



Universitätsbibliothek

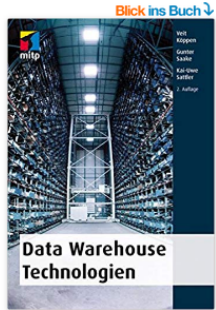
„Nutzer, die das Buch ausgeliehen haben, haben auch ...“
– Datengetriebene Empfehlungen für die Ausleihe

Veit Köppen, 18. März 2019, Leipzig

Überblick

- Motivation
- Datenschutz
- Ausleihverhalten & Empfehlung
- Data Mining
- System
- Evaluation & Ergebnisse

Motivation



Data Warehouse Technologien (mitp Professional) Broschiert – 26. Mai 2014

von Veit Köppen (Autor), Gunter Saake (Autor), Kai-Uwe Sattler (Autor)

Geben Sie die erste Bewertung für diesen Artikel ab

> Alle 2 Formate und Ausgaben anzeigen

Kindle Edition
EUR 25,99

Broschiert
EUR 29,99

Lesen Sie mit unserer **kostenfreien App**

Lieferung **Mittwoch, 1. Aug.**: Bestellen Sie innerhalb **1 Stunde und 26 Minuten** per **Premiumversand** an der Kasse. **Siehe Details.**

42 neu ab EUR 29,99 | 6 gebraucht ab EUR 21,00

Detailliert werden in diesem jetzt in einer aktualisierten Auflage vorliegenden Buch sowohl der Aufbau als auch die Nutzung von Data-Warehouse-Systemen beleuchtet. Dabei stehen Modellierungskonzepte und die Thematik der multidimensionalen Anfragen im Vordergrund. Zudem werden Interna wichtiger Systemlösungen von Oracle, IBM und Microsoft anhand zahlreicher Beispiele erläutert.



Dieses Bild anzeigen



Der Weltbestseller aus den USA

Basierend auf einer wahren Geschichte - der Roman "Unter blutrottem Himmel" von Bestsellerautor **Mark Sullivan** hier entdecken.

Hinweise und Aktionen

- Sie suchen preisreduzierte Fachbücher von Amazon Warehouse Deals? [Hier klicken.](#)
- Entdecken Sie die aktuellen **BILD Bestseller**. Jede Woche neu. [Hier klicken](#)

Wird oft zusammen gekauft



Gesamtpreis: **EUR 121,89**

Alle drei in den Einkaufswagen

Einige dieser Artikel sind schneller versandfertig als andere. [Details anzeigen](#)

- Dieser Artikel:** Data Warehouse Technologien (mitp Professional) von Veit Köppen Broschiert EUR 29,99
- Data-Warehouse-Systeme: Architektur, Entwicklung, Anwendung** von Andreas Bauer Gebundene Ausgabe EUR 49,90
- Data Warehouse Blueprints: Business Intelligence in der Praxis** von Dani Schneider Gebundene Ausgabe EUR 42,00

Kunden, die diesen Artikel gekauft haben, kauften auch



Data-Warehouse-Systeme: Architektur, Entwicklung, Anwendung
Andreas Bauer
★★★★★ 3
Gebundene Ausgabe
EUR 49,90 ✓prime



Data Warehouse Blueprints: Business Intelligence in der Praxis
Dani Schneider
Gebundene Ausgabe
EUR 42,00 ✓prime



Modellierung von Business-Intelligence-Systemen (Edition...
Michael Hahne
★★★★★ 1
Gebundene Ausgabe
EUR 69,90 ✓prime



HTML5 und CSS3: Das umfassende Handbuch zum Lernen und...
> Jürgen Wolf
★★★★★ 47
Gebundene Ausgabe
EUR 44,90 ✓prime



Der Oracle DBA: Handbuch für die Administration der Oracle Database 12c
Mirko Hotzy
★★★★★ 1
Gebundene Ausgabe
EUR 70,00 ✓prime

Gesponserte Produkte zu diesem Artikel [\(Was ist das?\)](#)

Suchen Suchergebnis Erweiterte Suche Zwischenablage Benutzerkonto Hilfe

suchen [und] [ALL] Alle Wörter sortiert nach Erscheinungsjahr unscharfe Suche

Universitätsbibliothek

Suchgeschichte Kurzliste Teildaten

Ihre Aktion

Titel: Data Warehouse Technologien / Veit Köppen, Gunter Saake, Kai-Uwe Sattler
Köppen, Veit, 1972.; Saake, Gunter, 1950.; Sattler, Kai-Uwe, 1958.
Ausgabe: Zweite Auflage
Sprache/n: Deutsch
Veröffentlichungsangabe: Heidelberg ; Hamburg ; München ; Landsberg, Frechen : mitp, Verlagsgruppe Hühig Jehle Rehm GmbH, 2014
Umfang: xiii, 337 Seiten : Illustrationen, Diagramme (schwarz-weiß) ; 24 cm
Literaturverzeichnis: Seite 319-337
Anmerkung: 3-8266-9485-6, 978-3-8266-9485-1 Pbk : 29,99 EUR
Schlagwörter: "Data-Warehouse-Konzept"
IS 51 Programmiermethodik
85.20 Betriebliche Information und Kommunikation
Dewey Dezimal-Klassifikation: (22/ger) 005.745 ; 004 ; 650
Regenburger Verbund-Klassifikation: 21.530
Inhaltsverzeichnis

Mehr zum Thema:
Mehr zum Titel:
Sachgebiete: Y.431.2

Standort: Freihand
Signatur: 2014.04597:1
Ausleihstatus: Ausleihbar
Bitte am Standort: Freihand entnehmen und vor Ort ausleihen

Standort: Freihand
Signatur: 2014.04597:2
Ausleihstatus: Ausleihbar
ausgeleihen bis 23-05-2018 [Vormerken](#)

Standort: Freihand
Signatur: 2014.04597:3
Ausleihstatus: Ausleihbar
Bitte am Standort: Freihand entnehmen und vor Ort ausleihen

