HARNESSING THE INTERNET INTO A KNOWLEDGE FRAMEWORK

Richard Longworth
Capilano College, Vancouver, Canada rlong@capcollege.bc.ca
iGEN Knowledge Solutions, Vancouver, Canada rlong@igeninc.com

Abstract: This paper identifies a critical problem for learners and teachers: the information overload associated with the Internet, and why it is not an effective tool for learning. The paper identifies two current learning theories, constructivism and multiple intelligences. It suggests a solution based on recent technology advances using XML, DOM, and the RDF framework. The proposed solution gives learners and teachers a way to retrieve information from the Internet, and to create an on-line knowledge framework for all learners..

1. THE PROBLEM: INFORMATION OVERLOAD

Excessive information on the Internet is particularly irksome for educators and students. So much raw material is received that, in most cases, it overwhelms the learner. The result is "negative information which causes the reader to know less because it cannot be integrated, applied, and transformed into knowledge"[4].

Acquiring information is not the same as knowledge. "We can mass produce information through the Internet, but we cannot mass produce knowledge which is created by individual minds "[6]. Jace Hargis identifies two pedagogical short-comings related to the seemingly infinite resources on the Internet. "The first is information overload and lack of useful instructional format, the second is the effective design and evaluation of different learning formats"[2].

A solution is to leverage the mass of information that is available on the Internet to create better knowledge frameworks.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35602-0_35

M. A. Musen et al. (eds.), *Intelligent Information Processing*

2. DEFINING A PEDAGOGICAL FRAMEWORK

In order to develop a new model for knowledge acquisition, we need to acquaint ourselves with current learning theories.

One modern approach is the constructivist framework. It implies that the knowledge received is not passively absorbed but integrated, or constructed, in a unique way by the learner.

Dirk Rodenberg suggests that learning must provide the context defined by the learner's experience, into which new content must be integrated. These cognitive models provide a ground onto which new information is integrated, which in turn extends and reshapes these models. Further, Rodenberg argues that the cognitive load that learners can carry is based on their level of expertise [5].

Thus learning is best accomplished when students are able to think about a topic and to "provide an opportunity of alternate sources, reconciling any differences and forming a reliable coherent individualized model of the topic in question"[5]. The basic premise of constructivism is that the learner must construct knowledge; the teacher cannot necessarily supply it.

3. A POTENTIAL SOLUTION

Some recent attempts to make the web smarter include work by Rob Barrett at the Almaden Research Center (Web Browsing Intelligence), and by Pattie Maes and her colleagues at MIT Multimedia Labs (collaborative filtering and intelligent agents) [3]. The acceptance of this work directs future software towards more personalization and more user-centred applications. Our solution is geared towards personalized services in a framework of XML, DOM and RDF.

XML

XML-based technology provides a means for identifying content. It gives meaning to the information, and if organized correctly, substantially reduces information overload. The XML-world speaks of XML Documents for specific constituents. Since our domain is knowledge, it will be named "XML General Knowledge Document." Here is our XML Document:

XML General Knowledge Document

```
<Knowledge>
  <Subject>
    <Topic>
```

<Title>

<Author>

<Reference>

<Language>

<Date>

<Comprehension>
<Level>

<Readability>

<CognitiveKeys>

<Content>

.... with suitable end tags

Most of the tags (elements) are self-explanatory, however some need annotation.

Knowledge is our root element. The nested elements include

- a Reference that identifies any credible sources such as World Encyclopedia or the Smithsonian Institute
- Language to identify the written content language.

The important Comprehension element contains nested elements to aid the learner at his/her comprehension level

- Level has three possible values: "Introductory" or "Intermediate" or "Advanced"
- Readability is based on either the Gunning-Fogg or Flesch-Kincaid index that identifies a grade level within a statistical norm.
- CognitiveKeys are key words and phrases (such as description, definition, outline, logical explanation) that correlate learning styles with Gardner's seven multiple intelligences [1] (verbal, visual, logical, kinesthetic, rhythmic, interpersonal, intra-personal).

DOM and RDF

The above schema can be enhanced through manipulating DOM objects using XPath and XPointers. We would be particularly interested in XPath to define the direction and destination through the object node, and XPointers to pick out partial content and reduce the excessive information.

RDF is a framework that provides an infrastructure for using and encoding metadata. Our XML General Knowledge Document is a product that would be accessed from the RDF. Since this framework works within the

XML domain, our General Knowledge Document would include a unique URI using a namespace convention.

4. APPLYING THE PROPOSED SOLUTION

4.1 A Curriculum Example

The learner wishes to enhance his/her knowledge of a topic by using the Internet. The learner sends a request identifying the Subject/Topic and any constraints (such as readability level, understanding level, and any cognitive keys) as a personalized request. An application program notifies a Search Engine of the Subject/Topic keyword search. The Search Engine, based on General Knowledge XML documents, finds sites meeting the Subject/Topic criteria. The URL and the XML document are returned and passed to the Application Server. The application program matches the learner's constraints against the XML tags. The Application Server alerts the Web Browser on the user-side of the matching filtered URL's.

The learner can now use the results of the refined search. User feedback is gathered and used for future intelligent agents and for data warehousing in a database server.

5. FUTURE POSSIBILITIES

Potential exists for disadvantaged learners; for personalized learning; for providing motivational resources; and for extending the model with technology enhancements.

Here are a few possibilities:

Adopting an XML Document allows us to use the same information across a variety of media, including speech synthesizers and Braille readers

Clustering on any CognitiveKeys would empower the learner to seek out his/her specific learning styles, for example, learning through "models" or "pictures" or "text"

Clustering on Level and Readability would empower the learner in the construction of a hierarchy of learning

Identifying personal preferences and behaviour patterns opening the potential of software agents

6. CONCLUSION

The problem defined in the introduction is the gap between learning and knowledge, and the difficulty of using the Internet as a meaningful information source. Using a constructivist approach to learning, we have suggested a way to turn the Internet into a learning framework, and a means of reducing information overload. The core solution is the XML General Knowledge Document. This involves the learner selecting keywords based on his/her level of knowledge and prior experiences; the effect is to open the possibility of constructing a hierarchy of learning that links new knowledge to existing patterns. The solution therefore offers repeated and relevant information which matches the experience of the learner, and provides a manner of learning through the Internet which is based on sound pedagogical principles.

7. REFERENCES

1. Gardner, Howard (1995). Multiple Intelligences. Arlington, Illinois: Skylight Training.

2. Hargis, Jace (Summer 2001). Can Students Learn Science Using the Internet? Journal of Research on Computing in Education. Volume 33 #4.

3. Maes Pattie. Agents that Reduce Work and Information Overload. Online: http://pattie.www.media.mit.edu/people/pattie/CACM-94/CACM-94.p1.html

4. Milton, B. B. (1989). Making Sense or Non-Sense: Key Issues in the Information Age. Canadian Vocational Journal 24, no 3 (February 1989).

5. Rodenberg, Dirk (December 1998). Shifting Perspectives in Education Technology. Online: http://horizon.unc.edu/TS/default.asp?show=article&id=59

6. Wurman, Richard (1989). Information Anxiety, New York: Bantam Doubleday Dell Publishing Group.