

XML, QPATH, XQUERY: NEUE KONZEPTE FÜR DATENBANKEN

Prüfungsvorbereitung

Hinweis

2

- Hier sind nur die Fragen zu den Originalfolien der Vorbereitungsveranstaltung aufgeführt. Diese stellen lediglich einen Ausschnitt möglicher Fragen dar. Die Prüfungsfragen werden sich über den gesamten Stoff erstrecken und sehr viel mehr ins Detail gehen.

Inhalt

3

- Datenbankkonzepte und XML
 - ▣ Theoretische Aspekte mit praktischer Sicht
- XML-basierte DBMS
 - ▣ Datenmodell und Anfragesprachen für XML
 - DTD, XSchema
 - XPath, XQuery
 - Realisierung in verschiedenen Möglichkeiten

Ziel der Vorlesung

4

- Grundlegendes Verständnis
 - ▣ Weiterführende (neue) (DB-) Konzepte
 - Im Bereich DB, IS, Web-Systeme
 - ▣ Basierend auf Datenbanktechnologie
 - Überblick und Details der Realisierung
 - ▣ Die neuen Anforderungen Rechnung tragen
 - Web, Verteilung
 - Unstrukturierte/semistrukturierte Daten
- Diskussion/Vorstellung
 - ▣ von Alternativen (pros und cons)
- Vergleich zu bisheriger DB-Technologie/Realisierung

Einführung in XML

5

Motivation für XML

- Trennung von Inhalt und Darstellung
- Daten und Informationen über die Daten (Struktur/Metadaten) im gleichen Dokument
- Durch das World Wide Web Consortium (W3C) entwickelt
 - ▣ Ziel des W3C: Schaffung von allgemeinen Protokollen für das WWW, um die Interoperabilität von Anwendungen zu sichern
 - ▣ Prozessphasen: Note, Working Draft, „Candidate Recommendation“, „Recommendation“
- Gut lesbar (human readable)
- Häufig eingesetztes Austauschformat
- Metasprache zur Definition von neuen Sprachen

Einführung in XML

6

- Beschreiben Sie Entwurfsziele für XML.
- Nennen Sie den Unterschied zwischen Datenzentriertem und Dokumentzentriertem XML.
- Was sind Merkmale semistrukturierter Daten?
- Was heißt die Aussage „XML ist selbstbeschreibend“?
- Vergleichen Sie Dokumente mit Datenbanken in Bezug auf Struktur, Beschreibung und Operationen.

XML vs. Datenbanken

7

□ XML & Datenbanken

- ▣ Welche Speicherungstechniken für XML-Dokumente?
- ▣ Welche Indizierung für gespeicherte Dokumente?
- ▣ Welche Anfragesprache und Updatesprache?
- ▣ Abbildung von XML zum relationalen Modell

□ Datenbanken & XML

- ▣ Neue Anforderungen an DBMS
- ▣ XML Erweiterungen in kommerziellen DBMS
- ▣ Abbildung von Relationen auf das XML Datenmodell

XML

Syntax

8

- Wie ist die Syntax eines XML-Elements?
- Mögliche Elementinhalte

XML

9

- Was ist XML?
- Nennen Sie Unterschiede zwischen XML und HTML.
- Was für eine Struktur liegt XML zugrunde?
- Wozu wird XML eingesetzt?
- Gibt es (verwandte) Standards zu XML?
- Werkzeuge
 - ▣ Welche Arten von XML-Werkzeugen gibt es?
- Formate/Zeichensätze
 - ▣ Warum wichtig?
 - ▣ Wie gelöst?

- Dokument
 - ▣ Wann ist ein Dokument wohlgeformt und wann gültig?
 - ▣ Nach welchem Prinzip sollte ein XML-Dokument strukturiert sein?
 - ▣ Struktur
 - ▣ Kommentare

XML-DTD

11

- Was ist eine XML-DTD?
- Was sind die Vorteile einer XML-DTD?
- Was sind die elementaren Konstrukte zur Strukturbeschreibung bei XML-DTD?

- Entitäten
 - allgemeine Entitäten
 - Zeichen-Entitäten (geparst)
 - Zahlendefinierte Entitäten
 - Namendefinierte Entitäten
 - XML-vordefinierte Entitäten
 - gemischter Inhalt (geparst)
 - interne Entitäten
 - externe Entitäten
 - unparsed Entitäten
 - Parameter-Entitäten (geparst)
 - interne Entitäten
 - externe Entitäten

Entitäten - allgemeine

13

- Wie wird eine Unicode-Zeichenentität referenziert?
- Was sind vordefinierte Zeichenentitäten?
- Wie wird eine Entität gemischten Inhalts deklariert?
- Wie werden Entitäten referenziert?
- Kann der Ersetzungstext Entitätenreferenzen enthalten?

Entitäten - allgemeine

14

- Wie werden externe geparste Entitäten deklariert?
- Wie werden externe ungeparste Entitäten deklariert?

Entitäten – Parameter-Entitäten

15

- Wie wird eine interne Parameterentität deklariert und wie referenziert?

Entitäten – Parameter-Entitäten

16

- Wie werden externe Parameterentitäten deklariert?

