

Semantic Web und Web 2.0

Unterschiede, Gemeinsamkeiten, Synergien

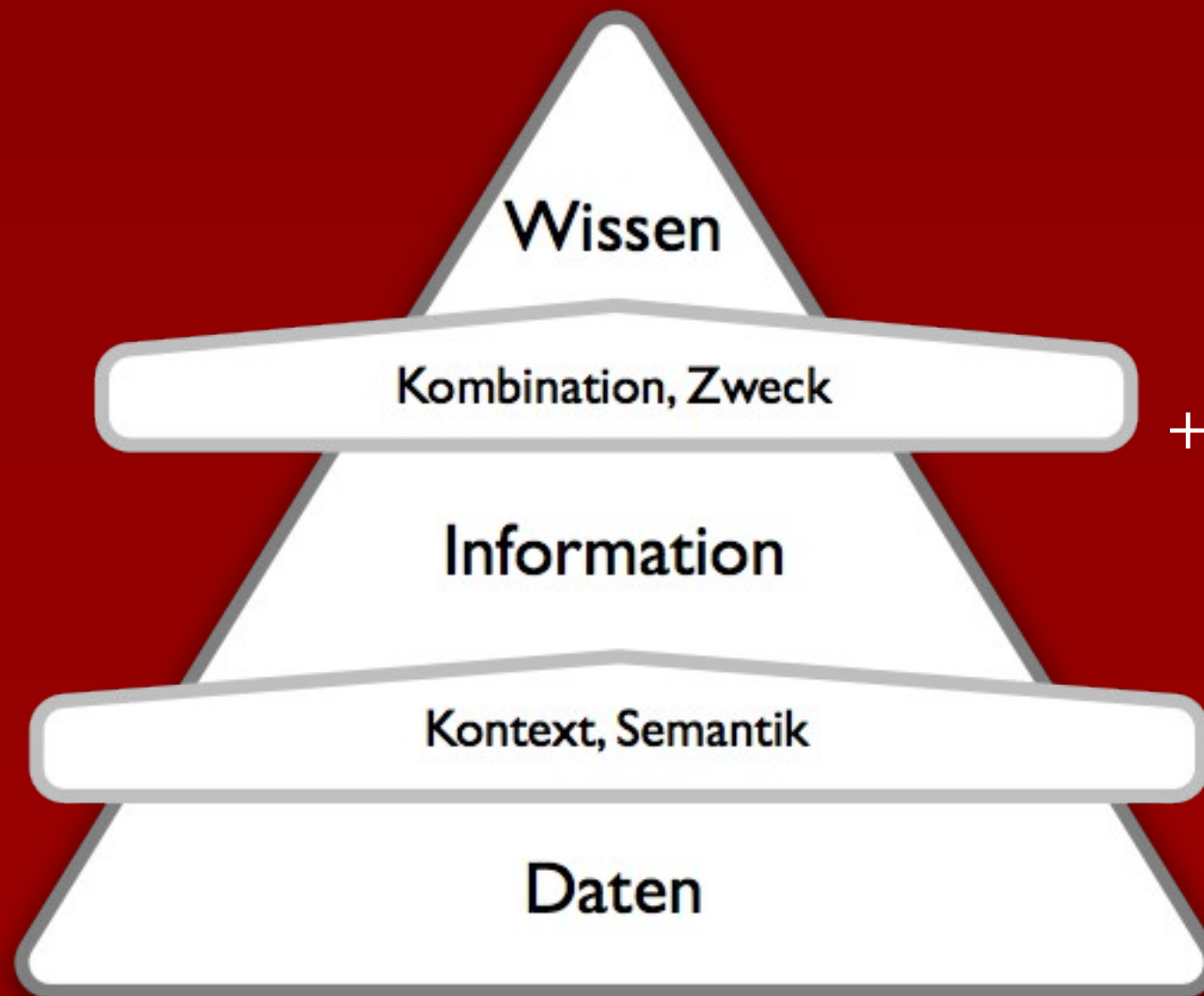
Benjamin Nowack

appmosphere web applications

http://www.appmosphere.com/2006/10/zgdv_2006_10_19.pdf

Kontext Wissenstechnologien

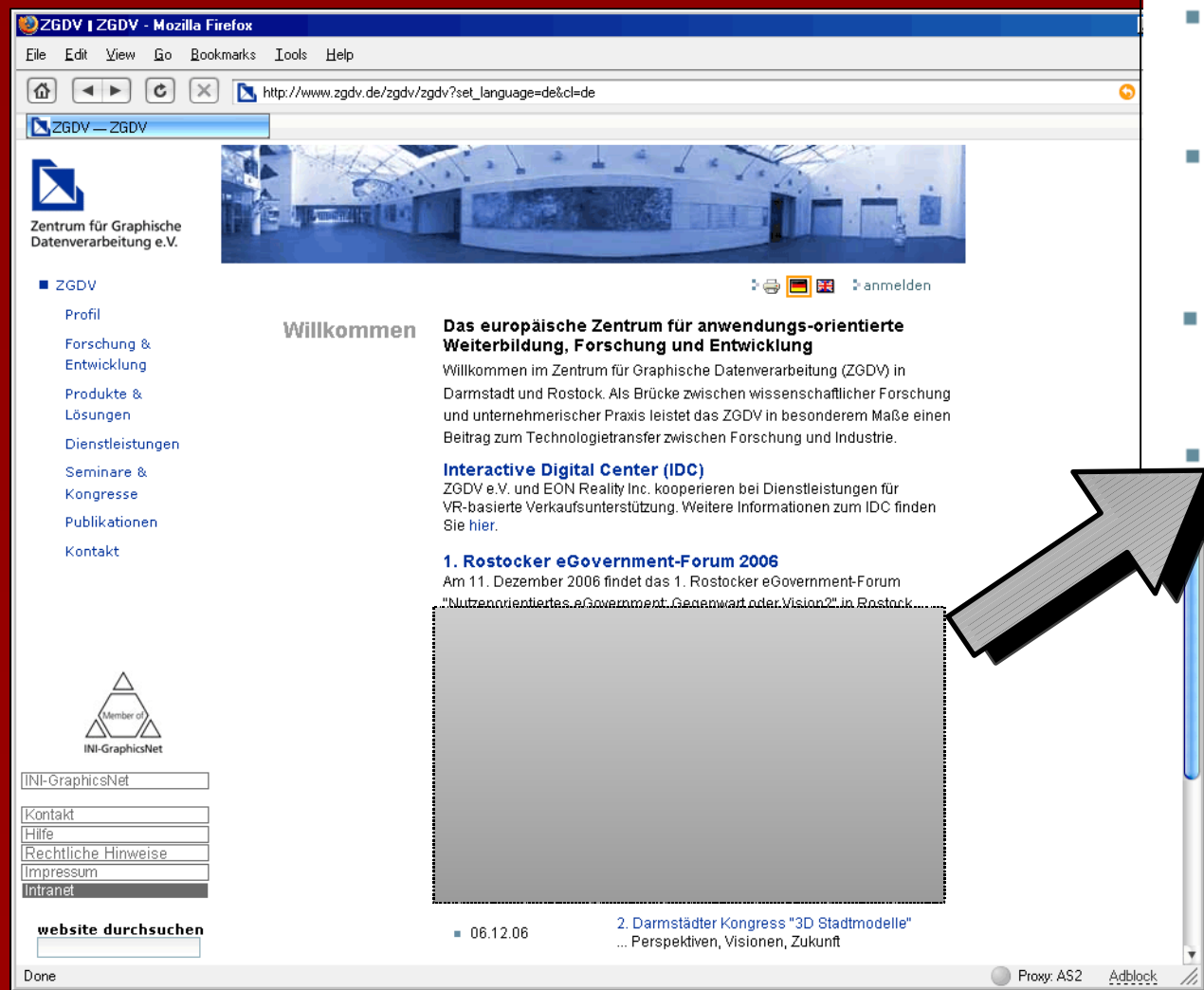
(vereinfachte Sichtweise, don't shoot me, please ...)



Alec Tronnik betreut Kunde X.
Kunde X hat PLZ "45145".
+ "45145" ist eine PLZ in NRW.
+ Den NRW-Vertrieb leitet Alec Tronnik.
"45145" ist eine deutsche Postleitzahl.
XML-*plz*-Tags kennzeichnen
deutsche Postleitzahlen.
... <plz>45145</plz> ...

Das Web als Wissensbasis

Das Web ist (und war schon immer) eine Wissensbasis, allerdings meist nur für Menschen zugänglich.



Kongresse und Veranstaltungen im ZGDV

- 19.10.06  3. Darmstädter Kongress "Semantic Web und Wissenstechnologien"
- 26.10.06  5. Darmstädter Kongress "Barrierefreies E-Government"
Aktuelle Trends in der Webentwicklung
- 02.11.06  5. Darmstädter Kongress "XML und ALKIS"
... nach der Pilotphase
- 06.12.06  2. Darmstädter Kongress "3D Stadtmodelle"
... Perspektiven, Visionen, Zukunft

Nur bedingte Unterstützung durch Maschinen (z.B. Suche auf der Datenebene.)

Informationsgewinn im Web (bisher)

Daten- Quellen	HTML, APIs (z.B. WSSs)
Informations- Gewinn	Programm-Code, Menschen
Informations- Verarbeitung	Programm-Code, Menschen

- HTTP: einheitlicher Datenzugriff
- HTML: vereinheitlichte Darstellung
- Informationsgewinn: nicht einheitlich automatisierbar (Ansätze: Web Services)
