

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



ĐINH QUANG ĐỊNH

NGHIÊN CỨU CÔNG NGHỆ WEB 3.0

(SEMANTIC WEB)

VÀ KHẢ NĂNG TRIỂN KHAI ÁP DỤNG

Chuyên ngành: **HỆ THỐNG THÔNG TIN**

Mã số: 60.48.01.04

TÓM TẮT LUẬN VĂN THẠC SĨ

HÀ NỘI – NĂM 2013

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: PGS.TS.Hoàng Minh

Phản biện 1:

Phản biện 2:

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: giờ ngày tháng năm

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỞ ĐẦU

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. Kể từ Web 1.0 chỉ là sự xuất bản thông tin một chiều đến Web 2.0 xuất hiện vào đầu những năm 2000 đã tạo ra một khuynh hướng mới với sự tham gia của lực lượng xây dựng nội dung web không chuyên ngày càng tăng, dẫn đến thể loại Web ngày càng phong phú.

Web 2.0 được đánh dấu bằng sự ra đời của các blogs, các mạng xã hội (social network), các trang web cho phép người dùng chia sẻ nội dung như Youtube. Facebook – mạng xã hội phổ biến nhất hiện nay trên thế giới với hơn 1 tỷ người sử dụng (số liệu công bố của Facebook ngày 4/10/2012) dù mới chỉ ra đời năm 2004, hẳn không xa lạ với rất nhiều người trong chúng ta.

Tuy nhiên Web 2.0 cũng xuất hiện một số nhược điểm như thông tin quá tải với người dùng, do nội dung được cung cấp nhiều nguồn hơn so với thế hệ trước. Kết quả tìm kiếm trên các công cụ search engine thường quá nhiều nội dung không liên quan tới chủ đề cần tìm kiếm. Các công cụ (thiết bị, trình duyệt) vẫn hiểu nội dung chỉ là các chuỗi byte 0101. Dữ liệu các website vẫn độc lập với nhau...

Để giải quyết các nhược điểm đó các nhà nghiên cứu đang xây dựng một thế hệ Web mới, Web 3.0. Web 3.0 là một tập hợp các công nghệ bao gồm các trang web ngữ nghĩa, dữ liệu liên quan, xử lý ngôn ngữ tự nhiên, trí thông minh nhân tạo, mashup, và các API... Với thế hệ mới, web đã có thể hiểu những gì bạn đang nghĩ.

Việc tìm hiểu công nghệ Web 3.0 sẽ giúp cho chúng ta có được cơ sở lý thuyết để định hướng cho việc áp dụng công nghệ mới này tại Việt nam.

1. Lý do chọn đề tài:

Web 3.0 sẽ dựa trên phần công nghệ quan trọng nhất là Semantic Web và dựa trên một số công nghệ Web khác. Web 3.0 hiện đang ở những bước sơ khai ban đầu được sáng tạo và phát triển nhưng đã thu hút sự quan tâm đặc biệt các nhà nghiên cứu, các tổ chức, các công ty, cộng đồng...

Xuất phát từ lý do đó luận văn đặt vấn đề nghiên cứu Web 3.0 nhằm chủ động nắm bắt được công nghệ Web mới và đề xuất khả năng triển khai áp dụng tại Việt Nam.

2. Mục đích, đối tượng và phạm vi nghiên cứu:

Mục đích của đề tài: Đưa ra được cái nhìn khái quát việc triển khai mô hình Web3.0 trên thế giới đồng thời đánh giá hiện trạng việc sử dụng Web 2.0 trong nước từ đó đưa ra lộ trình thực hiện áp dụng công nghệ Web 3.0 tại Việt nam.

Đối tượng nghiên cứu:

Lý thuyết mô hình kiến trúc Semantic Web, các công nghệ và các xu hướng phát triển của Web 3.0. Việc phát triển và triển khai Web 3.0 trên thế giới và đánh giá giá hiện trạng việc sử dụng Web tại Việt nam .

Phạm vi nghiên cứu:

Nghiên cứu lý thuyết về các công nghệ hỗ trợ và phát triển cho Web 3.0 thay thế cho Web 2.0 tại Việt nam.

Phương pháp nghiên cứu

Các tài liệu liên quan đến công nghệ Web 3.0 và các trang Web của W3C và các bài báo chuyên đề liên quan đến công nghệ này.

Bố cục của luận văn: gồm 3 chương và phần mở đầu, phần kết luận kiến nghị:

Phần mở đầu nêu được sự ra đời phát triển của các thế hệ Web từ 1.0 đến 2.0 đồng thời cũng đánh giá được những khuyết điểm hiện nay của Web 2.0 nhằm nêu bật được ý nghĩa sự cấp bách của việc áp dụng thay thế bằng Web 3.0 trên thế giới và tại Việt nam.

Chương 1: Nghiên cứu tổng quan về các công nghệ Web.

Chương 2: Nghiên cứu các công nghệ và xu hướng phát triển Web 3.0.

Chương 3: Khuyến nghị đề xuất khả năng triển khai áp dụng tại Việt nam.

