



openHPI: **Web-Technologien** Social, Semantic und Service Web

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

Willkommen zu Kurswoche 6

Zentrale Frage:

Das Web ist aus der modernen Gesellschaft nicht mehr wegzudenken. Welche Auswirkungen hat diese tiefgehende Durchdringung eigentlich und worauf müssen wir achten?

Beschäftigen uns in dieser Woche mit:

- Kurzem Überblick über Entwicklungsstadien des Webs
- **Social Web:** Soziale Netzwerk, Social Media und deren Verankerung in der Gesellschaft
- **Semantic Web:** Webseiten werden Maschinen-lesbar und -verstehbar
- **Service Web:** Technologische Grundlagen der digitalen Transformation der Gesellschaft und deren besondere Herausforderungen bei Datenschutz und Sicherheit

Das Web heute und morgen (1/3)

Seit seiner Einführung vor 25 Jahren unterlag das World Wide Web einem stetigen Entwicklungs- und Veränderungsprozess:

- **Web 1.0: Lesbar und statisch** (vor 2000)

- Nur einige (professionelle) Autoren konnten Inhalte veröffentlichen
- Web-Ressourcen waren ganz überwiegend statisch

- **Web 2.0: Kollaborativ und mitwirkend** (bis heute)

- Nutzer (auch Laien) können eigene Inhalte veröffentlichen dank Anwendungen im Bereich → **Social Web**
- Viele Webseiten werden dynamisch erzeugt und immer komplexer
- Neue Technologien und Frameworks ermöglichen es Entwicklern, immer komplexere Web-basierte Anwendungen zu bauen

Das Web heute und morgen (2/3)

Seit seiner Einführung vor 25 Jahren unterlag das World Wide Web einem stetigen Entwicklungs- und Veränderungsprozess:

- **Web 3.0: Semantisch und ausführbar** (noch im Entstehen)
 - Informationen werden maschinell lesbar und verständlich
 - Heutige Suchmaschinen nutzen Schlüsselwortbasierte Mechanismen
 - Such-Algorithmen vergleichen und zählen Zeichenketten, verstehen Bedeutung der Suchbegriffe und Dokumente nicht
 - **semantisches Web** bietet nicht länger nur Dokumente, sondern auch **Dinge** (Menschen, Orte, Ereignisse, ...), setzt sie in **Beziehung**
 - Web-Anwendungen können miteinander verwoben werden durch Web Services → **Service Web**

Das Web heute und morgen (3/3)

Seit seiner Einführung vor 25 Jahren unterlag das World Wide Web einem stetigen Entwicklungs- und Veränderungsprozess:

■ **Web 4.0: Mobil** (heute)

- parallele, alternative Version des existierenden Webs
- misst dem lokalen Kontext größere Bedeutung zu
- nutzt stark das → **Service Web** (Server-seitiger Teil von Apps)

■ **Und darüber hinaus?**

Wie könnte das **Web 5.0** aussehen?

→ Diskutieren Sie Ihre Ideen im Forum!

Social, Semantic und Service Web

Fokussieren uns auf drei Aspekte des Webs von heute und morgen:

Social Web

- Ursprung und Anwendungen des Social Web:
 - Blogs und Wikis
 - Social Networks und Social Media
- Meinungsbildung und Manipulation im Social Web
 - Twitter Bots
 - Fake News und Manipulation
 - Filterblasen und Echokammern
- Privatsphäre und Sicherheit im Social Web
 - „Privacy Paradoxon“
 - Gefahren durch Social Engineering

Social, **Semantic** und Service **Web**

Fokussieren uns auf drei Aspekte des Webs von heute und morgen:

Semantic Web

- Web bietet nicht mehr nur Sammlung von Dokumenten sondern von Entitäten (Personen, Orte, Ereignisse, ...) und setzt diese in Beziehung zueinander
- Maschinen „verstehen“ Bedeutung der Informationen im Web
- Anwendungsszenarien:
 - Semantik-unterstützte Suche
 - intelligente Agenten
- Technische Grundlagen des Semantic Web
- Linked Open Data als praktikabler Ansatz

Social, Semantic und **Service Web**

Fokussieren uns auf drei Aspekte des Webs von heute und morgen:

Service Web

- Web Services sind eine grundlegende Technologie für das **Internet der Dinge (IoT)**
- IoT Visionen: Smart Factory, Smart Home, Smart World
 - Industrie 4.0 – Ausprägung von IoT in Industrie und Wirtschaft
- Spektakuläre Sicherheitsvorfälle, z.B.
 - Chrysler Hack
 - Stuxnet
 - Identitätsdiebstahl → www.sec.hpi.de
- Sicherheit und Datenschutz als zentrale Herausforderung in einer hochvernetzten Welt



openHPI: **Web-Technologien** Social Web und Social Media

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

Entstehung des Social Web

Wahrnehmung und Nutzung des Internets änderte sich seit 2005:

- **Mitwirkungs-Prinzip** des Web 2.0
 - Nutzer sind Bearbeiter von Webseiten und nicht nur Leser von statischen Web Informationen
 - Nutzer teilen ihre Daten im Web, z.B. Lesezeichen, Fotos, berufliche Informationen, Videos, ...
 - Es entsteht ein (virtuelles) Gemeinschaftsgefühl
- Zunehmende Nutzung des Webs als Kommunikationsplattform, Grenzen zwischen Web Anwendungen und lokalen Anwendungen auf dem Rechner verwischen, z.B.
 - Nutzung von *Gmail* im Browser anstatt eines lokalen Mail Clients

Formen des Social Webs: Wikis

Wikis sind frühe Form von Anwendungen des Social Web

- „Wiki“ ist hawaiisch für kurz, schnell
- verwirklichen Idee einer kollaborativen Textverarbeitung und Kommentierung von Beiträge im WWW
- Erste Wiki-Installationen wurden bereits 1995 veröffentlicht (WikiWikiWeb)
- **Wikipedia** wurde 2001 gestartet und erhöhte Popularität und Verbreitung des Konzepts
 - Wikipedia ist unter den 10 populärsten Webseiten gelistet
- Wikis wurden auch bekannt durch eine Nutzung als Wissensdatenbanken in Unternehmen und Organisationen
 - Mitarbeiter können ihr Wissen teilen und speichern



Formen des Social Webs: Blogs

Web-Blogs, oder kurz **Blogs**, verbreiten sich Ende der 1990er Jahre und bieten weitere wichtige Form des Social Webs

- **Blogs** sind öffentliche persönliche Tagebücher bzw. Journale im Web
- Ein oder mehrere Autoren – die Web-Logger, kurz **Blogger**, veröffentlichen Beiträge (**Posts**) über ihr Leben oder ein spezielles Thema
- Blog-Leser können sich aktiv beteiligen durch Feedback in Form von Kommentaren, Bewertungen, ...
- Nutzer können Blogs abonnieren durch **Feeds** (z.B. RSS, Atom)
- Technisch sind Blogs realisiert durch
 - Blog Hosting Dienste, z.B. **Blogger**, **tumblr**, ...
 - Hosting auf eigenem Server, z.B. **WordPress**



tumblr.



Formen des Social Webs: Microblogging

Microblogging wurde für gelegentliche kurze persönliche oder unternehmerische Status-Updates eingeführt

- Posts haben meist Länge von nur 200 Zeichen oder weniger
- Posts können öffentlich oder privat erstellt werden mit einem/r Browser, App, E-Mail oder Instant Messaging
- Für Microblogs wurde Konzept der **Asynchronen Follower** eingeführt: User erhalten Status-Updates von Microloggern denen sie „folgen“, aber nicht umgekehrt
- Derzeit ist **Twitter** bekanntester Microblogging-Anbieter
 - wurde 2006 gestartet und ist heute wichtiger Mediankanal für öffentliche Personen, politische Aktivisten, ...
- Um aufkommende Themen zu identifizieren, werden **#Hashtags** im Social Web benutzt



Formen des Social Webs: Soziale Netzwerke

Soziale Netzwerkdienste, oder kurz **Soziale Netzwerke** entstanden zusammen mit dem Microblogging

- Soziale Netzwerke kombinieren Blogging / Microblogging mit Liste weiterer verbreiteter Funktionen (Features):
 - persönliches Profil
 - Freundesliste (im Gegensatz zum Follower-Konzept legen Soziale Netzwerke bidirektionale Nutzerbeziehungen an)
 - Aktivitäten-Verlauf (kombiniert Status-Updates der Freunde eines Nutzers und erlaubt Kommentare)
 - Privater Nachrichtenaustausch mit anderen Nutzern und Nutzergruppen
 - Ausdruck persönlicher Interessen, z.B. durch „likes“
 - Soziale Anwendungen (z.B. Spiele) → Soziale Netzwerke werden Plattform für eigene Anwendungen



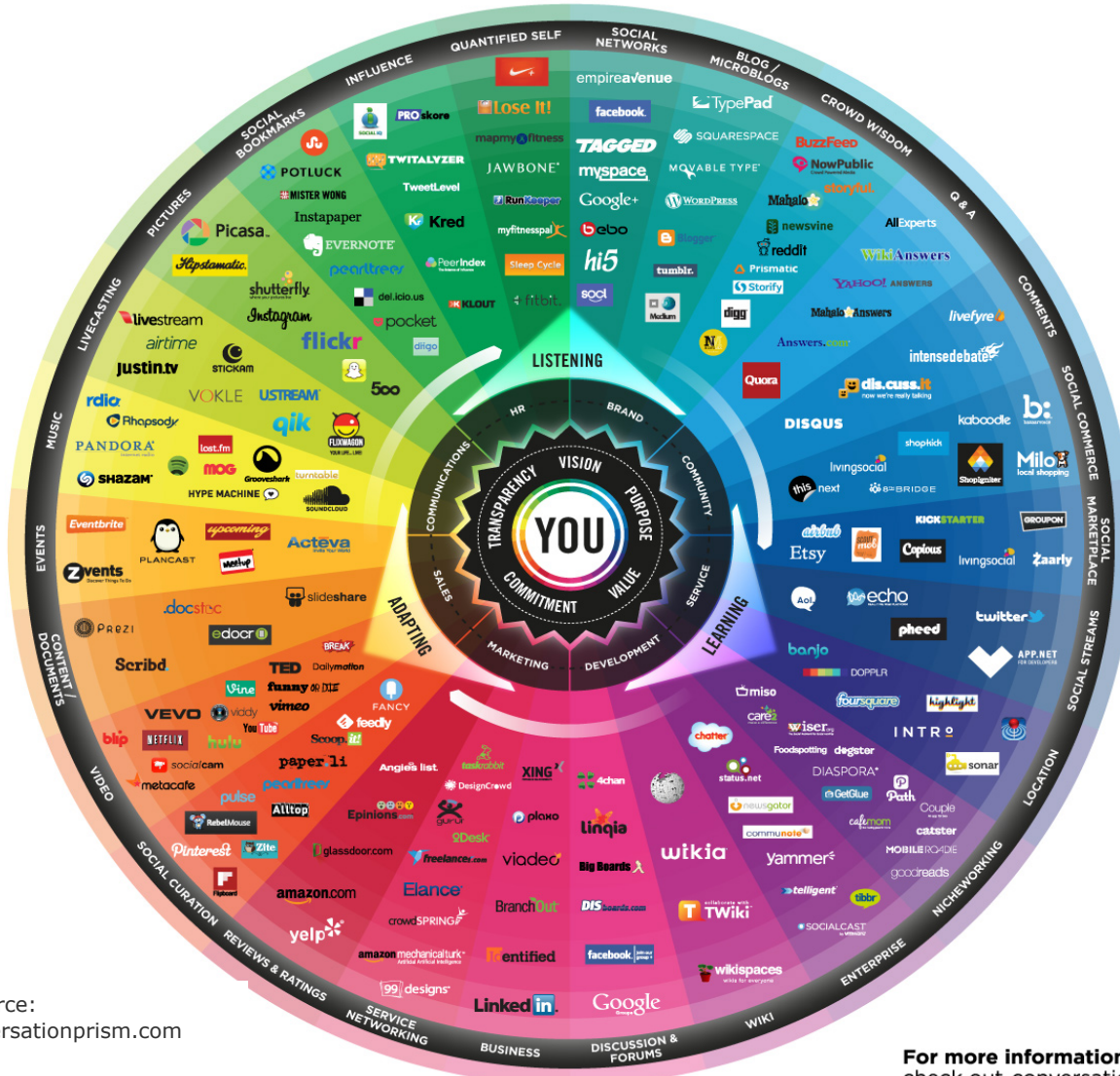
Formen des Social Webs: Soziale Medien

Durch stark fallende Preise für Speicher und hohe Bandbreiten entstanden seit 2005 zahlreiche Portale für Soziale Medien

- Soziale Medien bieten Nutzern Möglichkeit, Medien-Ressourcen wie Fotos, Videos oder Audios mit Freunden oder der Öffentlichkeit zu teilen
 - Autoren können Feedback für ihre Medien erhalten
 - Autoren können ihre Medien speichern, um sie z.B. in ihre persönliche Webseite einzubetten, ohne sich Gedanken um Speicherplatz zu machen
- Bekannte Plattformen für Soziale Medien sind heute:
 - Fotos – **Instagram, flickr, ...**
 - Video – **YouTube, Vimeo, ...**
 - Audio – **Soundcloud, ...**



Social Web – Unendliche Vielfalt und Variationen



Source:
<http://www.conversationprism.com>

For more information
 check out conversationprism.com

Kritik an Sozialen Medien

Privatsphäre (Privacy)

- Anbieter Sozialer Netzwerke fordern oft weitgehende Nutzungsrechte für von Nutzern bereitgestellte Inhalte
- Persönliche Informationen können genutzt werden für personalisierte Werbung und Verkauf an Werbetreibende zur Zielgruppenbestimmung

Vertrauenswürdigkeit

- Jeder kann wahrheitsgetreue oder falsche Inhalte erstellen, z.B. „Edit Wars“ auf Wikipedia oder die Diskussion um „**Fake News**“
- Unechte Accounts mit falschem Namen können erstellt werden

Verlässlichkeit

- Dienste können von Staaten geblockt werden, um den Zugriff auf Informationen zu beschränken

