

SEMANTIC WEB, QUO VADIS?

KEYNOTE @ SCAI 2006 / STEP 2006

Ora Lassila

Research Fellow

Nokia Research Center, Cambridge, MA

2006-10-26



WHO, WHAT, WHERE, ETC...?

- Who am I?
 - Research Fellow @ Nokia Research Center Cambridge
 - elected member of W3C's Advisory Board
 - past employment: HUT, CMU, MIT, etc.
 - researcher, engineer, entrepreneur, venture capitalist
 - (founding member of Finnish AI Society, back in 1986...)
- What is NRC Cambridge?
 - "open" research lab located near Boston
 - close ties with MIT
 - long-term focus on what handheld devices might look like 5+ years from now





WARNING! Contains Personal Opinions

HERE'S WHAT TO EXPECT ...

(I am attempting a synthesis of the messages from several past talks...)

1. Semantic Web
2. "Web 2.0"
3. Why Is the Semantic Web a Problematic Topic?
4. Serendipity
5. Moving Beyond the Semantic Web
6. NRCC's Project "SwapMe"
7. Q & A

NOKIA
Connecting People

ABOUT THE SEMANTIC WEB

NOKIA
Connecting People

OBSERVATIONS

- Modern information systems (such as Web sites) decide several things for you:
 - how information is presented
 - how information is used (combined, manipulated, etc.)
 - (and sometimes) what things are automated
- “Pre-determined” presentation makes it hard to
 - provide access to, say, visually impaired people
 - access the information using a “non-standard device”
- (I find this offensive!)
- In addition: combining information sources is hard
 - automating this is very hard
 - automating unforeseen situations is impossible

NOKIA
Connecting People

OBSERVATIONS

- Web (content) was built for humans
 - human interpretation is needed to accomplish tasks on the Web
- We need “machine-friendly” content
 - information w/ accessible formal semantics
 - allow machines to reason about information
- Semantic Web is an attempt to address this
 - initial goal: enabling automation
 - short term goal: interoperability
 - long term goal: make computers work on our behalf
 - (instead of using them like tools)
 - remove humans from the loop

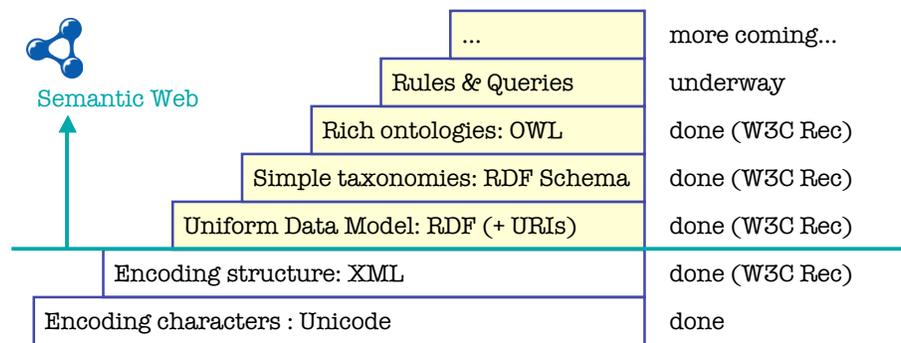
NOKIA
Connecting People

METADATA

- Data about data
- (Digital) library community a strong stakeholder
- Longer-term goal: alternatives for search
 - so far, search engine vendors have not bought into this
 - (chicken & egg problem, or something else...?)
- Current state: well established
 - Dublin Core
 - XMP (Adobe)
 - IPTC
- Social networks = metadata about people
 - “Friend of a Friend” (Foaf)

NOKIA
Connecting People

TOWARDS THE SEMANTIC WEB



- Semantic Web is built in a layered manner
- Not everybody needs all the layers
 - (i.e., it is OK if you don't want to use OWL...)

NOKIA
Connecting People

WHY SHOULD AI COMMUNITY CARE?