Phần kết luận và kiến nghị: đưa ra một số vấn đề tồn tại cần giải quyết và hướng nghiên cứu tiếp theo của đề tài.

Kết quả của luận văn đã nghiên cứu, tìm hiểu kỹ lưỡng về các công nghệ, các công cụ phát triển của Web 3.0. Việc triển khai nó trên thế giới đồng thời đánh giá thực trạng công nghệ Web 2.0 đang sử dụng trong nước dẫn đến việc đưa ra được khuyến nghị việc triển khai và áp dụng tại Việt nam

CHƯƠNG 1: TỔNG QUAN VỀ CÔNG NGHỆ WEB

Chương 1 giới thiệu tổng quan về lịch sử phát triển của các công nghệ Web giúp chúng ta có cái nhìn tổng thể về sự ra đời và phát triển của từng thế hệ Web.

Qua đó chương cũng đã đánh giá được những ưu và nhược điểm của các công nghệ Web 1.0, Web 2.0 dẫn đến sự cần thiết phải nghiên cứu phát triển thế hệ Web mới Web 3.0.

Phần 1.1 Nhìn lại lịch sử phát triển các công nghệ Web. Phần 1.2, Giới thiệu tổng quan và các đặc tính của Web3.0. Phần kết luận của chương được nêu trong mục 1.3.

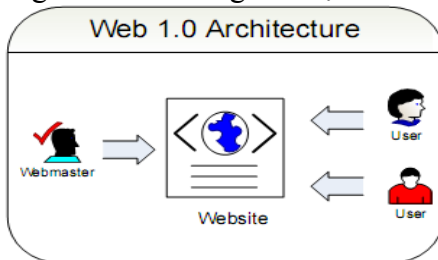
1.1. Lịch sử phát triển công nghệ Web

WWW được đề nghị lần đầu tiên vào những năm 1990 bởi Tim Berners-Lee và Robert Cailliau khi hai ông làm việc tại CERN ở Geneva, Thụy Sĩ.

Kiểu Web mà trong đó tác nhân người sử dụng mô phỏng theo phản ứng của con người, có thể đọc và hiểu thông tin sử dụng trí tuệ nhân tạo được gọi là Web ngữ nghĩa.

1.1.1. Web 1.0

Thế hệ đầu tiên của Web được gọi là “Web 1.0” hoặc đơn giản là “Web”. Web 1.0 còn có các tên gọi khác là “Read Web”, “Old Web” hoặc “Static Web”. Web 1.0 chủ yếu là môi trường xuất bản thông tin một chiều.



Hình 1.1. Kiến trúc Web 1.0 điển hình

Các đặc tính của Web 1.0 có thể được tổng kết như sau:

- Trong Web 1.0, Webmaster là người chịu trách nhiệm quản lý nội dung và duy trì cập nhật cho người sử dụng.

- Web 1.0 không hỗ trợ xuất bản thông tin rộng rãi
- . Web 1.0 sử dụng ngôn ngữ đánh dấu siêu văn bản cơ bản để xuất bản nội dung trên Internet;

- Web 1.0 không hỗ trợ nội dung có thể đọc bởi máy. Chỉ người đọc Web có thể hiểu được nội dung;

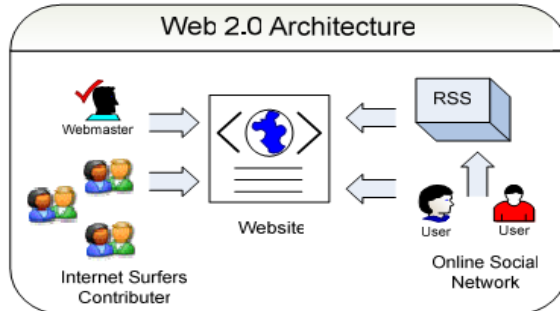
- Web 1.0 Người sử dụng vẫn phải sử dụng các công cụ không trực tuyến khác để truyền thông với thông tin liên lạc này;

- Trong Web 1.0, các trang Web được thiết kế để phản ứng theo bản năng dựa trên điều kiện được lập trình.

1.1.2. Web 2.0

Thuật ngữ Web 2.0 được chính thức định nghĩa vào năm 2004 bởi Dale Dougherty, phó chủ tịch của O'Reilly Media, trong hội nghị tổ chức bởi O'Reilly và MediaLive International. Tim O'Reilly định nghĩa Web 2.0 như sau:

Mọi thứ trên Web được gắn thẻ (tag), giúp cho việc điều hướng nhanh và dễ dàng hơn. Ngược với Web 1.0, Web 2.0 có trí tuệ tập thể của hàng triệu người sử dụng.



Hình 1.2 Kiến trúc Web 2.0 điển hình

Các đặc tính của Web 2.0 có thể tổng kết như sau:

- Web 2.0 là phiên bản thứ hai của Web cung cấp ứng dụng Internet giàu có RIA

- Kiến trúc hướng dịch vụ SOA là phần cơ bản trong Web 2.0.

