
Non-Standard-Datenbanken

Prof. Dr. Ralf Möller

Universität zu Lübeck

Institut für Informationssysteme

Felix Kuhr (Übungen)



Organisatorisches: Übungen

- **Start:** Montag, 16. Oktober 2017
- **Zwei Übungen:** ab 20.10. Fr. 8-9 Uhr, 9-10 Uhr , IFIS, Geb. 64, Raum 2035
(Anmeldung über Moodle nach dieser Veranstaltung)
 - Erste Ausgabe nach dieser Vorlesung
 - Erste Abgabe am 23.10. (siehe unten)
- **Übungsaufgaben** stehen jeweils nach der Vorlesung ca. ab 18 Uhr über Moodle bereit
- **Abgabe der Lösungen** erfolgt bis Montag 14 Uhr in der IFIS-Teeküche (jeweils in der Woche nach der Ausgabe, 1 Kasten pro Übungsgruppe)
- Aufgaben können in einer **2-er Gruppe** bearbeitet werden (also bitte Name(n), Matrikelnummer(n) und Übungsgruppennummer vermerken)
- In den Übungen am Freitag wird der Übungszettel besprochen, dessen Lösungen bis zum jeweils vorigen Montag abgegeben werden, und auch Fragen zum jeweils neuen Übungszettel geklärt (ggf. mit Präsenzaufgaben als Hilfestellung)

Organisatorisches: Prüfung

- Die **Eintragung in den Kurs** und in eine Übungsgruppe ist **Voraussetzung**, um an dem Modul Non-Standard-Datenbanken teilnehmen zu können
- Am Ende des Semesters findet eine **Klausur** statt
- **Voraussetzung** zur Teilnahme an der Klausur sind mindestens **50% der gesamtmöglichen Punkte aller Übungszettel**

Teilnehmerkreis und Voraussetzungen

Studiengänge

- Bachelor **Medizinische Informatik**
- Bachelor **Medieninformatik**
- Bachelor **Informatik**
- Bachelor **Mathematik in Medizin und Lebenswissenschaften**
- Master **Informatik**
- Master **MML**

Voraussetzungen (Bestehen der unten genannten Veranstaltungen ist keine Teilnahmebedingung für NDB-Klausur)

- Algorithmen und Datenstrukturen
- Lineare Algebra und Diskrete Strukturen 1
- Datenbanken
- Theoretische Informatik (kontextfreie Grammatiken)

Vorteilhaft

- Einführung in die Logik

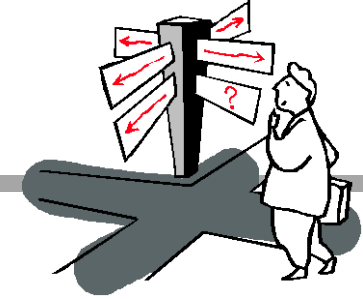
Was sind Merkmale von
Standard-Datenbanken?



Merkmale von Standard-Datenbanken

- Datenmodell: relational („Tabellen“ und „Tupel“)
- Annahmen:
 - Strukturen fix (Schemaänderung möglich aber aufwendig)
 - Verwerfen d. A. führt zu semistrukturierten Datenbanken und zu Graphdatenbanken
 - Abstrakte Entitäten mit abstr. Assoziationen (Tupel in Relationen)
 - Verwerfen führt zu temporalen, sequenzorientierten, räumlichen, und multimodalen¹ Datenbanken
 - Daten persistent, Anfragen bzgl. Schnappschuss beantwortet
 - Verwerfen führt zu stromorientierten Datenbanken
 - Daten enthalten feste Werte bzw. Referenzen
 - Verwerfen führt zu Datenbanken für unsichere und unvollständige Information

