

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



TIỂU LUẬN
MẠNG MÁY TÍNH VÀ TRUYỀN DỮ LIỆU NÂNG CAO

CHỦ ĐỀ
CÔNG NGHỆ WEB THẾ HỆ MỚI – WEB 3.0

Giáo viên hướng dẫn : TS Phạm Thế Quế
Học viên thực hiện : Đàm Thị Thanh Hằng
Mã Học Viên : CK15H.05
Lớp : CK15H

Hà Nội, 24 tháng 5 năm 2017

LỜI NÓI ĐẦU

World Wide Web gọi tắt là Web hay WWW là mạng thông tin toàn cầu chứa một lượng thông tin khổng lồ, được tạo ra từ các tổ chức, doanh nghiệp và nhiều cá nhân. Người sử dụng truy cập Web bằng URL và theo các liên kết để tìm ra tài nguyên liên quan. Tuy nhiên tính đơn giản của Web dẫn đến một số hạn chế, như một lượng thông tin rất lớn, không hợp lý và không liên quan được trả về khi tìm kiếm trên web.

Mặt khác, các ứng dụng Web phát triển dựa trên mô hình Client/Server. Mô hình này có ưu điểm nổi bật là mọi xử lý sẽ đều thực hiện trên Server, do đó tránh cho Client những tính toán nặng nề. Tuy nhiên, khi số lượng Client tăng thì nhu cầu tải và băng thông cũng tăng, dẫn đến việc Server không có khả năng cung cấp dịch vụ. Mô hình Client/Server không đáp ứng khi nhu cầu truy nhập Internet tăng. Đặc biệt khi IPv4 chuyển sang IPv6.

Câu hỏi đặt ra là làm thế nào có được kết quả tìm kiếm nhanh chóng và chính xác theo mong muốn ?

Để khắc phục các yếu điểm của Web hiện tại, khái niệm “Semantic Web” hay còn gọi là công nghệ Web 3.0 đã ra đời.

Trong bài tiểu luận này được chia thành 4 nội dung chính:

PHẦN I: TỔNG QUAN VỀ WEB

PHẦN II: CÔNG NGHỆ WEB 3.0

PHẦN III: LỢI ÍCH KHI SỬ DỤNG WEB 3.0 VÀ NHỮNG KHÓ KHĂN, THÁCH THỨC GẶP PHẢI.

PHẦN IV: KẾT LUẬN VÀ DỰ ĐOÁN TƯƠNG LAI

PHẦN I

TỔNG QUAN VỀ WEB

Trong những năm đầu thập kỷ 90 của thế kỷ trước, khái niệm Web được đưa ra để mô tả sự phát triển của một hệ thống các trang Web được xây dựng trên cơ sở mở rộng sự tham gia của cộng đồng.

World Wide Web là phương thức giao tiếp giữa con người với nhau trên Internet. Nó có vai trò to lớn trong cuộc cách mạng hướng tới một nền kinh tế tri thức, xã hội tri thức. Thông tin Web trình bày dưới dạng không cấu trúc, các tài liệu Web được lưu trữ trong một hệ thống siêu văn bản (Hypertext) trên Internet – được gọi là Website. Các hoạt động truy tìm thông tin theo các siêu liên kết được gọi là duyệt Web. Các công nghệ Web sử dụng đó là Web 1.0 và Web 2.0

Web 1.0 là thời kỳ của Netscape. Trong đó những người sử dụng thông tin đóng vai trò thụ động và sử dụng thông tin. Thông tin trên website có nội dung tĩnh. Mức tương tác giữa những người cung cấp và người sử dụng thông tin bị hạn chế. Các nhà cung cấp thông tin tập trung và công khai thông tin là chính. Thông tin trên Web 1.0 được tổ chức bởi taxonomies, là hệ thống phân loại từ trên xuống được thiết kế bởi những chuyên gia.

Web 2.0 có những đặc điểm:

- Công nghệ Web 2.0 đang sử dụng web như nền tảng: không yêu cầu cài đặt nhiều phần mềm trên phía người sử dụng- đó là mô hình lập trình nhẹ. Để làm được điều này bạn cần một chuẩn mở và giải pháp định hướng dịch vụ chứ không phải là việc lập trình theo yêu cầu và lắp ráp các dịch vụ. Việc tổ chức dữ liệu, video, hình ảnh và âm nhạc là những ví dụ. Một vài ứng dụng Web 2.0 là phát triển bên dưới và không theo các chu kỳ sản xuất phần mềm chuẩn. Điều này cũng liên quan đến thực tế mà họ không thể cài đặt vào các ứng dụng desktop nhưng phần mềm được cung cấp như dịch vụ qua Web.

