

Klartext.

Semantisches Netzwerk, Datenbank-Technologien, PIM-System, Content Authoring System, Content-Prozesse

www.klarso.com

Dr. Katharina Munk, Mitgründerin der **Klarso GmbH**. Das Startup wurde 2017 in Berlin gegründet und entwickelt innovative semantische Software-Technologie mit dem Kernprodukt klar:suite. Die klar:suite zur hoch-flexiblen, hoch-performanten Informationsablage und -Bearbeitung kommt mit ihrer flexiblen Software-Architektur in unterschiedlichsten Unternehmensbereichen und Branchen zur Anwendung.



Auf die Daten kommt es an. In vielen Unternehmen stellen die Handhabung und die Pflege komplex-strukturierter Inhalte wie Produktinformationen, Handbücher oder technische Dokumentationen eine große Herausforderung dar. Bei wachsenden Datenmengen und gleichzeitig veränderlichen Anforderungen müssen Softwaresysteme Daten für alle Business-Anwendungen bereithalten: Print, Web, App, schlaue Suche und Dialogsysteme.

Eine vergleichbare Aufgabe erfüllt das menschliche Gedächtnis: Es ist in hohem Maße fähig, Information und Erlerntes kontext- und bedarfsgemäß strukturiert bereitzustellen. Dabei erfolgen Ablage und Auswahl der Inhalte nicht über Ordner, Metadaten, Tabellen, Klassifikationen, XML-Tags und feste Algorithmen, bei denen die Daten in starre, einschränkende Strukturen gepresst werden.

In bestehenden IT-Systemen dagegen besteht ein Hauptteil der „Wissensarbeit“ darin, gedachte Schränke und Schubladen einzurichten und Daten zu sortieren ... in eine Ordnung und Datenstruktur, die für den nächsten Anwender oft schon nicht mehr nachzuvollziehen ist. Erschwerend kommt hinzu, dass die starren Strukturen für die Ablage einmalig zu Projektbeginn definiert werden müssen und sich nicht erweitern lassen. Strukturell abweichende Informationen müssen daher nachfolgend in (wildem) Sonderkonstruktionen untergebracht werden, mit denen das Verarbeitungssystem in der Regel inhaltlich nicht umgehen kann. Doch dieses Verfahren, unflexible Repräsentationen zu pflegen und gleichzeitig ihre Beschränkungen zu umgehen, ist nicht nachhaltig.

Orientierung am menschlichen Gedächtnis

Benötigt wird – in Anlehnung an die Funktionsweise des Gedächtnisses – ein neuer Ansatz, komplexe Daten zu pflegen,

zu bearbeiten, zu analysieren und bereit zu stellen. Diese besteht in einer natürlichen Datenrepräsentation und bildet möglichst viele Aspekte der Daten mit Kontext und Eigenheiten ab. So können langfristig und nachhaltig Zusammenhänge auch aus anderen neuen Kontexten heraus abgefragt werden – mit einer neuen Sicht auf die Daten. Um derartig frei strukturierte Daten jedoch handhaben zu können, müssen diese *smart* werden!

Das heißt weg von einer übergeordneten Verwaltung und zusätzlichen bedarfsgetriebenen Metadaten hin zu lokaler Selbstverwaltung und integrierten Strukturinformationen. Dazu benötigen die Daten-Elemente Informationen z.B. über ihre Quelle oder ihre Zusammengehörigkeit. Dieses direkt in den Daten abgelegte Wissen über die eigene Struktur ist nur ein Aspekt von *smarten* Daten.

„Natürliche“ Datenstruktur entspricht mehr der realen Welt

Semantische Software-Technologie der nächsten Generation bietet hier eine entsprechende Lösung für hoch flexible und komplexe Datenanwendungen und setzt den Gedanken von smarten Daten systematisch um. Das heißt, dass in Lösungen wie die der Klarso GmbH die Daten als frei zu definierende Entitäten/Elemente/Einträge zusammen mit ihren Merkmalen und Beziehungen in einem semantischen Netzwerk abgelegt werden. Das Neue an diesem Ansatz ist, dass dabei eine möglichst natürliche Inhaltsstruktur aufgebaut wird, deren Elemente und Beziehungen Entsprechungen in der realen Welt haben.

Durch eine erweiterte Sprachunterstützung mit Grammatikzerlegungen kommt die Lösung dabei der natürlichen Datenrepräsentation auch von Text-Content mit seiner sprachlichen Vielfalt noch einen Schritt näher. Mit zunehmendem Textverständnis erschließen sich Anwendungen von automatisierter Text-Klassifikation über echte kontextsensitive Suche hin zu einem bedarfsgemäß strukturierten Inhaltsmedium.

Semantisches Netzwerk mit angeschlossenen Anwendungen

Das semantische Netzwerk ist nicht nur eine semantische Datenbank, sondern eingebettet in ein vollständiges Software-System, die klar:suite. Die speziell designten Komponenten machen das semantische Netzwerk zugänglich und für Anwendungen produktiv, so beispielsweise für den Daten-Import/-Export, Editoren oder Suchfunktionen. Der Systemaufbau richtet sich dabei nach den folgenden Prinzipien: Die einzelnen Module sind leichtgewichtige Komponenten eines flexibel anpassbaren Systems. Jede Komponente dient dabei einem speziellen Aspekt, um maßgeschneiderte Anwendungen aus einfachen Bausteinen mit wenig Entwicklungsaufwand aufzubauen.

