

Introduction to Knowledge Graphs and the Semantic Web

Questions



- What are [Graph Databases](#)?
- What are [Knowledge Graphs](#)
- What is the [Semantic Web](#)?
- How are they all related?
- How are they being used today?
- What can we expect in the future?

Web is our greatest knowledge source



Designed for people, not machines



Designed for people, not machines

- Content is mostly text, spoken language, images and videos
- Easy for people to understand
- But hard for machines

Machines need access to this knowledge too



Access via information retrieval

A black and white photograph of Vannevar Bush in a laboratory setting. He is wearing a suit and glasses, leaning over a large, complex mechanical device with many horizontal rods and levers. The device appears to be a prototype for a memory or information retrieval system. The background shows shelves with various items, possibly books or documents.

Vannevar Bush envisioned a hypertext/IR system in 1945



Access is primarily via information retrieval

- Key-word queries→ranked document list

- We still need to read the documents or watch the videos

- We often want an answer to a question:

where is the Census Big Data Day event

And so do our machines and apps

Vannevar Bush envisioned a hypertext/IR

We need to add knowledge graphs



The background of the slide is a dense network graph. The nodes are represented by small, colorful icons of various shapes and sizes, including logos for companies like HP, Google, and Microsoft, as well as various abstract symbols. The edges are thin, light-colored lines that connect these nodes, creating a complex web of relationships. A semi-transparent black rectangular box is centered on the slide, containing white text. The text is the main message of the slide, which is to add knowledge graphs to the web. The overall aesthetic is technical and data-driven.

We need to add knowledge graphs

- High quality semi-structured information about entities and relations
- Represented and accessed via Web standards
- Easily integrated, fused and reasoned with

State of the Art?



Google is a good example, but Microsoft, IBM, Apple and Facebook all have similar capabilities

- 2010 Google acquired MediaWeb and its **Freebase** KB
- 2014: Freebase: 1.2B facts about 43M entities
- 2015+: Google knowledge graph, updated by text IE

DBpedia open source RDF KB is another

- 800M facts about 4.6M subjects from English **Wikipedia**, data available in 21 other languages
- Helps integrate 90B facts from 1000 RDF datasets in the linked data cloud

Ask: When was Tom Sawyer written?

when was tom sawyer written

About 501,000 results (0.56 seconds)

The Adventures of Tom Sawyer / Date written

1876

The Adventures of Tom Sawyer
MARK TWAIN

Aunt Polly (aunt), Sally Phelps (aunt), Mary (cousin), Sid (half-brother)
Thomas "Tom" Sawyer is the title character of the Mark Twain novel *The Adventures of Tom Sawyer* (1876). He appears in three other novels by Twain: *Adventures of Huckleberry Finn* (1884), *Tom Sawyer Abroad* (1894), and *Tom Sawyer, Detective* (1896).

[Tom Sawyer - Wikipedia](https://en.wikipedia.org/wiki/Tom_Sawyer)
https://en.wikipedia.org/wiki/Tom_Sawyer

Feedback

People also ask

- Where was *The Adventures of Tom Sawyer* first published?
- How old is Tom Sawyer in the book?
- What is the setting for *The Adventures of Tom Sawyer*?
- Who is Tom Sawyer in real life?

Feedback

The Adventures of Tom Sawyer
Novel by Mark Twain

[Preview book](#)

Originally published: 1876
Author: Mark Twain
Text: *The Adventures of Tom Sawyer* at Wikisource
Cover artist: Created by Mark Twain
Characters: Tom Sawyer, Huckleberry Finn, Becky Thatcher, Aunt Polly, Joe Harper, Sid Sawyer
Genres: Bildungsroman, Picaresque Fiction, Satire, Folklore, Children's literature
Followed by: *Wuthering Heights*, *The Prince and the Pauper*

Feedback

When was Tom Sawyer written?

- What does Tom Sawyer refer to?
- A real person? A fictional character? A film? A TV show? A book? A facility?
- Querying Wikidata returns more than 100 reasonable entities (See them [here](#))
- Systems can use word embeddings to decide we probably are referring to a book, but there are 17 possibilities in the first 100
- Simple estimates of prominence let us rank the results and predict it's Mark Twain's [The Adventures of Tom Sawyer](#)



BROWSE ▾

Find a recipe

Ingredient Search



Create a profile



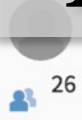
Home > Recipes > Desserts > Pies > Fruit Pies

Apple Pie by Grandma Ople



Grandma Ople's Apple Pie
★★★★★ 1930

Many commercial recipe sites on Web



26

"This was my grandmother's apple pie recipe. I have never seen another one quite like it. It will always be my favorite and has won me several first place prizes in local competitions. I hope it becomes one of your favorites as well!"



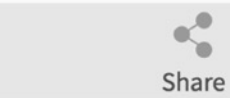
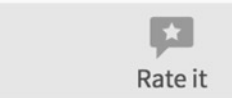
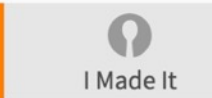
Recipes Videos Categories Articles



Blueberry Pie
★★★★★ 1K



All-Day Apple Butter
★★★★★ 883



SCHEMA.ORG (GooglePlus rich snippets)

```

itemprop:url (6) itemprop:mainEntityOfPage (1) itemprop:title (5) itemprop:image (2) itemprop:video (1) itemtype:http://schema.org/VideoObject (1) itemprop:thumbnailUrl (1) itemprop:embedUrl (1) itemprop:name (2) itemprop:description (2) itemprop:interactionCount (1) itemprop:uploadDate (1) itemprop:aggregateRating (1) itemtype:http://schema.org/AggregateRating (1) itemprop:ratingValue (12) itemprop:reviewCount (1) itemprop:author (12) itemprop:recipeYield (1) itemprop:nutrition (2) itemtype:http://schema.org/NutritionInformation (2) itemprop:calories (1) itemprop:fatContent (1) itemprop:carbohydrateContent (1) itemprop:proteinContent (1) itemprop:cholesterolContent (1) itemprop:sodiumContent (1) itemprop:ingredients (7) itemprop:prepTime (1) itemprop:cookTime (1) itemprop:totalTime (1) itemprop:recipeInstructions (1) itemprop:review (11) itemtype:http://schema.org/Review (11) itemprop:itemReviewed (11) itemprop:reviewRating (11) itemtype:http://schema.org/Rating (11) itemprop:dateCreated (11) itemprop:reviewBody (11)

```

9K made it | 6969 reviews

Most recipe sites embed **semantic data** about their recipes in an RDF-compatible form using terms from the **schema.org** ontology

Search engines read and use this data to better understand the semantics of the page content



Related

Recipes Videos Categories Articles

Day Apple Butter

★★★★★ 1K

★★★★★ 883

26

Save

I Made It

Rate it

Share

Print

Conversational Bots

Voice-driven conversational systems like Siri, Amazon Echo and Google Home use knowledge graphs to help understand our requests



Where does the knowledge come from?

- Knowledge graphs like [DBpedia](#) & Freebase started with **Wikipedia** data encoded in custom ontologies (aka schemas)
- Semantic Web technologies are an open source way to encode the knowledge
- They are and will continue to evolve, see [Wikidata](#)
- Current: extract data from text documents, e.g., articles, newswire, social media, etc.

Knowledge Graphs for AI

“Knowledge graphs” of one kind or another have been used for more than 60 years for many AI tasks, especially those involving language understanding or common-sense reasoning

How we got here

Over the decades much important and useful knowledge representation work has been done in support of AI

These are just a few familiar examples

- Micro-planner
- Semantic networks + logic
- Minsky Frames
- Schank Scripts
- Object oriented systems
- What's in a link?
- Logic programming
- KL-ONE
- Production systems
- Description Logic
- CYC
- Semantic Web
- OWL
- Linked Data
- Wikidata

Where are we, anyway?