Entitäten - Weiteres

17

- Können Entitäten mehrfach deklariert werden?
- Spielt die Position eine Rolle an der eine externe Entität referenziert wird?
- Wie kann die Dokumentstruktur (anhand von Entwicklungsphasen „draft“) mithilfe von Entitäten gesteuert werden?
 - ▣ Welche zusätzlichen Konstrukte sind notwendig?

Entitäten – DTD intern & extern

18

- Was ist der Unterschied zwischen externen und internen Entitäten? Wo tauchen diese auf?

XML

19

- Welche Techniken gibt es, um in XML zu navigieren, Elemente zu identifizieren bzw. Elemente zu referenzieren?

XPath

20

- Was macht XPath?
- Was sind die Bestandteile von XPath zum Lokalisieren von Knoten(-mengen/-sequenzen)?
- Wie wird die Auswahl eingeschränkt?
- Kann man XPath-Ausdrücke abkürzen?

XPath

21

- Warum sollte man beim Erstellen eines XPath-Ausdrucks eher von der Struktur (-beschreibung /Schema) und nicht von der Instanz ausgehen?
- Wozu kann Xpath eingesetzt werden?
- Welche Techniken basieren auf XPath?

XPointer

22

- Was macht XPointer?
- Worauf basiert XPointer?
- Auf was kann man mit XPointer verweisen?
- Welche Konzepte für Stellen gibt es in XPointer?
- Wie werden Namensräume in XPointer verarbeitet?

XLink

23

- Wozu werden XLinks verwendet?
- Wo können XLinks verwendet werden?
- Welche Form haben XLinks?
- Worauf basiert XLink?
- Können XLinks zusätzliche Informationen tragen?
- Welche Arten von XLinks gibt es?

XML - Schema

24

- Wozu braucht man XML-Schema?
- Vergleichen Sie XML-DTDs mit XML-Schemata.
 - ▣ Nennen Sie Vor- und Nachteile.
 - ▣ Welche Technik ist performanter?
 - Warum?
- Was sind die elementaren Konstrukte von XML-Schema?
- Wie wird objektorientierte Datenmodellierung in XML-Schema unterstützt?
- Wie können Elemente identifiziert, wie referenziert und wie eindeutig gemacht werden?

XML - Schema

25

- Was sind anonyme Typen?
- Wie können Vokabulare unterschieden und referenziert werden?
- Wie können Schemata erweitert/kombiniert werden?

Verarbeitung von XML

26

- Darstellungsformen für XML-Infoset
- API's & Bibliotheken
- Parser
- Sprachen
 - ▣ Abfrage
 - ▣ Veränderung
 - ▣ Skripte
 - ▣ XPath, XSLT, XQuery, ...
- Speichern
 - ▣ Repository
 - ▣ DBMS

Verarbeitung von XML

Parser, APIs, ...

27

- Welche Bestandteile können sinnvoll zurück gegeben werden?
- Deserialisierung
- Serialisierung
- Validität
- Flexibilität
- Robustheit
- Modellüberführung
 - ▣ Abstraktion

Verarbeitung von XML

Darstellungsformen für XML-Infoset

28

- Ereignisbasiert
- Baumbasiert
- Daten-/Objektorientiert

- Vergleichen Sie die Ansätze bzw. die konkreten umsetzenden APIs.

XQuery

29

- Was ist XQuery?

XQuery

30

- Anforderungen an eine Anfragesprache

XQuery

31

□ Grundoperationen

XQuery

32

- Was sind die Bestandteile von XQuery?
- Semantik
 - ▣ Normalisierung

XML zum Abbilden von DB-Inhalten

Motivation

33

- RDBMS
- XML

XML zum Abbilden von DB-Inhalten

Methoden

34

- verschiedene (standardisierte) Methoden siehe auch SQL/XML
- Varianten:
 - ▣ Vollständige Abbildung von Datenbankinhalten
 - ▣ Abbildung von Anfrageergebnissen oder Views
 - ▣ Einsatz individueller Transformationsregeln

XML zum Abbilden von DB-Inhalten

Bestandteile

35

- Welche Bestandteile des Relationenmodells können/müssen abgebildet werden?

Export mit SQL/XML

Transformation der Inhalte nach Abbildungsvorschriften

36

□ Bestandteile der Transformation

- Quelle
- Ziel

□ Hilfsmittel aus SQL/XML

- SQL-Anfragen
- ...