- Many AI techniques needed to “bootstrap” this
 - (not just KR)
- If we succeed, we will have built a “knowledge infrastructure”
 - easy sharing of knowledge (for various AI programs)
- (I believe Semantic Web is a component in the thawing of the “AI Winter”...)

NOKIA
Connecting People

ABOUT THE “WEB 2.0”

NOKIA
Connecting People

SEMANTIC WEB

- Data with formal semantics
 - RDF, OWL
 - SPARQL, RIF
- Spontaneous information integration (finally!)
- Semantic Web services, agents
- Strong emphasis on open standards

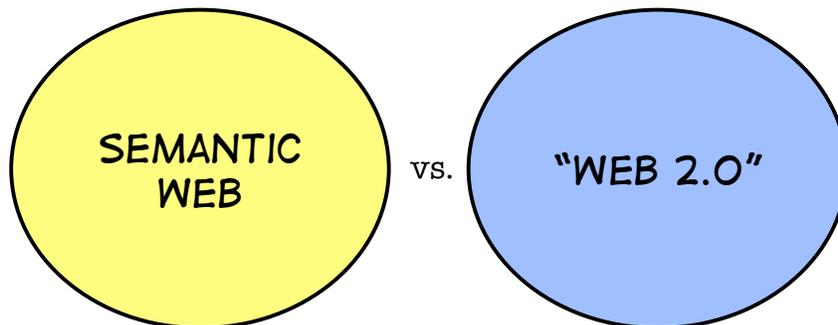
"WEB 2.0"

- New social phenomena: blogs, wikis, tagging, folksonomies
- New user interfaces
 - AJAX (or: "Rich User Experience")
- "New" kinds of data
 - microformats, RSS
 - "mash-ups"
 - Web services
- Plays "fast & loose" with standards

NOKIA
Connecting People

SEMANTIC WEB & "WEB 2.0"

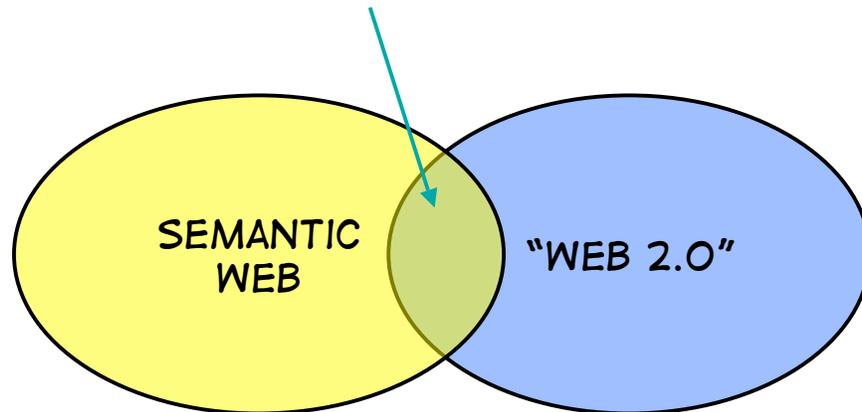
- What is their relationship?
- Will they stay separate? Does that even make sense?



NOKIA
Connecting People

SEMANTIC WEB & "WEB 2.0"

- NO! Considerable synergies exist



NOKIA
Connecting People

SOME CRITICISM OF WEB 2.0

- Name reflects "misguided philosophy"
 - "2.0" implies that this is a major new version
 - instead, the Web evolves
 - by the way, nothing really new has been introduced
 - (HTTP & HTML still rule...)
- Something good: transitioning the Web from a publishing medium to a communications infrastructure
 - wikis & blogs enable more people to be content producers

NOKIA
Connecting People

SOME CRITICISM OF WEB 2.0

- Semantic Web vs. “semantic web”
 - (neat vs. scruffy?)
 - microformats a step towards the real Semantic Web?
 - GRDDL to the rescue (really, XSLT)
 - misses some points (uniform metamodel, extensibility)
 - how do we connect folksonomies and ontologies?
- RSS (RDF Site Summary)
 - RSS 1.0 is based on RDF
 - integrates nicely with other RDF schemata (e.g., FoaF)
 - later versions miss many of the good features of the Semantic Web
 - XSLT to the rescue (the sole reason why, after all, it was a good idea to do RDF using XML)

NOKIA
Connecting People

YET ANOTHER VIEW...

