



Enterprise Compliance Search





Inhaltsübersicht

Einführung	3
Enterprise Compliance Search Best Practice	4
Suche ist nicht gleich Suche	5
Realisierung einer Compliance Search Architektur	7
Die wesentlichen Leistungsmerkmale	10
Ihr Hauptnutzen auf einen Blick ...	11

Ihr Ansprechpartner:

Dr. Andrea Galli

Phd. Nat. sci. ETHZ

Head of Intelligence Automation

Telefon +41 44 809 25 95

Mobile +41 79 456 76 32

E-Mail andrea.galli@scalaris.com

Einführung

Die Information Management-Infrastruktur eines Unternehmens umfasst in der Regel ganze Silos von Informationen, die bearbeitet und in unterschiedlichen Anwendungen, Aktenarchiven und Datenbanken abgelegt werden. Zu diesen so genannten **operativen Systemen** gehören das Enterprise Resource Planning (ERP), wie beispielsweise das Financial Management, das Customer Relationship Management (CRM), das Product Lifecycle Management und das Supply Chain Management (SCM). Das eigentliche Lebenselixier eines Unternehmens sind jedoch dessen **Kommunikationssysteme und -anwendungen**, wie z.B. E-Mail, Information Management (IM), VoIP, Website, Internet, Kundenkontakt Center und Video Conferencing Anwendungen. Diese Systeme unterstützen die geschäftlichen Arbeitsprozesse und enthalten aktuelle Daten in Hunderten, wenn nicht sogar Tausenden unterschiedlichen Dateiformaten.

Neben dieser Business-orientierten Systemlandschaft erfordern das Compliance und Litigation Management sowie die Vielzahl gesetzlicher Regulatorien heutzutage eine **parallel geführte, GRC-orientierte Systemfamilie**, die den Schutz der gespeicherten Daten sicherstellt. Dazu gehören das Enterprise Content Management (ECM) und Records Management, Archivsysteme und Daten-Repositories, das Policy Management und die Forensic Services (eDiscovery). Häufig existieren dadurch mehrere Kopien derselben oder fast identischer Daten, die in elektronischen Archiven oder auf digitalen Medien (normalerweise Backup-Tapes) gespeichert sind.

Dieser uneinheitliche, auf **zwei vollständig getrennten Systemwelten** basierende Ansatz ist nicht nur kostenintensiv und umständlich, sondern birgt auch zusätzliche Risiken, da Daten kopiert, verschoben und oftmals dauerhaft gespeichert werden. Hinzu kommt, dass jedes dieser Systeme im Rahmen von Anwendungsintegrationen, Unternehmensdurchsuchungen und forensischen Operationen immer wieder eingebunden werden muss. Dies erfordert immer wieder neue, zeitaufwändige Indizierungen und Analysen, die hohe Kosten verursachen. Ein weiterer großer Nachteil dieses Vorgehens ist, dass **Informationen in getrennten Systemen nicht automatisch miteinander verknüpft werden** können.

Bei der Vorbereitung von Rechtsstreitigkeiten fallen in Rechts- und IT-Abteilungen deshalb stets umfangreiche forensische Recherchen an. Insbesondere der Zugriff auf die Vielzahl von Datenquellen und Risikomanagement-Systeme erfordert viel Zeit und Geld. Angesichts der mit dem Information Management verbundenen Herausforderungen ist es nicht erstaunlich, dass die lückenlose Bereitstellung von gerichtsverwertbarem Beweismaterial mehrere Millionen Dollar an Kosten und Hunderte von Mann-Monate manueller Arbeit bedeuten kann. Das Risiko, ein entscheidendes Beweismittel zu übersehen oder bei Rechtsstreitigkeiten nicht alle notwendigen Schritte einzuleiten, kann sogar noch kostspieliger werden und neben gravierenden rechtlichen Auswirkungen auch der Reputation des Unternehmens immensen Schaden zufügen.