SQL/XML

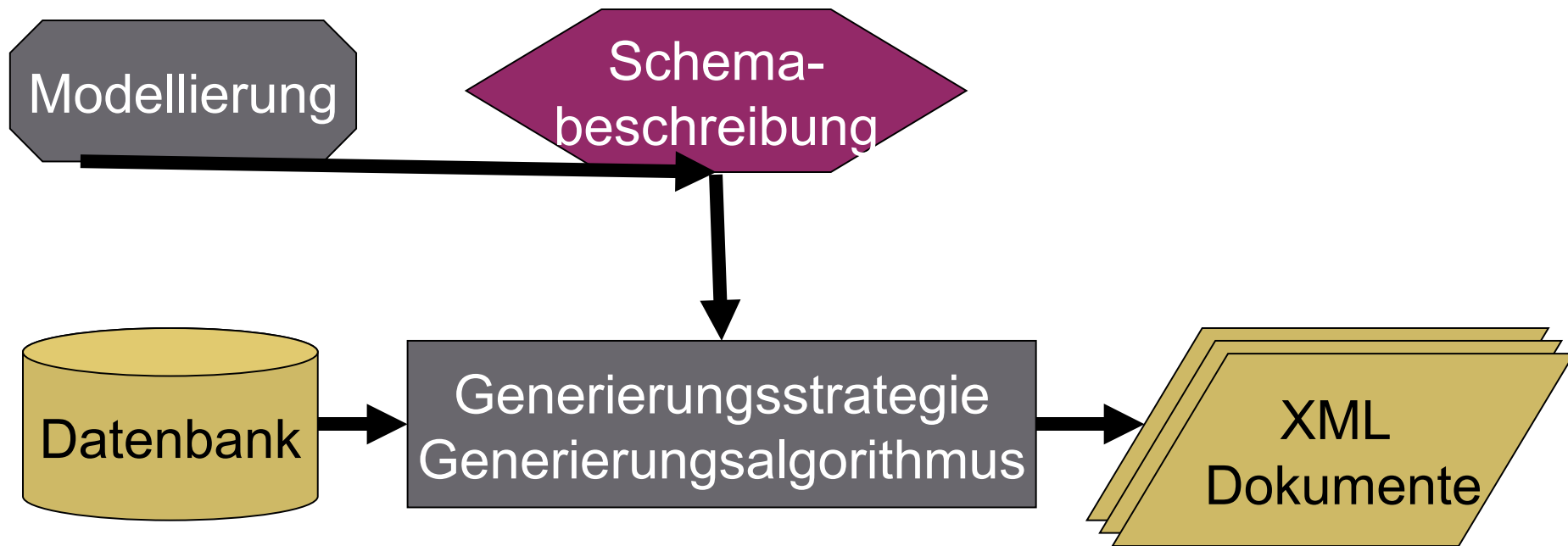
Grundidee

37

- Neuer Basisdatentyp „XML“
- Abbildung zwischen SQL und XML
 - ▣ Besonderheiten
 - ▣ Was wird wie abgebildet?
 - ▣ Auf was wird abgebildet?
- SQL/XML Funktionen
- SQL/XML 2005

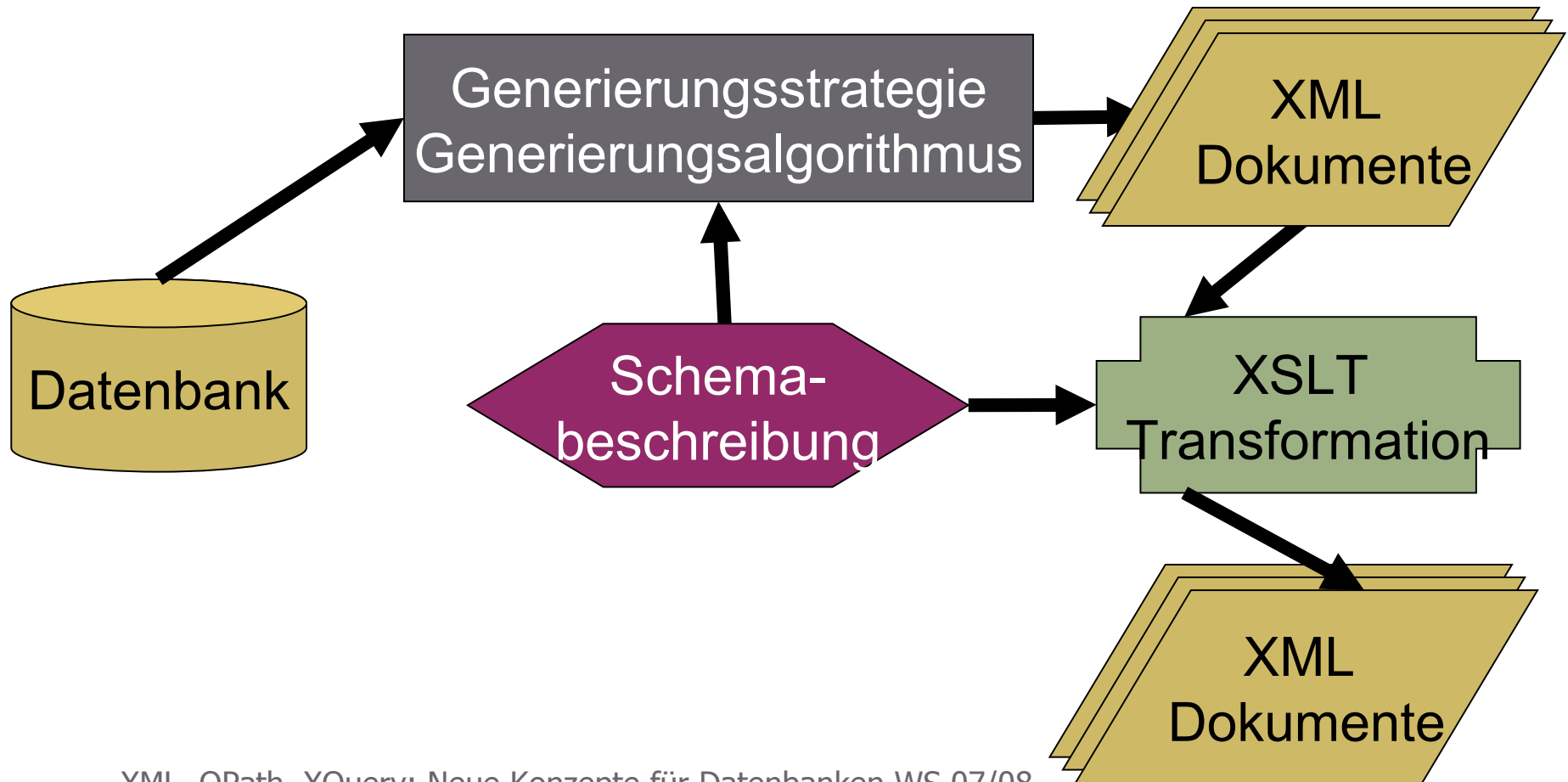
1. Generierung von XML-Dokumenten aus Datenbanken