Inhalt der Vorlesung

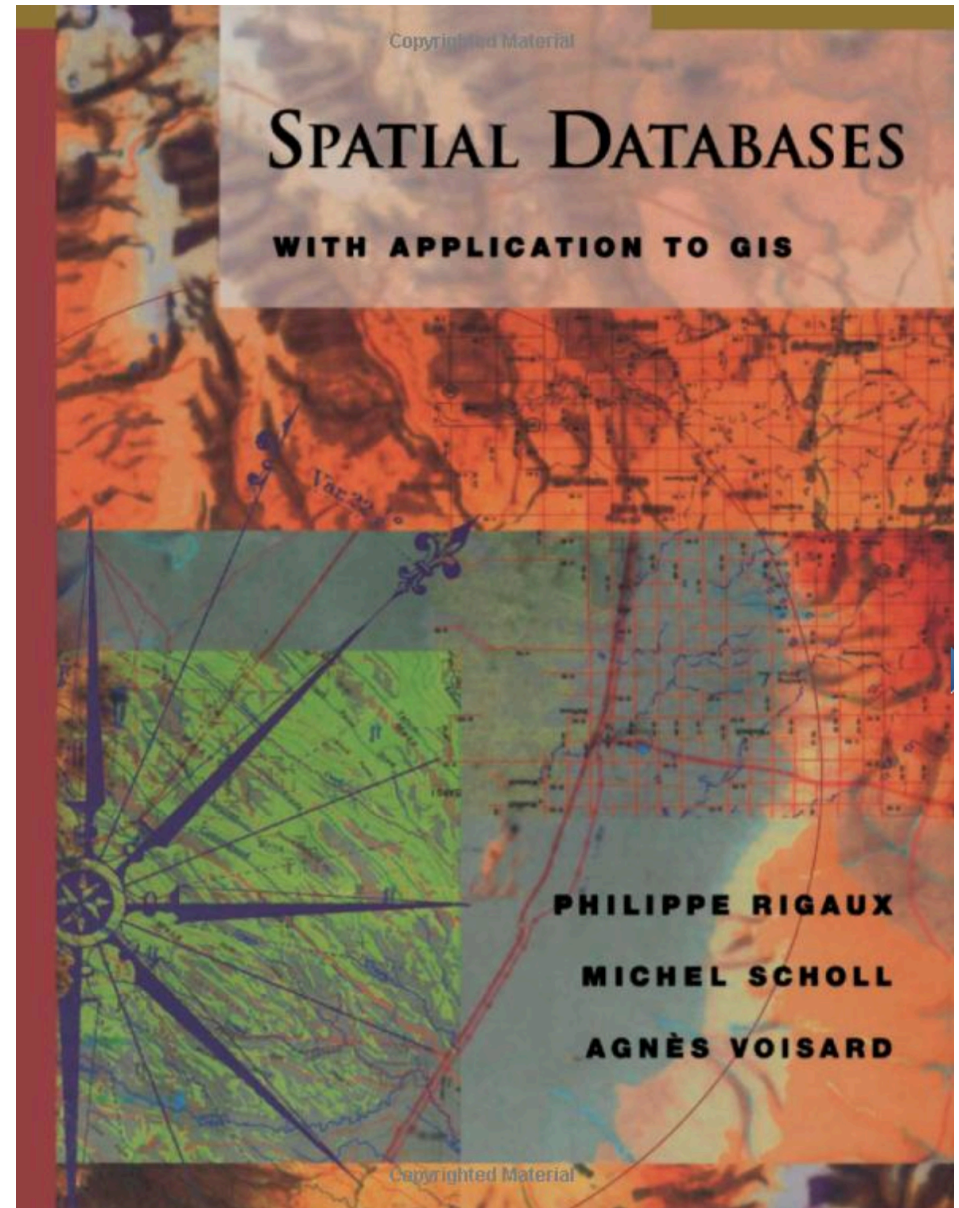