Dies führt dazu, dass immer mehr Unternehmen neue Wege gehen: Um die Prozesse für den Zugriff, die Verarbeitung und die Analyse von Unternehmensinformationen zu rationalisieren, **führen sie die beiden Systemwelten zusammen**.

Dies reduziert die Gefahr redundanter Daten ebenso sehr, wie die Kosten für die Beweismittelsicherung, bei der in der Regel Daten aus verschiedensten Systemen zusammengetragen werden. Auf diese Weise entsteht eine **Governance, Risk & Compliance**

Die Zusammenführung der beiden Systemwelten erschließt eine unternehmensweit nutzbare Informationsinfrastruktur, die alle Datenformate transparent indiziert und griffbereit vorhält. Dadurch ist es nicht länger notwendig, jede Datenquelle einzeln einzubinden.



(GRC) Plattform für das Audit Management, das Board/Entity Management, das Matter- & Litigation Management, das Claims Management, Software und Daten für Risikoanalysen, ausgelagerte Dienstleistungen zur Medienüberwachung, Corporate Performance Dashboards usw.

Eine integrierte und strategisch ausgerichtete GRC-Plattform bildet das verlässliche und zukunftsfähige «zentrale Nervensystem», das Unternehmen bei der Minimierung geschäftlicher Risiken maximal unterstützt.

Enterprise Compliance Search Best Practice

Heutzutage sind wir an einem Wendepunkt zwischen einer bürokratischen Industriegesellschaft und einer Gesellschaft des Informationszeitalters. Die Komplexität, der Umfang und die Zeitpläne für zukünftige Aufgaben erfordern neue, leistungsfähigere Tools und Prozesse für das Management von Business Information.

Durch die Computer-basierte Bearbeitung von komplexen Inhalten, wie z.B. Dokumente, Grafiken und Sprachmitteilungen sind Informationen heute weniger strukturiert. Dies ist der genaue Gegensatz zu den tabellarischen, strukturierten Daten, die herkömmlicherweise mit Computern bearbeitet werden. Ohne Frage haben wir ein großes Geschick bei der Bearbeitung dieser strukturierten Daten gezeigt und entsprechende Techniken und Methoden für deren Verarbeitung entwickelt. Dies war die Ära der relationalen Datenbanken, wie z.B. ORACLE und DB2.

Es ist jedoch ein Fehler zu glauben, dass wir dieselben Methoden zur Bearbeitung von unstrukturierten Daten verwenden können - auch wenn wir es immer wieder versuchen.

Anwendungen und IT-Systeme für deren Management, die während der letzten 50 Jahre entwickelt wurden, sind für die Lösung zukünftiger Probleme nicht geeignet.

Mithilfe von Metadaten, Tags und Schlagwörtern versuchen wir, unstrukturierte Inhalte zu strukturieren, damit wir die herkömmlichen Verarbeitungsmethoden einsetzen können. Die Lösung dieses Problems sind **Enterprise Search Anwendungen**, mit denen

- » unstrukturierte Daten direkt bearbeitet und
- » deren inhaltliche Bedeutung extrahiert werden können.

Wenn man bedenkt, dass heute **80 % der Informationen in einem Unternehmen in unstrukturierter Form gespeichert** sind, besteht die Herausforderung bei der Nutzung von unstrukturierten Daten darin,

- 1) sich an ständig ändernde Situationen und Anforderungen anzupassen;
- 2) bessere Technologien für das Management zukünftiger Entwicklungen zu haben;
- 3) flexiblere Datennutzungsprozesse und -infrastrukturen zu entwickeln, die vielfältig eingesetzt werden können.

Suche ist nicht gleich Suche

Ein Unternehmen kann und muss nicht die Suche per se erfolgreich handhaben, sondern die Informationsverarbeitung und die Unterstützung der Mitarbeiter, die diese Informationen zur Erledigung ihrer Aufgaben benötigen.