38



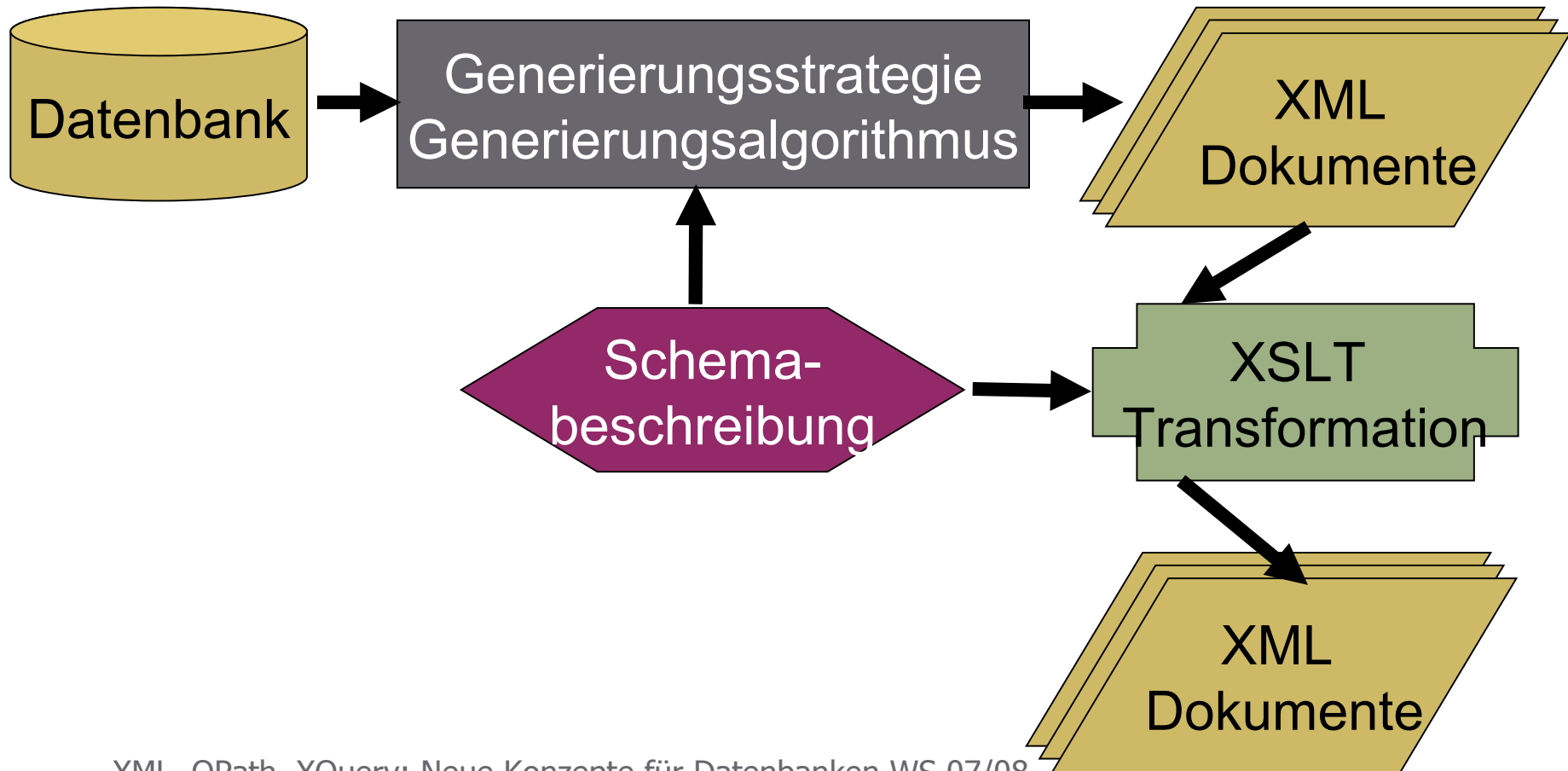
2. Generierung von XML-Dokumenten für festes Schema