- Verarbeitung und Integration: meist nur innerhalb einer (vorgegebenen) Anwendung

Bessere Informationsverarbeitung im Web

Ziel:

- Vereinfachte Informationsverknüpfung/-integration

Prinzipiell drei Anknüpfungspunkte:

- Daten (standardisierte Formate, explizite Semantik)
- Protokolle (standardisierte Schnittstellen)
- Tools (z.B. standardisierte Algorithmen)

Semantic Web “vs.” Web 2.0

Entstehung und Hintergrund

Semantic Web

- Initiative des W3C
- 1999, Tim Berners-Lee, “Weaving the Web”

Web 2.0

- geschützt als Service Mark für eine Konferenz
- 2004, O’Reilly & MediaLive, “Web 2.0 Conference”
- 2005, Tim O’Reilly, “What is Web 2.0”

Semantic Web “vs.” Web 2.0

Kurz-Charakterisierung

Semantic Web

Eine (zukunftsorientierte) Erweiterung der bestehenden Web-Technologien.

Web 2.0

Ein beschreibender und vermarktbarer Oberbegriff für (bereits sichtbare) Veränderungen im Web.

Semantic Web “vs.” Web 2.0

Kurz-Charakterisierung (ganz kurz ;)

Semantic Web

“A pipe dream.”

Web 2.0

“A piece of jargon.”

Nichtsdestotrotz:

- Es sind Communities entstanden.
- Es lassen sich durchaus Technologien identifizieren.
- Bessere Informationsverarbeitung ist jeweils “key”.

Versuch einer Kurz-Beschreibung

Semantic Web

“Das Web als Datenbank”: Web-Informationen werden in einer Form zur Verfügung gestellt, die eine Interpretation durch Computer ermöglicht.

Web 2.0

“Das Web als Plattform”: Die Anwender stehen im Mittelpunkt benutzerfreundlicher Systeme, sie sind Gestalter, Wertschöpfer und Nutznießer zugleich.

