

Glossar

Zusammengestellt von Adrian Pohl.¹

API steht für „Application Programming Interface“, dt. „Schnittstelle für die Anwendungsprogrammierung“. Im Allgemeinen ist eine API die Schnittstelle eines Softwaresystems, die anderen Programmen eine Anbindung an das System ermöglicht.

Im Kontext des -> World Wide Web und von Linked Open Data sind insbesondere *Web-APIs* relevant. Web-APIs sind Schnittstellen, die dem -> *REST*-Paradigma für Webanwendungen folgen. Web-APIs erlauben den Aufbau von sogenannten *Mashups*, das sind Webanwendungen, die selbst aus einer Kombination anderer Webanwendungen resultieren.

BEACON ist ein einfaches Dateiformat. Mit BEACON werden Links auf Webseiten angegeben, die Informationen enthalten zu in Normdateien verzeichneten Dingen. Derzeit wird das Format vor allem für Personen genutzt, die mittels ihres Identifikators in der Gemeinsamen Normdatei (GND) identifiziert werden. Es eignet sich in seiner jetzigen Form aber auch zur Verwendung für Körperschaften und Schlagwörter sowie grundsätzlich auch für weitere Normdaten.²

Bibliographische Daten bestehen aus bibliographischen Beschreibungen. Eine bibliographische Beschreibung beschreibt eine bibliographische Ressource (Artikel, Monographie etc. – ob gedruckt oder elektronisch), um sie a) eindeutig zu identifizieren und b) auffindbar zu machen. Traditionellerweise erfüllt eine Beschreibung beide Zwecke gleichzeitig, indem sie Information liefert über: Autor(en) und Herausgeber, Titel, Verlag, Veröffentlichungsdatum und -ort, Identifizierung des übergeordneten Werks (z.B. einer Zeitschrift), Seitenangaben. Eine eindeutige Identifikation bibliographischer Ressourcen findet heute häufig über -> URIs statt, weil zunehmend entsprechende Identifikatoren wie DOIs oder URNs verwendet werden.

Blank Nodes sind Knoten eines RDF-Graphen, die keinen -> URI als Identifier haben und damit nicht weltweit referenziert werden können. Sie werden beispielsweise für anonyme Ressourcen (analog der logischen Existenzaussage „es gibt mindestens ein x, das ...“) verwendet oder für bekannte Ressourcen, denen noch kein URI zugewiesen wurde. Wenn möglich sollten Blank Nodes vermieden werden, da sie die Nutzbarkeit von RDF-Daten einschränken.

CC0-> Public Domain Dedication

Content Negotiation ist ein -> HTTP-Mechanismus, der es ermöglicht, zwischen verschiedenen Repräsentationen einer Ressource auszuwählen. Verschiedene Repräsentationen können sich etwa anhand der Sprache oder des gelieferten Formats unterscheiden. Linked-Data-Services stellen beispielweise Daten in verschiedenen -> RDF-Serialisierungen (-> Turtle, N3, RDF/XML, JSON-LD) bereit, so dass Datennutzer das von ihnen bevorzugte Format wählen können.

1 Der Ersteller des Glossars bedankt sich bei Louise Rumpf (Universitätsbibliothek Bamberg), deren umfangreiche Vorschläge zur Verbesserung des Glossars sehr wertvoll waren.

2 Mit leichten Überarbeitungen übernommen von <http://de.wikipedia.org/w/index.php?title=Wikipedia:BEACON&oldid=114357945>.



Cool URI. Der Ausdruck „Cool URIs“ wurde von Tim Berners-Lee in seinem Text „Cool URIs don't change“ geprägt. Cool URIs sind demnach -> URIs, die über die Zeit stabil sind. Der an Berners-Lee anknüpfende Text „Cool URIs for the Semantic Web“ nennt zusammenfassend drei Anforderungen für Cool URIs: Einfachheit, Stabilität und Handhabbarkeit („simplicity, stability and manageability“). Darüber hinaus wird gefordert, HTTP-URIs zu verwenden („Be on the web“) und eindeutig zu sein, indem Dinge und Dokumente, die sie beschreiben, unterschiedliche URIs bekommen.

FRBR steht für „Functional Requirements for Bibliographic Records“ und ist ein Datenmodell zur Strukturierung des „bibliographischen Universums“. Die „funktionalen Anforderungen für bibliographische Datensätze“ wurden als Entity-Relationship-Modell entwickelt, das grundlegende Entitäten, Beziehungen und Attribute enthält, die für eine Beschreibung bibliographischer Ressourcen verwendet werden.

