

WEB NGŨ NGHĨA: NHỮNG THÁCH THỨC VÀ HƯỚNG TIẾP CẬN MỚI

Hoàng Hữu Hạnh

Ban Khoa học Công nghệ, Đại học Huế

TÓM TẮT

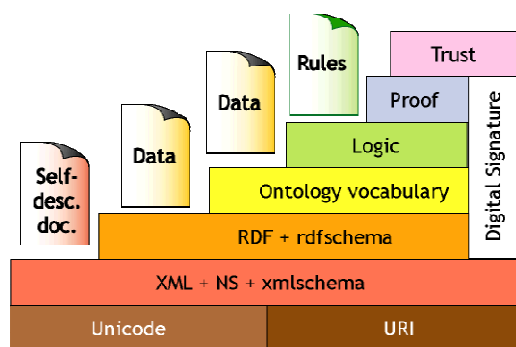
Khái niệm Web Ngữ nghĩa (WebNN–Semantic Web) được Tim Berners-Lee đưa ra năm 2001 [1]. Kể từ đó rất nhiều các nỗ lực trong nghiên cứu cũng như công nghiệp để hiện thực những mục tiêu của WebNN đã được tiến hành [2]. Theo quan điểm ban đầu và các phát biểu có tính chiến lược đó, WebNN sẽ đóng vai trò như một sự bổ sung cho web hiện tại, mà chúng hầu như chỉ con người mới hiểu được nội dung trình bày của nó. Các nhà nghiên cứu WebNN muốn vượt qua giới hạn này và mở rộng sự nhận thức của thông tin web vào các máy tính. WebNN cho phép biểu diễn thông tin của World Wide Web (WWW hay vắn tắt là Web), cơ sở dữ liệu và các kho dữ liệu có cấu trúc khác theo một thể thức thống nhất mà nó được hiểu và xử lý bởi các máy tính. Bài báo này điểm lại vắn tắt về thực trạng nghiên cứu WebNN cũng như vạch ra các hướng nghiên cứu mới của WebNN trong tương lai gần.

1. Một số khái niệm trong web ngữ nghĩa

1.1. Web Ngữ nghĩa

Theo định nghĩa của Tổ chức World Wide Web (W3C)¹, WebNN được hiểu như sau:

“WebNN là một cách nhìn về cách thức tổ chức dữ liệu: đó là ý tưởng về việc dữ liệu trên Web được định nghĩa và liên kết theo một cách mà nó có thể được sử dụng bởi máy tính với mục đích không chỉ cho việc hiển thị mà còn tự động hoá, tích hợp và sử dụng lại dữ liệu qua các ứng dụng khác nhau”.



Hình 1. Kiến trúc phân tầng của WebNN được đưa ra năm 2001

¹ <http://www.w3.org>

Kiến trúc phân tầng của WebNN được mô tả trong Hình 1. Trong đó, các tầng trên kế thừa các tầng thấp hơn với cơ sở là các chuẩn để mô tả siêu dữ liệu và các tài liệu bán cấu trúc. Do giới hạn của khuôn khổ bài báo, nên những giải thích chi tiết không được trình bày trong bài báo này. Những trình bày liên quan đến nó có thể tìm đọc ở tài liệu [3].

1.2. Khung Mô tả Tài nguyên

Khung mô tả tài nguyên (RDF) [4] là một ngôn ngữ siêu dữ liệu để biểu diễn dữ liệu trên Web và cung cấp một mô hình để mô tả và tạo các mối quan hệ giữa các tài nguyên. RDF định nghĩa một *nguồn tài nguyên* (resource) như một đối tượng bất kỳ có khả năng xác định duy nhất bởi một URI². Các nguồn tài nguyên có các thuộc tính đi kèm. Các *thuộc tính* (predicate/property) được xác định bởi các kiểu thuộc tính và các kiểu thuộc tính có các giá trị tương ứng. Kiểu thuộc tính biểu diễn các mối quan hệ của các giá trị được kết hợp với các tài nguyên.

Mô hình dữ liệu của RDF là các bộ ba (triple) gồm: $\langle \text{Chủ-thể}, \text{Thuộc-tính}, \text{Đối-tượng} \rangle$

- *Chủ-thể* (Subject): được xác định bởi URI cụ thể.
- *Thuộc-tính* (Predicate): thuộc tính của siêu dữ liệu, cũng được xác định bởi một URI.
- *Đối-tượng* (Object): giá trị của thuộc tính, có thể là một giá trị (literal) hoặc một URI.

