Diese Präsentation

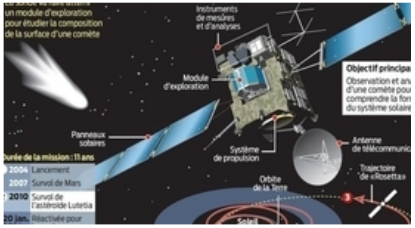
Linked Data - An Introduction on Speed

Adrian Gschwend, Zazuko GmbH

a glace fond lentement» à la nférence sur la Syrie

VE-2 — Les discussions entre les deux délégations nnes ont repris mercredi à Genève. Le représentant spécial NU et de la Ligue arabe Lakhdar Brahimi a souligné que le reste «très profond».

r à 18h50 1 Commentaire



Les infographies du «Matin»

Telles que publiées dans votre quotidien. Découvrez notre galerie.

Les plus belles photos de l'année 2013

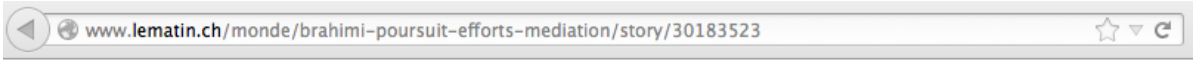
Découvrez les meilleures photos prises aux quatre coins de la planète.

Notre sélection d'images

Web of Documents

- Seit den 90ern
- Link (URL) pro Seite
- Eine Seite repräsentiert typischerweise ein Dokument
- Links auf weitere Seiten
- Für Menschen

URL als eindeutiger Identifikator



URL als eindeutiger Identifikator

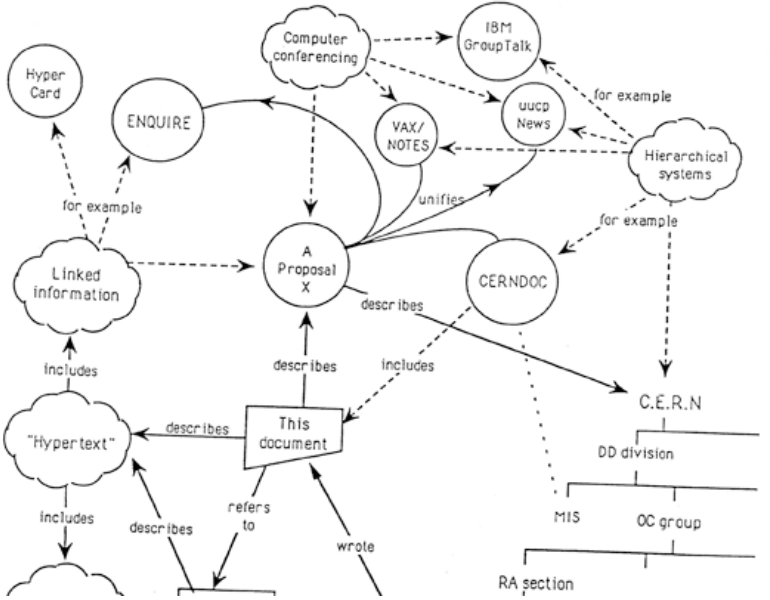
- Wir geben die URL händisch ein
- Benutzen eine Suchmaschine
- Oder folgen den Links innerhalb von Dokumenten
- Die Links sind nicht typisiert,
- sprich wir wissen nicht, was die Beziehung zum aktuellen Dokument ist

Wer hats erfunden?

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

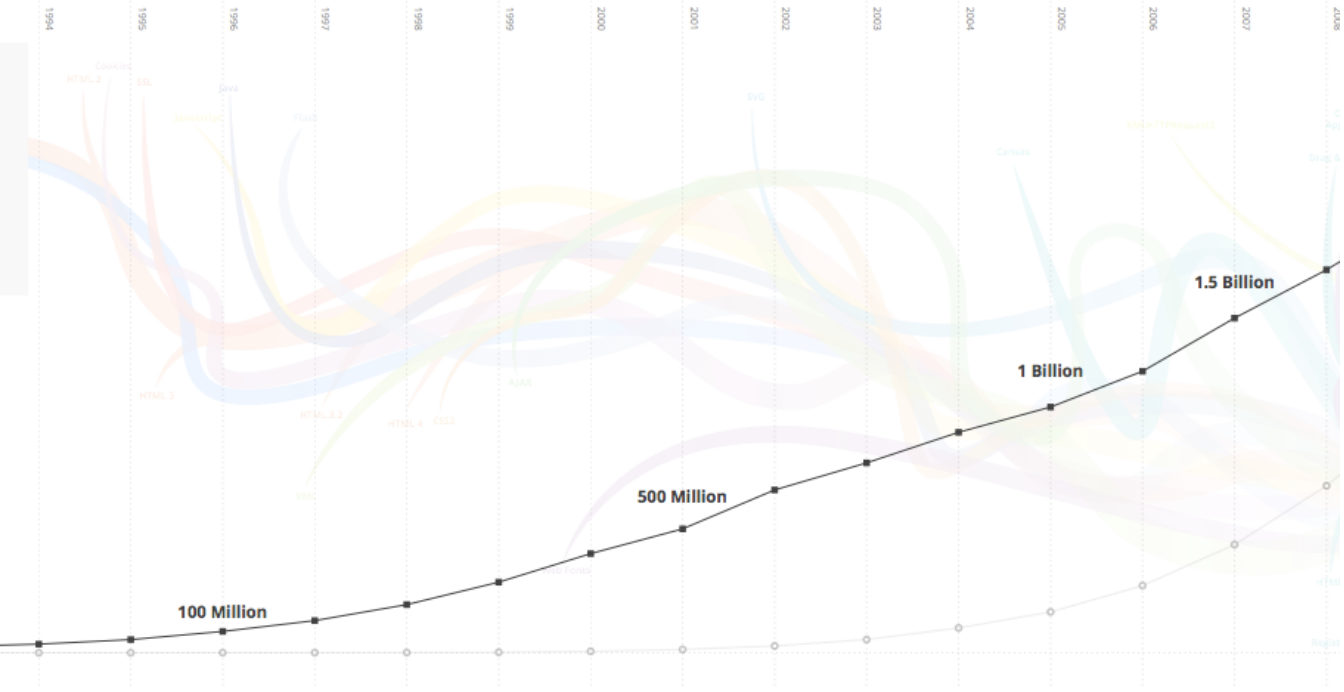
Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control