Hơn nữa, phần mềm độc lập với từng thiết bị luôn là một đặc điểm của Web nhưng với sự đa dạng lớn về các thiết bị được xảy ra và một vài người truy cập Internet không qua PCs thì việc hiển thị trở nên quan trọng hơn, ví dụ các thiết bị di động cầm tay. Một yếu tố quan trọng khác là các dịch vụ web được truy cập ngày nay không phải từ các Web clients (trình duyệt) mà còn từ phần mềm khác mà không định hướng cơ bản cho Web. Ví dụ như phần mềm iTunes từ Apple. Nó là một ứng dụng desktop nhưng nó được sử dụng để giao tiếp với thiết bị bên ngoài (iPhone hoặc iPod) và nó sử dụng RSS để có được thông tin. Sự kết hợp này thể hiện rõ ràng một thiết bị độc lập quan trọng đối với các ứng dụng Web 2.0.

- Tập trung vào việc kết nối con người chứ không phải máy tính: ở Mỹ những công cụ làm việc này được mô tả như “*folksonomies*” là các công cụ và kỹ thuật để tạo cho các wiki, blogs cũng như sử dụng *tagging* và *feeds*, mà tự động giúp tham gia trong mạng để chia sẻ những liên kết ưa thích. *Folksonomies* được phân biệt với *taxonomy* bằng việc sử dụng tags để phân loại một cách linh hoạt; *taxonomies* kiểm soát việc phân loại và tổ chức các khái niệm có liên quan với nhau. Một loại ứng dụng của Web 2.0 được gọi là *tag clouds* (hoặc *word clouds*) hoặc *tag maps*. Chúng tạo thành biểu diễn các từ trong một tài liệu hoặc các tags của trang (tập hợp các trang) và được sử dụng trong các trang tin tức để đưa ra tổng quan về những tin tức quan trọng nhất của ngày hoặc của một chủ đề chính. Dưới đây là một ví dụ về tag cloud cho Web 2.0. Một vài khái niệm và đặc điểm quan trọng là bản đồ và cỡ tương ứng với độ quan trọng hoặc thường xuyên sử dụng trong ngữ cảnh của Web 2.0.

- Làm giàu kinh nghiệm người sử dụng: Một trong những hứa hẹn và công nghệ sử dụng rộng rãi là AJAX (Asynchronous JavaScript and XML), có thể được mô tả tốt nhất bởi bản đồ Google và những bất đồng bộ tối thiểu

khi bạn di chuyển xung quanh hành tinh. Một vài dự án phát triển cho phép ‘*mash*’.

- Khai thác trí thông minh tập thể: các giải thuật, công nghệ *parsers* và ‘*mash up*’ được yêu cầu – Giải pháp thông minh thương mại được phát triển đồng thời với hệ thống thông minh để giám sát, bắt video, nhận dạng mẫu và phân giải mẫu. *Mashup* hay là kết hợp hai dữ liệu và/hoặc chức năng từ hai hay nhiều nguồn bên ngoài để tạo dịch vụ mới. Vài ví dụ tồn tại về những người phát triển kết hợp các dịch vụ, ví dụ, Google Maps (dịch vụ nhà đất, v.v.). API của Google có thể được sử dụng. Các công cụ chuyển đổi ngôn ngữ được cải thiện nhiều hơn và sẽ trở thành đặc điểm bình thường trong một vài năm tới. Bên cạnh đó, việc tham chiếu thực tế dữ liệu được tăng cường qua phần mềm và việc sở hữu dữ liệu sẽ xác định ai giám sát ứng dụng và có thể phân phối sản phẩm tốt nhất. Sở hữu dữ liệu có thể được tăng cường nhiều hơn qua kiến trúc tham gia mà cho phép người sử dụng để làm dữ liệu có giá trị hơn qua tương tác xã hội và đầu vào.