I just wanted to say how much I've come to dislike this "Web 2.0" faux-meme. It's not only vacuous marketing hype, it can't possibly be right. [...]

- Tim Bray (one of the creators of XML), August 2005

NOKIA
Connecting People

WHY IS THE SEMANTIC WEB A PROBLEMATIC TOPIC?

NOKIA
Connecting People

MANY ISSUES

- “Smacks of AI...”
- “We can do this with XML...”
- Many scenarios are predicated on “open data”
 - conflicts with existing business models, IPR, etc.
- But the biggest problem is...

NOKIA
Connecting People

A DIFFICULT MESSAGE

- Any specific problem (typically) has a specific solution that does not require Semantic Web technologies
- Q: Why then is the Semantic Web so attractive?
A: For future-proofing

Semantic Web can be a solution to those problems and situations that we are yet to define

(seriously, I am not kidding...)

NOKIA
Connecting People

ser•en•dip•i•ty | .serənˈdɪpɪtē |
noun
the occurrence and development of events by chance in a happy or beneficial way: a *fortunate stroke of serendipity* | a *series of small serendipities*

(Source: Oxford American Dictionary)

ABOUT SERENDIPITY

- The defining characteristic of the Semantic Web
- Serendipity in interoperability
 - can we interoperate with systems, devices and/or services we knew nothing about at design time?
- Serendipity in information reuse
 - when information has accessible semantics, this is easier...
- Serendipity in information integration
 - can information from independent sources be combined?
 - NB: issues of identity are amplified
 - even simple forms of reasoning can help

NOKIA
Connecting People



ON INTEROPERABILITY

- Dependent on sharing
 - sharing is difficult
- In today's world, interoperability increasingly matters...
- [Berners-Lee, Hendler & Lassila 2001] emphasizes agents
 - goal: "serendipitous interoperability"...

NOKIA
Connecting People

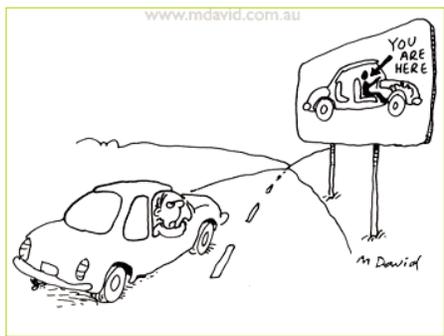
IDENTITY CRISIS?

- TimBL & W3C: “[...] *everything of importance deserves a URI*”
- Nice idea, but current reality is different
 - many things do not have URIs
 - it is unclear how to choose URIs
- Can we add something to RDF to help?
 - using something like owl:sameAs allows you to give “hints” to system (reasoner) on how to unify data
 - inverse functional properties (i.e., primary keys) allow automatic integration
 - use a reasoner to do this
- (I call this “RDF++”)

NOKIA
Connecting People

ONE BURNING QUESTION...

- Q: When will the Semantic Web arrive?
- A: When new “mash-ups” are done spontaneously, by ordinary people, not by nerds
 - (this is at least one possible answer)
 - promising things emerging (e.g., MIT’s “Piggy-Bank”)



While there are several useful mashups, many are simple novelties or gimmicks, with minimal practical utility.
- Wikipedia on “Mashup”

NOKIA
Connecting People

BEYOND THE SEMANTIC WEB

NOKIA
Connecting People

WHAT SHOULD WE DO NEXT?