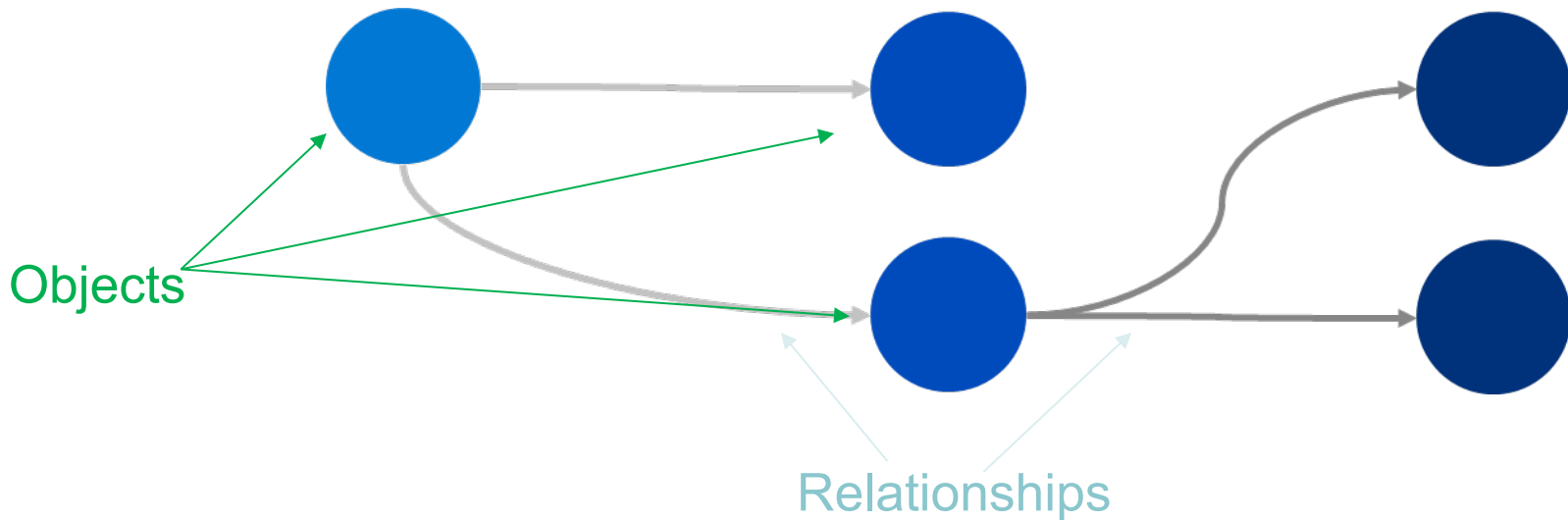
After decades of slow but steady advances, knowledge representation and AI have experienced a sea change that we can attribute to a convergence of multiple factors:

- Ubiquitous Internet and Web for sharing and accessing information
- Data availability, now that it's easy to share
- Increased computing power
- Machine learning advances

How relevant are knowledge graphs today?

Graph Database?

- A graph database represents information as a graph of nodes and edges
- Nodes represent entities (or concepts) and edges represent relations



What's wrong with SQL?

- Since the 1980s, the dominant database paradigm is the relational model that uses tables to represent data
- This is not a good fit for many kinds of knowledge
 - Wikidata has about 10,000 kinds of relations for it's ~100M entities
 - The Amazon Product knowledge graph has many possible properties for it's many products
- Using a table model would require many sparse tables

Who invented the Web?

Who invented the Web?



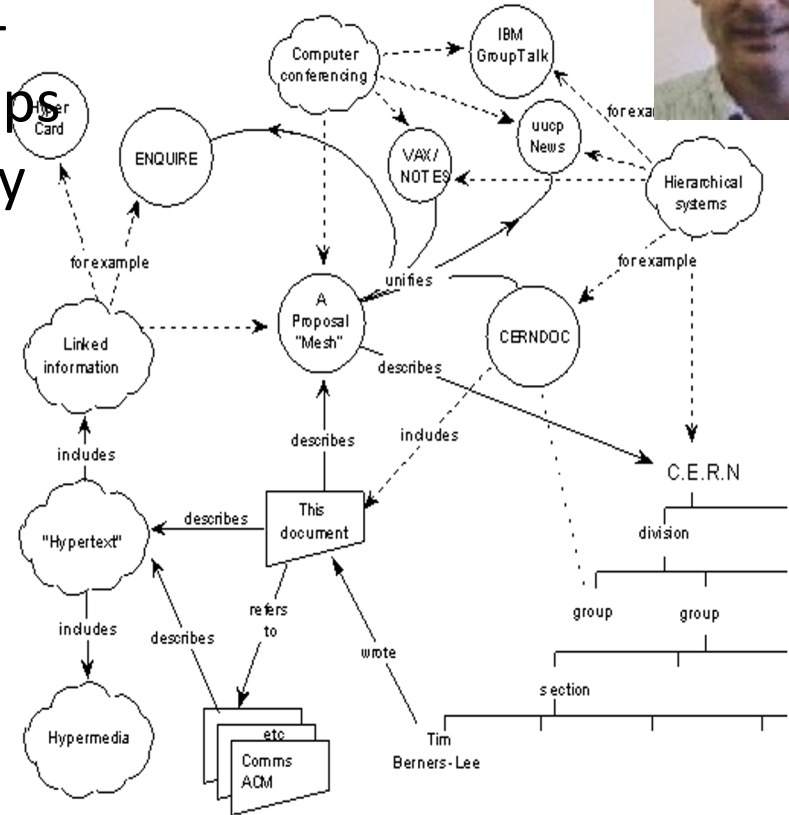
Semantic Web Origin



Tim Berners-Lee's original 1989 proposal described a web of relationships among named objects unifying many information management tasks

Capsule history

- Guha's MCF (~94)
- XML+MCF=>RDF (~96)
- RDF+OO=>RDFS (~99)
- RDFS+KR=>DAML+OIL (00)
- W3C's SW activity (01)
- W3C's OWL (03)
- ...



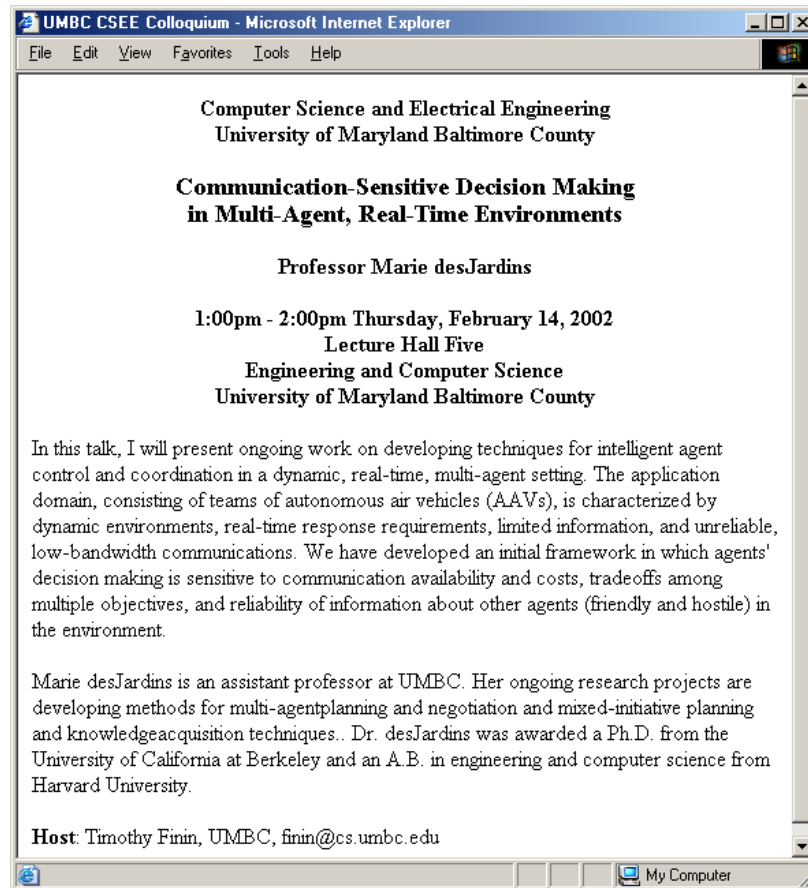
W3C's Semantic Web Goals

Focus on machine consumption:

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."

-- Berners-Lee, Hendler and Lassila, [The Semantic Web](#), Scientific American, 2001

Why is this hard?



UMBC CSEE Colloquium - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Computer Science and Electrical Engineering
University of Maryland Baltimore County

Communication-Sensitive Decision Making
in Multi-Agent, Real-Time Environments

Professor Marie desJardins

1:00pm - 2:00pm Thursday, February 14, 2002
Lecture Hall Five
Engineering and Computer Science
University of Maryland Baltimore County

In this talk, I will present ongoing work on developing techniques for intelligent agent control and coordination in a dynamic, real-time, multi-agent setting. The application domain, consisting of teams of autonomous air vehicles (AAVs), is characterized by dynamic environments, real-time response requirements, limited information, and unreliable, low-bandwidth communications. We have developed an initial framework in which agents' decision making is sensitive to communication availability and costs, tradeoffs among multiple objectives, and reliability of information about other agents (friendly and hostile) in the environment.

