Living in a Transparent Future: Search in a Wired World

Michael Wollowski
Computer Science and Software Engineering Department
Rose-Hulman Institute of Technology
5500 Wabash Ave.
Terre Haute, IN 47803, USA
wollowski@rose-hulman.edu

ABSTRACT. We present and discuss a likely future world in which advances in search and the wiring of our environment would provide us with ready access to much information. We outline the current status of search and the goals of the semantic web. We argue that much of the information already in electronic format will likely be made accessible through the web. We present a likely future scenario of a wired world, a world in which entities as simple as a bottle of pills to as complex as the human body are wired to the web. This adds much additional information to the web, information that by-and-large is currently not gathered in electronic format. We present and discuss three representative cases which highlight the benefits and drawbacks of ready access to information pertaining to individuals. We suggest that the future scenarios are not too far off in the making and suggest that a dialogue be started, attempting to develop enforceable privacy policies.

INTRODUCTION

Searching the web is a formidable challenge. Companies which develop search engine technologies are doing very well. Currently, information on the web is largely generated by people. Soon, devices connected to the web will generate vast amounts of contents. This will pose an immense challenge for search engine developers. In this paper, we attempt to envision a future scenario in which billions of devices are connected to the web. We will **not** focus on solving technological challenges that are posed by privacy concerns. Rather, we assume a happy world in which information is used responsibly. This way, we can focus our energies on the task of envisioning a connected future.

Attempting to envision future scenarios is a fascinating activity and we are not alone in this attempt [8, 9, 10, 11]. We reach beyond the visions presented there, attempting to provide a more complete picture of a connected world. We analyze several scenarios, arguing that the technology largely exists to make those visions a reality.

CURRENT WORK IN SEARCH

Currently, material on the web is largely edited by people and meant for the consumption of people. We can search for the birth date of Harry S Truman, a good deal on a vacation package, or the new 2005 Ford Lightning. Most information is contained on web-pages

or in databases. Current work on search engines aims to satisfy two kinds of clients: (i) the individual web-searcher and (ii) corporate clients. Work which focuses on the individual consumer aims to build a “genie” which can be prompted in natural language and who will give answers, rather than links to web-sites [13]. Work which focuses on corporate clients is largely concerned with locating information in the myriad of documents housed on corporate intranets [14]. This includes searching documents in formats other than HTML, such as PDF files and PowerPoint slides. In both cases, information that is searched is primarily designed for human consumption.

OLD GOALS AND NEW DEVELOPMENTS FOR THE WORLD-WIDE WEB

The most important current development for the world-wide web is the semantic web. The semantic web is designed to provide “data integration across application and organizational boundaries” [7]. A good example for such a scenario is depicted in figure 1. Figure 1 is a diagram from Tim Berners-Lee 1989 proposal for an information management system, which eventually started the world-wide web [5].

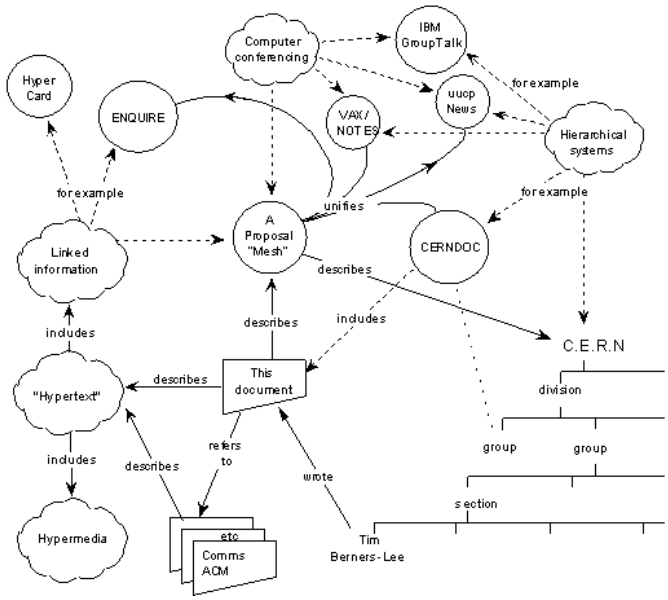

Figure 1: From Berners-Lee’s 1989 proposal for an information management tool.