Wer hats erfunden?

- Das Internet existiert schon seit den späten 70er Jahren (TCP/IP Stack)
- Mail (SMTP) seit 1981
- Ist gewachsen, aber nicht gross

Entwicklung des Web



Entwicklung des Web

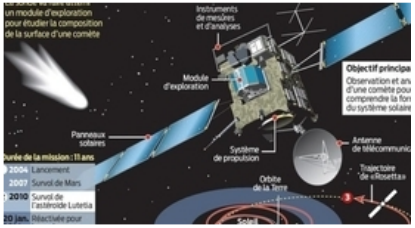
- Mit der Verbreitung des Web kam ein exponentielles Wachstum
- Das ist in erster Linie dem Link zu verdanken (und Hypertext/HTML)

Soweit so gut

a glace fond lentement» à la nférence sur la Syrie

VE-2 — Les discussions entre les deux délégations nnes ont repris mercredi à Genève. Le représentant spécial NU et de la Ligue arabe Lakhdar Brahimi a souligné que le reste «très profond».

r à 18h50 1 Commentaire



Les infographies du «Matin»

Telles que publiées dans votre quotidien. Découvrez notre galerie.

Les plus belles photos de l'année 2013

Découvrez les meilleures photos prises aux quatre coins de la planète.

Notre sélection d'images

奖电子表格及要求　关于互联网的行为规范　北京晨报新浪微博平台　北京晨报腾讯微博平台　北京晨报搜狐微博平台　北京晨报3G手

节目　PAD版

315

新增　ipho

整形美容医院　北京晨报开心社区入口　北京晨报网图片频道　北京晨报网易微博平台　北京晨报官方微信入口　非手术注射美容

Verstehen Sie Chinesisch?

奶牛放屁致甲烷超标引发爆炸 数头奶牛受伤　新疆塔城行署副专员操办婚礼动用公车9天

香港宣布对菲实施首阶段制裁

抢红包”游戏火如何提现引关...] [北京最低工资标准涨至1560...] [萧山一家三口相继染H7N9]

国内 | 国际 | 文化 | 体育 | 时尚 | 娱乐

要闻　更多 >

七常委出席春节团拜会

天价拖车费问责9人　违停公务员夫妻被严处

银行理财产品收益再上6% 普京：欧盟请别搅和

多名H7N9谣言散布者被拘 长寿乡巴马变灾难地

日本教科书“夺岛” 塞西晋升 获准竞选总统

医生收红包回扣可撤职 雍和宫大街初一有限行

紫竹桥至四季青桥将设潮汐车道

“广西贺州医生感染H7N9死亡”属谣言

春运前最高峰出行人次破亿 女童年前遭母遗弃

河北一乡政府打白条十年不还

合肥房叔女儿女婿贪污被公诉 打击窃听犯罪

检方对衡阳贿选嫌犯立案侦查 唆使父母自焚

北晨评论　调查　更

娶赦

在长沙天心区法院执行...[详细]