Standort: Freihand
Signatur: 2014.04597:4
Ausleihstatus: Ausleihbar
Bitte am Standort: Freihand entnehmen und vor Ort ausleihen

Standort: Freihand
Signatur: 2014.04597:5
Ausleihstatus: Ausleihbar
ausgeleihen bis 15-05-2018 [Vormerken](#)

Standort: Freihand
Signatur: 2014.04597:6

Data Warehouse Technologien
Veit Köppen, Gunter Saake, Kai-Uwe Sattler
Zweite Auflage

1

Titel Ähnliche Publikationen

ID 780642740
Titel Data Warehouse Technologien
Verfasser Köppen, Veit % | Saake, Gunter % | Sattler, Kai-Uwe %
Format Buch
Ausgabe Zweite Auflage
Veröffentlicht Landsberg, Frechen : mitp, Verlagsgruppe Hühig Jehle Rehm GmbH %, 2014 %
Sprache Deutsch
Umfang xiii, 337 Seiten, Illustrationen, Diagramme (schwarz-weiß), 24 cm
Hinweise Literaturverzeichnis: Seite 319-337
ISBN 3826694856 (Pbk) | 3-8266-9485-6 | 9783826694851 (Pbk) | 978-3-8266-9485-1
Zusatzinformationen Data-Warehouse-Konzept %
Klassifikation 54.51 – Programmiermethodik – 85.20 – Betriebliche Information und Kommunikation – ST 530
Quelle Gemeinsamer Verbundkatalog / CVK
Sachgebiet Y.431.2 %

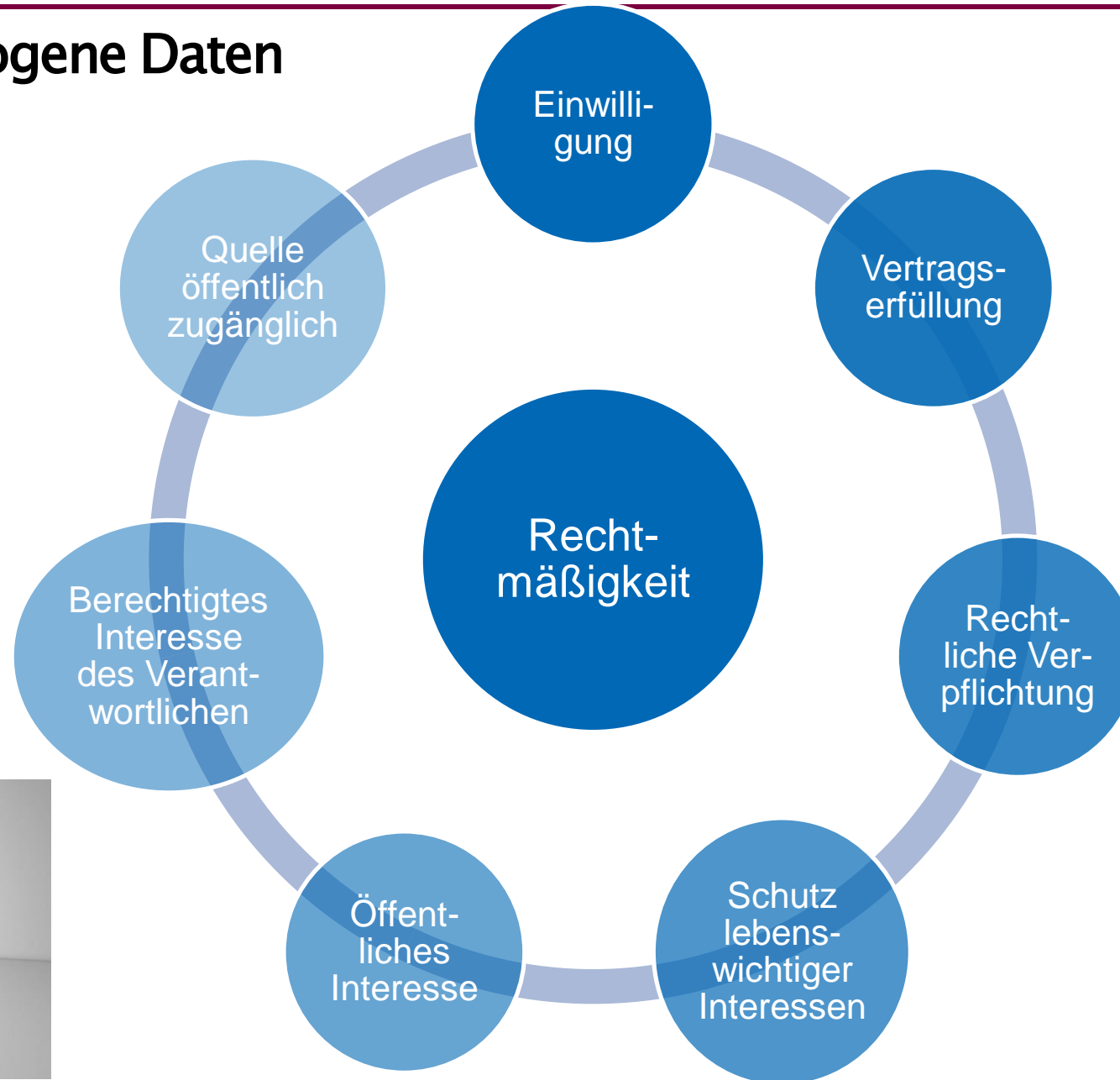
Exemplare

Freihand Signatur 2014.04597:1 Ausleihbar	Freihand Signatur 2014.04597:2 Ausgeleihen bis 23.05.2018	Freihand Signatur 2014.04597:3 Ausleihbar
Freihand Signatur 2014.04597:4 Ausleihbar	Freihand Signatur 2014.04597:5 Ausgeleihen bis 15.05.2018	Freihand Signatur 2014.04597:6 Ausleihbar
Freihand Signatur 2014.04597:7 Semesterapparat	Freihand Signatur 2014.04597:8 Ausleihbar	Freihand Signatur 2014.04597:9 Ausgeleihen bis 17.05.2018
Freihand Signatur 2014.04597:10 Ausleihbar		