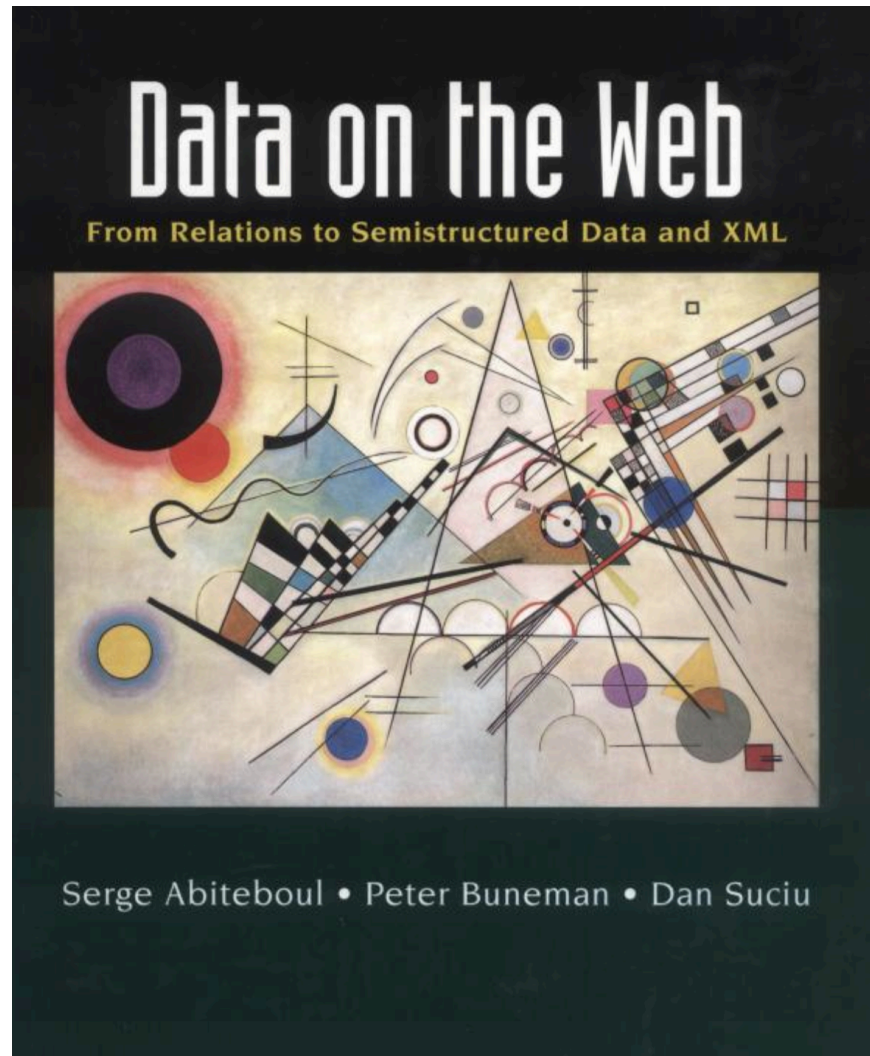