Die wichtigsten Entitäten sind in der Gruppe 1 der FRBR enthalten, die das Ergebnis einer schöpferischen, intellektuellen oder künstlerischen Tätigkeit darstellen (-> WEMI). Die Gruppe 2 enthält die Entitäten Personen (person) und Körperschaften (corporate body). Gruppe 3 widmet sich mit den Entitäten concept, object, event und place schließlich der inhaltlichen Erschließung.

Freie Software (englisch „free software“) ist Software, die sicherstellt, dass die Endbenutzer (Privatpersonen, Organisationen, Firmen) die Freiheit haben, die Software zu verwenden, zu untersuchen, zu kopieren, weiterzugeben und zu modifizieren. Es geht daher nicht um den Preis der Software, sondern um die Freiheiten der Endnutzer; z. B. werden die Freiheitsrechte, die Software zu untersuchen und zu modifizieren, bei Freier Software immer durch Verfügbarkeit des Quellcodes garantiert.³

Freie Inhalte / Open Content. Als freie Inhalte (englisch „free content“), auch Open Content genannt, bezeichnet man Inhalte, deren kostenlose Nutzung und Weiterverbreitung urheberrechtlich erlaubt ist. Dies kann nach Ablauf von gesetzlichen Schutzfristen zutreffen, womit ursprünglich geschützte Werke gemeinfrei werden. Alternativ werden Inhalte als frei bezeichnet, wenn der Urheber oder Inhaber der vollumfänglichen Nutzungsrechte ein Werk unter eine freie -> Lizenz gestellt hat.⁴

Graph. Ein Graph im Allgemeinen ist eine abstrakte Struktur, die eine Menge von Objekten zusammen mit den zwischen diesen Objekten bestehenden Verbindungen repräsentiert. Die mathematischen Abstraktionen der Objekte werden dabei Knoten des Graphen genannt, die Verbindungen zwischen Knoten heißen Kanten. Ein Beispiel für einen Graphen ist der Plan eines U-Bahn-Netzes.

Ein -> RDF-Graph im Speziellen ist ein benannter gerichteter Graph. „Gerichtet“ ist der Graph, weil die Richtung der Kanten durch Pfeile gekennzeichnet ist. „Benannt“ ist der Graph, weil die Kanten typisiert sind, das heißt mit URIs versehen werden. Tripel-Subjekte und -Objekte bilden die Knoten eines RDF-Graphen, RDF-Prädikate sind in einem RDF-Graphen die Kanten.

Der Beispiel-RDF-Graph in graphischer Darstellung (nach Konvention werden Ressourcen durch Ellipsen und Literale durch Rechtecke symbolisiert) benutzt folgende -> Namensraum-Präfixe:

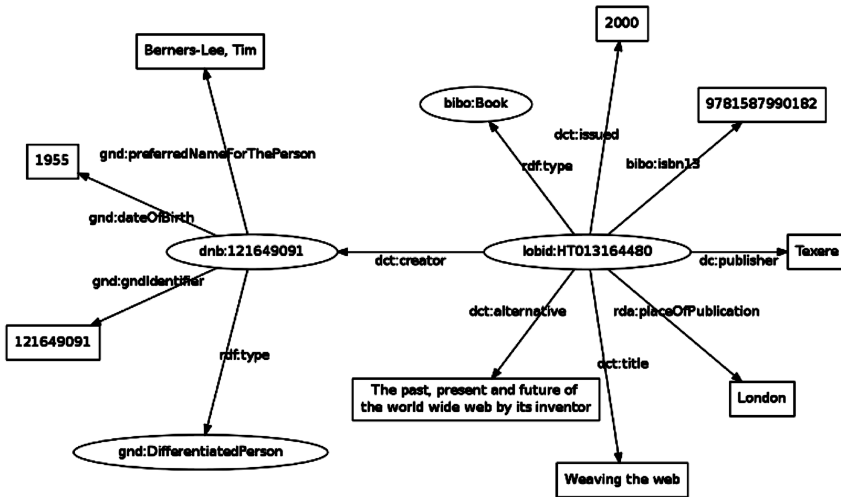
³ Mit leichten Überarbeitungen übernommen von http://de.wikipedia.org/w/index.php?title=Freie_Software&oldid=112515943.

⁴ Quelle: http://de.wikipedia.org/w/index.php?title=Freie_Inhalte&oldid=112075189

```

.
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix dnb: <http://d-nb.info/gnd/> .
@prefix gnd: <http://d-nb.info/standards/elementset/gnd#> .
prefix lobid: <http://lobid.org/resource/> .
@prefix rda: <http://rdvocab.info/Elements/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

```



HTTP. Das Hypertext Transfer Protocol (HTTP) ist ein Protokoll zur Übertragung von Daten im -> Word Wide Web. Web-Agenten (z. B. Web-Browser) kommunizieren über dieses Protokoll mit Web-Servern.

