By Jason Ohler

THE
Power
AND Peril OF
Web 3.0

IT'S
MORE
THAN
JUST
SEMANTICS

If the Semantic Web lives up to its promise, intelligent tagging technology will enable educators and students to spend less time searching and sifting through the information overload and more time thinking and participating.

The Information Age has been built, in part, on the belief that more information is always better. True to that sentiment, we have found ways to make a lot of information available to the masses—perhaps more than anyone ever imagined.

If you Google *global warming*, for example, you'll receive more than 33 million hits in less than a second. Unfortunately, 33 million is really too much for most of us. If you're anything like me, you'll tend to follow the first 10 or 20 links, read a bit at each location, and unconsciously cultivate the illusion of being informed. But this is much more dangerous than it seems, as Google's page-rank algorithm orders information by popularity rather than by, say, source credibility.

It's time to face facts: We have gone from information underload to overload in 25 years, and I have to wonder if we're any better off than we were before. There simply isn't time to read a fraction of the content available, let alone to reflect on it or determine whether it's fact, political spin, or total hooey.

Enter the Semantic Web, often called Web 3.0, which has the potential to come much closer to World Wide Web inventor Tim Berners-Lee's original conception of a universal network in which computers adapt to humans rather than the other way around.

How the Semantic Web Works

Right now when you search the Web, you find pages, not data. You are then left to slog through the pages to determine how relevant your findings are, and then manually cross-reference whatever you want to use. The goal of the Semantic Web is for users to spend less time looking for information and more time using and participating with what they find. In other words, less slogging, more blogging.

Before we can have a Semantic Web, though, we must first make informa-tion much more understandable to ma-chines. That means making three leaps over the current version of the Web.

Leap 1: Use intelligent tagging. Be-hind every webpage are lines of code that tell your computer how to dis-play information on your screen. To verify this, go to the menu bar of your favorite Web browser and select the Page Source option. What you see is mostly in hypertext markup language (HTML), the language of the Web.

In HTML, the only tags that would likely modify my phone number text (besides font and color tags) are <p> and </p>, which do no more than tell the computer to display, and then to stop displaying, the text on the screen. In this system, a human being must discern the familiar pattern of the phone number, because computers are not programmed to do this. So if you want to find my phone number on my website, you need to root around in my bio information until you find it. I can help your search by putting the word "phone" nearby, but I can't make my phone number understandable as such to machines. They just see characters.

Could we program search software to find phone numbers? We do have tools for this, such as microformats and Plain Old Semantic HTML (POSH). But these are stop-gap work-arounds until the real Semantic Web gets here, as they don't offer the exten-sible interconnectivity and scalability that true semantic data offer.

Now imagine a Web in which your computer displays phone numbers on your screen using this command: <phone>555-555-5555</phone>. This kind of "intelligent" tagging would change everything. A com-puter could "know" a phone number when it sees it. If enough people in the standards business agree on us-ing a <phone> tag for phone numbers and enough Web programmers adopt those standards, then finding my phone number with a semantic Web search will become very simple.

Intelligent tagging also takes care of another thorny problem: phone num-bers in different formats. You can use an international phone format with a country code or a local phone number with no area code. You can even use an encrypted version of your phone number to confuse spam autodial-ers. As long as it is tagged with the <phone> tag, the Web can still recog-nize it as a phone number.

Leap 2: Group intelligently tagged in-formation into ontologies. Can we ex-tend our success in finding my phone number to finding my town or my ZIP code? Sure. But more important, we can tag each component of my address and then group them into something called an ontology. In this case, a per-sonal information ontology might consist of <name>, <add1>, <add2>, <zip>, <phone>, <favoriteMusic>, or whatever the greater digital community has agreed upon as a standard. So, rath-er than searching for each piece of my address separately, you simply search for my address, which would return all of the information that is semantically associated with that address.

Leap 3: Use shared ontologies and data-bases. Now imagine sharing not just tags, but whole ontologies. This would allow you to search and cross-refer-ence databases. For example, if there is a personal information database and a journal publication database, and both use my address ontology, then these two data sets can share information. As a researcher, you can use one data set as a portal into another, allowing you to merge, borrow from, and cross-reference them.

Now extrapolate this. Imagine that we can convince others to use the address ontology across the many do-mains of our lives, including Amazon. com, our state motor vehicles depart-ment, our health insurer, Facebook, a database for local musicians—whatev-er intersects with your life. When you

search for my phone number, instead of receiving a list of pages that might have my phone number, you receive my phone number, set within the context of the other information you asked for that is associated with my phone number. And because information is intelligently cross-referenced, a semantic search can "make sense of the information," telling searchers that I am a guitar player (based on my local musician database) with a van (according to the DMV database) who loves fusion jazz (extrapolated from my iTunes selections).

If this makes you worry about your privacy, it ought to. But it will also make finding the information you really want much easier and faster.

Education 3.0

The Semantic Web may not be mainstream yet, but many are using the technology right now. According to industry expert Liam Ó Móráin from the Digital Enterprise Research Institute (DERI), Google, Wikipedia, Microsoft, and Yahoo are exploring the use of semantic data, and news media giants, such as Reuters and the *Financial Times,* leverage semantics in their search engines. Large amounts of semantic or linked data are also being created to satisfy domain-specific data needs for the U.S. Census, among others. And some scientists have adopted the Semantic Web because scientific information, which is so hierarchical and classified in nature, is well suited to the nature of semantic structure. For a good example of this, look at "Rewerse: Adding Semantics to the Bioinformatics Web" (http://rewerse. net/A2/Overview.htm).