- Now forget that we are talking about the Web...
- Modern PC applications are essentially just repositories for information (typically) in proprietary formats
 - combining or sharing information across application boundaries is impossible or difficult at best
 - any two applications can be engineered to enable information exchange, but we cannot anticipate all possible "pairings"

NOKIA
Connecting People

IMPLICIT → EXPLICIT

- In addition to the explicitly represented information, these systems (e.g., PC apps) hold a lot of implicit information
- Implicit information is largely inaccessible to current applications
- For example, your calendar may indicate that you have a flight reservation from Boston to Helsinki
 - implying that if you take the flight, you will then be in Helsinki
 - (this will be useful)

NOKIA
Connecting People

BRAVE NEW APPLICATIONS

- Exhibit robustness in the face of
 - changing, inconsistent and unexpected data
 - variations in reliability, trust
- Operate autonomously in “unanticipated” situations
- Capable of serendipitous behavior, opportunism

NOKIA
Connecting People

"SMART DATA" ...

- Information, in more "raw" form, is appealing
 - not tied to specific rendering, specific device, specific browser, etc.
 - most importantly, not tied to a specific task
- Presentation can be determined dynamically
 - based on user, device, ...
 - based on context (task, environment, etc.)

NOKIA
Connecting People

"SMART DATA" ...

- New approach: separate data from applications
 - data carries declarative descriptions of its semantics
 - manipulate any data with (almost) any application
 - e.g., browse photos using your calendar
- Combining information could be automated
 - Web 2.0 "mash-ups" are a step in the right direction, but they still require engineering
- All data available in Semantic Web formalisms
 - shared local and distributed repositories ("triple stores")
 - legacy data sources "exposed" as RDF (e.g., via XSLT)
 - query data via SPARQL, etc.

NOKIA
Connecting People

MOBILE & UBIQUITOUS COMPUTING

NOKIA
Connecting People

MOBILE WEB ACCESS TODAY

- Web access on mobile devices is available today
 - (in fact, some browsers are quite good...)
- Some technical limitations exist
 - network (narrow bandwidth, high latency)
 - display (typically small)
 - input (often no full keyboard)
- Content is designed for “standard devices”
 - (= PCs: high bandwidth, large display)
 - most (commercial) content is rendering-oriented

NOKIA
Connecting People

ISSUES WITH MOBILE WEB ACCESS

- Real limitations are of different nature...
 - mobile users are attention-constrained
- Mobile use happens in “unusual” situations
 - when laptops are not viable (e.g., in the car)
 - when paying attention to something else
 - browsing is not the ideal paradigm for information access
- What do we need?
 - content that’s not rendering-oriented
 - more automation (now, humans essentially do all the work)



NOKIA
Connecting People

UBIQUITOUS COMPUTING

- An interoperability nightmare!
 - instead of occasionally connecting a handful of devices, dynamically connect/disconnect/reconnect possibly hundreds of devices
- Traditional approach to interoperability: standardization
 - anticipate everything about the future
 - and a priori agree on how to act
 - (or: force all interactions to a restricted set)
- What about unanticipated situations?
 - how do you agree dynamically on how to behave in a situation that wasn’t covered by a standard?
 - not “future-proof”

NOKIA
Connecting People

UBIQUITOUS COMPUTING

- Connections with public and/or untrusted devices
- Need to “borrow” functionality from other devices
 - uniform representation of functionality is useful
 - this implies that we need to be able to represent and reason about contracts, payments, etc.
- The vision is largely contingent on getting unanticipated “encounters” of devices to work
- Today, high cost of ensuring interoperation
 - any interaction has to be specifically designed/engineered
 - heavy emphasis on application-specific standardization
 - spontaneous interoperability is next to impossible

NOKIA
Connecting People

NRCC'S PROJECT "SWAPME"



NOKIA
Connecting People

SOME ASSUMPTIONS...

