

PROFI.info

IBM WATSON EXPLORER AC

WERTVOLLE INHALTE
UNSTRUKTURIERTER DATEN
NUTZEN

UNSTRUKTURIERTE VS. STRUKTURIERTE DATEN

Das Potenzial der in unstrukturierten Daten enthaltenen Informationen wird bei Weitem nicht ausgeschöpft.

Die Ursache hierfür ist, dass der „Inhalt“ der Information durch die Art der Verwendung von Text, mehr oder weniger versteckt wird und es der „Aufbereitung“ durch kognitive Prozesse bedarf. Der lesende Mensch kann dies gut – und kognitive Computersysteme sind auf dem Weg dies auch immer besser zu können.

Im Folgenden konzentrieren wir uns hierbei auf „Texte“ in Office-Dokumenten, „Freitexte“ in Datenbanken, User Generated Content in den Social Media, Websites, E-Mails usw. als Quellen unstrukturierter Daten (ohne prinzipiell darauf eingeschränkt zu sein).

Daten mit einer gleichartigen Struktur (= strukturierte Daten) besitzen durch ihre Position in einem System und ihre leichte(re) „Lesbarkeit“ einen direkteren Zugang für das Benutzen dieser Informationen, um deren Bedeutung und deren Wert zu erkennen.

Man kann relativ rasch ermitteln, **wie viele** E-Mails sich zwei Menschen zusenden, dies ggf. auch auf Zeitstrahlen legen usw. – das ist bereits Arbeiten mit unstrukturierten Daten. Die Fragen WER, WANN, WIE VIEL sind mit diesen unstrukturierten Daten zu beantworten.

Man kann sich auch fragen: **Worüber** sprechen zwei Menschen in ihren E-Mails? Für die Antwort muss man „anders“ mit den unstrukturierten Daten arbeiten: Man muss die Texte inhaltlich analysieren und Themenbereichen zuordnen. Die Antworten auf die Fragen WIE, WIESO, WESHALB, WARUM sind zumeist „nur“ in den unstrukturierten Daten zu finden.

Spannend wird es nun, wenn man z. B.

- a) die Themen in ihrem zeitlichen Verlauf sehen möchte
- b) dies mit den Häufigkeiten zusammenbringt
- c) versucht, gewisse Muster in Beziehung zu anderen Gegebenheiten zu erkennen.

Um dies zu erreichen sind in den „beiden Welten“ jeweils typische Aufgaben zu lösen, um damit strukturierte und unstrukturierte Daten geeignet zu verbinden:

Unstrukturierte Daten

- Die „Original-Daten“ (in allen Formen und in ggf. hohem Volumen) **müssen an ihren Quellen abgeholt werden und für die weitere Bearbeitung bereitgestellt werden.**
- **Aus den Daten sind Entitäten und Beziehungen zwischen den Entitäten auf Basis anzuwendender Regeln zu extrahieren.**
- **Diese Entitäten (und die Beziehungen zwischen den Entitäten) werden bei der folgenden Analyse der Daten verwendet.**

Strukturierte Daten

- **Datenberichte in Listenform aller Art**
- **Visualisieren der Daten in Diagrammen und Beziehungsgeflechten**
- **Weiteres Erforschen der Daten in ihrem jeweiligen Kontext**
- **Muster erkennen**
- **Aufdecken von „versteckten“ Beziehungen**

LÖSUNGSANSATZ

Mit **IBM Watson Explorer Analytical Components** (WEX AC) liegt eine Lösung vor, mit der Entitäten auf Basis von Wörterverzeichnissen und Regeln sowie für die eigene Wissensdomäne spezifisch trainierte Intelligenzbausteine extrahiert werden können.

WEX AC liest die Daten von den Quellen ein, extrahiert die Entitäten sowie die Beziehungen zwischen den Entitäten und stellt diese für die weitere Verarbeitung bereit.