- Web 2.0 là Web xã hội. Ứng dụng Web 2.0 hướng tới tương tác nhiều hơn với người sử dụng đầu cuối.

- Trong thuật ngữ và chiến lược Web 2.0 thì “Web là môi trường mở”.

- Trong Web 2.0, dữ liệu là động lực. Một trong những công nghệ quan trọng là AJAX, công nghệ này hỗ trợ sự phát triển kinh nghiệm của người sử dụng tiềm năng.

- Các công nghệ và dịch vụ chính của Web 2.0 bao gồm các blog, tổ chức cung cấp đơn giản thực sự RSS, wiki, mashup, tag, folksonomy và các đám mây gắn thẻ.

Như vậy, Web 2.0 là phiên bản cải tiến của World Wide Web, Web 2.0 vẫn còn tồn tại nhiều hạn chế, có thể tổng kết như sau [9, 10]:

- Các hạn chế của ngôn ngữ HTML

- Các ứng dụng được phát triển tập trung vào sự tiện nghi cho con người, do đó làm cho việc cơ giới hóa các nhiệm vụ trở nên phức tạp.

- Việc truy nhập tới hầu hết thông tin là miễn phí và phần mềm là mở, dẫn tới các dịch vụ phải đương đầu với các thách thức về lợi nhuận.

- Thiếu các Web server thông minh để tránh hiện tượng tắc nghẽn nút cổ chai (bottle-neck)

- Thiếu các phương pháp và mô hình hóa để hỗ trợ việc thiết kế RIA UI, không hỗ trợ sự tương thích ngữ nghĩa từ UI của Web 1.0 tới UI của Web 2.0.

- Thiếu tổ chức thẩm quyền trung tâm thực hiện việc tổ chức và tiêu chuẩn hóa phương thức mà Web được quản lý.

- Các thách thức về an ninh và bảo mật do sự phơi bày thông tin cá nhân/tổ chức trên Web 2.0.

- Cung cấp các khả năng truy vấn tài: thiếu sự biểu diễn dữ liệu tổng quát.

- Quá tải thông tin: Quá tải thông tin phân tán với chất lượng không đáng tin cậy được xem là vấn đề nghiêm trọng.

- Chu kỳ lặp lại không đổi của việc thay đổi và nâng cấp các dịch vụ.

- Các vấn đề nguyên tắc trong việc xây dựng và sử dụng Web 2.0: Các công nghệ và các dịch vụ mới của Web 2.0 bắt đầu cho thấy sự hạn chế theo thuật ngữ sự riêng tư và bản quyền.

- Vấn đề liên kết nối: Sự liên kết nối và kiến thức chia sẻ giữa các nền tảng (platform) qua các ranh giới giữa cộng đồng vẫn còn bị hạn chế.

- Sự không hiệu quả của các hệ thống chia sẻ thông tin trong các ứng dụng Web.

- Sự tin cậy của các Website và các nội dung bên trong chúng:

- Truy nhập toàn cầu: Một thách thức mà Web 2.0 phải đương đầu là đảm bảo rằng tất cả các nhà phát triển Web và các nhà thiết kế Web tuân theo một nguyên tắc truy nhập trong việc cung cấp sự mô tả, tối ưu hóa việc truy nhập tới tất cả người sử dụng Web, đặc biệt là những người khuyết tật.

1.1.3. Web thế hệ kế tiếp

Ý tưởng chính của công nghệ Web 3.0 thế hệ kế tiếp là tạo ra nội dung Web bằng cách không sử dụng ngôn ngữ tự nhiên mà ở dạng tập lệnh (script) có thể hiểu được và phán đoán được bởi các agent phần mềm để cho phép chúng tìm kiếm, chia sẻ hoặc tích hợp thông tin dễ dàng hơn và hiệu quả hơn, hướng tới các ứng dụng thông minh. Mục đích chủ yếu của công nghệ Web 3.0 là hỗ trợ người sử dụng đóng góp thông tin theo các phương thức mà máy tính có thể hiểu được, xử lý và trao đổi.

1.2. Tổng quan về công nghệ Web 3.0

1.2.1. Giới thiệu

Thuật ngữ “Web 3.0” lần đầu tiên được đề nghị bởi John Markoff trên thời báo “New York Times” vào năm 2006 [15, 19], và lần đầu tiên xuất hiện nổi bật vào đầu năm 2006 trong bài báo Blog “*Critical of Web 2.0 and associated technologies such as Ajax*” viết bởi Jeffrey Zeldman.

Đầu tiên phải kể đến quan điểm: Web 3.0 là sự chuyển đổi Web thành cơ sở dữ liệu.

Quan điểm thứ hai nhấn mạnh Web 3.0 với sự phát triển hướng tới trí tuệ nhân tạo.

Quan điểm thứ ba cho rằng Web 3.0 thực hiện Web ngữ nghĩa và kiến trúc hướng dịch vụ SOA.

Và một quan điểm khác đó là Web 3.0 là sự phát triển hướng đến 3D.