39



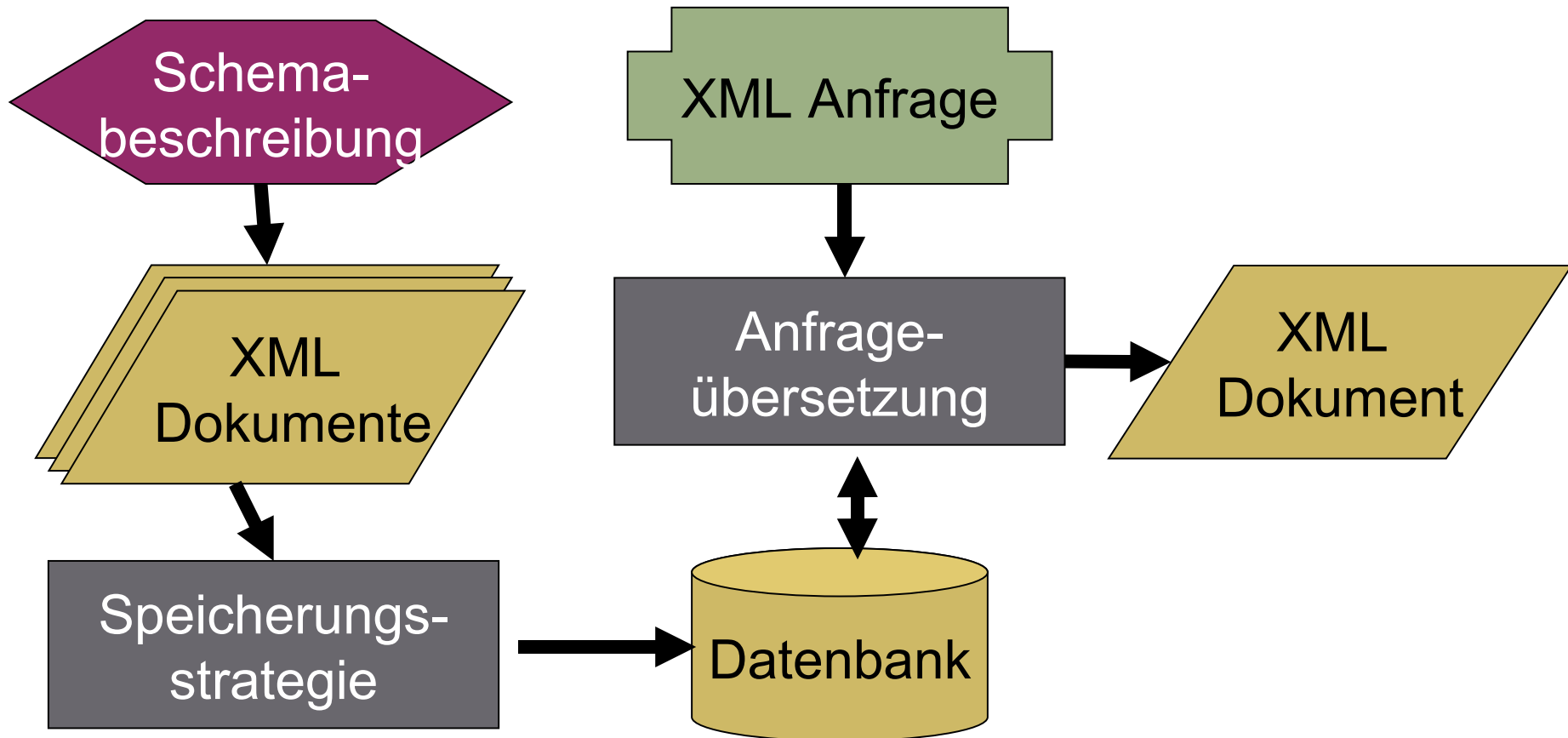
3. Visualisierung von Datenbankinhalten

40



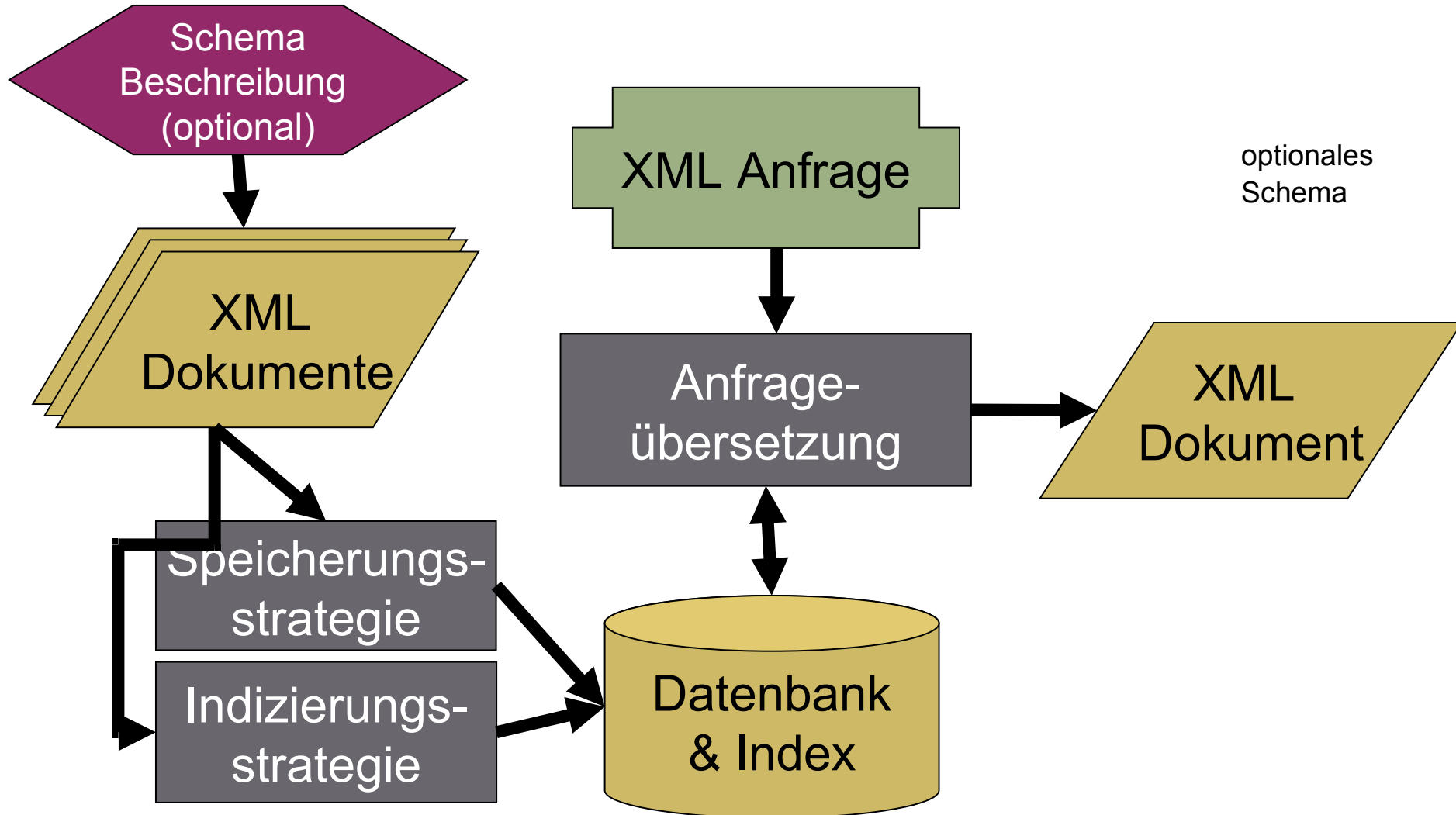