Mit **IBM i2 Enterprise Insight Analysis** (i2 EIA) liegt eine Lösung vor, die quasi als Standard bei Ermittlungsbehörden im Einsatz ist um Datenbeziehungen zu untersuchen und zu visualisieren. Ein Datenanalyst arbeitet mit dem i2 Analyze Portal oder dem i2 Analyst's Notebook gegen die für i2 verfügbaren Daten.

Mit der „**Cogniware i2WEX-Bridge**“ können die strukturierten und unstrukturierten Daten zusammen betrachtet werden, indem sie die geeignete Unterstützung beim Erforschen der Daten bietet. Hierbei werden verschiedene Modi unterstützt:

- Arbeiten nur mit dem Frontend i2 Analyst's Notebook, hierbei werden die Daten im WEX AC gehalten mit einem On Demand-Zugriff gemäß Abfrage im Frontend
- Arbeiten mit einer kompletten i2-Infrastruktur, hierbei werden die WEX AC-Daten der Datenhaltung im i2-Repository gemäß deren Definitionen zugeführt

PROFI LEISTUNG

Wenn bei Ihnen ein hohes Volumen unstrukturierter Daten (neben strukturierten Daten) auftritt und das Auswerten dieser unstrukturierten Daten für Ihr Geschäft wichtig ist z. B. wegen

- Vermeidung von Verlusten durch das betrügerische Handeln anderer
- Verbesserung der Wettbewerbsposition (Reputation, Frühwarnsysteme etc.)
- Kostenreduzierung nach dem Erkennen von kostenrelevanten Faktoren
- Compliance einhalten

und wenn Sie dies bisher nicht im erforderlichen Maße umgesetzt haben, dann können wir mit Ihnen IHREN möglichen Lösungsansatz diskutieren.

Vorgehensweise

- **Auf Basis eines Project Definition Workshops können wir mit Ihnen einen Lösungsvorschlag erarbeiten, der als Roadmap (inkl. Darstellung der Kosten) zu verstehen ist.**
- **Im Rahmen eines Proof of Concepts erhalten Sie die Möglichkeit, den erarbeiteten Lösungsvorschlag zunächst prinzipiell zu bewerten.**

WEX AC UND I2 EIA IM KONTEXT DER AUFBEREITUNG UND PRÄSENTATION VON DATEN

Watson Explorer Content Analytics

Lösung für unstrukturierte Inhalte / Daten

Der Inhalt ist im fließenden Text „versteckt“, z.B. in Office-Dokumenten, Freitexten in Datenbanken, Websites, Social Media, E-Mails etc.



Extrahiert Entitäten und Beziehungen



Verarbeitet viele Dokumente gleichzeitig



Analysiert Daten auf Entitäten-Level



Volltextsuche

i2 Enterprise Insight Analysis

Lösung für strukturierte Inhalte / Daten

Daten mit bekannter Struktur und Bedeutung, i.allg. in Datenbanken wie ERP, CRM, Finanzen usw.



Durchsuchen unbekannter Daten nach möglichen Mustern



Daten in Kontext(en) und verschiedenen Sichten untersuchen



Identifizieren verborgener & unbekannter Beziehungen



Visualisieren der Daten und Ihrer Beziehungen

WEX AC und I2 EIA sorgen mit ihren jeweiligen spezifischen Möglichkeiten für das geeignete Aufbereiten und Präsentieren der Daten. Cogniware i2-Bridge sorgt dafür, dass "die beiden Welten" in einem Analyse-Client zusammengeführt und gleichzeitig genutzt werden können.

USE CASES

Cogniware WEX-i2 Bridge use cases*

Im Folgenden sind einige Use Cases aufgelistet, die durch die Integration von IBM Watson Explorer Analytical Components und IBM i2 Analyst's Notebook gelöst werden können.

Use Case 1 – Öffentliche Sicherheit und Aufklärung

Eine Ermittlungsbehörde (oder eine mit Ermittlungen beauftragte Organisation) extrahiert verschiedene Entitäten wie Personen, Orte, Organisationen oder Ereignisse z.B. aus Berichten oder Web-Quellen (Social Media, Diskussionsforen etc.).

