

The search for meaning on the web

Yorick Wilks

Florida Institute of Human and Machine Cognition
and

Internet Institute, University of Oxford

<http://staffwww.dcs.shef.ac.uk/people/Y.Wilks/>

Gresham College, London, 2019



What the talk is about:

- What is the Semantic Web (SW)?—Tim Berners Lee's original vision---and how does it compare to the WWW he gave us first.
- The SW as a huge philosophical experiment: a web that knows what's on it, and « grounds » meaning.
- The Artificial Companion as an interface to the web: to our whole life's information.
- Identity shielding from the state and corporations.
- Distributing our life information, anonymity and the Companion as an agent that protects and manages our web life.
- Safeguarding that information after death.



Berners-Lee developed the World Wide Web in 1989

- It was carried by the Internet—then a US defence computer network
- It was “serendiptious” : it turned out to be something quite different from what Tim Berners-Lee intended originally
- It was a revolutionary tool for the whole world not just his colleagues
- The WWW was not the *Semantic Web* he intended originally



There was a preweb Internet: specialised groups and data

- ftp: (pre http:)
- Usenet Interest groups –such as *alt.sex.com*--with images
- Downloadable repositories
- Email(first “civilian” email 1986?)
- Mailing out book length files (Minsky 1970)
- Early blogging and diaries
- Usenet had no central servers---it anticipated peer-to-peer information transfer
- But all this was “owned” by defense-related US computer scientists



Berners-Lee on the origins of the WWW

- *...in 1989, while working at the European Particle Physics Laboratory, I proposed that a global hypertext space be created in which any network-accessible information could be referred to by a single "Universal Document Identifier".*
- *I wrote in 1990 a program called "WorldWideWeb", a point and click hypertext editor which ran on the "NeXT" machine. This, together with the first Web server, I released to the High Energy Physics community at first, and to the hypertext and NeXT communities in the summer of 1991.*



What did Berners Lee originally want?

- A system of documents readable anywhere
- Accidental features not designed as the basis for universal knowledge and services
- Initially owned not by CS but physicists
- The sex wave can have been no surprise by then (the Minitel experience in France)
- Later attempts to go back to his vision---we shall see below with the Semantic Web



The Semantic Web

(Berners-Lee, Hendler, and Lassila, *Scientific American* 2001)

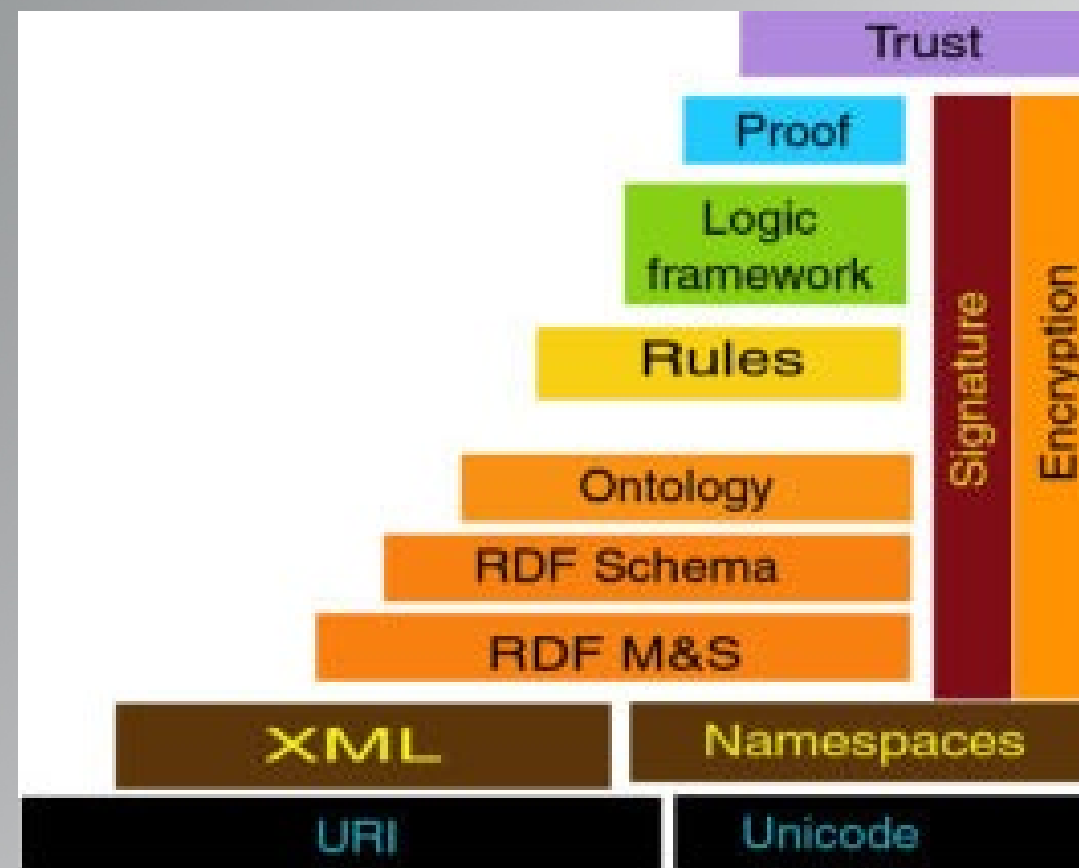
- A vision of making the Internet as readable by computers (agents) as it is by us.
- A similar notion to the “ascent” in the semantic web pyramid---meaning/interpretation somehow trickling UP it from the bottom (cf. Braithwaite’s view of scientific theories--neutrinos linked to experiment)
- Is this last what SW people mean/want, or do they assume that the higher-level structures are self-interpreting?
- The SW as the basis of eScience? Ontologies.