...



openHPI: **Web-Technologien** Meinungsbildung und Manipulation im Social Web

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

Bedeutung Sozialer Netzwerke

Soziale Netzwerke sind sehr populär und werden von immer mehr Menschen genutzt, z.B.:

- **Facebook:** 1,79 Milliarden aktive Nutzer pro Monat (Sep. 2016)
- **Instagram:** 600 Millionen Nutzer (Dez. 2016)
- **Twitter:** 313 Millionen aktive Nutzer pro Monat (Juni 2016)
- Für **Verbreitung von Nachrichten und Informationen** ist die Bedeutung sozialer Netzwerke enorm gestiegen
- **Quelle** und **Korrektheit** von Informationen lassen sich für Nutzer nur **schwer überprüfen** → Stichwort ***Fake-News***
- Jeder Nutzer, ob Mensch oder Maschine, kann Inhalte erstellen und verbreiten
- Potenzial der Manipulation der öffentlichen Meinung ist durch soziale Netzwerke dramatisch gestiegen

Twitter-Bots (1/2)

Ein **Bot**, kurz für Roboter, ist Programm, das weitgehend automatisiert eine Aufgabe abarbeitet

- Twitter-Bots täuschen gezielt **menschliches Verhalten** vor mithilfe von Machine Learning Technologien
- Erste Generation wurde zum Aufbau der Reichweite des eigenen Profils eingesetzt, heute geht es vermehrt um **Beeinflussung**
- Bots folgen dabei vorher festgelegten Interessen und twittern nur gelegentlich, um nicht automatisch als Spam erkannt zu werden
- Durch Tweets und Retweets (teilen von fremden Tweets) wird öffentliche Aufmerksamkeit für bestimmte Inhalte verstärkt
- Es existieren aber auch „gute“ Bots, die nützliche Automatisierungen durchführen, z.B.
 - Chat-Bots im Facebook Messenger

Twitter-Bots (2/2)

Aktuelle Beispiele der Meinungsbeeinflussung durch Bots:

■ **Brexit**

- Untersuchung von 1,5 Millionen Tweets von mehr als 300.000 Twitter-Accounts zum Thema Brexit
 - Bedeutender Anteil der Tweets wurde von Bots verfasst
 - Großteil der Bots sprach sich für Brexit aus

■ **US-Präsidentenwahl**

- Mehr als ein Drittel der Pro-Trump Tweets und fast ein Fünftel der Pro-Clinton Tweets zwischen erstem und zweitem Fernsehduell wurden von Bots abgesetzt
- Insgesamt mehr als 1 Million Tweets von Bots

➔ Wieviel Beeinflussung Bots tatsächlich erreichen, lässt sich schwer messen, aber Bots spielen starke Rolle vor allem in politischen Diskussionen

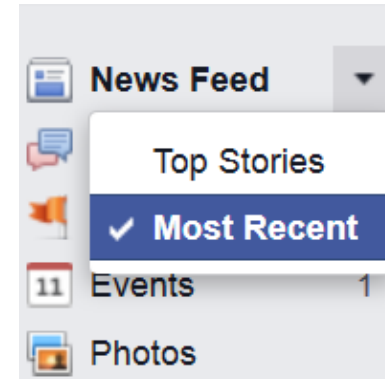
Filterblasen (1/2)

- Soziale Netzwerke versuchen mithilfe von Algorithmen vorausszusagen, welche Informationen einen Nutzer besonders interessieren
- Diese Inhalte werden dann prominent platziert, als uninteressant eingestufte Inhalte werden teilweise sogar ausgeblendet mit dem Ziel, das Nutzer länger auf Plattform verweilt und sich wohlfühlt
- Nutzern werden so allerdings auch Informationen vorenthalten, die vermeintlich nicht seinen Interessen und Standpunkten entsprechen
- Verstärkend wirkt hierbei noch das ***Echokammern***-Phänomen:
 - Nutzer vernetzen sich im virtuellen Raum verstärkt mit Gleichgesinnten
 - Dies kann zu einer fatalen Verengung der Weltsicht führen

Filterblasen (2/2)

■ Beispiele:

- Personalisierte Suche von Google
- Personalisierter Aktivitäts-Feed von Facebook und Instagram



- Diese Informationsblasen können in verschiedenen Themenbereichen Auswirkungen auf den öffentlichen Diskurs haben
- Aber auch hier gilt: Tatsächliche Beeinflussung ist bisher nicht einwandfrei nachgewiesen ...

Fake-News (1/2)

- Ob Twitter-Bots oder echte Menschen, jeder kann online Beiträge erstellen und verbreiten
- Zurzeit ist Diskussion um Fake-News in aller Munde, also absichtlich über das Internet verbreitete Falschmeldungen zur Beeinflussung der öffentlichen Meinung zu bestimmten Themen
- Im Europäischen Auswärtigen Dienst (EAD) in Brüssel arbeiten inzwischen 11 Beamte ganztags daran, Fake-News zu enttarnen und Gegendarstellungen zu veröffentlichen
- Dazu durchforstet ein Netzwerk aus rund 400 Journalisten, Universitätsangestellten, Beamten, NGO-Mitarbeitern und Einzelpersonen in 30 Ländern das Internet nach Fake-News

Fake-News (2/2)

- Nachdem Politik Druck auf Facebook ausgeübt hat, will Facebook neue Netzwerkfunktionen einführen, um gegen Fake-News vorzugehen
 - Nutzer sollen Möglichkeit erhalten, einen Beitrag als potenzielle Falschmeldung zu melden
 - Externes Recherchezentrum soll die gemeldeten Inhalte prüfen und gegebenenfalls als zweifelhaft markieren
 - Am Ende soll neben Beitrag mit falschen Informationen ein Warnhinweis erscheinen sowie Link zu einem Artikel mit den tatsächlichen Fakten

Kann man Fake-News erkennen? (1/2)

Wie kann man selbst Fake-News identifizieren?

- Seriosität der Quelle:
 - Sind Impressum und Kontaktmöglichkeiten vorhanden?
 - Wie lange existiert Seite schon?
 - Was verbreitet Quelle ansonsten für Nachrichten?
- Handelt es sich um Satire?
- In welchem Kontext steht die Seite?
 - Veröffentlicht Seite ansonsten viele Quatschmeldungen und Unsinn?
- Was steht im Artikel und was nur in Vorschau?
 - Auf Facebook können Titel eines verlinkten Artikels einfach geändert werden



Kann man Fake-News erkennen? (2/2)

- Wo kommen Informationen ursprünglich her? Seriöse Journalisten nennen und verlinken Informationen eines Artikels, Meldungen von Nachrichtenagenturen enden mit Kürzeln wie „dpa“, „Reuters“ oder „AFP“
- Wurde Quelle richtig wiedergegeben? Bei Fake-News werden verlinkte Quellen oft veraltet oder falsch wiedergegeben
- Zeigt Foto/Video wirklich beschriebene Situation, oder wurde einfach ein altes genommen und als Augenzeugen- oder Skandalclip inszeniert? (Lässt sich leider nur schwer nachvollziehen für Laien)

Macht der Algorithmen und Big Data (1/3)

- Unternehmen setzen zunehmend auf das Sammeln, Verknüpfen und Auswerten (persönlicher) Daten
- Über Smartphone oder Browser können permanent Daten über Nutzungsverhalten gesammelt und auch an Dritte weitergegeben werden
- Auf Basis dieser Daten wird personalisierte Werbung geschaltet
- Analyse dieser „Big Data“ bringt Unternehmen starken Wettbewerbsvorteil:
 - Kunden können besser verstanden und bedient werden
 - interne Prozesse können effizienter gestaltet werden
- Diese Praxis bietet allerdings auch viel Potenzial für Missbrauch!

Macht der Algorithmen und Big Data (2/3)

Beispiele für Nutzung von Big Data:

- US-Supermarktkette Target: Analyse des Einkaufsverhaltens kann schwangere Frauen identifizieren und deren Geburtstermine hochrechnen
- Facebook-Likes: Persönliche Eigenschaften wie Alter, Geschlecht, sexuelle Orientierung, ethnische Zugehörigkeit, politische Einstellung, Beziehungsstatus oder Alkoholkonsum lassen sich mit hoher Treffsicherheit erschließen
- ZestFinance: Berechnung der Kreditwürdigkeit von Kunden
- Predictive Policing: Steuerung des Einsatzes von Polizeikräften nach Wahrscheinlichkeit, wo und wann zukünftig Straftaten stattfinden, auf Basis der Daten von Fallakten
- ...

Macht der Algorithmen und Big Data (3/3)

- Wie kann man sich vor der ungewollter Datensammelei schützen?
 - AdBlocker/ScriptBlocker für Browser installieren
 - Berechtigungen in Apps kontrollieren und ggf. entziehen
- Vollständige Abschottung jedoch nicht möglich
- Gesellschaftlicher und politischer Diskurs ist nötig, was Algorithmen dürfen und was nicht
- Algorithmen der Unternehmen müssen transparenter werden



openHPI: **Web-Technologien** Privatsphäre und Sicherheit im Social Web

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

Das „Privacy Paradoxon“ (1/2)

- Soziale Netzwerke ermutigen Nutzer, **private Informationen preiszugeben**
 - Profilinformationen: Voller Name, Wohnort, Telefonnummer, Geburtsdatum, Arbeitgeber, Schulabschlüsse, Foto, ...
 - Interessen: Facebook „Likes“
 - Freunde und Familienmitglieder
 - Persönliche Meinungen („Timeline“)
- **Privacy Paradoxon**
 - Nutzer sozialer Netzwerke geben **bereitwillig** ihre private Informationen preis („**Selbstoffenbarung**“)
 - Dennoch haben diese Nutzer **gleiches (hohes!) Bedürfnis nach Privatsphäre**, wie diejenigen, die soziale Netzwerke ablehnen

Das „Privacy Paradoxon“ (2/2)

Gründe für Privacy Paradoxon bei Nutzung sozialer Netzwerke

- Gefühlt viele Vorteile, Nachteile nicht offensichtlich
 - Nutzer sozialer Netzwerke machen überwiegend Erfahrung, dass Preisgabe privater Informationen fast immer ohne (direkte) negative Konsequenzen bleibt
 - Veröffentlichte Details aus Privatleben führen oft zu positivem Feedback („Likes“), Zuspruch oder Anteilnahme aus der Community
- Mangelndes Problembewusstsein
 - Nutzer wissen oft nicht, was mit ihren Daten geschieht
 - Bewusstsein, dass mit Daten für „kostenlose“ Leistung bezahlt wird, ist oft nicht vorhanden



Verletzung der Privatsphäre durch Dritte

Soziale Netzwerke erlauben oft Einschränkung der Sichtbarkeit preisgegebener Informationen auf bestimmte Empfängergruppen

- Aber: Durch Interaktion berechtigter Empfänger mit geteilten Inhalten können auch weitere Personen Zugriff erlangen
 - Beispiel:
 - Nutzer X teilt Information mit seinen Freunden
 - Nutzer Y aus Freundesliste kommentiert Information
 - Nutzer Z ist mit Y bekannt aber nicht mit X
 - Nutzer Z sieht: Y hat Information von X kommentiert
 - Z erlangt so Zugriff auf Information auch von X ($X \rightarrow Y \rightarrow Z$)
- **Vorsicht:** häufig nicht einfach zu durchschauen, wer auf welche preisgegeben Information Zugriff erhält

Mögliche Nachteile durch Selbstoffenbarung

Preisgabe privater Informationen kann **handfeste Nachteile** haben

- Zugang „unerwünschter“ Personen zu privaten Informationen
 - 60% der US-Arbeitgeber recherchieren Bewerber in sozialen Netzwerken
 - 49% dieser Arbeitgeber fanden Informationen, die dazu führten, Bewerber nicht einzustellen

(Quelle, www.careerbuilder.com, Apr. 2016)
- Informationsressource für Angriffe durch → **Social Engineering**
 - Erraten von Passwörtern (→ **Passwortsicherheit**)
 - Diebstahl digitaler Identitäten
- Überwachung durch Betreiber des sozialen Netzwerks
 - Betreiber kann Internet-Bewegungsprofil des Nutzers erstellen
 - eventuelle Weitergabe des Bewegungsprofils an Dritte

Social Engineering

Grundidee: Angreifer kontaktiert Opfer, gibt sich als Techniker oder Systemadministrator aus und versucht, Zugangsdaten (Passwörter) zu erlangen

Beispiel: Angreifer erlangten Zugang auf privaten Email-Account von ehemaligem CIA-Direktor John Brennan (Identitätsdiebstahl)