Danach verwenden sie diese Daten in i2 Analyst's Notebook, um Verbindungen zwischen den Entitäten zu visualisieren und ggf. verborgene Beziehungen aufzudecken. Dies ermöglicht den Ermittlungsorganisationen, aus verschiedenen Sichten die Entitäten und deren Beziehungen untereinander zu betrachten und anspruchsvolle Diagrammlayouts zu erstellen, um in einer Benutzeroberfläche mit höherer Effizienz zu arbeiten

Use Case 2 – Kundendienst

Marketing oder PR-Abteilungen können ihr Kundennetzwerk über alle verfügbaren Kommunikations- und Verkaufskanäle visualisieren und deren „Texte“ analysieren. Hiermit ist es möglich, zufriedene Kunden zu clustern und sog. „Hater“ zu identifizieren, um ggf. geeignete Aktionen abzuleiten.

Use Case 3 – Gesundheitswesen (Insbesondere Unikliniken)

Ärzte füllen in allen Kliniken Tausende von Patientenakten aus. Eine große Menge an Informationen ist in medizinischen Zeitschriften, Artikeln und sogar in strukturierten Datenbanken enthalten.

IBM Watson Explorer kann verwendet werden, um nach Themen wie Krankheiten, Behandlungsoptionen usw. in diesen offiziellen Quellen zu suchen und alle aktuellen

Patientendaten dabei mit einzubeziehen, um nach Symptomen, Medikamenten, Alter, Geschlecht etc. zu suchen.

Die i2 WEX-Bridge ermöglicht es, alle diese Daten in einer leicht verständlichen grafischen Aufbereitung zu zeigen – verständlich nicht nur für Analysten, sondern auch für Ärzte. Neue Verbindungen zwischen Krankheiten, Behandlungen, Reaktionen, Alter, Geschlecht etc. können aufgedeckt werden.

Use Case 4 – Versicherungsbetrug

Manche Kriminelle schaffen komplizierte Systeme, um zu verbergen, dass ein angemeldeter Versicherungsanspruch tatsächlich betrügerisch ist. Viele Details dieser Fälle sind in den Berichten von Versicherungsfachleuten enthalten, meist jedoch „versteckt“. Spätere Detektivarbeit ist lang und teuer.

IBM Watson Explorer kann verwendet werden, um den Inhalt dieser Berichte zu strukturieren, indem Entitäten identifiziert werden: Autos, Orte, Personen, Nummernschilder, Telefonnummern, Spitznamen, Firmen usw.

Die i2WEX-Bridge überträgt alle diese Entitäten und die Beziehungen zwischen ihnen in das beste „detective IT-Tool“ – IBM i2 Analyst's Notebook. Hier kann der Versicherungssachbearbeiter, der sich mit Betrugserkennung beschäftigt, nach auffälligen Mustern suchen und diese zielgerichtet untersuchen.

Use Case 5 – Wirtschaftskriminalität

Viele Firmen suchen Wege für eine „Steuroptimierung“, einige gehen dabei zu weit. Wenn es erforderlich wird z.B. Geldströme (oder auch Informationsströme, Produkt- und Warenströme etc.) zu visualisieren, dann ist i2 Analyst's Notebook ein mächtiges Tool. Allerdings können einige Informationen auch außerhalb der strukturierten Daten - z. B. in

Dokumenten oder in Social Media Quellen - liegen.

Mit IBM WEX können darin Personen, Orte, Firmen usw. gefunden werden, die mit der i2WEX-Bridge an das i2 Analyst's Notebook übergeben werden können, um damit die strukturierten Daten zu ergänzen. Damit ist einiges mehr aus den Beziehungsgraphen abzulesen.

Use Case 6 – General Data Visualization (Marketing, Sales, Management, GDPR)

Der letzte Anwendungsfall ist ein allgemein formulierter: Im Tagesgeschäft verbringen Manager, Verkaufsteams und Marketing-Teams viel Zeit damit, ihre Verkaufs-, Finanz- und sonstigen Daten zu analysieren. Es könnte für sie interessant sein, diese Daten grafisch zu betrachten:

- Was hängt mit wem zusammen?
- Was und wie verändern sich diese Verbindungen im Laufe der Zeit?