His own later view of his creation

- *“The great need for information about information, to help us categorize, sort, pay for, own information is driving the design of languages for the web designed for processing by machines, rather than people. The web of human-readable document is being merged with a web of machine-understandable data. The potential of the mixture of humans and machines working together and communicating through the web could be immense.”*





15/01/2019

The structure of the Semantic Web



Forms of knowledge in the SW

- 1) Universal Resource Indicators (URIs)
- 2) Resource Description Format (RDF)
- RDF triples---means putting all facts in the form:
John-LOVES-Mary
- Not quite logic yet! Basically output of simple analysis; explicitly called « subject » « object »!
- 3) Ontologies---trees of concepts in hierarchical relations: again like Canary-ISA-Bird
- 4) Plus DAML/OIL reasoning languages



The bottom levels of the SW always sit on NLP *annotations* of objects and actions

- Classic Information Extraction (IE) detects named entities--populates SW's "namespace"
- It does *semantic type annotation*: **kinds** of things marked into text –like humanities scholars.
- Behind it all is an *ontology* (*knowledge of classes of things*)



Press release on Slashdot

- “Reuters just opened access to their corporate semantic technology crown jewels. For free. For anyone. Their Calais API lets you turn unstructured text into a formal RDF graph in about one second. I ran about 5,000 documents through it and played with a subset of them in RDF-Gravity. The results were impressive overall. Is this the start of the semantic web getting real? “



The SW is the WWW of text **plus meanings**:

- 1) An updating of the old AI dream of representing everything in logic for reasoning over the world (GOFAI); actually the SW is much less sophisticated than that--it has traded representation power for tractability.
- 2) It is also the apotheosis of annotation in IE, attempting to build up to concepts in ontologies for e.g. scientific knowledge, by very large shallow computations over texts:
- 3) A system of trusted data bases that ground meanings of those terms in something close to objects (TB-Ls own view?) But are they really OBJECTS?



Putnam's scientists: grounding meaning

Who knows what words **really** mean?

- Aluminium and Molybdenum look the same ----people cannot tell them apart but scientists can.
- Therefore scientists really know the meanings of these terms –which objects they refer to in the world.
- But they should not let this “leak out” to the general populace or the meanings might change.
- Compare: heavy water vs. water
- BUT (Hugh Mellor): we call heavy water “water” because it is water-----and that's simpler and more democratic than the alternative view (Scientists as guardians)
- You cant lock away meanings and keep them safe anywhere (as Wittgenstein might have said)--it's all usage in the end.