Klasse. RDF-Klassen werden in einem -> Vokabular definiert. Ihnen können Dinge unter Verwendung der -> Property `rdf:type` zugewiesen werden. Ein Vokabular kann zudem Relationen festlegen, die Klassen mit anderen Klassen und mit Properties haben.

Linked Data. Während -> *Open Data* keine Datenformate vorgibt, solange es sich um offen dokumentierte Formate handelt, geht es bei Linked Data um die Etablierung von Best Practices für im Web integrierte Daten. Diese Best Practices werden im Rahmen des -> *World Wide Web Consortium (W3C)* entwickelt.

Die vier von Tim Berners-Lee formulierten Linked-Data-Prinzipien lauten:

1. Benutze URIs als Namen für Dinge.
2. Benutze HTTP-URIs, damit Menschen die Namen nachschlagen können.
3. Wenn jemand eine URI nachschlägt, liefere nützliche Informationen auf Basis der Standards (RDF, SPARQL).
4. Verlinke zu anderen URIs, so dass mehr Dinge entdeckt werden können.

Linked Data baut also auf den bestehenden Web-Standards Uniform Resource Identifiers (-> URI) und Hypertext Transfer Protocol (-> HTTP) auf und ergänzt diese durch das

Datenmodell -> RDF zur Repräsentation von Information, -> SPARQL zu Abfrage von RDF-Daten und RDFS sowie OWL zur Schaffung von ->Vokabularen/Ontologien.

Linked Open Data sind offene Daten (-> *Open Data*), die gemäß Linked-Data-Prinzipien (-> Linked Data) unter Nutzung der entsprechenden offenen -> W3C-Standards bereitgestellt werden.

Literale sind Unicode-Zeichenketten, die mit einem Datentyp versehen sind und optional mit einem Sprachcode markiert werden können. In einem RDF-Tripel sind Literale allein in Objektstellung erlaubt.

Lizenz. Allgemein ist eine Lizenz (v. lat. licet, „es ist erlaubt“) eine Erlaubnis, Dinge zu tun, die ansonsten verboten sind. Durch einen Lizenzvertrag erteilt der Inhaber eines geschützten Rechts (z. B. eines Urheberrechts) dem Lizenznehmer ein definiertes Nutzungsrecht. Besonders bekannte Lizenzverträge im World Wide Web sind die Musterverträge von Creative Commons. Da es - insbesondere in Europa - Probleme mit der Anwendung von Creative-Commons-Lizenzen auf Daten und Datenbanken gibt, wurden verschiedene Open-Data-Lizenzen entwickelt, u. a. von Open Data Commons.

Es gibt unter anderem folgende Lizenztypen, die sich in den Nutzungsbedingungen unterscheiden:

1. *Attribution.* Der Name des Urhebers muss genannt werden.
2. *Share-Alike.* Das Werk muss nach Veränderungen (z.B. Bearbeitungen, Zusammenführungen mit anderen Daten etc.) unter der gleichen Lizenz weitergegeben werden.
3. *Non-Commercial.* Das Werk darf nicht für kommerzielle Zwecke verwendet werden. Darüber hinaus gibt es die -> Public Domain Dedication, bei deren Nutzung sämtliche Rechte - soweit möglich - an einem Werk aufgegeben werden.

LOD-Cloud. Die Summe aller im Web publizierten und untereinander verlinkten RDF-Datenquellen wird als LOD-Cloud bezeichnet. In der Regel werden auch die nicht-offen lizenzierten Linked-Data-Quellen dazugezählt, weshalb der Name „Linked-Data-Cloud“ treffender wäre.

Illustriert wird die LOD-Cloud durch die von Richard Cyganiak und Anja Jentzsch seit 2007 erstellte Graphik, auf der Linked-Data-Quellen im Netz und ihre Verlinkungen untereinander dargestellt werden. Siehe die Abbildungen der Linked-Data-Cloud in Abschnitt 2.1 des Beitrags von Pohl/Danowski in diesem Sammelband.

Named Graph. Ein Named Graph ist ein -> RDF-Graph (d. h. eine Menge von Tripeln), der durch einen URI identifiziert ist. Dies ermöglicht es, den Graphen zu referenzieren, ihn separat zu bearbeiten und weiterzuverwenden. Momentan wird am -> W3C daran gearbeitet, Named Graphs zu einem offiziellen Standard zu erheben. Für den Datenaustausch zwischen verteilten Quellen und für Provenienzinformationen sind Named Graphs sehr nützlich.