Beim Einsatz von Suchtechnologien kann die Verarbeitung von Informationen, d.h. Erfassung, Indizierung, Speicherung, Zugriff und Verteilung, durch die Verwendung von modernen **Enterprise Search-Plattformen** effektiv unterstützt werden. In den Compliance-Abteilungen basiert dieser Prozess auf dem **Intelligence & Compliance Management (ICM)-Modell**.

ICM optimiert den Nutzungsprozess von unstrukturierten Daten durch eine fokussierte Unterstützung auf allen Compliance- (strategisch, betrieblich, taktisch) und Unternehmensebenen (z.B. Rechtsabteilungen, Archive, Forensic Services, Anti Money Laundering (AML), Risikomanagement) konsistent, skalierbar und kostengünstig.

Dies basiert auf dem wieder verwendbaren Konzept des ICM-Modells, das z.B. damit verglichen werden kann, dass für eine neue «Tankstelle» auch nicht jedes Mal eine neue «Raffinerie» gebaut werden muss.

Das ICM-Modell



Das ICM-Modell kann folgendermaßen mit der Wertschöpfungskette der Öl-industrie verglichen werden:

Zunächst werden wichtige Informationsquellen systematisch identifiziert und «angezapft» (Beschaffung). Die gefundenen Daten werden anschließend in verwertbare Informationen umgewandelt (Verarbeitung – Aufbereitung/Raffinerie) und effizient verteilt. Das Endprodukt sind personalisierte und transparent präsentierte Inhalte (Nutzung - Tankstelle), die für die Unterstützung der Geschäftsprozesse und -anwendungen des Benutzers konzipiert sind. Zudem können diese Inhalte im Handumdrehen mit ähnlichen Referenzbegriffen über dieselbe «Raffinerie» auch anderen Benutzern und Anwendungen zur Verfügung gestellt werden.



Die ICM-Systematik umfasst die folgenden drei Prozesse:

- Beschaffung:** Identifizieren aller relevanten Quellen durch die Untersuchung jeder potenziell wichtigen Informationsquelle (intern oder extern) und die Bereitstellung aller erforderlichen Verbindungen und Zugriffsrechte.
- Aufbereitung:** Herausfiltern, Indizieren, Modellieren und Verknüpfen der ausgewählten Informationen und deren Bereitstellung für Benutzer und Anwendungen in der benötigten Präsentationsform, Transparenz und Struktur.
- Nutzung:** Liefern von Informationen gemäß den Anwendungsanforderungen und Bereitstellen dieser Informationen für alle autorisierten Benutzer. Erstellen eines Netzwerks von inhaltsorientierten Anwendungen auf der Basis einer einzigen «Raffinerie»-Infrastruktur.

Fast alle internen Vorgänge und Prozesse in einem Unternehmen sind miteinander verbunden. Auf dieselbe Weise sollte die Technologie (unabhängig von den Back-End-Datenquellen) die Kommunikation zwischen Anwendungen ermöglichen.

Zusammen mit seinem Partner Autonomy bietet Scalaris AG den **Intelligent Data Operating Layer (IDOL)-Server** an, mit dem Unternehmen digitale Inhalte automatisch verarbeiten und Anwendungen miteinander kommunizieren können.

Die IDOL-Technologie liefert eine vollständige Software-Infrastrukturlösung, die Inhalte jeder Art (text- oder sprachbasiert, unstrukturiert oder strukturiert), unabhängig vom Speicherort, dem Format oder den dazugehörigen Anwendungen, «verstehen» kann. Auf diese Weise ermöglicht die Technologie **«Integration durch Verstehen» (Meaning Based Computing)**. Dadurch können Anwendungen ohne Metadaten und vorher manuell festgelegte Konnektoren miteinander kommunizieren. In der Folge können viele wichtige Prozesse und Aufgaben, die bisher manuell durchgeführt wurden, automatisiert werden.

Das umfassende Funktionsspektrum von IDOL bietet Unternehmen echte geschäftliche Vorteile, wie z.B. gesteigerte Produktivität und optimierte Informationsspeicherung. Dadurch werden die mit der Datenverwaltung und der Nachbearbeitung von menschlichen Fehlern verbundenen Kosten erheblich reduziert.