Prof. Karen Sparck Jones



A critique of the SW

- Sparck Jones, K. (2004) What's new about the Semantic Web? ACM SIGIR Forum.
- “words stand for themselves”: the basis of successful IR search in WWW and elsewhere.
- Content cannot be recoded in a general way especially if it is general content—Information Retrieval has gained from “decreasing ontological expressiveness”
- Successful Question Answering and Information Extraction technologies are “superficial”
- Note: philosophical critics, like Lewis, want just OBJECTS, whereas Sparck Jones wants just WORDS.
- The SW ends up in data or more words , not ever in actual objects



The problem of Recoding content

- (KSJ auction catalog example)"A Charles II parcel-gilt cagework cup, circa 1670"
- What, she asks, can be recoded beyond some empty piece of logic like {object type: CUP}?
- The rest, she says, IS THE ENGLISH
- Yet, dictionaries explain meaning with more words and people seem happy!
- That doesn't make logicians happy!



Things offered to *ground meaning*

- OBJECTS (Adam and the apple in Genesis)
- ABSTRACT NUMBER POINTS in computer systems #98267KJV
- LOGIC EXPRESSIONS in classic AI
<object type: CUP>
- DATA BASE ITEMS Berners-Lee's SW?
- **MORE WORDS Dictionary definitions**
- SPECIAL CODE WORDS linguistics and some AI



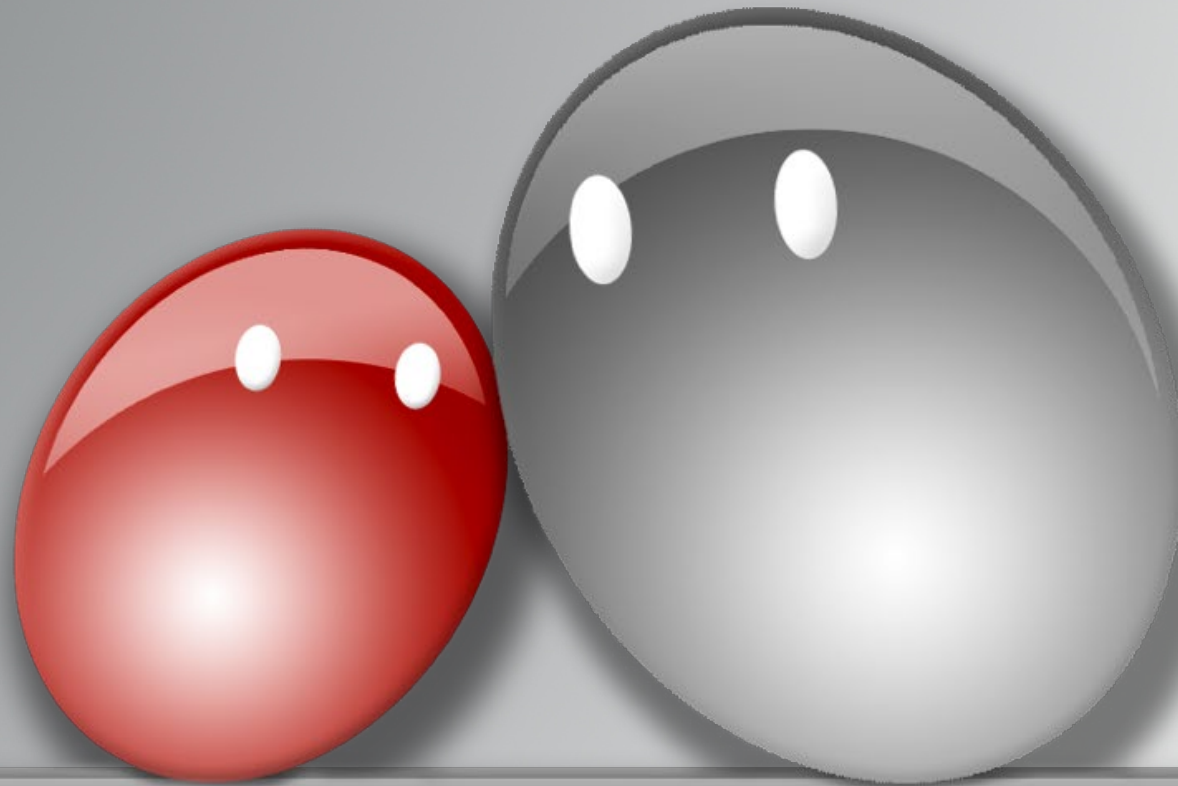
The Semantic Web as a huge philosophical experiment