Namensraum. Der Namensraum ist im Kontext von Linked Data ein -> HTTP-URI, der wiederkehrender Bestandteil von URIs einzelner Dinge (etwa der Elemente eines Vokabulars) ist. Zum Beispiel haben die Elemente des Dublin-Core-Element-Set den Namensraum <http://purl.org/dc/elements/1.1/>. Namensräume werden von dem Besitzer der jeweiligen Domain kontrolliert.

In RDF-Serialisierungen wie -> Turtle oder RDF/XML können häufiger verwendete Namensräume durch ein Präfix (z.B. dc:) abgekürzt werden, um den Umfang des Dokuments zu verkürzen und/oder die Lesbarkeit zu erleichtern. So steht dc:creator für den URI <http://purl.org/dc/elements/1.1/creator>, wenn dc: als Präfix für den Namensraum

<http://purl.org/dc/elements/1.1/> festgelegt wurde. (Siehe auch -> Turtle, -> Reasoning und -> Graph für den beispielhaften Gebrauch von Präfixen für Namensräume.) Unter prefix.cc können gebräuchliche Namensraum-Präfixe insbesondere von Vokabularen nachgeschlagen werden.

Ontologie/Vokabular. In der Tatsache, dass die Prädikate zur Beschreibung der Beziehung zwischen Subjekt und Objekt selbst URIs sind (-> RDF), liegt ein grundlegender Unterschied zwischen Linked Data und dem World Wide Web. Zwar basiert auch das Web auf Links, allerdings sind diese nicht typisiert. Das heißt, in einem HTML-Dokument steht zwar, dass dieses Dokument auf ein anderes verlinkt, welcher Art diese Verlinkung ist, bleibt allerdings allenfalls für Menschen nachvollziehbar. Ob auf das Dokument zum Zweck eines Hinweises, einer Rezension etc. verlinkt wird, ist nicht explizit angegeben. Im Linked-Data-Netz ist dies anders, hier ist jede Verlinkung typisiert (siehe das Werk-Autor-Beziehungs-Beispiel unter ->RDF).

Ontologien und Vokabulare definieren sogenannte Properties, die als Prädikate in RDF-Tripeln benutzt werden können sowie deren Beschreibungen und Beziehungen untereinander. Darüber hinaus legt ein Vokabular Klassen fest, denen Dinge zugeordnet werden können, und deren Hierarchien (Unter- und Überklassen). Auf Basis einer Menge von Tripeln und der darin verwendeten Ontologien können implizite Informationen explizit gemacht werden, d.h. weitere Tripel generiert werden (-> Reasoning).

Open Data / Offene Daten. Open Data ist ein sehr weiter Begriff, der sich auf den Teil von -> Open Knowledge bezieht, der aus strukturierten, möglichst maschinell prozessierbaren Daten besteht. Entwicklungen hin zu Open Data gibt es bereits in vielen verschiedenen Bereichen, u. a. veröffentlichen Regierungen und öffentliche Verwaltungen Daten in Open-Data-Portalen (Open Government Data), Wissenschaftler publizieren Forschungsdaten wie z. B. Klimadaten oder Genome unter offenen Lizenzen (Open Science Data) und Bibliotheken veröffentlichen ihre Titel-, Norm-, Bestandsdaten etc. (Open Library Data).

Open Knowledge ist ein Oberbegriff für verschiedene spezifischere Konzepte wie zum Beispiel -> Freie Inhalte / Open Content, -> Open Data / Offene Daten, Open Government Data (Offene Informationen der Regierung und öffentlichen Verwaltung).

Eine Definition von „Open Knowledge“ liefert die Open Knowledge Definition (<<http://opendefinition.org/okd/>>) der Open Knowledge Foundation, die folgende zentrale Bedingungen an offenes Wissen stellt:

- *Offener Zugang* zum Wissen, d.h. es muss offen und kostenlos als Gesamtheit (vorzugsweise im Internet) zugänglich sein.
- *Offene Standards*, d.h. das Wissen muss in einem offen dokumentierten und nicht-proprietären Format vorliegen.
- *Offene Lizenzen*, d.h. die Daten müssen (als Einzeldatum und als Sammlung) unter einer offenen Lizenz publiziert werden. Als einzige Einschränkungen sind erlaubt: die Forderung nach einer Attribution der Quelle sowie die Share Alike, womit eine Lizenzierung abgeleiteter Werke unter denselben Bedingungen gefordert wird.