Aber nicht alle Unternehmensdaten sind strukturiert, tatsächlich sind nach übereinstimmender Ansicht mehr als 80% unstrukturiert.

IBM Watson Explorer kann verwendet werden, um (interne und externe) Daten zu analysieren, um nach „Themen“ zu suchen: Produkte, Verkäufe, Kampagnen, Unternehmen, Personen usw. Diese können mit der i2WEX-Bridge in das i2 Analyst's Notebook übertragen werden, und mit der Visualisierung und Strukturierung der Daten kann dann mehr als zuvor gesehen aus den verschiedenen Darstellungen gefunden werden.

*Quelle: Cogniware, Redaktionelle Überarbeitung: PROFI

HABEN WIR IHR INTERESSE GEWECKT?

Dann sprechen Sie mich an.



Omar Alaoui

Software-Architekt

+49 6151 8290-7562

o.alaoui@profi-ag.de

DIE PROFI ENGINEERING SYSTEMS AG

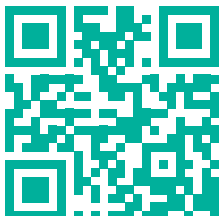
Wir sind ein mittelständisches, inhabergeführtes und finanzkräftiges IT-Lösungshaus mit Hauptsitz in Darmstadt. Innovationskraft und Kundenorientierung sind wesentliche Säulen unserer Unternehmensstrategie. Seit über 35 Jahren unterstützen wir unsere Kunden mit individuellen hochwertigen Lösungen zur Optimierung von IT-Prozessen und Systemlandschaften.

Wir begleiten Ihr Unternehmen bei der digitalen Transformation – von der Strategie über die Umsetzung bis zum Betrieb. Unsere erfahrenen Berater und Architekten beschäftigen sich seit vielen Jahren intensiv mit der Digitalisierung aller Geschäftsabläufe und Unternehmensbereiche. Gestalten Sie mit den PROFI-Fokusthemen schon heute Ihre digitale Zukunft. Profitieren Sie von unserem Know-how vor allem im Kontext von Managed Service Solutions, VDI & Digital Workplace, SAP HANA, Business Continuity, Agile Software-Entwicklung & DevOps, Netzwerk & Security, SDDC & Agile Plattformen, Speicherlösungen und Server-Lösungen. Wir übernehmen für Sie Projektmanagement und Implementierung, einschließlich dem Betrieb Ihrer Systeme und Plattformen. Unser Anspruch ist höchste Kompetenz, Zuverlässigkeit und Qualität, mit messbarem Erfolg und direktem Beitrag zur Wertschöpfung und Wettbewerbsfähigkeit unserer Kunden.

Wir beschäftigen rund 320 Mitarbeiterinnen und Mitarbeiter bundesweit an 12 Standorten. Seit vielen Jahren gehören wir zu Deutschlands erfolgreichsten Systemhäusern und pflegen langjährige Partnerschaften mit allen führenden IT-Herstellern.

Unsere IT-Lösungen für Ihren Erfolg

- **Business Continuity**
- **Cloud & Hybrid**
- **DevOps & Agile Software-Entwicklung**
- **Managed Services & Managed Service Solutions**
- **SAP HANA**
- **SDDC & Agile Plattformen**
- **Security & Netzwerk**
- **Server-Lösungen**
- **Storage**
- **VDI & Digital Workplace**



PROFI Engineering Systems AG

Otto-Röhm-Straße 18
64293 Darmstadt
Telefon: +49 6151 8290-0
Telefax: +49 6151 8290-7610
E-Mail: profi@profi-ag.de
www.profi-ag.de

UNSERE PARTNER

Gemeinsam mit unseren starken Partnern setzen wir Ihre optimalen Lösungen um.



Gold Partner



12/2019