Web 2.0 là thời kỳ của Google và *weblogs* đóng vai trò quan trọng. Ở thời điểm này, người sử dụng thông tin đóng vai trò tích cực. Họ không chỉ sử dụng thông tin, mà còn thêm và chia sẻ thông tin. Wikis đóng vai trò quan trọng để phát và chia sẻ thông tin. Như vậy, người sử dụng cũng có thể trở thành người sản xuất. Do đó, mà các ứng dụng web trở nên khác nhau nhờ việc sử dụng. Web 2.0 là hệ thống Web động, được phát triển dựa trên ngôn ngữ lập trình và hệ cơ sở dữ liệu SQL. Nó hoạt động trong môi trường Client/Server và cung cấp các dịch vụ và giao tiếp lập trình ứng dụng, có tính mềm dẻo, có thể đọc và viết, có mức tương tác trên web là cao. Chúng ta khó có thể phân biệt giữa người sử dụng và người tạo thông tin. Web 2.0 nổi bật hơn Web 1.0, là một tầng mới về phát triển kỹ thuật và không thay thế Web 1.0. Web 2.0 được hiểu như pha mới nhất trong cải tiến của Internet và Web.

Tuy nhiên Web 2.0 cũng xuất hiện một số nhược điểm như

- Khi lượng thông tin trên Internet tăng, đồng nghĩa với việc tìm kiếm, khai thác, tổ chức và duy trì thông tin ngày càng trở nên khó khăn hơn.

- Độ chính xác và chứng thực của thông tin trong quá trình lướt Web thường không chính xác.

Nguyên nhân:

Dữ liệu Web thiếu ngữ nghĩa, máy không hiểu. Máy chỉ là công cụ xử lý thông tin cho các ứng dụng

- Tìm kiếm thông tin bằng cách gõ các từ hoặc cụm từ (Keyword) truy vấn. Nó sẽ trả về danh sách kết quả các trang web liên quan đến cụm từ truy vấn. Đa số các máy tìm kiếm trên Web truyền thống đều tìm kiếm dựa trên việc đối sánh với từ khóa.

- Tài nguyên trên Web chủ yếu phi cấu trúc. Vì vậy, có rất nhiều kết quả không mong muốn trả về. Các nhà cung cấp cũng đã đề xuất nhiều giải pháp nhằm cải thiện hiệu quả tìm kiếm, tuy nhiên, kết quả tìm kiếm vẫn chưa thỏa mãn được người dùng.

- Xét trên phương diện người dùng, các máy tìm kiếm truyền thống gặp phải hai vấn đề chính:

+ Mỗi từ khóa có thể có nhiều nghĩa. Máy tìm kiếm không hiểu nghĩa mối quan hệ giữa các từ khóa.

+ Rất nhiều trang Web có liên quan đến từ khóa nhưng nội dung lại không chứa từ khóa. Vì vậy kết quả tìm kiếm là rỗng, vì dữ liệu Web không được tổ chức theo ngữ nghĩa

- Các máy tìm kiếm truyền thống dựa trên từ khóa có độ chính xác còn thấp. Phù hợp các ứng dụng cô lập.

Xét ví dụ: *Tìm thông tin cá nhân của một người, dựa trên các thông tin ban đầu như sau:*

- ☐ Tên của người cần tìm là “Bác sĩ Lan”
- ☐ Bác sĩ Nhi của một bệnh viện nổi tiếng
- ☐ Có chồng làm nhà báo rất nổi tiếng của báo Gia đình xã hội

Nếu tìm kiếm theo từ khoá “Lan” thì kết quả sẽ trả về hàng nghìn trang Web liên quan đến “Lan”. Nếu tìm kiếm theo cụm từ “Bác sĩ nhi” thì có kết quả rất nhiều bác sĩ....

Tóm lại thông tin trên Web cần phải được biểu diễn theo một quy cách khác mà các máy tính có thể hiểu và xử lý được chúng. Hay nói cách khác cần trang bị ngữ nghĩa cho các dữ liệu trên Web. Và giải pháp cho vấn đề này chính là Semantic Web hay Web ngữ nghĩa, cái mà ta sẽ làm rõ trong các phần dưới đây.

PHẦN II

CÔNG NGHỆ WEB 3.0

2.1. Web 3.0 là gì?

Web 3.0 còn được gọi là Semantic Web ra đời năm 2001 bởi Berners-Lee và hai đồng tác giả của cuốn sách Scientific American là James Hendler và Ora Lassila.

