



# DOK

Technologien, Strategien & Services für das digitale Dokument

**Lean Management im Umgang mit Wissen**

**Web 2.0 trifft E-Government**

**Social Media & Content Management**

## Technologien des Information Mining: Business Intelligence, Semantik & Co.

Special: Output-Management



# Best Practices, um aus unstrukturierten Daten den größtmöglichen Nutzen zu ziehen

Informationsmanagement, unstrukturierte Daten, Geschäftsdaten, Business Intelligence, Daten-Policies, Best Practices

Informationsmanagement ist angesichts der Datenmengen eine der größten Herausforderungen im Unternehmen. Die meisten Informationen in Unternehmen existieren in Form von unstrukturierten Daten: Word- oder PDF-Dateien, E-Mails, Webseiten oder PowerPoint-Präsentationen. Terabytes dieser unstrukturierten Geschäftsdaten, darunter zahlreiche Kopien, verschiedene Versionen oder identische Dokumente sind typischerweise auf unterschiedliche Speicherorten verstreut. Diese Artefakte auf lokalen FileShares oder Servern stellen ein Datenchaos dar, das es Unternehmen unmöglich macht, die eigentlichen Informationen ernsthaft zu nutzen. Deshalb benötigen Unternehmen Best Practices, um den Wert ihrer unstrukturierten Daten überhaupt erkennen und nutzen zu können. Ohne eine derartige Methodik und die dazugehörigen Lösungen ist es unmöglich, Lösungen für Business Intelligence (BI), Content Management, IT-Sicherheit oder Compliance überhaupt in Betracht zu ziehen. Denn es ist schlicht und einfach nicht möglich zu erkennen, wo wichtige Unternehmensdaten aufbewahrt werden, die analysiert, verwaltet und weiterverarbeitet werden sollen.

[www.nogacom.eu](http://www.nogacom.eu)

**Thomas Quednau** ist CEO der **Nogacom**. Das Unternehmen mit Sitz in Israel, Deutschland und der Schweiz bietet Software und Dienstleistungen für Information Management und Governance, mit denen sich unstrukturierte Daten handhaben lassen, unter anderem mit einer eigenen Plattform für die Klassifikation von Daten. Nogacom beschäftigt unternehmensweit rund 30 Mitarbeiter.

## Eine bewährte Methode, um Geschäftsdaten zu verstehen

Wie kann nun eine vernünftige Methodik aussehen? Um Geschäftsdaten über sämtliche Geschäftsfelder hinweg zu verstehen und zu verwalten, sollte die Methodik aus vier Phasen bestehen: Bewerten, Definieren der Policies und Parameter, Durchsetzen und Überprüfung und schließlich als letzte Phase das Messen des Erfolgs.

### Phase 1: Bewertung

Zuallererst muss sich ein Unternehmen einen vollständigen Überblick über den aktuellen Stand der Daten verschaffen. Dazu

BEI UNS HAT DIE  
**AKTE X**  
EINEN NAMEN:

DIE  
DIGITALE AKTE  
**CENIT  
ECLISO**

gehört das Wissen, wie diese Daten genutzt werden, wie darauf zugegriffen wird, wie sie gespeichert und verteilt werden. Ein vollständiger Überblick beinhaltet auch die Beantwortung der folgenden Fragen:

- Wo liegen die Daten und warum? Warum zum Beispiel speichern Mitarbeiter Daten auf ihrem Desktop, obwohl sie angehalten werden, diese zentral abzulegen?
- Wer fertigt Kopien und Versionen von Dokumenten an und warum? Warum legt jemand in der Finanzabteilung verschiedene Versionen der Kundenverträge an und speichert diese alle im Ordner der Verkaufsabteilung?
- An wen werden Dokumente gemailt und warum? Warum beispielsweise umging jemand das Unternehmens-VPN und sendete Dokumente an private E-Mail-Konten, obwohl dies eine Verletzung der firmeninternen DLP-Policy (Data Loss Prevention) bedeutet?
- Wer hat Zugriff auf sensible Unternehmensdateien und warum? Warum zum Beispiel können Mitarbeiter auf sensible Daten zugreifen, obwohl sie mittlerweile mit anderen Aufgaben betraut sind?