Was unterscheidet Meaning Based Computing von herkömmlichen Methoden, wie z.B. die Schlagwörtersuche?

«Meaning Based Computing» geht weit über die herkömmlichen Methoden hinaus, die dem Benutzer nur das simple Suchen und Finden von Daten ermöglichen.

Schlagwort-Suchmaschinen sind beispielsweise nicht in der Lage, die inhaltliche Bedeutung von Informationen zu erfassen. Sie finden lediglich Dokumente, die ein bestimmtes Wort enthalten. Dieses Unvermögen führt leider dazu, dass inhaltsverwandte - und oftmals wichtige - Dokumente, in denen andere Wörter verwendet werden, übersehen werden. Dagegen werden häufig Dokumente mit einer völlig anderen Bedeutung angezeigt, so dass der Benutzer seine Suchanfrage entsprechend den Regeln der Suchmaschine ändern muss.

Bei der Entwicklung des IDOL-Servers wurde die Wichtigkeit von Schlagworttechnologien berücksichtigt und vollständig implementiert. Diese erweiterten Funktionen sind der Schlüssel zu seinem Erfolg. Zudem fehlen bei den Schlagwort-Suchmaschinen einige der Kernfunktionen des Meaning Based Computing, wie beispielsweise die erweiterte Phonetik, Fuzzy, automatisches Hyperlinking und Clustering.

BEISPIEL

Für das automatische Hyperlinking, das dem Benutzer verschiedene Dokumente, Personennamen oder Datensätze anzeigt, die inhaltlich mit dem Originaltext verknüpft sind, ist es erforderlich, dass der Inhalt des Originaldokuments vollständig verstanden wird. Für das automatische Sammeln, Analysieren und Organisieren von Informationen müssen Computer in der Lage sein, die inhaltliche Bedeutung einer Information zu verstehen und zu extrahieren.

Das können nur «Meaning Based Computing»-Systeme!



Realisierung einer Compliance Search-Architektur

Mit dem ICM-Modell kann jede Compliance Search-Roadmap durch die Auswahl der richtigen Komponenten in den Bereichen Beschaffung, Aufbereitung und Nutzung, Schritt für Schritt entwickelt werden - und das ohne Investitionen in große IT-Projekte für eine unternehmensweite Suche im Vorfeld.

Der erste Schritt für die Gestaltung einer Compliance Search-Architektur ist die Definition einer zentralen, skalierbaren, flexiblen und wiederverwendbaren **«Raffinerie» (IDOL)** sowie die Verbindung der **«Ölfelder» (Informationsquellen)** und **«Tankstellen» (aufbereiteter Inhalt)** gemäß dem Roadmap-Zeitplan.



Dies kann mit einer **«Einheitlichen Suche»** für wenige Quellen und Benutzer beginnen und mit einem Anwendungsnetzwerk für die **«Erweiterte Suche»** nach strategisch und betrieblich wichtigeren Domains enden.

Zwei wichtige Schritte sind erforderlich, um die vorhandenen Suchprozesse in vielen Intranets in eine ICM-konforme Architektur umzuwandeln:

Einheitliche Suche

Die meisten Benutzer stellen sich unter einer Suche eine weiße Leiste und Ergebnisse vor. Dies ist auch in vielen Unternehmen der Fall, bei denen einige Intranetanwendungen über eine eingebettete Suche verfügen. Viele andere Anwendungen weisen nicht einmal diese Funktion auf. Momentan können die verschiedenen existierenden Suchtechnologien nicht aufeinander abgestimmt werden. Eine Angleichung verschiedener Technologien auf eine einzige Technologie würde die IT-Betriebs- und -Verwaltungskosten jedoch erheblich senken und den Weg für den nächsten Entwicklungsschritt zu einer «Erweiterten Suche» frei machen.