Web 3.0 được chính tác giả định nghĩa như sau:

“Bước đầu tiên là đặt dữ liệu trên Web theo một định dạng mà máy tính có thể hiểu được, hoặc chuyển thành định dạng mà máy tính có thể hiểu được. Điều này tạo ra một loại Web gọi là Semantic Web - là một Web dữ liệu mà có thể được xử lý được trực tiếp hoặc gián tiếp bằng máy tính.”

Nghĩa là: Web ngữ nghĩa là một loại Web dữ liệu, có thể được xử lý trực tiếp hoặc gián tiếp bằng máy tính.

- Semantic Web hay Web ngữ nghĩa thực chất là một sự mở rộng của Web hiện hành nhằm mục đích khai thác tốt nhất công nghệ Web. Sự mở rộng của Semantic Web chính là việc thêm vào trong Web hiện hành yếu tố ngữ nghĩa, để cho phép máy tính khai thác và khai thác tốt hơn các thông tin trên Web. Trên Semantic Web, tài nguyên được đưa ra với sự xác định rõ ràng về ngữ nghĩa và thuận tiện để máy tính và con người có thể làm việc một cách cộng tác.

Tóm lại: Semantic Web đơn thuần chỉ là một sự mở rộng của Web hiện hành mà không phải là một sự đột phá thay thế công nghệ Web cũ. Ngược lại Semantic Web kế thừa từ Web hiện hành và cho phép khai thác Web hiện tại trong một con đường mới, con đường mà máy và người có thể làm việc cộng tác trong khai thác tài nguyên Web.

Khác với Web hiện hành, Semantic Web hướng tới đối tượng phục vụ chính của mình là máy, song nó vẫn kế thừa cái đích con người từ Web hiện hành

và kết quả là tạo ra một môi trường làm việc cộng tác người máy trong khai thác thông tin trên Web.

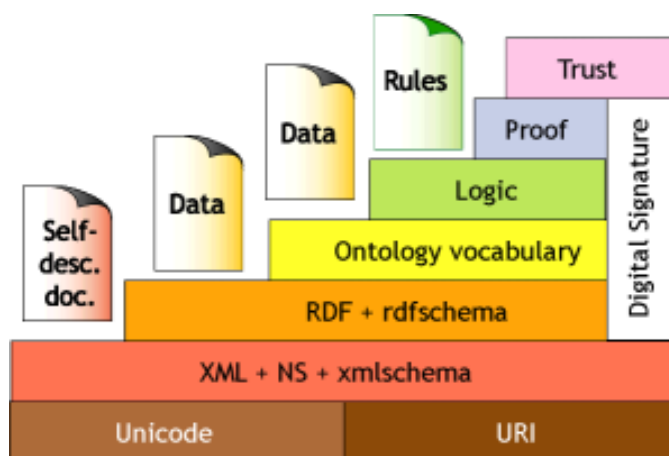
Ví dụ: với cụm từ “Nguyễn Ngọc Cương”, trong không gian Semantic Web, được suy luận đến các thuộc tính như: năm sinh, địa chỉ, số điện thoại, email...

- Siêu dữ liệu (Metadata) - dữ liệu trong dữ liệu, một đặc trưng quan trọng của Semantic Web, không chỉ giúp máy tính có thể xử lý thông tin trên Web một cách thông minh, mà còn làm cho dữ liệu “thông minh” hơn

2.2 . Kiến trúc của Semantic Web.

Kiến trúc của Semantic Web được cho bởi sơ đồ sau.

Theo kiến trúc này Semantic Web được phân thành các tầng. Cụ thể mỗi tầng được miêu tả như sau.



2.2.1. Tầng Unicode và URI.

Tầng Unicode và URI là tầng đầu tiên của kiến trúc Semantic Web. Đây là hạ tầng đầu tiên cho xây dựng Semantic Web, là nền tảng để mã hóa, định vị và truyền vận thông tin. Với :

- Unicode là chuẩn mã hóa dữ liệu để vận chuyển thông tin. Unicode là chuẩn mã hóa quốc tế, nó cho phép mã hóa mọi ngôn ngữ.
- URI-Uniform Resource Identifier là nền tảng để xác định vị trí cho các tài nguyên Web, cũng chính là việc xác định tài nguyên Web.