- Medical and scientific ontologies are now vast and essential to research
- Google's Knowledge Graph behind search
- The WWW grounding issue: bottoming out in books and objects, the URI/URL
- Meaning and whole systems (Quine, Wittgenstein)



COMPANIONS

INTELLIGENT, PERSISTENT, PERSONALISED
MULTIMODAL INTERFACES TO THE INTERNET



COMPANIONS

www.companions-project.org



Your life on the web and a Companion as a life interface

- Memories for Life (M4L) and 28 terabytes
- Companies want access to your life: out of the PC and into Google storage=the Cloud
- It will then no longer matter that you leave your laptop in the taxi, but.....
- How can you control all that material?
- For you, the Companion as an organizer and autobiographer
- But will you trust it? Will the Government or the Company that made it secretly “own” it?



This will affect your own memory processes

- Children no longer bother to remember phone numbers.
- Will you need to remember who anyone is or even your own life?

Jeff Bernard is thinking of writing his autobiography. If anyone can tell him what he was doing between 1960 and 1985 will they please write to him at Box 1234.



Incompetence on the web

- People aren't good at search and need help
- They cant use the facilities that are there
- Even thouh they get enough by searching sex, holidatys, investment and murder
- The Companion as the better Web interface.



The need for a companion- assistant to manage your personal or corporate data

- The web will get harder to manage in some ways not easier, unless we all have access to help
- Help won't be « IT people » (a dying and deskilled breed) but more expert automated assistants and companions who know all about you and your affairs.
- So *much* knowledge, so *much* of your value!



Companions will know all about you

- And that will be basically a good thing....
- But they must be yours--and whoever you leave your Companion TO---not the State's
- Nor than of a company, but you might wish to make concessions there.....my accountant already knows virtually all about me..
- The Companion will have to know what to discuss with whom, which will include other Companions (*Discretion*, again).



The State, as well as individuals, may have problems!

- **The Fading Memory of the State**
- By David Talbot July 2005

Imagine losing all your tax records, your high school and college yearbooks, and your child's baby pictures and videos. Now multiply such a loss across every federal agency storing terabytes of information, much of which must be preserved by law. That's the disaster The National Archives and Records Administration (NARA) is racing to prevent.But NARA has hired two contractors--Harris Corporation and Lockheed Martin--to attempt that miracle.



Companions and the Internet

- A new kind of agent as the answer to a passive web
- The web/internet must become more personal to be tractable, as it gets bigger (and more structured or unstructured?)
- Personal agents will need to be autonomous and trusted (like space craft on missions)
- But also personal and persistent, particularly for large sections of populations now largely excluded from the web.
- Personal knowledge, over a life-time, will be on the web (where else), so Companions must also be web agents.



Here's how simple Companion interfaces look now



ALEXA is like this:

- *User:* Alexa, open Travel Helper.
- *Alexa:* Which country are you going to?
- *User:* France.
- *Alexa:* The average cost of accomodation is one hundred dollars and the average price of a meal for two is fifty dollars.



SIRI (Apple) and ALEXA (Amazon)

- We do not know what technical approaches the leading commercial systems (SIRI and ALEXA) are pursuing
- Probably a mix of handcoded responses (at least the witty ones!) and machine learning—they were initially just handcoded with huge amounts of human work.



Remember how slow AI is not how fast!

- ALEXA is not that different from PARRY 50 years ago, but with an internet connection.
- And has taken 15 years to get onto the market from the research Companions of 10 years ago and still knows much less about you than they did.



Remember ALEXA is always listening

- To be installed by major hotel chains in rooms to replace the concierge
- Is feeding back what you say for advertising purposes
- Precisely captures the Orwell 1984 scenario
- In a decade or two there may be many ALEXAS as people
- But she doesn't know about you at all, only what you say and ask for.



The crucial role of discretion:

- When Companions share with those of others, how will they know what not to say?
- As the Victorians did—or lost their jobs
- Some social sites already allow: “show no postings by Bobbie to Helen”
- Will Companions need such instructions about what other Companions to talk to?



Safeguarding Companions' content

- Remember the discretion issue!
- Who can they talk to; gossip to other people's Companions?
- Protecting content from crime, governments, named-individuals (“NEVER tell my children about my life in Rome”)
- Companions and wills—purges and letter burning (porn and death sites!)



