

Wirtschaftswissenschaftliches Zentrum der Universität Basel
Seminar zur Wirtschaftsinformatik – Sommersemester 2000

Data Mining unter dem Gesichtspunkt des Datenschutzes

Ein Data Mining ohne Daten ist unmöglich. Kann man
Daten trotzdem mittels Data Mining analysieren, wenn
die Daten geschützt sind?

Matthias Leisi leisi@astrum.ch

24. Mai 2000

Was ist Datenschutz?

- *Personendaten*: Alle Angaben zu einer *bestimmten* oder *bestimmbaren* Person.
- *besonders schützenswerte Daten*: Daten über religiöse, weltanschauliche, politische, . . . Ansichten; Intimsphäre, Gesundheit, . . .
- *Persönlichkeitsprofil*: Zusammenstellung wesentlicher Aspekte einer Persönlichkeit.
- *Bearbeiten*: Weit gefasster Begriff des gesamten Lebenszyklus' der Daten.
- *Zweckbindung*: Bearbeitung für den Zweck, der bei der Beschaffung angegeben wurde.

Für die Schweiz gilt das Datenschutzgesetz (DSG) von 1992 und die zugehörige Verordnung (VDSG). Für EU-Staaten gilt die EG-Datenschutzrichtlinie 95/46/EG von 1995.

Was fällt unter den Datenschutz?

Jede Datensammlung und -bearbeitung kann unter das Datenschutzgesetz fallen.

- Alle Personendaten, egal ob (in-) direkt erhoben oder mittels Data Mining analysiert.
- Strengere Regeln für besonders schützenswerte Daten und Persönlichkeitsprofile.

Ausnahmen:

- Bewusste Einwilligung der Betroffenen.
- Enger Zusammenhang mit Vertragshandlungen.
- Überwiegendes öffentliches oder privates Interesse.

Bei Persönlichkeitsprofilen oder "besonders schützenswerten Daten" werden strengere Maßstäbe an den Datenschutz angelegt. Beide Kategorien werden im Gesetz ungefähr gleich behandelt. Regelungen wie das Bankgeheimnis oder das Berufsgeheimnis von Ärzten gehen dem Datenschutz vor.

- Staatliche Behörden unterstehen ebenfalls dem DSG.

Zielkonflikt zwischen Data Mining und Datenschutz

- Ziel des Einsatzes von Data Mining-Methoden ist u.a. die Aufdeckung bisher unbekannter Zusammenhänge.
- Der Datenschutz bestimmt, dass Personendaten nur für den Zweck verwendet werden dürfen, der bei der Beschaffung angegeben wurde.

Implikationen:

- Die Kombination von einzeln harmlosen Daten kann datenschutzrechtlich relevant sein.
- Datensparsamkeit und Datenvermeidung als Grundprinzipien.
- Datenbedarf aus den Zielen des Data Mining ableiten.
- Betroffene müssen in die Anwendung von Data Mining-Methoden eingewilligt haben.

Datenerwerb und -beschaffung

- Einwilligung der Betroffenen, z.B. im Rahmen eines Vertragsabschlusses (AGB!).
- Zukauf externer Adressen (vertragliche Zusicherung datenschutzrechtlicher Unbedenklichkeit).
- Öffentliche Register und Verzeichnisse (Adressbücher, Handelsregister, . . .).
- Betroffene für die Preisgabe von Daten entschädigen.
- Rechtmässige Beschaffung: Treu und Glauben, Verhältnismässigkeit.

Datenbearbeitung und -richtigkeit

- Jeder Bearbeitungsschritt verändert die Qualität der Daten.
- Implikationen auch für Data Warehouse-Architekten.
- Implikationen für Qualitätssicherung, Auswahl der anzuwendenden Methoden und Bedeutung von Stichprobenkontrollen.
- trash in – trash out
- Gefahr der Verletzung von Persönlichkeitsrechten minimieren.
- Recht auf Richtigkeit respektive Feststellung der Unrichtigkeit.

Löschen Sie die nicht mehr benötigten Daten!
(EDSB, Leitfaden für Inhaber von Datensammlungen)

Datenhandel

- Bekanntgabe von "besonders schützenswerten Daten" nicht möglich.
- Bekanntgabe ins Ausland nur unter gewissen Bedingungen ("Safe Harbour").
- Einwilligung der Betroffenen.
- Eingeschränkte Outsourcing-Möglichkeiten.
- Problematische gemeinsame Daten-Pools von Unternehmen.
- Auch Data Mining-Resultatsanalysen sind Handelsobjekte.
- Private kommerzielle Verwertung von öffentlichen Registerdaten ist nicht zulässig, da die Zweckbindung verletzt wird.

Datensicherheit

Ohne Datensicherheit gibt es keinen Datenschutz.
(EDSB, Leitfaden zu den technischen und organisatorischen
Massnahmen des Datenschutzes)

- Schutz vor unbefugter Einsicht / unbefugtem Zugriff.
- Verschlüsselung.
- "Need to know"-Prinzip gegenüber Mitarbeitern.
- Protokollierung der Verwendung.
- Sicherstellung der Datenintegrität als Massnahme gegen datenschutz-relevante Unrichtigkeit der Daten.

Folgerungen für den praktischen Einsatz

Datenschutz bedeutet nicht, dass Unternehmen auf Marktforschungsmechanismen oder auf die Bildung von massgeschneiderten Marketingprofilen gänzlich verzichten müssen.

(EDSB, 6. Tätigkeitsbericht 1998/1999)

- Bewusstsein und Sensibilisierung für Datenschutz schaffen und fördern.
- Ausgebauter Datenschutz als vertrauensbildende Massnahme dem Kunden gegenüber
- Definition von Zielen beim Einsatz von Data Mining.
- Überprüfen der vorhandenen Datenbasis auf (Un-)Bedenklichkeit (v. a. Zweckbindung).
- Clustering et al. – Verzicht auf individuell zuordenbare Methoden, Einsatz von datenschutzfreundlichen Technologien.

Datenschutz ist kein Zustand, sondern ein *Prozess*. Die Sensibilisierung der Inhaber von Datensammlungen und das Nachdenken über den sinnvollen, verhältnismässigen Einsatz von Data Mining-Technologien sind dabei entscheidende Einflussfaktoren.

Weiterführende Fragen

- Selbstregulierung der Industrie: "Sorgfaltsvereinbarung", Normensetzung, Durchsetzung und Sanktionen. Welches Verhältnis von staatlicher zu privater Regulierung?
- Betrieb von Datenbanken: Datenschutzrechtliche Probleme beim Outsourcing?
- Welche zusätzlichen Erkenntnisse über Personen kann man mit Data Mining-Methoden gewinnen?
- Werbefinanzierte Internetdienstleistungen: ex- / implizite Einwilligung, Kontrollmöglichkeiten?
- Liste "200 Reichste Schweizer" – Ist das zulässig?
- "Online-Identität": Unter welchen Umständen sind Mailadressen besonders schützenswerte Daten?