Thực tế tầng này đã được hoàn thiện và sử dụng trong nền Web hiện tại. Các URI được miêu tả với các giao thức khác nhau như : HTTP, FTP, SMTP, ... hiện đang được sử dụng rộng rãi trên Internet. Sự xuất hiện của tầng này cho

thấy được sự kế thừa thực sự của Semantic Web. Semantic Web thực sự chỉ là một sự mở rộng của Web hiện tại nó giữ lại hầu hết những đặc điểm thiết kế bên dưới của Web hiện tại và chỉ mở rộng thêm phần ngữ nghĩa ở những mức bên trên nhằm tạo thêm một khung nhìn mới cho Web hiện tại, đó là khung nhìn của các ứng dụng (hay của máy tính).

2.2.2. Tầng XML, NS và XMLSchema.

Đây là tầng thứ hai của kiến trúc phân tầng Semantic Web. XML và các chuẩn liên quan tới nó, cung cấp cho ta một con đường để diễn đạt cấu trúc thông tin bất kỳ và đây cũng là một chuẩn thực tế để truyền dữ liệu giữa các ứng dụng. Do vậy mà XML được hỗ trợ rộng rãi về cả các công cụ và người dùng. Đó cũng là lý do mà XML tồn tại ở tầng thứ hai này với vai trò làm một điểm trung gian giữa những dạng biểu diễn giàu ngữ nghĩa hơn và các dạng dữ liệu thô giàu cấu trúc.

XML cung cấp một cú pháp chung cho biểu diễn dữ liệu trong môi trường Internet. XML Schema cung cấp các định nghĩa kiểu dữ liệu và các cấu trúc cho tài liệu XML. Các không gian tên(name space-NS) cũng được sử dụng như một giải pháp đã được áp dụng cho các tài liệu XML.

Thực sự thì ở tầng thứ hai này mọi cái vẫn được kế thừa từ những gì mà Web hiện tại đã làm được. Các chuẩn cú pháp XML, không gian tên và XML Schema vẫn được chấp nhận trong Semantic Web. Sự kế thừa này chính là một cơ sở để tích hợp các định nghĩa Semantic Web với các chuẩn XML khác.

2.2.3. Tầng RDF và RDFSchema.

Bắt đầu từ tầng này, các đặc điểm mới của Web thực sự được bộc lộ làm cơ sở để khẳng định rằng đó là Semantic Web. Sự thực thì tầng này có vai trò như một mô hình, một ngôn ngữ để biểu diễn ngữ nghĩa hay tạo ra các khung nhìn đơn giản tới máy tính.

RDF là viết tắt của **R**esource **D**escription **L**anguage là một ngôn ngữ, một mô hình dữ liệu cho phép biểu diễn các siêu dữ liệu hay các phát biểu

ngữ nghĩa về dữ liệu. Trong ngữ cảnh của Web thì các dữ liệu này còn được hiểu là các tài nguyên (resource).

RDFS-RDF Vocabulary Description Language là một ngôn ngữ để miêu tả các từ vựng được sử dụng trong tài liệu RDF. Sự xuất hiện của RDFS giúp tăng cường ngữ nghĩa cho mô hình dữ liệu RDF.

2.2.4. Tầng từ vựng Ontology.

Thông qua việc miêu tả các từ vựng dưới dạng cây hay dạng phân lớp, RDFS đã góp phần mở rộng ngữ nghĩa cho dữ liệu RDF. Tuy nhiên nếu dừng lại ở đây, ngữ nghĩa mà dữ liệu RDF cung cấp thực sự chưa đủ để đạt được những gì mà Semantic Web mong đợi. Từ đó mà tầng Ontology cần thiết được xây dựng.

Ontology cho phép mở rộng các từ vựng để miêu tả những thuộc tính và những lớp, chẳng hạn như các mối quan hệ khác giữa các lớp mà không chỉ đơn thuần là quan hệ cha con.

Có thể nói rằng RDFS giống như các kiểu tài nguyên sử dụng trong các tài liệu RDF. Nhờ định nghĩa về kiểu một cách thống nhất này mà phần nào đó ngữ nghĩa được thêm vào trong RDF. Đến lượt Ontology thì ngữ nghĩa được thêm vào đúng với mức từ vựng và được chia sẻ trên Internet.