Personal information in a Companion and the state

- Paradox in countries like UK: we have strong constraints on our use of people's information
- Cricket teams and students' marks!
- But the state has total freedom for government scrutiny of us
- « Transparency for them, none for us » !



Identity and its shielding

- The companion as an internet agent (for an incompetent user?) could also shield and negotiate the owner's identity. Power of Attorney.
- UK law, unlike Europe, specifically allows multiple identities and names if not fraudulent.
- Identification minimization (DVLA? WHY must you identify yourself to license your car?)



Identification minimization

- Problem for crime and security with less identification
- Public support for intrusion, but.....
- There will be growing market in identification protection: TOR, Bitcoin and paedophiles, but.....
- **The finger machine on the street** neednt say WHO you are but THAT you are—that you are on the national database of citizens perhaps?



Identification minimization

- In *The Return of Martin Guerre* the wife doesn't care in some sense WHO he is, though she knows who he isn't.
- Are women right to complain that police spies who are lovers lie about their names and occupations?
- Declining importance of names in a WWW world—the web can find you *without your name*.
- Berners-Lee has argued real anonymity impossible in web transactions—your name just one more feature.



Identification minimization

- Deciding how much to reveal and to whom may be the answer.
- Your Companion might do that for you as your attorney, your PR person, and keep track of *who knows what about you on the web*.
- The Companion then becomes your identifier—as the phone is becoming now---it says where you are as all TV dramas know and as the bank puts you and you card in the same place. Biology or Finger print not name!



Identification Minimization

- Giving legal incentives to ISPs that require LESS identification.
- Berners-Lee is now exploring these ideas with **SOLID** but his emphasis is also ethical in wanting to reduce hatred and lack of self control he believes is associated with anonymity
- The downside of anonymity



Managing your life information will include being your "information executor" after your death

- Memorials/tribute sites for self and others
- "Locked boxes" of passwords etc. for survivors—protecting assets.
- Legacy wishes and emails to send after death



Tribute sites, memorials, shrines



Tribute Times – Remembering Lives


http://www.tributetimes.co.uk/ Google

Fake Captian...ost(tm) Bot Grand Challe...d – The IET Boats for sa...Apollo Duck Broaden you...ng software Female geni...ncyclopedia InstaFlex Fr...nstaflex.com

Tribute Times
Remembering Lives


Home Obituaries Memorials Funeral Homes Our Services Help & Support Community Forum Login

Try A Memorial Site For Free
Example Obituary
Example Memorial Site
Bereavement Advice
Words of Comfort
FAQ



Memorial Websites & Online Obituaries

Remembering Lives...

Memorial Websites		Online Obituaries
No PC experience required		No PC experience required
Add up to 200 pictures		No limit on obituary length
Choose from a selection of background songs		Interactive condolence book for friends & family
Write and share their life story		Online immediately
Interactive remembrance book		
Private or public		
On-line immediately		
Unique internet address		



Facebook memorial-site shift

BBC Low graphics Help Search [Explore](#)

NEWS **LIVE BBC NEWS CHANNEL**

News Front Page
[World](#)
[UK](#)
[England](#)
[Northern Ireland](#)
[Scotland](#)
[Wales](#)
[Business](#)
[Politics](#)
[Health](#)
[Education](#)
[Science & Environment](#)
[Technology](#)
[Entertainment](#)
[Also in the news](#)
[Video and Audio](#)
[Have Your Say](#)
[Magazine](#)
[In Pictures](#)
[Comment & Discussion](#)

Page last updated at 10:49 GMT, Tuesday, 27 October 2009
[E-mail this to a friend](#) [Printable version](#)

Facebook 'memorialises' profiles

Facebook has announced that it will be giving friends and family the option to "memorialise" the profiles of members who have died.

It follows some cases of members receiving updates about dead friends.

If a user is reported as deceased, Facebook will remove sensitive information such as status updates and contacts.

When reporting a death, users must offer "proof" by submitting either an obituary or news article.