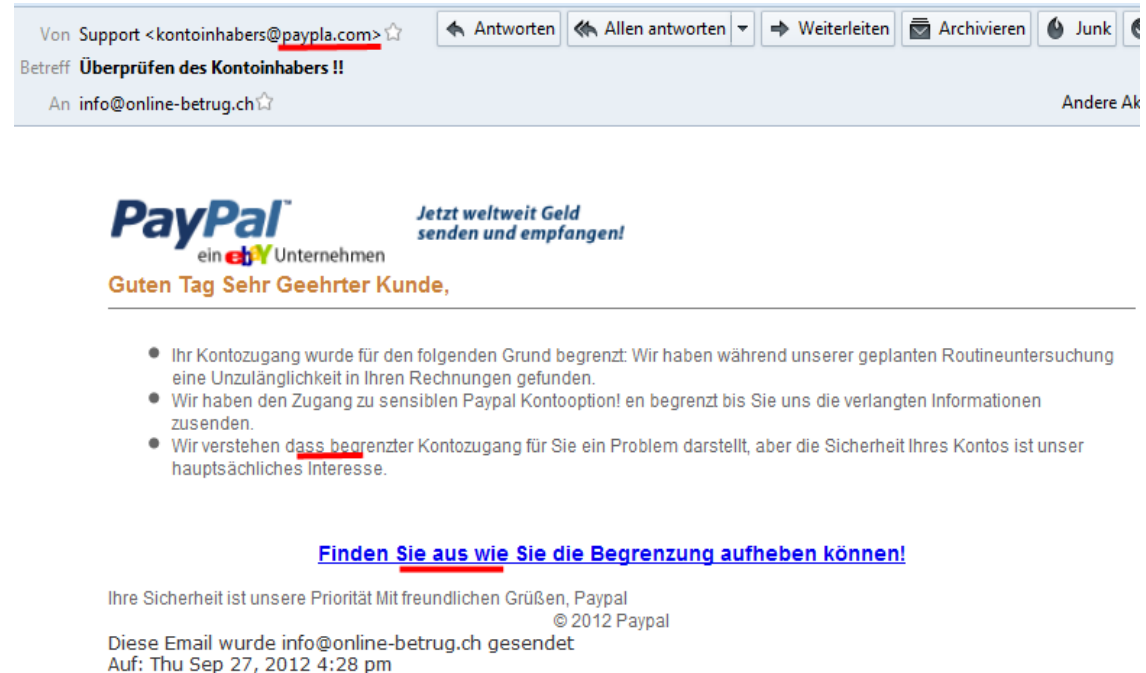
- Hacker kennt Handynummer des Opfers, stellt fest, zu welchem Provider (Verizon) dieser gehört (durch Reverse Lookup)
- Hacker gibt sich bei Verizon-Support erfolgreich als Verizon-Techniker aus, der Kundendaten benötigt – erhält z.B. letzte vier Stellen des Bankkontos
- Hacker gibt bei AOL-Support an, Passwort vergessen zu haben und kann mit Verizon-Daten die Sicherheitsfragen beantworten
- Passwort wird zurückgesetzt
- Angreifer hat 3 Tage lang Zugriff auf Email-Konto

Phishing = Password Fishing

- Technik zur betrügerischen Erlangung von sensiblen Informationen
- Typisches Vorgehen: Versand betrügerischer Emails
 - Versand mit vertrauenswürdig wirkender Absenderadresse, z.B. Bank, Unternehmen, Behörde, ...
 - Absenderadresse wird ausgewählt nach Information, die erschwindelt werden soll
 - Emailadressen können leicht gefälscht oder vorgetäuscht werden („Email Spoofing“)
 - Mit dem dadurch erlangten Vertrauen wird erwünschte Nutzerreaktion hervorgerufen
 - Herausgabe von Daten, z.B. Benutzernamen und Passwörter
 - Installation von Schadsoftware

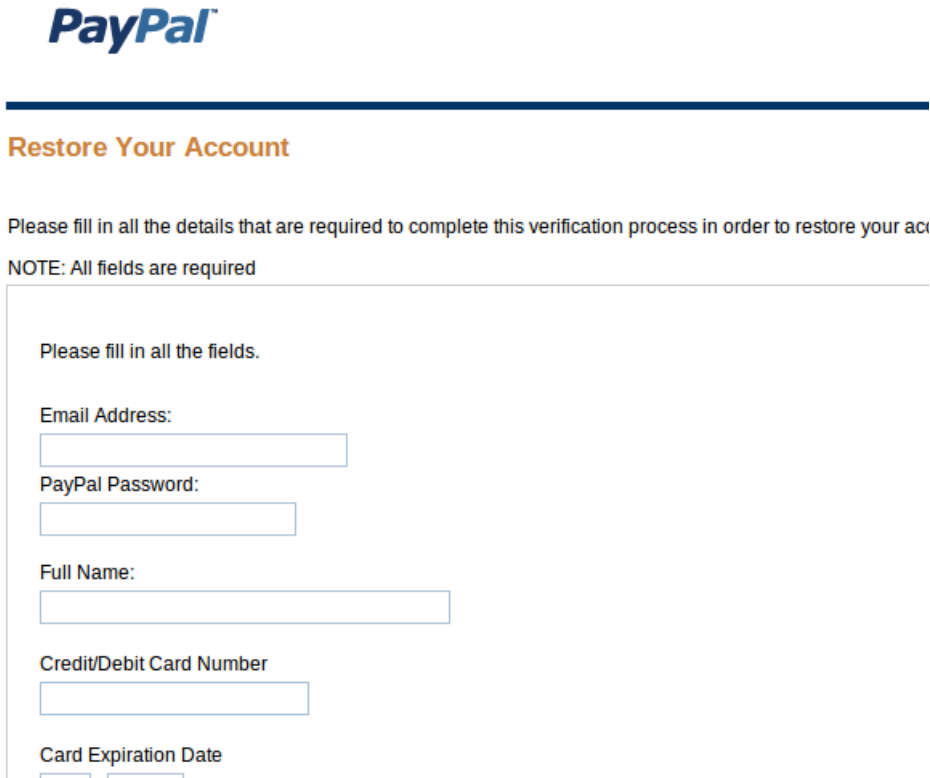
Beispiel:

- Versand einer Email mit vorgetäushtem Absender
- Nachricht stammt angeblich von Bankinstitut (Paypal) und hat entsprechend ähnliche Email-Adresse (paypla.com)
- Ziel: Empfänger soll veranlasst werden, seine Kontodetails preiszugeben, um Konto zurückzusetzen
- Phishing oft erkennbar an vielen sprachlichen Fehlern oder fehlender persönlicher Ansprache („Guten Tag Sehr Geehrter Kunde“)



- Email leitet weiter auf Webseite mit Eingabeaufforderung für Kontodaten
- Opfer wird aufgefordert, sämtliche Details über sein Konto preiszugeben
- Angreifer gelangt in Besitz der Daten

Gute Nachricht: Diese Art „herkömmlicher“ Phishing-Angriffe wird heute von vielen Internetnutzern als solcher erkannt und läuft dadurch ins Leere



The screenshot shows a PayPal login page titled "Restore Your Account". It contains a form with the following fields: "Email Address:", "PayPal Password:", "Full Name:", "Credit/Debit Card Number", and "Card Expiration Date". The form is styled with a light blue border and a white background. The PayPal logo is visible at the top left of the form area.

Quelle: <http://labs.m86security.com/>

Schlechte Nachricht: Personalisierte Phishing-Angriffe, sogenanntes **Spear Phishing** ist auf dem Vormarsch

- Angreifer sammelt detaillierte Informationen über sein Opfer
- Hilfreich für vorgetäuschte Vertrauenswürdigkeit, z.B. Agieren unter Identität einer Person, die Opfer kennt und der es vertraut
- Opfer denkt, dass Email von vorgetäuschem Absender kommt, weil Nachricht Informationen enthält, die eigentlich nur dieser Absender kennen kann
- Erfolgsaussichten bei dieser Masche sind sehr hoch, aber auch für Angreifer ist Aufwand deutlich höher als beim herkömmlichen Phishing

Grundidee von Social Engineering-Attacken basiert auf Missbrauch von **Vertrauen**

- Opfer vertraut Angreifer, da dieser **vertrauliche Informationen** hat, die eigentlich nur Mitglieder eines **vertrauenswürdigen Kreises** haben
- Masche zieht genau so, wie beim Spear Phishing und Telefon-Betrug im Fall „John Brennan“

Wie erlangt Angreifer vertrauliche Informationen?

- Klassische Technik: **Dumpster Crawling** – Angreifer durchsucht Mülltonne des Opfers nach Briefen, Abrechnungen, Kontoauszügen, ...
- Informationstechnischer Ansatz: Hacken des Computers eines Opfers
- **Viel einfacher:** Einsammeln der freiwillig preisgegebenen Informationen aus den sozialen Netzwerken

Passwortsicherheit (1/2)

Auch **Angriffe auf Passwörter** können durch persönliche Informationen aus Social-Media-Profilen erleichtert werden

- Verbreitet: **Password Guessing** – Erraten von Passwörtern
 - Angreifer probiert wahrscheinliche Passwörter zu einem Account durch
 - Angriffstools arbeiten mit Wortlisten und Wörterbüchern
- **Problem:** Nutzer wählen gerne (immer gleiche/ähnliche) Passwörter, die sie sich leicht merken können, z.B.
 - Begriffe oder Daten aus persönlichem Umfeld
 - Geburtsdaten, Namen, Begriffe im Zusammenhang mit Interessen (Lieblings-Romanfigur, -Band, -Filmtitel, ...)
 - Persönlich Informationen der Opfer helfen Angreifer, kann z.B. **Wortlisten personalisieren** → Trefferwahrscheinlichkeit steigt

Passwortsicherheit (2/2)

Beispiel:

- Nutzer gibt auf Facebook an, Fan von „Star Trek“ zu sein
- Angreifer kann spezielle (fertig erhältliche) Wortliste mit allen Begriffen aus Star Trek-Universum verwenden
- Ähnliche Wortlisten auch verfügbar z.B. für „Herr der Ringe“, „Game of Thrones“, Monty Python-Filme und -Serien, ...

Informationen aus Social Networks nicht nur hilfreich beim Knacken (Raten) von Passwörtern, sondern insbesondere auch für sogenannte **Sicherheitsfragen** zum Zurücksetzen von Passwörtern, z.B.

- Mädchenname der Mutter
- Erstes Automodell
- Straße der ersten Wohnung
- Teile von Kreditkarten- oder Kontonummern (siehe John Brennan)



openHPI: **Web-Technologien** Einführung in das Semantic Web

Prof. Dr. Christoph Meinel

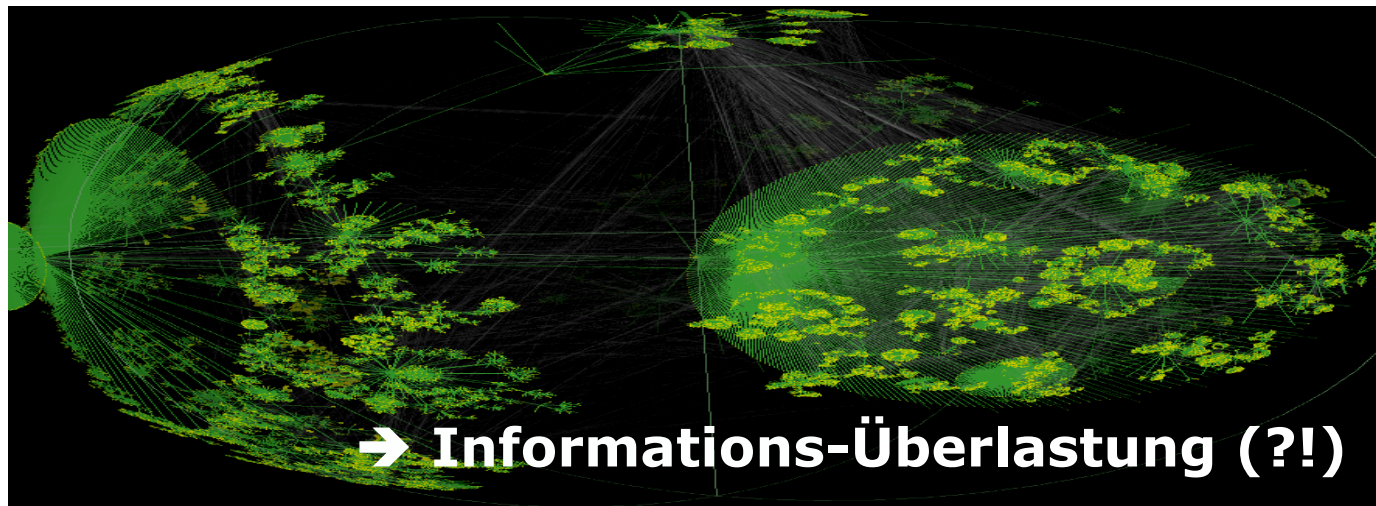
Hasso-Plattner-Institut

Universität Potsdam

Beschränkungen des World Wide Web (1/7)

Erinnerung: Das WWW erlaubt Zugriff auf enorme Menge an Informationen

- Es gibt Milliarden von WWW-Dokumenten – allerdings ist nur ein kleiner Teil von Suchmaschinen indexiert
- Außerdem können Milliarden von Dokumenten im *Deep Web* gefunden werden
- Die Menge der Dokumente im WWW verdoppelt sich alle 6 Monate



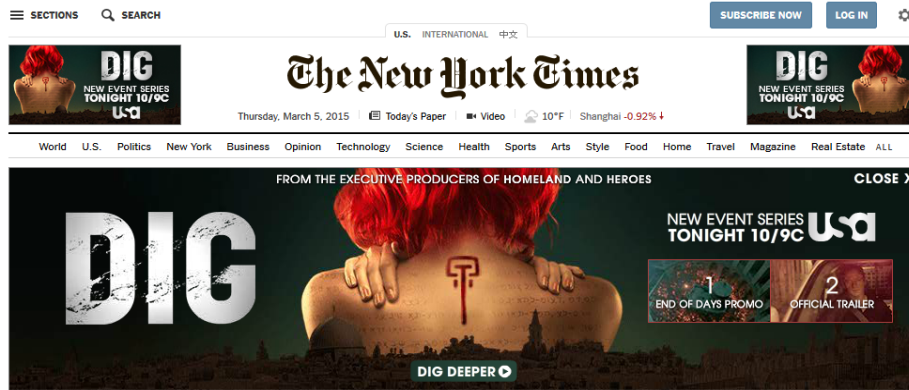
Beschränkungen des World Wide Web (2/7)

The screenshot shows the New York Times homepage with a prominent 'DIG' event banner. The banner includes the text 'FROM THE EXECUTIVE PRODUCERS OF HOMELAND AND HEROES' and 'NEW EVENT SERIES TONIGHT 10/9C'. Below the banner, the page is organized into several columns of article teasers. The first column features a large article about Ferguson, Missouri, with a sub-headline 'Policing, Revenue and Discrimination Went Hand in Hand'. The second column has a section titled 'A Mark of Status for Clinton's Inner Circle' with a photo of Hillary Clinton. The third column includes a section about 'Jawbone's Discovery' with a photo of a fossil. The fourth column has a section titled 'A Chilling Portrait of Ferguson' and a 'Watching' section with various news snippets. The fifth column has a section about 'Chinese Premier Wen Jiabao's Remarks'.

Relevante Informationen?

- Was ist **wichtig**, was nicht?
- Was ist **Information**, was ist Werbung?
- Was ist **Bedeutung** der Information?
- Wie **glaubwürdig** ist die Information?
- Was gehört **zusammen**?
- Welche Teile sind eigentlich **überflüssig**?