2.2.5. Tầng Logic.

Tầng Logic hiện tại vẫn đang được phát triển và chưa có một công bố nào có tính chất chuẩn và hoàn thiện về tầng này. Nhóm phát triển Semantic Web của W3C thực sự vẫn đang phát triển tầng này và cũng vẫn chưa hề đưa ra những miêu tả hoàn chỉnh nào về tầng này.

Bản chất của tầng logic là cung cấp những cơ sở để siêu dữ liệu RDF có thể trở thành tri thức, cái được áp dụng để thực hiện các suy luận logic nhằm chứng minh hoặc đưa ra các thông tin mới từ các thông tin đã có.

2.2.6. Tầng Proof và tầng Trust.

Proof và Trust là những tầng cuối cùng trong kiến trúc của Semantic Web. Hiện tại, chưa có nhiều miêu tả về các tầng này cũng như giải pháp thực sự cho chúng. Chúng là những điểm đáng xấu hổ của hiện tại và là những điểm quan trọng của Semantic Web trong tương lai.

Một điều đơn giản để hiểu sự khắc nghiệt thật sự của về vấn đề này đó là sự mâu thuẫn của thông tin. Chẳng hạn như có người nói rằng x có màu xanh, lại có người nói rằng x có màu khác xanh, phải chăng Semantic Web sẽ sụp đổ với những trường hợp như thế này?

Câu trả lời dĩ nhiên là không, bởi vì hai lý do cơ bản sau:

- Một là ứng dụng trên Semantic Web ở hiện tại thường dựa trên nền tảng một ngữ cảnh cụ thể.
- Hai là các ứng dụng trong tương lai sẽ thường chứa đựng các kỹ thuật kiểm tra các chứng cứ và xác thực điện tử (digital signatures).

Khái lược về ngữ cảnh(Context).

Các ứng dụng trên Semantic Web dựa trên một ngữ cảnh thường để mọi người xác thực sự đúng đắn của dữ liệu.

Ngữ cảnh là một cơ sở tốt bởi vì chúng ta có thể tin tưởng được mà không cần phải nhờ vào những sự thẩm định phức tạp và các hệ thống kiểm tra. Tuy nhiên vấn đề xuất hiện đối tượng thứ ba, kẻ giả mạo là không tránh khỏi và đó là đòi hỏi sự ra đời của digital signature hay chữ ký số.

2.2.7. Chữ ký số - Digital Signature.

Việc áp dụng công nghệ mã hóa và ký điện tử trong RDF sẽ đảm bảo rằng nguồn tài liệu mà ta đang sử dụng là do chính xác một nhà cung cấp nào đó mà ta tin tưởng. Điều này giống như sự xác thực điện tử mà trong an toàn thông tin đã đề cập đến.

Trong kiến trúc của Semantic Web, Digital signature đóng một vai trò rất quan trọng. Nó gắn liền với các tầng của kiến trúc Semantic Web kể từ tầng thứ ba RDF, với vai trò là mở rộng cho các tầng này để đảm bảo rằng

những thông tin trong các tài liệu này là xác thực do một nhà cung cấp nào đó. Điều này giúp ngăn chặn sự sụp đổ của Semantic Web do chính đặc tính đơn giản và phổ cập của nó mang đến. Với Digital Signature, các ứng dụng sẽ có căn cứ để sử dụng các thông tin chính xác do những nhà cung cấp mà ứng dụng đó tin tưởng.

PHẦN III

LỢI ÍCH KHI SỬ DỤNG WEB 3.0 VÀ NHỮNG KHÓ KHĂN, THÁCH THỨC GẶP PHẢI.

3.1. Lợi ích khi sử dụng Web 3.0

Qua những phân tích ở trên, chúng ta cũng có thể thấy được những lợi ích khi sử dụng công nghệ Web mới này.

- *Máy có thể hiểu được thông tin trên Web*

Web hiện hành chỉ phục vụ cho con người. Semantic Web sẽ cung cấp ý nghĩa cho máy hiểu vì nó chứa cấu trúc ngữ nghĩa thông tin.

- *Tìm kiếm thông tin nhanh chóng và chính xác hơn*

Với Semantic Web, việc tìm kiếm sẽ dễ dàng nếu mọi thứ được đặt trong ngữ cảnh người sử dụng biết đến. Mục tiêu của Semantic Web là phát triển kỹ thuật để giúp máy hiểu nhiều thông tin trên Web.