Ngoài ra, còn tồn tại một số định nghĩa mở rộng được đề nghị đối với Web 3.0 khác bao gồm:

- Kết nối toàn cầu
- Tính toán mạng
- Các công nghệ mở
- Nhận dạng mở
- Web thông minh

1.2.2. Tổng quan Web 3.0

Web 3.0, ngoài việc có một số tính năng bao gồm tính năng của Web 2.0, sẽ có một số mục tiêu chính là: tìm kiếm thông tin, tìm kiếm hiệu lực, và tìm kiếm giải trí. Việc tìm kiếm thông tin sẽ gọn hơn trong Web 3.0.

Giải trí, xu hướng phổ biến nhất của Web 2.0, sẽ tiên tiến hơn trong Web 3.0, bởi vì nó dựa trên việc lựa chọn của cá nhân.

Sức mạnh thực sự của Web 3.0 sẽ là ở việc tạo ra dữ liệu và truyền tải nó hiệu quả. Web 3.0 sẽ sử dụng các công nghệ máy tính và Internet khác nhau hỗ trợ sau đây:

- Trí tuệ nhân tạo
- Suy diễn tự động
- Kiến trúc nhận thức
- Các ứng dụng tổng hợp
- Tính toán phân tán
- Biểu diễn tri thức

- Ontology
- Văn bản tổ hợp
- Vector đồ họa
- Web ngữ nghĩa
- Wiki ngữ nghĩa
- Các agent phần mềm

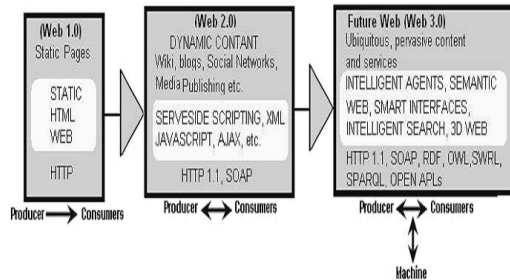
1.2.3. Các đặc tính của Web 3.0

Đặc tính thông minh

Đặc tính cá nhân hóa

Đặc tính hiển thị

1.2.4. Các xu hướng công nghệ đối với Web 3.0



Hình 1.5. Sự phát triển công nghệ Web 3.0

Các xu hướng nổi bật của Web 3.0 có thể kể đến là Web ngữ nghĩa, Web 3D, Web mạng xã hội, Web tập trung dữ liệu đa phương tiện, và Web toàn cầu.

1.3. Kết luận chương

Chương đầu của luận văn đã khái quát toàn bộ lịch sử của các công nghệ Web. Đầu những năm của thập kỷ 90 của thế kỷ trước Web 1.0 ra đời, vào thừa sơ khai nó chỉ là dạng thông tin một chiều trên Internet. Người sử dụng chỉ có thể đọc các thông tin có trên đó.

Web 2.0 xuất hiện vào đầu những năm 2000 người sử dụng đã có thể truyền thông với nhau. Web 2.0 là phiên bản thứ hai của Web cung cấp ứng dụng Internet giàu có RIA với kiến trúc hướng dịch vụ SOA là phần cơ bản trong Web 2.0 đồng thời nó được đánh dấu bằng sự ra đời của các blogs, các mạng xã hội (social network), các trang Web cho phép người dùng chia sẻ nội dung như Youtube, Facebook. Tuy nhiên Web 2.0

cũng xuất hiện một số nhược điểm như tác giả đã đánh giá trong phần mở đầu ” *Web 2.0 cũng xuất hiện một số nhược điểm như thông tin quá tải với người dùng, do nội dung được cung cấp nhiều nguồn hơn so với thế hệ trước. Kết quả tìm kiếm trên các công cụ search engine thường quá nhiều nội dung không liên quan tới chủ đề cần tìm kiếm. Các công cụ (thiết bị, trình duyệt) vẫn hiểu nội dung chỉ là các chuỗi byte 0101. Dữ liệu các website vẫn độc lập với nhau...*” .

Web 3.0 đang được nghiên cứu và triển khai để khắc phục những nhược điểm của Web 2,0. Với ý tưởng là tạo ra nội dung Web bằng cách không sử dụng ngôn ngữ tự nhiên mà ở dạng tập lệnh (script) có thể hiểu được và phán đoán được bởi các agent phần mềm để cho phép chúng tìm kiếm, chia sẻ hoặc tích hợp thông tin dễ dàng hơn và hiệu quả hơn, hướng tới các ứng dụng thông minh. Các đặc tính chủ yếu của Web 3.0 như đặc tính thông minh, đặc tính cá nhân hóa và hiển thị đều sẽ được các nhà nghiên cứu tích hợp trong công cụ phát triển của Web 3.0.

CHƯƠNG 2: NGHIÊN CỨU CÁC CÔNG NGHỆ VÀ XU HƯỚNG PHÁT TRIỂN CỦA WEB 3.0

Chương 2: Tập trung vào nghiên cứu lý thuyết về các công nghệ Web 3.0 hay Web ngữ nghĩa (Semantic Web).

