



Vernetzen schafft Nutzen

Wie können Unternehmen ihre exponentiell wachsenden Daten nachhaltig pflegen und nutzen? Die Antwort liegt in der Vernetzung der Daten.

→ VON CHRISTIAN HEUER

Langsam folgten die Anforderungen an die Datenhaltung den Business-Prozessen. Daraus entwickelten sich Datensilos, die sich nur langsam wieder aufbrechen lassen. Verändert sich das Unternehmen durch Akquisitionen, Fusionen oder Verkäufe macht das den Prozess noch schwieriger. Daten zu harmonisieren oder zentral zu halten, gleicht einer Sisyphusarbeit.

Führende Marktforschungsinstitute sind sich einig, dass sich die Datenmenge in den nächsten Jahren verfünffachen wird. Dabei ist es schon heute schwierig, Daten aus verschiedenen Quellen zu nutzen, um Geschäftsprozesse zu unterstützen oder regulatorischen Anforderungen zu genügen. Hinzu kommt, dass nicht nur Daten aus dem Unternehmen zu integrieren sind, sondern auch aus dem erweiterten Ökosystem, z. B. von Kunden oder Zulieferern.

Es stellt sich also die Frage: Wie ist es möglich, grosse Datenvolumen intelligent zu verwalten und bereitzustellen, um damit die digitale Transformation voranzutreiben?

EIN NETZ FÜR ALLES

Google, Facebook und LinkedIn haben es vorgemacht. Alle Datenobjekte werden dort in sogenannten «Graphen» vernetzt: Wo bisher heterogene Datendomänen in unterschiedlichen Datenbanken lagen, wird nun alles miteinander verbunden. Die Basis dafür ist das Semantische Web, manchmal auch Web 3.0 genannt.

Die Grundidee besteht darin, Daten und Inhalte unabhängig von der Darstellung in offene und maschinenlesbare Formate zu bringen. Durch Ontologien, d. h. durch strukturierte Wissensabbildungen, können homogene Modelle geschaffen werden, die globale Standardisierung und vereinfachten Datenaustausch ermöglichen. Google setzt beispielsweise auf die Ontologie «schema.org», die zur Aufbereitung von Inhalten in der Suchergebnisseite dient (sogenannte Rich Snippets). Ein wichtiger Punkt: Die Vernetzung legt neue, bisher nicht sichtbare Zusammenhänge offen, die sich für Vorhersagen, also prädiktive Analysen, nutzen lassen.

UNKOMPLIZIERT, ABER INTELLIGENT

Das Semantische Web bietet eine Reihe von Vorteilen gegenüber traditionellen Technologien wie BI, MDM, SOA oder Federated Search. Einer davon ist der «uneindringliche» Charakter, d. h., Anwendungen lassen sich öffnen, ohne APIs oder besondere Middleware zu benötigen. Ein anderer Vorteil: Es ist möglich, strukturierte und unstrukturierte Daten sowohl aus relationalen Datenbanken als auch aus Dokumenten zu vernetzen. Dokumente werden nicht wie in Suchmaschinen nur mit Schlüsselwörtern indiziert, sondern auch verstanden. Das System ist also in der Lage, zwischen einem Apfel und der gleichnamigen Firma zu unterscheiden. Das ist wichtig für das Verständnis und die Analyse grosser Datensätze.

Mit heute verfügbaren Technologien können immer grössere Datenmengen gehalten und aggregiert werden. Um diese zu verstehen und so tatsächlich Nutzen daraus zu ziehen, müssen die Datensätze in Zusammenhang miteinander gebracht werden – ähnlich wie unser Gehirn durch

die Vernetzung einzelner Synapsen Schlussfolgerungen zieht. Aus Big Data wird Smart Data.

Langsam entdecken Unternehmen diese Funktionsweise für sich. Das beginnt mit der Einführung sogenannter Thesauri, einer semantischen Metaebene zur Homogenisierung von

so weit wie möglich auszudehnen. Das bietet folgende Zeit-, Kosten- und Qualitätsvorteile:

- **Transparenz:** Aufgrund der durchgängigen Vernetzung von Unternehmensdaten werden diese global sichtbar. Jeder Berechtigte ist in der Lage, auf die Daten jederzeit zuzugreifen.