Ví dụ: Dữ liệu về tên họ của một cá nhân có mã số (ID) xác định

Chủ-thể	Thuộc-tính	Đối-tượng
http://www.hueuni.edu.vn/StaffID/1234	ns:hasFullName	Hoang Huu Hanh

được biểu diễn như Hình 2:



Hình 2. Đồ thị RDF biểu diễn ví dụ trên

1.3. Ontology

Thuật ngữ Ontology bắt nguồn từ triết học, nó được sử dụng như tên của một lĩnh vực nghiên cứu về sự tồn tại của tự nhiên, xác định các vật thể trong tự nhiên và làm thế nào để mô tả chúng. Chẳng hạn như quan sát thế giới thực, xác định các đối tượng và sau đó nhóm chúng lại thành các lớp trừu tượng dựa trên thuộc tính chung [3].

Tuy nhiên, trong những năm gần đây, Ontology đã trở thành một thuật ngữ được biết đến nhiều trong lĩnh vực khoa học máy tính và có ý nghĩa khác xa so với nghĩa ban

² Uniform Resource Identifier: Định danh Tài nguyên đồng nhất

đầu của nó. Ontology được xem như là “linh hồn” của WebNN. Chúng giúp con người và máy có thể hợp tác, cùng nhau làm việc, giúp máy có thể “hiểu” và có khả năng xử lý thông tin hiệu quả. Các Ontology được phát triển trong lĩnh vực trí tuệ nhân tạo để sử dụng lại và chia sẻ tri thức được thuận tiện hơn [3, 5, 6]. Đầu những năm 1990, Ontology trở thành một chủ đề nghiên cứu phổ biến. Ontology được nghiên cứu bởi một số cộng đồng nghiên cứu trí tuệ nhân tạo, bao gồm kỹ sư tri thức, xử lý ngôn ngữ tự nhiên và biểu diễn tri thức.

Một định nghĩa chung cho ontology là: *Ontology là một đặc tả hình thức của sự khái niệm hóa về một lĩnh vực ứng dụng cụ thể* [3]. Định nghĩa này nhấn mạnh hai điểm chính: sự khái niệm hóa (conceptualisation) là hình thức và do đó cho phép suy diễn bởi máy tính; và một ontology trên thực tế được thiết kế cho một miền ứng dụng cụ thể nào đó. Các ontology bao gồm các khái niệm (các lớp - classes), các quan hệ (các thuộc tính - properties), các thể hiện (instances) và các tiên đề (axioms).

2. Web ngữ nghĩa trong ngữ cảnh của hôm nay

2.1. Web của hôm nay và Web Ngữ nghĩa

Ngày nay, thông tin trên World Wide Web (WWW) chủ yếu được biểu diễn ở dạng HTML³ (Hyper-Text Markup Language), một ngôn ngữ phổ dụng để trình diễn thông tin. XML⁴ (Extensible Markup Language) ra đời và trở thành một công cụ trao đổi dữ liệu không có cấu trúc, bán cấu trúc và có cấu trúc giữa các hệ thống, nâng cao sự tích hợp của các ứng dụng. Tuy nhiên, các giải pháp dựa trên XML cho quá trình tích hợp của các ứng dụng và các hệ thống chưa đủ, do dữ liệu được chuyển đổi thiếu mô tả tường minh về ngữ nghĩa của nó. Sự tích hợp của các ứng dụng cũng phải bao gồm sự tích hợp cả về ngữ nghĩa.

WWW hiện tại có thể được xem như là một cơ sở dữ liệu toàn cầu lớn nhất tuy nhiên sự tồn tại về các cấu trúc ngữ nghĩa để giữ cho các thành phần của nó có quan hệ lẫn nhau mà vẫn mang tính độc lập là rất hạn chế. Điều này làm cho thông tin hiện có trên WWW hầu như chỉ cho con người hiểu và xử lý. WebNN cung cấp một số ngôn ngữ nhằm biểu diễn dữ liệu và thông tin theo định dạng mà máy tính có thể xử lý được.