- Semantic Web technologies are useful
 - role of serendipity is critical in “ephemeral” use cases (e.g., in ubiquitous computing)
 - Many problems in mobile and ubiquitous computing are (ultimately) problems of representation
- Context-awareness is useful
- Policy-awareness is useful

NOKIA
Connecting People

CONTEXT-AWARENESS

- “context” = information about “current situation”
- Can guide decisions about selection and prioritization
 - (of what information gets presented and how)
- Automation & autonomy
 - knowing the context can ease service discovery
 - contextual information can be used to limit choices in planning
- Semantic Web formalisms and techniques are beneficial in implementing context-awareness [Lassila & Khushraj 2005]

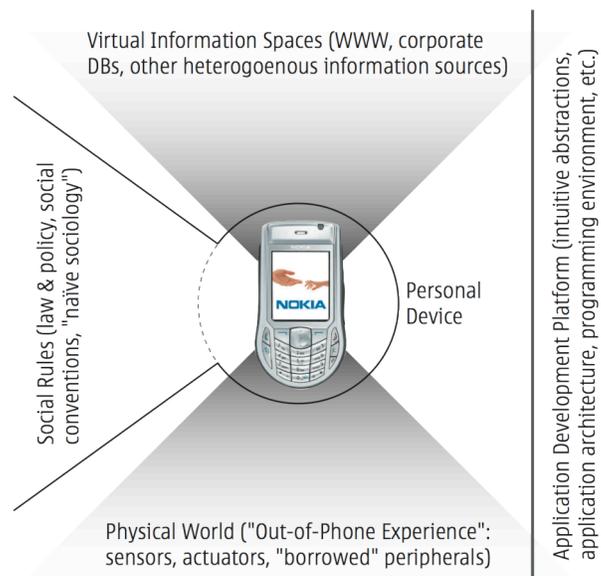
NOKIA
Connecting People

POLICY-AWARENESS

- Ability to represent, reason about, and enforce policies
- Policies: representations on how to act in a future situation
 - can control data access and usage (security & privacy)
 - support autonomous behavior
- Trying to determine context benefits from policy-awareness
 - e.g., access to some information permitted given that it is only used for context-determination
- Semantic Web languages have been used to build sophisticated policy solutions [Kagal 2004]

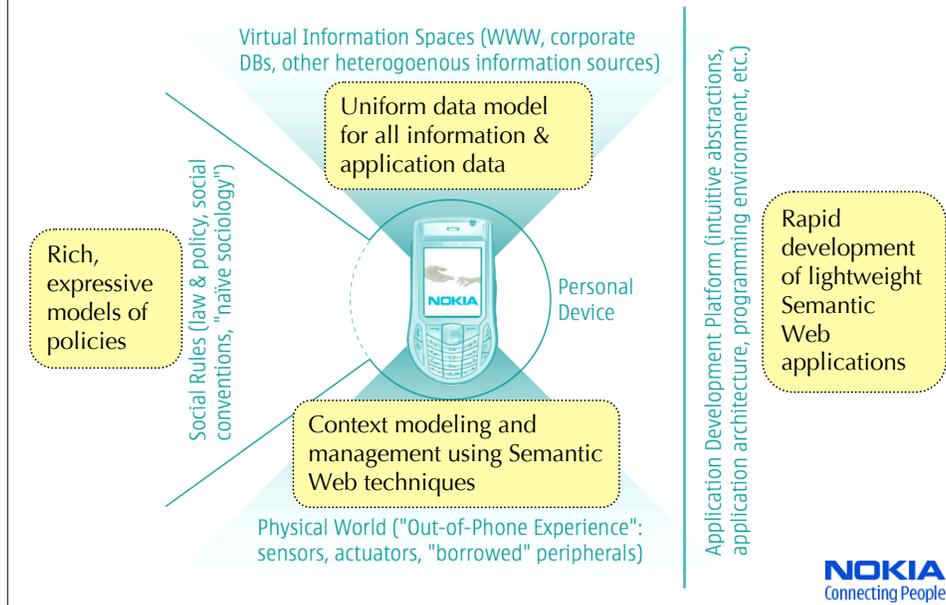
NOKIA
Connecting People

MOBILE ECOSYSTEM



NOKIA
Connecting People

MOBILE ECOSYSTEM & "SWAPME"