- 李娜获奖官员为啥凑热闹
- 人间至美乃羞涩
- 春节 环保是一种态度

96101 热线新闻　更

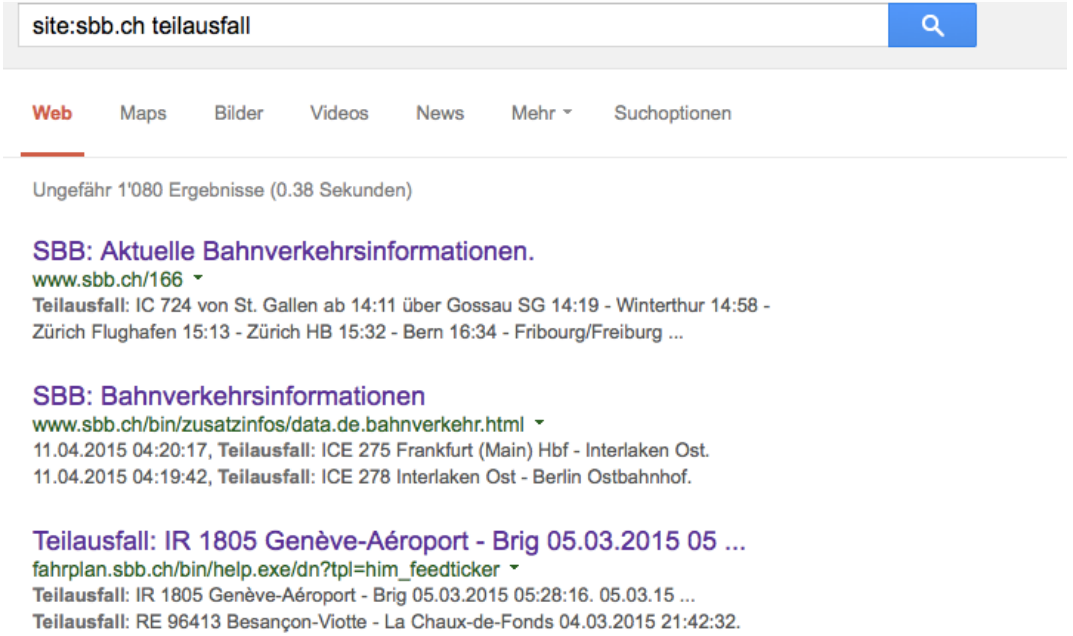
别报道　更多 >

Verstehen Sie Chinesisch?

- Die Semantik einer Seite ist für (sehende) Menschen verständlich,
- so lange mir die Schrift, Sprache und Thematik vertraut ist

Semantik für Maschinen schwer erfassbar

Zum Beispiel für Google



Zum Beispiel für Google

- Die Semantik ist für Maschinen nur schwer erfassbar
- Der Kontext fehlt oftmals völlig
- Beispiel: Jaguar das Tier, das Auto oder der MacOS 10.2 Release?

Babylonische Sprachverwirrung

Alles wird besser mit Datenbanken?

- Daten in Silos
- Teilweise auf dem Web verfügbar
- Zwischen den Datenbanken herrscht Sprachverwirrung
- Sowohl technisch wie inhaltlich (Semantik)
- Datensatzübergreifende Suche und Verlinkung ist nicht möglich

Zentrale Datenhaltung als Lösung?

Kaum realistisch:

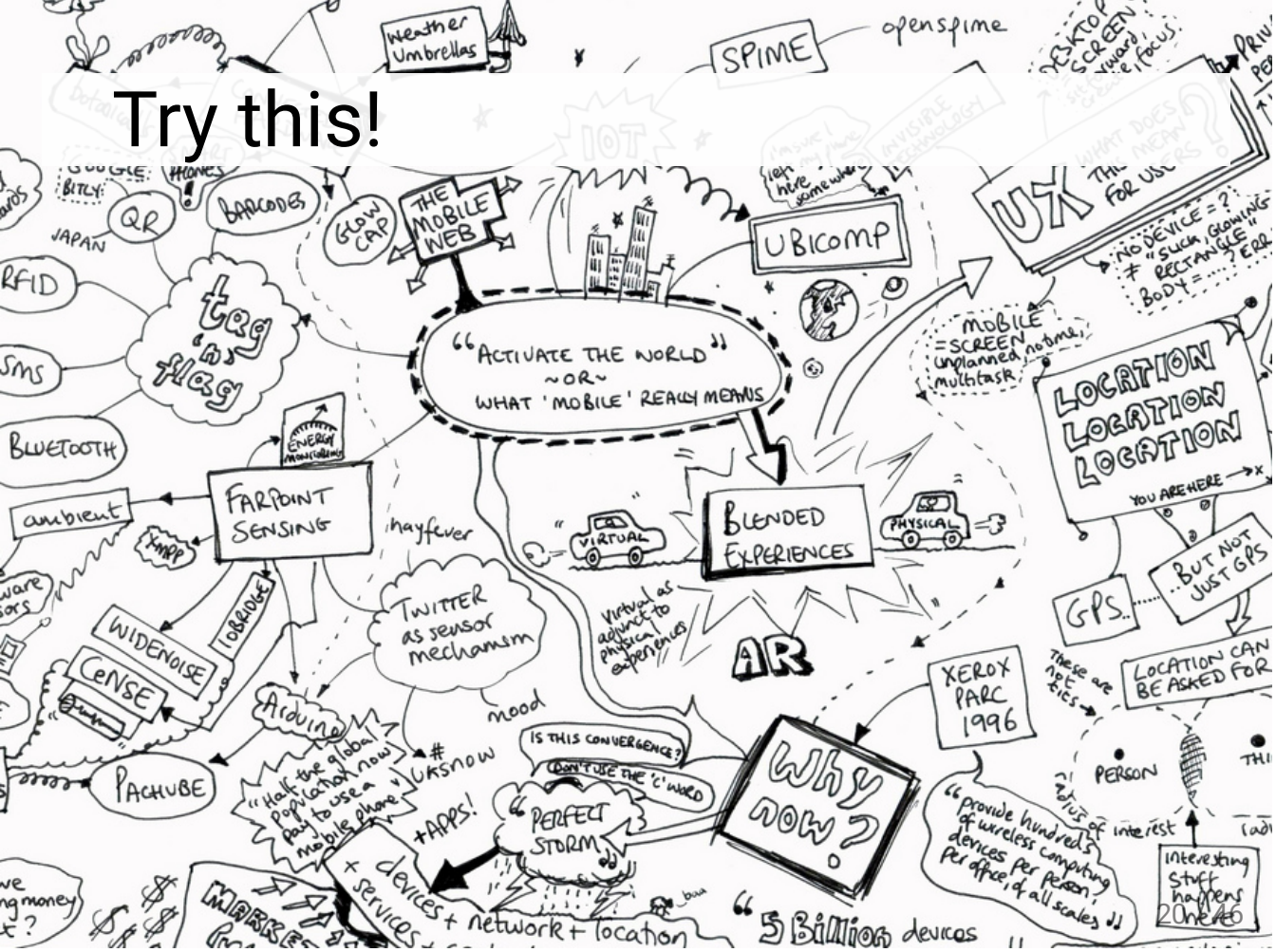
- Es gibt nicht eine Wahrheit
- Entspricht nicht föderalen und pluralistischen Strukturen
- Funktioniert selbst in zentralistischen Systemen eher selten
- Inhärente Hürden bei der Skalierbarkeit
- Technisch wie inhaltlich anspruchsvoll