← →

↑ Zurück QR-Code Druck Mail Export Merkleiste

Personenbezogene Daten



Datenschutzgrundsätze der DSGVO

Datenschutzgrundsatz	Was bedeutet das?
Verbot mit Erlaubnisvorbehalt, Zweckbindung	Datenverarbeitung (DV) nur nach Einwilligung oder mit eindeutig vorher definierten legitimen Zwecken
Transparenzgrundsatz	Informationen zu DV-Prozessen, eingesetzten Technologien und Datenströmen
Datenminimierung	Nur die Daten erheben, die zur DV benötigt werden
Richtigkeit und auf neuestem Stand	Daten sollten aktuell, richtig und abänderbar sein
Speicherbegrenzung	Daten nur so lange speichern, wie sie für die DV gebraucht sind
Integrität	Daten sollen korrekt (unversehrt) sein und durch korrekt funktionierende Systeme verarbeitet werden.
Datensicherheit	Legitimität, Vertraulichkeit und Verfügbarkeit muss gewährleistet sein
Rechenschaftspflicht	Nachweisfähigkeit über Einhaltung der Vorgaben

Personenbezogene Daten?

Persönliche Verhältnisse	Sachliche Verhältnisse	Verhalten
<ul style="list-style-type: none"> • Name • Alter / Geburtsdatum • Ausbildung, Beruf • Adresse • Telefonnummer • Systembenutzerdaten • Familienstand • Interessen 	<ul style="list-style-type: none"> • KFZ-Kennzeichen • Gehaltsgruppe • Kreditdaten • Kreditwürdigkeit • Steuern • Mietverhältnis • Vertragsbeziehungen 	<ul style="list-style-type: none"> • Kaufverhalten • Logdaten <ul style="list-style-type: none"> • Internet • Telefonie • Aufenthaltsdaten (Geo-Daten) • Mail-Informationen

Besonders schützenswert:

- Gewerkschaftszugehörigkeit
- Religion
- Gesundheitsdaten
- Sexuelleben

Identifizierbar: Identität einer Person ist durch Kombination mehrerer Daten feststellbar

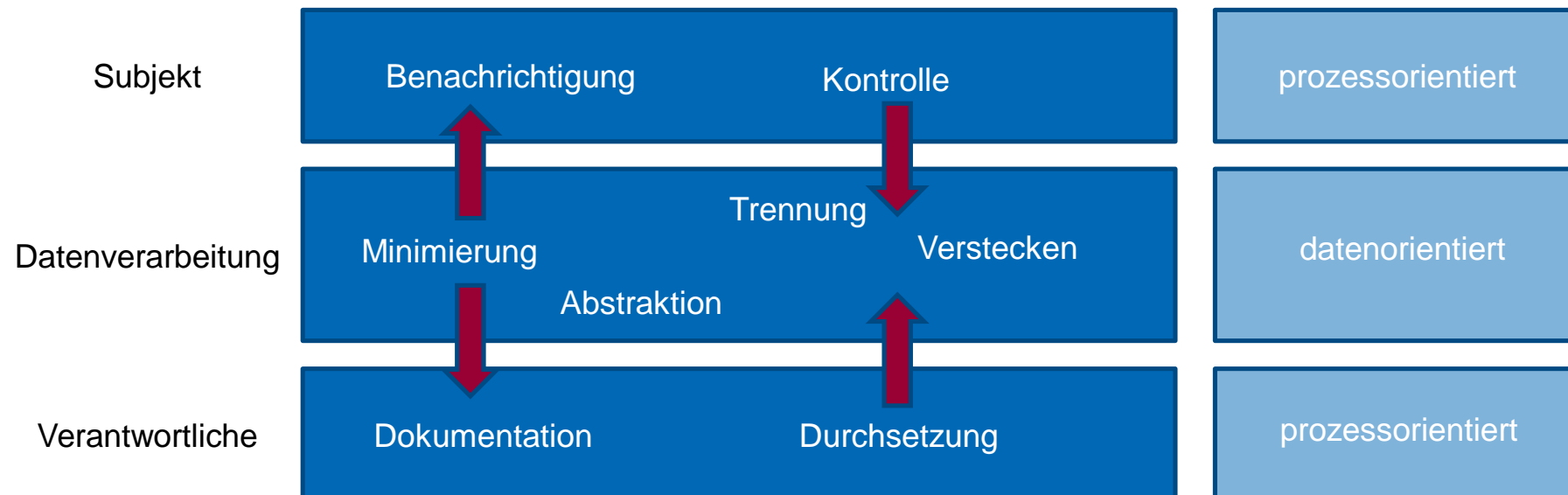
Wie viel Prozent der Amerikaner lassen sich durch Geschlecht, Geburtsdatum und Postleitzahl identifizieren?



www.menti.com

Privacy by Design

- “Datenschutz durch Technikgestaltung”
- In der DSGVO gefordert!
- Frühzeitiges Ergreifen von technischen und organisatorischen Maßnahmen



Privacy by Design ist ein Prozess über alle Phasen der Entwicklung – ein ganzheitlicher Ansatz!