Các phát triển gần đây của công nghệ thông tin và truyền thông đã tạo ra những khả năng để thu thập được một số lượng lớn dữ liệu mà chúng có liên quan đến nhau về mặt khái niệm, tuy nhiên, đa số những mối quan hệ này chỉ được con người “nhớ” chứ không được lưu trữ theo một cách mà giúp các máy tính có thể hiểu để xử lý. Thách thức này đã chỉ ra một hướng nghiên cứu đó là tạo ra khả năng cho phép con người tạo, lưu giữ, sắp xếp, ghi phụ chú và truy xuất kho dữ liệu cá nhân rất lớn của mỗi người

³ <http://www.w3.org/MarkUp/>

⁴ <http://www.w3.org/XML/>

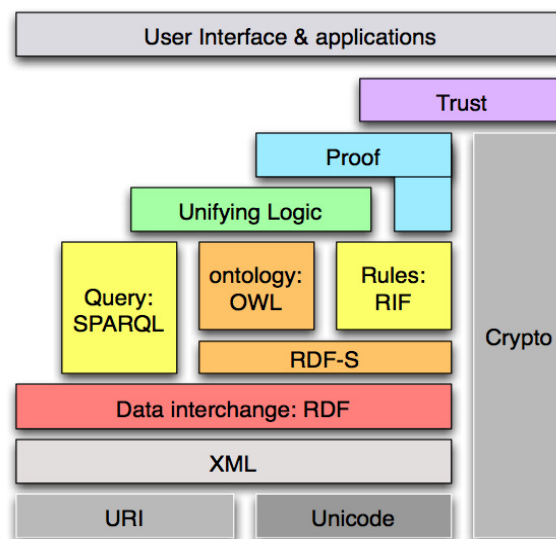
trong quá khứ theo hình thức như một nhật ký cuộc sống được cá thể hoá và sẽ trở thành một sự bổ sung và trợ giúp cho bộ nhớ con người [7].

WebNN hứa hẹn cung cấp một mô hình để liên kết những nguồn thông tin khác nhau, chẳng hạn như những trang web, các cơ sở dữ liệu hay ngay cả luồng dữ liệu và thông tin trong cuộc sống hàng ngày. Các ứng dụng của WebNN trong các lĩnh vực khác nhau đã thu hút nhiều nhóm nghiên cứu. Do đó, các ứng dụng được thiết kế dựa trên các khái niệm đã đề cập và sử dụng các thông tin có thể xử lý được bởi máy tính để tạo ra động lực lớn cho việc phát triển của một thể hệ các công cụ và các ứng dụng Web mới.

2.2. Web Ngữ nghĩa trong ngữ cảnh hôm nay

Trên cơ sở tầm nhìn ban đầu về WebNN, rất nhiều thành phần của kiến trúc nguyên thủy của WebNN (Hình 1) đã được giải quyết, cũng như các ca sử dụng (use-case) của các ứng dụng ngữ nghĩa đang được tiến triển hơn như được mô tả trong kiến trúc của WebNN được hoàn thiện hơn cho đến năm 2006 như trong Hình 3.

Như được mô tả trong Hình 3, các tầng của kiến trúc WebNN đã được quy chuẩn với các chuẩn đã được W3C đề xuất cũng như cộng đồng nghiên cứu WebNN thống nhất sử dụng trên thực tế (de facto). Theo đó, dữ liệu trong WebNN dựa cơ sở trên XML, và được mô hình hoá bằng RDF. RDF cũng được chọn là chuẩn trao đổi dữ liệu trong WebNN. Ngôn ngữ Ontology được chuẩn hoá là OWL dựa trên cơ sở của RDFS; trong khi đó, chuẩn ngôn ngữ cho các luật vẫn đang còn chưa đồng nhất được cho đến bây giờ. Bên cạnh đó, ngôn ngữ truy vấn SPARQL [8] đã được sử dụng rộng rãi và là khuyến nghị của W3C, tuy nó chưa trở thành chuẩn thật sự. Ở ba tầng trên vẫn đang là những vấn đề cần được nghiên cứu trong tương lai.



