

Heben Sie Ihren Datenschatz!



Ralph Tonzler
Partner AWK Group

Aus Daten gewinnt man Informationen, die in unserem Zeitalter der Informationsgesellschaft in grossem Überfluss vorhanden sind. Leicht verliert man den Überblick und verpasst es, die relevanten Informationen zu extrahieren und daraus Wissen zu gewinnen, um schliesslich Weisheit zu erlangen.

Mittel und Tools sind in vielen Unternehmen in reichlichem Masse im Einsatz, um aus den Geschäftsdaten mehr Informationen und ein besseres Abbild der Realität zu erhalten.

In den Projekten stellen wir oft fest, dass Lücken in der Auswertungsmethodik dieser Daten bestehen oder dass die Möglichkeiten der bestehenden Tools nicht optimal ausgeschöpft werden.

Gerne möchten wir Ihnen mit diesem AWK-Fokus einen Überblick über Methoden und Tools der Datenanalyse geben und Sie anhand beispielhafter Einsatzmöglichkeiten motivieren, allfälliges Potenzial auch in Ihrem Unternehmen zu nutzen.

Herzlich, Ihr

Mine your data – Ein Weg zum besseren Verständnis des Unternehmens

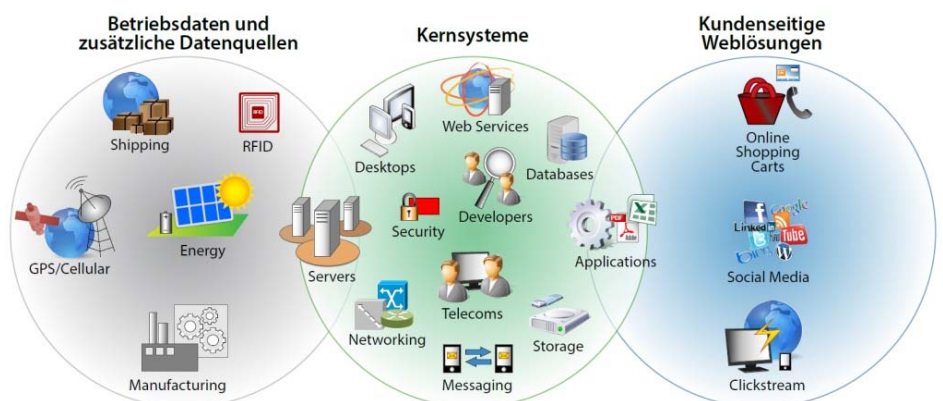
Fast alle wichtigen Prozesse werden heute durch IT-Systeme unterstützt und elektronisch verwaltet – zumindest auf der Datenebene. Dennoch werden die Daten für strategische Entscheide, Prozessoptimierungen oder für die Umsetzung neuer Geschäftsideen zu wenig genutzt. Wir wollen nachfolgend Möglichkeiten aufzeigen, wie Datenanalysen zum Nutzen des Unternehmens eingesetzt werden können.

Patrik Rüegge, Dr. Jonas Dischl, Dr. Dimitrios Tombros

Die Datenmengen in Unternehmen steigen im Schnitt um 40% jährlich. Diese Daten werden meistens nur innerhalb dedizierter Geschäftsanwendungen genutzt. Neben den klassischen Anwendungsdaten und den unstrukturierten Dokumentsammlungen haben auch die maschinengenerierten Daten aus Geräten und Systemen stark zugenommen. Im Zusammenhang mit Social Media und E-Business sind in den letzten Jahren neuartige Werkzeuge für die Datenanalyse und -interpretation auf den Markt gekommen, die ohne ein Datawarehouse auskommen.

Eine bessere Datenanalyse und die daraus abgeleitete Nutzung bringen verschiedene Vorteile:

- Besseres Verständnis der Wertschöpfungskette: Aufbereitete Daten zeigen die effektiven Abläufe und sind Ideenquellen für Verbesserungen und effizientere Prozessgestaltung.
- Grundlage für strategische Planung: Eine Visualisierung der Daten zeigt den Entscheidungsträgern komplexe Zusammenhänge einfacher auf, womit strategische Anpassungen zielgerichteter und effizienter durchgeführt werden können.
- Bessere Entscheidungsgrundlagen: Nicht das Bauchgefühl, sondern fundierte Fakten sind die Basis für Entscheide.