Ausgangssituation

Lösungen für die Warenkorbanalyse vorhanden:

Exemplarisch: BibTip

- Nutzung des Suchverhaltens im OPAC
- Austausch über unterschiedliche Kataloge möglich
- Suche \neq Ausleihe
- Keine Berücksichtigung unseres Discovery-Systems

Empfehlungssystem für wiss. Bibliotheken (z.B. Huaqiao University (China))

Informationen zur Ausleihe vorhanden

Anonymisierung notwendig für Data Privacy Anforderungen

→ Keine personenbezogenen Daten

Fragen: Parametrisierung & Vergessen von Daten

Ausgangslage

UB Magdeburg

1.220.000 ausleihbare Exemplare

ca. 980.000 Suchanfragen 2017

200.121 Ausleihen (DBS, 2017)

21.776 Nutzer (DBS, 2017)



Pica-LBS, d.h. „Vergessen“ des Nutzers bei Rückgabe

OPAC & Lukida-Discoverysystem

Technologie-Stack: SQL, Java, R, PHP

Hardware: Kein Datentransfer → Arbeiten auf dem datenhaltenden System

Wie können Data-Privacy-Anforderungen für ein Buchausleih-Empfehlungssystem umgesetzt werden?

Vorgehen

1. Datenaufbereitung

Privacy mittels Hash-Verfahren im Warenkorb (SHA-256/512)

SQL (vom LBS) & Java für Erzeugung der Warenkörbe (Flat-Files)



2. Data Mining (Mustererkennung):

Assoziationsverfahren: Apriori-Algorithmus (R, arules)

Regelaufbereitung in Java

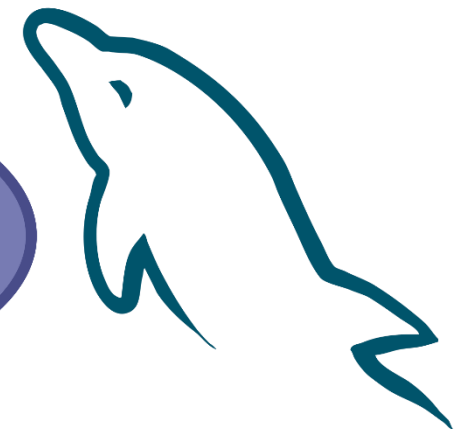


3. Präsentation

Datenbanksystem inkl. Index

Integration mittels iFrame im OPAC (PHP)

oder im Discoverysystem



Wo kann die Empfehlung gegeben werden?

Suchen Suchergebnis Erweiterte Suche Zwischenablage Benutzerkonto Hilfe

Universitätsbibliothek

suchen [und] [ALL] Alle Wörter sortiert nach Erscheinungsjahr unscharfe Suche

Suchen

Katalogmenü Abmelden

Online-Katalog der Universitätsbibliothek Magdeburg

- Bücher, Zeitschriften und andere Medien der UB
Bei Titeln bis 1990, die Sie nicht im OPAC finden,
benutzen Sie bitte auch die "Digitalisierten Zettelkataloge"
zur Absicherung Ihrer Recherche
- Zugriff auf elektronische Zeitschriften über die EZB

Leihfrist verlängern? Konto abfragen?
 Klicken Sie im Katalog oben auf "Benutzerkonto"

Sie haben Fragen zur Katalognutzung?
 Lernen Sie [online](#) die Benutzung kennen.

https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63884-p0011-7-0001-0

Suchen Suchergebnis Erweiterte Suche Zwischenablage Benutzerkonto Hilfe

Universitätsbibliothek

suchen [und] [ALL] Alle Wörter sortiert nach Erscheinungsjahr unscharfe Suche

Suchen

Ihre Aktion

Titel: [Data Warehouse Technologien / Veit Köpper, Gunter Saake, Kai-Uwe Sattler](#)
 Erscheinungsjahr: 2014
 Ausgabe: Zweite Auflage
 Sprache: Deutsch
 Veröffentlichungsort: Heidelberg : Heidelberg : München : Landberg, Frechen : mitp, Verlagsgruppe Hühig Jehle Rehm GmbH, 2014
 Umfang: xiv, 337 Seiten : Illustrationen, Diagramme (Schwarz-weiß) : 24 cm
 Literaturverzeichnis: Seite 319-337
 ISBN: 3-8266-9485-6 ; 3-8266-9485-1* Pbk. : 29,99 EUR
 Schlagwörter: [Data-Warehouse-Konzept](#)
 Sachgebiete: [IS-23 Betriebsliche Informationen und Kommunikation](#)
 Demy Decimal Klassifikation: (22)ger 004.751 : 004 : 630
 Regensburger Verbundklassifikation: ST 530
[Info/Wissensbereiche](#)

Sachgebiete: [Y.431.2](#)

Standort: [Freihand](#)
 Signatur: [2014.04597:1](#)
 Ausleihstatus: [Ausleihbar](#)
 Bitte am Standort: Freihand entnehmen und vor Ort ausleihen

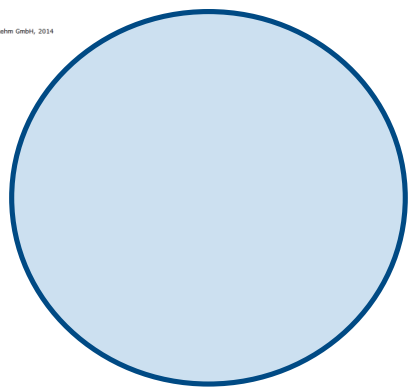
Standort: [Freihand](#)
 Signatur: [2014.04597:2](#)
 Ausleihstatus: [Ausleihbar](#)
 ausgeliehen bis 23-05-2018 [Anmelden](#)

Standort: [Freihand](#)
 Signatur: [2014.04597:3](#)
 Ausleihstatus: [Ausleihbar](#)
 Bitte am Standort: Freihand entnehmen und vor Ort ausleihen

Standort: [Freihand](#)
 Signatur: [2014.04597:4](#)
 Ausleihstatus: [Ausleihbar](#)
 Bitte am Standort: Freihand entnehmen und vor Ort ausleihen

Standort: [Freihand](#)
 Signatur: [2014.04597:5](#)
 Ausleihstatus: [Ausleihbar](#)
 ausgeliehen bis 15-05-2018 [Anmelden](#)

Standort: [Freihand](#)
 Signatur: [2014.04597:6](#)



https://ubfind.evgu.de

UBfind Ihre Suche...