Beschränkungen des World Wide Web (3/7)



U.S. Calls on Ferguson to Overhaul Its Justice System

By MATT APUZZO and JOHN ELSON
A report examining the deadly shooting of an unarmed black teenager by a white police officer in Ferguson, Mo., said that the city had engaged in multiple constitutional violations.

177 Comments

Policing, Revenue and Discrimination Went Hand in Hand

By CAMPBELL ROBERTSON
9:50 PM ET
Ferguson, a report finds, acts less like a municipality and more like a business, extracting money from poor blacks that it uses to sustain the city's budget.

RELATED COVERAGE
U.S. Clears Ferguson Officer in Inquiry 8:55 PM ET
Report: What Happened in Ferguson?

Chinese Premier Weighs Options to



Hillary Rodham Clinton on Tuesday in Washington. An aide set up clintonemail.com and registered it to her New York home.
Jabin Botsford/The New York Times

A Mark of Status for Clinton's Inner Circle

By AMY CHOIRICE and STEVE EDER 12:43 AM ET
A domain name set up for Hillary Rodham Clinton shielded her correspondence and was used for coveted accounts that conferred prestige and closeness to the secretary of state.

RELATED COVERAGE

- Clinton Asks State Dept. to Release Emails 50 minutes ago
- Democrats Rally to Clinton's Defense After Email Disclosure
- Associated Press Threatens Legal Action Over FOIA Requests

Jawbone's Discovery Fills Barren Evolutionary Period

By JOHN NOBLE WILFORD
The 2.8-million-year-old mandible of a Homo habilis predates by at least 400,000 years any previously known Homo fossils.



A Chilling Portrait of Ferguson

By THE EDITORIAL BOARD
The Missouri town must move far more aggressively to correct its dangerous problem of entrenched racism.

MORE IN OPINION

- Editorial: The Fate of American Health Care
- Op-Ed: Shortchanging New Jersey by Billions
- Taking Note: High Rent, Low Wages

The Opinion Pages

CHARLES M. BLOW

The Feds vs. Ferguson

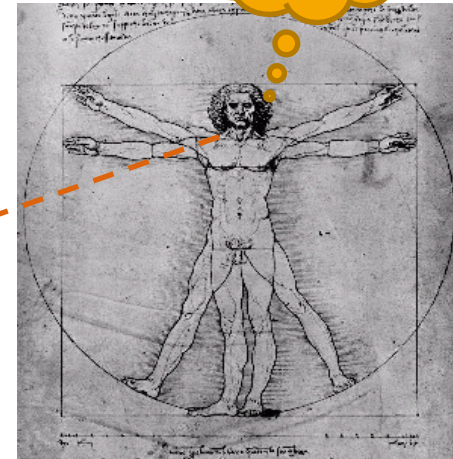
The Ferguson Police Department and the municipal courts treated citizens like a revenue stream, violating their constitutional rights in the process.

- Bruni: Hillary's Messy Habits
- Friedman: What Bibi Didn't Say
- Edsall: Establishment Populism Rising

Watching

- 4m The family of a Canadian pastor says it has received confirmation that he has been **detained in North Korea** after traveling there on a humanitarian trip.
CNN >
- 2h A Singapore court **sentenced two Germans to canings and jail time** after they admitted spray-painting a train with graffiti.
Reuters >
- 3h The number of people worth \$30 million or more **rose to nearly 173,000 worldwide** last year, a study found.
The Guardian >
- 3h 

Ist diese große Menge an Informationen eigentlich nützlich?

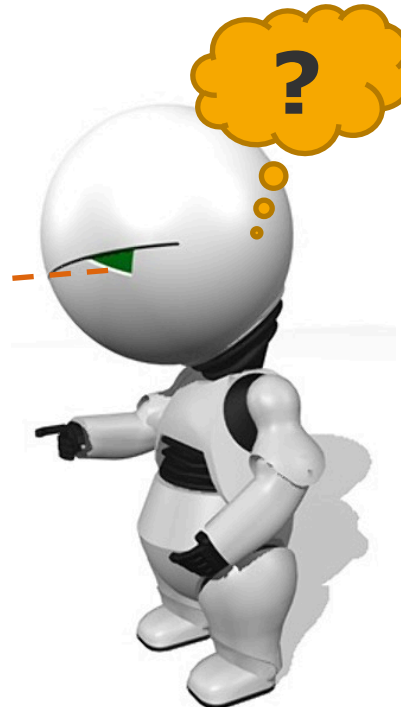


Menschliche Nutzer haben **kontextuelles Wissen / Allgemeinwissen** und können daher die (meisten) Informationen korrekt interpretieren

Beschränkungen des World Wide Web (4/7)

The screenshot shows the New York Times homepage. The main article is 'U.S. Calls on Ferguson to Overhaul Its Justice System'. Below it, a sub-article 'Policing, Revenue and Discrimination Went Hand in Hand' is highlighted. To the right, the 'Opinion Pages' section features 'A Chilling Portrait of Ferguson' and 'The Feds vs. Ferguson'. At the bottom, the 'Watching' section lists a video about a Canadian pastor's family. A red arrow points from the sub-article title to the robot's thought bubble.

Ein **Roboter** (Programm) kann nicht zwischen wichtigen und unbedeutenden Informationen unterscheiden, da es **Bedeutung nicht versteht**



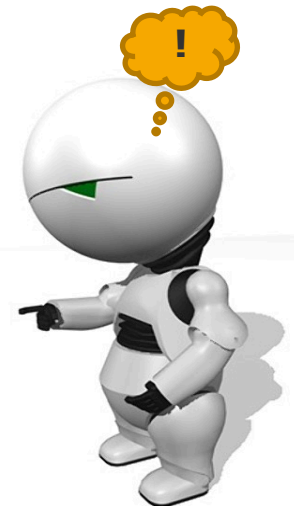
Beschränkungen des World Wide Web (5/7)

The screenshot shows the New York Times homepage. Red arrows point from the text 'Roboter braucht Informationen über Bedeutung eines Dokuments und seiner verschiedenen Teile' to the 'DIG' event series banner and the 'U.S. Calls on Ferguson to Overhaul Its Justice System' article. Another red arrow points from the text 'Werbung' to the 'The Feds vs. Ferguson' article. A third red arrow points from the text 'Allerdings bietet HTML nur sehr begrenzte Möglichkeiten, Metadaten zu beschreiben' to the 'Watching' section.

Roboter braucht Informationen über Bedeutung eines Dokuments und seiner verschiedenen Teile

→ Metadaten

Werbung



Allerdings bietet HTML nur sehr begrenzte Möglichkeiten, Metadaten zu beschreiben

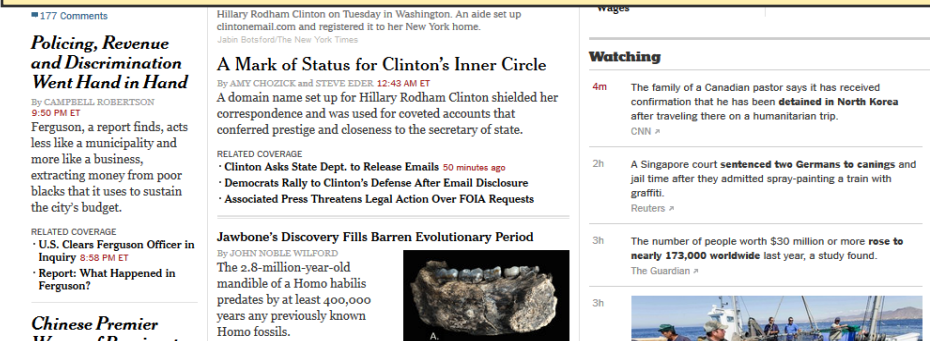
Beschränkungen des World Wide Web (5/7)



Roboter braucht Informationen

```
<meta charset="UTF-8" />
<meta http-equiv="cache-control" content="no-cache" />
<meta http-equiv="content-language" content="en-US" />
<meta http-equiv="content-type" content="text/html; />
<meta name="author" content="Jane Doe" />
<meta name="description" content="...summary of web page..." />
<meta name="robots" content="all" />
<meta name="keywords" content="meta tags meaning" />
```

Alle Metadaten müssen vom Autor der Webseite erstellt werden



Allerdings bietet HTML nur sehr begrenzte Möglichkeiten Metadaten zu beschreiben

Beschränkungen des World Wide Web (7/7)

- WWW wurde ursprünglich **für Menschen entwickelt**
- Technologie hinter dem WWW ist Markup-Sprache HTML
- HTML und CSS beschreiben
 - wie Informationen **strukturiert** und **dargestellt** sind
 - wie Informationen **verlinkt** werden können
 - aber **nicht Bedeutung** der Informationen



Semantik?

Semantik

Semantik ist Teilgebiet der **Linguistik**, dass sich mit

- **Sinn** und
- **Bedeutung**

von **Sprachen** und **linguistischen Symbolen** beschäftigt

Semantik versucht Frage zu beantworten, wie

- Sinn und Bedeutung von komplexen Begriffen abgeleitet werden können von Sinn und Bedeutung einfacher Begriffe

Semantik baut auf **Syntax** auf

- **Kommunikation** benötigt ein allgemeines Verständnis von **Syntax** und **Semantik** der ausgetauschten Sprachsymbole und Zeichenketten

→ Gegenseitiges **Verständnis** ist nur möglich, wenn Sender und Empfänger die **identische Semantik** vereinbart haben

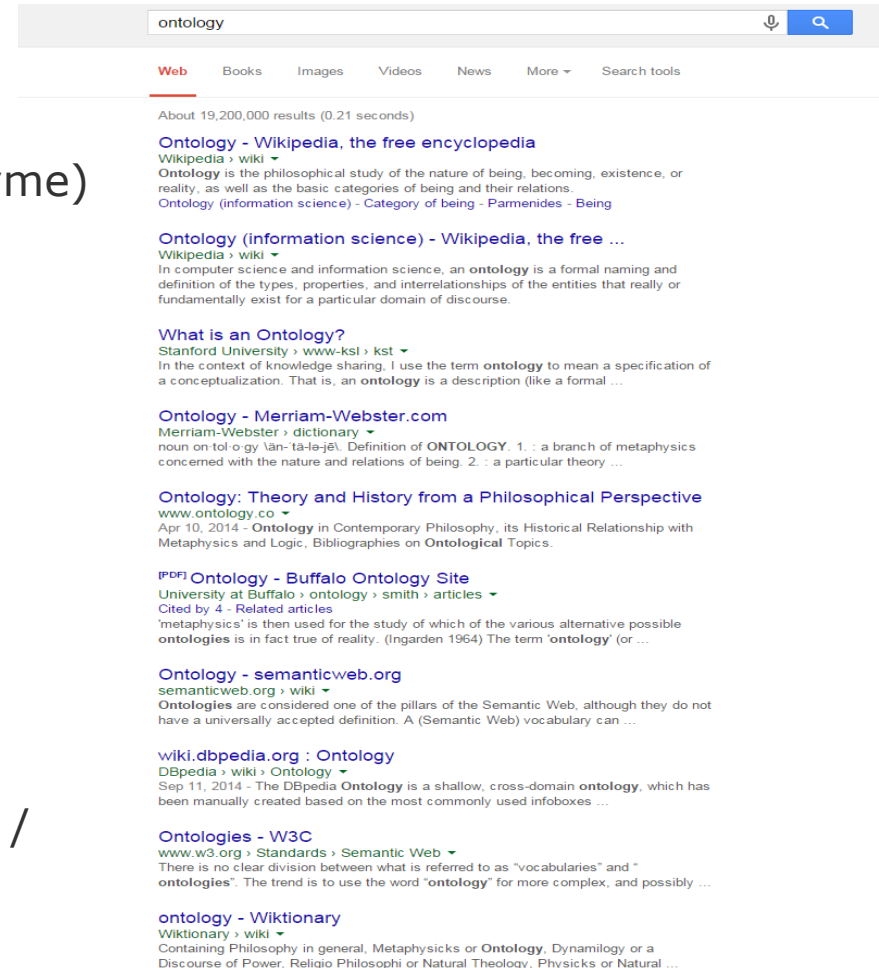
Problembereich 1: Suche nach Informationen

Schlüsselwort-basierte Suche

- Ruft viele **irrelevante** Ergebnisse ab
 - verschiedene Bedeutungen (Homonyme)
 - anderer Kontext
- Kann nicht **alle relevanten** Ergebnisse finden
 - Synonyme
 - Mangel an präzisiertem Kontext

Anforderung:

- Einführung einer formalen Terminologie /
Ontologie



Problembereich 1: Suche nach Informationen

Schlüsselwort-basierte Suche

- Ruft viele **irrelevante** Ergebnisse ab
 - Verschiedene Bedeutungen (Homonyme)
 - Anderer Kontext
- Kann nicht die gewünschte Information finden
 - Synonym
 - Mangelt an Kontext