Phần đầu của chương sẽ giới thiệu kiến trúc bảy lớp được thiết kế bởi Tim Berners-Lee trong đó tập chung phân tích kỹ các lớp cơ bản lớp 2, lớp 3 và 4.

Phần 2.2 của chương đưa ra những đánh giá về xu hướng phát triển của các công nghệ Web 3.0.

Phần kết luận của chương được nêu trong mục 2.3.

2.1. Các công nghệ sử dụng trong Web 3.0

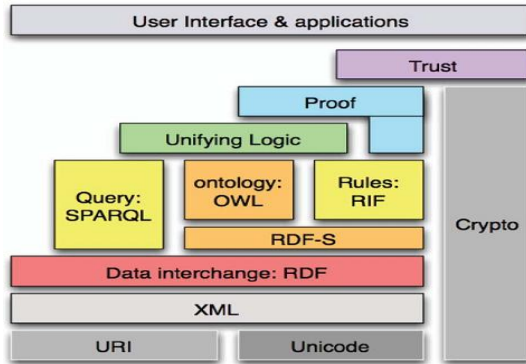
2.1.1 Giới thiệu

Công nghệ Web ngữ nghĩa (Semantic Web) được đề cập trong bài báo tầm nhìn của Tim Berners-Lee về Web như là môi trường toàn cầu để trao đổi dữ liệu, thông tin và kiến thức. Web ngữ nghĩa khuyến khích người sử dụng xuất bản và chia sẻ dữ liệu của mình và bổ sung các liên kết tới dữ liệu khác.

Một số ưu điểm của việc đạt được mục tiêu này bao gồm:

- Khả năng xác định thông tin dựa trên nghĩa của nó, ví dụ biết khi nào hai câu là tương đương, hoặc biết rằng một tham chiếu tới một người ở các trang Web khác nhau là đang đề cập tới cùng một cá nhân;
- Tích hợp thông tin qua các nguồn khác nhau – bằng cách tạo ra các ánh xạ qua ứng dụng và các giới hạn thuật ngữ chúng ta có thể mô tả các khái niệm đồng nhất và liên quan;
- Cải tiến phương thức thông tin được biểu diễn tới người sử dụng, ví dụ như tập hợp thông tin từ các nguồn khác nhau, loại bỏ sự trùng lặp, và tổng kết dữ liệu.

2.1.2. Kiến trúc của Web ngữ nghĩa



Hình 2.2. Kiến trúc của Web ngữ nghĩa

Kiến trúc của Web ngữ nghĩa được thiết kế bởi Tim Berners-Lee, và được chấp nhận bởi W3C, được mô tả trên Hình 2.2. Có bảy lớp trong kiến trúc của Web ngữ nghĩa, trong đó lớp thứ hai, lớp thứ ba và lớp thứ tư là những lớp cơ bản của Web ngữ nghĩa.

2.1.3. Dữ liệu được liên kết và dữ liệu mở

2.1.3.1. Định nghĩa dữ liệu mở được liên kết

Dữ liệu được liên kết là kết quả của một nỗ lực cộng đồng. Dự án Dữ liệu Mở Liên kết được thực hiện bởi W3C nhằm mục đích tăng lượng Web có dữ liệu được liên kết bằng cách xuất bản các tập dữ liệu mở khác nhau ở dạng RDF trên Web và kết nối chúng tới các nguồn dữ liệu khác nhau.

2.1.3.2. Các nguyên lý cơ bản của dữ liệu được liên kết

Dữ liệu được liên kết sử dụng URI để liên kết tới một đối tượng dữ liệu hơn là một tài liệu. Tim Berners-Lee, người phát minh ra dữ liệu được liên kết, đã đưa ra bốn nguyên lý của dữ liệu được liên kết:

- Sử dụng URI để mô tả mọi thứ biểu diễn trên Web như các nguồn tài nguyên;

- Sử dụng các HTTP URIs, vì vậy người sử dụng có thể tìm kiếm các tên này, có thể xác định đối tượng cụ thể;
- Khi người sử dụng tìm kiếm URI, cung cấp thông tin hữu ích, sử dụng tiêu chuẩn RDF;
- Bao gồm liên kết tới các URI khác, để người sử dụng có thể khám phá nhiều thông tin hơn.

Dữ liệu được liên kết tốt nên gồm các đặc điểm sau [43]:

- Có lẽ ưu điểm quan trọng nhất của dữ liệu được liên kết là thuận tiện và đơn giản. Với bốn nguyên lý hoạt động trên, dữ liệu được liên kết có thể dễ dàng được hình thành và ngay lập tức được áp dụng;
- Bởi vì dữ liệu được liên kết sử dụng khuôn dạng RDF, dữ liệu có cấu trúc và dữ liệu không có cấu trúc có thể áp dụng được như nhau;
- Dữ liệu được liên kết có thể kết nối tất cả dữ liệu riêng lẻ;
- Linh hoạt và dễ dàng nâng cấp;
- Truy nhập, phân tích và thao tác dữ liệu có thể thực hiện bởi người sử dụng.