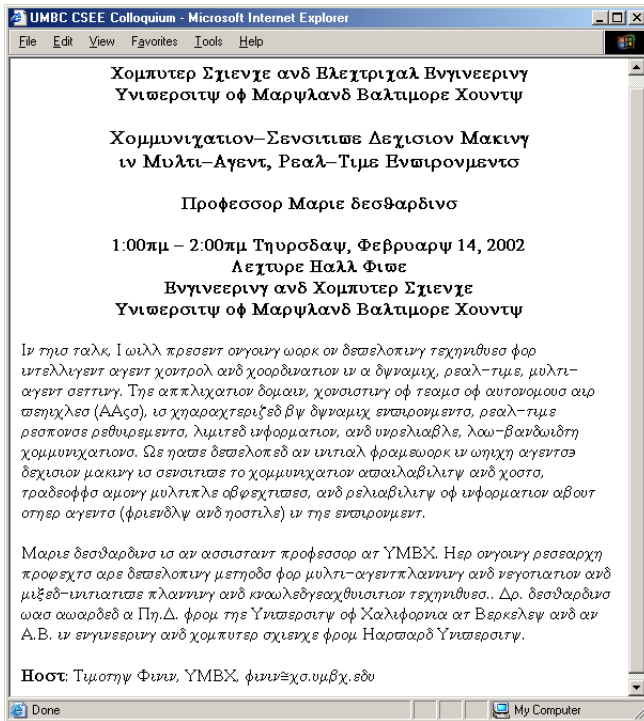
Marie desJardins is an assistant professor at UMBC. Her ongoing research projects are developing methods for multi-agent planning and negotiation and mixed-initiative planning and knowledge acquisition techniques. Dr. desJardins was awarded a Ph.D. from the University of California at Berkeley and an A.B. in engineering and computer science from Harvard University.

Host: Timothy Finin, UMBC, finin@cs.umbc.edu

My Computer

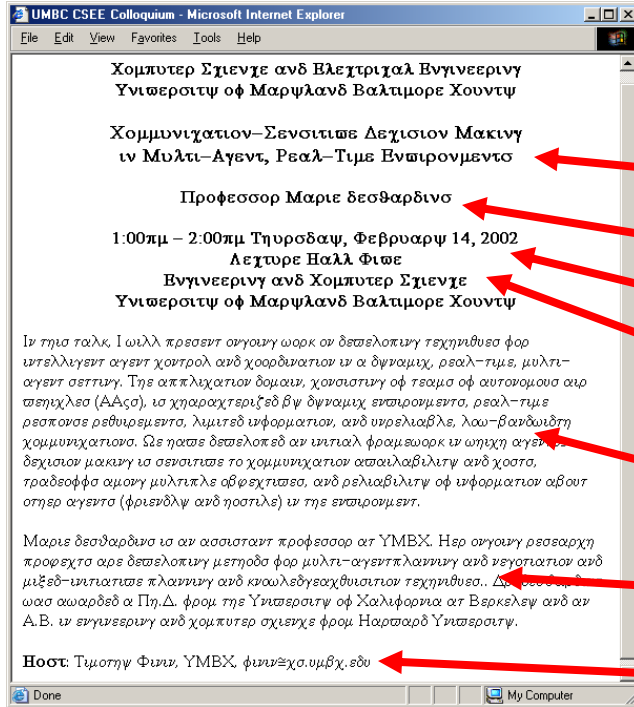
after Frank van Harmelen
and Jim Hendler

What this looks like to a machine...



after Frank van Harmelen
and Jim Hendler

OK, so HTML is not helpful



Maybe we can tell the machine what the different parts of the text represent?

title

speaker

time

location

abstract

biosketch

host

XML to the rescue?

UMBC CSEE Colloquium - Microsoft Internet Explorer

Χομπυτερ Σχιενιχε ανδ Ηλεκτριχαλ Ενγινεερινγ
Υνιωερσιτυ οφ Μαρψλανδ Βαλτιμορε Χουντυ

<title> Χομμυνηχατιον-Σενσιτιβε Δεχισιον Μακινγ **</title>**
ιν Μυλτι-Αγεντ, Ρεαλ-Τιμε Ενβαριονμεντο

<speaker> Προφεσορ Μαριε δεσθαριντο **</speaker>**

<time> 1:00πμ - 2:00πμ Τηυροδαψ, Φεβρουαρη 14, 2002 **</time>**
Λεχτυρε Χαλλ Φιωε **</location>**
Ενγινεερινγ ανδ Χομπυτερ Σχιενιχε **</location>**
Υνιωερσιτυ οφ Μαρψλανδ Βαλτιμορε Χουντυ

<abstract> Ιν τηισ ταλλκ Ι ωιλλ πρεσεντ ονγουνγ ωορκ ον δεσελοπιγγ τεχηνηθεσ φορ
μυλτι-αγεντ χοντρολ ανδ χοορδινατιον ιν α δυναμιχ, ρεαλ-τιμε, μυλτι-
αγεντ σετινγ. Τηε αππλιχατιον δομαιν, χονοιστιγγ οφ τεαμο οφ αυτονομουσ αιω
σπειχλεσ (ΑΑςσ), ισ χηαροχαχτεριζεδ βγ δυναμιχ ενβαριονμεντα, ρεαλ-τιμε
ρεσπονσε ρεθιουρμεντα, λιμιτεδ ινφορματιον, ανδ υνρελιαβλε, λω-βανδωιδθη
χομμυνηχατιονα. Ωε ηαστε δεσελοπεδ αν ιντιαλ φορμωορκ ιν ωηιχη αγεντασ
δεχισιον μακινγ ισ σενσιτιβε το χομμυνηχατιον ασαιλαβιλιγγ ανδ χαστο,
προδεοφοσ αμογγ μυλτιπλεσ οβφεχτιωεσ, ανδ ρελιαβιλιγγ οφ ινφορματιον αβουτ
στηρε αγεντα (φοιενδλγ ανδ ηοστιλε) ιν τηε ενβαριονμεντ. **</abstract>**

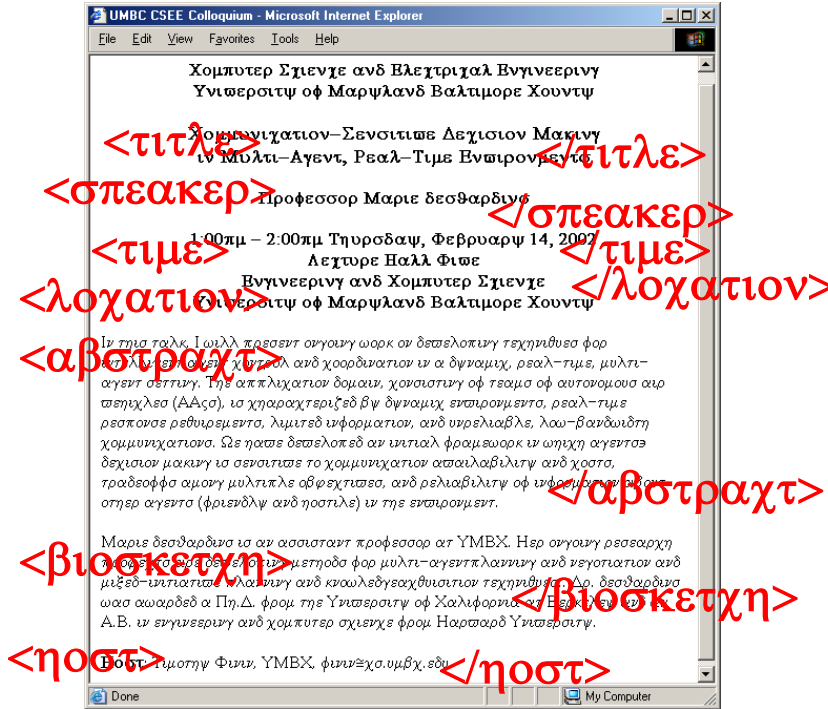
<biosketch> Μαριε δεσθαριντο ισ αν ασισταντ προφεσορ ατ ΥΜΒΧ. Ηερ ονγουνγ ρεσαρχη
εστιν ενασ δεσελοπιγγ μετηοδ φορ μυλτι-αγεντπλανινγγ ανδ νεγοτιατιον ανδ
μιξεδ-ιντιατιωεσ πλανινγγ ανδ κνωλεδωεσραχθισιτιον τεχηνηθεσ.. Δο, δεσθαριντο
ωασ ακωαρδεδ α Πη.Δ. φορμ τηε Υνιωερσιτυ οφ Χαλιφορνια οφ Βερμολεγγ ανδ ατ
Α.Β. ιν ενγινεερινγ ανδ χομπυτερ σχιενιχε φορμ Ηαρσααρδ Υνιωερσιτυ.

<host> Ηοστ: Τιμοτην Φινν, ΥΜΒΧ, φινηνχσ.υμβχ.εδυ **</host>**

XML fans propose creating a XML tag set to use for each application.

For talks, we can choose **<title>**, **<speaker>**, etc.

XML ≠ machine accessible meaning



But, to your machine, the tags still look like this....

The tag names carry no meaning.