Property. Properties beschreiben die Relation zwischen Subjekt und Objekt eines RDF-Tripels, wobei sie als Prädikat dieses Tripels fungieren. Properties und ihre Relationen zu -> Klassen und anderen Properties werden in einem -> Vokabular definiert.

Provenienz meint die Herkunft eines bestimmten Dinges - eines Kunstwerks, eines Dokuments oder von Daten. Provenienzinformationen geben Aufschluss über Entitäten und Prozesse,

die an der Produktion oder Auslieferung einer Ressource beteiligt waren oder sie anderweitig beeinflusst haben. Die Provenienz liefert die entscheidende Grundlage, um die Authentizität zu bewerten und Vertrauen und Reproduzierbarkeit zu ermöglichen. Provenienzinformationen sind eine Form von Metadaten und können selbst zu wichtigen Datensätzen mit eigener Provenienz werden. (Vgl. in diesem Band: Eckert, Die Provenienz von Linked Data.)

Public Domain Dedication. Mit einer Public Domain Dedication - oftmals auch als „Waiver“ bezeichnet - können Anwender ihre eigenen Werke in die Gemeinfreiheit überführen, d. h. eventuell an dem Werk bestehende geistige Eigentumsrechte abgeben. Wenn dies - wie in Deutschland - rechtlich nicht möglich ist, fungiert eine Public Domain Dedication i. d. R. als Lizenzvertrag ohne einschränkende Lizenzbedingungen. Die Creative Commons Public Domain Dedication (CC0) und die Public Domain Dedication and License von Open Data Commons sind die bekanntesten Waiver.

RDF. Das Resource Description Framework (RDF) ist ein Datenmodell zur Repräsentation von Information in der elementarsten Form: als einzelne Aussage. RDF-Aussagen bestehen aus drei Teilen, aus diesem Grund spricht man auch von einem RDF-Tripel. Die drei Teile werden ‚Subjekt‘, ‚Prädikat‘ und ‚Objekt‘ genannt. Die Aussage „Dieser Text wurde verfasst von Adrian Pohl.“ besteht aus Subjekt („Dieser Text“), Prädikat („wurde verfasst von“) und Objekt („Adrian Pohl“). In RDF lässt sich das Ganze wie folgt ausdrücken.

```
<> <http://purl.org/dc/elements/1.1/contributor> „Adrian Pohl“ .
```

Dies ist eine gültige RDF-Repräsentation der obigen Aussage. Das Subjekt „<>“ bezieht sich auf das Dokument, in dem das Tripel vorkommt, die Prädikat-URI steht für die creator-Property des Dublin Core Element Sets und „Adrian Pohl“ ist selbst keine URI, sondern ein sogenanntes -> Literal.

Literale dürfen in RDF-Tripeln nur im Objekt vorkommen. Subjekt und Prädikat müssen -> URIs sein. Um die Aussage des obigen Tripels eindeutig zu machen, sollte statt einem Namensliteral ein URI für die entsprechende Person (z.B. <http://d-nb.info/gnd/14326723X>) angegeben werden.

RDF ist ein abstraktes Modell, das in verschiedenen Formaten (Serialisierungen) ausgedrückt werden kann. Bekannte RDF-Serialisierungen sind NTriples und RDF/XML sowie für menschliche Lesbarkeit besser geeignete Serialisierungen wie -> Turtle und Notation3 (N3). RDF-Tripel können auch graphisch dargestellt werden: Nach Konvention werden Ressourcen, die Subjekt oder Objekt eines Tripels sind, durch Ellipsen und Literale durch Rechtecke symbolisiert (vgl. -> Graph).

RDFa (für „RDF in Attributes“) ist eine -> W3C-Empfehlung, die das Einbetten von RDF-Aussagen in Webseiten ermöglicht. Gemeinsam mit Mikroformaten und Microdata zählt es zu den gebräuchlichsten Methoden, (X)HTML-Seiten mit maschinenlesbaren Zusatzinformationen aufzuwerten.

Reasoning oder Inferencing (Inferenzieren) bezeichnet den Vorgang, auf Basis bestehender RDF-Daten neue Verlinkungen zu generieren. Die Grundlage bilden die den verwendeten Klassen und Properties zugrundeliegenden -> *Vokabulare/Ontologien*. Mit Reasoning wird gewissermaßen implizites Wissen, das in der Kombination von einer Datenbasis und den darin genutzten Ontologien liegt, explizit gemacht.

Im Folgenden ein einfaches Beispiel, geschrieben in der -> Turtle-Syntax.