Erweiterte Suche

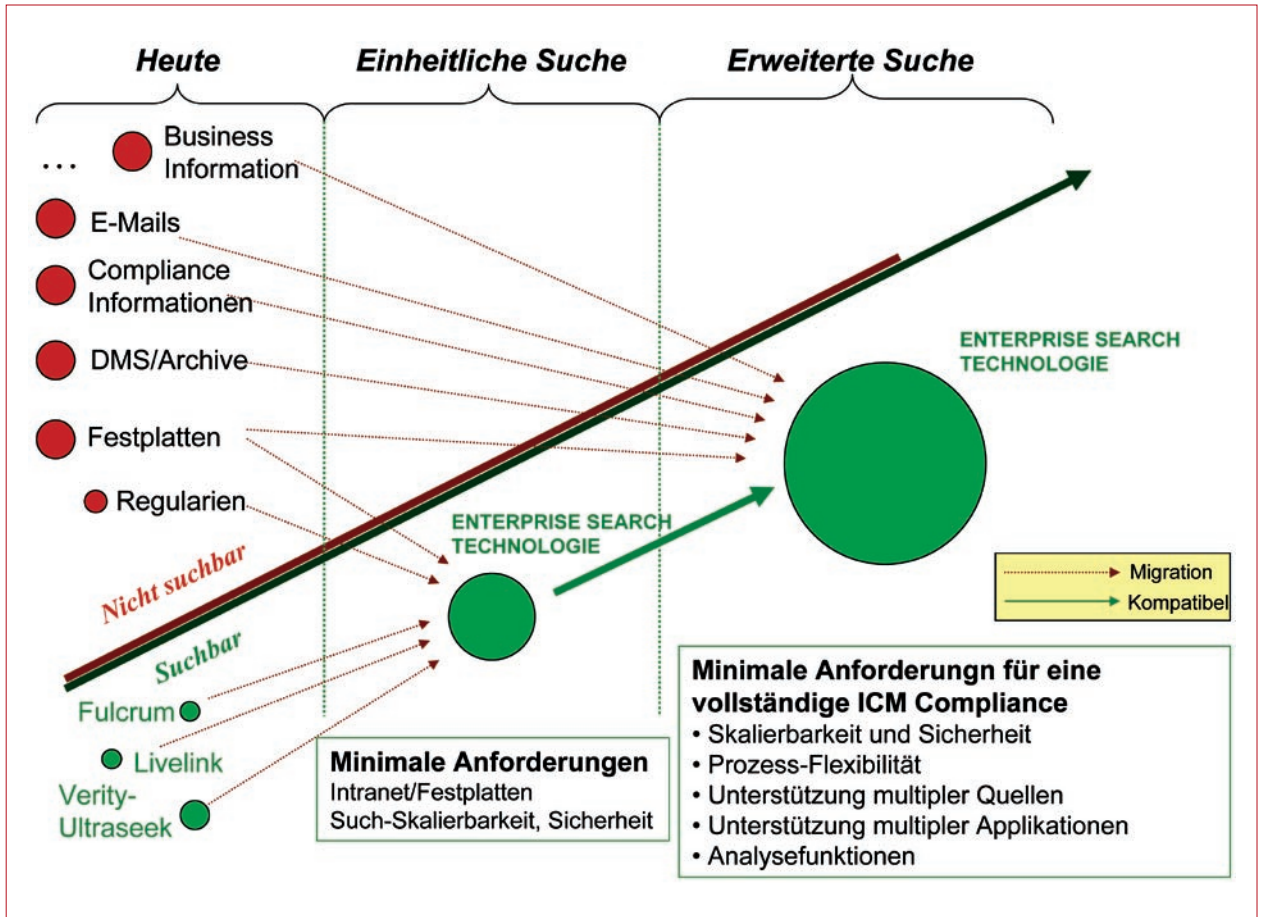
Die Technologie hat sich weiterentwickelt und spielt jetzt eine wichtige Rolle für Legal und Compliance-Abteilungen, Business Information, Marktforschung und viele weitere wichtige strategische und operative Aufgaben. Um komplexe Geschäftsprozesse zu unterstützen, die auf unstrukturierten Informationen basieren, verfügen Anwendungen für eine «Erweiterte Suche» über umfassende analytische Funktionen. Dadurch werden die betriebliche Effizienz gesteigert und die rechtlichen und geschäftlichen Risiken verringert.

