Weitere Aspekte

- Informationsintegration
- Metdadaten
- Sicherheit
- Zahlungssysteme

Informationsintegration

- Kombination von Daten aus verschiedenen Quellen in einer globalen (virtuellen) Datenbank und Ableitung von Informationen
- Ansätze
  - Föderierte Datenbanken
  - Data Warehouses
  - Mediatoren
- Anwendungen
  - Informationsagenten
  - OLAP, Data Mining, ...

Informationsintegration: Probleme

- Probleme durch Heterogenität der Quellen
- Szenario:
  - Quelle #1
    Cars (serialNo, model, color, autoTrans, cdPlayer, ...)
  - Quelle #2
    Autos (serial, model, color)
    Options (serial, option)

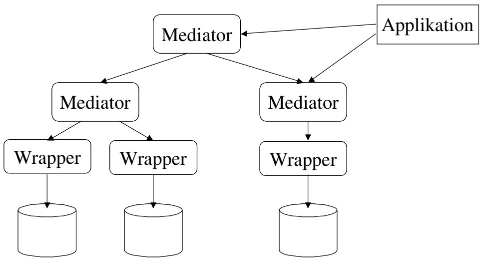
Informationsintegration: Probleme

- unterschiedliche Datentypen
  - Seriennummer als Zeichenkette oder Integer-Wert
- unterschiedliche Werte
  - Farbe als „BLACK“ und „BL“
- semantische Unterschiede
  - LKWs in Cars, nur PKWs in Autos
- fehlende Werte
  - Automatik/Handschaltung nicht angegeben in Quelle #1

Mediatoren: Prinzip

- Bereitstellung einer integrierten Sicht über verschiedene heterogene Informationsquellen (Datenbanken, semistrukturierte Quellen)
- Quellen mit unterschiedlichen und eingeschränkten Anfrageschnittstellen
- Nutzung von gebietsspezifisches Wissen zur Auswahl der Quellen und der Beantwortung der Anfragen
  - Aggregierungen, Abstraktionen, Datenbereinigung
  - Einbeziehung von Semantik

Mediatoren: Prinzip



Mediatoren: Aufgaben

- Auswahl der Quellen
  - Auswahl der Datenbanken mit relevanten Daten
  - erfordert Quellenbeschreibung (Inhalt, Fähigkeiten)
- Anfrageübersetzung
  - Zerlegung in Teilanfragen für die einzelnen Quellen
  - Transformation in lokale Anfragesprache
- Komposition der Ergebnisse
  - Ergebnisse der Teilanfragen vereinigen (Fusion)
  - Duplikate entfernen, Ranking normalisieren

Mediatoren: Ansätze

- Aufgabe:
  - Abbildung zwischen Schemata der lokalen Quellen und der globalen Sicht
- „Global As View"
  - Globales Schema als Sicht über Quellenschemata
- „Local As View"
  - Lokale Schemata als Sichten über globales Schema

Wrapper

- Softwarekomponente zum Zugriff auf Quelle
  - Realisierung einer Anfrageschnittstelle
  - Verbergen der Heterogenität
- Beispiel:
  - Mediator:
    - AutosMed (serialNo, model, color, autoTrans, dealer)
    - SELECT serialNo, model FROM AutosMed WHERE color = "red"
  - Wrapper für Quelle1:
    - SELECT serialNo, model FROM Cars WHERE color="red"
  - Wrapper für Quelle 2:
    - SELECT serial, model FROM Autos WHERE color="red"

Metadaten

- Überblick
- Resource Description Framework
  - Modell und Syntax
  - Schema
  - Anwendungen
- Vokabulare
  - PICS
  - Dublin Core

Metadaten - Überblick

- Metadaten
  - Informationen über andere Daten (Dokumente, Datensammlungen, Server)
  - in einer Form, die Recherche, Retrieval und Nutzung der Primärdaten ermöglichen und erleichtern
  - Kennzeichnung, Katalogisierung und Beschreibung zur automatisierten Verarbeitung
  - erfordern Modellierung und Kodierung