2.1.4. Xuất bản dữ liệu được liên kết trên Web

Dữ liệu được liên kết sử dụng URI để đặt tên các đối tượng trên thế giới, các nguồn tài nguyên này không phải là các nguồn tài nguyên thông tin. File nguồn tài nguyên thông tin tổng quát có hai loại: nêu yêu cầu đến từ các trình duyệt (trong tiêu đề bao gồm yêu cầu text/html), chúng ta trả về file HTML. Nếu yêu cầu là ứng dụng/rdf+xml, chúng ta trả về file RDF.

2.1.5. Các công nghệ Web 3D

2.1.5.1. Giới thiệu

Hiện nay cả công nghệ phần cứng và phần mềm hỗ trợ việc hiển thị đồ họa 3D thực đều đang phát triển. Công nghệ Web 3D thường gắn với thuật ngữ công nghệ thực tế ảo VR.

2.1.5.2. VRML**2.1.5.3. X3D****2.1.5.4. DMLW****2.1.5.5. XML3D****2.1.5.6. WebGL****2.1.5.7. O3D****2.1.5.8. U3D****2.1.5.9. COLLADA****2.1.6. Các công nghệ Media-Centric Web**

Media-Centric Web là công nghệ Web trong đó media là thành phần cơ bản. Media-Centric Web có thể được hỗ trợ phát triển bởi các công nghệ sau:

- Công nghệ chuyển đổi tiếng nói thành văn bản (voice-to-text) sẽ là thành phần quan trọng của Media-Centric Web.

- Các giao diện dựa trên các chuyển động sẽ trở nên quan trọng trong việc tính toán và Web tập trung vào media.

2.2. Các xu hướng phát triển của Web 3.0**2.2.1 Ubiquitous Web****2.2.2. Soccial Semantic Web***2.2.2.1. Tổng quan về Social web**2.2.2.2. Cầu nối các công nghệ web Semantic và Social**Web**2.2.2.3. Mô hình hóa người dùng trong Social Semantic**Web**2.2.2.4. Giám sát truy nhập trong Social Semantic Web**2.2.2.5. Tái phát hành dữ liệu**2.2.2.6. Những vấn đề liên quan đến bảo mật***2.2.3. Web 3D**

Trải qua quãng thời gian dài phát triển và không ngừng hoàn thiện, giờ đây 3D đã chinh phục rất nhiều lĩnh vực và đang tạo nên một cuộc cách mạng công nghệ mới.

Chính vì vậy mà thông tin sản phẩm nhận được từ Web3D là rất lớn.

2.3. Kết luận chương

Nội dung của chương 2 đi sâu vào các công nghệ của Web 3.0 cũng như từng bước miêu tả kiến trúc bảy lớp của mô hình Tim Berners-Lee. Các khái niệm mới như dữ liệu được liên kết (Linked Data) và Sematic Web cũng được nêu rõ trong chương này. Dữ liệu được liên kết sử dụng URI để liên kết tới một đối tượng dữ liệu hơn là một tài liệu. Đối tượng dữ liệu được mô tả bởi RDF (đảm bảo dữ liệu có ngữ nghĩa), và file RDF nên gồm nhiều hơn các nguồn tài nguyên khác được đánh dấu bởi URI. Việc xuất bản trên Web dựa vào File nguồn tài nguyên thông tin tổng quát có hai loại: nếu yêu cầu đến từ các trình duyệt (trong tiêu đề bao gồm yêu cầu text/html), chúng ta trả về file HTML. Các công nghệ như Ubiquitous Web, Media-Centric Web, Social Web, Web 3D đã được phân tích đánh giá cụ thể trong chương này.

Trên nền lý thuyết của phần đầu, phần cuối của chương tổng hợp, phân tích đánh giá khả năng và xu hướng phát triển của các công nghệ Web 3.0 trên thế giới để làm tiền đề cho chương tiếp theo.

CHƯƠNG 3: KHUYẾN NGHỊ ĐỀ XUẤT VỀ KHẢ NĂNG TRIỂN KHAI ÁP DỤNG CÔNG NGHỆ WEB 3.0 TẠI VIỆT NAM

Chương 3 Trên quan điểm cách tiếp cận của công nghệ Viễn thông và Công nghệ Thông tin nội dung chương đã có cách nhìn tổng thể đánh giá việc tiếp cận công nghệ của các Công ty lớn trên thế giới để từ đó đánh giá tình hình sử dụng Web 2.0 tại Việt nam nhằm đưa ra khuyến nghị triển khai Web 3.0 tại Việt nam.

Phần 3.1 Phân tích tình hình triển khai Web 3.0 trên thế giới từ đó nhìn lại việc sử dụng công nghệ Web 2.0 tại Việt nam phần 3.2. Cách tiếp cận sẽ được phân tích trong phần 3.3. Phần 3.4 là phần quan trọng của luận văn khuyến nghị triển khai Web 3.0 tại Việt nam. Phần 3.5 Kết luận của chương.