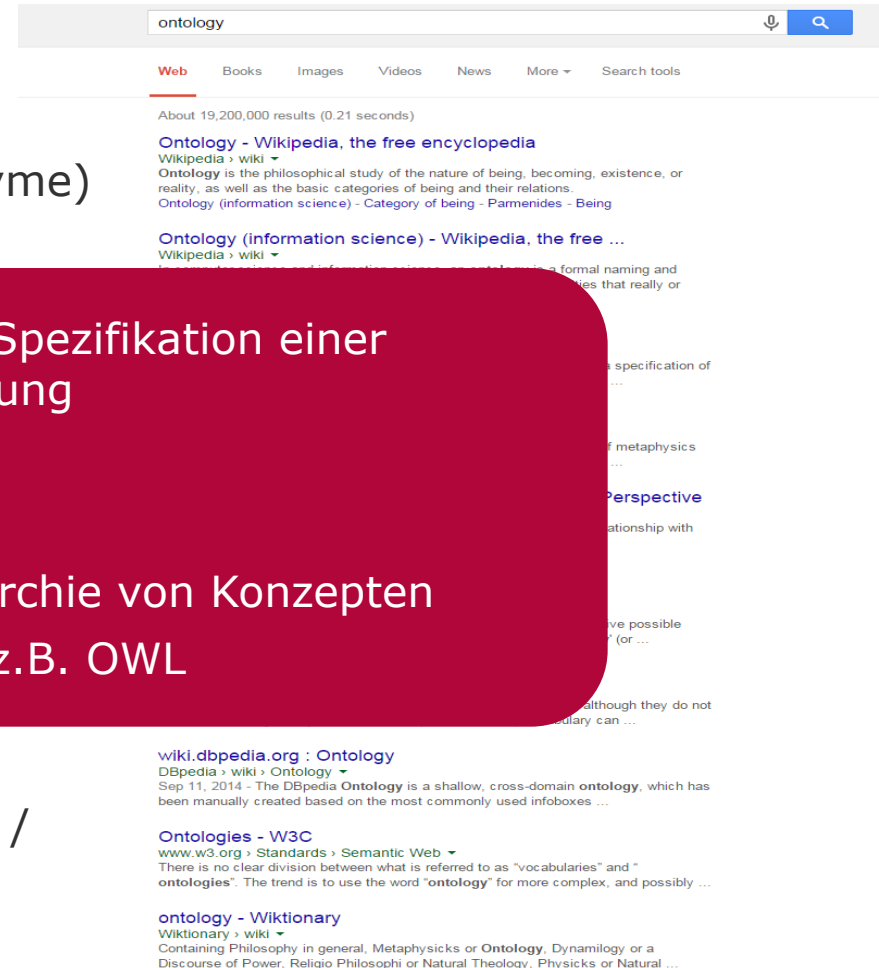
Ontologie: Explizite formale Spezifikation einer gemeinsamen Konzeptualisierung

Eine Ontologie besteht aus:

- **Taxonomie**, d.h. eine Hierarchie von Konzepten
- sprachlichen Beschreibung, z.B. OWL

Anforderung:

- Einführung einer formalen Terminologie / **Ontologie**

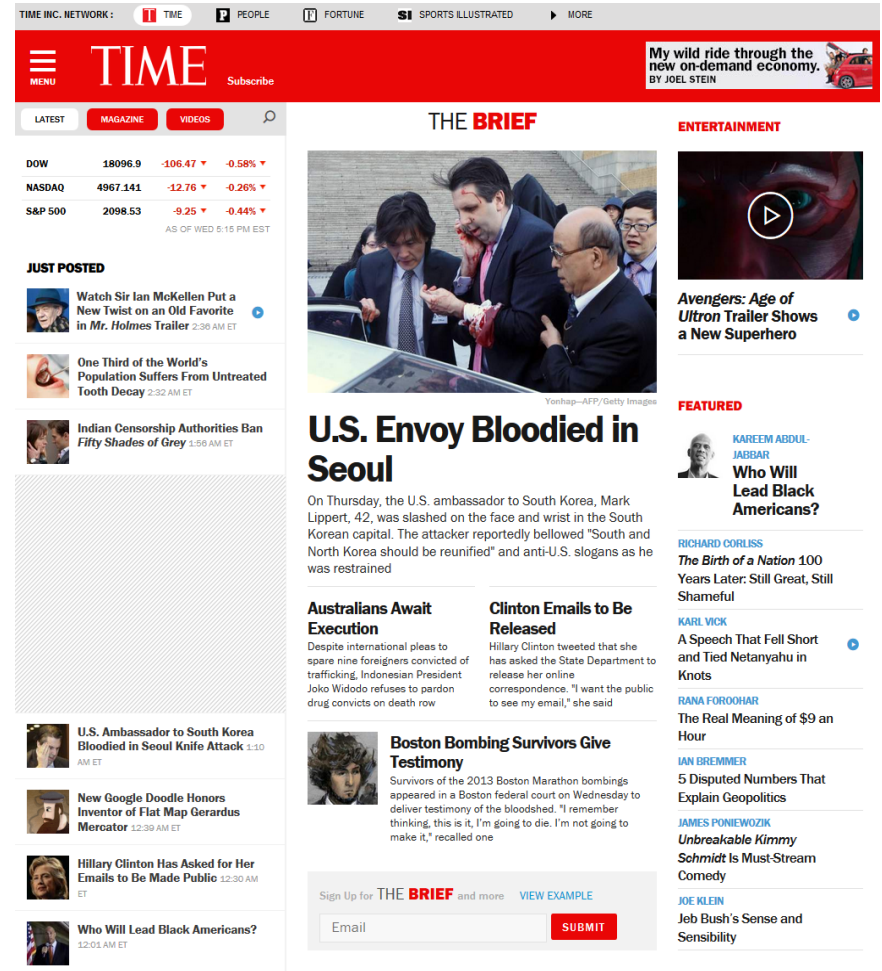


Problembereich 2: Informationsextraktion

- Nur Menschen können Informationen korrekt extrahieren
 - Software (Roboter) fehlt
 - kontextuelles Wissen
 - Allgemeinwissen
- wird benötigt, um Informationen aus ihrer Text- oder Bild-basierten Darstellung zu extrahieren

Anforderung:

→ **Aggregation und Integration von Informationen** aus verschiedenen Quellen

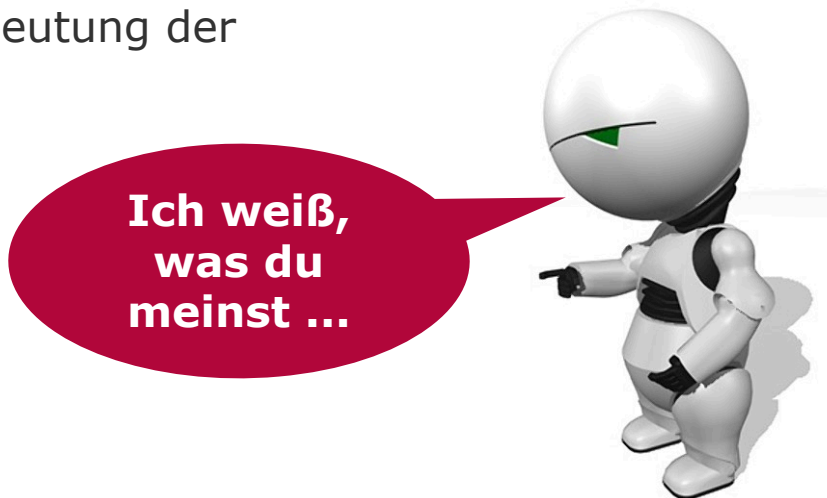


The screenshot shows the TIME website interface. At the top, there's a navigation bar with links to TIME INC. NETWORK, TIME, PEOPLE, FORTUNE, SPORTS ILLUSTRATED, and MORE. Below this is a red header with the TIME logo and a 'Subscribe' button. A sidebar on the left displays financial data for DOW, NASDAQ, and S&P 500, along with a 'JUST POSTED' section featuring articles like 'Watch Sir Ian McKellen Put a New Twist on an Old Favorite in Mr. Holmes Trailer' and 'One Third of the World's Population Suffers From Untreated Tooth Decay'. The main content area is titled 'THE BRIEF' and features a large article 'U.S. Envoy Bloodied in Seoul' with a photo of Mark Lippert. Other articles include 'Australians Await Execution', 'Clinton Emails to Be Released', 'Boston Bombing Survivors Give Testimony', and 'Who Will Lead Black Americans?'. A 'FEATURED' section highlights 'The Birth of a Nation 100 Years Later: Still Great, Still Shameful'. At the bottom, there's a sign-up form for 'THE BRIEF' and more content.

Vision für das Semantic Web

Tim Berners-Lee – Erfinder des WWW – im September 1998:

- „Das Web war als Informationsraum entworfen worden, mit dem Ziel nicht nur für Mensch-zu-Mensch Kommunikation nützlich zu sein, sondern auch, dass es Maschinen möglich ist, sich zu beteiligen und zu helfen.“
- **HTML** als Sprache zur Strukturierung von Informationen im WWW fehlt es an Möglichkeiten, die Bedeutung der Informationen auszudrücken



Ich weiß,
was du
meinst ...

Potenziale des Semantic Web (1/3)

Beispiel 1: Suchmaschinen heute

Suchanfrage: „Golf von 2015“

Nutzer muss Anfrage verfeinern, da Ergebnisse mehrdeutig sind

- Anfrage zu „Golf“:
 - Auto oder Sport oder sonstiges? → Verfeinerung: Auto
- Anfrage zu „2015“:
 - Jahr der Konstruktion oder technischer Parameter?
→ Verfeinerung: Jahr der Konstruktion
- Anfrage zu „Golf von 2015“:
 - Verkaufsangebot, Handbuch oder sonstiges?
→ Verfeinerung: Handbuch
- ...

Suchmaschine muss Nutzer nach Bedeutung der Anfrage fragen

Potenziale des Semantic Web (2/3)

Beispiel 1: Suchmaschinen von morgen

Suchanfrage: „Golf von 2015“

Suchmaschine „kennt“ den Nutzer

- Nutzer überprüft in letzter Zeit regelmäßig Ergebnisse von professionellen Golf-Turnieren
- Nutzer hat online im letzten Jahr Golfschläger gekauft
- Nutzer hat News-Gruppe „en.rec.golf.balls“ abonniert
- Nutzer hat Bildschirmhintergrund „PGA Tour 2014“ gewählt
- Nutzer besitzt die Domain www.thegolffanatic.com
- ...

Suchmaschine kann anhand dieses Wissens Bedeutung der Anfrage erschließen und richtig beantworten ...

Potenziale des Semantic Web (3/3)

Beispiel 2: Planung eines Familienurlaubs

Programm-Eingabe: Urlaub am Meer, ≥ 10 Tage, Familie, ≤ 2000 €

- Programm überprüft Kalender aller Familienmitglieder und sucht nach freien Zeiträumen
- Programm überprüft Urlaubsangebote für Zeiträume anhand der Anforderungen
 - geografische Parameter
 - familienfreundlich?
 - Golfplatz in der Nähe?
 - ...
- Programm überprüft freie Kapazitäten und stellt Auswahl aus passenden Angeboten zusammen
- Programm organisiert Transport anhand des ausgewählten Ortes

Ziele des Semantic Web

- Automatische Organisation von Wissen in der jeweiligen Domäne
- Bereitstellung automatisierter Werkzeuge für Wartung, Entfernung von Inkonsistenzen oder Extraktion von neuem Wissen
- Ersetzung der Schlüsselwort-basierten Suche durch Inhalts-basierte Suche mit „intelligenten“ Frage- und Antwort-Zyklen
- Nutzerfreundliche Suche
- Extraktion und Darstellung von Wissen
- Beantwortung von Fragen zu über verschiedenen Dokumente verteilte Informationen
- Exakte Bestimmung, wer auf welche Teile einer bestimmten Information zugreifen kann

Technische Realisierung des Semantic Web (1/2)

Realisierung des Semantic Web benötigt Anzahl von Technologien für jeweils verschiedene Zwecke:

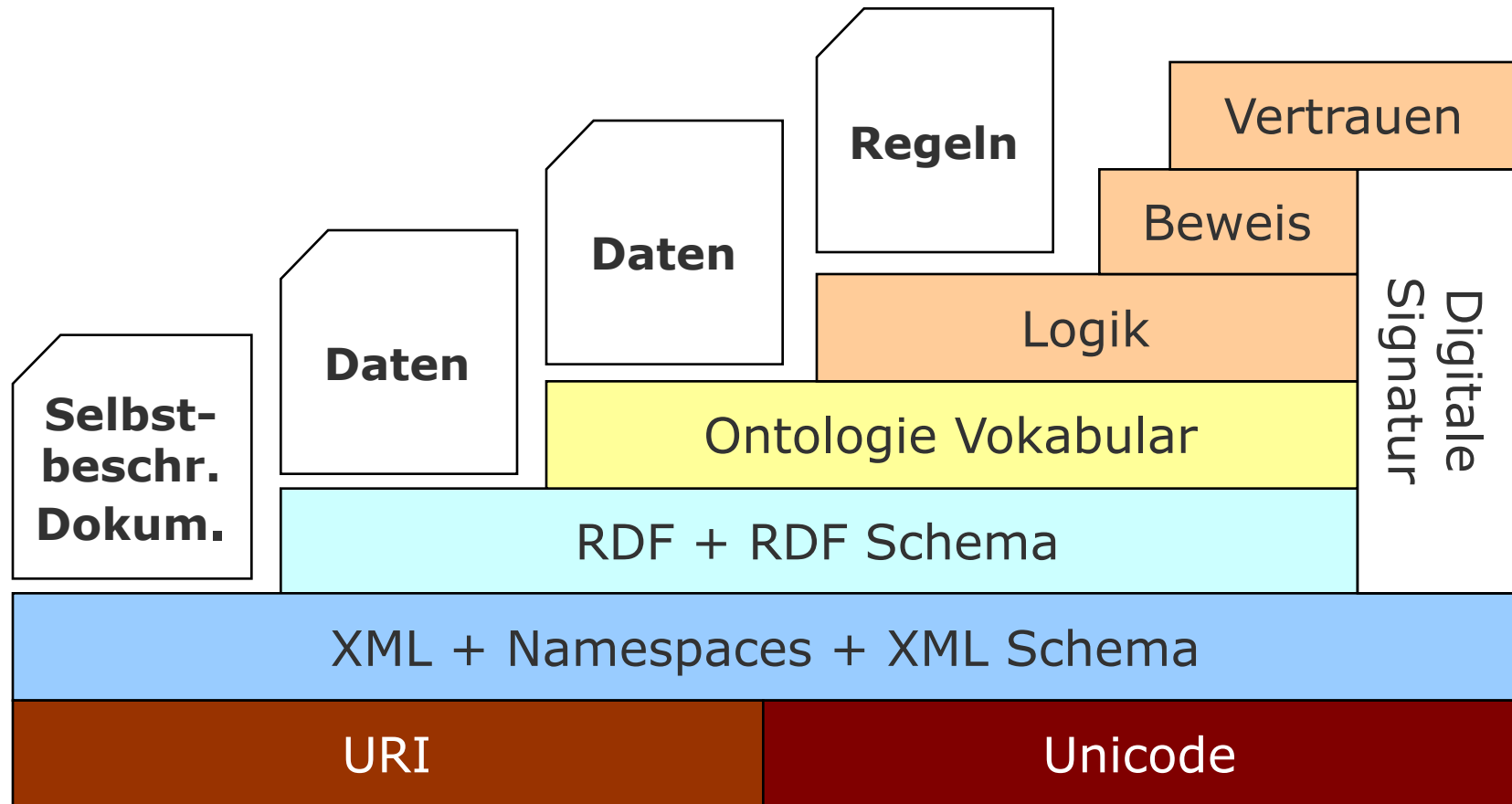
- Jede Informations-Quelle im Semantic Web benötigt eine eindeutige Kennung
→ **URI**
- Informationen im Semantic Web müssen formal ausgedrückt werden, um automatisiert verarbeitet werden zu können; dazu wird einheitliche Syntax gebraucht
→ **XML, XMLSchema**
- Einfache semantische Beziehungen zwischen Informations-Entitäten müssen ausgedrückt werden können
→ **RDF (Resource Description Framework), RDFSchema**
- ...