Dezentraler Lösungsansatz

Was wären Anforderungen?

- Ein gemeinsames Austauschformat
- Semantik muss Teil der Daten sein
- Mehrsprachigkeit muss im Kern unterstützt sein
- Muss jegliche Beziehungen zwischen Daten ausdrücken können
- Beziehungen zwischen Datensätzen/Silos können ausgedrückt werden
- Dezentrale Datenhaltung/Pflege muss möglich sein
- Datensatzübergreifende Abfragen müssen möglich sein

Try this!



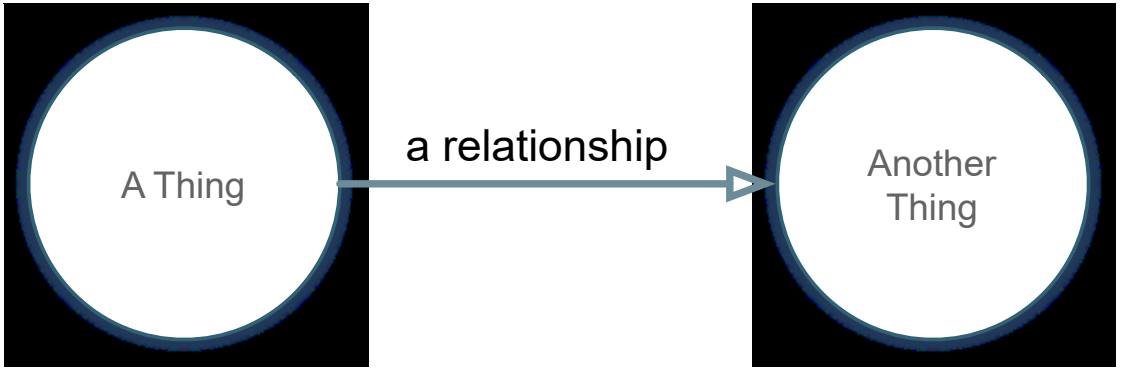
Weather Umbrellas
SPIME
openspime
IOT
BITLY
QR
JAPAN
BARCODES
GLOW CAP
THE MOBILE WEB
UBICOMP
UX
WHAT DOES THIS MEAN FOR USERS?
NO DEVICE = ? ≠ "SLICK, GLOWING RECTANGLE" BODY = ? ERR
RFID
tag 'n' flag
SMS
"ACTIVATE THE WORLD" ~OR~ WHAT 'MOBILE' REALLY MEANS
MOBILE = SCREEN unplanned, no time, multitask
LOCATION LOCATION LOCATION
YOU ARE HERE
BLUETOOTH
ENERGY MONITORING
FARPOINT SENSING
ambient
XMPP
hayfever
VIRTUAL
BLENDED EXPERIENCES
PHYSICAL
BUT NOT JUST GPS
GPS
WIDENOISE
LOBRIDGE
CeNSE
TWITTER as sensor mechanism
virtual as adjunct to physical experiences
AR
XEROX PARC 1996
These are not tics
LOCATION CAN BE ASKED FOR
Arduino
mood
IS THIS CONVERGENCE?
DON'T USE THE 'C' WORD
PACHUBE
"Half the global population now pay to use a mobile phone"
#UKSNOW
+APPS!
"PERFECT STORM"
Why now?
PERSON
radius of interest
"provide hundreds of wireless computing devices per person, per office, of all scales"
interesting stuff happens here
devices + network + location
+ services
"5 Billion devices