Metadaten im Internet

- Motivation: Probleme bei der gezielten (Volltext-)Suche nach Dokumenten im Internet
- Qualität abhängig
  - vom inhaltlichen und formalen Erschließungsgrad der Quellen,
  - deren wissenschaftlichen Niveau,
  - den Selektionskriterien sowie
  - den Retrievalmodalitäten
- Metadaten können
  - Erschließungsgrad steigern
  - Retrievalmodalitäten verbessern und Selektionskriterien präzisieren

Metadaten im Internet

- Herausforderungen/Probleme
  - Traditionelle Indexierungsmethoden nicht anwendbar
  - Metadaten müssen von Autoren erstellt und maschinell verarbeitbar sein
  - Standards für formale und inhaltliche Erschließung (Modelle, Syntax, Vokabulare)

Resource Description Framework (RDF)

- Technologie des W3C
- Graphenbasiertes **Modell**, das Internet-Ressourcen und deren Beziehungen untereinander beschreibt
- **Syntax** für Serialisierung von graphenbasierten Modellen
- Grundlage für Erstellung von Vokabularen (**Schemaspezifikationssprache**)

RDF: Modell

- Repräsentation von bezeichneten Eigenschaften und Werten (Attribute von Ressourcen) sowie von Beziehungen
- syntaxneutraler Mechanismus zur Repräsentation von RDF-Ausdrücken
- Nutzung zur Bestimmung von (semantischer) Äquivalzenz zwischen Ausdrücken
- Objekttypen:
  - Ressourcen, Properties, Statements

RDF: Objekttypen

- Ressource
  - „Dinge“, die durch RDF-Ausdrücke beschrieben werden
  - bezeichnet durch URI (URL)
  - Bsp.: Web-Seite, Elemente einer Seite, Sammlung von Web-Seiten
- Property
  - Aspekt, Attribut oder Relation zur Beschreibung einer Ressource
  - spezifiziert Bedeutung, erlaubte Werte, den Typ der beschriebenen Ressource sowie Beziehung zu anderen Properties

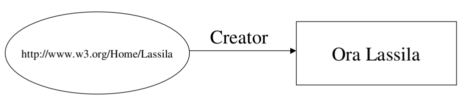
RDF: Objekttypen

- Statement
  - Ressource + Property + Wert dieser Eigenschaft
  - Subjekt, Prädikat, Objekt
  - Objekt (Eigenschaftswert): Ressource, Literal
- Beispiel:

  Ora Lassila is the creator of the resource http://www.w3.org/Home/Lassila.

RDF: Beispiel

Subjekt (Ressource)	http://www.w3.org/Home/Lassila
Prädikat (Property)	Creator
Objekt (Literal)	„Ora Lassila“



RDF: Syntax

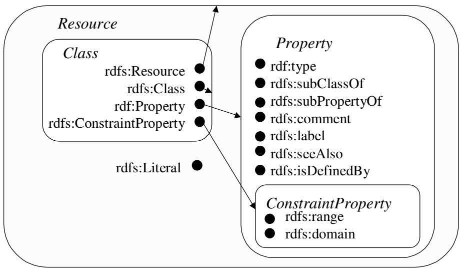
- XML-Syntax
- Gruppierung in **Description**-Element
- Nutzung von XML-Namensräumen
- Beispiel:

```
<rdf:RDF>
<rdf:Description
     about="http://www.w3.org/Home/Lassila">
     <s:Creator>Ora Lassila</s:Creator>
</rdf:Description>
</rdf:RDF>
```

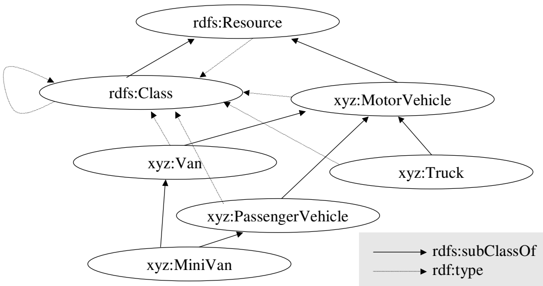

RDF: Schema

- Schema
  - Deklaration der Eigenschaften (Attribute) und ihrer Semantik
  - sowie der Ressourcen, die durch diese beschrieben werden
- RDF-Schema
  - Schemaspezifikatíonssprache
  - Typsystem (Klasse, Subklassen-Beziehung, ...)
  - Property-zentriert
    - Klassen kennen nicht ihre Attribute, aber Attribute kennen beschriebenen Klassen