Schon jetzt spielt die Einführung einer Infrastruktur für die «Einheitliche Suche» eine Schlüsselrolle bei dem Aufbau eines neuen **Intelligence & Compliance Management (ICM)-Layers**, der Inhalte und Daten miteinander verknüpft. Dieser Layer dient zur Entwicklung neuer Anwendungen für Ursachenanalysen, Forensic-Dokumentationen und inhaltsorientierte Prozesse der «Erweiterten Suche» (unstrukturierte Informationen strukturieren, und sie mit strukturierten Daten, Textanalysen, maschineller Sprachverarbeitung, automatisierter Klassifizierung, Ergebnisvisualisierung usw. verknüpfen).



Die einfache eingebettete Suche der heute existierenden Systeme kann nicht für die komplexe Verarbeitung von unstrukturierten Informationen eingesetzt werden. Mit den umfassenden Funktionen des ICM-Modells können Sie in die Zukunft starten!

Die hoch entwickelte Roadmap für die Intranet Suchmigration



Vollständige ICM-Compliance schrittweise erreichen und zukünftige Herausforderungen bei der Nutzung von unstrukturierten Informationen meistern.



Die wesentlichen Leistungsmerkmale

EINHEITLICHE SUCHE

- | | |
|-------------------------|--|
| Suche | <ul style="list-style-type: none">» Schlagwortsuche» Inhaltsorientierte Suche |
| Such-Konnektoren | <ul style="list-style-type: none">» Indizierung von HTTP- und Daten-Repositories» Unterstützung von mindestens 200 der bekanntesten Datenformate |
| Sicherheit | <ul style="list-style-type: none">» Präzise definierbare Zugriffsrechte» Integration mit dem existierenden Autorisierungssystem (Sphinx) via standardisierte LDAP Protokolle» Datenverschlüsselung |
| Skalierbarkeit | <ul style="list-style-type: none">» Skalierbarkeit für Millionen von Dokumenten» Skalierbarkeit für Tausende von Benutzern» Globale Sprachunabhängigkeit |

ICM-KONFORME, ERWEITERTE SUCHE

- | | |
|-----------------------------------|--|
| Prozessflexibilität | <ul style="list-style-type: none">» Modellierungsfunktion für zahlreiche, miteinander verknüpfbare Indextabellen mit freier Definition der Metadatenfeldern» Unterstützungsfunktionen für Verteilung, Ausfallsicherung und unterschiedliche Hardware |
| Unterstützte Suchquellen | <ul style="list-style-type: none">» Standardschnittstellen für die meisten herkömmlichen kommerziellen RDBM-, DMS-, GroupWare- und CMS-Systeme» Sicherheitsintegration mit diesen Systemen» OEM-Komponente dieser Systeme» Allgemeines Abstract API für kundenspezifische Indexschnittstellen |
| Unterstützte Applikationen | <ul style="list-style-type: none">» Unterstützung von API für unterschiedliche Standard-Entwicklungssprachen (z.B. ASP, ASP+, PHP, JSP, JAVA)» Unterstützung von API für verteilte Anwendungen unter Verwendung von Standard-Kommunikationsprotokollen (TCPIP, HTTP, HTTPS)» Modellierung von Informationsflussprozessen und Verknüpfung von Indizes» Unterstützung der Standard-Suchanfragesyntax, einschließlich boolesche Verknüpfungen, freie Sprachwahl, Fuzzy, musterbasierte Suche und Metadatenuche |

**Erweiterte
Suchfunktionen**

- » Unterstützung einer erweiterten Sprachverarbeitung, Textmining, Suchanfrageanalyse, erweiterte phonetische Fuzzy-Suche
- » Unterstützung von Echtzeit-Visualisierung großer Datensätze mittels Clustering und intelligenter Rauschminderungs-Algorithmen



Ihr Hauptnutzen auf einen Blick ...