«Das Semantische Web vereint strukturierte und unstrukturierte Daten aus unterschiedlichsten Quellen»

Christian Heuer

Begriffen. Thesauri sorgen für eine gemeinsame Sprache, optimieren Suchbegriffe durch intelligentes Tagging von Daten in Systemen und können als Referenzdatensystem dienen. Sie lassen sich sowohl unternehmensintern als auch über Unternehmensgrenzen hinweg einsetzen.

BREITES ANWENDUNGSGEBIET

Die Anwendungsmöglichkeiten semantischer Technologien sind vielfältig. Durch die offene Struktur der Daten können bestimmte Informationen für Kunden transparent und verfügbar gemacht werden, etwa Beipackzettel für Medikamente. Banken nutzen semantische Systeme, um ihre Datenströme zu verbinden und so Geldwäscherei auf die Spur zu kommen. Durch Vernetzung haben Unternehmen die Möglichkeit, ihre Produkte gemeinsam mit Zulieferern zu entwickeln, indem sie standardisierte Daten austauschen und gleichzeitig mit externen Informationen anreichern. Durch vernetztes Datenmanagement können neue Daten direkt mit externen Daten abgeglichen oder angereichert werden, um die Datenqualität zu erhöhen oder um Datenmigration effizienter zu gestalten.

Einige Unternehmen verfolgen bereits die Strategie, die Prinzipien des Semantischen Webs

Compliance-Anforderungen müssen dabei selbstverständlich berücksichtigt werden.

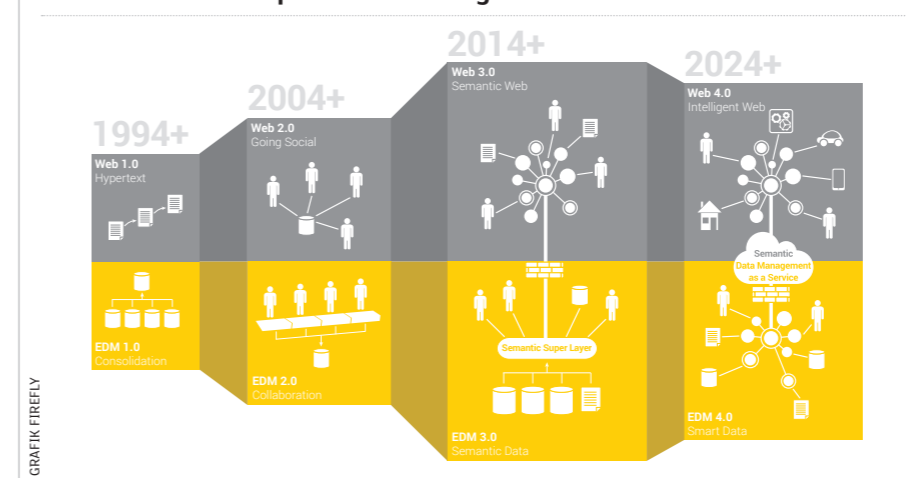
- **Mehr Flexibilität:** Vernetzte Daten sind darauf ausgerichtet, in möglichst vielen Applikationen verwendet zu werden, ohne bestehende Systeme zu beeinträchtigen. Zudem kann ein vernetztes Graphendatenmodell problemlos erweitert werden.

Reduzierte Komplexität: Durch gemeinsame Standards können Daten über Applikationen und Unternehmensgrenzen hinweg geteilt und akquiriert werden. Von der Öffnung von Datensilos für Benutzer und Applikationen profitieren auch externe Stakeholder, beispielsweise Kunden, Zulieferer oder sonstige Nutzer im Ökosystem des Unternehmens. Inhalte werden dementsprechend entweder visuell für den Enduser oder strukturell (im RDF-Format) für die maschinelle Verarbeitung aufbereitet.

- **Kosten:** Direkte Kostenvorteile ergeben sich in der IT dadurch, dass die Integration von Systemen und die Harmonisierung von Daten grundlegend vereinfacht wird. Daraus folgen weitere Einsparungen bei Material- und Personalausgaben. Denn wie bei vielen Dateninitiativen sind nicht alle Einsparungen direkt ersichtlich. ←

Christian Heuer ist Co-Founder und Head of Product Management bei Firefly Information Management AG → www.fireflyim.com

Evolution des Enterprise Data Management



Entwicklungsvergleich: Web und Enterprise Data Management