Linked Data als Lösungsansatz

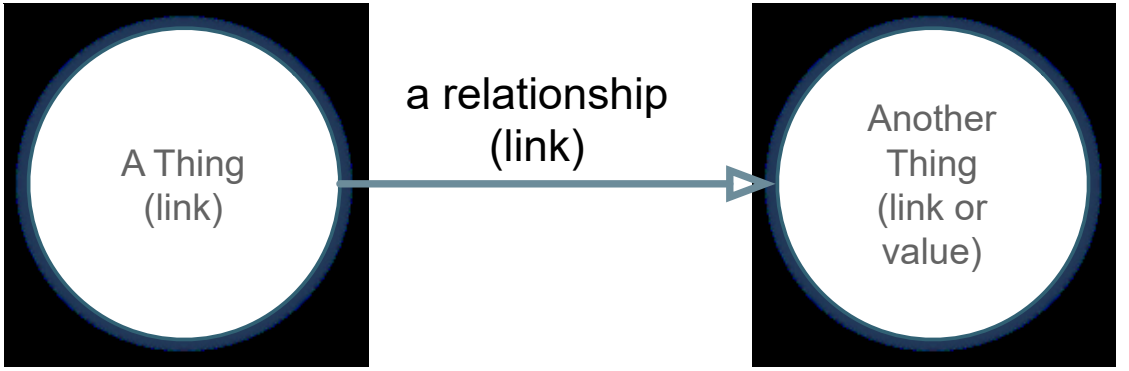
- RDF als allgemeines Datenmodell
- Well known Schemas & Ontologien als Lingua Franca
- Web (HTTP) als Transportmedium
- Links (URIs) als (dezentrale) Identifikatoren
- Mehrsprachigkeit im Datenmodell
- SPARQL als standardisierte Abfragesprache
- "Agiles" Datenmodell

Maschinenlesbarkeit

Datenmodell



RDF Datenmodell



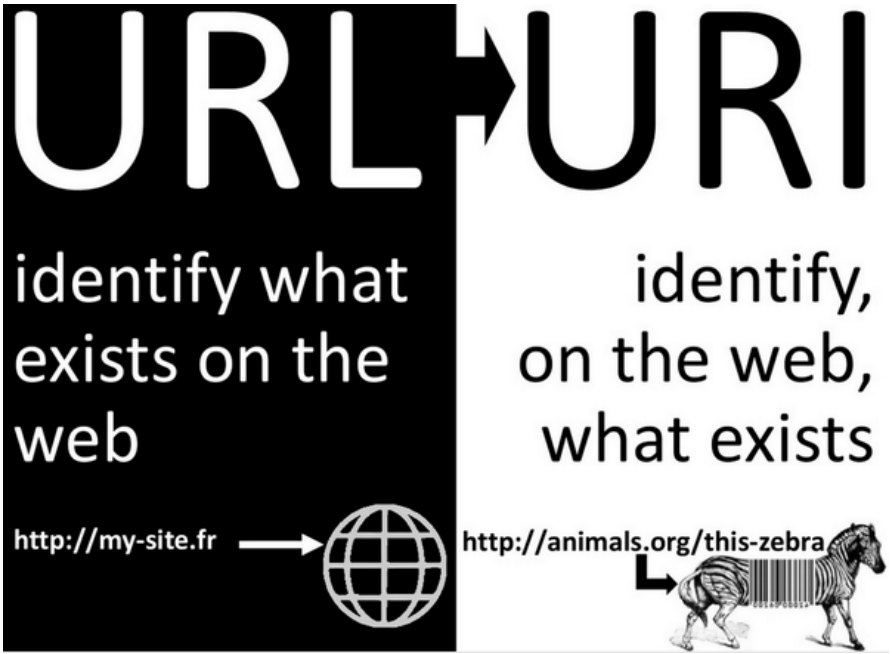
RDF Datenmodell

- Statt ein Dokument beschreibe ich eine einzelne Information

Beispiel:

```
<ktk> <vorname> "Adrian"
<ktk> <familienname> "Gschwend"
```

Der Unterschied zum bisherigen Web



"Now! … *That* should clear up
a few things around here!"