Bekannte Beispiel-Anwendungen

Semantic Web

Tools: ja (Parser, Frameworks etc.),

Anwendungen im WWW: (noch) eher wenig:

Facetted Browser, Search (z. B. CC), Geo, BBC
(aber Tool- und Entwickler-Nachfrage steigt)

Web 2.0

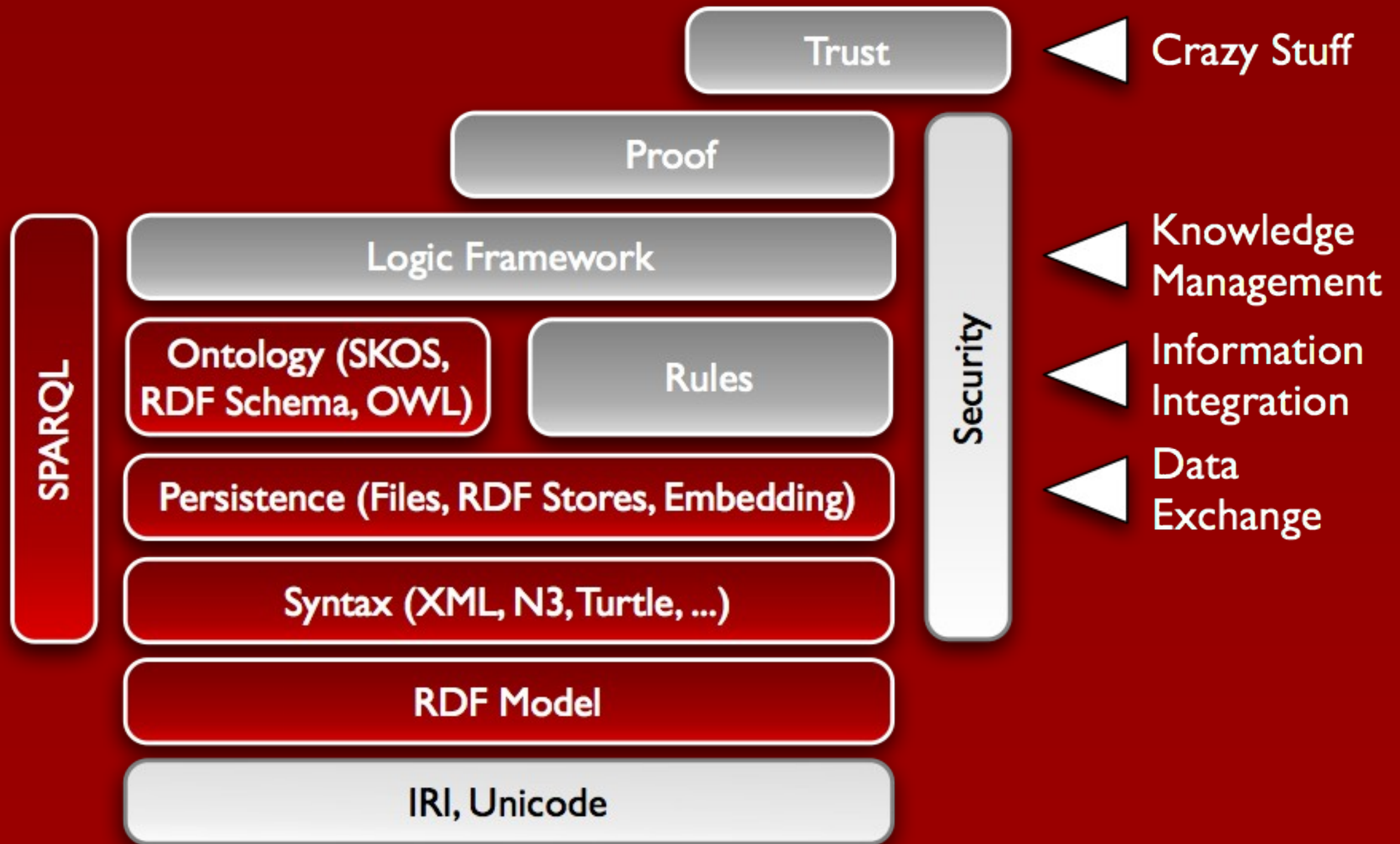
Amazon (Bewertungen, Kaufempfehlungen),

del.icio.us (Bookmark Sharing), flickr (Photo Sharing),

Google (PageRank), eBay (Bewertungen, Long Tail),

Digg, Netvibes, Wikipedia, Blogs, ...

Der Semantic Web "Layer Cake"



Semantic Web Technik: Unterbau

IRIs und Unicode

- Internationale Darstellung von URIs (Uniform Resource Identifier) und Text
- Identifikation, Benennung, Addressierung

Namespaces (Namensräume)