Hình 3. Kiến trúc phân tầng của WebNN được hoàn thiện năm 2006

Trong suốt vài năm trở lại đây, các công nghệ liên quan đến WebNN đã xuất hiện hoặc đã được hiệu chỉnh nhiều. Một trong những phần quan trọng của những tiến triển này là các ngôn ngữ phát triển ontology. Trong thời gian đầu, hàng loạt các ngôn ngữ đã đua nhau xuất hiện, tuy nhiên chúng đã ổn định thời gian gần đây với các chuẩn đã được chấp thuận. Tổ chức W3C đã làm việc tích cực để chuẩn hoá các chuẩn và đã phê chuẩn RDF và ngôn ngữ ontology OWL [9, 10] là hai chuẩn quan trọng nhất trong số các công nghệ WebNN. Điều này đã cung cấp một cơ sở vững chắc để triển khai các ứng dụng WebNN trên diện rộng, cũng như làm đòn bẩy cho việc đưa WebNN từ nghiên cứu sang ứng dụng trong công nghiệp cũng như các chuẩn cho nó để xây dựng các thể hệ ứng dụng mới.

So sánh với tình hình vài năm trước đây, các công cụ để tạo và xuất bản thông tin mang ngữ nghĩa đã bùng phát mạnh mẽ và điều này giúp cho những người không chuyên trong việc ứng dụng công nghệ WebNN có thể sử dụng vào lĩnh vực của riêng họ. WebNN đã mở ra một cái nhìn mới cho hàng loạt ứng dụng và hệ thống mà chính chúng đã hưởng lợi từ việc thông tin có thể được máy tính hiểu và xử lý được.

3. Công nghệ Ontology và Ontology tham chiếu

Các ontology là các thành phần cơ bản để xây dựng các ứng dụng dựa trên WebNN. Việc tạo ra các ontology không phải là việc đơn giản, và rõ ràng là không có một ontology đúng duy nhất cho bất kỳ một miền ứng dụng nào. Giá trị thực tế của một ontology chỉ có thể được đánh giá bằng các ứng dụng thực tế sử dụng nó. Các ontology được xây dựng với các miền ứng dụng cụ thể chính là những viên gạch đặt nền móng cho các ứng dụng dựa trên ngữ nghĩa đang xuất hiện ngày càng nhiều và nhanh chóng làm cho những giấc mơ về WebNN toàn cầu đang trở thành sự thật.

Hiện tại còn có rất nhiều các vấn đề quan trọng khác khi đề cập đến mảng công nghệ ontology; một trong số các hướng này có thể kể đến như: tích hợp ontology (ontology integration), ánh xạ ontology (ontology mapping), tái sử dụng ontology (ontology reuse), và kiểm tra tính chặt chẽ của ontology (ontology consistency). Song song với các vấn đề này, một số điểm vẫn đang còn bỏ ngỏ, chẳng hạn các ontology hạt nhân (kernel) dùng để tham chiếu trong các miền ứng dụng khác nhau nên được tạo ra để tạo sự thuận lợi cho các mở rộng có ý nghĩa của chúng. Việc tạo ra các ontology tham chiếu trong mỗi lĩnh vực như vậy sẽ tránh không bị “tái đầu tư” để tạo lại các lược đồ, URI và các ontology đã tồn tại.

Ngay cả khi chúng ta có một tập đầy đủ các ontology đó thì tiến trình công nghệ ontology vẫn được tiếp tục, và vấn đề chúng ta phải đối phó đó là “tuổi tác” của ontology (ontology aging). Các kết quả được trích rút từ một ontology không cập nhật không thể được sử dụng một cách hoàn toàn có ý nghĩa. Ta nên có một vài cơ chế để phát hiện vấn đề “tuổi” của ontology và buộc các hệ thống dựa trên ngữ nghĩa phải thay đổi các ontology theo các thay đổi của các tham số môi trường. Nói một cách khác, các ontology nên “nhận thức” được “ý niệm thời gian” và thích ứng với tiến trình mà nó

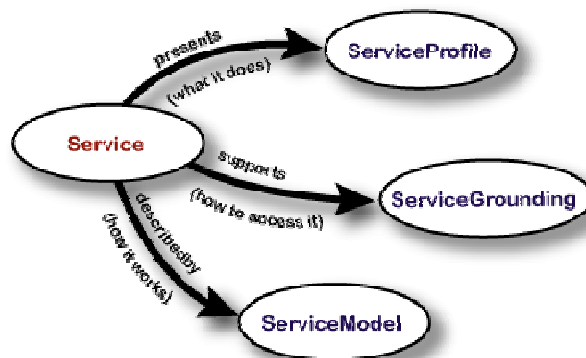
tham gia vào, tức là các ontology nên phát triển dựa trên các tham số môi trường như thời gian. Do vậy, các yêu cầu trong tương lai trong lĩnh vực công nghệ ontology ít nhất cũng bao gồm các chủ đề sau:

- Sử dụng lại, ánh xạ và tích hợp ontology;
- Phát triển các ontology hạt nhân chuẩn trong các lĩnh vực khác nhau;
- Tích hợp các ontology trong các tiến trình và có sử dụng ý niệm về thời gian.