RDF-Schema: Klassen und Ressourcen



RDF-Schema: Beispiel



RDF-Schema: Beispiel

```
<rdf:Description ID="MotorVehicle">
  <rdf:type resource="#Class" />
  <rdfs:subClassOf rdf:resource="#Resource" />
</rdf:Description>
<rdf:Description ID="PassengerVehicle">
  <rdf:type resource="#Class" />
  <rdfs:subClassOf rdf:resource="#MotorVehicle" />
</rdf:Description>
...
<rdf:Description ID="MiniVan">
  <rdf:type resource="#Class" />
  <rdfs:subClassOf rdf:resource="#Van" />
  <rdfs:subClassOf rdf:resource="#PassengerVehicle" />
</rdf:Description>
```


RDF: Anwendungen

- Thesauri und bibliographische Klassifikationssysteme
- Web-Sitemaps
- Beschreibungen des Inhaltes von Web-Seiten
- Beschreibung von Gerätefähigkeiten
- Rating-Systeme
- Digitale Signaturen

PICS

- Platform for Internet Content Selection
- Ziele
  - Finden von geeigneten Inhalten
  - Unterdrücken von ungeeigneten oder ungewollten Inhalten
- Rating-Dienst: Kennzeichnung von Inhalten
  - Self-rating
  - Third-party rating

PICS

- Rating-Dienst
  - definiert Dimensionen für Kennzeichnung, Skala für mögliche Werte und Beschreibung der Kriterien
  - Bsp.: MPAA (US-Kino): G, PG, PG-13, R, NC-17
- PICS-Rating-Dienst
  - Spezieller Dokumententyp: application/pics-service
  - definiert Kategorien Dimensionen)

PICS - Beispiel

```
((PICS-version 1.1)
    (rating-system "http://www.gcf.org/ratings")
    (rating-service "http://www.gcf.org/v1.0/")
    (icon "icons/gcf.gif")
    (name "The Good Clean Fun Rating System")
    (description "Everything you ever wanted to know about soap,
     cleaners, and related products.  For demonstration purposes only.")
(category
    (transmit-as "suds")
    (name "Soapsuds Index")
    (min   0.0) (max   1.0))
(category
    (transmit-as "density")
    (name "suds density")
    (label (name "none") (value 0) (icon "icons/none.gif"))
    (label (name "lots") (value 1) (icon "icons/lots.gif")))
```

Dublin Core

- Satz von Erschließungsmerkmalen (Vokabular) zur Verbesserung des Retrievals digitaler Dokumente bei Internet-Recherchen
- Ziele:
  - interdisziplinäre, eindeutige und intuitive Anwendung
  - angesiedelt zwischen automatisch erstellten Suchindexen und von professionellen Anbietern erzeugten Katalogen

Dublin Core - Elemente

- Elemente als Metatags in Dokumenten
- Elementgruppen
  - **Content:** Title, Subject, Description, Type, Source, Relation, Coverage
  - **Intellectual Property:** Creator, Publisher, Contributor, Rights
  - **Instantiation:** Date, Format, Identifier, Language

Dublin Core - Elemente

- **Title**: Name der Resource
- **Creator**: Verfasser, Urheber
- **Subject**: Thema, Urheber
- **Description**: Beschreibung, Abstrakt
- **Publisher**: Herausgeber, Verleger
- **Identifier**: URL, URN, ISBN
- **Format**: Datenformat, z.B. MIME
- **Type**: Kategorie (text, image, sound, data, software, interactive, physical object)

Dublin Core - Relationen

- Relation: Identifikation einer zweiten Ressource und die Art der Beziehung dazu
  - Beziehungstypen: IsPartOf, HasPart, References, IsReferencedBy, Requires, IsRequiredBy
  - Notation:
    Title="the present resource"
    Relation="relationship-type [space] unique identifier"
  - Beispiel:
    Title="The Memory Remains"
    Creator="Metallica"
    Date="1997"
    Type="sound"
    Relation="IsPartOf ReLoad"

