informationen verbringt. Fehler, Verzögerungen und letztlich auch Frustration aufgrund eines Mangels an personellen Ressourcen oder Expertise sind typischerweise der Hauptschlüssel für die Realisierung von effizienten und erfolgreichen Geschäftsprozessen.

## Personalisierung

Die Verschiebung von Anbieter orientierten hin zu Nachfrage orientierten Geschäftsmodellen erfordert die Unterstützung von mehreren Sichten auf ein und dasselbe Dokument, entsprechend den individuellen Bedürfnissen, Präferenzen und der Expertise jedes Nutzers oder zumindest der unterschiedlichen Nutzergruppen. Hierdurch werden Werkzeuge für die automatische Generierung von Taxonomien und für die semantische Personalisierung hoch relevante Geschäftsressourcen und ermöglichen erst Wissens intensive Organisationen.

Um den neuen Herausforderungen zu begegnen sind ausgefeilte und intelligente Werkzeuge notwendig, die dem Mitarbeiter helfen oder ihn an manchen Stellen sogar ganz ersetzen. Die Werkzeuge werden ein unbezahlbarer Bestandteil eines jeden Dokumenten Management Prozesses.

SDMS ist fokussiert auf eine integrierte statistische und semantische Analyse natürlich sprachlicher (Natural Language Understanding and Processing – NLU / NLP) Dokumente und realisiert ein präzises, effizientes und flexibles Dokumenten Management System für eine Vielzahl von Anwendungsgebieten.

## Anatomie des SDMS

Das SDMS hat eine drei Schichten Architektur: die SDMS Kernschicht, die SDMS Mantelschicht und die SDMS Applikationsschicht.

Der SDMS Kern realisiert die Basisfunktionen für das Dokumenten Management unter Benutzung von semantischem linguistischem Wissen und statistischen Maschinen-Lern- und Daten-Verarbeitungs-Technologien. Drei Gruppen von Funktionen werden hier unterstützt:

- Klassifikation von Dokumenten, das die automatische Assoziation von existierenden Dokumenten zu einigen vordefinierten hierarchischen ontologischen Kategorien erlaubt.
- Clustering von Informationen, das hauptsächlich die Gruppierung von Informationen oder Dokumenten in Cluster unterstützt, basierend auf semantischen Gemeinsamkeiten der Dokumentinhalte. Eine solche Gruppierung unter anderem zur Unterstützung der dynamischen Erzeugung von Ontologien/Kategorien und Taxonomien für Dokumente oder Dokumentsammlungen genutzt werden.
- Informationsauszug oder Zusammenfassungen, hierbei werden automatisch Auszüge oder Zusammenfassungen der wichtigen Informationen eines Dokumentes erzeugt, was die effiziente Erfassung von Dokumentinhalten auf der konzeptionellen Ebenen ermöglicht.

Der SDMS Mantel bietet Schnittstellen für die Konfiguration, die Kombination und die Verwendung der Basis Management Funktionen in verschiedenen Kontexten. Eine graphische Schnittstelle und Java Programmierschnittstellen spielen eine wichtige Rolle für die Integration und Anwendung des SDMS in heterogenen Umgebungen. Das SDMS kann als Desktop Anwendung oder als Server Anwendung, in der es über zahlreiche Protokolle (http, RMI, etc.) angesprochen werden kann, eingesetzt werden.



SDMS Anwendungen können basierend auf dem SDMS Mantel implementiert werden. Einige typische Anwendungsfelder in diesem Kontext sind:

- Auffinden von Dokumenten, wobei der Ergebnis des Clustering und der Klassifikation benutzt werden kann um die Dokumentinformationen entsprechen ontologischen Konzepten zu organisieren. Das erlaubt dem Benutzer zu Blättern und relevante Informationen anhand von semantischen Konzepten zu finden. Clustering und Klassifikation können auch benutzt werden um Modelle der Dokumente in einer spezifischen Kategorie zu kreieren. So ein Modell kann in einer erweiterten Suchanfrage nach Informationen in einer größeren Datenbank, z.B. dem Web, genutzt werden.
- Lernen von Profilen, wobei die Kategorisierung vorgenommen wird, um inhaltliche Kategorien zu bilden über die Dokumente, die der Benutzer häufig benötigt oder schreibt (siehe Agentscape Cyb/CyMON Agenten Technologie), anhand derer dann die Benutzerpräferenzen gelernt werden. Diese Funktion kann erweitert werden um Produktions- und Veröffentlichungsaktivitäten zur Erstellung von Expertisen über einen Benutzer.
- Antwort Generierung, wobei das Ergebnis der Klassifikation und die Informationsextraktionsfunktion benutzt werden um die automatische Verarbeitung von Dokumenten und somit die initiale Generierung von Antworten zu unterstützen. Z.B. kann die Belastung von Call Centern damit signifikant gesenkt werden indem der Software Agent in der Lage vorab schon einige Antworten für den Benutzer zu generieren und/oder die eingehenden Nachrichten auch nur an den richtigen Platz zur Bearbeitung weiterzuleiten.