Willkommen in der Universitätsbibliothek Magdeburg!

Suche im Bestand der UB und zugleich

- in den Beständen von über 400 Bibliotheken Nord- und Mitteldeutschlands,
- im Bestand der Bibliotheken der Stiftung Preußischer Kulturbesitz,
- in mehr als 57 Mio. gedruckten Artikeln (Online Contents und Directory of Open Access Journals),
- in mehr als 23 Mio. Objekten und Nachweisen aus den Nationalbibliotheken,
- in mehr als 11 Mio. elektronischen Ressourcen (eBooks, eArticle, eJournals).

Mehr Informationen auf Info-Seite [UBfind](#)

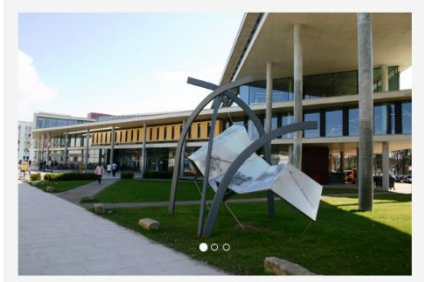
Sie können hier

- Ihre Medien finden und den Standort der Exemplare einsehen
- Ihre Medien zur Merkliste hinzufügen
- Ihre Merkliste einsehen
- Ihre Online-Medien direkt öffnen
- Ihre Suche speichern und später wieder ausführen (Anmeldung erforderlich)
- Ihre Medien vormerken (Anmeldung erforderlich)
- Ihre Medien bestellen (Anmeldung erforderlich)
- Ihre Medien verlängern (Anmeldung erforderlich)
- Ihre Gebühren einsehen (Anmeldung erforderlich)
- Ihr Passwort ändern (Anmeldung erforderlich)

Aktuelle Informationen der Bibliothek

Die Homepage der Bibliothek finden Sie hier

Impressum · Lukida 45.06.58 © 201



Co

Data Warehouse Technologien
 Veit Köpper, Gunter Saake, Kai-Uwe Sattler
 Zweite Auflage

1

Titel Ähnliche Publikationen

ID: 780642740

Titel: [Data Warehouse Technologien](#)
 Verfasser: [Köpper, Veit % | Saake, Gunter % | Sattler, Kai-Uwe %](#)
 Format: Buch
 Ausgabe: Zweite Auflage
 Veröffentlicht: [Landsberg, Frechen : mitp, Verlagsgruppe Hühig Jehle Rehm GmbH % 2014 %](#)
 Sprache: Deutsch
 Umfang: [xiv, 337 Seiten, Illustrationen, Diagramme \(Schwarz-weiß\), 24 cm](#)
 Hinweise: [Literaturverzeichnis: Seite 319-337](#)
 ISBN: [3826694856 \(Pbk\) | 3-8266-9485-6 | 9783826694851 \(Pbk\) | 978-3-8266-9485-1](#)
 Zusatzinformationen
 Schlagwort: [Data-Warehouse-Konzept %](#)
 Klassifikation: [54.51 - Programmiermethodik - 85.20 - Betriebliche Information und Kommunikation - ST 530](#)
 Quelle: [Gemeinsamer Verbundkatalog / CVK](#)
 Sachgebiet: [Y.431.2 %](#)

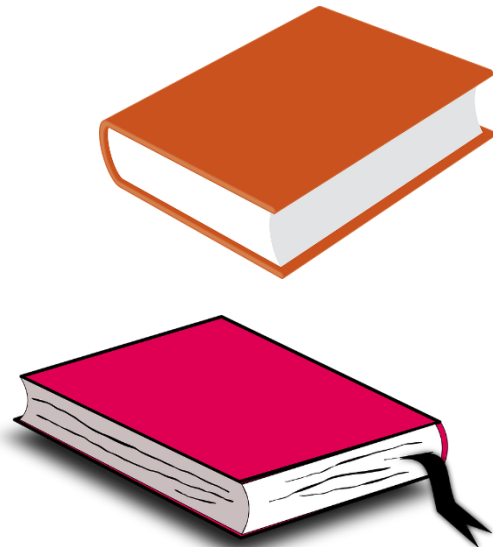
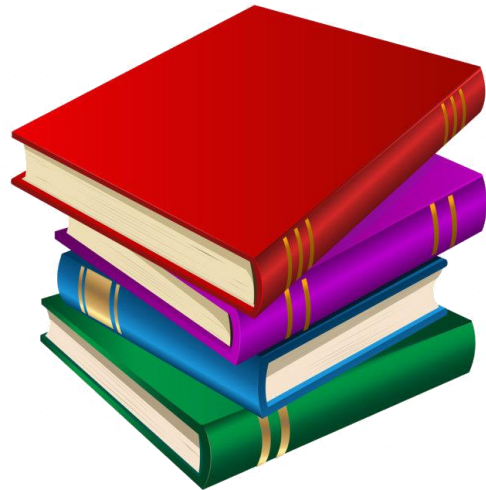
15
7

Freihand Signatur: 2014.04597:1 Ausleihbar	Freihand Signatur: 2014.04597:2 Ausgeliehen bis 23.05.2018	Freihand Signatur: 2014.04597:3 Ausleihbar
Freihand Signatur: 2014.04597:4 Ausleihbar	Freihand Signatur: 2014.04597:5 Ausgeliehen bis 15.05.2018	Freihand Signatur: 2014.04597:6 Ausleihbar
Freihand Signatur: 2014.04597:7 Semesterapparat	Freihand Signatur: 2014.04597:8 Ausleihbar	Freihand Signatur: 2014.04597:9 Ausgeliehen bis 17.05.2018
Freihand Signatur: 2014.04597:10 Ausleihbar		

← → ^ Zurück QR-Code Druck Mail Export Merkliste

Apriori Algorithmus

- Entwickelt von Agrawal und Srikant (1994)
- Transaktionen innerhalb eines Warenkorbes
- Fragestellung: Welche Dinge sind gemeinsam im Warenkorb vorhanden
- Betrachtung von Auftrittshäufigkeit (in allen und in den gefilterten Warenkörben)
- Zahlreiche Weiterentwicklungen der Algorithmen, effizienter Zugriff