Wir haben eine Datenbasis bestehend aus dem Tripel:

```
<http://viaf.org/viaf/85312226> a <http://xmlns.com/foaf/0.1/Person> .
```

Mit diesem Tripel wird das durch <http://viaf.org/viaf/85312226> identifizierte Ding der Klasse `foaf:Person` zugeordnet. Im FOAF-Voklabular wird die Klasse `:Person` wie folgt definiert:

```
.
@prefix owl: <[http://www.w3.org/2002/07/owl#]> .
@prefix foaf: <[http://xmlns.com/foaf/0.1/]> .
foaf:Person
  a rdfs:Class, owl:Class ;
  rdfs:label „Person“ ;
rdfs:subClassOf
<http://www.w3.org/2000/10/swap/pim/contact#Person>,
<http://www.w3.org/2003/01/geo/wgs84_pos#SpatialThing>, foaf:Agent .
```

Das FOAF-Vokabular sagt unter anderem über `foaf:Person` aus, dass es sich um eine ->Klasse (`rdfs:Class`, `owl:Class`) handelt, die mit dem menschenlesbaren Etikett „Person“ versehen ist (`rdfs:label`). Für das Reasoning ist hier allerdings allein die Aussage mit der Property `rdfs:subClassOf` relevant. Damit wird ausgesagt, dass jede Instanz der Klasse `foaf:Person` automatisch auch eine Instanz der drei aufgezählten Klassen ist. Somit können wir auf Basis des FOAF-Vokabulars und unserer Datenbasis folgende Tripel inferieren, wodurch die Datenbasis vervierfacht wird:

```
<http://viaf.org/viaf/85312226> a <http://www.w3.org/2000/10/
swap/pim/contact#Person>, <http://www.w3.org/2003/01/geo/wgs84_
pos#SpatialThing>, foaf:Agent .
```

Ressource / Repräsentation. Der Begriff ‚Ressource‘ wird sehr allgemein definiert (vgl. etwa RFC 3986). Eine Ressource (engl. resource) ist all jenes, was durch einen -> URI identifiziert werden kann. Neben Webseiten, Services, Fotos und anderen Dingen, die über -> HTTP abgerufen werden können, gehören auch raumzeitliche Dinge dazu wie Personen und Orte sowie abstrakte Dinge. Somit wird der Ausdruck „Ressource“ verwendet wie die Ausdrücke „Ding“ oder „Entität“.

Die *Repräsentation* einer Ressource im Web ist das, was zurückgeliefert wird, wenn ich auf den URI der Ressource zugreife. Repräsentationen derselben Ressource können sich sowohl im Laufe der Zeit als auch zu einem konkreten Zeitpunkt unterscheiden (vgl. -> Content Negotiation).

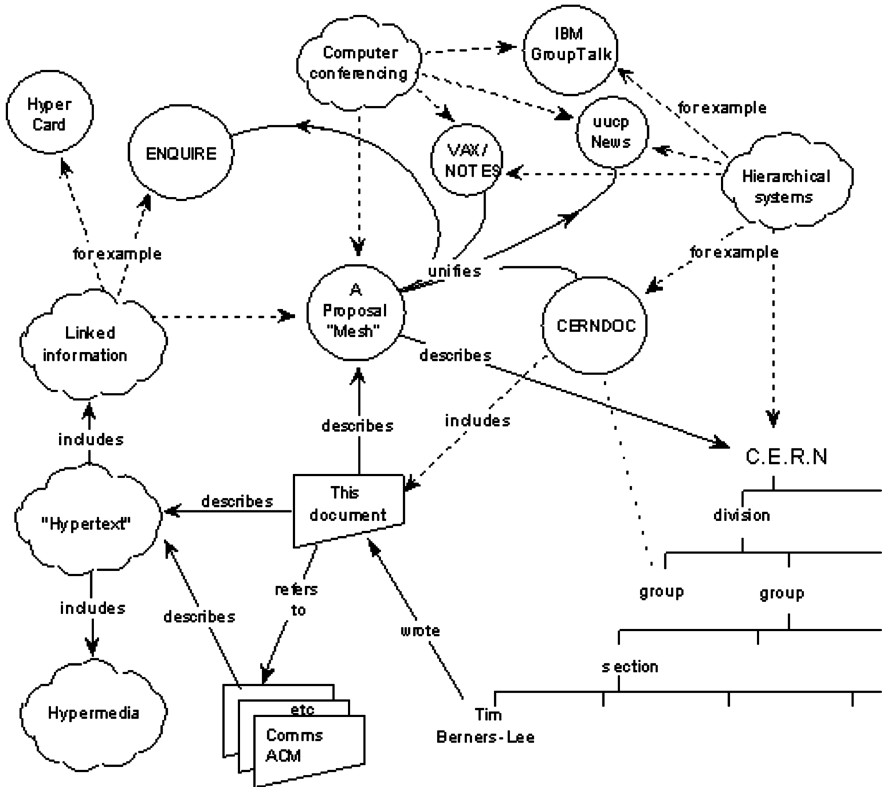
REST ist ein Paradigma für Software-Architekturen in verteilten Systemen wie dem World Wide Web, die *HTTP* oder ähnliche Protokolle nutzen. Jede *Ressource* wird dabei mit einem *Cool URI* angesprochen, weshalb Cool URIs stark mit dem REST-Ansatz verknüpft sind. REST vereinheitlicht Schnittstellen auf eine überschaubare und standardisierte Menge von Aktionen (für HTTP-basierte Schnittstellen sind dies GET, POST, PUT, DELETE).