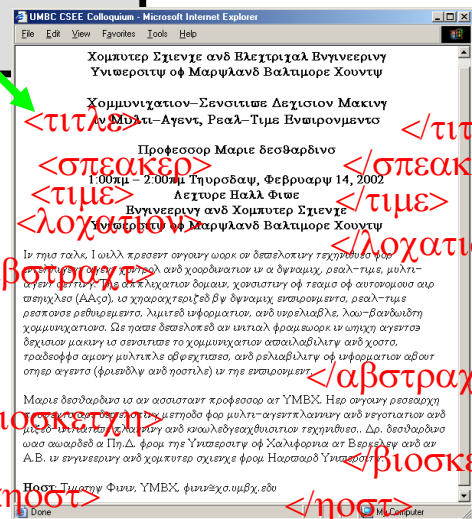
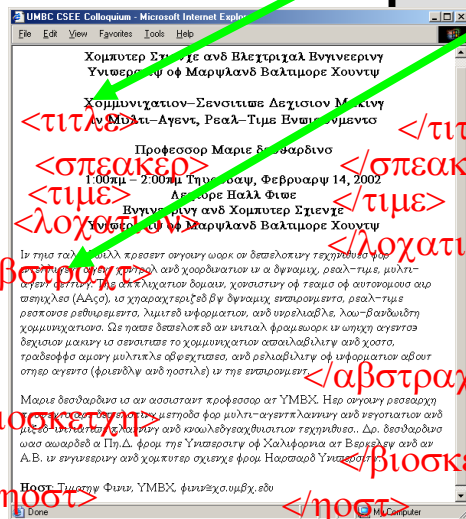
XML DTDs and Schemas have little or no semantics.

XML Schema helps

XML Schema file

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" >
  <xs:element name="book">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="title" type="xs:string"/>
        <xs:element name="author" type="xs:string"/>
        <xs:element name="character" minOccurs="0" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="name" type="xs:string"/>
              <xs:element name="friend-of" type="xs:string" minOccurs="0"
                maxOccurs="unbounded"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        <xs:element name="price" type="xs.date"/>
        <xs:element name="qualification" type="xs:string"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

XML Schemas provide a simple mechanism to define shared vocabularies.



But there are many schemas

XML Schema file 1

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="book">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="title" type="xs:string"/>
        <xs:element name="author" type="xs:string"/>
        <xs:element name="character" minOccurs="0" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="name" type="xs:string"/>
              <xs:element name="friend-of" type="xs:string" minOccurs="0"
                maxOccurs="unbounded"/>
              <xs:element name="date" type="xs:date"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

XML Schema file 2

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="book">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="title" type="xs:string"/>
        <xs:element name="author" type="xs:string"/>
        <xs:element name="character" minOccurs="0" maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="name" type="xs:string"/>
              <xs:element name="friend-of" type="xs:string" minOccurs="0"
                maxOccurs="unbounded"/>
              <xs:element name="since" type="xs:date"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Χομπίουτερ Στιενγε ανδ Βλεχτριχαλ Βηγγερευν
 Υνιερστυν οφ Μαριφλανδ Βαλτιμορε Χουτυν

Χομμουναχτιον-Σενσουατσε Δεχισιον Μακιτυν
 Ρεαλ-Τυμε Βηπυρομμεντο

Προφασσορ Μαριε δεσθαριδνο

14, 2002μ. ηηροδαυ, Φεβρουαρη 14, 2002
 Λεχτυρε Βαλλ Φιυε

Βηγγερευν ανδ Χομπιουτερ Στιενγε
 Υνιερστυν οφ Μαριφλανδ Βαλτιμορε Χουτυν

Ιν ηηια τοκλ, Ι ωλλ κρεσεντ οργου ωοκ αν δεσπελοτυνη τεχνητυνη φη
 ανδ χροσδινατυν ιν α θηνατυ, ρεαλ-τυμε, μωλτι-
 αγγετυκλατυνη δυατυ, χροστυνη οφ τωαο μφ αυτουνοτυο ανρ
 απηχλαο (Αηχο), ιο χροσαχτυρετυδ βηθ θηνατυ, ενωρτυομεντο, ρεαλ-τυμε
 ρεστυοσ οφ ενωρτυομεντο, λυμτυδ ινφορτυοτυν, ανδ ινρελυαβηλε, λωυ-βαδωδωτη
 χρομμουναχτυοτυο. Ωε ηηασε δεσπελοτυδ αν ιντυαλ φορμωωοκ ιν ιντυνη αγεντυσ
 δεχισιον μωλτι ιν σεντυατυσε το χρομμουναχτυον απωλυαβητυνη ανδ χροστυ,
 τραδωσφο αμωρτυ μωλτυκλε ωβρεχτυατυ, ανδ ρελυαβητυνη οφ ινφορτυοτυν αβωτ
 στυρη αγεντυ (φωρετυδ ανδ ηωστυλε) ιν ηηε ενωρτυομεντυ.

Μαριε δεσθαριδνο ιν αν απωστυατυν προφασσορ ατυ ΥΜΒΧ. Ηεο οργουνη ρεσωσχυ
 μεθωδο φωρ μωλτι-αγγετυκλατυνη ανδ ενωρτυοτυν ανδ
 μωλτι-αγγετυκλατυνη ανδ κωυλεδωγεαχτυοτυνη τεχνητυνη. Δω, δεσθαριδνο
 ωωω απωρωδω αν Πη.Δ. φωρ ηηε Υνιερστυν οφ Χαλιφορνια ατ Βερκελυ ανδ αν
 Α.Β. ιν ενηγγερευν ανδ χρομπιουτερ στυνηχε φοορ Ηωρσωαδ Υνιερστυν.

Πωστ Τυποτυη Φιτυν, ΥΜΒΧ, φηινεχσο.ομβχ, εδω

Host

Χομπίουτερ Στιενγε ανδ Βλεχτριχαλ Βηγγερευν
 Υνιερστυν οφ Μαριφλανδ Βαλτιμορε Χουτυν

Χομμουναχτιον-Σενσουατσε Δεχισιον Μακιτυν
 Ρεαλ-Τυμε Βηπυρομμεντο

Προφασσορ Μαριε δεσθαριδνο

14, 2002μ. ηηροδαυ, Φεβρουαρη 14, 2002
 Λεχτυρε Βαλλ Φιυε

Βηγγερευν ανδ Χομπιουτερ Στιενγε
 Υνιερστυν οφ Μαριφλανδ Βαλτιμορε Χουτυν

Ιν ηηια τοκλ, Ι ωλλ κρεσεντ οργου ωοκ αν δεσπελοτυνη τεχνητυνη φη
 ανδ χροσδινατυν ιν α θηνατυ, ρεαλ-τυμε, μωλτι-
 αγγετυκλατυνη δυατυ, χροστυνη οφ τωαο μφ αυτουνοτυο ανρ
 απηχλαο (Αηχο), ιο χροσαχτυρετυδ βηθ θηνατυ, ενωρτυομεντο, ρεαλ-τυμε
 ρεστυοσ οφ ενωρτυομεντο, λυμτυδ ινφορτυοτυν, ανδ ινρελυαβηλε, λωυ-βαδωδωτη
 χρομμουναχτυοτυο. Ωε ηηασε δεσπελοτυδ αν ιντυαλ φορμωωοκ ιν ιντυνη αγεντυσ
 δεχισιον μωλτι ιν σεντυατυσε το χρομμουναχτυον απωλυαβητυνη ανδ χροστυ,
 τραδωσφο αμωρτυ μωλτυκλε ωβρεχτυατυ, ανδ ρελυαβητυνη οφ ινφορτυοτυν αβωτ
 στυρη αγεντυ (φωρετυδ ανδ ηωστυλε) ιν ηηε ενωρτυομεντυ.

Μαριε δεσθαριδνο ιν αν απωστυατυν προφασσορ ατυ ΥΜΒΧ. Ηεο οργουνη ρεσωσχυ
 μεθωδο φωρ μωλτι-αγγετυκλατυνη ανδ ενωρτυοτυν ανδ
 μωλτι-αγγετυκλατυνη ανδ κωυλεδωγεαχτυοτυνη τεχνητυνη. Δω, δεσθαριδνο
 ωωω απωρωδω αν Πη.Δ. φωρ ηηε Υνιερστυν οφ Χαλιφορνια ατ Βερκελυ ανδ αν
 Α.Β. ιν ενηγγερευν ανδ χρομπιουτερ στυνηχε φοορ Ηωρσωαδ Υνιερστυν.