Technologiestack

- W3C Standard, wie das Web selber
- SPARQL als Abfragesprache
- Implementationen des Stacks in allen gängigen Programmiersprachen

SPARQL

- Hersteller von Datenbanken halten sich an den Standard
- Abfragesprache für Linked Data
- Standardisiert
- Skaliert (in memory)
- Unterstützt Federated Queries
- Kommerzielle Anbieter, stark wachsender Markt
- Open Source Alternativen

SPARQL Basics

- At home: SPARQL in 11 minutes
- SELECT für Auswahl, WHERE für Konditionen
- WHERE kombiniert typischerweise verschiedene Kombinationen miteinander
- Entspricht in SQL einem JOIN
- Aber ohne die Kopfschmerzen, die JOIN verursacht

Historisiertes Gemeindeverzeichnis

- Via Bundesamt für Statistik, Daten & Doku
- eCH-0071 Standard
- Kantone, Bezirke und Gemeinden
- Namensraum des Schemas: `PREFIX gont: <https://gont.ch/>`
- Nun kann per verkürzter URI darauf zugegriffen werden
- Klassen: `gont:Municipality`, `gont:District`, `gont:Canton`
- Details zur Dokumentation: https://github.com/zazuko/fso-lod/tree/master/doc/eCH0071

SPARQL by Example

- Siehe Github Dokumentation

Vorteile

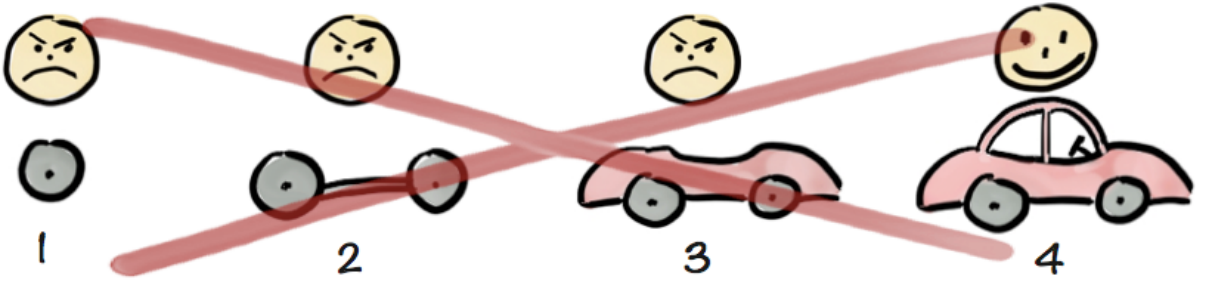
- Wiederverwendbarkeit ausserhalb des Anwendungsfalles
- Daten bleiben beim Datenhalter (wie Webseiten auch)
- Zugriffe werden sichtbar
- Komplexere Abfragen über SPARQL möglich
- ⇀ Antworten als CSV, XML, JSON, RDF
- die Daten sind das API
- ⇀ Antworten in Text, XML, JSON
- keine Probleme wie API versioning & maintenance

Zusammenfassung

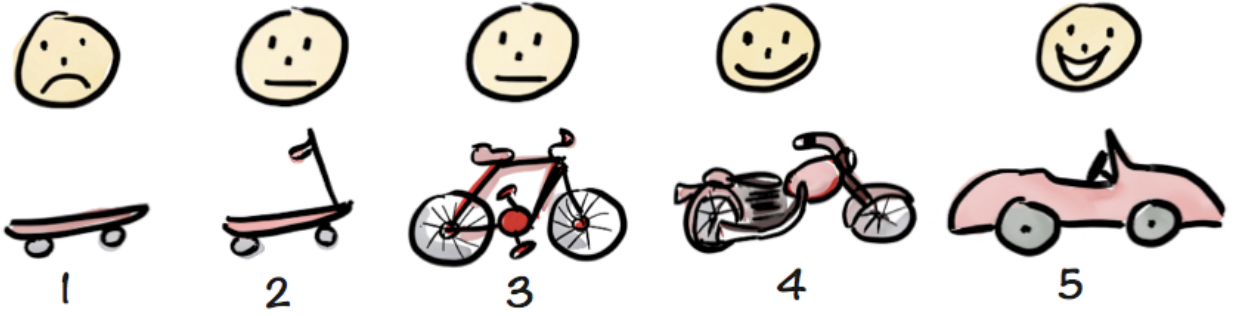
- Web of Data statt Web of Documents
- Gleicher Technologiestack
- Offene, klar definierte Standards
- Semantische Interoperabilität durch Vokabulare & Ontologien
- Dezentral, wie das Web
- Web Scale Database