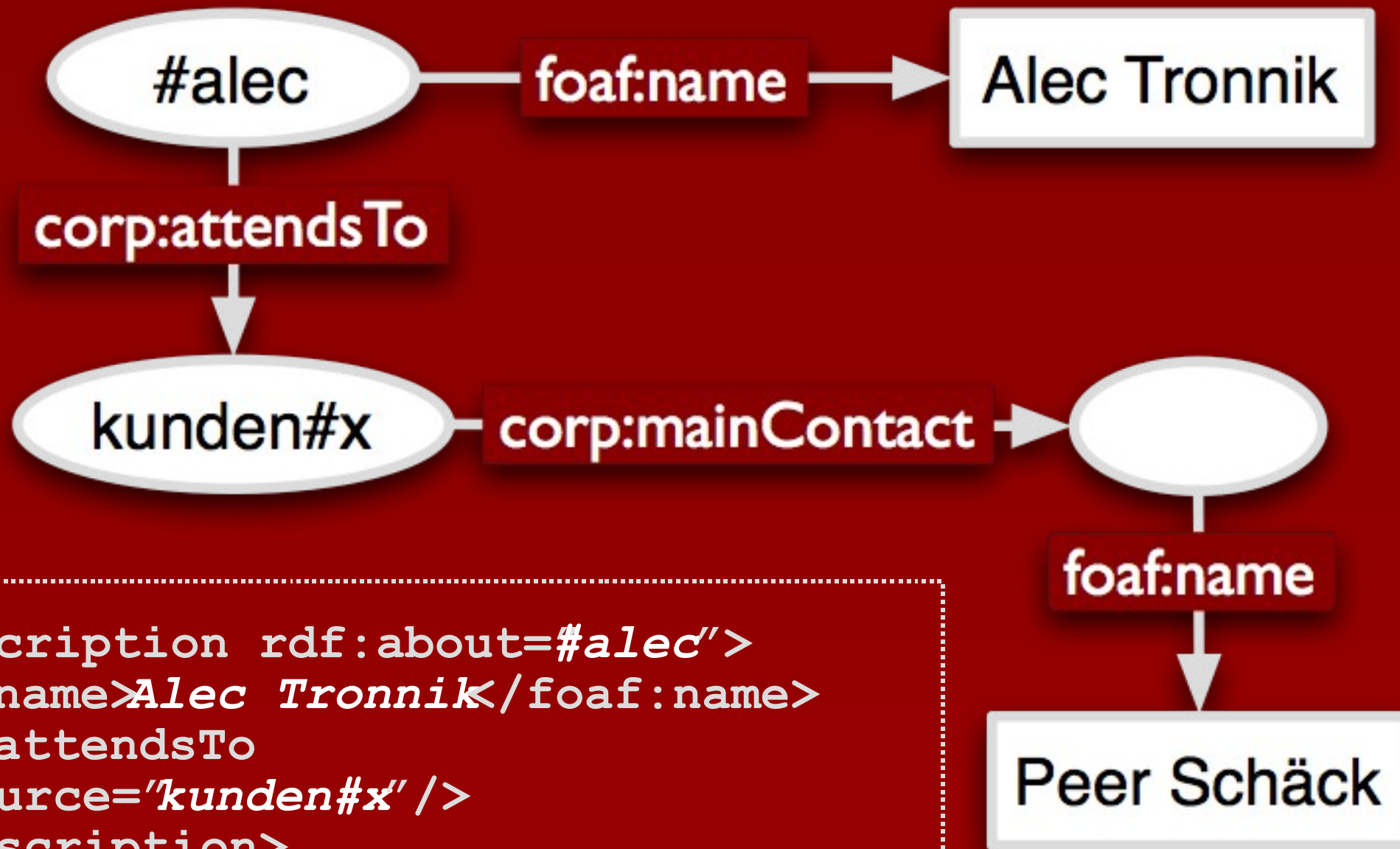
- Gruppierung und Abkürzung von IRIs (z.B. dc:creator für <http://purl.org/dc/elements/1.1/creator>)

Semantic Web Technik: RDF

Das Resource Description Framework (RDF)

- Einfaches Prinzip zur Beschreibung von Ressourcen
- Tripel aus Subjekt - Prädikat - Objekt:
- `</home.html> dc:creator "John Doe" .`
- Alle Terme können IRIs sein.
- Objekte können das Subjekt anderer Tripel sein.
- Tripel bilden einen Graph
- Terminologie: Resource, Triple/Statement, Node, Subject, Predicate, Object, Term

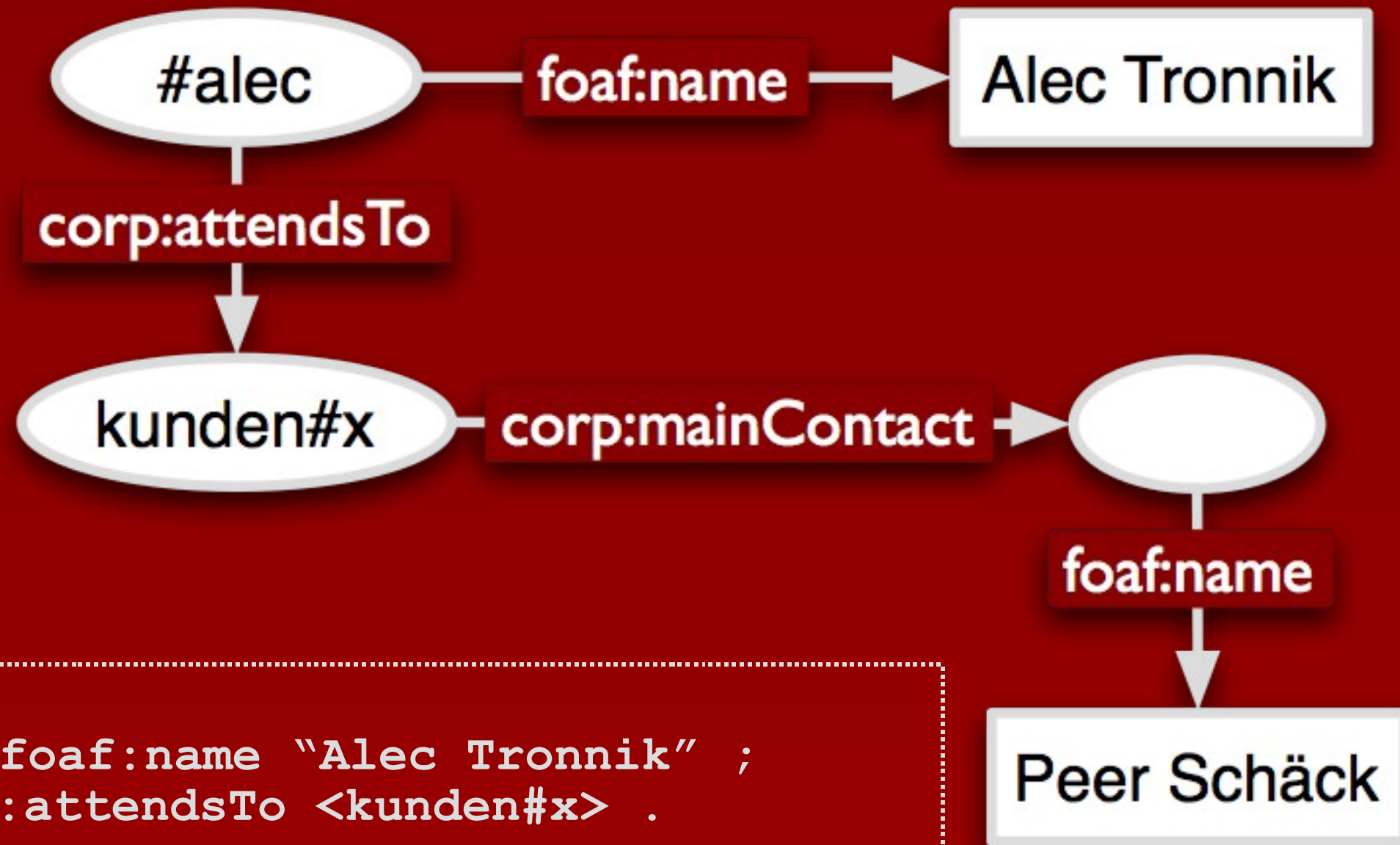
Semantic Web Technik: RDF + Syntax



```
...  
<rdf:Description rdf:about=#alec'>  
  <foaf:name>Alec Tronnik</foaf:name>  
  <corp:attendsTo  
rdf:resource="kunden#x" />  
</rdf:Description>  
  
<rdf:Description rdf:about=kunden#x">  
<corp:mainContact rdf:nodeID=bn1" />  
</rdf:Description>
```