The semantic web consists of several technologies, including XML, XML Schema, RDF, RDF Schema, and OWL. See [6] for the current stand of development.

TOWARDS A WIRED FUTURE

The semantic web promises to do great things for web-pages containing information meant for human consumption. Once we have a reasonable implementation for the “semantic layer cake” [15], we will be able to realize the core features of the “genie”

making web-searches a convenience rather than a chore. However, there is another aspect of the web for which the semantic web will be very useful. There is a desire to connect devices such as pill bottles, cars, homes, and even people to the web. While the term *wired* is an often used catch phrase, most devices will likely be connected in a wireless fashion.

In principle, we can connect just about any device to the web. The device may have its own IP number, or it could be connected to a device, such as a computer, which does have an IP number. Currently, one can purchase refrigerators which can be connected to the web [19]. Our departmental soda machine is an example of the second category, it has several sensors which are connected to a computer which itself is connected to the web [18]. Microsoft is working on devices which take regular images of our field of view [16]. In general, people can soon record a lifetime of conversation [4]. In principle, cars can be wired to the web [11, 12].

LIVING IN A CONNECTED WORLD

In order to explore how to use the devices from the prior section to construct a connected world, we present three increasingly wired scenarios. After each scenario, we analyze the underlying technology which is necessary to make the scenario happen.

Example one: A connected car

Traffic jams are a constant challenge of a driver's patience. It would be useful to obtain updated traffic information, without having to catch that information on a radio broadcast. A solution for this problem will be available in the next few months [20]. XM Satellite radio will show drivers on the navigation system which parts on their planned route of travel have traffic delays. The technology necessary to make this scenario happen are: (i) a GPS system, (ii) electronic maps of the area of travel, and (iii) a digital way to communicate which roads and streets have a traffic jam. The key challenge is to pinpoint areas of traffic delays. Within cities, street numbers can be used, and on highways, highway markers can be used. The system does not suggest alternate routes yet. Such a feature would require search techniques, possibly involving Artificial Intelligence (AI) techniques such as heuristic search to suggest good alternate routes.

Another challenge to drivers is the task of locating parking spaces, especially in highly congested areas. MIT's media lab proposed "smart curbs" as a way to address this problem [17]. In order to accomplish this vision, we need to implant chips in curbs which are connected to a central computer. In order to uniquely identify a car, one could use the OnStar system or better yet, electronic toll collection systems, such as E-Z Pass. Electronic toll collection could also be used to pay for parking, eliminating the need for parking meters. A car's computer would employ search to locate the nearest empty parking space based on their current location and their destination. One of the problems with a fairly straight-forward implementation of such a system is that likely, by the time one reaches the "empty" parking space, it will be taken by someone else. One could

possibly implement some sort of reservation system. Alternatively, AI analysis techniques could be used to identify areas which will likely contain several available parking spaces. If the process of finding a parking space takes longer than expected, a note could be sent to colleagues or spouses who may be waiting for a meeting or rendezvous.

In order to facilitate the use of different toll systems, it may be desirable to employ some non-proprietary standards for smart curbs, possibly involving the semantic web so as to identify users and curbs across systems.

Example two: A connected house

Some of us are likely using home automation devices, such as X-10 based light switches. The house of the future will be heavily wired. Builders are interested in automating houses so as to obtain "smart houses." In a smart house, select lights and the furnace could be turned on when one of the owners' car drives up the drive-way. All lights could be automatically turned off when the last person leaves the house. The smart house could monitor their inhabitants' progress when getting up in the morning, so as to "talk" to the wired car, when things take longer than usual. The outcome of a house-car chat could be that the car gets started so as to warm up the engine on a cold winter morning. AI techniques could be used to estimate the time it takes to get to the office so as to alert co-workers about a meeting that may start late.

Wired houses already exist. In order to facilitate that house and car interact, a common communication protocol and a common representation of information are necessary. The semantic web will prove useful for this. In order to achieve cross-platform compatibility, a language such as Java may provide useful. There has been talk about a Java-enabled car [11, 12], why not build a Java-enabled house?