Πωστ Τυποτυη Φιτυν, ΥΜΒΧ, φηινεχσο.ομβχ, εδω

Host

There's no way to relate schema

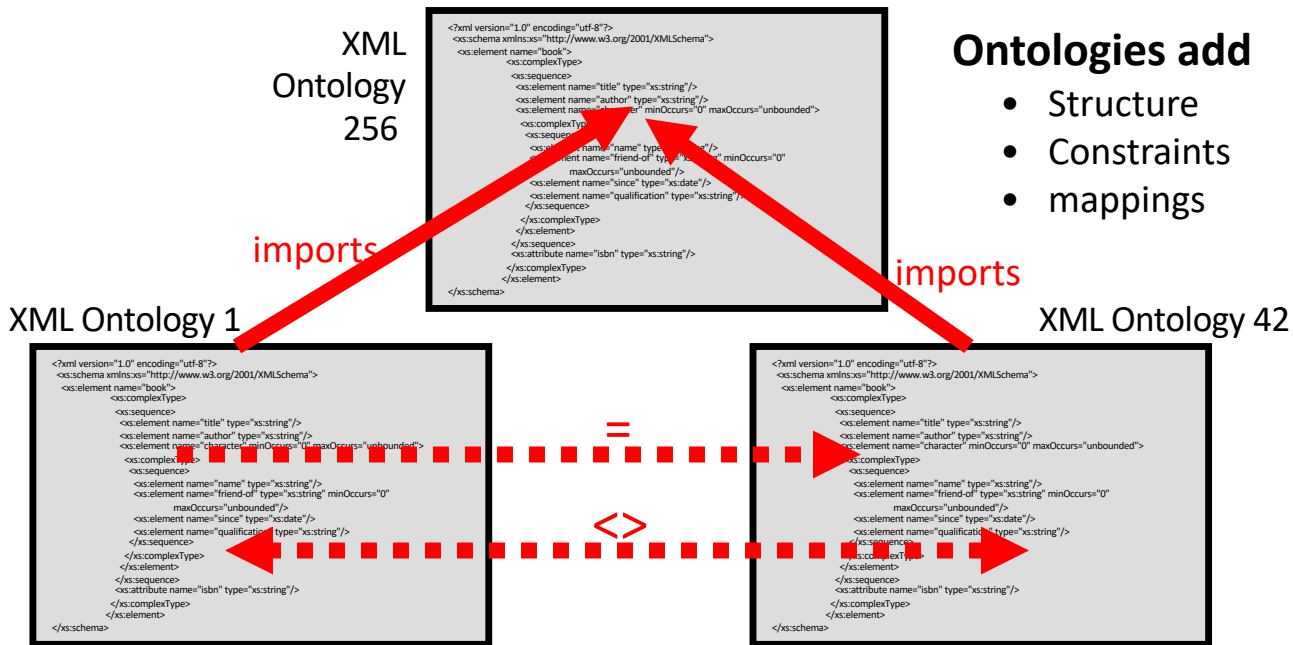
The diagram illustrates the difficulty of relating XML schemas. At the top, two XML Schema files are shown side-by-side, labeled "XML Schema file 1" and "XML Schema file 42". Red 'X' marks are placed between them, indicating a lack of relationship. Below each schema file is a browser window displaying XML content. Red annotations map schema elements to content elements:

- `<title>` maps to `<title>`
- `<speaker>` maps to `<speaker>`
- `<time>` maps to `<time>`
- `<location>` maps to `<location>`
- `<abstract>` maps to `<abstract>`
- `<biosketch>` maps to `<biosketch>`
- `<host>` maps to `</host>`

At the bottom, a blue box contains the text: "Either manually or automatically. XML Schema is weak on semantics."

Either manually or automatically.
XML Schema is weak on semantics.

An Ontology level is needed



We need a way to define ontologies in XML

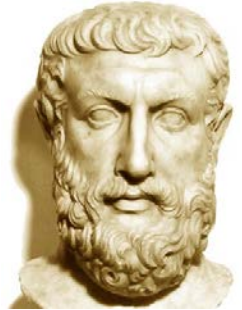
So we can relate them

So machines can understand (to some degree) their meaning

What is an ontology (1)

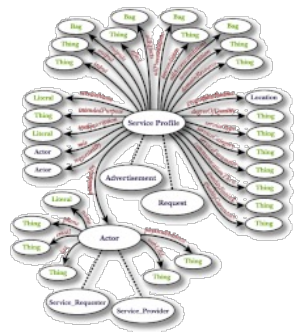
- In philosophy

- A branch of philosophy that studies concepts such as existence, being, becoming, and reality
- Specifies how entities are grouped into basic categories and the entities that exist on the most fundamental level



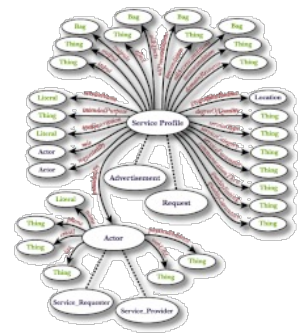
- In information systems

- Specifies the representation, formal naming, and definition of categories, properties, and relations between the concepts, data, and entities in a domains of discourse.



What is an ontology (2)

- Creating ontologies for computing systems is not unlike designing object oriented systems or databases
- But is typically declarative, i.e.,
 - Grounded in logic
 - Without procedural elements
 - Intended to support information sharing



Semantic Web

Use Semantic Web Technology
to publish shared data &
knowledge

Semantic web technologies
allow machines to share data
and knowledge using common
web language and protocols.

[Weaving the Web](#), TBL, 1999

The node in the center is DBpedia

 Browse using ▾ Formats ▾ Faceted Browser Sparql Endpoint

About: [University of Maryland, Baltimore County](http://dbpedia.org/page/University_of_Maryland,_Baltimore_County)

An Entity of Type: [university](#), from Named Graph: <http://dbpedia.org>, within Data Space: [dbpedia.org](#)

The University of Maryland, Baltimore County (UMBC) is a public research university in Baltimore County, Maryland. It has a fall 2020 enrollment of 13,497 students, 61 undergraduate majors, over 92 graduate programs (38 master, 25 doctoral, and 29 graduate certificate programs) and the first university research park in Maryland. It is classified among "R1: Doctoral Universities – Very High Research Activity". 

Property	Value
dbo:abstract	<ul style="list-style-type: none">The University of Maryland, Baltimore County (UMBC) is a public research university in Baltimore County, Maryland. It has a fall 2020 enrollment of 13,497 students, 61 undergraduate majors, over 92 graduate programs (38 master, 25 doctoral, and 29 graduate certificate programs) and the first university research park in Maryland. It is classified among "R1: Doctoral Universities – Very High Research Activity". Established as a part of the University System of Maryland in 1966, the university became the first public college or university in Maryland to be inclusive of all races. UMBC has the fourth