4. Các dịch vụ Web ngữ nghĩa

Các dịch vụ Web⁵ (Web Services - WS) cung cấp một nền tảng cho kiến trúc phần mềm mới được gọi là *Kiến trúc hướng Dịch vụ* (Service-Oriented Architecture – SOA). Kiểu kiến trúc này đã phát triển một cách nhanh chóng và trở thành như một giải pháp cho những khó khăn trong việc phát triển phần mềm như: phần mềm sử dụng lại, lập trình phân tán và các môi trường ứng dụng đa dạng. Đặc biệt, Ngôn ngữ Mô tả Các dịch vụ Web (WSDL) [11] tiếp tục là một phần quan trọng của WS đã được chuẩn hóa bởi W3C.

WS cho phép chúng ta truy cập các ứng dụng liên quan, cho dù vậy các quá trình như khám phá (discovery), biên tập (composition) và thực thi (execution) WS vẫn cần sự hỗ trợ từ sự tương tác của con người. Đây chính là điểm mà ta cần WebNN để hỗ trợ WS với các ontology như là các giá trị gia tăng (Hình 4). Việc kết hợp sức mạnh của WS và các giá trị gia tăng của WebNN sẽ cho một nền tảng cụ thể cho các ứng dụng doanh nghiệp.



Hình 4. *Ontology được sử dụng để mô tả ngữ nghĩa của dịch vụ*

Hiện có rất nhiều hoạt động diễn ra song song để hiện thực hóa mục tiêu này. Chương trình Các Công nghệ cho Xã hội Thông tin (Information Society Technologies) của Ủy ban Châu Âu⁶ đã khởi động nhiều dự án về WS sử dụng các khái niệm WebNN.

⁵ <http://www.w3.org/2002/ws/>

⁶ <http://cordis.europa.eu/ist/>

Một số các vấn đề thách thức cần được các dự án này thực hiện là:

- Tạo lập các ontology cho việc mô tả và phân loại WS;
- Sự tin cậy và chứng thực các dịch vụ web ngữ nghĩa (Semantic Web Services);
- Biểu diễn tri thức cho các dịch vụ web ngữ nghĩa;
- Ngữ nghĩa cho sự ủy thác dịch vụ và kết hợp tri thức.

5. Vấn đề đa ngôn ngữ

Tính đa ngôn ngữ được nhìn nhận như là một trong những thách thức quan trọng nhất của WebNN. Hàng triệu người sử dụng Internet thường muốn nội dung mà họ yêu cầu được thể hiện bằng ngôn ngữ của họ và điều này có thể sinh ra khả năng hoặc là tạo các công cụ để hỗ trợ việc phụ chú (annotation) các nội dung theo các ngôn ngữ khác nhau; hoặc là khởi tạo các bộ dịch giữa những ontology liên quan. Bộ dịch này ánh xạ các ontology và nội dung sang một ngôn ngữ khác.

Một vấn đề quan trọng hơn nữa đó là xử lý với việc thích ứng với các ontology đã phát triển trước đó với các yêu cầu văn hóa của các vùng miền khác nhau. Các dân tộc khác nhau sẽ có lối suy nghĩ về một khái niệm dù đơn giản cũng khác nhau, và điều này có căn nguyên từ gốc của văn hóa và lịch sử. Như K. Veltman đã nói trong bài báo “Challenges for a Semantic Web - Những Thách thức của WebNN” của ông: “WebNN, như khi nó khởi sinh, phản ánh một cách tuyệt diệu nhu cầu của khoa học và công nghệ hiện đại. Nhưng nó vẫn không đáp ứng được những yêu cầu quá phức tạp về khía cạnh văn hóa. Một số người có thể tranh cãi rằng điều này là không quan trọng và khá xa xỉ. Trong một thế giới mà các đặc tính thiên cận của các nhóm chính thống đang đe dọa tính cấu trúc của xã hội, nhu cầu về những bản sắc đa chiều trở thành hy vọng duy nhất cho sự tồn tại lâu dài như một nền văn minh” [12].