4. Speicherung datenzentrierter XML-Dokumente

41



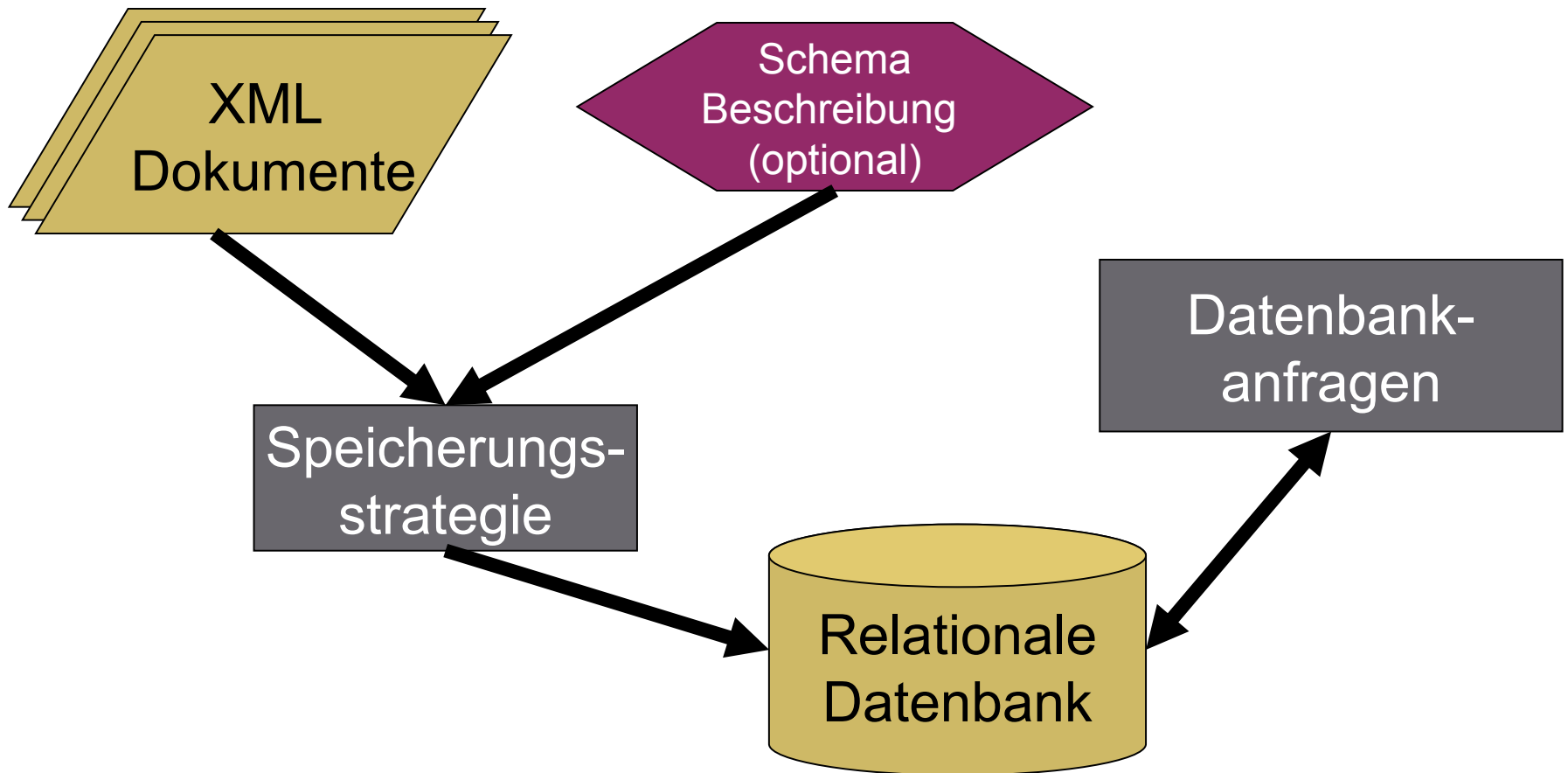
5. Speicherung dokument-zentrierter XML-Dokumente