Um sich einen vollständigen Überblick zu verschaffen, müssen sämtliche unstrukturierten Daten zuerst auf ihren geschäftlichen Kontext hin identifiziert und klassifiziert werden. Ganz gleich, wie eine Datei heißt, wo sie gespeichert ist und in welchem Format sie vorliegt. Dieser „geschäftliche Kontext“ bezieht sich auf die Sprache, in der ein Unternehmen sich selbst beschreibt. Dazu gehören beispielsweise Kunden, Produkte, Abteilungen, Angestellte und Lieferanten. Der geschäftliche Kontext eines Dokumentes ist von entscheidender Bedeutung. Damit können tabellarische Übersichten, welche Zeiten und Preise eines Zugfahrplanes enthalten, automatisch von einem Finanzbericht unterschieden werden – beide von der Assistenz des CFOs angelegt.

- NACHHALTIGE REDUKTION  
DER BEARBEITUNGSZEITEN
- VIELFÄLTIGER EINSATZ:  
KUNDENAKTE, PROJEKTAKTE,  
STEUERAKTE U.V.M.

## Phase 2: Definition der Policies und Kontrollen

Sobald eine granulare, kontextbasierte Übersicht zum aktuellen Stand der Inhalte vorliegt, können Unternehmen die geeigneten Regularien definieren, geeignete Kontrollen implementieren und Best Practices definieren, die es den Nutzern ermöglichen, den maximalen Geschäftswert aus den Daten zu ziehen. Gleichzeitig müssen die Anwender sicherstellen, dass die Anforderungen von Content Management, der IT Security und Compliance kostengünstig umgesetzt werden. So kann ein Unternehmen beispielsweise eine Policy festlegen, die verlangt, sämtliche Dokumente wie E-Mails, Korrespondenz und Kalkulationen, die sich auf die wichtigsten 1.000 Kunden beziehen, automatisch in einen speziellen Ordner zu verschieben, um für weiter gehende BI-Analysen zur Verfügung zu stehen. Ein weiteres Beispiel wäre eine Policy, die es verbietet, sensible Inhalte über Web-basierte E-Mail-Konten zu versenden. Um diese Policies definieren zu können, muss ein klares Verständnis darüber bestehen, in welchem Zustand sich die Daten befinden und wie sie im Unternehmen genutzt werden. Und natürlich darüber, wie sich diese Situation verbessern lässt.

## 3. Durchsetzung und Überprüfung

Nachdem die Policies und Kontrollen festgelegt worden sind, müssen sie im nächsten Schritt durchgesetzt und deren Einhaltung überwacht werden. Idealerweise sollte dazu unterstützend eine Technologie implementiert sein, die solche Arbeitsabläufe automatisiert durchführt und in der Lage ist, bestimmte Arten von Inhalten (beispielsweise Dateien, die nach einem Datum X zu Produkt Y angelegt wurden) zu erkennen und daraufhin die weiteren notwendigen Schritte automatisch auf die Dokumente anzuwenden. Mit „Vorgängen“ ist beispielsweise das Verschieben oder Löschen von Dokumenten oder das Ändern der Zugriffsrechte gemeint. Solche Technologien müssen beispielsweise die Kreditkartennummern der Kunden erkennen, sobald diese in Word-Dateien, E-Mails oder anderen unstrukturierten Daten vorhanden sind. Schließlich müssen Zugriff und Verteilungsrechte für sämtliche dieser Daten automatisch begrenzt sein – ein entscheidender Schutz zur Einhaltung regulatorischer Pflichten beim Datenschutz sensibler Kundendaten. Ebenso wichtig ist es, Anomalien zu erkennen. Beispielsweise, wenn Kopien bestimmter Dokumente auf dem Server an einem

Ort auftreten, auf dem sie bisher noch nicht abgespeichert wurden. In diesem Fall müssen Mechanismen greifen, die autorisierte Content Manager darauf aufmerksam machen. Diese können dann die Anomalie untersuchen und festlegen, ob es sich um ein neu auftretendes Problem oder nur um das Ergebnis einer neuen, legitimen Nutzung der Informationen handelt.