Aufbauend auf einem agilen Datenmodell

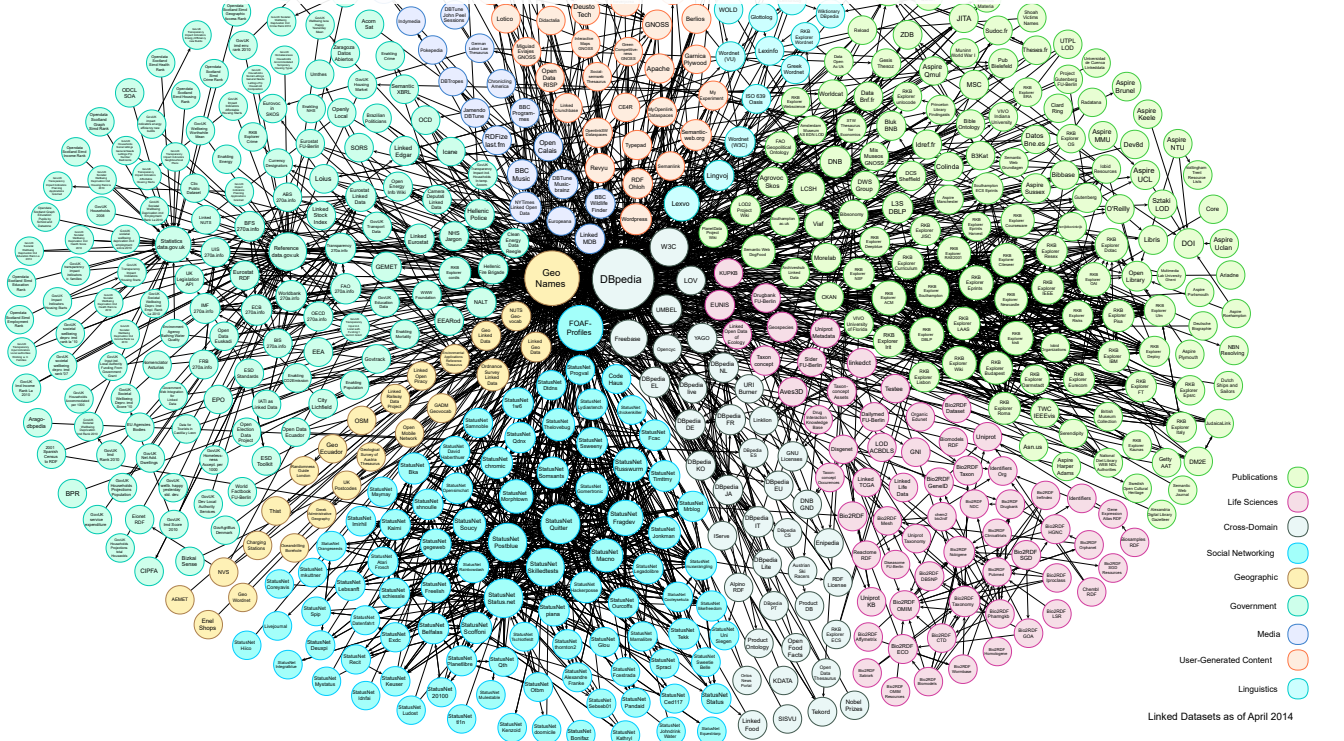
Not like this....



Like this!



Linked Open Data Cloud



Linked Open Data Cloud

- Siehe lod-cloud.net
- Beinhaltet die grössten öffentlichen Linked Open Data Repositories
- Linked Open Vocabularies LOV für gängige Vokabulare
- Prefix.cc für Prefixe & Shortcuts

Wikidata

- Wikipedia für Rohdaten
- Laaange Geschichte um/gegen RDF & Linked Data
- Heute mit SPARQL Endpunkt!
- Via korrektem Prefix `http://www.wikidata.org/entity/Qxyz` auch Linked Data
- Beispielabfrage Flughafencodes & Labels: https://t.co/gyWQ7MRzL6
- Neu auch mit Geospatial Queries

Google Suche

- SEO, Suchmaschinenoptimierung
- Heute via schema.org
- Komplett Linked Data
- JSON-LD oder RDFa Serialisierung

Ungefähr 114'000 Ergebnisse (0.59 Sekunden)

Vorstellungen für "The Angry Birds Movie"