Các yêu cầu trong tương lai liên quan đến các vấn đề quốc tế hoá sẽ phải bao gồm:

- Phát triển các công cụ cho phép người sử dụng đầu cuối phát triển các ontology của chính họ;
- Đưa ra được những ánh xạ và tính liên thông nhau của các ontology theo các ngôn ngữ khác nhau;
- Tích hợp ý niệm văn hóa vào các ontology quốc gia.

6. Các ứng dụng của Web ngữ nghĩa

WebNN cho phép các ứng dụng làm cầu nối giữa những nhu cầu của người sử dụng và các tài nguyên thông tin sẵn có. Điều này sẽ thiết lập một cơ sở cho việc phát triển các ứng dụng. Trong thời gian gần đây, các tiếp cận WebNN đã trở nên ổn định hơn và các thành phần thiết yếu đã được cung cấp theo các chuẩn. Bước tiếp theo sẽ là ứng dụng công nghệ đang phát triển mạnh này vào cuộc sống hàng ngày của chúng ta.

Một số thử nghiệm các ứng dụng dựa trên ngữ nghĩa đầu tiên đã được cài đặt thành công và rất nhiều các ứng dụng nữa đang được phát triển. Do miền ứng dụng rộng của WebNN, nên ta có thể có được rất nhiều các ca sử dụng của WebNN cho các lĩnh vực khác nhau. Hầu hết các nhà nghiên cứu WebNN đều có chung một thoả thuận ngầm với nhau để tránh nhắc đến những điều thuộc về trí tuệ nhân tạo (TTNT). Dù rằng WebNN có thể rơi vào một vài vấn đề của TTNT nhưng một điều chắc chắn rằng nó không phải là một thể hệ kế tiếp của TTNT. Chúng ta không thể trông đợi WebNN sẽ là chìa khoá cho mọi vấn đề. Một giải pháp tốt của WebNN có được bởi nguồn gốc chắc chắn của nó là một ontology được thiết kế tỉ mỉ.

Một số ứng dụng mang tính thách thức mà chúng có thể làm thay đổi cách mà chúng ta sử dụng máy tính:

- “Các bộ nhớ Cá nhân Tăng cường” là một ứng cử viên tiềm năng cho các ứng dụng WebNN [13]. Đã có rất nhiều nỗ lực của các nhóm nghiên cứu, các dự án cho vấn đề này chẳng hạn như [14], [15], and [16].
- Tích hợp tri thức của thế giới vào máy tính desktop (semantic desktop) theo một cách có thể truy nhập dễ dàng và hiệu quả [17].
- Tích hợp các dịch vụ web ngữ nghĩa vào các ứng dụng thường được sử dụng (các trình thư điện tử, các trình duyệt web,...) để cung cấp nhiều hơn tính hoạt động liên thông trên web cũng như giữa các ứng dụng [18, 19].
- Ứng dụng WebNN vào thương mại điện tử, đặc biệt là các nghiên cứu liên quan đến Quản lý tiến trình nghiệp vụ (Business Process Management) trong thương mại điện tử và B2B (business-to-business). Đây là một trong những ứng dụng mới của WebNN nhằm làm giàu ngữ nghĩa cho các tiến trình nghiệp vụ trong môi trường cộng tác doanh nghiệp [20, 21], [22].

7. Tổng kết

Sự tiến triển của WebNN đã mở ra một khung nhìn mới trong CNTT và đặc biệt trong các lĩnh vực về công nghệ dữ liệu. Trong suốt vài năm qua hầu hết các chuẩn và đặc tả liên quan đến WebNN đã được sửa đổi tỉ mỉ và nó nhanh chóng thu hút được sự chú ý của nền công nghiệp CNTT trong việc phát triển các giải pháp doanh nghiệp dựa trên WebNN.