Before we can start using the Semantic Web in schools, however, we'll have to work through a couple caveats. First, Web 3.0 relates only to Web-accessible information. Textbooks and historical documents not on the Web would need to be made semantically

This diagram shows a section of the data sets that have been published and interlinked by the Linking Open Data Project of the World Wide Web Consortium (W3C), an international community that develops standards to ensure the long-term growth of the Web. To see the full graphic and for a clickable version, go to: http://esw.w3.org/SweoIG/TaskForces/CommunityProjects/LinkingOpenData.

accessible or be excluded from the database of materials. Second, social studies interfaces with a number of other areas. So everyone with an interest in this area—which could include anything from Wikipedia to news agencies to the Smithsonian—would need to agree on tags such as <historicalText> or <patrioticSpeech>. There will be a tremendous amount of data to organize.

For now, let's consider more general potential impacts of the Semantic Web on education from the learner's point of view.

Knowledge construction. To start with, basic Internet searches would become much more effective. When a student searches for *causes of the Civil War*, for example, rather than receiving a list of pages that merely contain the words *cause* and *Civil War*, he would receive information in which the word *cause* is specifically related to the cause of the Civil War.

In one vision of a well-developed Semantic Web, that student's search would return, rather than a list of page hits, a multimedia report drawing from websites, articles, electronic book chapters, blog dialogue, YouTube presentations, cell phone memory, virtual reality resources—anything that is available via the Web. The report would compare, collate, and synthesize the information in the report and present it perhaps in a wiki-like structure, where it could further analyze the information by points of agreement and disagreement, or in light of political positions or contrasting research. Information could also be personalized, alerting users to preferred and local resources based on their profiles. Again, the goal is to reduce the amount of time we spend searching and sifting so that we can spend more time thinking and participating.

Personal learning network (PLN) maintenance. Many of us spend far too

much time searching the Web, trolling blogs, wading through long podcasts, and so on just to find the few nuggets we can use to augment our PLNs. Besides being inefficient, this approach to managing our own educational resources can often lead to inaccuracies simply because we run out of the time or motivation to do a thorough job.

One of the primary improvements that Web 3.0 offers is the ability to build PLNs around subjects and information relationships rather than around specific tools and services. Specialized computer "knowbot" programs acting as personal learning agents could identify relevant information from any source that is semantically accessible and provide an information synthesis tailored to a personal learning objective.

Personal educational administration. Schools and other educational institutions tend to be isolated entities that don't play well together. Students who transfer between schools can bear witness to how difficult it can be to do something as basic as transfer credit for a course from one institution to another. But even if we decided to develop a student-centered, multi-institutional approach to education, we would discover that doing so is logistically impossible for a very practical reason: Education providers typically do not share common languages to describe course or degree requirements.

Semantic Web technologies have the potential to challenge institution-centric education with the same force that distance-learning technologies challenge place-centric education. At some point, institutions will describe courses and degrees semantically to help their own internal functioning. This will have the secondary effect of making many of the components of education somewhat comparable across institutions. The result will be that students may be able to identify comparable coursework from sev-

eral education providers and, in the process, perhaps meet the graduation requirements at more than one. This will force the potential of intra-university degrees and institutional cooperation. Smart schools will get out ahead of this now.

The Road Ahead

How the Semantic Web will be created is still under debate. No doubt it will be the result of many advancements in a number of areas. Artificial intelligence may be able to parse Web text on the fly, turning it into useful semantic information in real time. Knowbots may troll the Web 24/7, inferring the semantics of the information it finds. And we will surely create some aspects of Web 3.0 the old-fashioned way—by writing new code. Although recoding the entire Web seems an impossible task, bear in mind that it has already been recoded a number of times. Advanced tools and the efforts of thousands, if not millions, of Web-savvy people will help regenerate the Web as new semantic capabilities become available. We will never stop recoding the Web.

Should we worry? Always. Every technology has the potential to both connect and disconnect, and the more powerful the technology, the more magnified these properties. As I explain in *Digital Communities, Digital Citizenship*, the problem is that connective properties are obvious, immediate, and helpful, whereas disconnective properties are often camouflaged, delayed, and disappointing. We all bought microwave ovens because they connected us to an easier life. Only in retrospect do we understand that they disconnected us from the benefits of family dinners by enabling children to cook dinner on their own.

After a public beta testing period—which could last for many years and become rather frustrating at times—the Semantic Web's connective properties will become obvious. It will

provide clearer, more interconnected, informed searches as well as ways for us to create our own resources and connect them much more coherently to the greater world of information.

Of course, the disconnective properties will have immense power as well. The potential to access information about each other will skyrocket. The pressure to make our information available to marketers, health care companies, public agencies, and other interested groups may become intense. The Semantic Web could well go rogue unless there is a legal framework to contain it.

But the truly insidious part of the Semantic Web lies in the fact that, just below its surface, in code that very few of us will understand, is a rewriting of the world as we know it. Whoever determines how information is linked in the Semantic Web will introduce perspective and bias into a global rendering of reality, through the information that we have access to, that will be too subtle to notice. Because we will be so busy enjoying the dexterity and clarity the Semantic Web offers, we will simply go on about our business, blissfully unaware of the revisionist reality we have tacitly accepted.

Regardless of what form it takes, the Semantic Web is one of our most important next steps in the development of distributed intelligence. And because Web 3.0 changes the Web, it changes everything. In my three decades of involvement in educational technology, this is one of those rare times when we can see a fundamental change coming far enough in advance that we can actually plan for it. That's why the world desperately needs education to join the discussion about what the Semantic Web can be.

Jason Ohler is professor emeritus of educational technology at the University of Alaska, Juneau, and an instructor with the Media Psychology graduate program at Fielding Graduate University. He has been writing, speaking, teaching, and researching in the area of technological change for 25 years.