Beispiel

Transaktion	Bücher
T1	Der Fremde, Der Prozess, Alkohol
T2	Der Fremde, Der Prozess, Paroles
T3	Der Fremde, 1984
T4	Der Prozess, Alkohol
T5	Der Fremde, Der Prozess, Alkohol, Paroles
T6	1984
T7	Ulysses, Der Prozess, Alibi
T8	Ulysses, Nadja, Lolita
T9	Lolita, Alkohol

Parameterfestlegung
 Minimaler Support: 2
 Minimale Konfidenz: 75%



Mentimeter: Für wie viele Bücher können wir eine Empfehlung geben?



www.menti.com

Beispiel

Transaktion	Bücher
T1	Der Fremde, Der Prozess, Alkohol
T2	Der Fremde, Der Prozess, Paroles
T3	Der Fremde, 1984
T4	Der Prozess, Alkohol
T5	Der Fremde, Der Prozess, Alkohol, Paroles
T6	1984
T7	Ulysses, Der Prozess, Alibi
T8	Ulysses, Nadja, Lolita
T9	Lolita, Alkohol

Büchermenge	Support
{Der Prozess}	5
{Der Fremde}	4
{Alkohol}	4
{Paroles}	2
{1984}	2
{Ulysses}	2
{Lolita}	2
{Alibi}	1
{Nadja}	1

Parameterfestlegung
 Minimaler Support: 2
 Minimale Konfidenz: 75%



Beispiel (weiter)

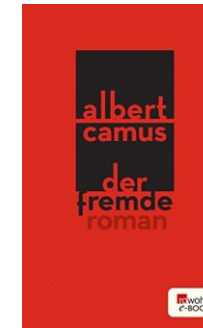
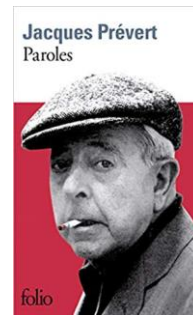
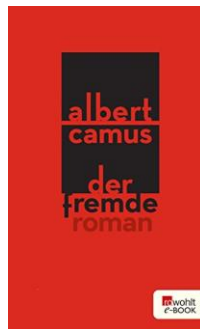
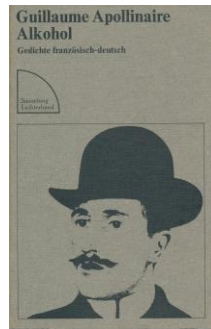
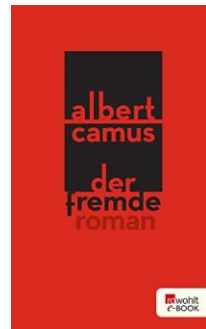
Kombinationen aus den verbliebenen Kandidaten

Büchermenge	Support
{Der Prozess, Der Fremde}	3
{Der Prozess, Alkohol}	3
{Der Prozess, Paroles}	2
{Der Prozess, 1984}	0
{Der Prozess, Ulysses}	1
{Der Prozess, Lolita}	0
{Der Fremde, Alkohol}	2
{Der Fremde, Paroles}	2
{Der Fremde, 1984}	1
{Der Fremde, Ulysses}	0
{Der Fremde, Lolita}	0

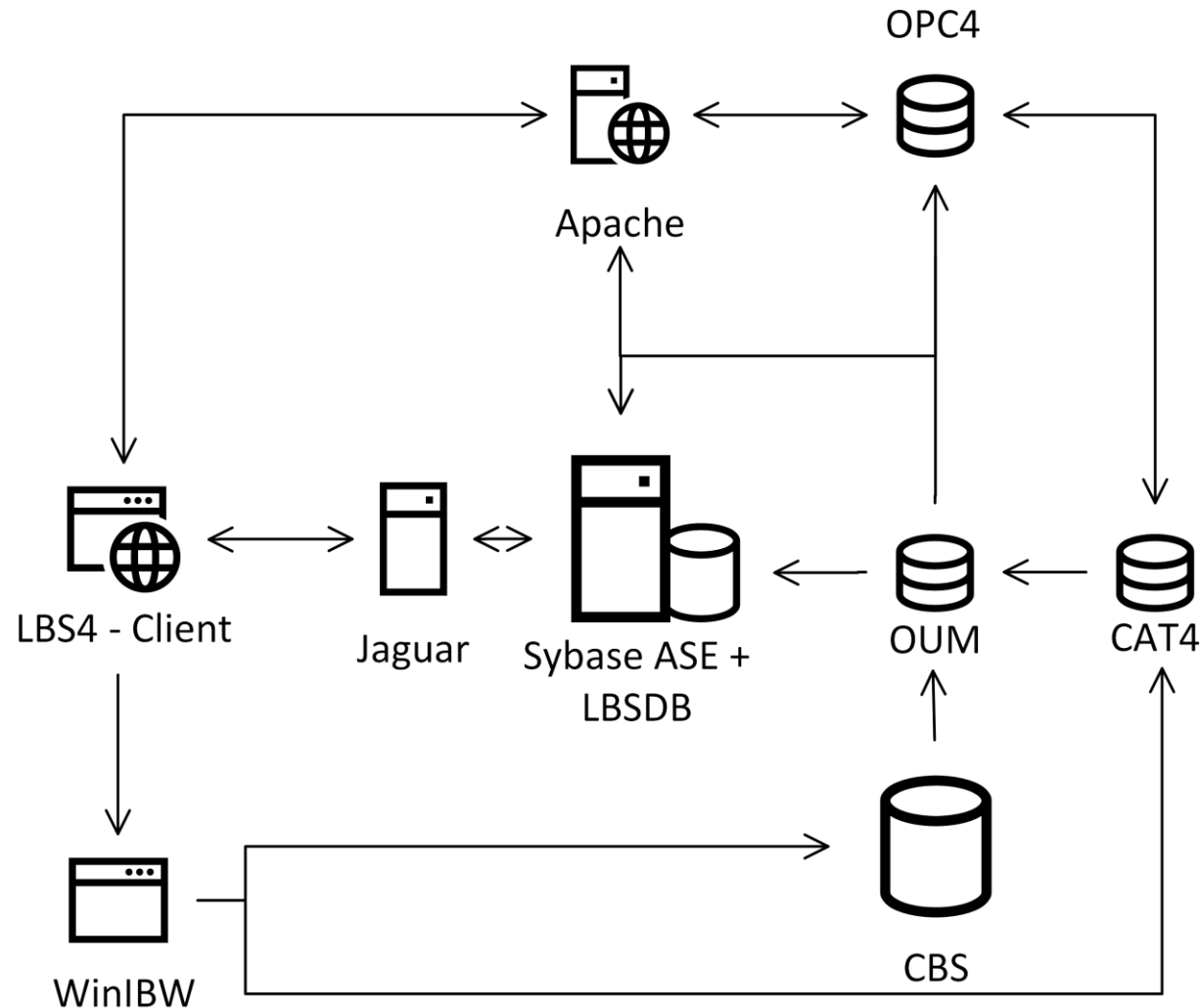
Büchermenge	Support
{Alkohol, Paroles}	1
{Alkohol, 1984}	0
{Alkohol, Ulysses}	0
{Alkohol, Lolita}	1
{Paroles, 1984}	0
{Paroles, Ulysses}	0
{Paroles, Lolita}	0
{1984, Ulysses}	0
{1984, Lolita}	0
{Ulysses, Lolita}	1