Tuy nhiên cũng nên nhắc lại rằng WebNN có thể không hứa hẹn một sự thành công nhanh với việc định dạng dữ liệu phù hợp với nó. Berners-Lee, Giám đốc của W3C đã xác nhận khái niệm đó là “khá khó để giải thích”. Tuy nhiên, ông cũng đã kinh qua vấn đề tương tự khi đã cố gắng giải thích về World Wide Web 18 năm trước: “Các trang siêu văn bản: vấn đề lớn! mọi người đã nói như vậy. Họ không thể nhìn nhận ra được làm sao chúng có thể liên kết đến hầu như mọi thứ và ý nghĩa của nó là gì?” [23].

TÀI LIỆU THAM KHẢO

1. T. Berners-Lee, J. Hendler, and O. Lassila, *The Semantic Web*, Scientific American, vol. 284, (2001), 28-37.
2. W3C, *The Semantic Web Activity*, (2001).
3. G. Antoniou and F. Van Harmelen, *A Semantic Web Primer*, MIT Press, 2004.
4. W3C, *Resource Description Framework (RDF)*, in Semantic Web, 2004.
5. T. B. Passin, *Explorer's Guide to the Semantic Web*, Manning, 2005.
6. D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster, *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, MIT Press, 2003.
7. W. Wahlster, *Grand Challenges in the Evolution of the Information Society*, Technical Report, 2004.
8. E. Prud'hommeaux and A. Seaborne, *SPARQL Query Language for RDF*, W3 Recommendation, January 2008.
9. D. L. McGuinness and F. van Harmelen, *OWL Web Ontology Language Overview*, W3C Recommendation, February 2004.
10. W3C, *Web Ontology Language (OWL)*, 2004.
11. E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana, *Web Services Description Language (WSDL) 1.1*, 2001.
12. K. Veltman, *Challenges for a Semantic Web*, Cultivate Interactive, vol. 7, 2002.
13. A. Fitzgibbon and E. Reiter, *Memories for Life - Managing Information over a Human Lifetime*, Grand Challenges in Computing Workshop, UK Computing Research Committee, May 2003.
14. D. R. Karger, K. Bakshi, D. Huynh, D. Quan, and V. Sinha, *Haystack: A General Purpose Information Management Tool for End Users of Semistructured Data*, Proceedings of 2nd Biennial Conference Innovative Data Systems Research Asilomar, CA, USA, (2005), 13-26.
15. J. Gemmel, G. Bell, and R. Lueder, *MyLifeBits: Living With a Lifetime Store*, in *ATR Workshop on Ubiquitous Experience Media*, Kyoto, Japan, 2003.
16. M. Ahmed, H. H. Hoang, S. Karim, S. Khusro, M. Lanzenberger, K. Latif, E. Michlmayr, K. Mustofa, T. H. Nguyen, A. Rauber, A. Schatten, T. M. Nguyen, and A. M. Tjoa, *SemanticLIFE - A Framework for Managing Information of A Human Lifetime*, in Proceedings of the 6th International

- Conference on Information Integration and Web-based Applications and Services Jakarta, Indonesia, 2004.
17. CEUR, *The Semantic Desktop Search*, 2005.
 18. J. Hendler, *An interview with James Hendler*, AIS SIGSEMIS Bulletin, vol. 1, (2004), 2-4.
 19. W. König, *Interview with James Hendler on the Semantic Web*, Wirtschaftsinformatik, vol. 44, (2002), 481-483.
 20. D. E. Jenz, *Business Process Ontologies: Speeding up Business Process Implementation*, Jenz & Partner GmbH, 2003.
 21. D. E. Jenz, *Ontology-Based Business Process Management: The Vision Statement*, Jenz & Partner GmbH, Erlensee, Germany, 2003.
 22. SUPER-Project, *Semantics Utilized for Process Management within and between Enterprises*, 2006.
 23. T. Berners-Lee, *Fourth Annual Bio-IT World Conference*, 2005.

THE SEMANTIC WEB RESEARCH: CHALLENGING ISSUES AND NEW DIRECTIONS

*Hoang Huu Hanh
Hue University*

SUMMARY

According to initial vision and strategy statements, the Semantic Web will play the role of a complement for the current web which is mostly understandable by human users. Semantic Web researchers intend to overcome this limitation and extend the recognition of the web information to machines by making it processable. Currently, the Semantic Web has been used in many fields of ICT and e-commerce, and it has inspired researchers in finding a new paradigm for a new generation of applications. This paper provides a view on challenges in the Semantic Web research as well as the state of the art of applications and research directions in the field.