42



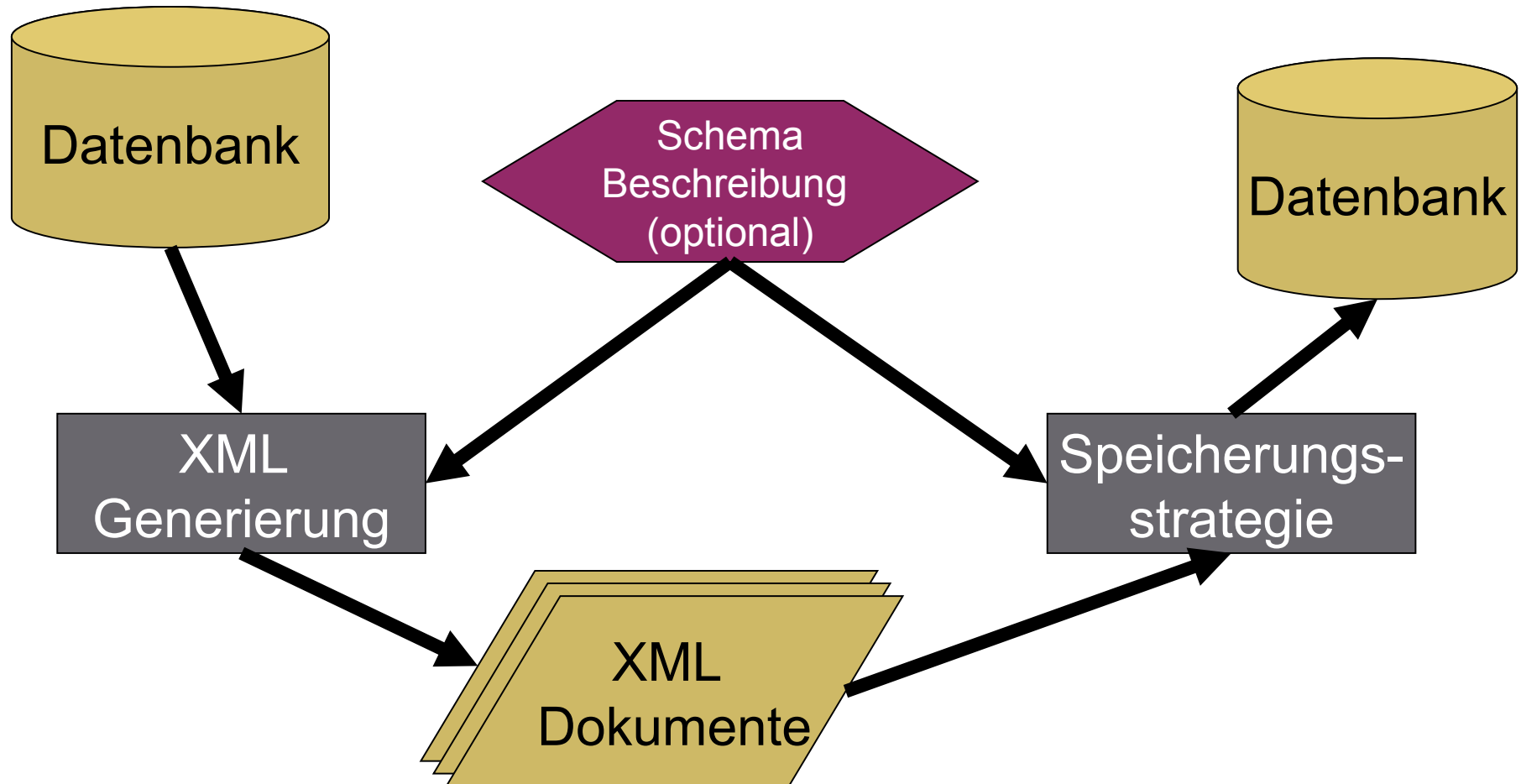
6. Speicherung von XML-Dokumenten in bestehende DB