Dublin Core - Syntax

- Elemente sind optional und können mehrfach vorkommen
- Elemente können durch *Qualifier* präzisiert werden
  - SCHEME: ISO-Norm, bibliographischen Normsatz, Datumsformat, ...
  - LANG: Sprache des Metatags
  - SUB-ELEMENT: präzisiert Element
- Für jedes Element nur festgelegte Qualifier

Dublin Core in HTML

- DC „integriert“ in HTML
  - Elemente in Meta-Tag <META>
    - Syntax: DC.*Element[.Subelement]*
    - Beispiel: DC.Creator.PersonalName
- Mit <LINK> Verweise zu Informationen über die Metadatenelemente

Dublin Core in HTML - Beispiel

```
<HTML>
<HEAD><TITLE>Song of the Open Road</TITLE>
<META NAME="DC.Title" CONTENT="Song of the Open Road">
<META NAME="DC.Creator" CONTENT="Nash, Ogden">
<META NAME="DC.Type" CONTENT="text">
<META NAME="DC.Date" CONTENT="1939">
<META NAME="DC.Format" CONTENT="text/html">
<META NAME="DC.Identifier" CONTENT="http://www.poetry.com/nash/open.html">
</HEAD>
<BODY><PRE>
I think that I shall never see. A billboard lovely as a tree.
Indeed, unless the billboards fall. I'll never see a tree at all.
</PRE></BODY>
</HTML>
```


Sicherheit

- **Voraussetzung für Einsatz von Web-Anwendungen**
  - Vertraulichkeit der benutzerspezifischen Daten (Profil, Passwörter, Kreditkartendaten)
  - keine Möglichkeit zur Manipulation bzw. zum Mißbrauch durch Dritte

Sicherheit

- **Risiken**
  - Identifizierung von Teilnehmern, Maskerade
  - unbefugtes Mitlesen oder Modifizieren von Informationen
  - Bestreiten des Sendens bzw. Empfangens von Informationen
  - Bekanntwerden vertraulicher Informationen
  - wiederholtes Einspielen von Informationen (z.B. finanzieller Transaktionen)

Aspekte der Sicherheit

- Schutz von Hosts und Datenbankobjekten
- Verschlüsselung von Informationen
- Authentifizierung
- Digitale Signaturen
- Abrechnung, Durchführung finanzieller Transaktionen

Datenbanksicherheit

- Rechtevergabe
  - Zugriffsrecht: *(AutorisierungsID, DB-Ausschnitt, Operation)*
  - AutorisierungsID: interne Kennung eines Datenbanknutzers
  - Datenbank-Ausschnitt: Relationen und Sichten
  - DB-Operation: Lesen, Einfügen, Ändern, Löschen

Datenbanksicherheit

- Rechtevergabe in SQL
  `grant` *Rechte* **`on`** *Tabelle* **`to`** *Benutzerliste*
  [**`with grant option`**]
  - *Rechte-Liste*: **`all`** bzw. Langform **`all privileges`** oder Liste aus **`select`**, **`insert`**, **`update`**, **`delete`**
  - *Tabelle*: Relationen- oder Schemaname
  - *Benutzerliste*: Autorisierungsidentifikatoren (auch **`public`**, **`group`**)
  - spezielles Recht: Recht auf Weitergabe von Rechten (**`with grant option`**)

Datenbanksicherheit

- Beispiel: Autorisierung

```
create view MeineAuftraege as
select * from Auftrag
where Kname = user;

grant select, insert on
MeineAuftraege to public;
```

- Jeder Benutzer kann seine Aufträge sehen und neue Aufträge einfügen (aber nicht löschen) !