RDF/XML

Semantic Web Technik: RDF + Syntax



```
...  
<#alec> foaf:name "Alec Tronnik" ;  
  corp:attendsTo <kunden#x> .  
  
<kunden#x> corp:mainContact _:bn1 .  
  
_:bn1 foaf:name "Peer Schäck" .  
...
```

Turtle / N3

Semantic Web Technik: Modelle

SKOS (Simple Knowledge Organization System)

- “Concept” als Basiskonzept
- Verschlagwortung (“Tagging”, “Folksonomies”)
- einfache Taxonomien (broader, narrower)
- Thesauri (broader, narrower, related to, altLabel)

RDFS (RDF Schema / Vocabulary Description Language)

- Basiskonzepte: “Class”, “Property”
- Klassen- und Eigenschafts-Hierarchien (subClassOf)
- Zuordnung von Eigenschaften zu Klassen
- Ermöglicht Klassifizierung durch Inferenz

Semantic Web Technik: Modelle

OWL (Web Ontology Language)

- Basiskonzepte: “Class”, “Property”, “Restriction”
- Weitergehende Beschreibungsstrukturen
- symmetrische, inverse, transitive, typisierte

Eigenschaften

- complementOf, disjointWith, union/intersectionOf
- min/maxCardinality
- Mapping (equivalentClass, sameAs)
- Versionierung
- OWL-Ontologien sind kombinierbar (inkl. RDFS)
- Verschiedene Subsets (Lite, DL)

Semantic Web Technik: SPARQL

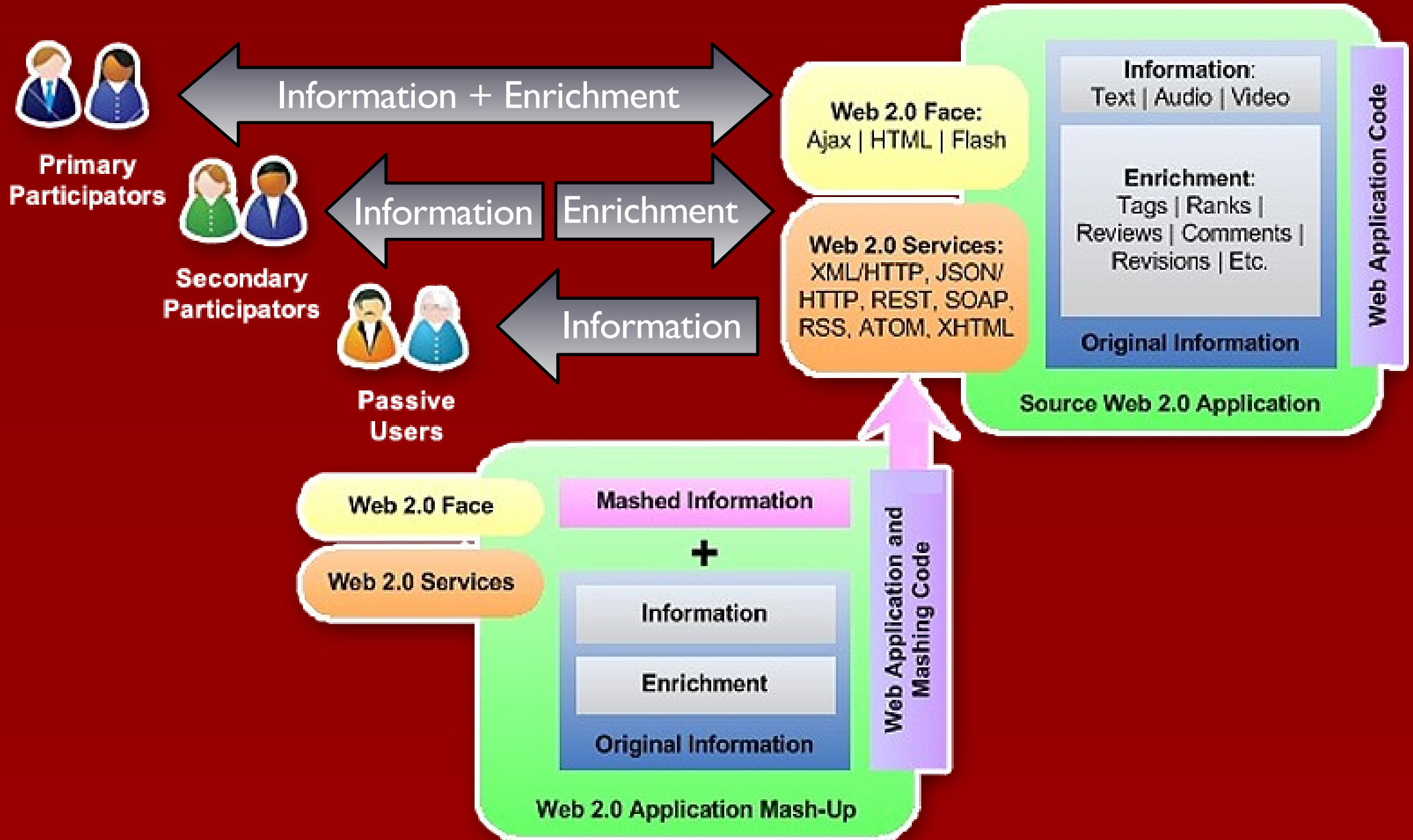
SPARQL (SPARQL Protocol And RDF Query Language)