"When someone leaves us, they don't leave our memories or our social network," Max Kelly, head of security at the firm, wrote in the



Privacy settings of memorialised profiles will be changed

SEE ALSO
[Facebook](#) 16 Sep 09
[Club attac](#) 08 Oct 09
[Facebook](#) 23 Sep 09

RELATED IN
[Facebook](#)
The BBC is no
Internet sites

TOP TECHN
[Microsoft](#)
[Mobile fir](#)
[Intel and](#)



Preparing for the digital afterlife

How should we deal with web users' Facebook, PayPal and other accounts when they log off for good?

Duncan Jefferies

guardian.co.uk, Wednesday 30 September 2009 19.30 BST

[Article history](#)



Legacy Locker – The safe and secure way to pass your online accounts to your friends and loved ones.

http://legacylocker.com/ Google

Fake Captian...ost(tm) Bot Grand Challe...d – The IET Boats for sa...Apollo Duck Broaden you...ng software Female geni...ncyclopedia InstaFlex Fr...nstaflex.com

Sign Up Login

Legacy Locker

Features Why Legacy Locker? Pricing & Sign Up Support

The safe and secure way to pass your online accounts to your friends and loved ones.

Legacy Locker is a safe, secure repository for your vital digital property that lets you grant access to online assets for friends and loved ones in the event of loss, death, or disability.



The value of web accounts	Protect your online assets	Privacy and security	Family and loved ones	Financial and estate planners	Sign Up Now in 60 seconds or less
---------------------------	----------------------------	----------------------	-----------------------	-------------------------------	---

In the event you need to report a Legacy Locker user's passing or incapacitation. >

Report A Passing



Final message sites

<http://www.mylastemail.com/>

- Leave last emails, text, letters, book of condolence, bad aesthetics, write your own obituary.
- letterfrombeyond.com
- Leaving words and pictures to be sent to loved ones
- FinalThoughts.com
- Postmortem messages (10 free) after your “Guardian angel” gets the news/trigger (Gone—now an “entertainment resource”)
- SlightlyMorbidity.com
- Sends a list of last emails when triggered by a “trusted representative”

- 1/15/2019



[Home](#)
[How It Works](#)
[F.A.Q.](#)
[Contact Us](#)
[Tell A Friend](#)
[Letter Release](#)
[Free Account Offer](#)
[Create Your Account](#)
[Sign In](#)

Bookmark with:

- [Facebook](#)
- [StumbleUpon](#)
- [Reddit](#)
- [Digg](#)
- [del.icio.us](#)
- [Newsvine](#)
- [YahooMyWeb](#)
- [soc.lalizer.us](#)
- [What are these?](#)

Letter from Beyond

"Keeping your thoughts and memories alive"

CLICK HERE FOR DETAILS

AS SEEN IN THE NEW YORK TIMES

At Letter From Beyond, you have a precious opportunity to leave your thoughts and memories behind for when you pass on. A gift your loved ones will treasure for a lifetime.

Use your own personal space to leave a letter for those you care about. You have a blank canvas to fill with your words and pictures, and our promise to deliver your letters when the time comes.

Whether you want to remind your loved ones of your favourite times together, tell them how proud they have made you, or inspire them for the future, you've come to the right place.



To start using your ONE YEAR FREE ACCOUNT, [Create Your Account](#) with us today.

Please note that we take your privacy and security seriously. All transactions are processed over a secure, industry-standard SSL connection. Your card details are NOT stored by us once your payment has been taken.

[Ads by Google](#)
[Letter](#)
[Sympathy Card](#)
[Sympathy Poems](#)
[Clear Writing](#)



1/15/2019



Serious legacy Issues

- What is the death “trigger”
- “suicide sites”
- Permanence of information ---the British Library permanent “British Lives” inventory?
(Problem of any data permanence
cf.cryogenics)
- So MUCH personal information—the
quantity, sorting and management of it???
- What is the output of a whole life?



Conclusions

- Tim Berners Lee's original vision was the Semantic Web that understood its own content.
- It is coming into being as structured texts and data, above all in the areas of science and medicine where it is already indispensable, as well as commercial forms like Google's Knowledge Graph.
- The SW is a huge philosophical experiment: a web that knows what's on it, and « grounds » meaning.
- The Artificial Companion will be an interface to the web: to our whole life's information.
- Distributing our life information, anonymity and the Companion as an agent that protects and manages our web life.
- Identity shielding from the state and corporations.
- Safeguarding that information after death.



The End