In diesem System besteht eine enge Beziehung zwischen Daten und ihren Ansichten/Editoren auf der Benutzeroberfläche, um schnell und flexibel neue Datentypen und ihre Bedingungen zu implementieren. Die Software-Komponenten sind als Skripte selbst in der Datenbank eingebettet. Die gemeinsame Speicherung und Verwaltung von Daten und Skripten stellt ►

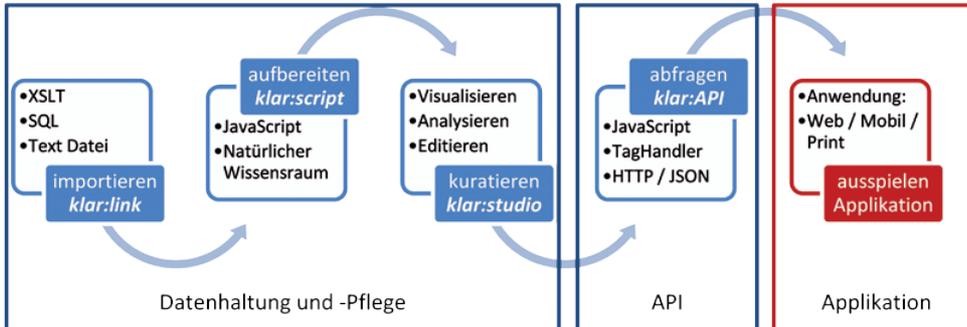
Technische Dokumentation.

- Entlastung in Spitzenzeiten
- Expertenwissen auf Abruf
- Zusammenarbeit auf Dauer

Wir machen Ihrer Doku Beine!

München • Frankfurt am Main • www.people-text.de

{ people text



Semantisches Netzwerk mit möglichen Anwendungen

selbst Teil der smarten Daten dar. Die Datenbank synchronisiert sich für Backups, Versionierungen und dezentrale Bearbeitung. Damit stehen Skript-basierte neue Funktionen und Erweiterungen sofort allen Nutzern zur Verfügung. Besonders während der Einführungsphase lässt sich die Benutzeroberfläche so schnell erweitern und agil anpassen.

Variable Einsatz-Szenarien

Das Software-System kann in verschiedenen Kontexten eingesetzt werden: So wird die klar:suite beispielsweise als PIM-System zur redundanzfreien, zentralen Pflege und Ausgabe komplexer Produktinformationen und Preislisten genutzt. Deren Komplexität entsteht z.B. durch Spezialanpassungen der Produkte, Abhängigkeiten der Preise von den jeweiligen Serien,

verschachtelte Optionen, Merkmalen und erlaubten Bestellkombinationen. Durch spezielle Export-Skripte und entsprechende Editoren lässt sich die Ausgabe von Preislisten flexibel kundenspezifisch anpassen (Rapid-Prototyping).

Auch Content-Anbieter nutzen das System – als Content Authoring System. Damit lassen sich komplexe Inhalte, z.B. aus der Wissensdomäne Medizin, pflegen, bearbeiten und kontextabhängig und präzise kundenspezifisch bereitstellen oder bei einer Webplattform auf eine Such-Anfrage hin aufbereiten. In einem Markt, in dem Informationen heutzutage über viele Quellen erhältlich sind, bieten Content-Anbieter ihren zahlenden Kunden so Mehrwert.

Und last, but not least: Unternehmen mit mehreren Insel-Lösungen können alle zusammengehörenden Content-Prozesse in einem zentralen semantischen Netzwerk vereinen. Separate Systeme führen zu unzusammenhängenden Informationen und meist einer redundanten Datenhaltung. Der Zugriff auf zusammengehörende und verteilte Informationen ist kompliziert, die Pflege der redundanten Daten aufwändig und Inkonsistenzen und Fehler sind kaum zu vermeiden. Zur schnelleren Umsetzung lässt sich in einem Bereich beginnen und der Content anderer Bereiche durch die flexible Datenstruktur sukzessive ergänzen. Die Einführung einer solchen Lösung amortisiert sich schnell. Der Übergang in das gemeinsame Datennetz ist eine Chance zur Datenbereinigung und Vereinheitlichung – und zu neuen Business-Anwendungen.

Fazit

Software für semantische Netzwerke bieten Lösungen für hoch flexible und komplexe Datenanwendungen. Auf der Basis von smarten Daten legen diese die Grundlage für Produktinformationen, Kataloge, wissenschaftliche Fakten, Fachbuchtexte, strukturierte Dokumente wie Verträge, Handbücher, Qualitätsdokumentationen und vieles weitere. ■

Aspekte smarter Daten

- **Selbst-organisierend:** sie kennen selbst ihre Quelle, Besitz- und Zugriffsberechtigungen, Lebenszeit und Verwendung und versionieren sich selbst
- **Selbst-verarbeitend:** sie wissen selbst, wie und nach welchen Regeln sie bearbeitet werden
- **Selbst-pflegend:** sie kennen selbst ihre begrenzte Lebenszeit, Gültigkeit oder Relevanz, um ggf. vergessen werden zu können
- **Selbst-verständlich:** sie enthalten eine in sich abgeschlossene Beschreibung im Kontext einer bestimmten Applikation oder eines Wissens-Korpus
- **Selbst-reflektierend:** sie prozessieren sich selbst, aktualisieren Schlussfolgerungen und strukturelle Analysen
- **Selbst-redend:** sie beschreiben und verorten sich selbst im natürlichen Sprachraum.