Not just builders are interested in wired houses, organizations such as the Alzheimer's Association are actively pushing for the next step [1]. The Alzheimer's Association is interested in wiring everything from pill bottles, to refrigerators [10]. A sensor on a pill bottle can be used to determine whether a patient has taken their medication. If the medication has not been taken in a specified time window, medical personnel can be alerted to that fact. This example is within technical feasibility. As a matter of fact, patients with pace makers, can already download pacemaker data to a computer where physicians analyze the data. In the future, it would be desirable to use more sophisticated AI techniques to analyze data provided by pace-makers or sensors in the home, so as to remove medical personnel from low-level analysis. The semantic web and languages such as Java would again be very useful to provide cross-system and cross-platform compatibility.

Example three: A java-enabled person

In search of a juicy burger. A guy, let's call him Bob, has high cholesterol. His doctor strictly prohibits him from eating foods high in cholesterol. One day, while Bob goes

about his daily business, he starts to have this nagging desire for a juicy hamburger. After much agonizing, he cannot resist his urges any longer and drives over to the nearest Burger King. Not one who is prone to exercise, Bob cruises down the drive-through lane to order a burger with fries. In the process of wirelessly deducting the cost of the meal from Bob's E-Z Pass account, Bob's computer realizes that the bill comes from a fast food establishment. The computer quickly studies the bill to notice that the order is for a hamburger and fries and promptly refuses payment and the order is cancelled. Bob's health insurance company made him install a piece of software which monitors his purchases of high-cholesterol foods. Muttering to himself inaudibly, Bob drives home.

Pulling into the garage, he suddenly has a bright idea; let's take the old Triumph TR 6 out for a spin. It doesn't have an E-Z Pass system! He heads to the local burger joint, where cash is still accepted. Waiting in line and chatting to the people around him, he suddenly gets shoed away by an off-duty paramedic. The paramedic received an alarm from a centralized system which was able to pinpoint Bob's location based on the health monitor implanted in his body. The health monitor includes a GPS system so as to alert paramedics to his location, in case of emergency. Bob's location was pinpointed to be in a high-cholesterol zone, alerting the medic. After receiving an admonition from the medic, Bob heads home again.

His friendly neighbor, noticing him all dejected, shouts: "Hey Neighbor! Have a beer, I just turned on the grill, interested in joining me for a cookout?" Teary eyed and weak-legged, Bob agrees. While the hamburgers are sizzling on the grill, an alarm goes off in the central system. Bob had to agree that as part of a lower rate on his health insurance, the insurance company would get access to the images which his digital camera takes of his point of view every few seconds. The advanced image analysis software recognized flames, smoke, and juicy hamburgers on the grill. This time however, no one will keep Bob from his hamburger, as the personnel monitoring the alarms are socializing outside, over tofu burgers.

While some aspects of this example may seem far-fetched, they can in principle be realized. We already mentioned that people wearing pace makers can download data to central computers. GPS systems are getting smaller. It may be beneficial for everyone to have some health monitoring system implanted in their body, thereby improving diagnosis of illnesses. Life insurance companies regularly offer lower rates for non-smokers. Recently a car insurance company offered discounted insurance rate for drivers willing to install a black box in their car which is periodically monitored by the insurance company [3]. The Microsoft research center in Cambridge, England developed a prototype of a personal image and data recall system which can record up to 2000 VGA images a day [16]. At this point, pictures need to be annotated with meta-data. However, picture analysis software is getting better and better. McDonald's operates a restaurant in Long Island where patrons can pay at the drive-through window with their wireless E-Z Pass account, traditionally used to pay tolls [2]. We have already argued the possibility of wired buildings and cars in the prior sections.

CONCLUSIONS

We showed that most of the hardware required to make a connected world a reality exists already and will soon be sufficiently small and cheap to make them viable. While the semantic web still has a ways to go in order to realize the "genie" which gives us instant access to good information, the connected world does not necessarily have to rely on the most powerful methods that the semantic web has to offer. Search in the connected world goes on behind the scenes. Much of the information necessary to make the connected world a reality is already stored in electronic format. It is mostly located in databases. Information that will be placed on the web through sensors can be specified with a well-defined syntax, and as such, database techniques such as keys will be sufficient to pull information from different sources together, while preserving the ability to uniquely identify individuals. A platform independent language such as Java should be sufficient to ensure interaction between many kinds of devices.

