

Datenschutz in PLM &amp; Co. heben

# Unternehmensweite semantische Suche spinnt digitalen Faden

Der Datenschatz, der in allen zu einem Produkt abgelegten Daten steckt, ist nicht einfach zu heben. Auch das Fischen im Data Lake oder das Data Warehousing kommen bei unstrukturierten Daten an ihre Grenzen. Doch mit der semantischen Suche im Datenbestand des Unternehmens per Enterprise Search von Sinequa eröffnen sich neue Möglichkeiten, nicht zuletzt durch die Large Language Models (LLMs).

Frank Zscheile, IT-Journalist, München

Mit den Tools für die unternehmensweite Suche von Sinequa (hier Workplace Search) lassen sich die Grenzen zwischen den verschiedenen Daten-silos durchbrechen.

**D**ata Warehousing funktioniert nicht bei unstrukturierten Daten. Um Daten aus verschiedenen Silos an einem Ort zu konsolidieren und den digitalen Faden zu spinnen, braucht man deshalb Softwarelösungen für die intelligente Unternehmenssuche. Im Lebenszyklus eines Produkts verwendet jedes beteiligte Team unterschiedliche Prozesse und Werkzeuge. Der Vertrieb legt Spezifikationen für ein neues Produkt fest und speichert sie in seiner Projekt- oder Vertriebsmanagementsoftware. Die Konstruktion verwaltet Teile, Stücklisten und Zeichnungen, dazu das Ände-

**»In der Produktentwicklung werden Konstruktionsdaten ermittelt, die besser zur Rolle der fragenden Person und zu früheren Suchanfragen passen.«**

rungsmanagement seine in PLM- und CAD-Tools. Nach Fertigstellung und Freigabe werden die Produktdaten erneut übertragen, diesmal an das ERP-System, bei komplexeren Produkten u.U. in ein Manufacturing Execution System (MES). Nach Auslieferung an den Kunden übernehmen schließlich Wartungs-, Reparatur- und Betriebsteams (MRO) die Verwaltung der Projektdaten in den von ihnen verwendeten Tools.

Vier isolierte Projektmanagement-Tools also für jeweils einzelne Geschäftsaufgaben. Und es gibt keine Garantie, dass diese effektiv kommunizieren, weil die Informationen in ihnen unstrukturiert und nur vom jeweiligen Tool aus überhaupt einsehbar sind. Die Lücken zwischen diesen verhindern, dass der digitale Faden vollständig und zusammenhängend ist.

Die Herausforderung liegt also darin, Silos miteinander zu verbinden, die in ihnen verwalteten Informationen für alle sichtbar zu machen und so den digitalen Faden zu spinnen. Einst galt Data Warehousing als Mittel dafür. Hierbei werden Daten an einem Ort, dem Data Lake, konsolidiert, um den Zugriff zu erleichtern. Bei strukturierten Daten (relationale Datenbanken, Tabellenkalkulationen) funktioniert dies leidlich, weniger aber bei unstrukturierten, also Bildern, Diagrammen, Inhalten von Web-Chats und E-Mails, Kommentaren auf dem Kundenauftrag etc.

## Auf dem Weg zum echten PLM

Der weitaus größere Teil geschäftlicher Informationen (nach IDC-Schätzungen rund 80 – 90 %) liegt jedoch in unstrukturierter Form vor; als Office- oder PDF-Dokument, Webseite, E-Mail oder in Form von Bildern und Videos. Je unstrukturierter die Daten, desto höher die Anforderungen an die zugrundeliegenden Algorithmen zu deren Auswertung. Hier kommen Softwarelösungen für die intelligente Unternehmenssuche ins Spiel, die alle Datenquellen

im Unternehmen verbinden, strukturierte wie unstrukturierte – auch CAD-Programme, deren Inhalte oft proprietär und nicht quelloffen sind. Je mehr solcher vorgefertigter Konnektoren eine Such-Software mitbringt, desto breiter ihr Einsatzgebiet. Bei der Plattform des Herstellers Sinequa etwa sind es rund 200 Out-of-the-Box-Konnektoren, darunter zu den Tools der vdR Group sowie denen anderer US-amerikanischer Anbieter von Datenmanagement- und PLM-Lösungen für die Fertigungsindustrie.

Enterprise-Search-Plattformen indizieren die Inhalte der angeschlossenen Datenquellen, erstellen daraus eine Kerndatenbank und führen die Suchergebnisse in einer einheitlichen Ansicht zusammen. Durch KI-Funktionalitäten wie die Verarbeitung natürlicher Sprache (Natural Language Processing = NLP) und Algorithmen für maschinelles Lernen erkennen Enterprise-Search-Engines auch den Kontext des gesuchten Begriffs. Relevante Informationen können damit auch solche sein, in denen der gesuchte Begriff gar nicht vorkommt.


### ChatGPT & Co. bieten weitere Chancen


Durch das Aufkommen von großen Sprachmodellen (Large Language Models = LLMs) öffnet sich derzeit die Tür für die Vektorsuche, bei der Informationen auf der Grundlage ihrer Bedeutung abgerufen werden. Leistungsfähige Enterprise-Search-Engines integrieren die semantische Vektorsuche mit einer statistischen Schlüsselwortsuche, um die beste Relevanz zu erzielen. Sucht zum Beispiel ein Servicetechniker nach Informationen zur Teilenummer 808, gibt er diese in das Suchfeld ein und erhält Treffer, die nicht nur die Teilenummer als solches enthalten, sondern auf die Rolle des Suchenden und frühere Suchergebnisse zugeschnitten sind – ein „Google für Unternehmen“.

- Gibt man als Frage ein „Wie installiert man Teil 808 im Feld?“, würde die Enterprise-Search-Engine Informationen wie Zugangscodes für die Wartung, Benutzerdokumente, Statistiken usw. ausgeben, häufige Fehlersuchverfahren oder Installationsprozesse, die dem Suchbegriff entsprechen.
- In der Produktentwicklung hingegen könnte derselbe Suchbegriff Konstruktionsdaten ermitteln, die besser zur Rolle der fragenden Person und zur Auswahl aus deren früheren Suchanfragen passen.

Das Ergebnis ist ein vollständiger, zusammenhängender und durchsuchbarer digitaler Faden. Produktdaten zu jedem Prozessschritt sind für alle Beteiligten sofort im Zugriff.

[www.sinequa.com](http://www.sinequa.com)

**INFO**

Mehr zu den Suchmöglichkeiten von Sinequa:  
 [hier.pro/JdjnQ](https://hier.pro/JdjnQ)