Semantic Web. Die Idee des Semantic Web steckte bereits in Tim Berners-Lees Vorschlag für das -> World Wide Web.

Es geht dabei darum, nicht nur ganze Dokumente im Web zu veröffentlichen, sondern auch Daten in Form einzelner Aussagen (Tripel).

„Semantisch“ soll dieses Web sein, weil Ontologien es Maschinen erlauben sollen, auf Basis bestehender Informationen durch automatisierte Schlussfolgerungen neue Informationen zu generieren, die in den Daten bisher nicht enthalten sind. Da allerdings von Menschen generierte Daten immer fehlerhaft sind, handelt es sich hier um eine Utopie, weshalb oft der weniger verheißungsvolle Ausdruck „Linked Data“ bevorzugt wird,

der den Vernetzungscharakter betont und die semantische Ebene der Schlussfolgerungen und Wissensgenerierung außen vor lässt.



SKOS steht für „Simple Knowledge Organization System“ und ist das standardmäßig genutzte ->Vokabular für die RDF-Repräsentation von Thesauri und anderen Dokumentations-sprachen. SKOS beinhaltet Relationen wie „broader“, „narrower“ und „related terms“ und ermöglicht die Spezifikation bevorzugter und alternativer Ausdrücke für die Präsentation z.B. einer Klassifikationsstelle in verschiedenen Sprachen.

SPARQL. Das Akronym „SPARQL“ steht für „SPARQL Protocol And RDF Query Language“. Bei SPARQL handelt es sich um die Abfragesprache für ->Triple Stores.

Triple Store. Ein Triple-Store ist eine Datenbank, die auf ->RDF-Tripeln basiert, im Gegensatz z.B. zu einer relationalen Datenbank, die auf einer Tabellenstruktur aufgebaut ist.

Turtle ist ein einfaches, menschenlesbares RDF-Format. Die wesentlichen Merkmale seien hier kurz dargestellt.

Die Deklaration von *Präfixen* am Anfang eines Turtle-Dokuments erlaubt es, URIs im weiteren Verlauf des Dokuments abzukürzen. Präfixe werden wie folgt deklariert:

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
```


@prefix dcterms: <http://purl.org/dc/terms/> .

Nach Definition dieser Präfixe kann man die in der Deklaration genannte Vokabular-URI im Folgenden „unterschlagen“ und durch das jeweilige Präfix ersetzen. Beispiel („<>“ verweist immer auf das Dokument, in dem das Tripel gespeichert ist):

<> dcterms:contributor <http://d-nb.info/gnd/14326723X> .

Neben URIs lassen sich auch Tripel mit gleichem Subjekt vereinfachen, indem das Subjekt nur einmal genannt wird. Zu diesem Zweck findet das Semikolon Einsatz, um diese gekürzten Tripel aufzulisten. Zum Beispiel:

```
<http://d-nb.info/gnd/14326723X>
  a foaf:Person ;
  foaf:name „Adrian Pohl“ .
```

Ebenso können Tripel vereinfacht werden, die das gleiche Subjekt und Prädikat aufweisen. Die verschiedenen aufgezählten Objekte werden durch Kommas getrennt. Zum Beispiel:

```
<> dcterms:contributor <http://d-nb.info/gnd/14326723X>, <http://
d-nb.info/gnd/1026364728> .
```

URI. Ein Uniform Resource Identifier (URI) ist ein Identifikator, der im RFC 3986 definiert ist und einem vorgegebenen Schema folgt. Allseits bekannte URIs sind URLs, DOIs oder URNs.

Beispiele für URIs sind:

- http://de.wikipedia.org/wiki/Uniform_Resource_Identifier
- <ftp://example.org/resource.txt>
- pohl@hbz-nrw.de
- <urn:isbn:0-06-251586-1>
- <file:///home/adrian/bsp.txt>

Für ->Linked Data wird allein die Benutzung von URIs empfohlen, die auf dem HTTP-Protokoll basieren (HTTP-URIs). Diese HTTP-URIs werden somit auch zur Identifikation raum-zeitlicher Dinge wie Personen, Bauwerken etc. oder abstrakten Dingen wie Klassifikationsstellen, Konzepten etc. genutzt und nicht nur zur Identifikation von Online-Ressourcen wie etwa HTML-Seiten.

Waiver -> Public Domain Dedication

WEMI oder „WEMI-Modell“ bezieht sich auf die Entitäten der Gruppe 1 in *FRBR*. Das Akronym „WEMI“ steht dabei für die Anfangsbuchstaben dieser vier Entitäten, die das Ergebnis einer schöpferischen, intellektuellen oder künstlerischen Tätigkeit darstellen:

- **Work** (dt. Werk): Ein Werk, beispielsweise ein Roman als abstraktes Objekt (z. B. *Die Leiden des jungen Werther* von Goethe)
- **Expression** (dt. Expression): Der Ausdruck eines Werkes, beispielsweise eine Übersetzung (z. B. die Übersetzung des *Werther* von Pierre Leroux)
- **Manifestation** (dt. Manifestation): Eine manifestierte Ausgabe einer Expression, beispielsweise eine Auflage eines Buches (z. B. die in Paris im Jahr 1841 erschienene Leroux'sche Übersetzung)
- **Item** (dt. Exemplar): Ein konkretes raum-zeitliches Ding, beispielsweise ein Exemplar eines Buches (z. B. das Exemplar mit der Signatur „Yv 7991/1“ in der Staatsbibliothek zu Berlin)⁵

⁵ Mit leichten Anpassungen übernommen von http://de.wikipedia.org/w/index.php?title=Functional_Requirements_for_Bibliographic_Records&oldid=113340752.

World Wide Web. Das World Wide Web (kurz Web oder WWW aus dem Englischen für: „Weltweites Netz“) ist ein über das Internet abrufbares System von elektronischen Hypertext-Dokumenten, die durch Hyperlinks miteinander verknüpft sind und über die Protokolle HTTP bzw. HTTPS übertragen werden.

Zur Nutzung des World Wide Web wird ein Webbrowser benötigt, welcher die Daten vom Webserver holt und zum Beispiel auf dem Bildschirm anzeigt. Der Benutzer kann den Hyperlinks im Dokument folgen, die auf andere Dokumente verweisen, gleichgültig, ob sie auf demselben Webserver oder einem anderen gespeichert sind. Dadurch ergibt sich ein weltweites Netz aus Webseiten. Das Verfolgen der Hyperlinks wird oft als Internetsurfen bezeichnet.

Das WWW wird im allgemeinen Sprachgebrauch oft mit dem Internet gleichgesetzt, obwohl es jünger ist und nur eine von mehreren möglichen Nutzungen des Internets darstellt. Andere Internet-Dienste wie E-Mail, IRC und Telnet sind nicht in das WWW integriert.⁶

World Wide Web Consortium (W3C). Das World Wide Web Consortium (kurz W3C) ist das Gremium zur Standardisierung der das World Wide Web betreffenden Techniken. Es wurde am 1. Oktober 1994 am MIT Laboratory for Computer Science in Cambridge (Massachusetts) gegründet. Das W3C ist eine Mitgliedsorganisation. Mitglieder sind Softwarefirmen, Telekommunikationsunternehmen, Hochschulen, Forschungseinrichtungen, Normungsorganisationen und öffentliche Einrichtungen. (Zu den W3C-Mitgliedern gehören: Microsoft, IBM, Google, Apple, Samsung, Deutsche Telekom, ebay und aus der Bibliothekswelt: OCLC, Library of Congress, Deutsche Nationalbibliothek, hbz.)

Gründer und Vorsitzender des W3C ist Tim Berners-Lee, der auch als der Erfinder des World Wide Web bekannt ist. Das W3C entwickelt technische Spezifikationen und Richtlinien mittels eines ausgereiften, transparenten Prozesses, um maximalen Konsens über den Inhalt eines technischen Protokolls, hohe technische und redaktionelle Qualität und Zustimmung durch das W3C und seiner Anhängerschaft zu erzielen.⁷

⁶ Quelle: http://de.wikipedia.org/w/index.php?title=World_Wide_Web&oldid=111959628

⁷ Mit leichten Anpassungen übernommen von http://de.wikipedia.org/w/index.php?title=World_Wide_Web_Consortium&oldid=112632072.