- » Nutzung von Informationen in jeder Sprache, zu jedem Thema ohne Zeit raubenden Schulungs- und Verwaltungsaufwand
- » Umgehendes Auffinden von Daten in nahezu allen Formaten - PDF, MS Office, E-Mails etc.
- » Sekundenschnelle Anzeige aller themenbezogenen Dokumente aus dem Netzwerk der angeschlossenen Quellen
- » Automatisierte Recherchen in Dokumentenmanagement-Systemen, Datenbanken, Informationsdiensten, Daten- und Web-Servern
- » Erkennen von Zusammenhängen zwischen Dokumenten aufgrund der Fähigkeit des Systems, den Inhalt der verwalteten Dokumente zu «verstehen»
- » Pünktliche Bereitstellung aller entscheidungsrelevanten Fakten an der richtigen Stelle
- » Extraktion der relevanten Fakten für schnellere und bessere Entscheidungsprozesse
- » Zeitgewinn für die effektive Denkarbeit und wichtige Kernaufgaben
- » Kosteneinsparungen bei der Beschaffung und gebrauchsfertigen Aufbereitung relevanter Business Information

Die Scalaris AG auf einen Blick

Scalaris AG, mit Sitz in der Schweiz und Niederlassungen in Villingen-Schwenningen und Glattbrugg, bietet Beratung, Lösungen und Dienstleistungen für ein unternehmensweites Management von Information, Wissen und Geschäftsprozessen an.

Mit unserer **Intelligent Document Processing (IDP)** Gesamtlösung adressieren wir Unternehmen und Organisationen aus allen Branchen, für die eine schlanke, effiziente und wirtschaftliche Informationslogistik von entscheidender Bedeutung ist. Unsere IDP Lösungen automatisieren die Bearbeitungsprozesse an Dokumenten ebenso umfassend, wie das Management Dokumenten- und Akten-basierter Geschäftsvorgänge. Damit erschließen wir unseren Kunden ein durchgängig optimiertes Dokumenten Lifecycle Management aus einem Guss - vom Input der Dokumente bis zu deren revisionssicherer Archivierung:

- » **IMA:** Input Management Automation
- » **DMS:** Document Management Solutions
- » **IAR:** Intelligent Archiving & Retrieval
- » **IAS:** Intelligence Automation Solutions

Was Sie von uns erwarten dürfen

Unser Leistungsangebot verbindet bewährte Technik mit hohen Service- und Qualitäts-Standards. Im Mittelpunkt stehen Lösungen und Software as a Service (SaaS) Angebote, die unseren Kunden auch bei komplexen Aufgabenstellungen schnell den erwarteten Nutzen bringen. Die offene Architektur unseres IDP Portfolios vereinfacht sowohl die Integration mit bestehenden IT-Infrastrukturen als auch mit Drittsystemen. Dies sichert unseren Kunden einen langfristigen Schutz der getätigten Investitionen.

Unsere Referenzen

Zu unserem Kundenkreis gehören namhafte Unternehmen und Organisationen aus den Bereichen Bank und Finanzdienstleistung, Versicherung, Transport- und Logistik, Pharma, Biotech, Dienstleistung, Versandhandel und Government.



Möchten Sie mehr erfahren? - Wir sind gerne für Sie da.

In der Schweiz

Scalaris AG
Feldeggstrasse 5
8152 Glattbrugg

Telefon +41 44 809 25 00
Telefax +41 44 809 25 03
E-Mail info@scalaris.ch

In Deutschland

Scalaris AG
Roggenbachstraße 1
78050 Villingen-Schwenningen

Telefon +49 7721 9175-0
Telefax +49 7721 9175-800
E-Mail info@scalaris.com