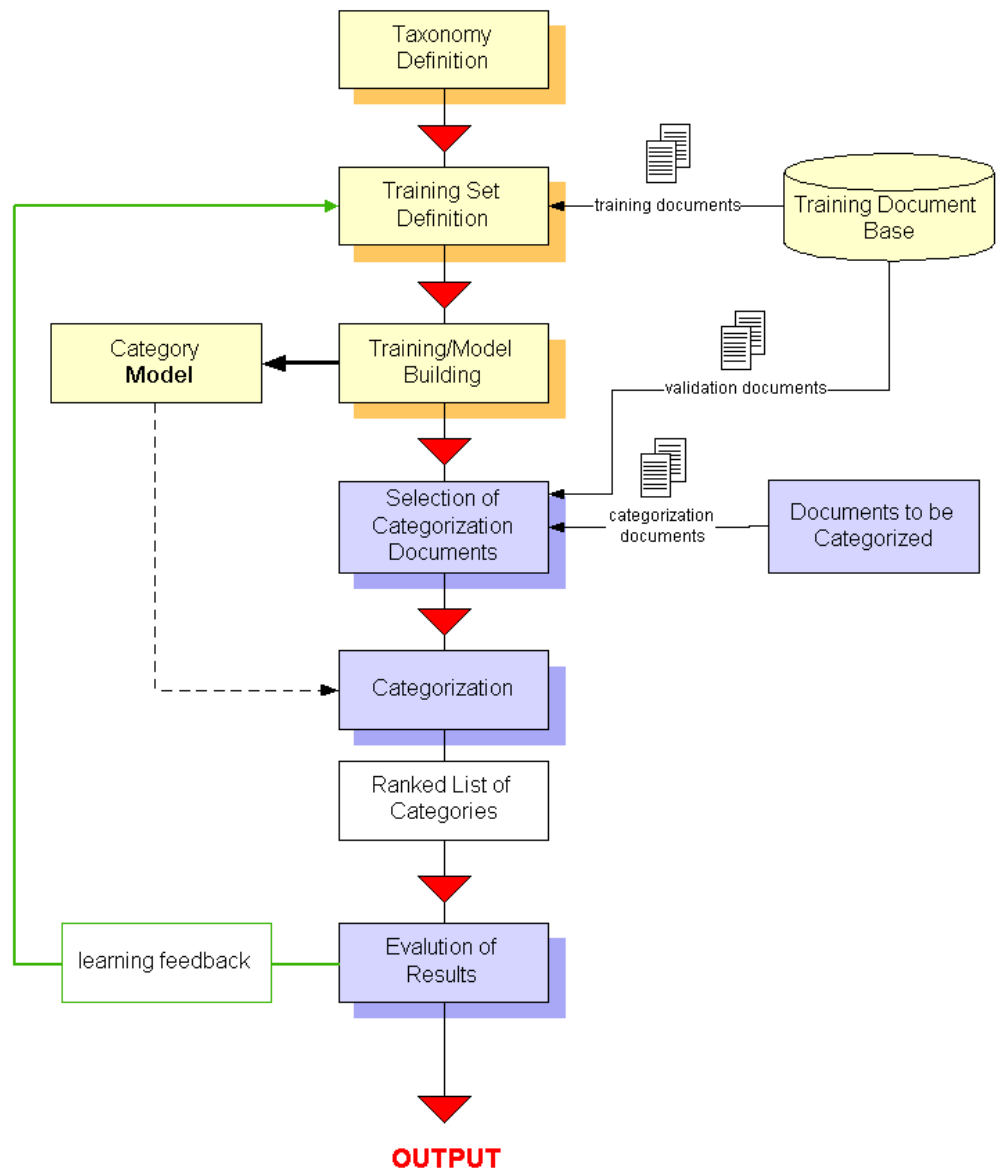


## Kategorisierer Modul

Der Kategorisierer kann als ein Wissensdefinitions- und Indexingwerkzeug angesehen werden. Die Informationen des Kategorisierungsmoduls können aus verschiedenen Typen von Wissen kommen.

Der Kategorisierungsprozess startet mit einer vordefinierten Taxonomie von Kategorien und baut im Trainingsprozess ein Modell für diese Taxonomie auf.

Nach dieser Modellbildungsphase können neue Dokumente kategorisiert werden. Das ist der eigentliche Laufzeitprozess in dem typischerweise mehr als 100.000 Dokumente pro Stunde verarbeitet werden.



*main workflow for the SDMS*



## Wer sollte SDMS benutzen?

Die folgenden Kundengruppen werden den größten Nutzen aus der aktuellen Version des SDMS ziehen:

Firmen, die nach einer intelligenten und flexiblen Lösung für das Wissensmanagement im Unternehmen suchen, wo das SDMS z.B. als Serverlösung eingesetzt werden kann, um Textbasierte Dokumente zu kategorisieren, zu organisieren und zu unterhalten sowie zur Verteilung der relevanten Informationen in einer geeigneten Form (z.B. als Zusammenfassung) zu den geeigneten Mitarbeitern entsprechend der Unternehmenspolitik.

Portalanbieter welche das SDMS zum automatischen Veröffentlichen und Organisieren der Webseiten nutzen können sowie auch zur Unterstützung von intelligenten Blätter- und Suchfunktionen.

Call Center welche das SDMS nutzen können zur Kategorisierung der eingehenden Nachrichten um sie zu dem geeigneten/verantwortlichen Mitarbeiter weiterzuleiten. Außerdem kann das SDMS die automatische Generierung von initialen Antworten basierend auf der semantischen Wissensextraktion, dem Profil und Historie des Benutzers und basierend auf vorhandenem Domänen spezifischem Wissen (z.B. FAQs) übernehmen.

**X**treme  
**P**ersonalisation



### Wie können Sie uns erreichen?

#### Adresse

Agentscape AG  
Bülowstr. 66  
10783 Berlin

#### Kontakt

Tel.: 030-59 00 478-0  
Fax: 030-59 00 478-99

<http://www.agentscape.de>  
[info@agentscape.de](mailto:info@agentscape.de)