While a connected world will make certain aspects of our life more convenient, there is the possibility of becoming enslaved to a system that is hard to control. Most of the search and processing will go on behind the scenes, initiated by devices rather than the user. This makes the use of our information hard to control.

The extent and the ramifications of a connected work are hard to fathom. In this paper, we merely scratched the surface of a wired future. It seems that we will have to wait and see how different aspects of a connected world will evolve. If both the semantic web and a connected world become a reality, it is hard to see which piece of information could not be gathered with more or less effort. One may not wish to make an assumption of a happy wired world. In this case, it is high time to seriously work on building security into the systems that we are about to develop, and to develop acceptable norms for confidentiality.

REFERENCES

[1] Alzheimer's Association. *Everyday Technologies for Alzheimer Care Research Grants*. http://www.alz.org/Research/RGP/ETAC.asp

[2] The Associated Press. *McDonald's testing e-payment system.* http://www.usatoday.com/tech/news/2001-05-29-mcdonalds-e-payments.htm, May 2001.

[3] The Associated Press. *Drivers trade privacy for insurance discounts.* http://www.cnn.com/2004/TECH/biztech/09/03/blackboxdriving.ap/, September 2004.

[4] Gordon Bell and Jim Gray. *Digital Immortality*. Communications of the ACM, pp 29-31, Vol. 44, No. 3, March 2001.

[5] Tim Berners-Lee. *The Semantic Web: The Original Web Realized.* http://www.w3.org/Talks/2001/07/30-swws/slide5-0.html, July 2001.

[6] Tim Berners-Lee. *The Semantic Web Wave*. http://www.w3.org/2003/Talks/01-siia-tbl/slide19-0.html, January 2003.

[7] Tim Berners-Lee. *What is the Semantic Web?* http://www.w3.org/2003/Talks/0521-www-keynote-tbl/slide8-0.html, May 2003.
[8] Tim Berners-Les. *Semantic web is for connecting things.* http://www.w3.org/2004/Talks/0519-tbl-keynote/slide17-0.html, May 2004.
[8] Tim Berners-Lee, James Hendler and Ora Lassila. *The Semantic Web.* Scientific American, May 2001.
[10] Jill Glomstad. *Everyday Technologies*. Advance for Speech-Language Pathologists & Audiologists, p. 24, March 1, 2004.
[11] Janice J. Heiss. *The Network is the car.* http://java.sun.com/features/1999/06/concept_car.html, June 1999.
[12] Kammie Kayl. *The Networked Car: Where the Rubber Meets the Road.* http://java.sun.com/features/2000/10/convergence.html, October 2000.
[13] Elizabeth Liddy. *Why Settle for a List, when you Want an Answer?* http://www.infonortics.com/searchengines/sh02/02slides/liddy.pdf, April 2002.
[14] Andrew Littlefield. *Effective Information Retrieval Across New Content Forms*. http://www.infonortics.com/searchengines/sh02/02slides/littlefield.pdf, April 2002.
[15] Eric Miller. *Semantic Web Architectural Dependencies*. http://www.w3.org/Talks/2001/07/30-swws/slide21-0.html, July 2001.
[16] Microsoft. *SenseCam Personal Image & Data Recall* http://research.microsoft.com/research/hwsystems/
[17] Phil Patton. *At M.I.T., Rethinking the Car for City Life.* NYTimes, p D9, September 6, 2004.
[18] Rose-Hulman Institute of Technology, Computer Science and Software Engineering Department, Web-based Soda Machine. http://soda.cs.rose-hulman.edu
[19] Samsung refrigerator with internet connection. http://www.samsung.com/Products/Refrigerator/HomePADRefrigerator/Refrigerator_HomePADRefrigerator_RH2777AT.htm
[10] Eric Taub. *As a Business gets Bigger, its Radios get Smaller*. NYTimes, p C3, August 16, 2004.