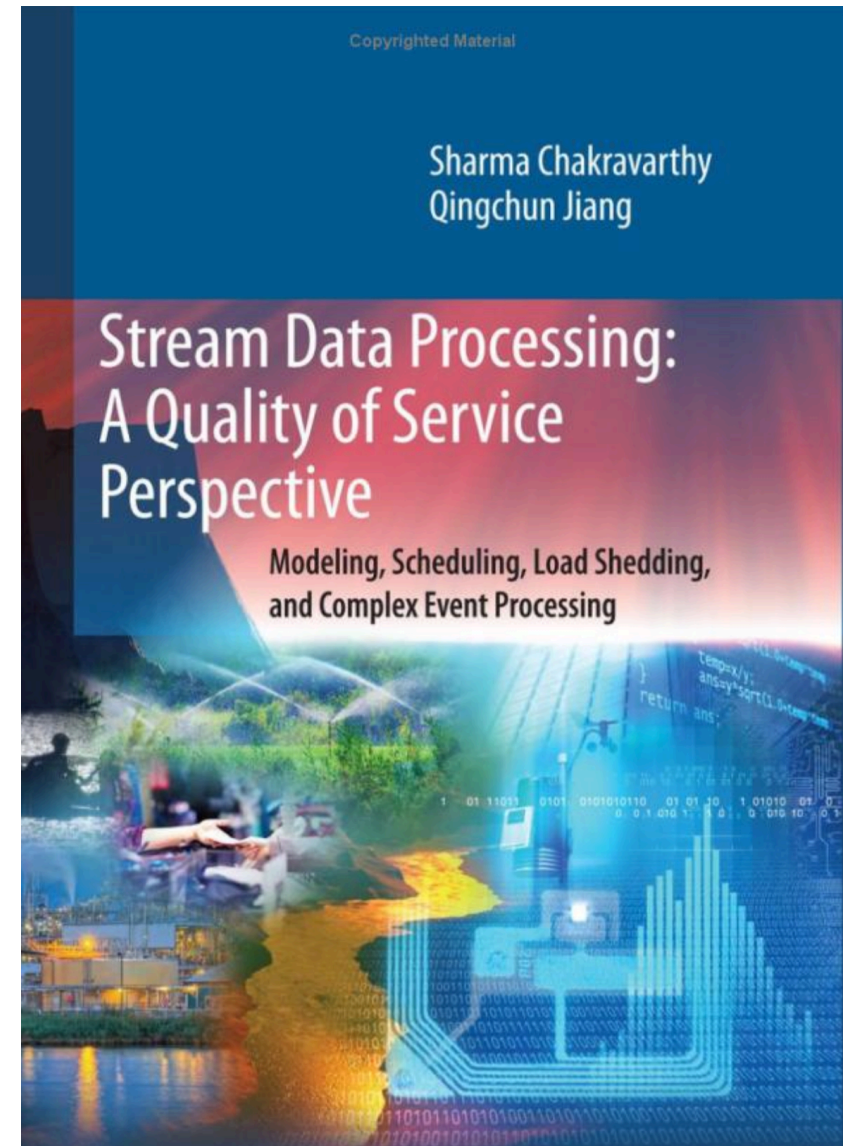
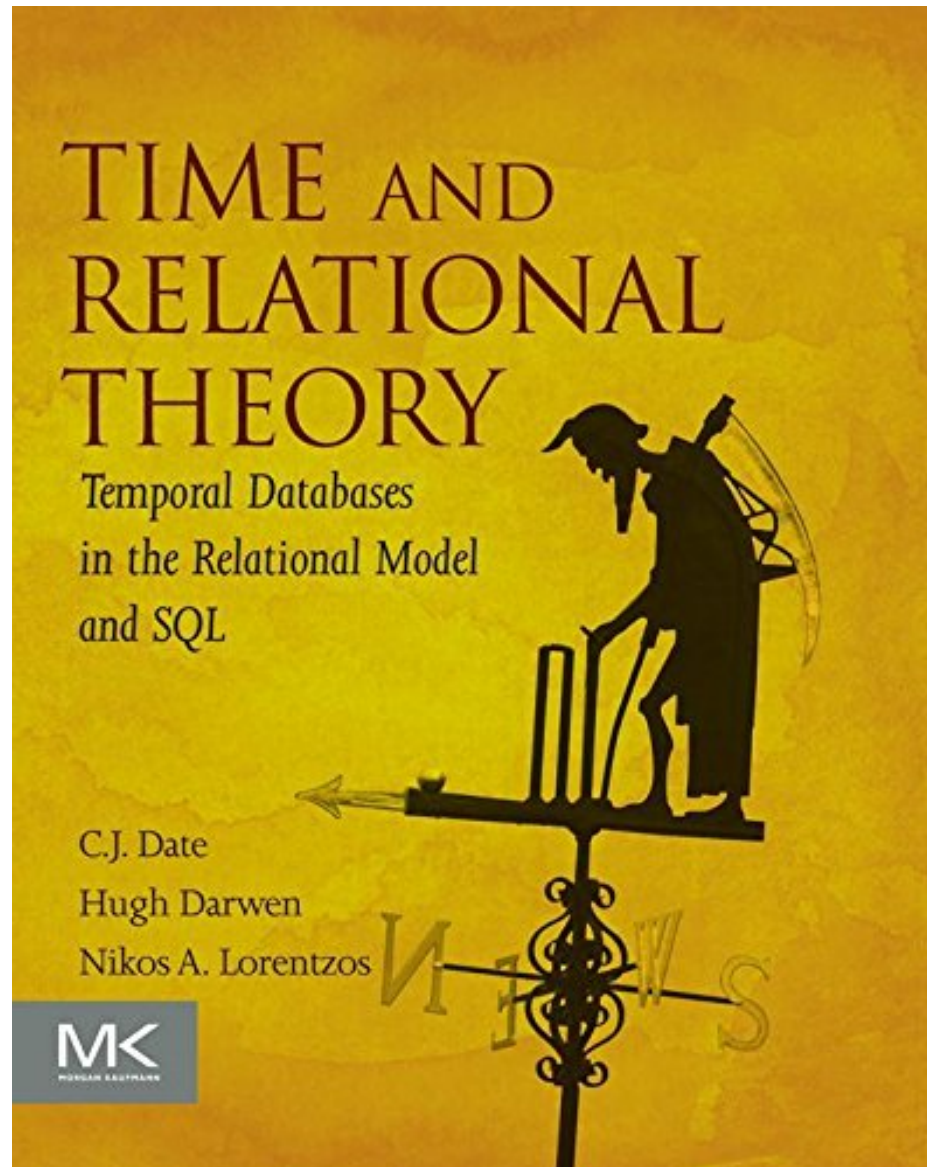
- **Semistrukturierte** Datenbanken (XML)
 - **Räumliche und multimodale** Datenbanken
 - lineare und mehrdimensionale Strukturen
 - **Temporale Datenbanken**
 - zeitlich beschränkte Gültigkeiten
 - Datenbanken für **Datenströme** (Fensterkonzept)
 - **Bewertung** von Antworten (Top-k-Anfragen)
 - **Probabilistische** Datenbanken zur Repräsentation **unsicherer Information**
 - Next Generation Databases, **Graphdatenbanken**
- Behandlung **unvollständiger Information** im Master