Quellen für die Datenanalyse

Datenformate und Analysetools

Daten liegen in der Regel aus verschiedenen Quellen und in unterschiedlichen Formaten vor. Neue Tools und Technologien erschliessen diese Daten für die Analyse.

Beispiele für Datenformate sind:

- Strukturierte Daten in Tabellen
- Semistrukturierte Daten, wie Logs und XML-Dateien
- Unstrukturierte Daten (Dokumente oder Freitext-Kommentare)
- Graphen bzw. Relationendaten

Für diese Daten können die hier vorgestellten Analysemethoden grundsätzlich angewandt werden, wenn die passenden Tools verfügbar sind und die Daten richtig aufbereitet werden.

Für die Datenanalyse gibt es ein breites Spektrum an Tools – von einfachen SQL-Interfaces für die manuelle Datenanalyse bis hin zu umfangreichen Data Mining Tools.







Eine Auswahl an verschiedenen Tools für Datenanalysen

Methoden der Datenanalyse

Es gibt sehr unterschiedliche Methoden für die Datenanalyse. Wer hier den Überblick hat, kann mit der richtigen Methodenwahl optimale Ergebnisse erzielen. Bewährt, bekannt und weit verbreitet sind die "manuellen" Methoden: Der Anwender nutzt den Computer, um Daten auf verschiedene Arten darzustellen und zieht seine Schlüsse aus dem, was er sieht.

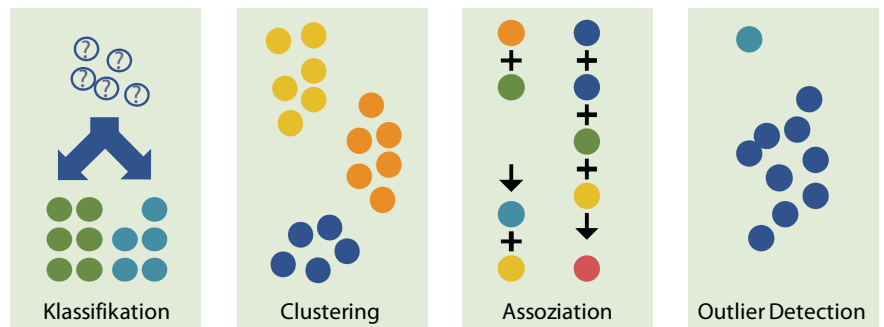
Bei Daten mit vielen Attributen und komplexen Zusammenhängen stossen die manuellen Methoden jedoch an ihre Grenzen. Hier setzt "Data Mining" an – der Computer durchpflügt die Daten automatisiert nach relevanten Mustern und Zusammenhängen.

Manuelle Datenanalyse	Data Mining
 Gruppieren, Isolieren Anordnen, Darstellen	 Muster, Strukturen suchen Modelle rechnen
 Muster, Strukturen suchen Hypothesen bestätigen	 Suchalgorithmen steuern Ergebnisse interpretieren

Zwei Arten von Datenanalysen

Data Mining wird für zahlreiche Problemstellungen seit vielen Jahren erfolgreich angewendet und gewinnt mit dem aktuellen Thema "Big Data Analytics" weiter an Bedeutung. Es lohnt sich deshalb, die wichtigsten Problemstellungen zu betrachten, die mit Data Mining gelöst werden können.

- In der **Klassifikation** werden neue Daten auf Grund der bisher gemachten Erfahrungen automatisch kategorisiert. Mit dieser Methode entscheiden beispielsweise Banken, ob ein neuer Kunde kreditwürdig ist oder nicht.
- Durch **Clustering** werden Gruppen von ähnlichen Daten sichtbar gemacht. Unternehmen lernen die verschiedenen Kundentypen besser kennen und können z.B. Werbung gezielter einsetzen.
- In der **Assoziation** wird erkannt, welche Daten häufig zusammen vorkommen. So wird beispielsweise erkennbar, welche Produkte oft zusammen gekauft werden, um das Produktportfolio oder das Layout von Ladenlokalen zu optimieren.
- Die **Outlier Detection** erkennt auffällige Abweichungen – auch dort, wo sie mit manuellen Methoden übersehen werden. Dies ist beispielsweise in der Betrugserkennung von grosser Bedeutung.