Technische Realisierung des Semantic Web (2/2)

Realisierung des Semantic Web benötigt eine Anzahl von **Technologien** für verschiedene Zwecke:

- ...
- Verschiedene Semantiken müssen beschreib- und erklärbar sein, um auch scheinbar inkonsistente Informationen kombinieren zu können
 - ➔ **Ontologien**
- Neue Information (Wissen) muss aus vorhandenen Informationen abgeleitet werden können
 - ➔ **Inferenz-Mechanismen**
- Datenschutz (in Bezug auf Vertraulichkeit und Integrität) muss gewährleistet werden können
 - ➔ **XMLEncryption and XMLSignature**

Semantic Web – Technology Stack



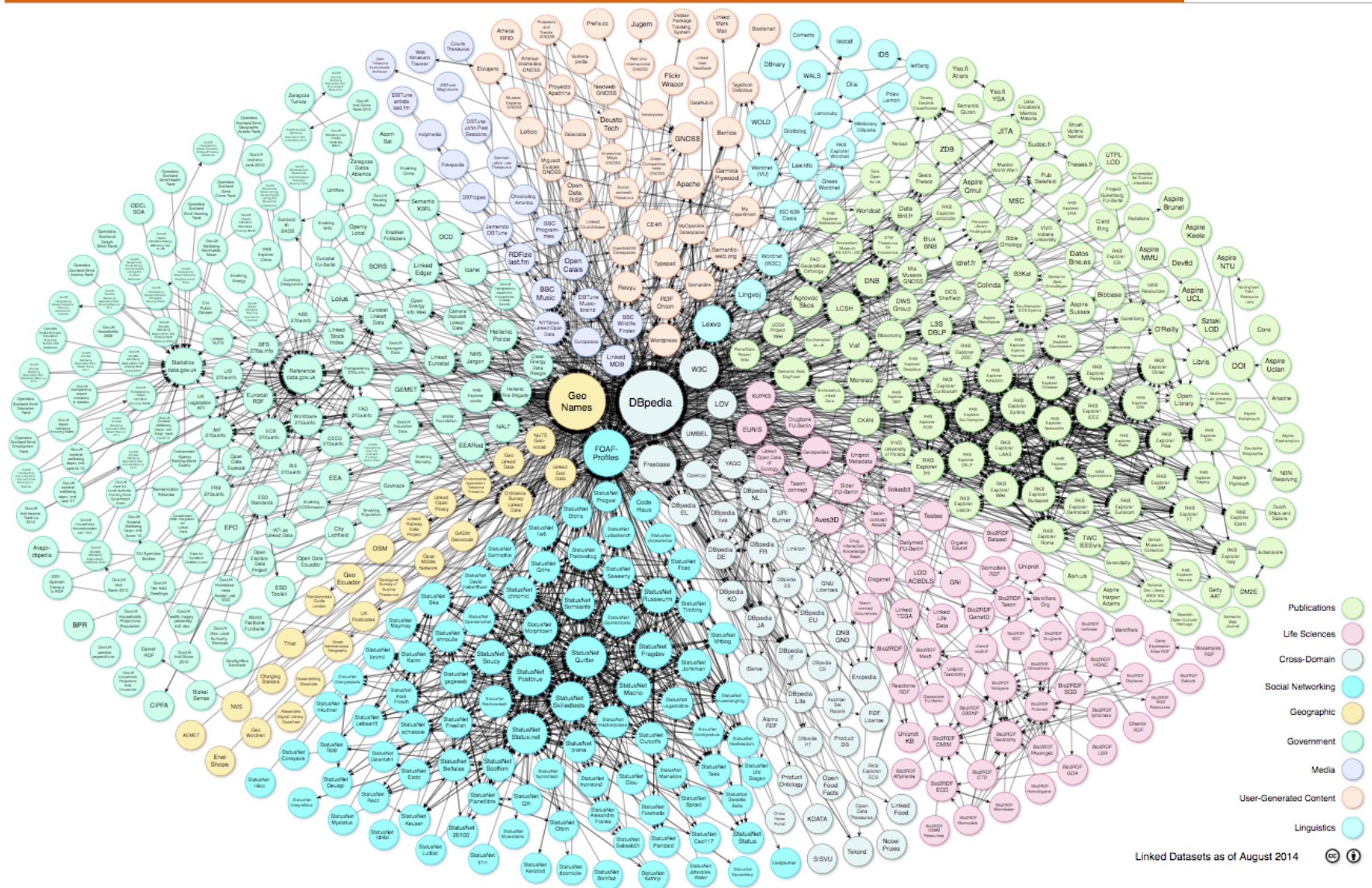
Weitergehende Informationen:

➔ openHPI Kurse zu **Semantic Web Technologien** von Prof. Dr. Harald Sack

Linked Open Data (1/2)

- Weltweites Netzwerk im Web frei verfügbarer Daten, die automatisiert von Programmen verarbeitet werden können
- Werden per URI und HTTP abgefragt
- Meist per RDF beschrieben und verlinkt, semantisch abfragbar per SPARQL
- Damit auch Teil des Semantic Web
- Bekannte Datensätze:
 - DBpedia: Extrahierte Informationen aus Wikipedia
 - FOAF: Datensatz über Personen und ihre Beziehungen
 - GeoNames: Informationen über Orte und deren Position
 - ...

Linked Open Data (2/2)





openHPI: **Web-Technologien** IoT – Internet der Dinge und Webservices

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

Digitale Transformation (1/2)

Entwicklung der digitalen Technologien – „Digitale Revolution“ – treibt **digitale Transformation** in allen Bereichen unserer Gesellschaft

- **Digitale Technologien** sind charakterisiert durch
 - breite Verfügbarkeit von Computern, Smartphones, Tablets, Programmierbaren Controllers und vielen neuen Smart Devices
 - **Internet** – als universelle globale Kommunikationsplattform
- Internet entwickelt sich immer mehr zum **Internet der Dinge – IoT**, in dem nicht nur Menschen sondern auch Dinge angesprochen und untereinander interagieren können
 - **Dinge werden smart** – bekommen digitale Hülle –, so dass sich neben der physikalischen Interaktionsebene eine neue, digitale Interaktionsebene etabliert

Digitale Transformation (2/2)

Dinge werden smart – bekommen digitale Hülle –, so dass sich neben der physikalischen Interaktionsebene eine neue, digitale Interaktionsebene etabliert

- Können mit Dingen nicht mehr nur physikalisch interagieren
 - berühren, verschieben, heben, anstoßen, zerdrücken, ...
- Können mit Dingen **digital** interagieren, d.h.
 - mit Lichtgeschwindigkeit und
 - über beliebig große Entfernungen hinweg
- Vermittels ihrer digitalen Hüllen können erstmals auch Dinge miteinander interagieren über einen digitalen Kanal („Internet“)
 - Bestands- und Sensordaten austauschen
 - Warnungen ausgeben
 - Aktivitäten und Reaktionen auslösen ...

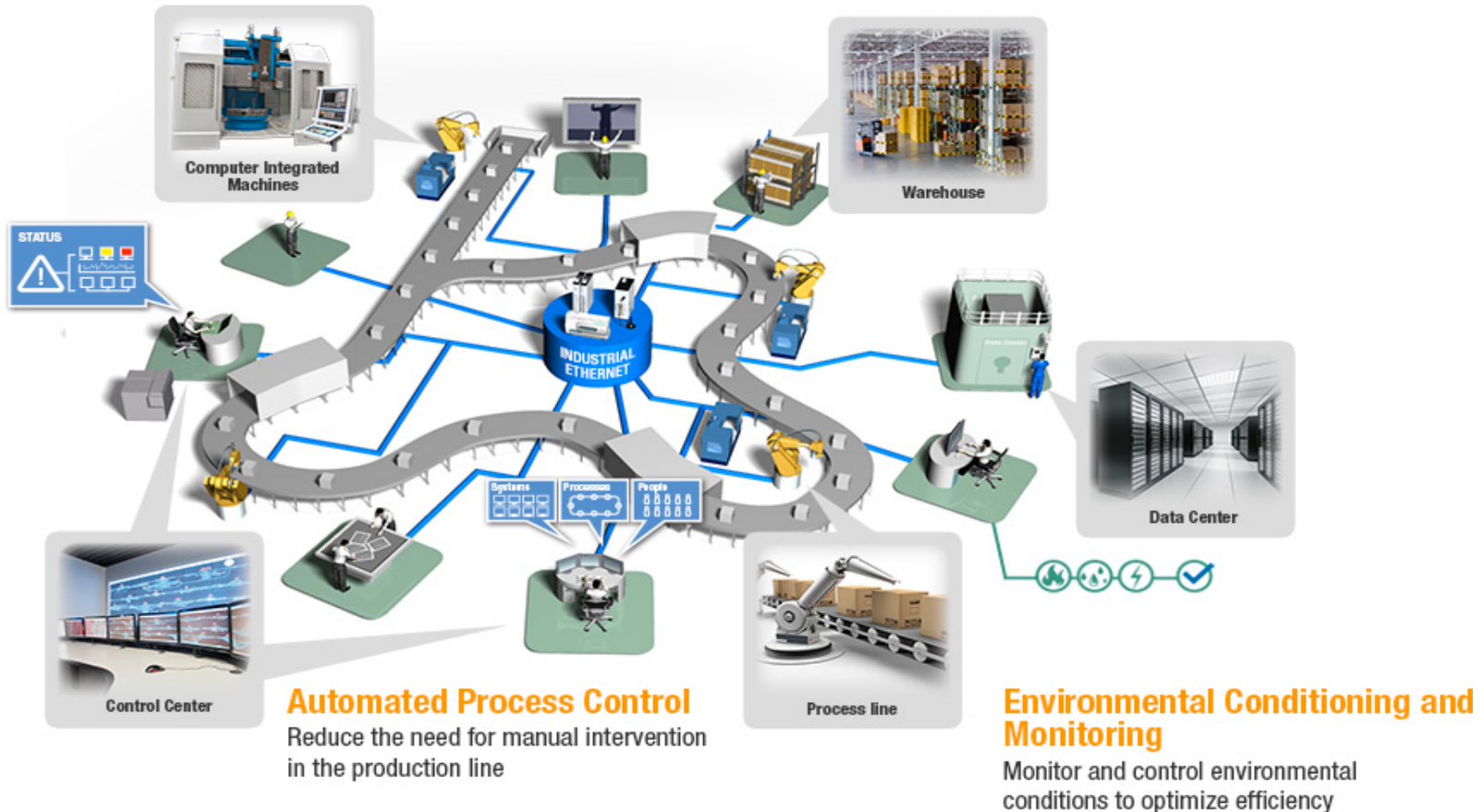
Digitale Transformation in der Wirtschaft: Smart Factory und Industrie 4.0

Computer-Integrated Manufacturing

Real-time and accurate collection of production line data

Real-time Production Monitoring

Greater control over the production process



http://www.moxa.com/Event/DAC/2013/Factory_Automation_IO/images/Factory_Automation.jpg

■ Trends

- Verbindung von industriellen Netzwerken mit dem Internet
- Zusätzliche, oft drahtlose, Sensoren und Aktuatoren

■ Herausforderungen

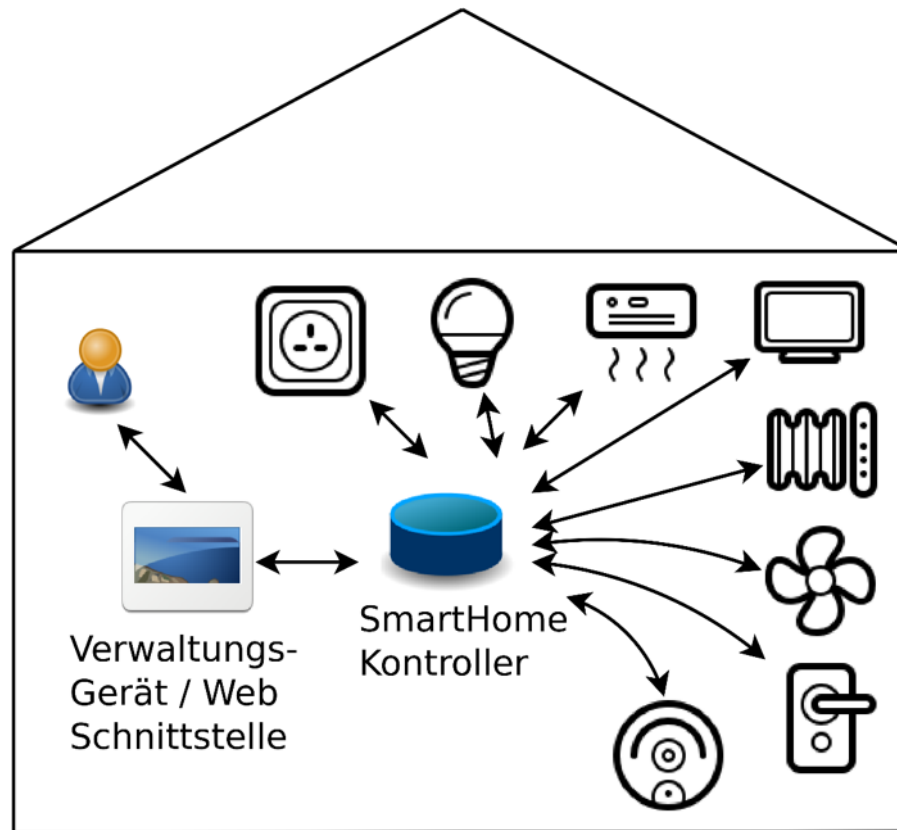
- Umgang mit Ressourcenbeschränkungen
 - um Größenordnungen unterschiedliche Reaktionszeiten
 - Sicherheit in allen Aspekten
 - Kombination von Informatik mit Elektrotechnik, Maschinenbau, Automatisierungstechnik, ...
- Interaktion mit industrieller Technik (genauer: deren digitaler Hülle) erfordert Schnittstellen, auf die über Internet zugegriffen werden kann → **Webservices** als mögliche Technologie