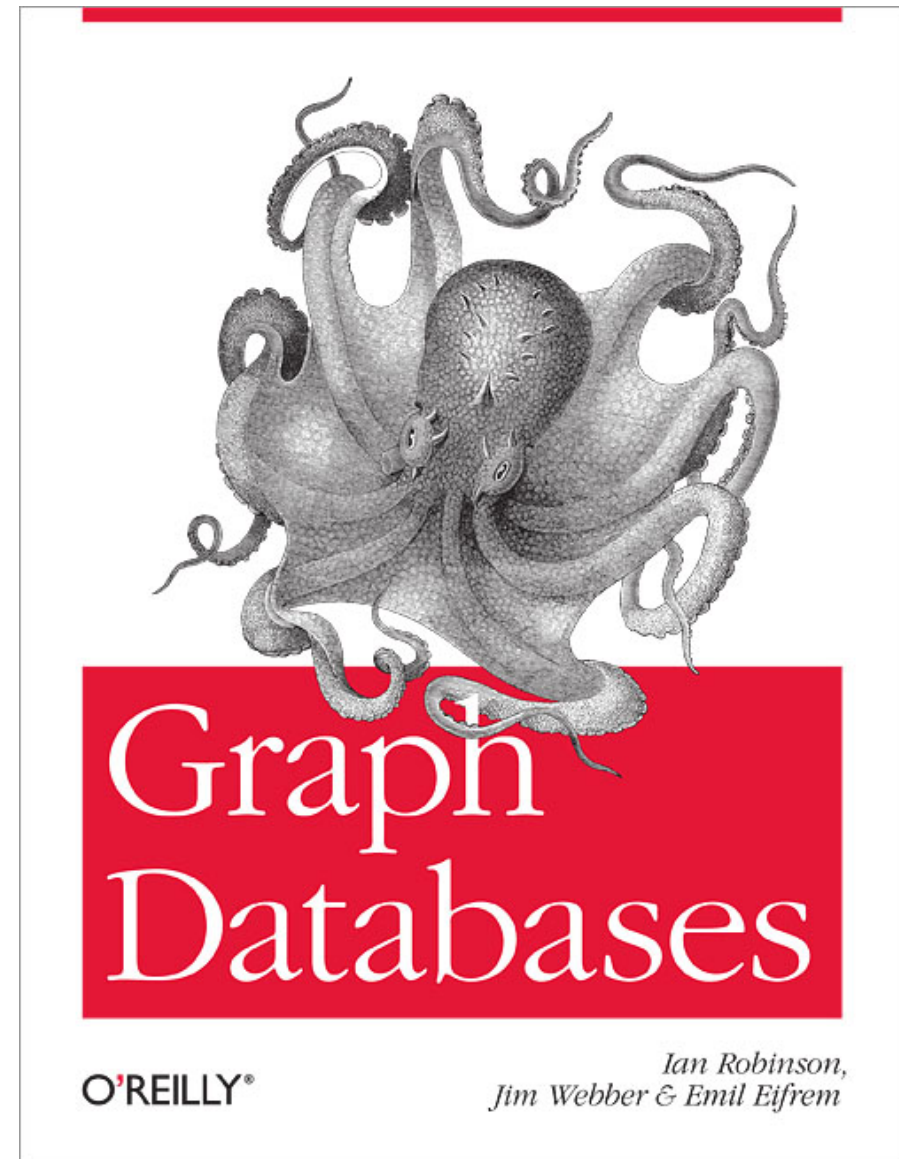
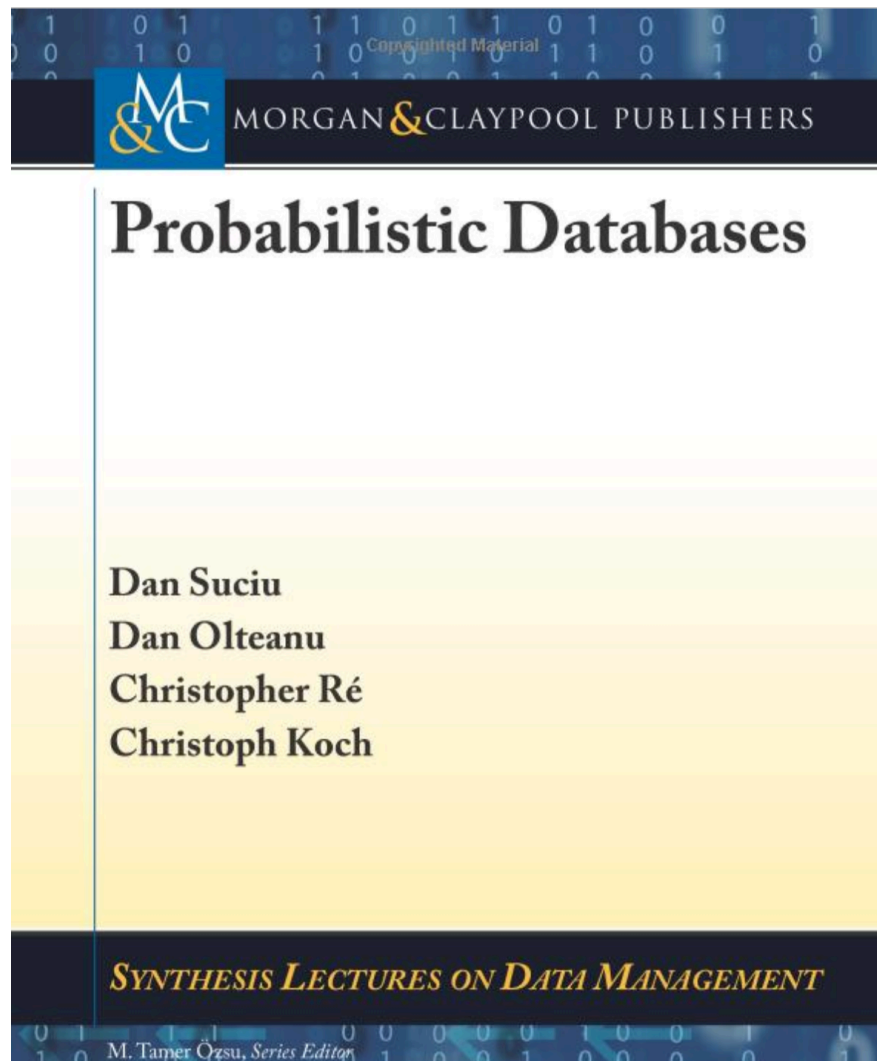
Literatur