- Erweitert das RDF-Triple-Modell um Graphen
- (Triple-Kontext/-Speicherort/-Sammlung)
- Externer Zugriff auf RDF-Speicher (bisher nur lesend)
- SELECT, DESCRIBE, CONSTRUCT, ASK
- Kann als einfaches Regelsystem verwendet werden

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX corp: <http://example.com/ns/corp#>
```

```
SELECT DISTINCT ?quelle ?k_name
WHERE {
  GRAPH ?quelle {
    ?mitarbeiter foaf:name "Alec Tronnik" .
    ?mitarbeiter corp:attendsTo ?kunde .
    ?kunde foaf:name ?k_name .
  }
}
```

Das Web 2.0 Modell



Quelle: web2.wsj2.com (Dion Hinchcliffe)

Web 2.0 Technik: Grundprinzipien

Permalinks und URL-Konventionen

- Stabile Web-Adressen (äquiv. zu IRIs im SemWeb)
- z.B. von Blog Posts
- `http://your.server.com/tag/darmstadt`

XML-Feeds für alles (hauptsächlich noch als RSS 2.0)

- Standard-Export-Format (item, title, description)
- implizite Semantik (z.B. foto-feed von flickr)

REST-ful APIs (Representational State Transfer)

- HTTP-Prinzipien (GET, POST, PUT, DELETE), z.B. APP
- Jeder Transfer enthält alle notwendigen Informationen

Web 2.0 Technik: Grundprinzipien

AJAX (Asynchronous JavaScript And XML)

- Interaktive Benutzeroberflächen (z.B. In-Page-Bewertung)
- Nachladen externer Skript-Bibliotheken

JSON (JavaScript Object Notation)

- Alternative zu XML
- Kann direkt im Browser ausgewertet werden

Mashups (Kombination von mehreren Applikationen)

- Minimaler Programmieraufwand
- Wiederverwendung ganzer Frontends (z.B. GMaps)

Web 2.0 Technik: Vereinfachte Mitwirkung

Blogs

- Trackbacks
- Aggregation

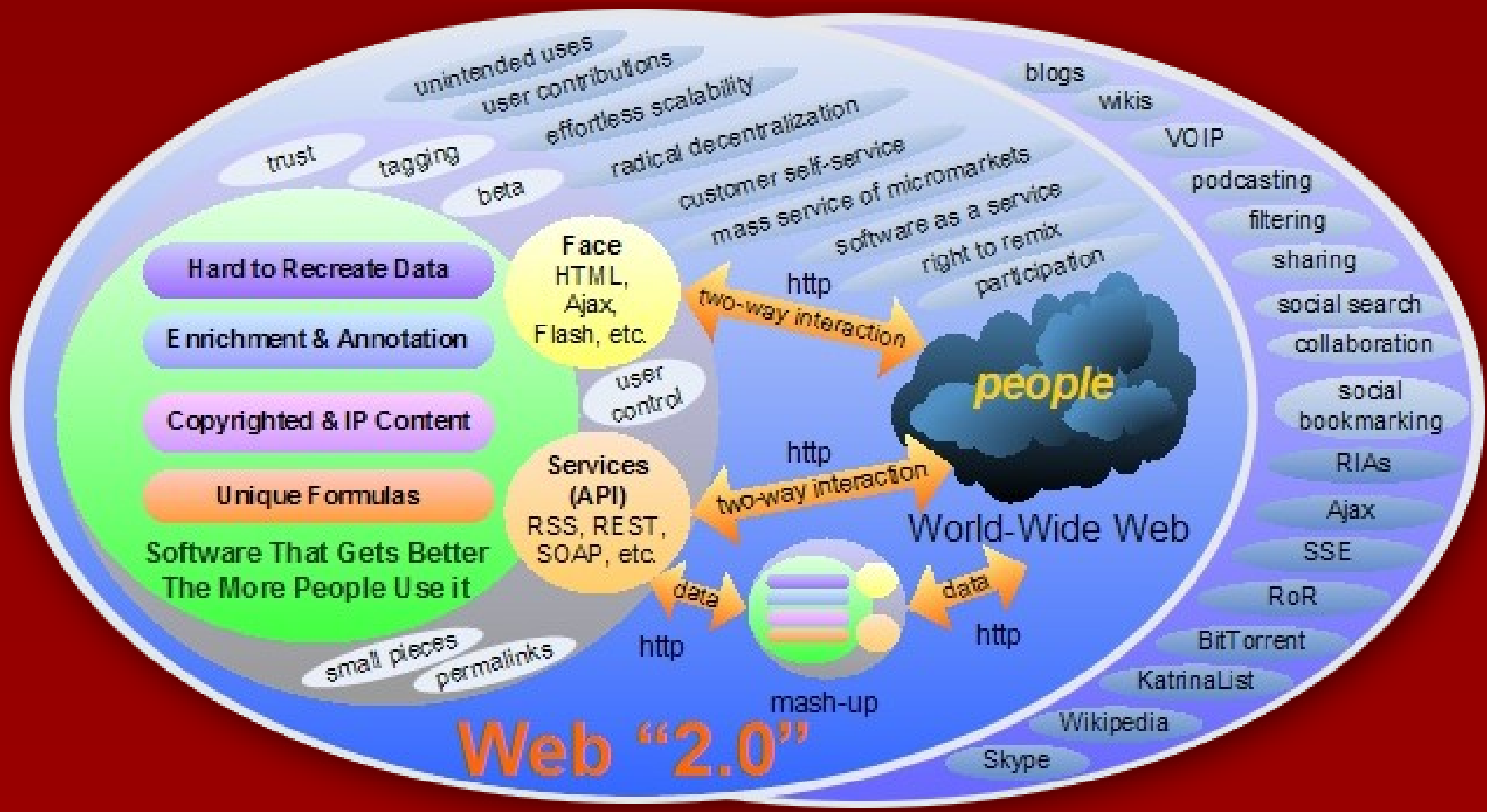
Tags/Folksonomien

- uneingeschränkte Massen-Indizierung
- Analyse von Häufigkeiten (“Collective Intelligence”)

Beschränkte, vorgegebene Annotierungen

- Bewertungen, Kommentare
- neu: Geo-Informationen
- Structured Blogging (einfache Erzeugung von z.B. μ Fs)

Web 2.0 = Alles, was trendy ist



Quelle: web2.wsj2.com (Dion Hinchcliffe)

Web 2.0: Jetzt auch mit Semantik!

Microformats

- Nutzen Semantik bestehender HTML-Elemente und -Attribute (abbr, dl, h1-h6, rel, class)
- Top-Down-Ansatz (vorgegebene Menge an Formaten)
- hCard (vCard), hCalendar (iCal), rel-tag, XFN (social relations)
- keine Namespaces (gewollt)
- “Humans first” (sichtbare Metadaten)