Datenbanksicherheit

- Zurücknahme von Rechten
  `revoke` *Rechte* **`on`** *Tabelle* **`from`** *Benutzerliste* [**`restrict`** | **`cascade`**]
- **`restrict`**: falls Recht bereits an Dritte weitergegeben: Abbruch von **`revoke`**
- **`cascade`**: Rücknahme des Rechts mittels **`revoke`** an alle Benutzer propagieren, die es von diesem Benutzer mit **`grant`** erhalten haben

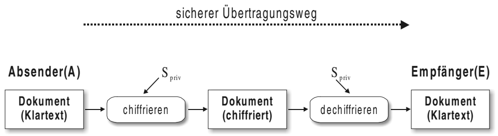
Verschlüsselung: Grundlagen

- Übertragungssicherheit (Abhörsicherheit) von Informationen durch Verschlüsselung
  - Kodierung mittels Geheiminformation (Schlüssel) die späteres Entschlüsseln ermöglicht
- Weitere Anwendungen
  - Authentifizierung:
    - Kontrolle, ob jemand derjenige ist, der er zu sein vorgibt
    - Verwendung von Geheiminformationen (Passwort)
  - Digitale Signaturen:
    - Unterschreiben von Dokumenten

Symmetrische Verfahren

- Private Key-Verfahren
- Kodierung der Nachrichten mit privatem Schlüssel
- Verfahren:
  - DES (Data Encryption Standard)
  - IDEA, RC4, RC5
- sichere Kommunikation erst nach Austausch des Schlüssels außerhalb des Mediums möglich

Symmetrische Verfahren



Asymmetrische Verfahren

- Public Key-Verfahren
- jeder Teilnehmer besitzt *Schlüsselpaar*
  - öffentlicher Schlüssel
  - privater Schlüssel
- Nachricht, die mit einem Schlüssel chiffriert wurde, kann nur mit anderem Schlüssel dechiffriert werden
- Verschlüsselung: Chiffrieren mit öffentlichem Schlüssel des Empfängers
- Verfahren:
  - RSA (Rivest, Shamir, Adleman): langsamer als DES (Faktor 10 ... 10000)

Assymetrische Verfahren

Digitale Unterschrift

- Unterschreiben einer Nachricht
- Prinzip:
  - Absender chiffriert Nachricht mit privatem Schlüssel
  - Empfänger dechiffriert Nachricht mit öffentlichem Schlüssel des Senders
  - Reduzierung des Aufwandes:
    - Ableitung einer Signatur (Digest) aus Nachricht und anschließende Verschlüsselung
- Verfahren: MD4, MD5

Zertifizierung

- Beglaubigung von öffentlichen Schlüsseln durch Zertifizierungsstelle
- Zertifikat:
  - Name
  - Schlüssel
  - von Zertifizierungsstelle digital unterzeichnet
- Prüfung mittels öffentlichem Schlüssel der Zertifizierungsstelle

Zahlungssysteme: Überblick

- E-Commerce-Anwendungen
- Realisierung sicherer finanzieller Transaktionen (z.B. mit Kreditkarten)
- Bezahlen mit Kleinstbeträgen im Internet (Micro Payment)
- Aspekte
  - Vertraulichkeit von Informationen
  - Integrität von Zahlungen
  - Identität der beteiligten Partner

SET: Secure Electronic Transaction

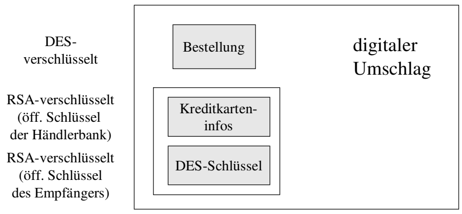
- Industriestandard zur Kreditkartenzahlung im Internet
- im Auftrag und unter Beteiligung von Visa und MasterCard entwickelt
- Abbildung der Transaktionen einer Kreditkartenzahlung
- Nutzung von
  - Zertifikaten
  - digitalen Unterschriften
  - Verschlüsselungstechniken

SET: Prinzip

- Anforderung: *Gewährleistung der Zahlungssicherheit*
  - Geheimhaltung der Karteninfos durch Händler
  - Authentifizierung des Händlers
  - Gültigkeit der Karte
  - Authentifizierung des Karteninhabers
- Partner:
  - kartenausgebende Bank (Issuer)
  - Händlerbank (acquirer)
- Transaktionen:
  - Bestellung, Abrechnung, Autorisierung

SET: Prinzip

- Kreditkarteninformationen in digitalem Umschlag
  → Händler kann Informationen nicht lesen