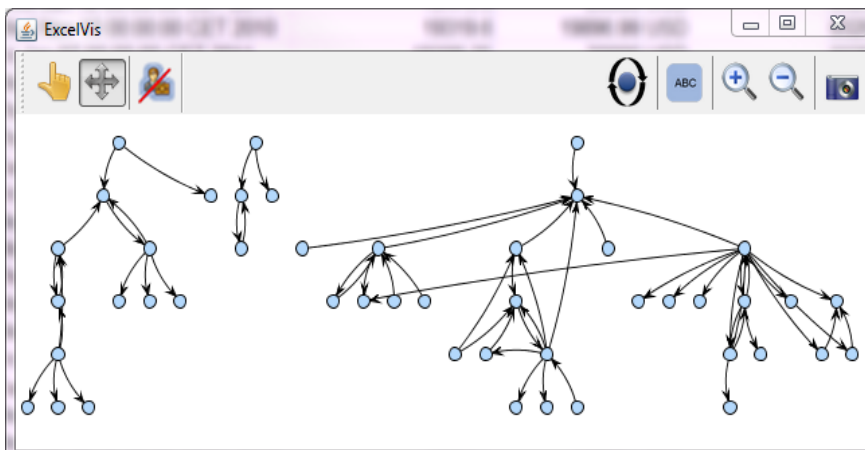
Die wichtigsten Problemstellungen im Data Mining

Einsatzmöglichkeiten

Die Möglichkeiten der Datenanalyse sind so vielfältig wie die Daten, die in den Unternehmen verfügbar sind. Die folgenden konkreten Einsatzbeispiele decken nur einen kleinen Teil des Spektrums ab:

- **Cross-Selling-Analysen** bewähren sich in vielen verschiedenen Branchen. Beispielsweise konnte damit die Information über ein neues Lebensversicherungsprodukt eines grossen Versicherers gezielt an potenzielle Kunden kommuniziert werden.
- Eine **Social Media-Analyse** kann z.B. zeigen, inwiefern sich Tweets von Kunden eines Kommunikationsdienstleisters für die Früherkennung von Incidents eignen. Darüber hinaus bietet sich Social Media für eine „Sentiment Analysis“ an, um einen Überblick der Kundenstimmungslage in Bezug auf verschiedene Angebote und Produkte zu gewinnen.
- Mit der **Auswertung von Service-Requests** können Produkte mit erhöhtem Support-Aufwand identifiziert und das Produktportfolio optimiert werden. Text-Mining sucht in den Freitextfeldern von Service-Requests nach Mustern, die auf bestehende, zuvor unbekannte Probleme hinweisen.
- In der **Betrugserkennung** werden verschiedene Methoden der Datenanalyse eingesetzt, um bisher unbekannte Betrugsmuster zu identifizieren oder verdächtige Ereignisse zu klassifizieren. Auf diese Weise werden z.B. verdächtige Schadensmeldungen bei Versicherungen oder Zusammenhänge im Ereignisfall automatisiert erkannt.
- Marktanalysen helfen beispielsweise beim **Fiber-To-The-Home-Ausbau** das Potenzial verschiedener Gebiete zu identifizieren und dementsprechend den Ausbau zu priorisieren.

Eine besonders attraktive, intuitiv erfassbare Anwendung ist das sogenannte **Visual Data Mining**. Diese Methode visualisiert die Daten optimal für die Mustererkennungsfähigkeit des menschlichen Gehirns. Sogenannte Transaktionsgraphen helfen beispielsweise, verdächtige Konstellationen von Zahlungsströmen in einer grossen Menge von Daten auf einen Blick zu erkennen – die entsprechende Erfahrung für die korrekte Interpretation vorausgesetzt. Excelvis ist ein solches Visual Data Mining Tool, das von AWK entwickelt wurde und in Banken für verschiedene Analyse-Aufgaben eingesetzt wird.



Das Visual Data Mining Tool "Excelvis" von AWK

Beispiel: Graph-Analyse in der Betrugsbekämpfung

Traditionelle Datenanalysen fokussieren auf isoliert betrachtete Objekte bzw. Personen und ihre Eigenschaften (z.B. Kunden mit ihren Einkäufen oder Filialen mit ihren Umsätzen).

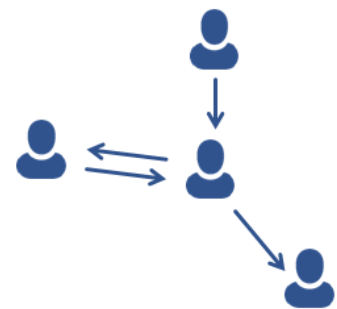
In unserer vernetzten Welt wird es immer zentraler, auch die Verknüpfungen zwischen Objekten bzw. Personen in der Analyse zu berücksichtigen (z.B. Transaktionen zwischen verschiedenen Kunden).