43



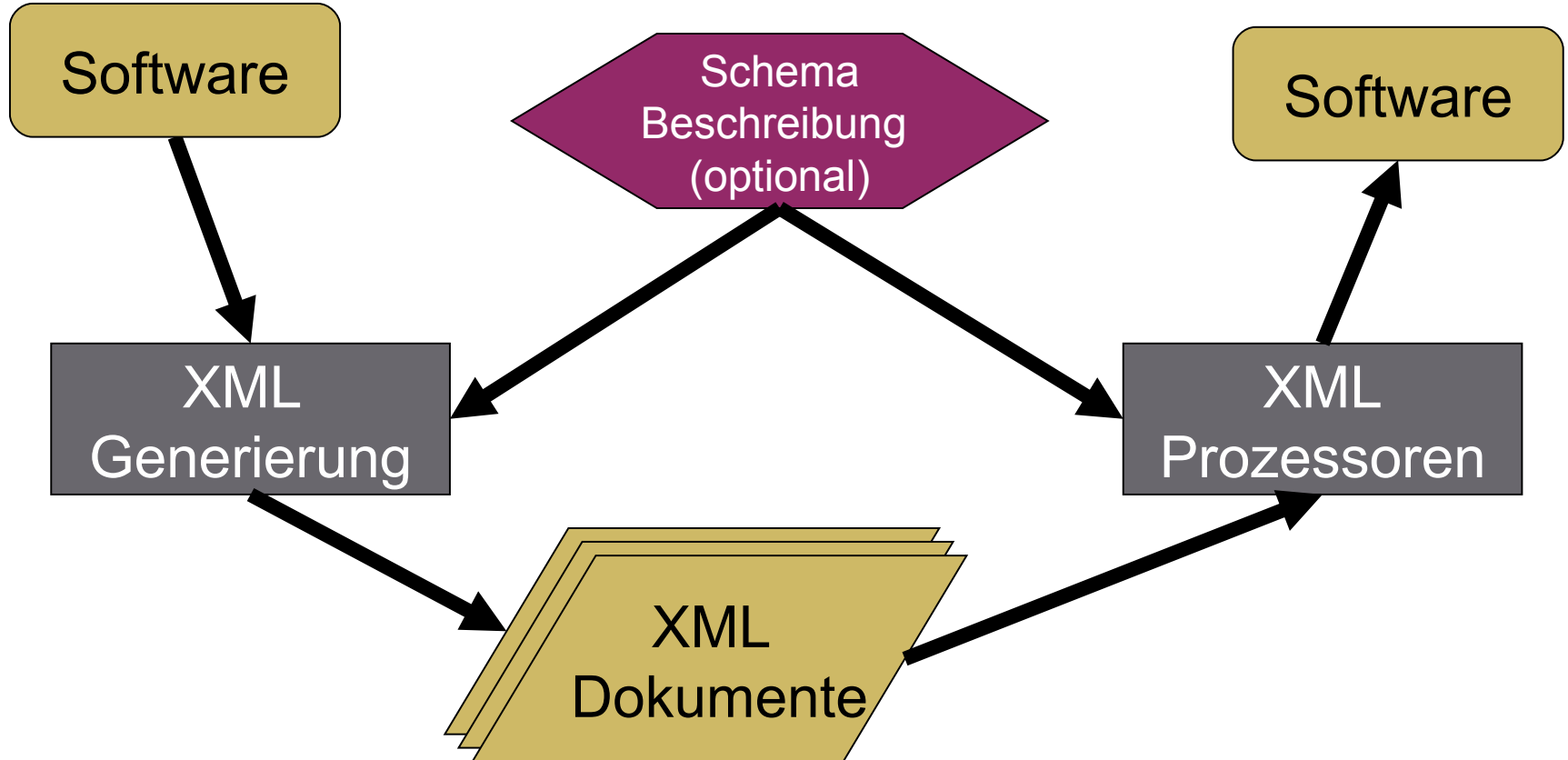
7. Austausch von Datenbankinhalten

44



9. Kommunikation von Softwarekomponenten

45



Themen

46

- Einführung
 - ▣ XML Datenmodell & Syntax
- XML Prozessoren
 - ▣ SAX / DOM
- XML Schemata
 - ▣ DTD
 - ▣ XML Schema
- Anfragesprachen
 - ▣ XPath, XQuery, SQL/XML, XSLT, Xupdate
- XML Generierung aus DBMS

Themen

Forts.

47

- Architekturen zur Speicherung von XML
- Indizierung
 - ▣ PrePost Order Index, Volltext Indices
- XML Query Execution Engine (XEE)
 - ▣ Seitenorientierte Speicherung von XML-Dokumenten
- Relationale DBMS für XML
 - ▣ DB2, Oracle, SQLServer
- XML Benchmarks