## 4. Messung des Erfolgs

Die beschriebene Methodik wird schließlich durch die Messung komplettiert, wie effektiv die Policies und Kontrollen sind und ob die Anwender die Policies einhalten oder nicht. All das wird durch detaillierte Reports erreicht, die die Datennutzung vor und nach der Einführung der Policies und Best Practices analysieren. Besonderes Augenmerk liegt auf der Nachvollziehbarkeit, wohin die Kopien und Versionen der Dokumente gelangen – eine der häufigsten Schwachstellen in Unternehmen. Auch die ungewöhnliche Nutzung oder Verteilung der Daten spielt eine Rolle, um bei Bedarf die Policies anzupassen und Anwender zu schulen. Alternativ können Unternehmen vorsätzlich bestimmte Risiken in Kauf nehmen, um Vorteile zu erzielen. Ein Beispiel dafür wäre, Verkäufern den Zugriff auf eine breitere Palette an Dokumenten zu gewähren, um ihnen zusätzlichen Einblick für noch mehr Geschäftsabschlüsse zu ermöglichen. Darüber hinaus ist es wichtig zu verstehen, dass sich die Umgebung der Daten und die Unternehmen selbst ständig wandeln. Unternehmen müssen also ständig neu festlegen, wie die Daten zu nutzen sind, ein stetiger Kreislauf. Dementsprechend sind die Policies anzupassen, um Änderungen wie M&A, neue gesetzliche Bestimmungen und veränderte geschäftliche Abläufe durchzuführen.

## Reale Auswirkungen auf das Geschäft und den informationsbasierten Markt

Die umfassende, iterative Methode erlaubt es Unternehmen, kontinuierlich die Nutzung ihrer Informationen zu optimieren, die Einhaltung regulatorischer und firmeninterner Vorgaben durchzusetzen und gleichzeitig Kosten für Storage-Architektur und Verwaltung zu reduzieren. Die genannten Vorteile kommen bereits heute zum Tragen und werden künftig noch wichtiger, da Unternehmen jeden Tag mehr und mehr unstrukturierte Daten erzeugen.

Bei der praktischen Umsetzung sollte zum einen zwischen Best Practices unterschieden werden, die geeignet sind, im unstrukturierten Datenchaos Ordnung zu schaffen, und zum anderen Einzelanwendungen oder Tools, die Unternehmen nutzen, um Business Intelligence einzuführen bzw. Datenschutzrichtlinien durchzusetzen. Einzelanwendungen wirken sich schlicht und einfach nur auf einzelne Bereiche der unstrukturierten Daten sinnvoll aus. Sie befassen sich jedoch nicht mit der grundlegenden Frage, über welche Informationen, über welches Wissen das Unternehmen wirklich verfügt. Deshalb ist der Einsatz von Einzelanwendungen nur begrenzt sinnvoll, solange diese Anwendungen nicht durch eine praktikable Methode ergänzt werden, die auf einer Informa-

tion-Management-Lösung basiert, die sämtliche Vermögenswerte eines Unternehmens „versteht“.

Jedes Unternehmen sollte sich deshalb eine performante Methodik aneignen, um seine unstrukturierten Daten zu verstehen und zu verwalten. Eine solche Methode hilft, die Business Performance in einem informationsgetriebenen Markt zu steigern, ganz im Sinne des Grundsatzes, dass Informationen und Wissen unsere wertvollsten Rohstoffe sind. ■



Bibliothek des Klosters Strahov, Prag (Tschschchien) · www.kraas-lachmann.com

## Als wär man da.

Ihre Nutzer wollen noch vor dem Aufstehen auf die Inhalte Ihrer wertvollen Originalausgaben zugreifen? Kein Problem! Wir beherrschen mit unseren Digital- und Analogsystemen alle Prozesse der Dokumenten-Erfassung, -Archivierung, -Verarbeitung und -Bereitstellung. Seit mehr als 40 Jahren.

**Zeutschel, die Zukunft der Vergangenheit.**



Zeutschel GmbH · Heerweg 2 · 72070 Tübingen · Tel.: +49 7071 9706-0  
Fax: +49 7071 9706-44 · info@zeutschel.de · [www.zeutschel.de](http://www.zeutschel.de)