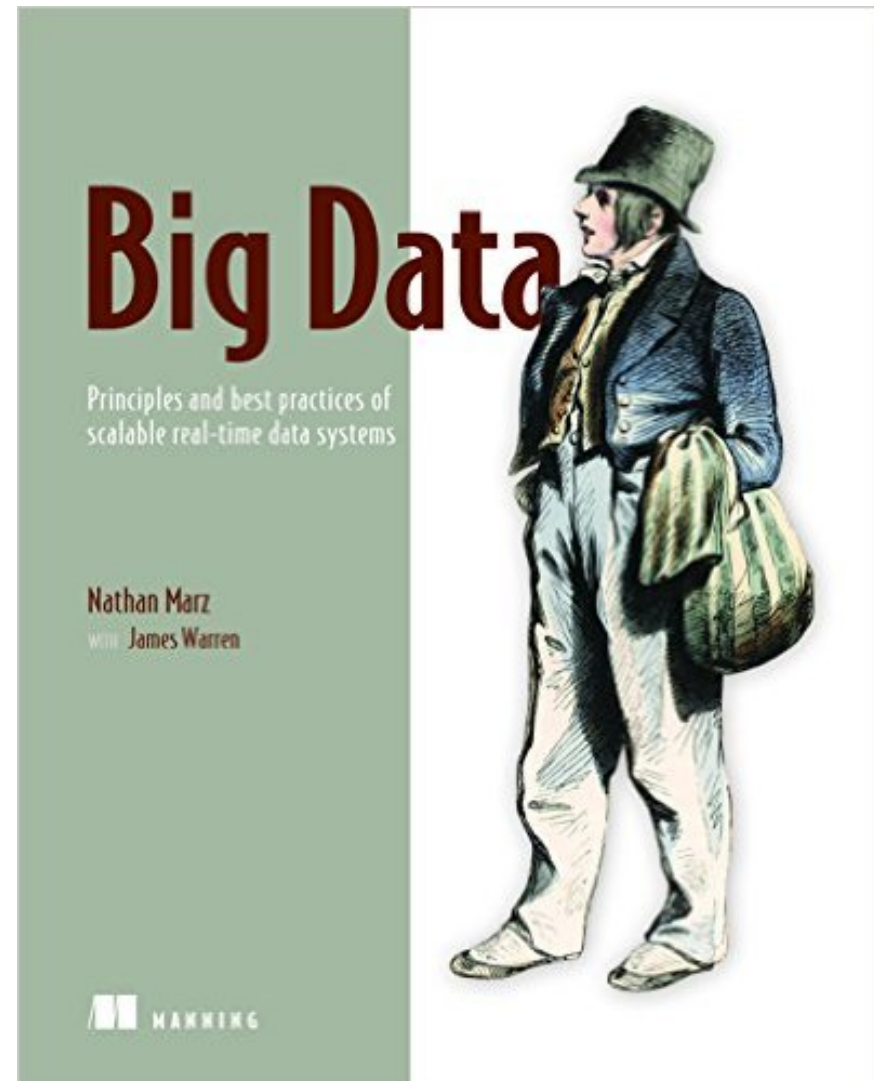
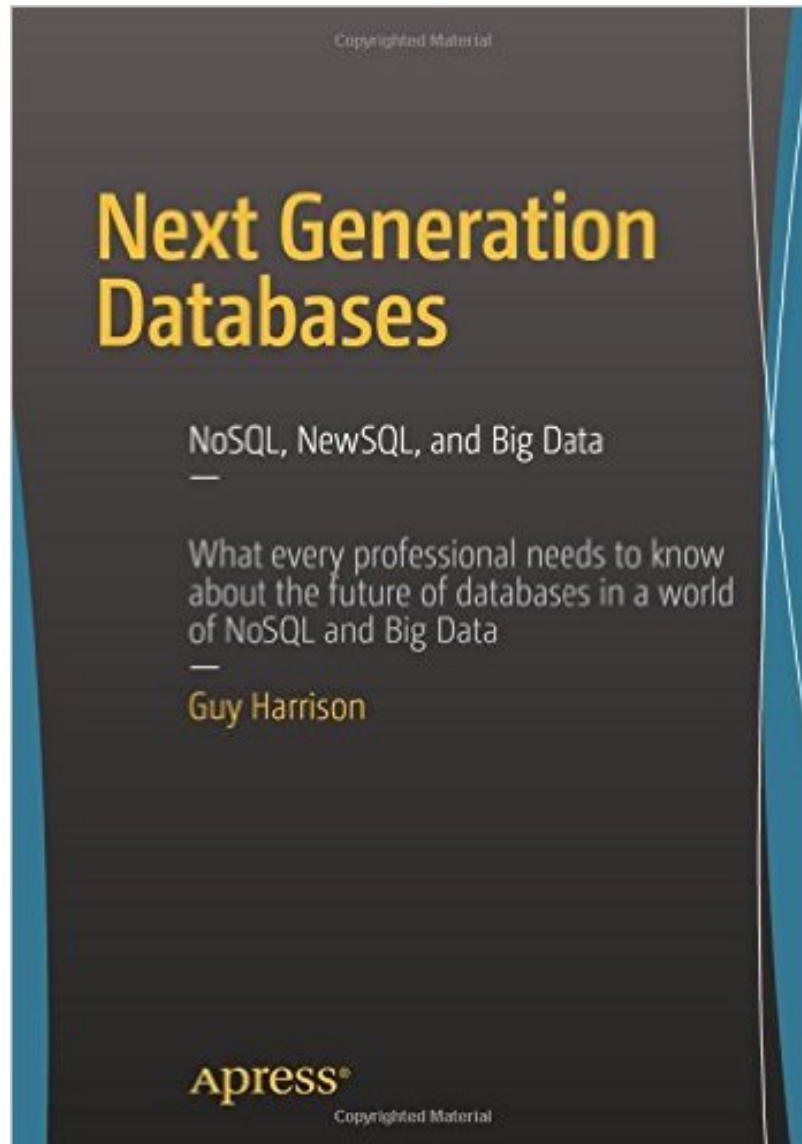
1999