- *Hỗ trợ công cụ tự động hóa:*

- + Hỗ trợ cho các phần mềm thông minh

- + Cung cấp các dịch vụ: gia đình và các thư viện điện tử, kinh doanh điện tử, dịch vụ sức khỏe.v.v.

- + Tự động cập nhật các thông tin chi tiết lên Web nhằm hỗ trợ sự tự động hóa cho các dịch vụ

- *Dữ liệu liên kết động*

- + Semantic Web có thể kết hợp các thông tin đã được mô tả và giàu ngữ nghĩa với bất kỳ nguồn dữ liệu nào

- + Có thể tích hợp thông tin từ nhiều nguồn khác nhau, tìm kiếm, duyệt dữ liệu và tự động hóa các thao tác

Ví dụ: thêm các Metadata (dữ liệu về dữ liệu) cho các tài liệu khi tạo ra nó, có thể tìm kiếm các tài liệu mà Metadata cho biết tác giả là Eric Miller

- + Semantic Web, không chỉ cung cấp các URI cho tài liệu như đã làm trong quá khứ mà còn cho con người, các khái niệm và các mối liên hệ.

3.2. Những thách thức hiện tại

Vấn đề gì cũng có hai mặt, và Web 3.0 cũng thế. Một số những thách thức mà Web 3.0 phải đối mặt bao gồm:

- Khối lượng dữ liệu khổng lồ: World Wide Web hiện có cả tỉ trang web trên đó, mỗi trang web lại tạo ra một lượng dữ liệu to nhỏ khác nhau. Các cảm biến, thiết bị di động, hành vi của người dùng lại tạo thêm một đống nữa. Chính vì thế, nếu các nhà thiết kế, nhà phát triển web không kịp thay đổi công nghệ thì sẽ không thể đảm đương được hết những dữ liệu này, và không mang lại hiệu quả tối đa cho người dùng. Những hệ thống tự động hóa cũng phải được thiết kế lại để xử lý lượng dữ liệu đầu vào rất rất lớn. Dữ liệu trùng lặp cũng là một vấn đề cần được cân nhắc.

- Dữ liệu không rõ ràng: những từ như "trẻ", "cao" hoặc "lớn" có thể được xem là dữ liệu không rõ ràng, hay nói cách khác là quá chung chung. Như thế nào là trẻ? Như thế nào là cao, cao bao nhiêu? Lớn là gì, lớn ra sao, so với cái gì? Kết quả là những truy vấn tìm kiếm của người dùng cũng sẽ chung chung như thế, ngoài ra còn có sự chung chung trong dữ liệu được cung cấp bởi các công ty / tổ chức, hoặc sự chung chung trong các cơ sở dữ liệu kiến thức sẽ khiến việc xử lý, giao tiếp trở nên khó khăn.

- Sự thiếu nhất quán: Mặc dù đã có các chuẩn chung nhưng chắc chắn sẽ luôn có sự khác biệt xuất hiện. Có thể là vì giải pháp 1 tốt hơn giải pháp 2 nhưng giải pháp 2 lại dễ dùng hơn nên cả hai tồn tại song song. Điều đó tạo ra sự thiếu nhất quán trong thế giới Word Wide Web. Như hiện nay cũng thế, Flash và Silverlight, JSON và XML,...

- Bảo mật: khi dữ liệu được chia sẻ dễ dàng thì cũng mang theo nguy cơ cao về an toàn thông tin. Ví dụ, thông tin đó có thể bị giả mạo, có thể bị thay thế giữa lúc đang truyền đi, thông tin mang theo mã độc, thông tin không được mã hóa...

Vẫn còn nhiều nhiều những vấn đề khác, và tất cả đều cần đến một hoặc một số giải pháp đã có sẵn hoặc chưa có. Nhưng tóm lại, Web 3.0 là xu hướng không thể chối bỏ, và nó sẽ dần rõ nét hơn trong thời gian tới.

PHẦN IV

KẾT LUẬN VÀ DỰ ĐOÁN TƯƠNG LAI

Web 3.0 chính là tương lai mà các tổ chức, các công ty, cộng đồng muốn hướng đến, Web 3.0 sẽ giúp cho việc sử dụng Internet dễ dàng và tiện ích hơn.

Mặc dù Web 3.0 chỉ mang tính lý thuyết nhưng không không gì có thể bắt mọi người dừng những suy đoán về những gì sẽ xảy ra tiếp theo.