3.1. Tình hình triển khai trên thế giới

Bên cạnh sự phát triển của công nghệ viễn thông, công nghệ thông tin nói chung, công nghệ Web cũng có những bước phát tiến vượt bậc trong thời gian vừa qua, mang lại rất nhiều lợi ích không chỉ đối với người cung cấp dịch vụ mà còn đối với những người sử dụng dịch vụ.

Thế hệ Web kế tiếp đón nhận sự gia nhập của các thiết bị cá nhân khác vào Web ngoài computer,

Nền kinh tế phát triển, chúng ta đang sống trong một thế giới phẳng. Lượng người tham gia vào Internet sẽ chỉ có tăng chứ không hề giảm.

Có thể kể ra một vài lĩnh vực áp dụng công nghệ Semantic Web thành công như: Hệ thống lọc ảnh Kodak, Chế tạo thuốc Eli Lilly...

Web 3.0 có cấu trúc phức tạp hơn, bao gồm những công nghệ mạnh hơn qua đó cũng yêu cầu nhiều dung lượng mạng hơn, công nghệ truyền tải cũng phải mạnh mẽ và tối ưu

hóa hơn, đây cũng có thể coi là nền tảng và yêu cầu thiết yếu cho sự phát triển của Web 3.0.

3.2. Tình hình phát triển Web 2.0 tại Việt Nam và những vấn đề cần giải quyết

Web 2.0 được dùng rộng rãi tại Việt Nam, các trong mạng xã hội như Facebook, chia sẻ hình ảnh như Flickr, tìm kiếm bằng Google, chia sẻ video bằng Youtube, định nghĩa ở Wikipedia. Bên cạnh đó cũng có những ứng dụng có thể coi là "được Việt hóa" của những ứng dụng trên, như các mạng xã hội với các trang web tiêu biểu như Yobanbe, Cyworld, Faceviet, . . . ; lĩnh vực tìm kiếm với các trang web tiêu biểu như Baambo, Tim nhanh, Zing Search,...; mảng dịch vụ giải trí với các trang xem video, nghe nhạc và các dịch vụ khác như Clip vn, Yeuanhac.com, Mp3zing, . . . ; lĩnh vực lưu địa chỉ trang web, hay còn gọi là Bookmark với các trang web tiêu biểu như Lin khay.com, Tagvn.com, Vietkicks.com,... Từ đó có thể thấy được quá trình áp dụng triển khai các ứng dụng dịch vụ trên nền tảng công nghệ Web 2.0 đã đạt được những thành công bước đầu tại Việt Nam.

Tuy nhiên có một câu hỏi đặt ra là tại sao số lượng người sử dụng những ứng dụng "Việt hóa" nói trên tại Việt Nam ít hơn số lượng người sử dụng những trang tương tự nhưng "không được Việt hóa

Có thể đưa ra một vài lý do dễ nhận thấy để lý giải những thắc mắc trên.

Thứ nhất giao diện của những ứng dụng Web Việt Nam không thật sự thân thiện và tạo được sự tiện dụng với người dùng.

Thứ hai, tính bảo mật ở những trang web hoặc mạng xã hội Việt Nam chưa cao.

Cuối cùng, chúng ta thường gặp trường hợp tải về hoặc đăng một bức ảnh hoặc video lên những trang mạng của Việt Nam chậm hơn những trang mạng của quốc tế mặc dù khoảng cách địa lý là rất khác nhau.

Việt Nam đang trên con đường hội nhập và phát triển cùng với nền công nghệ thông tin trên thế giới. Bên cạnh những công nghệ khác thì Web 2.0 là lĩnh vực gặt hái được nhiều thành công cũng như rất gần gũi trong đời sống.

Web không chỉ phụ thuộc và độ sẵn sàng của máy chủ mà còn phụ thuộc vào chất lượng của đường truyền.

Tuy nhiên quá trình áp dụng triển khai sẽ mang đặc trưng của Việt Nam.

3.3. Xu hướng tiếp cận Web 3.0 tại Việt Nam

Để có được và khai thác triệt để những lợi ích nêu trên do Semantic Web mang lại và có thể triển khai ứng dụng tại Việt Nam đồng thời mang lại lợi ích cho cộng đồng không chỉ cộng đồng mạng mà còn là cộng đồng xã hội. Thời điểm hiện tại chúng ta đang kế thừa một khối lượng dữ liệu khổng lồ qua giai đoạn phát triển vượt bậc của thế hệ Web 2.0, đây là điều kiện cần để có thể triển khai Web 3.0. Tuy nhiên điều kiện đủ để có thể triển khai đó là dữ liệu của chúng ta phải có cùng một khuôn dạng chung cho phép chia sẻ và tái sử dụng qua các ứng dụng.

Bên cạnh việc xây dựng dữ liệu theo cấu trúc chuẩn, còn cần phải xây dựng các ontology đại diện cho các vùng miền dữ liệu theo từng lĩnh vực.

Để triển khai Web 3.0 tại Việt Nam chúng ta cũng có thể thực hiện một giai đoạn trung gian như tiếp cận phương pháp Data mining.

