



## Safety-Reports: Tiefere Einblicke dank Datenanalyse

### AUF EINEN BLICK

#### Software / Die Lösung

- Planung, Einführung und Umsetzung der MIOvantage-Daten-Management-Plattform
- Implementierung einer maschinellen Lerntechnik zur Bewertung der Datenqualität und Definition eines neuen Kategorisierungsschemas mittels der MIOsoft Lösung MIOvantage

#### Vorteile

- Daten-Identifizierung und Klassifizierungsmustern mit hoher Qualität umgesetzt
- Erhöhung der Datenqualität bei gleichzeitiger Reduzierung des Zeitaufwandes

#### Architektur

- tiefe Erkenntnisse über potenzielle Gefahrenquellen
- Definition frischer KPIs
- verbesserte Sicherheitsstandards



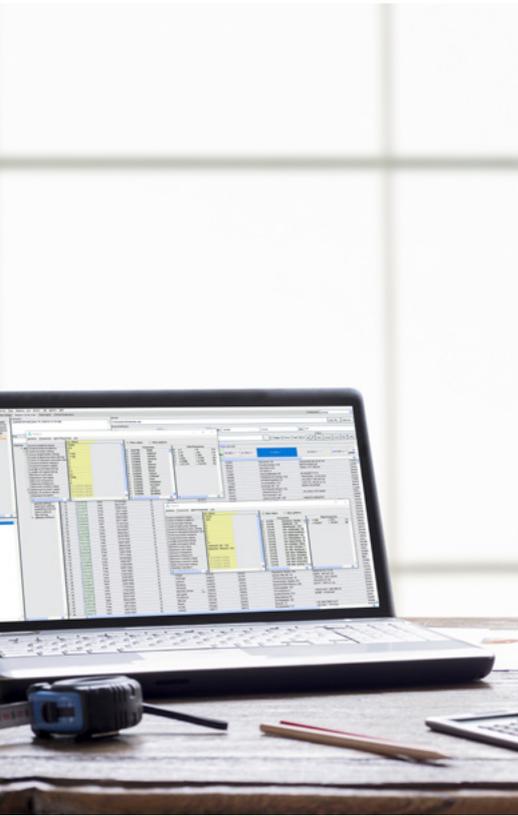
### Effiziente Datenqualitätsbewertung

Industriekonzerne, wie Unternehmen aus der Öl- und Gas-Branche nutzen oftmals ein HSSE (Health, Safety, Security & Environment)-Leistungsberichtstool, um Erkenntnisse, Gefahren und Vorfälle zu verfolgen. Die enthaltenen Daten werden in der Regel laufend vom Management analysiert und überprüft. Während der Dateneingabe klassifizieren die Benutzer die Informationen manuell für eine spätere Wiederverwendung. Statistische und KPI-Berichte werden genutzt, um positive Lerneffekte zu erzielen, die Sicherheitsstandards zu erhöhen und eine Verhaltensänderung bei der täglichen Arbeit voranzutreiben. Das alles hat das übergeordnete Ziel, Vorfälle zu reduzieren. Diese Eingaben werden manuell vollzogen. Oftmals kann die Verlorenen-Verletzungsrate stark reduziert werden – aber eben nicht bis auf null.

### HERAUSFORDERUNG

#### Gesucht: Erkenntnisse aus vorhandenen Daten

Aktuelle Methoden, um Datensätze zu analysieren und so Erkenntnisse aus diesen zu gewinnen, beinhalten eine große menschliche Komponente. Die Herausforderung vieler Unternehmen dieser Branche und Größe, besteht darin, eine Methode zu finden, die die menschliche Komponente so weit wie möglich reduziert. Ziel ist dabei, die möglichen Optionen der Verwendung eines maschinenbasierten Klassifikationssystems gegenüber dem bestehenden menschlichen Klassifikationsansatz zu verifizieren, um mögliche Verbesserungsbereiche zu identifizieren. ►



Um das zu erreichen, sollten maschinelle Lerntechniken zur Bewertung der Datenqualität verwendet werden. Zudem sollte ein neues Kategorisierungsschema definiert werden, um die Dateneingabe effizienter zu gestalten und die Datengenauigkeit und -konsistenz zu erhöhen.

## LÖSUNG

### MIOvantage – Datenanalyse Out of the Box

Eine Optimale Lösung für Unternehmen mit diesen Herausforderungen und Ansprüchen ist MIOvantage von MIOsoft. MIOvantage bietet einen Data First-Ansatz zur Kontextualisierung strukturierter und unstrukturierter Daten, um exakt jene tieferen Einblicke in vorhandene Daten zu erhalten. Zudem bietet die Lösung eine Verbesserung der Datenqualität und eine höhere Vertrauenswürdigkeit der Berichte. So wird zudem eine Ausgangsbasis für neue KPIs zur Messung der führenden Leistung im Bereich Sicherheit geschaffen.

## UMSETZUNG

### Datenclustering mit MIOvantage

Zunächst wurden die Daten normalisiert, Rechtschreibfehler erkannt und Fragmente mit Hilfe eines Porter-Stemming-Algorithmus erstellt.

Dann wurden die Daten in der Tiefe analysiert, in dem Techniken wie visuelle Dateninspektion, Tokenisierung, statistische Wortzählungen, Wortkombinations-/Mustererkennung, Drilldown, Fuzzy-Suche, allgemeine Ausdrücke und kontextbezogene Datenanalyse eingesetzt wurden. Mit Hilfe der kontextuellen Datenclustering wurden Datenobjekte kategorisiert und Übereinstimmungslisten erstellt.

Um die bestehende Klassifikation zu normalisieren, wurde ein regelbasierter Algorithmus zum Clustern von Vorfall- und Suchdatensätzen auf der Grundlage der Freitextfelder innerhalb der definierten Klassifikationsmatrix verwendet. Um die Daten nun in verschiedenen Themenbereichen zu clustern, wurden mehrere Algorithmen des maschinellen Lernens evaluiert, schließlich wurde der Algorithmus des maschinellen Lernens TF-IDF & k-Medoids auf die Clusterdaten angewendet.

## ERGEBNIS

### Gefunden: Tiefe Einblicke in vorhandene Daten

Durch den Einsatz von MIOvantage wurde die Identifizierung der Daten und die Klassifizierungsmustern in unstrukturiertem Text mit hoher Qualität umgesetzt. Die Klassifikationsschemata konnten nun voll auf die Bedürfnisse des Benutzers abgestimmt werden. Im Zuge dessen kam es zu einer Erhöhung der Datenqualität bei gleichzeitiger Reduzierung des Zeitaufwandes für die Dateneingabe. Dadurch können nun neue Erkenntnisse über potenzielle Ursachen für Vorfälle dazu dienen, frische KPI-Sätze zu definieren, um Sicherheitsstandards in Unternehmen der Gas- und Öl-Branche weiter zu verbessern.

## DEXTRADATA GMBH

### STANDORT ESSEN

Girardetstraße 4  
45131 Essen

Telefon +49 201 9 59 75 0  
info@dextradata.com  
www.dextradata.com

### STANDORT BERLIN

Am Treptower Park 75  
12435 Berlin

### STANDORT HAMBURG

Heidenkampsweg 99  
20097 Hamburg

### STANDORT MÜNCHEN

Oskar-Schlemmer-Straße 3  
80807 München