Heute | Morgen | Fr., 20. Mai | Sa., 21. Mai

Alle Zeiten | Vormittags | Nachmittags | Abends | Nachts

KITAG CINEMAS Jura - Karte

Mit Untertiteln in Deutsch

Standard 17:30 20:30

Mit Untertiteln in Französisch

Standard 17:30 20:30

Standard 18:45 21:00

KITAG CINEMAS City - Karte

Mit Untertiteln in Deutsch

Standard 17:45 20:45

Mit Untertiteln in Französisch

Standard 17:45 20:45

CineBubenberg - Karte

Mit Untertiteln in Deutsch

Standard 18:30

Mit Untertiteln in Französisch

Standard 18:30

Alle Zeitangaben: MEZ

Mehr Vorstellungszeiten

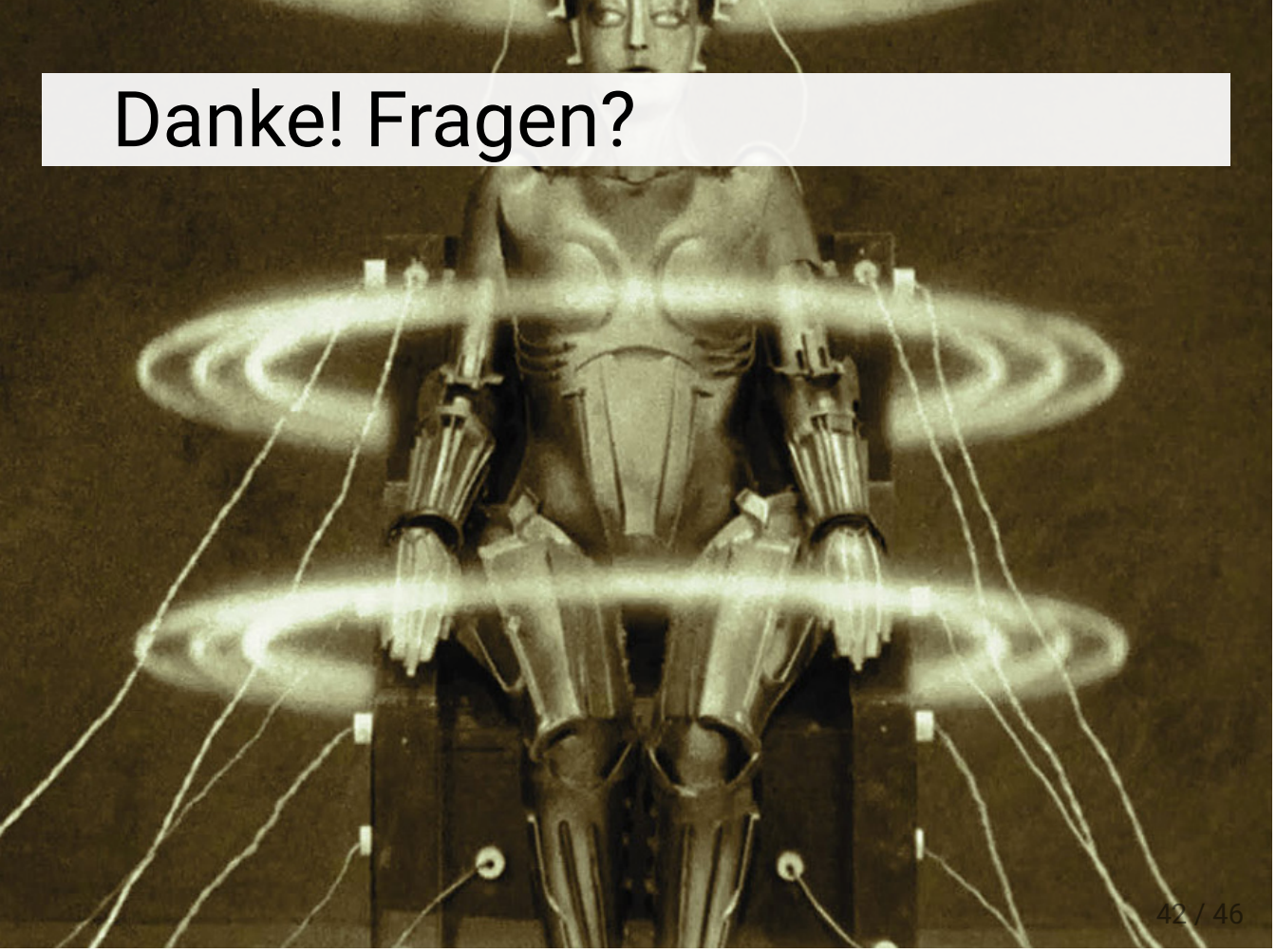
The Angr

Action/Trickfilm/Kom

Offshore Leaks als Linked Data

Gibt zwei verschiedene Datasets aktuell:

- https://titanpad.com/gdVduv14jD
- https://rawgit.com/Ontotext-AD/leaks/master/README.html

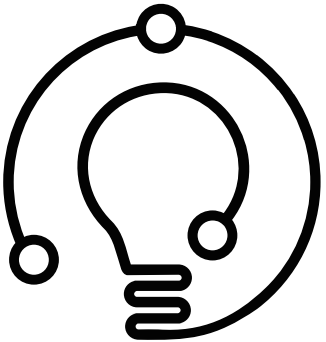


Danke! Fragen?

Kontakt

adrian.gschwend@zazuko.com

http://www.zazuko.com



zazuko

Referenzen: Standards

- RDF 1.1 Primer
- Turtle Serialisierung
- JSON for Linking Data (JSON-LD)
- SPARQL 1.1 Query Language

Bücher

- Learning SPARQL, Homepage, Safari Books Online
- Semantic Web for the Working Ontologist Homepage, Safari Books Online

Programmiersprachen

- JavaScript: RDF-Ext, Übersicht mit Alternativen
- JavaScript Visualisierung: Uduvudu, RDF2h
- Java: Apache Jena, Eclipse RDF4J (ehemals Sesame)
- C#.Net: dotNetRDF
- Ruby: Ruby RDF Project
- Python: RDFLib
- C: Redland RDF Libraries
- PHP: EasyRdf (Google for more)
- Perl: Perl and RDF & Attean