Digitale Transformation: Smart Home (1/2)

Smart Home – weiterer Anwendungsfall für Internet der Dinge: Steuerung von vernetzter Haustechnik mithilfe von Controllern

■ Physikalische Systeme in Häusern / Wohnungen erhalten digitale Hülle

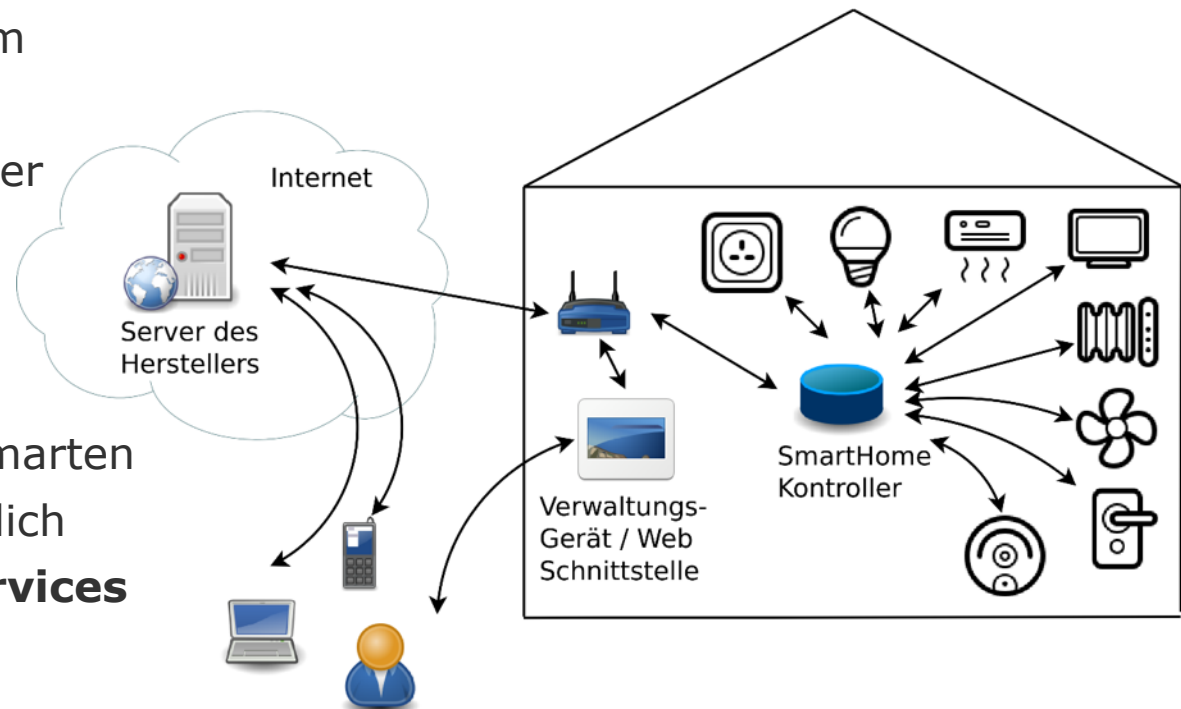
- Licht- und Unterhaltselektronik
- Heizung und Klimaanlage
- Fenster und Türen
- Alarmanlage
- Jalousien
- ...



Digitale Transformation: Smart Home (2/2)

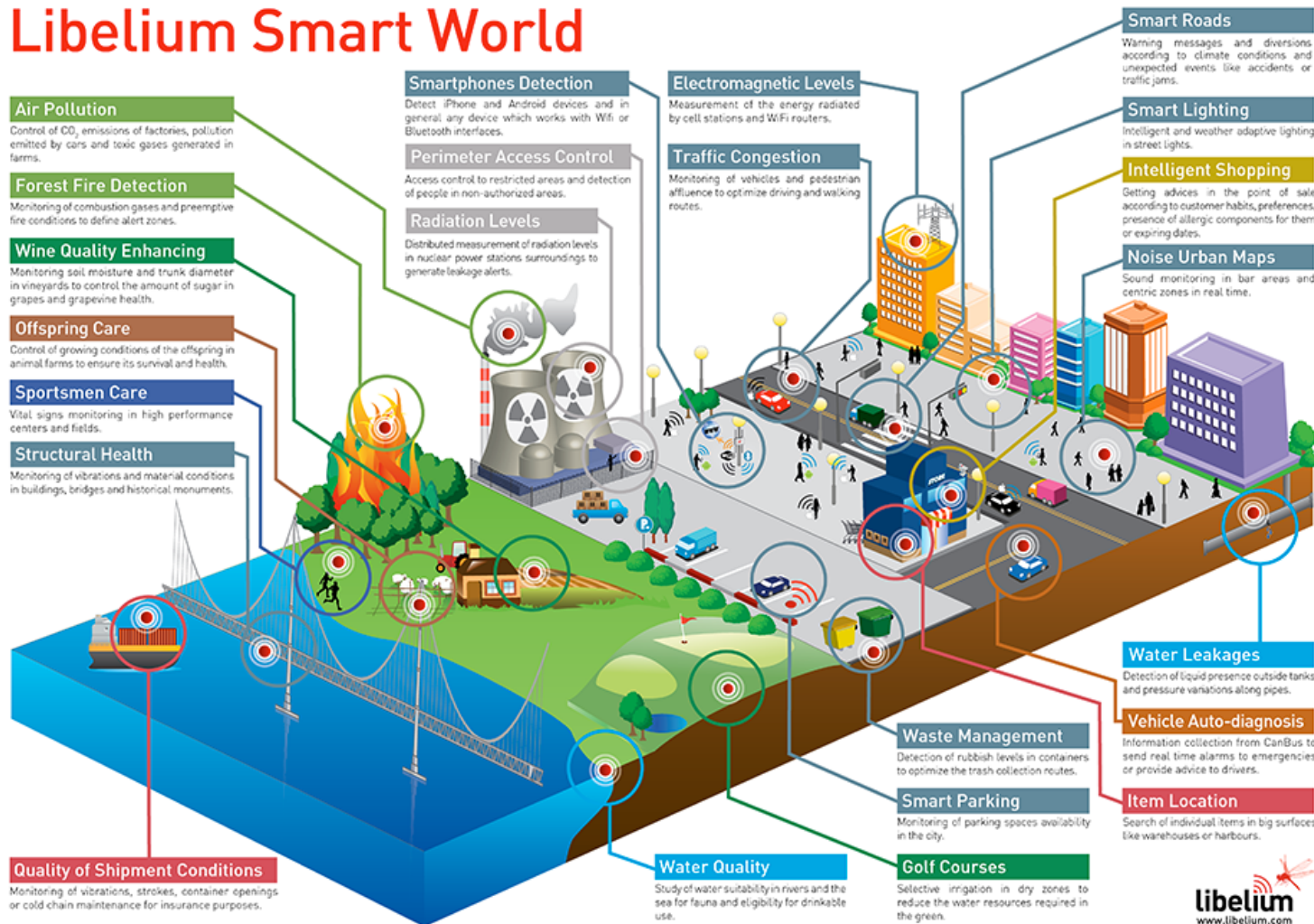
Viele Smart Home Lösungen nutzen Cloud-Technologie, um Verwaltung der Haustechnik für Nutzer zu erleichtern

- Kontroller des Smart Home System ist über Internet ständig mit Server des Herstellers verbunden
- Nutzer kann von jedem beliebigen Ort über Internet und den Server des Herstellers sein Smart Home System verwalten
- Kommunikation mit smarten Geräten im Haus möglich wieder mit → **Webservices**



Digitale Transformation: Alles mündet in die Vision einer Smart World

Libelium Smart World



http://www.libelium.com/resources/top_50_iot_sensor_applications_ranking

Internet der Dinge und Webservices

Ressourcen im Web sind nicht mehr nur Dokumente, sondern verstärkt auch **physische Geräte**, die über **Webservice** Technologien **digitale Schnittstellen** anbieten

- Menschen können über (Web-)Anwendungen mit Geräten interagieren, z.B. Steuerung der Heizung im Smart Home
- Maschinen interagieren auch untereinander, z.B. kann Jalousiensteuerung im Smart Home den Webservice eines Wetterdienstes nach Vorhersagen abfragen

Digitale Transformation eröffnet **riesige Potenziale** neuer Möglichkeiten, bringt aber auch **Herausforderungen** mit sich

- Kleine Smart Devices (z.B. Sensoren) müssen **energiesparend** gebaut werden → **schlanke Kommunikationsprotokolle**
- Nachdenken über **Sicherheit** und **Schutz der Privatsphäre** muss vor Hintergrund der Smart World noch mehr in den Fokus rücken



openHPI: **Web-Technologien** Privatsphäre und Sicherheit in einer hochvernetzten Welt

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

Smart Cars – ein spektakulärer Angriff

Spektakuläre Angriffsdemonstration aus 2015:
Hacker übernehmen **über das Internet**
ein Auto während der Fahrt



Sicherheitsforscher Miller und Valasek demonstrieren für WIRED Magazin Übernahme eines modernen SUV

1. Angreifer können Internet-Adresse des Autos feststellen und verbinden sich damit über dessen Internetanbindung
2. Über verschiedene Sicherheitslücken und Schwächen im Design gelingt es ihnen, das Car Entertainment System zu übernehmen und von dort auf internes Bordnetzwerk zuzugreifen
3. Weitergehende Manipulation der Software des Fahrzeugs erlaubt umfassende „Fernsteuerung“



Den Angreifern war es möglich:

- Klimaanlage zu manipulieren
- Soundsystem zu bedienen
- Wischer, Lichter, etc. zu steuern
- GPS-Position und Geschwindigkeit auszulesen
- **Gaspedal und Bremse zu bedienen**
- Motor abzuwürgen
- Lenkung teilweise zu bedienen
(derzeit nur im Rückwärtsgang)

(Vernetzte) Computersysteme haben Schwachstellen und können über das Internet massiv manipuliert werden!



- Früher zielten Hacker-Angriffe meist auf rein digitale Systeme oder auf Nutzer des Webs ab
- Integration digitaler Komponenten in physikalische Systeme (etwa bei Smart Factory oder im Smart Home) macht auch Cyberangriffe auf physikalische Systeme möglich
 - unberechtigtes An-/Abschalten
 - Auslesen von Daten oder Veränderung von Konfigurationen
 - Beschädigung oder Zerstörung der Systeme
 - ...
- Angriffe können nicht mehr „nur“ Daten, Reputation oder finanzielle Ressourcen treffen, sondern auch Maschinen oder gar die Unversehrtheit von Menschen – wie im Fall des Auto-Hacks – werden zu direkten oder mittelbaren Zielen von Cyberattacken

Risiken für das Smart Home

SmartHome-Controller, die Fernsteuerung der Haustechnik erlauben, bieten (noch) zahlreiche **gefährliche Angriffspunkte**

Beispiel:

- MiCasaVerde VeraLite hat Schwachstellen, die es Angreifer erlauben
 - **beliebige Daten** aus Dateisystem auszulesen (inklusive gehashter Passwörter)
 - **Firmware zu ändern** und **Einstellungen herunterzuladen**
 - aus dem LAN heraus **als Administrator jede Art von Schadsoftware** auszuführen
 - mit Kenntnis einer gültigen GeräteID **vollständigen Zugriff** auf den SmartHome-Controller zu erlangen

Risiken für das Smart Home – Smart-TV

Nicht nur SmartHome-Controller, sondern auch andere „Smart“ Geräte haben **Probleme mit Privacy** und **Sicherheit**

■ Samsung Smart-TV

- hört alle Gespräche (und andere Geräusche) im Raum mit, falls Sprachsteuerungsfunktion aktiviert ist
- gesprochene Worte werden “an Drittanbieter übertragen”

■ LG Fernseher

- senden verschiedene Nutzer-Daten an Hersteller, um Smart Ad (gezielte Werbung) zu zeigen
- Informationen werden auch übertragen, wenn entsprechende Einstellung deaktiviert ist (!)
- übertragene Daten: angesehene TV-Kanäle, Datei-Listen von angeschlossenen USB Sticks und Festplatten, Suchanfragen, ...