Beispiel

Büchermenge	Support
{Der Fremde, Der Prozess, Alkohol}	2
{Der Fremde, Der Prozess, Paroles}	2



Systemarchitektur: LBS



Betriebssystem: Unix-basiert

Datenbank: Sybase
SQL-Zugriff

Anonymisierung mittels Java

Data Mining: R mit Paket arules

Nachbereitung mittels Java

Repräsentation: Apache-Server

MySQL-Datenbank

Anonymisierung: Kryptologische
Hashfunktionen

Re-Identifikation verhindern:

Support

nur Regel-Teilmenge in DB

randomisierte Teilausgabe

Data Mining

Apriori: R-Paket arules nach Borgelt und Kruse (2002)

```
1 require(arules)
2 info <- read.table(paste0(dateiname, "_trans_items.csv"), sep = ",", header = T)
3 supportNR <- 2
4 trans_data <- read.transactions(file = paste0(dateiname, "_INT-baskets.csv"), format='
   basket', sep=',', cols=NULL)
5 rules = apriori(trans_data, parameter=list(support=supportNR / info$transactionNR,
   confidence=1/info$transactionNR, maxlen=2, minlen=2))
6 write(rules, file = paste0(dateiname, "_", supportNR, "_rules.csv"), quote=TRUE, sep = "
   ", col.names = NA)
```

```
1 " ", "rules", "support", "confidence", "lift", "count"
2 "1", "{6442} => {6443}", 0.000687757909215956, 1, 1246.28571428571, 6
3 "2", "{6443} => {6442}", 0.000687757909215956, 0.857142857142857, 1246.28571428571, 6
4 "3", "{23485} => {8142}", 0.000687757909215956, 0.857142857142857, 115.041758241758, 6
5 "4", "{8142} => {23485}", 0.000687757909215956, 0.0923076923076923, 115.041758241758, 6
6 "5", "{8641} => {7370}", 0.000802384227418615, 0.538461538461538, 151.533498759305, 7
7 "6", "{7370} => {8641}", 0.000802384227418615, 0.225806451612903, 151.533498759305, 7
8 "7", "{3722} => {3723}", 0.000687757909215956, 0.666666666666667, 290.8, 6
9 "8", "{3723} => {3722}", 0.000687757909215956, 0.3, 290.8, 6
```