http://dbpedia.org/page/University_of_Maryland,_Baltimore_County

Semantic Web => Linked Open Data

Use Semantic Web Technology
to publish shared data &
knowledge



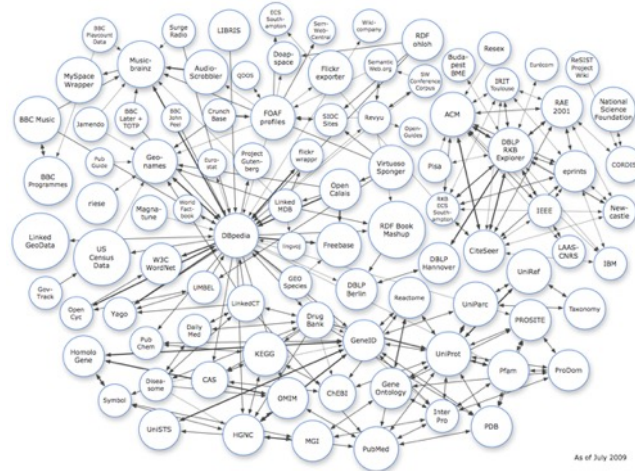
2008

Data is inter-
linked to support inte-
gration and fusion of knowledge

LOD growing

Semantic Web => Linked Open Data

Use Semantic Web Technology
to publish shared data &
knowledge



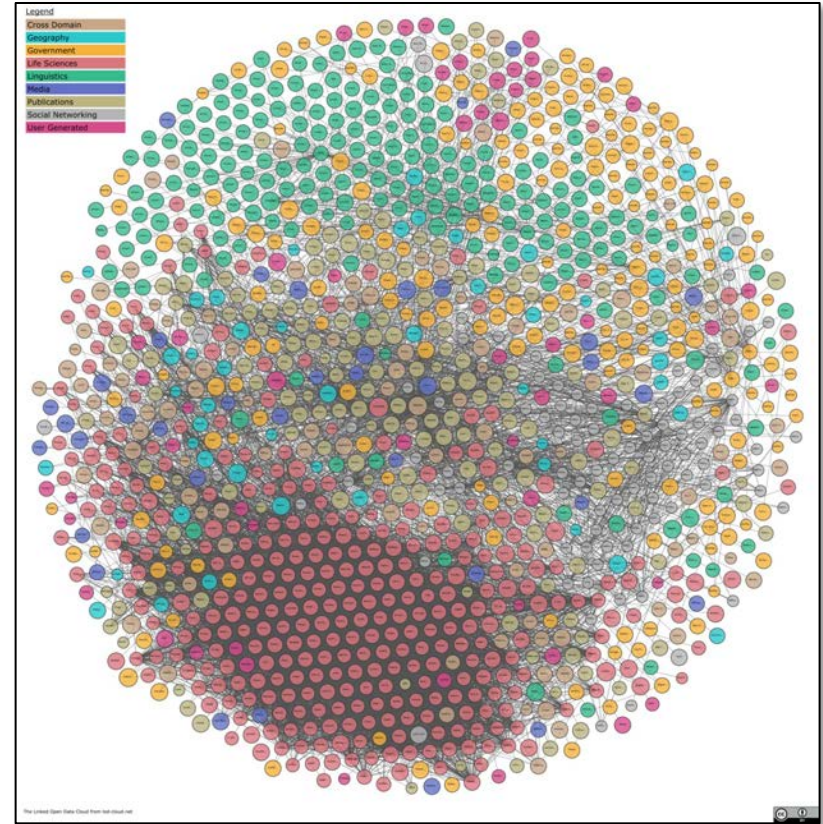
2009

Data is inter-
linked to support inte-
gration and fusion of knowledge

... and growing

Linked Open Data Cloud 2020

- The [LOD cloud](#) currently contains **1,255** datasets with **16,174** links (as of May 2020)
- [Wikidata](#) is another good example of an important background knowledge graph
- It has a billion facts about 100M entities supported by a large ontology
- Links to underlying Wikipedia sources in many languages



Data sharing requires standards

- Standards can be set in many ways
 - A company: Java is owned by Oracle
 - A non-profit organization: IEEE manages many standards
 - A community group: Python (but with Guido van Rossum as benevolent dictator for life for most of its history)
- W3C has defined **most** semantic web standards
 - Some failed to reach consensus, but are widely used (e.g., SWRL: [Semantic Web Rule Language](#))
 - Others were developed by key parties, e.g., [Schema.org](#)

Semantic Web

- The [semantic web](#) rests on a set of standards, most managed by the [World Wide Web Consortium](#) (W3C)
 - TBL is director of the WTC and a professor at MIT & Oxford
- The W3C has over [450 members](#), which include many companies and non-profit organizations
- It defined many semantic web standards including [RDF](#), [OWL](#), [SPARQL](#), [SHACL](#), [PROV](#), and [SKOS](#)
- **RDF** is the most basic and we'll look first at that

Semantic Web: 1, 2, 3

Languages typically divided into three parts:

- 1. Syntax:** legal forms that make up the sentences in a language
- 2. Semantics:** mapping of sentences to meaning (perhaps truth theoretic)
- 3. Pragmatics:** everything else (how to do things with language, knowledge of world, etc.)

1: Syntax

- **URIs** denote classes, properties, objects, relations
 - http://live.dbpedia.org/resource/Alan_Turing
 - <http://schema.org/Person>
 - <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
 - Use strings for literals
 - Use **triples** to make statements
 - dbpedia:Alan_Turing rdfs:type schema:Person .
 - “Alan Turing is a Person”
- *URI = [Uniform Resource Identifier](#)

2: Semantics

- Semantics maps URIs to the things they denote in “the world”
- Some of this is in your mind or in how you write your program
- Meaning of some URIs allow **inference**
 - **parentOf** relation is **inverse** of the **childOf** relation
 - schema:parentOf owl:inverse schema:childOf

3: Pragmatics

- Semantics is more than just about truth (statements that assert things)
- Must account for commands, requests, questions, context, etc.
 - Some handled by Web protocols (GET, POST)
 - Some by special protocols (e.g., SPARQL queries)
 - Some by having reference KBs of the world (e.g., Wikidata) to help identify common entities

Where are we

- The W3C version of the open semantic web has grown over the past 20+ years
- Languages and standards are being used, e.g.,
 - BBC uses RDF to make up much of its content online
 - Google and Facebook detect AND MAKE USE OF (some) RDF embedded in html pages
 - Google, Yahoo, Microsoft and Yandex formed Schema.org to develop useful vocabularies
 - Data.gov has many datasets in RDF

Wikipedia data
in RDF

DBpedia

dbpedia.org/page/Alan_Turing

dbpedia-owl:almaMater	<ul style="list-style-type: none">dbpedia:King's_College,_Cambridgedbpedia:Princeton_University
dbpedia-owl:award	<ul style="list-style-type: none">dbpedia:Royal_Societydbpedia:Order_of_the_British_Empiredbpedia:Fellow_of_the_Royal_Societydbpedia:Officer_of_the_Order_of_the_British_Empire
dbpedia-owl:birthDate	<ul style="list-style-type: none">1912-06-23 (xsd:date)1912-06-23 (xsd:date)
dbpedia-owl:birthName	<ul style="list-style-type: none">Alan Mathison Turing
dbpedia-owl:birthPlace	<ul style="list-style-type: none">dbpedia:Paddingtondbpedia:Maida_Vale
dbpedia-owl:deathDate	<ul style="list-style-type: none">1954-06-07 (xsd:date)
dbpedia-owl:deathPlace	<ul style="list-style-type: none">dbpedia:Wilmslow
dbpedia-owl:doctoralAdvisor	<ul style="list-style-type: none">dbpedia:Alonzo_Church
dbpedia-owl:doctoralStudent	<ul style="list-style-type: none">dbpedia:Robin_Gandy
dbpedia-owl:field	<ul style="list-style-type: none">dbpedia:Computer_sciencedbpedia:Mathematicsdbpedia:Cryptanalysis
dbpedia-owl:individualisedPnd	<ul style="list-style-type: none">118802976
dbpedia-owl:knownFor	<ul style="list-style-type: none">dbpedia:Turing_machinedbpedia:Cryptanalysis_of_the_Enigmadbpedia:Automatic_Computing_Enginedbpedia:Turing_test

dbpedia:Alan_Turing dbpedia-owl:doctoralAdvisor dbpedia:Alonzo_Church .

Freebase

Acquired by
Google in 2010

The screenshot shows the Freebase website interface for the entry 'Alan Turing'. The browser address bar shows 'www.freebase.com/view/en/alan_turing'. The page features a navigation bar with 'Data', 'Schema', 'Apps', and 'Docs' tabs. A left sidebar lists various categories like 'People', 'Literature Subject', and 'Computers'. The main content area includes a portrait of Alan Turing, a biographical paragraph, and a list of key facts such as 'Date of birth: Jun 23, 1912' and 'Profession: Mathematician, Philosopher, Computer Scientist, Logician'. A 'Related Topics' section on the right lists other individuals like Edward Thomas Hall and William Kingdon Clifford. At the bottom, there are sections for 'People' and 'Alan Turing Quotes', with a notable quote: '« Mathematical reasoning may be regarded... »' and another: '« No, I'm not interested in developing a powerful brain... »'.

Alan Turing

Scroll to:





- People
- Literature Subject
- Influence Node
- Computers
- Name source
- Academic
- Chivalric Order Member
- Inventor
- Author
- Person Or Being In Fiction
- Organization member
- Film subject
- TV subject
- More...

These people have edited this topic:



[Edit this topic](#)

Last edited Dec 12, 2012 See all topic history »

Related Topics

-  [Edward Thomas Hall](#)
-  [Alan Turing](#)
-  [William Kingdon Clifford](#)
-  [Ada Lovelace](#)

Alan Turing elsewhere on the web

-  [Official Website](#)
-  [Wikipedia](#)

People

Place of birth: [Maida Vale, United Kingdom](#)