Risiken für das Smart Home – Smart Meter

Smart Meter sind Verbrauchsmesser („Meter“) für Strom, Gas, Wasser mit Netzanschluss

- Analoge Verbrauchsmesser werden 1-2 mal pro Jahr abgelesen
- Smart Meter sind ständig mit Energie-Anbieter verbunden und schicken ihm Messwerte zu
- Anbieter erkennt zeitnah Energiebedarf und kann diesen z.B. über Tarifgestaltung optimieren (Strom in der Nacht günstiger, ...)
- Bilanzierungsprozesse, Prognosen, Rechnungserstellung, ... werden erleichtert
- Typischerweise eröffnen Smart-Tarife dem Verbraucher auch Sparmöglichkeiten



EVB Energie AG, via Wikimedia Commons

Risiken für das Smart Home – Smart Meter

- Detaillierte Informationen über Energie-Nutzung in einem Haushalt sagt viel über Bewohner aus
 - Schlafzeiten
 - Waschzeiten
 - wie viele Personen sind gerade zu Hause
 - Kochzeiten
 - ...
- Polizei in Texas (USA) nutzt Strom-Daten, um Marihuana-Plantagen zu erkennen
- In Australien wurden solche Daten an Inkasso-Unternehmen weitergegeben
- Smart Meter ermöglicht dem Energie-Anbieter – und damit auch Angreifern – auch Fernabschaltung der Energieversorgung

Risiken für Smart Factory und Industrie 4.0

Industrielle Anlagen waren bisher **isolierte Systeme**

- Steuerungsanlagen (PLCs) für physikalische Maschinen und industrielle Netzwerke auf Zuverlässigkeit und sicheren Betrieb ausgelegt
- Keine Verbindung mit anderen Unternehmensnetzwerken und dem Internet

Industrie 4.0 und Smart Factory bedingen **hochvernetzte Anlagen**

→ dadurch **anfälliger für Cyberattacken**

- Datendiebstahl
- Spionage
- DoS-Attacken (Denial of Service – Ziel: Betriebsschädigung und Außerbetriebnahme von Anlagen)
- Malware (Viren, Würmer, Trojaner)
- ...

Risiken für Smart Factory und Industrie 4.0

Beispiel für cyber-physikalischen Angriff: **Stuxnet**

- **Stuxnet** ist **Schadsoftware**, die 2010 auf den Computern von iranischen Anlagen zur Urananreicherung entdeckt wurde
- Schlägt in zwei Phasen zu:
 - Infektion von Windows-Systemen
 - Stuxnet verbreitet sich über Netzwerk und USB-Sticks
 - Ist Windows-Rechner infiziert, prüft Stuxnet, ob Software zur Programmierung von Steueranlagen installiert ist
 - Wenn ja, modifiziert Stuxnet diese Software
 - Infektion von PLCs (Programmable Logic Controller, Steueranlagen)
 - Schadsoftware verändert einzelne Codeblöcke in den Steueranlagen
 - ...

Risiken für die Smart Factory und Industrie 4.0

Beispiel für cyber-physikalischen Angriff: **Stuxnet**

- **Stuxnet** ist **Schadsoftware**, die 2010 auf den Computern von iranischen Anlagen zur Urananreicherung entdeckt wurde
- Schlägt in zwei Phasen zu:
 - Infektion von PLCs (Programmable Logic Controller, Steueranlagen)
 - ...
 - Zusätzlich verschleiert Stuxnet Manipulationen, dem Betreiber werden keinerlei Abweichungen im Betriebsablauf gemeldet
→ so blieb Stuxnet lange unentdeckt ...
- 2010 waren Angriffe auf Steuerungsanlagen nur Fiktion, die in einigen Forschungspapieren behandelt wurde
- Zunehmende Vernetzung, auch mit Internet und Web Services, **vergrößert Angriffsflächen** deutlich

Identitätsdiebstahl – ein spektakulärer Fall (1/3)

Angriff auf Mat Honan, Autor bei Wired.com:

- Wired.com: Bekanntes US-amerikanisches Technologiemagazin
- Mat Honan (@mat) ist Senior Staff Writer bei Wired
- 2012: Angreifer haben mit einer mehrstufigen Attacke erreicht, dass alle Daten sowohl auf Honans MacBook, iPhone und iPad sowie den Online-Backups gelöscht wurden
- Als Warnung veröffentlichte Honan Ablauf der Ereignisse in einem Artikel



Identitätsdiebstahl – ein spektakulärer Fall (2/3)

Angriff auf Mat Honan, Autor bei Wired.com:

- Angreifer wollten Twitter-Konto von Mat Honan übernehmen
- Twitter-Konto enthält Information über persönliche Webseite
- Mithilfe des „Whois“-Dienstes finden Angreifer Honans Postadresse und private Gmail-Adresse heraus
 - „Whois“ erlaubt Auslesen hinterlegter Daten der Domaininhaber (hier: von Honans persönlicher Website)
- Durch Passwort-Wiederherstellung im Gmail-Account finden Angreifer Email-Adresse der Apple ID heraus
 - Gmail: „Wir haben eine Email an [m****n@me.com](#) gesendet“
 - Angreifer rät vollständige Email-Adresse ...

Identitätsdiebstahl – ein spektakulärer Fall (3/3)

Angriff auf Mat Honan, Autor bei Wired.com:

Angreifer kennen nun Name, Postadresse, private Gmail-Adresse und Email der Apple ID

- kontaktieren Amazon und lassen eine Kredit-Karte mit Honans Konto verbinden
- kontaktieren Amazon erneut und lassen Passwort wiederherstellen (möglich durch Kenntnis der zuvor hinterlegten Kreditkarten-Nummer)
- finden im Amazon-Konto letzte vier Ziffern echter Kreditkarte von Honan
- kontaktieren Apple, erhalten Zugang zur Apple ID (durch Kenntnis von Name, Adresse, Email und letzten vier Ziffern der Kreditkarte)
- mit Apple ID bzw. Zugriff auf Apple-Mailkonto setzten Angreifer Gmail-Passwort (und anschließend auch Twitter-Passwort) zurück und löschen alle Daten von Honan's iPhone, iPad und MacBook...

Zusammenfassung

- IoT, Smart World, hochgradige Vernetzung und weitergehende Durchdringung unserer Gesellschaft mit (Web-)Anwendungen läuten **neues technologisches Zeitalter** ein (4. industrielle Revolution)
- Digitale Transformation birgt aber gleichzeitig **hohes Gefahrenpotential** für Industrie und Gesellschaft
 - neue Arten von Angriffen, nicht nur auf digitale Systeme
 - inhärent Verletzung der (digitalen) Privatsphäre, bis hin zu deren faktischer Abschaffung
- Sicherheit und Datenschutz sind **zentrale Herausforderungen**, denen sich Forschung, Politik, Wirtschaft – aber auch **jeder einzelne** – stellen muss
 - Schaffung geeigneter politischer Rahmenbedingungen
 - Bereitstellung technischer Hilfsmittel
 - Sicherheitsbewusstsein → openHPI-Kurs „**Sicherheit im Internet**“



openHPI: **Web-Technologien** Kursresumé

Prof. Dr. Christoph Meinel

Hasso-Plattner-Institut

Universität Potsdam

World Wide Web ist **verteiltes Hypermedia-System**, untereinander verlinkte Web-Dokumente liegen weltweit auf verschiedenen Servern

- Zum Abruf der Dokumente wird eindeutiges Adressierungsschema benötigt
 - Jedes Dokument braucht (mindestens) einen **Uniform Resource Identifier – URI**
 - URI meist als **Uniform Resource Locator – URL** – realisiert, Identifikation von Ressourcen erfolgt über deren „Adresse“ (Domain Name des Server, Pfad auf dem Server, Dateiname)
- Anforderung und Übertragung der Dokumente erfolgt mit **Hypertext Transfer Protocol – HTTP**
 - HTTP bietet einfaches Frage-Antwort-Schema (Request / Response), mit dem Clients (Browser) bei Servern Dokumente anfordern können

World Wide Web ist **verteiltes Hypermedia-System**, untereinander verlinkte Web-Dokumente liegen weltweit auf verschiedenen Servern

- Erweiterungen für HTTP verbessern Performance:
 - **Persistente Verbindungen** und Pipelining
 - **Kompression**
 - **Caching** (in Clients, Servern und Zwischensystemen – Proxies)
- Überwindung der Zustandslosigkeit von HTTP mittels **Sessions und Cookies** für Web-Anwendungen möglich
- Browser und Server können Auslieferung von Content-Varianten aushandeln – **Content Negotiation**
- Nach über 25 Jahren wird mit **HTTP/2** derzeit Nachfolger für das zentrale Web-Protokoll eingeführt

Dokumente im WWW werden mit HTML und CSS beschrieben

- Trennung von Struktur und Gestaltung
 - **Hypertext Markup Language – HTML** – für strukturierte Aufbereitung des Dokumenteninhalts
 - **Cascading Stylesheets – CSS** – für Gestaltung der Darstellung von Dokumenten auf verschiedenen Ausgabegeräten
- HTML führt **Hyperlink**-Konzept ein
- HTML-Dokumente bestehen aus
 - Head: Informationen *über* den Inhalt (Metainformationen)
 - Body: eigentlicher Inhalte
- HTML ist **Markup-Sprache**: Strukturgebende Markups („Tags“) sind Teil des Dokumenteninhalts und werden durch Markup-Trenner (<...>) gekennzeichnet

Dokumente im WWW werden mit HTML und CSS beschrieben

- Aktuelle Version **HTML 5** führt „sprechende“ Strukturelemente, Multimedia-Tags, neue Formular-Funktionen und Zeichenfläche (Canvas) ein
- CSS-Regeln legen fest, welche HTML-Elemente wie gestaltet werden sollen
 - **Selektoren** wählen zu gestaltendes Element aus
 - **Deklarationen** weisen bestimmten Gestaltungsmerkmalen einen Wert zu
- **Responsives Webdesign** sorgt dafür, dass Webseiten automatisch an spezifische Anforderungen der verschiedenen Endgeräte (Desktop, Smartphone, Tablet, ...) angepasst werden
- **Extensible Markup Language – XML** ist (wie SGML) Meta-Markupsprache, also Sprache zur Beschreibung von Sprachen

Komplexere Nutzerinteraktion im Web und Web-Anwendungen werden mittels Web-Programmierung realisiert

- Wir unterscheiden **clientseitige** und **serverseitige** Web-Programmierung
- Bei clientseitiger Web-Programmierung wird Programmcode mit Webseite ausgeliefert und dann **im Browser ausgeführt**
- **Document Object Model – DOM** – bietet Schnittstelle für die Manipulation von HTML-Elementen
- Bei serverseitiger Web-Programmierung wird **Programm auf dem Server ausgeführt** und Ausgabe (z.B. HTML-Dokument) an Browser zur Anzeige ausgeliefert
- Für **Kodierung von Informationen** (Text, Grafik, Multimedia) zur Übermittlung im Web gibt es spezifische Anforderungen und Formate

JavaScript ist das wichtigste Werkzeug für clientseitige Web-Programmierung

- JavaScript (standardisiert als ECMAScript) ist Skriptsprache zur Manipulation von HTML-Elementen
- Für viele gängige Anwendungsfälle gibt es umfangreiche **Bibliotheken** und **Frameworks**, z.B. jQuery oder D3
- Auch für CSS gibt es Reihe an Frameworks, die wiederkehrende Aufgaben vereinfachen
 - **CSS-Präprozessoren** erweitern CSS-Syntax, z.B. um Variablen und Funktionen
- **AJAX – Asynchronous JavaScript and XML** – erlaubt Kommunikation zwischen clientseitiger Web-Anwendung und Servern, z.B. zur hintergründigen Anfrage zusätzlicher Ressourcen

Viele Aufgaben der Web-Programmierung – insbesondere solche, bei denen große Datenmengen zu verarbeiten sind – werden auf dem Server gelöst

- Webserver ruft **serverseitiges Programm** auf (z.B. über **CGI-Schnittstelle**) und übergibt Anfrageparameter aus dem HTTP-Request
- Programm auf dem Server wird ausgeführt, erzeugt **parameterabhängige Ausgabe**, z.B. als **HTML** oder **JSON**, und gibt diese an Webserver zurück (der diese als HTTP-Response an Client ausliefert)
- Serverseitige Web-Programmierung basiert heute fast immer auf mächtigen **Web-Frameworks**, z.B. Ruby on Rails oder Django

Viele Aufgaben der Web-Programmierung – insbesondere solche, bei denen große Datenmengen verarbeitet werden – werden auf dem Server gelöst

- Daten auf dem Server werden **in Datenbanken persistiert** und über Web-Framework mit **Object Relational Mapper – ORM** – abgefragt
 - **Relationale Datenbanken:** Fixe Datenstruktur (Tabellen), hohe Datenkonsistenz
 - **NoSQL-Datenbanken:** Flexible Datenstruktur, leichter zu skalieren, Konsistenz schwieriger zu gewährleisten
- **Web Services** ermöglichen **Kommunikation zwischen** unterschiedlichen **Web-Anwendungen** und erlauben so Realisierung verteilter Applikationen

Soziale Netzwerke, das Semantic Web und Anwendungen für Web Services (z.B. im IoT Bereich) sind wichtige Themen für das Web von heute und morgen

- **Soziale Netzwerke** und **Social Media** sind aktuelle Killer-Applikationen des Webs
- Kritisch zu betrachten sind trotz oder gerade wegen der weitgehenden Verbreitung
 - Manipulationsmöglichkeiten bei der Meinungsbildung
 - Implikationen für Datenschutz und Sicherheit
- **Semantisches Web** verknüpft nicht länger nur Dokumente, sondern auch **Dinge** (Menschen, Orten, Ereignissen, ...) und setzt sie in **Beziehung**
 - **Bedeutung** einer Information wird auch erkennbar für Maschine

Soziale Netzwerke, das Semantic Web und Anwendungen für Web Services (z.B. im IoT-Bereich) sind wichtige Themen für das Web von heute und morgen

- **Internet of Things:** „Dinge“ (physikalische Geräte, Sensoren, ...) erhalten „digitale Hülle“ – können über Entfernung angesprochen werden oder auch miteinander kommunizieren
- **Smart Home, Smart Factory, Smart Traffic, ...** zeigen Visionen auf für mögliche Veränderungen dank digitaler Technologien
- **Sicherheit** und **Datenschutz** sind zentrale Herausforderungen für tiefergehende Verankerung der Smart-Technologien in der Gesellschaft

Wir verabschieden uns und ...

- ... hoffen, Sie hatten Freude im Kurs und konnten etwas dazulernen
- Alle Kursinhalte stehen auch nach Ende des Kurses weiter zur Verfügung



Prof. Dr.
Christoph Meinel



Christian
Willems



Matthias
Bauer



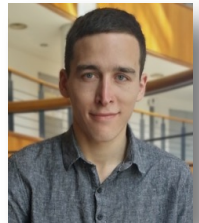
Martin
Malchow



Thomas
Staubitz



Tobias
Rohloff



Franz
Liedke

Wir wünschen Ihnen viel Erfolg bei der
Abschlussprüfung!