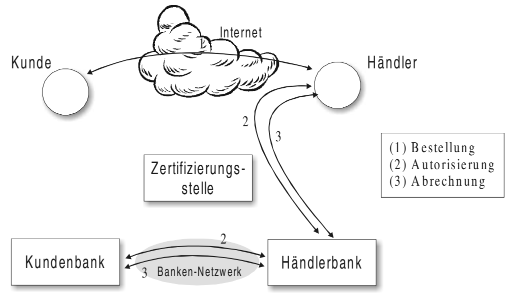
SET: Ablauf einer Bestellung

1. Initialisierungsnachricht von Kunden an Händler:
   Kaufwunsch
2. Antwort mit Zertifikaten:
   eigener RSA-Schlüssel + öffentlicher Schlüssel der Händlerbank
3. Kunde:
   - Prüfung der Zertifikate
   - Verschlüsselung der Bestellung mit DES-Schlüssels
   - Verpacken der Kreditkarteninfos in digitalem Umschlag mit öffentlichem Schlüssel der Händlerbank

SET: Ablauf einer Bestellung

4. Händler
   - Auspacken des DES-Schlüssels mit privatem RSA-Schlüssel
   - Dekodierung der Bestellung
   - Weiterleitung der verschlüsseltem Kreditkarteninfos an Bank
   - Empfang der Autorisierung

SET: Ablauf



Micro Payment

- Online-Zahlung von Kleinstbeträgen im Internet (Pfennigbruchteile bis wenige DM)
- Anwendung:
  - Zahlung mittels Karte oder Rechnung zu aufwendig
  - Systeme mit hohem Durchsatz (10 Mill. Transakt./Tag)
- Elektronisches Geld zur Bezahlung von
  - Online-Artikeln
  - Anfragen an Datenbanken oder Suchmaschinen
  - Software-Downloads

Micro Payment

- Prinzipien
  - direkter Münztransfer
    - virtuelle Münzen
    - händler- oder institutsbezogen
  - Schattenkonto bei Institut
- Projekte
  - Millicent (Digital)
  - ecash (Deutsche Bank)
  - Cybercash (Dresdner Bank, Cybercash)

Millicent

- Rollen:
  - Kunde mit Millicent-Wallet (elektronische Geldbörse)
  - Händler:
    - unterhält für definierten Zeitraum Scrip-Konto für jeden Kunden
  - Broker:
    - Instanz zum Kauf/Verkauf von Scrips (Provider, Bank)
    - übernimmt Tausch von Geld für verschiedene Händler → kein Sammeln von Händler-Scrips notwendig
- keine zentrale Kontoführungseinrichtung

Millicent

- Scrips: elektronisches, virtuelles Geld
  - verschlüsselte Nachricht mit Account, Kontostand, Verfallsdatum
  - definieren vorausbezahlten Wert
  - Gültigkeit für Zahlung mit einem bestimmten Händler → Händler kann Gültigkeit selbst prüfen

Millicent: Ablauf

1. Kunde kauft beim Broker Scrips für bestimmten Händler
   - Broker kann bei Händler Scrips kaufen
   - Bezahlung: Ecash, Kreditkarte, Telefonrechnung
2. Broker sendet Händler-Scrip und Wechsel-Scrip (Broker-Scrip) an Kunden
3. Kunde bezahlt Produkt mit Händler-Scrips
   - Händler sendet Produkt und Wechsel-Scrip (Händler-Scrip)
4. Weitere Käufe bei diesem Händler möglich

CyberCash

- Internet-Zahlungssystem (Dresdner Bank, Sachsen LB)
- Unterstützte Zahlungsverfahren
  - edd: electronic direct debit (Lastschriftverfahren)
  - Kreditkarte
  - CyberCoins (Schattenkonto)
- Komponenten
  - Kunden-Wallet
  - CashRegister (Händler)
  - CyberCash Gateway-Server

CyberCash

- CyberCoins :=
  keine Münzen, sondern „Kontostandsanzeige“ für Schattenkonto bei einer Bank
- Prinzip
  - Kunde erhält digitale, verschlüsselte Ware
  - Empfang des Schlüssels zum „Auspacken“ erst nach vollständiger Bezahlung
- Anwendungsbereiche: digitale Waren (Software, Bilder, ...)