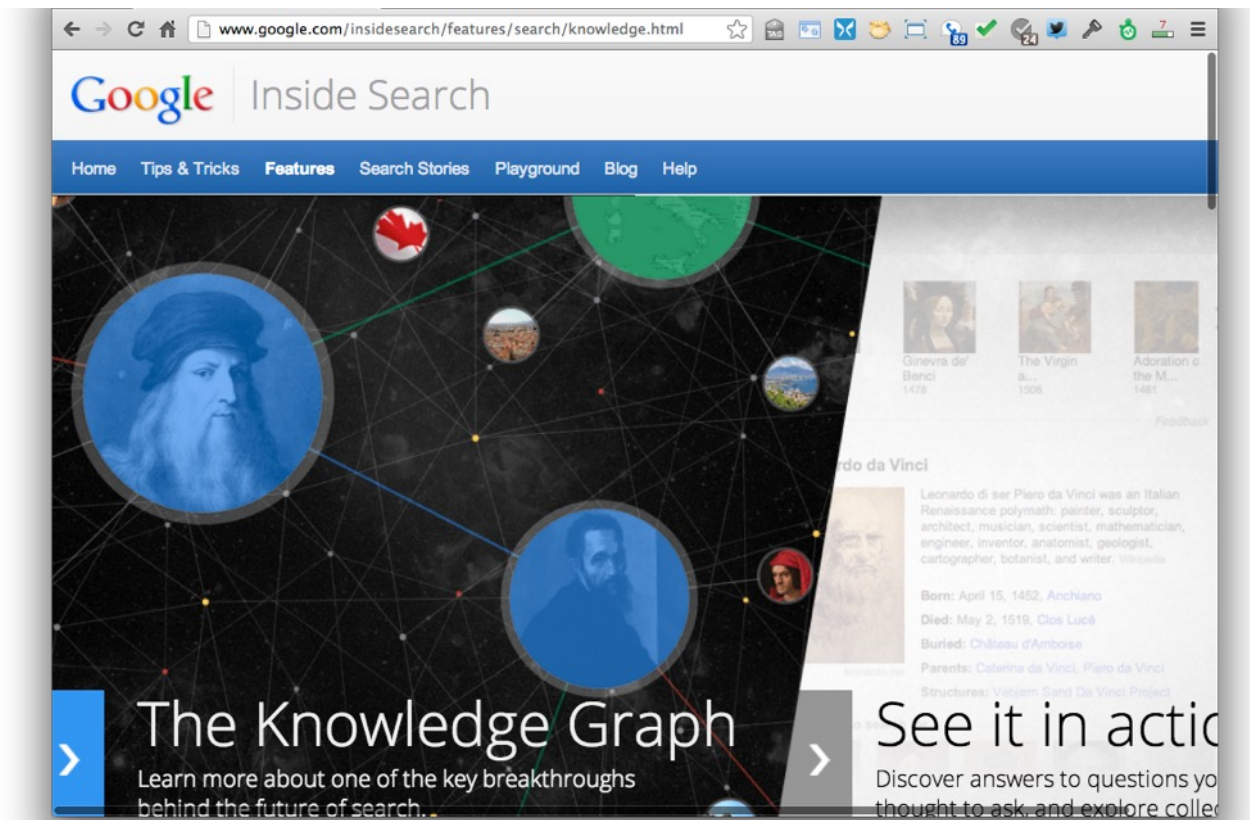
Alan Turing Quotes

« Mathematical reasoning may be regarded... »

« No, I'm not interested in developing a powerful brain... »

“An entity graph of people, places and things, built by a community that loves open data”

Google Knowledge Graph



The screenshot shows the Google Knowledge Graph interface. At the top, the Google logo and "Inside Search" are visible. Below the navigation bar, a large network graph is displayed with various nodes connected by lines. A prominent node is a portrait of Leonardo da Vinci. To the right, a detailed profile for Leonardo da Vinci is shown, including his birth and death dates, parents, and a list of structures he designed. The profile text reads: "Leonardo di ser Piero da Vinci was an Italian Renaissance polymath, painter, sculptor, architect, musician, scientist, mathematician, engineer, inventor, anatomist, geologist, cartographer, botanist, and writer." Below the profile, there are three small thumbnail images of artworks: "Ginevra de' Benci" (1478), "The Virgin & Child" (1506), and "Adoration of the M..." (1481). At the bottom, there are two call-to-action buttons: "The Knowledge Graph" and "See it in action".

Google Inside Search

Home Tips & Tricks **Features** Search Stories Playground Blog Help

The Knowledge Graph

Learn more about one of the key breakthroughs behind the future of search.

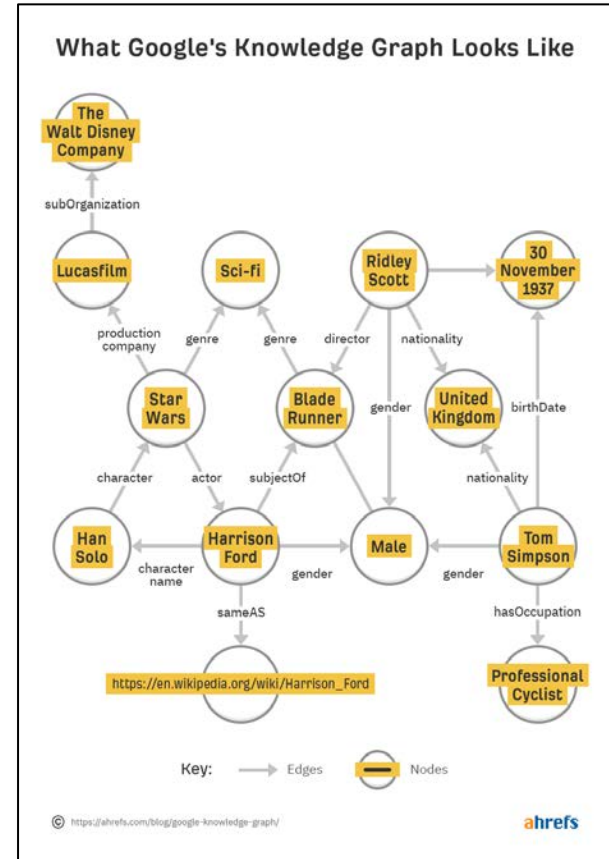
See it in action

Discover answers to questions you thought to ask, and explore collections

Google's slogan for the knowledge graph: "things, not strings"

SEO & the Google Knowledge Graph

- SEO is Search Engine Optimization
- Companies provide SEO services to make web pages rank more highly
- One technique is to ensure your web page is recognized by search engines as referencing the right entities
- This can include embedding metadata in the web pages
- See [Google's Knowledge Graph Explained: How It Influences SEO](#)



Wikidata

- [Wikidata](#) aims to create an RDF-compatible KG that can be read/edited by humans & machines
 - Wikimedia project started in April 2012
- Wikidata clients use the repository, e.g., to populate Web pages or Wikipedia infoboxes
- Based on ideas from [Semantic MediaWiki](#) and [Freebase](#)

Wikidata

- **Knowledge graph** with ~1B statements about ~100M items
- Fine-grained **ontology** has ~2M types and ~10K properties
- Strings tagged with **language id**
- Entities have a canonical **name**, **aliases** and **description** in multiple languages
- UMBC= [Q735049](#), a *university* with English name *University of Maryland, Baltimore County* and alias *UMBC*



University of Maryland, Baltimore County (Q735049)

public university in Maryland
UMBC

- In more languages Configure

Language	Label	Description	Also known as
English	University of Maryland, Baltimore County	public university in Maryland	UMBC
Spanish	No label defined	No description defined	
Traditional Chinese	馬里蘭大學巴爾的摩縣分校	No description defined	
Chinese	马里兰大学巴尔的摩县分校	No description defined	馬里蘭大學巴爾