```
<span class="vevent">  
  <a class="url" href="http://iswc2006.semanticweb.org/">  
    <span class="summary">ISWC 2006</span> starts on  
    <abbr class="dtstart" title="2006-11-05">November 5</abbr>  
    in <span class="location">Athens, GA</span>  
  </a>  
</span>
```


Web 2.0: Jetzt auch mit Datenaustausch!

Live Clipboard

- Teil von Microsofts "Live" Initiative
- Erlaubt System-weites Copy&Paste von microformats
- Copy-By-Reference über Simple Sharing Extensions (SSE)

```
<liveclipboard version="0.92" source="http://..."  
xmlns:lc="http://www.microsoft.com/schemas/liveclipboard">  
<lc:data>  
<lc:format type="vcard" contenttype="application/xhtml+xml">  
<lc:item><div class="vcard uid">...</div></lc:item>  
</lc:format>  
</lc:data>  
</liveclipboard>
```

Semantic Web: Jetzt auch mit HTML!

RDFa

- Ähnlich zu Microformats, aber mit Namespaces und voller RDF-Funktionalität
- Nicht XHTML 1.x-konform

eRDF

- Ähnlich zu RDFa, aber mit XHTML 1.0-Konformität
- Nur ein RDF-Subset

GRDDL

- Mechanismus, um z.B. Microformats über XSLT in RDF/XML zu transformieren

SemWeb und Web 2.0: Daten-Ebene

	SemWeb Das Web als Datenbank	Web 2.0 Das Web als Plattform
Daten	<p>Strukturen:</p> <ul style="list-style-type: none">•IRIs, Tripel, Graphen, HTML Elemente <p>Formate:</p> <ul style="list-style-type: none">•RDF/XML, Turtle/N3, HTML (eRDF/a), XML <p>Formale Beschreibung:</p> <ul style="list-style-type: none">•RDFS, OWL	<p>Strukturen:</p> <ul style="list-style-type: none">•Feeds, Tags, HTML Elemente, JSON-Obj. <p>Formate:</p> <ul style="list-style-type: none">•RSS/ATOM, μFormats, OPML, JSON, XML <p>Formale Beschreibung</p> <ul style="list-style-type: none">• XMDP (μF)

SemWeb und Web 2.0: Protokoll-Ebene

	SemWeb Das Web als Datenbank	Web 2.0 Das Web als Plattform
Protokolle	<ul style="list-style-type: none">•Content-Negotiation•SPARQL¹•GRDDL²-Konventionen•HTML “Hooks”	<ul style="list-style-type: none">•einfache APIs (REST³ful)•Atom Publishing Protocol (APP)•Live Clipboard•URL-Konventionen

¹ SPARQL Protocol And RDF Query Language

² Gleaning Resource Descriptions from Dialects of Language

³ Representational State Transfer

SemWeb und Web 2.0: Tool-Ebene

	SemWeb Das Web als Datenbank	Web 2.0 Das Web als Plattform