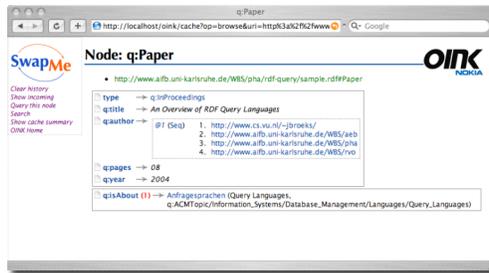
"SWAPME" APPLICATION PLATFORM

- All DATA uses Semantic Web formalisms
 - this may require various types of adapters and transformations to allow legacy data sources to be connected to the system
- All FUNCTIONALITY described using ontologies
 - e.g., OWL-S
 - local and remote services
- All ACCESS to data and services enabled and limited by pervasive policy-enforcement
 - policies themselves are expressed in a policy language based on Semantic Web technologies (e.g., REI)
- All BEHAVIOR is sensitive to the current context
 - e.g., focus, prioritization

NRC'S "WILBUR" SEMANTIC WEB TOOLKIT

- Open-source, written in Common Lisp
 - triple store
 - path query language
 - rewrite rule engine
 - RDF++ reasoner
- Semantic Web browser "OINK"
 - querying-by-browsing

LATE-BREAKING NEWS:
NOW ALSO IN PYTHON!



Ora Lassila. **Enabling Semantic Web Programming by Integrating RDF and Common Lisp**, *First Semantic Web Working Symposium*, 2001.

Ora Lassila. **Taking the RDF Model Theory Out for a Spin**, *ISWC 2002*

Ora Lassila, Deepali Khushraj, Ralph R. Swick: **Spontaneous Collaboration via Browsing of Semantic Data on Mobile Devices**, *ISWC 2006* workshop on Semantic Desktop and Social Semantic Collaboration

Ora Lassila: **Generating Rewrite Rules by Browsing RDF Data**, *RuleML 2006*

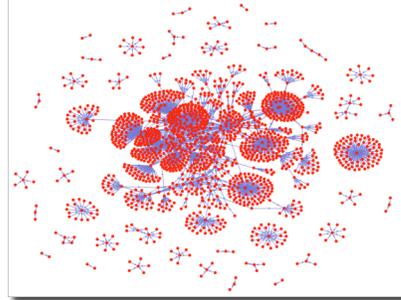
NOKIA
Connecting People

FINALLY...

NOKIA
Connecting People

"A LITTLE SEMANTICS GOES A LONG WAY"

- RDF: Uniform metamodel
 - the importance of this cannot be overemphasized
- Useful, general schemata
 - DC, IPTC4XMP, SKOS
 - RSS, Annotea
 - FoaF, CoaC, DoaP
 - (all these play nicely together)
- We have no way of knowing how people are going to use and combine information
 - enabling and encouraging linkages helps



FoaF clusters (source: MindLab, Univ. of Maryland)

NOKIA
Connecting People

SOME RANDOM CONCLUSIONS

- Semantic Web is a response to
 - certain shortcomings of the "classical" Web
 - a desire to have computers do more on our behalf
- "Web 2.0" will help
- Evangelizing the Semantic Web has been hard
 - AI community should care
- Ontological techniques useful more generally
- Mobile & Ubiquitous Computing have many problems we think Semantic Web applies to
 - NRCC's project "SwapMe" studies how

NOKIA
Connecting People

QUESTIONS?

- <http://wiki.nrcc.noklab.com/SwapMe>
- <mailto:ora.lassila@nokia.com>

- Thanks to:
 - Deepali Khushraj
 - Nokia Technology Platforms
 - ...

NOKIA
Connecting People

***SEMANTIC WEB
NEEDS YOU!***

- The Semantic Web will emerge from the serendipitous reuse of information
- Action points for you:
 1. make information available (use RDF, OWL)
 2. do not "reinvent" - instead, borrow from others (i.e., use existing schemata)

Source: Library of Congress



NOKIA
Connecting People