2001









Literatur

S. Abiteboul, P. Buneman, D. Suciu, Data on the Web - From Relations to Semistructured Data and XML, Morgan Kaufmann, 1999

P. Rigaux, M. Scholl, A. Voisard, Spatial Databases With Applications to GIS, Morgan Kaufmann, 2001

C. J. Date, H. Darwen, N.A. Lorentzos, Time and Relational Theory: Temporal Databases in the Relational Model and SQL, Morgan Kaufmann, 2014

S. Chakravarthy, Q. Jiang, Stream Data Processing A Quality of Service Perspective, Springer, 2009

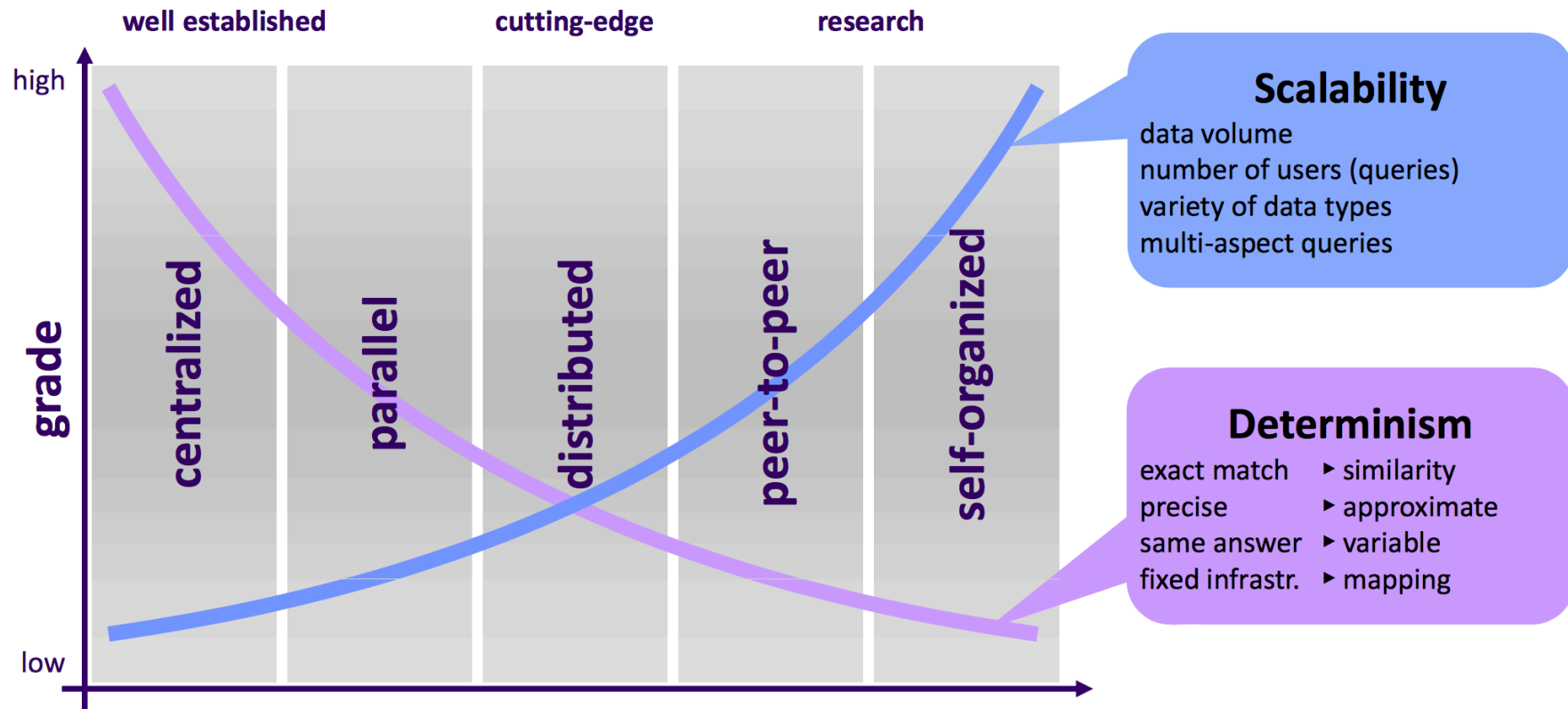
D. Suciu, D. Olteanu, Chr. Re, Chr. Koch, Probabilistic Databases, Morgan & Claypool, 2011

G. Harrison, Next Generation Databases: NoSQL, NewSQL, and Big Data, Apress, 2015

I. Robinson, J. Webber, E. Eifrem, Graph Databases, O'Reilly, 2016

N. Marz, J. Warren, Big Data: Principles and best practices of scalable realtime data systems; Manning, 2015

Datenhaltungstypen und Anfragebeantwortung



Datenbanken als Forschungsgebiet

- Semantik und Ausdrucksstärke
 - Was braucht man in der Anwendung?
- Ausdrucksstärke vs. Skalierbarkeit
 - Betrachtung der Komplexität des Anfragebeantwortungsproblems für eine gegebene Anfragesprache
 - Datenkomplexität
 - Wie wirkt sich eine Verdopplung des Datenbestandes bei fixer Anfrage auf die Worst-Case-Laufzeit der besten Anfragebeantwortungsalgorithmen aus?
 - Kombinierte Komplexität
 - Laufzeit bezogen auf Anfragelänge und Daten (relevant aber selten betrachtet, da Anfragelänge klein)