Tools	<ul style="list-style-type: none">•Parser•RDF Stores•SPARQL Engines•Converter/XSLTs•Aggregatoren•Browser•Inferenz-Maschinen	<ul style="list-style-type: none">•API Wrapper•UI-Bibliotheken•AJAX•Widgets•Parser/XSLTs•Statistische Auswertungen

Unterschiedliche Ansätze (Zusammengefasst)

Semantic Web

Standardisierung erfolgt auf allen Ebenen, offene Daten und das Back-End (weniger die Anwendungen) stehen im Mittelpunkt.

Web 2.0

Kaum Standardisierung, keine explizite Semantik, aber einfache APIs (und Anwendungen!) zur Kombination und Erweiterung von Informationen mit (impliziter) Semantik (z.B. Landkarten-Mashups wie frappr)

Stärken

Semantic Web

- Universelles Datenbeschreibungsmodell
- Vereinfachter Datenzugriff
- Standardisierter Informationsgewinn
- Automatisierbare Informationsverarbeitung
- Mehrfache Wiederverwendung (von Informationen)

Web 2.0

- Vereinfachte Mitwirkung (z.B. In-Page Annotierung)
- Teil-Standardisierte Anwendungsprinzipien (z.B. Tags)
- Wiederverwendung ganzer Anwendungen
- Verfügbarkeit wertvoller Informationen

Problembereiche

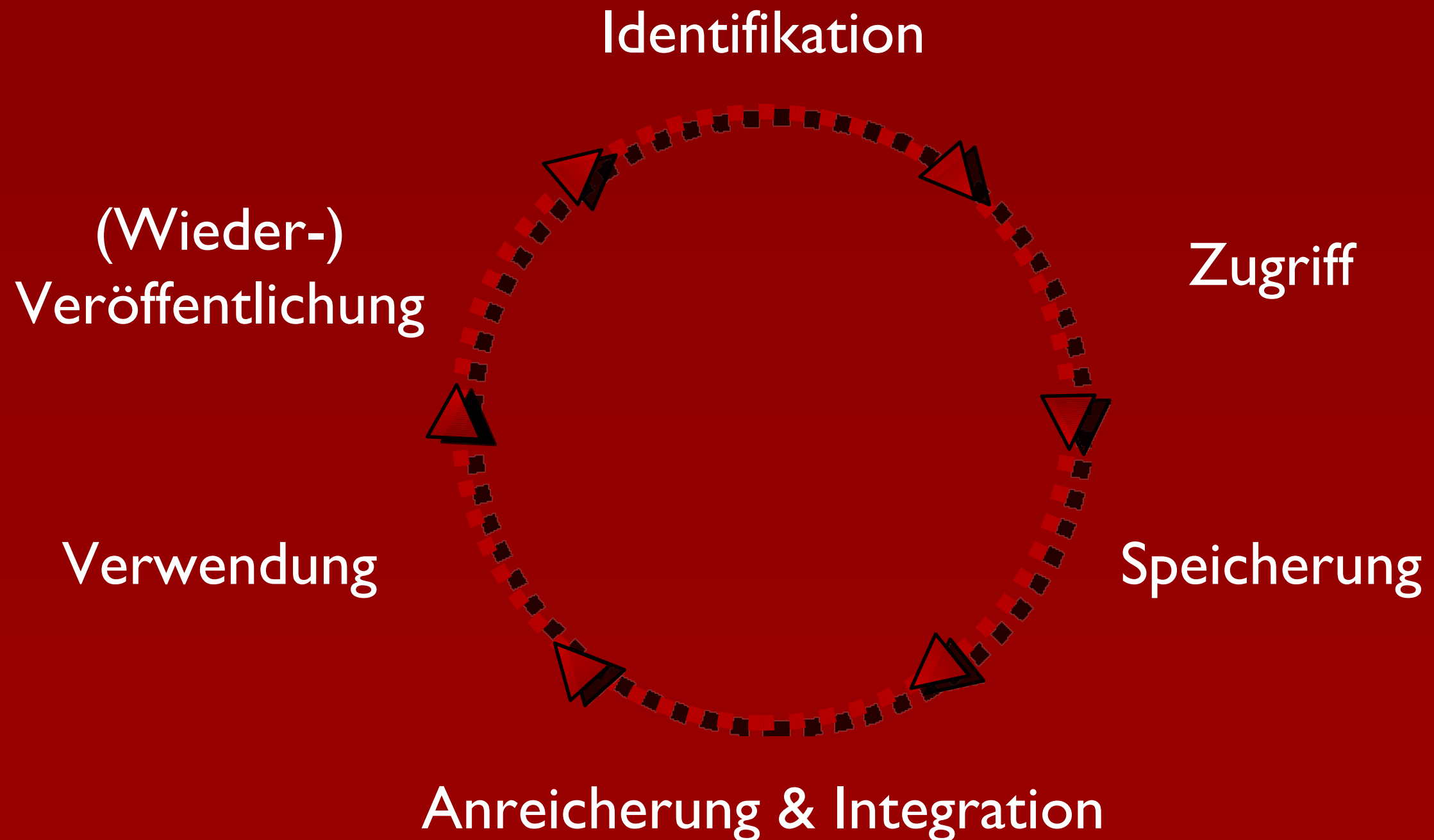
Semantic Web

- Time2Market (komplexe Technik / Eigenentwicklung?)
- Skalierbarkeit (keine anwendungsspez. Optimierung)
- Fehlende Entwickler und Experten
- Noch sehr akademisch, teilweise ungewohnt (OWA, KI-Touch, anderes Verständnis von Validierung)
- Das RDF/XML-Format

Web 2.0

- Keine direkte Wiederverwendung von kombinierten Informationen (kein “Mashup Chaining”, auch: Rechte)
- Performance (permanente Aktualisierungen)

Informations-Life-Cycle



Ideal-Kombination (Vorschlag)

Identifikation	•IRIs, HTML-Hooks, μ F, eRDF
Zugriff	•Feeds, LiveClipboard, GRDDL, Content Negotiation, REST, SPARQL
Speicherung	•RDF Store
Anreicherung & Integration	•RDFS, OWL, Tags, Web 2.0-Tools
Verwendung	•SemWeb-Tools, Web 2.0-ToolsProgramm-Code, Widgets, APIs
(Wieder-) Veröffentlichung	•RDF/XML, HTML (μ F, eRDF), SPARQL-Endpoints, (RSS), spezifische Formate

Fazit

Semantic Web und Web 2.0:

- beschreiben Entwicklungen für das “Web der nächsten Generation”
- teilen sich den Web-Technologie-Markt
- haben sowohl gegensätzliche, überschneidende und komplementäre Elemente
- können bei einer Kombination
- die Menge an nutzbaren Informationen vergrößern
- Entwicklungszeiten verkürzen
- Mehrwert-Generierung vereinfachen
- einfache Technologien als Brücken verwenden
- Wissenstechnologien “on Demand” einbeziehen