Dự đoán vượt xa Web 3.0

Mọi thứ mà chúng ta gọi là thế hệ kế tiếp của web thì những gì sẽ xảy ra sau đó? Một loại các lý thuyết từ các dự đoán nhằm đoán trước nghe gần giống như các bộ phim khoa học viễn tưởng.

Đây là một trong số các phán đoán đó:

- Theo một chuyên gia về công nghệ và cũng là chủ một doanh nghiệp, Nova Spivack, sự phát triển của Web sẽ thay đổi cứ 10 năm một lần. Trong thập kỷ đầu tiên của web, hầu hết sự phát triển chỉ tập trung vào back end, hoặc cơ sở hạ tầng của Web. Các nhà lập trình đã tạo ra các giao thức và các ngôn ngữ mã để chúng ta có thể sử dụng nhằm tạo ra các trang web. Trong thập kỷ thứ hai, những tập trung được hướng vào front end và kỷ nguyên Web 2.0 bắt đầu. Giờ đây, mọi người sử dụng các website như một nền tảng để chạy các ứng dụng của họ. Họ cũng tạo các mashup và thử nghiệm nhiều cách để làm cho cảm nhận về web mang tính tương tác hơn. Hiện chúng ta đang nằm trong những năm cuối của chu kỳ Web 2.0. Chu kỳ tiếp theo sẽ là Web 3.0 và sự tập trung của nó sẽ hướng trở về back end. Các lập trình viên sẽ định nghĩa lại cơ sở hạ tầng Internet để hỗ trợ thêm những tính năng ưu việt của các trình duyệt web 3.0. Khi giai đoạn đó kết thúc, chúng ta sẽ chuyển sang thời đại web 4.0. Sự tập trung lúc này lại trở về front end, và chúng ta sẽ thấy hàng nghìn các chương trình mới sử dụng Web 3.0 như một nền cơ sở.

- Web sẽ chuyển sang môi trường ba chiều. Đúng hơn là Web 3.0, chúng ta sẽ thấy một Web 3D. Việc kết hợp các thành phần ảo với các thế giới trực tuyến của các trò chơi trực tuyến, Web có thể trở thành một “Landscape” số để kết hợp chặt chẽ vấn đề ảo giác. Bạn sẽ điều hướng web từ một phối cảnh người đầu tiên hoặc thông qua sự hiện diện của bản thân mình được gọi là avatar.
- Web sẽ xây dựng trên những sự phát triển của điện toán phân phối và tạo ra một trí óc nhân tạo đích thực. Trong việc điện toán phân phối, một vài máy tính xử lý một công việc xử lý lớn. Mỗi một máy tính quản lý một phần nhỏ của toàn bộ nhiệm vụ. Một số người tin tưởng rằng web sẽ có thể suy nghĩ bằng cách phân phối workload trên hàng nghìn máy tính và việc tham chiếu sâu các ontology. Web sẽ trở thành một bộ não khổng lồ có khả năng phân tích dữ liệu và ngoại suy các ý tưởng mới dựa trên các thông tin đó.
- Web sẽ mở rộng vượt ra khỏi các máy tính và điện thoại di động. Mọi thứ từ xem tivi đến y phục sẽ kết nối Internet. Người dùng sẽ có một kết nối “constant” với web, ngược lại. Mỗi software agent của người dùng sẽ học thêm nhiều về phối cảnh người dùng của nó bằng cách quan sát các hành động của anh ta. Điều này có thể dẫn đến các tranh luận về sự cân bằng giữa vấn đề riêng tư và lợi ích của một cảm nhận duyệt web mang tính cá nhân hóa.
- Web sẽ hợp nhất với các biểu mẫu giải trí khác cho tới khi tất cả các sự khác biệt giữa các biểu mẫu về truyền thông mất đi. Các chương trình radio, Tivi show và các phim sẽ dựa trên web với tư cách là một hệ thống phân phối.

Vẫn còn quá sớm để nói về phiên bản tương lai nào của web sẽ diễn ra. Có thể rằng tương lai thực của web lại vượt xa hơn nhiều những gì đã được dự đoán. Rõ ràng chúng ta chỉ có thể hy vọng rằng theo thời gian và tương lai của web sẽ ngày càng tốt đẹp.