Webpräsentation: Beispiel OPAC

<iframe src="https://<%variable(SERVER)>/empfehlung.php?PPN=<%variable(PPN)>" >

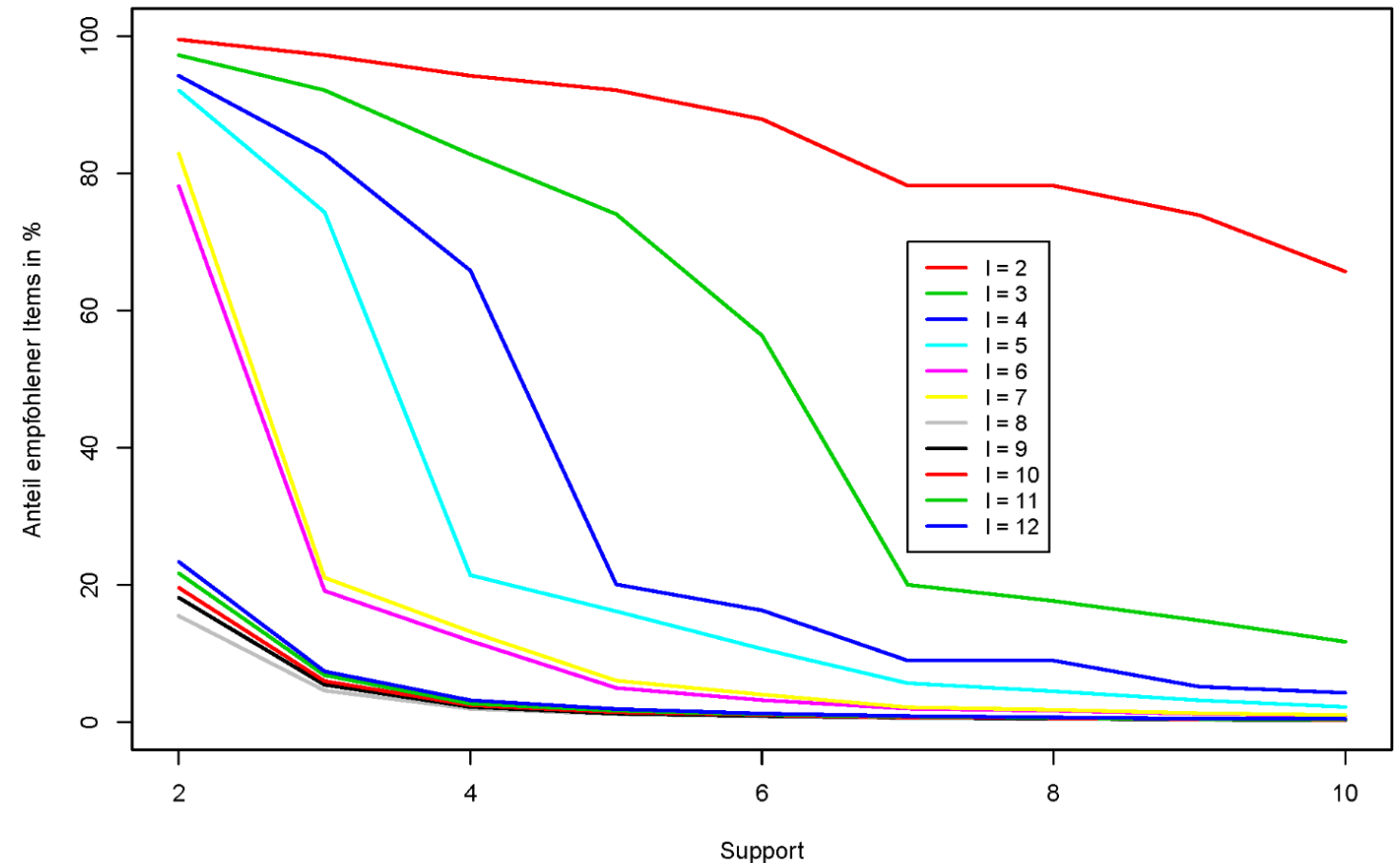
```
1 <?php
2 require_once ('DB-Konfiguration.php');
3 $db_link = mysqli_connect (
4     MYSQL_HOST,
5     MYSQL_BENUTZER,
6     MYSQL_KENNWORT,
7     MYSQL_DATENBANK
8 );
9 $ppn = $_GET['PPN'];
10 echo "Ergebnisempfehlungen f&uuml;r " . $ppn;
11
12
13 $sqlAnfrage = "SELECT Rumpf FROM 'RegelImport' WHERE KOPF = " . $ppn . " ORDER BY RAND()
14     LIMIT 3";
15
16 $dbErgebnisrg = mysqli_query( $db_link , $sqlAnfrage );
17 if ( ! $dbErgebnis )
18 {
19     die('Ung&uuml;ltige Abfrage: ' . mysqli_error());
20 }
21 echo '<table border="1">';
22 while ($zeile = mysqli_fetch_array( $dbErgebnis))
23 {
24     echo "<tr>";
25     echo "<td>". $zeile['Rumpf'] . "</td>";
26     echo "<td> <a href=\"https://[REDACTED]/CMD?ACT=SRCHA&IKT=1016&TRM=ppn+\" . $
27         zeile['Rumpf'] . \"%3F\"> zu f&uuml;hrende Empfehlung </a></td>";
28 }
29 echo "</table>";
30
31 mysqli_free_result( $db_erg );
32 ?>
```

Evaluation: Vorhersage

Datenbestand Februar 2017 bis Februar 2018,
Vorhersage aller Buchausleihen für März 2018

Intervalllänge I:
Wie viele Monate bilden einen
Warenkorb?

Abdeckung der März-Items



Zusammenfassung und Ausblick

- Ausleihdaten im System vorhanden
 - Empfehlung auf Ausleihen möglich
 - Data Privacy by Design Ansatz
 - Technologie-Stack erforderlich
- Parametertuning
 - Andere Algorithmen:
 - Data Mining Fragestellung
 - Data Privacy Berücksichtigung
 - Automatisierung
 - Vergessen von Nutzerdaten
 - Berücksichtigung
 - neuer Bücher
 - Hierarchieebenen

Fragen?

www.menti.com

Literatur & Referenzen

- C. Borgelt und R. Kruse. Induction of association rules: Apriori implementation. In W. Härdle und B. Rönz, Hrsg., *Proceedings in Computational Statistics*, Seiten 395–400. Springer, 2002.
- BSI. Kryptographische Verfahren: Empfehlungen und Schlüssellängen. Technical Report TR–02102–1, Bundesamt für Sicherheit in der Informationstechnik, Februar 2017.
- C. Dwork. Differential privacy. In M. Bugliesi, B. Preneel, V. Sassone, und I. Wegener, Hrsg., *Automata, Languages and Programming*, Seiten 1–12, Berlin, Heidelberg, 2006. Springer.
- S. Fu, Y. Zhang, und S. Minn. On the recommender system for university library. In *International Association for Development of the Information Society*, Seiten 215–222. ERIC, 2013.
- M. Hahsler, B. Grün, und K. Hornik. A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software*, 14(15): 1–25, 2005.
- M. Mönnich und M. Spiering. Einsatz von Bibtip als Recommendersystem im Bibliothekskatalog. *Bibliotheksdienst*, 42(1):54–59, 2008.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL <https://www.R-project.org/>.
- L. Sweeney. Simple demographics often identify people uniquely. Technical Report 3, Carnegie Mellon University, Pittsburgh, 2000.
- L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.

Dr. Veit Köppen

Abteilungsleiter IT-Anwendungen

Otto-von-Guericke Universität Magdeburg

Universitätsbibliothek

Anschrift: Universitätsplatz 2

39106 Magdeburg

E-Mail: veit.koepfen@ovgu.de