Insbesondere in der Betrugserkennung gewinnen sogenannte Graph-Analysen zunehmend an Bedeutung und halten in kommerziellen Tools Einzug. Diese Methoden werden auch als „Link“- oder „Network“-Analysis bezeichnet und helfen, Muster zu erkennen, die mit herkömmlichen Methoden unsichtbar bleiben.



Geschlecht: m
Alter: 42
Kundentyp: X

Traditioneller Ansatz: Kunde mit seinen Eigenschaften



*Graph-Ansatz: Kunden mit ihren Interaktionen
(z.B. Geldflüsse oder Telefonate)*

Interaktionsmuster machen verdächtiges Verhalten sichtbar, indem sie z.B. ungewöhnliche (indirekte) Verbindungen von verschiedenen Personen oder Organisationen über Zahlungsströme zeigen. Entsprechende Lösungen hat AWK im Bankenumfeld bereits umgesetzt.

Vorgehen in der Datenanalyse

Die verschiedenen Methoden der manuellen Datenanalyse können – richtig angewandt – oft schon in kurzer Zeit wertvolle Informationen zu Tage fördern. Je nach Ausgangssituation bringt aber erst der richtige Data Mining-Algorithmus die wirklich wertvollen Informationen zum Vorschein.

Ob manuell oder mit Hilfe von Data Mining Tools – Datenanalysen bieten viel Potenzial, sind aber auch ein anspruchsvolles Unterfangen. Fehlinterpretationen und andere Fallstricke können zu falschen Entscheidungen führen. Ein vertieftes Verständnis der theoretischen und statistischen Grundlagen der verschiedenen Methoden, gepaart mit langjähriger, pragmatischer Praxiserfahrung, hilft, den in den Daten verborgenen wertvollen Wissensschatz zu heben.

Das Vorgehensmodell von AWK für Datenanalysen legt grossen Wert auf das Verständnis der Rohdaten (Business Understanding), da dieses für die Gestaltung der Analyse und die richtige Interpretation zentral ist.

Die Rohdaten müssen für die Analyse aufbereitet und in eine geeignete Form gebracht werden. Je nach Format und Qualität der vorliegenden Daten macht dieser Schritt einen grossen Teil des Aufwandes aus. Effizienz dank Best Practices macht sich hier bezahlt. Eine stetige, kritische Plausibilisierung der Ergebnisse nach jedem Schritt sichert die Qualität und Fehlerfreiheit. Geeignete Visualisierungen helfen, die zentralen Erkenntnisse zielgruppengerecht zu kommunizieren.



AWK-Vorgehensmodell für die Datenanalyse

AWK führt Datenanalysen durch und bietet Unterstützung für die Evaluation und den optimalen Einsatz von Datenanalyzelösungen beim Kunden an. Unsere Erfahrung zeigt, dass oft kostspielige Tools im Einsatz sind, deren Potenzial aber nur teilweise genutzt wird. Quer durch alle Branchen haben wir festgestellt, dass moderne, kostengünstige Tools vielen Anwendern nicht bekannt sind oder nicht genutzt werden. Hier kann eine Schulung das entsprechende Know-how vermitteln und die Organisation befähigen, auch ohne kostspielige Tools mehr Erkenntnisse aus den Daten zu ziehen.

Die Autoren



Patrik Rüegge
Bereichsleiter,
Dipl. El.-Ing. ETH

Patrik Rüegge ist bei AWK verantwortlich für den Marktbereich Banken und Versicherungen



Dr. Jonas Dischl
Consultant,
Dr. Inform., Dipl. Inform.

Jonas Dischl ist bei AWK Experte im Bereich Datenanalyse und Betrugserkennung



Dr. Dimitrios Tombros
Bereichsleiter,
Dr. Inform., MScS,
lic. oec. publ.

Dimitrios Tombros ist bei AWK verantwortlich für den Kompetenzbereich Informationsmanagement

Über die AWK Group

AWK ist mit rund 130 Mitarbeitenden eines der grössten Schweizer Beratungsunternehmen für Informationstechnologie.

Wir sind schweizweit tätig mit Standorten in Zürich und Bern.

Unsere Dienstleistungen umfassen Consulting, Engineering und Projektmanagement.



AWK GROUP
Consulting | Engineering | Project Management