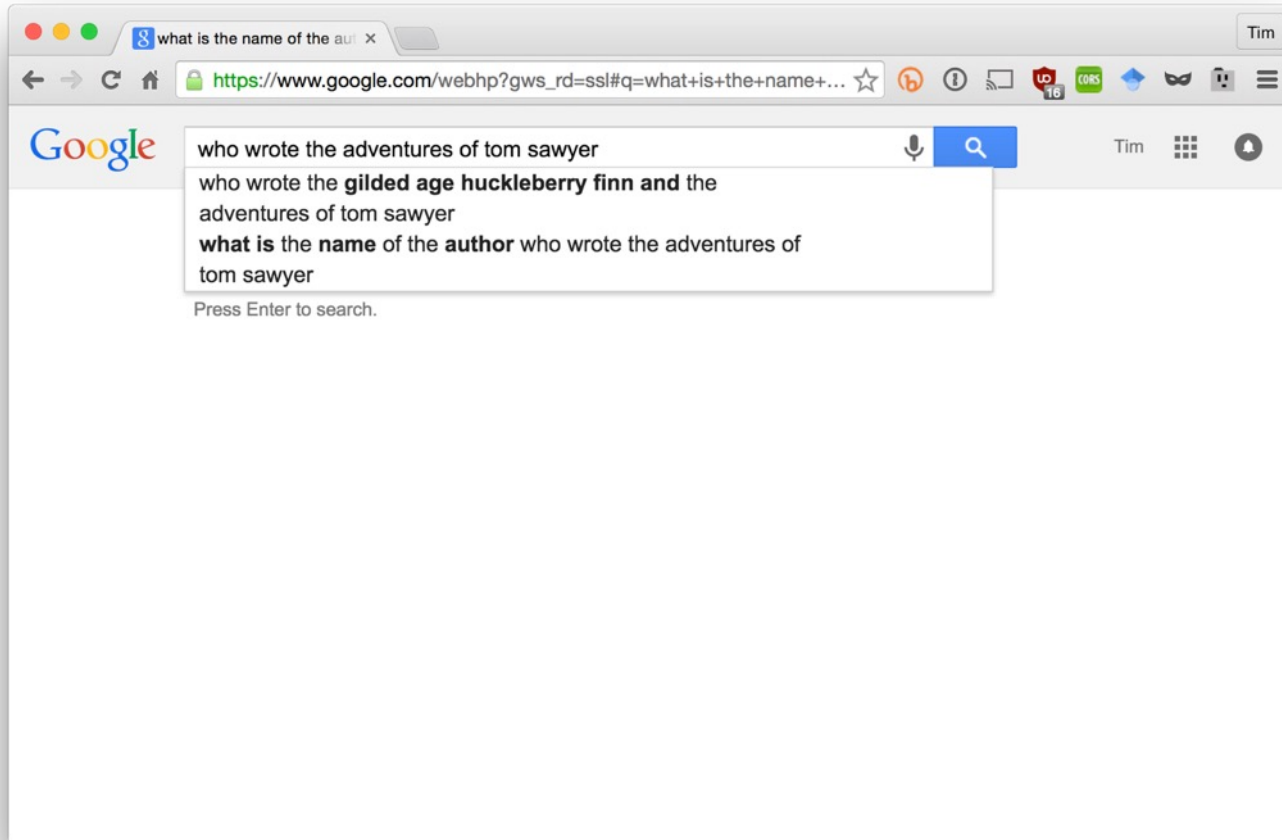
All entered languages

Statements

instance of	<div><div> university edit</div><div> + 1 reference</div></div>
	<div><div> public educational institution of the United States edit</div><div> + 1 reference</div></div>
	<div><div> research university edit</div><div> + 1 reference</div></div>
	<div><div> + add value</div></div>

logo image	<div><div>  edit</div><div>University of Maryland, Baltimore County logo.svg 512 x 118; 7 KB</div><div> + 1 reference</div></div>
	<div><div> + add value</div></div>

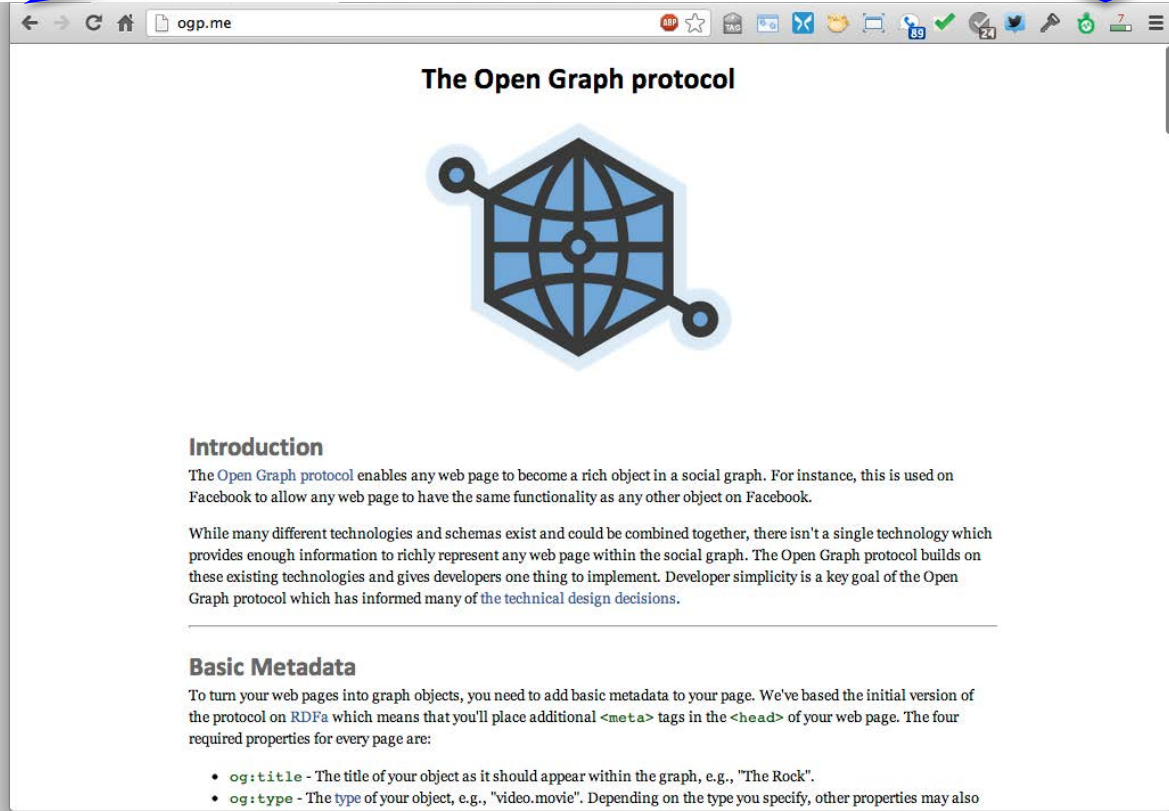
Who wrote Tom Sawyer?



Annotate your web pages in RDFa


Facebook Open Graph

=> object in the FB graph



The screenshot shows a web browser window with the URL ogp.me. The page title is "The Open Graph protocol". Below the title is a logo consisting of a blue hexagon with a globe inside, and two nodes connected to the globe. The page content includes an "Introduction" section and a "Basic Metadata" section.

The Open Graph protocol



Introduction

The Open Graph protocol enables any web page to become a rich object in a social graph. For instance, this is used on Facebook to allow any web page to have the same functionality as any other object on Facebook.

While many different technologies and schemas exist and could be combined together, there isn't a single technology which provides enough information to richly represent any web page within the social graph. The Open Graph protocol builds on these existing technologies and gives developers one thing to implement. Developer simplicity is a key goal of the Open Graph protocol which has informed many of the technical design decisions.

Basic Metadata

To turn your web pages into graph objects, you need to add basic metadata to your page. We've based the initial version of the protocol on RDFa which means that you'll place additional `<meta>` tags in the `<head>` of your web page. The four required properties for every page are:

- `og:title` - The title of your object as it should appear within the graph, e.g., "The Rock".
- `og:type` - The type of your object, e.g., "video.movie". Depending on the type you specify, other properties may also

speech => text =>
entities => task

Apple's SIRI


SIRI engineers
from AI/SW
community

www.apple.com/ios/siri/

Store Mac iPod iPhone iPad iTunes Support

iOS

Overview What's New What is iOS

 **Siri.** Beta

Your wish is its command.

Siri lets you use your voice to send messages, schedule meetings, place phone calls, and more.* Ask Siri to do things just by talking the way you talk. Siri is so easy to use and does so much, you'll keep finding more and more ways to use it.

iPhone: "What's the weather like today?"
The weather's looking good today... up to 81° and partly sunny:
66° H: 81° L: 59°
12:00 PM 75°
1:00 PM 75°
2:00 PM 79°
3:00 PM 79°

iPad: "OK, I can send a text to Cory Quinn for you... what would you like it to say?"
"Be there in 30 minutes"
I updated your message. Ready to send it?
To: Cory Quinn
Be there in 30 minutes

SIRI needs lots of semantic data about entities in the world

A collection of useful ontologies

Schema.org

Embed in HTML using RDFa to make machine understand-able statements

The screenshot shows a web browser window with the URL 'schema.org/Person'. The page has a red header with the 'schema.org' logo, a search bar, and navigation links for 'Home', 'Schemas', and 'Documentation'. The main content area is titled 'Thing > Person' and includes a brief description: 'A person (alive, dead, undead, or fictional)'. Below this is a table listing properties, their expected types, and their descriptions. The table is divided into two sections: 'Properties from Thing' and 'Properties from Person'.

Property	Expected Type	Description
Properties from Thing		
<u>additionalType</u>	URL	An additional type for the item, typically used for adding more specific types from external vocabularies in microdata syntax. This is a relationship between something and a class that the thing is in. In RDFa syntax, it is better to use the native RDFa syntax - the 'typeof' attribute - for multiple types. Schema.org tools may have only weaker understanding of extra types, in particular those defined externally.
<u>alternateName</u>	Text	An alias for the item.
<u>description</u>	Text	A short description of the item.
<u>image</u>	URL	URL of an image of the item.
<u>name</u>	Text	The name of the item.
<u>sameAs</u>	URL	URL of a reference Web page that unambiguously indicates the item's identity. E.g. the URL of the item's Wikipedia page, Freebase page, or official website.
<u>url</u>	URL	URL of the item.
Properties from Person		
<u>additionalName</u>	Text	An additional name for a Person, can be used for a middle name.
<u>address</u>	<u>PostalAddress</u>	Physical address of the item.
<u>affiliation</u>	<u>Organization</u>	An organization that this person is affiliated with. For example, a school/university, a club, or a team.
<u>alumniOf</u>	<u>EducationalOrganization</u>	An educational organizations that the person is an alumni of.

Summary

- Web has made us smarter by sharing information and knowledge as text, audio and images
- Machines should also be able to use the Web to publish & retrieve information & knowledge
- Human forms of knowledge are hard for machines to understand and generate
- The Semantic Web is a collection of languages, ontologies, software tools, services and KBs that are designed to support machines