Công nghệ Web 3D sẽ là một công nghệ hứa hẹn cho nhiều ứng dụng cung cấp dịch vụ trong tương lai tại Việt Nam như dịch vụ cung cấp thông tin thực về sản phẩm tìm kiếm, dịch vụ du lịch trực tuyến ... Tiêu biểu có thể nhắc đến là trang web www.hello3dworld.com (do Công ty Cổ phần Phúc Lê Gia (PLG) phát triển). Sau một năm ra mắt (8/2010), đến nay, PLG đã xây dựng thành công sàn giao dịch thương mại điện tử B2C dưới hình thức tổ chức hội chợ triển lãm trực tuyến.

3.4. Khuyến nghị nghiên cứu, triển khai, áp dụng Web 3.0 (Semantic Web) tại Việt Nam

Mặc dù các công nghệ Semantic Web đã bước đầu được sử dụng, tuy nhiên chúng sẽ không được tiếp nhận rộng rãi cho đến khi các công cụ trở nên hoàn thiện và các vấn đề kinh doanh được giải quyết. Hơn nữa, việc phát triển các ontology có thể yêu cầu sự đầu tư đáng kể. Do đó, về khả năng bước đầu ứng dụng công nghệ Semantic Web ở Việt Nam, chúng tôi khuyến nghị như sau:

3.4.1. Đối với định hướng các hoạt động nghiên cứu phát triển

- *Cần tiếp tục nghiên cứu sự phát triển của các công cụ Semantic Web và sự chuẩn hóa Web ngữ nghĩa.*

- *Nắm bắt các ngữ nghĩa trong một ontology.*

- *Mô hình hóa dữ liệu dựa trên thế giới thực.*

- *Tạo ra nhiều giao diện tới dữ liệu..*

- *Đầu tư vào việc nghiên cứu các thực tế tốt nhất về Semantic Web và các mẫu thiết kế.*

- *Khuyến khích các hoạt động nghiên cứu R&D trong các lĩnh vực Web ngữ nghĩa.*

3.4.2. Đối với việc đào tạo tìm hiểu thông tin về các công nghệ Semantic Web tại Việt Nam, chúng tôi khuyến nghị về các nguồn tài liệu tham chiếu như sau:

+ Các công cụ để phát triển Web ngữ nghĩa: Có một số danh sách trên Web đưa ra tổng quan toàn diện về các công cụ khả dụng khác nhau. Trang Web có thể tham khảo là trang Wiki trên W3C ESW được duy trì bởi W3C ESW cũng như nhiều cộng đồng.

+ Danh sách các sách được tổng hợp ở trang Wiki của W3C, bao gồm hơn 40 quyển sách về Semantic Web được xuất bản bởi các nhà xuất bản uy tín như O'Reilly, MIT Press, Cambridge University Press, Springer Verlag...

+ Các bài báo/xuất bản về Semantic Web có thể tham khảo ở các nguồn sau:

* "*International Semantic Web Conference* " series là sự kiện hàng năm xuất bản các bài viết bởi nhà xuất bản Springer. Trong khi các hội nghị này thường lưu hành trên toàn cầu, "European Semantic Web Conference" và "Asian Semantic Web Conference" series được tổ chức ở Châu Âu và Châu Á.

* "*International World Wide Web Conference* " là hội nghị hàng năm về các công nghệ World Wide Web nói chung, trong đó luôn có một mảng quan trọng về Semantic Web cho cả những nhà học thuật và các cộng đồng phát triển Web. Có thể tham khảo trang của cộng đồng tổ chức... để biết thêm chi tiết về các hội nghị này và các liên kết tới các bài viết.

* Hội nghị *Semantic Technologies* hàng năm trở thành sự kiện quan trọng. Hội nghị này ít tập trung vào các khía cạnh nghiên cứu về Web ngữ nghĩa, mà tập trung vào các khía cạnh kinh doanh, công nghiệp, sự phát triển và các ứng dụng mới.

+ Tìm kiếm các ontology, các thuật ngữ hay các tập dữ liệu cho các ứng dụng: Có một số portal thu thập thông tin về các ontology hiện tại. Ví dụ điển hình là Schemaweb. Ví dụ khác là dịch vụ "Pingthesemanticweb" thu thập thông tin về các tài liệu RDF mới trên Web dựa trên lệnh "pings" được gửi bởi các ứng dụng tạo ra dữ liệu và dựa trên các liên kết phát hiện tự động RDF được tìm thấy bởi những người duyệt Web. Cũng có các engine tìm kiếm khác như Falcon, Sindice hoặc Watson,... chuyên về tìm kiếm các tài liệu Web ngữ nghĩa.

+ Cộng đồng những nhà phát triển Semantic Web chúng ta có thể tham gia: Nhóm W3C Semantic Web Interest Group là một trong những nhóm tốt nhất để tham gia trước tiên. Đồng thời cũng có nhiều cộng đồng khác tập trung vào các lĩnh vực hoặc mục đích cụ thể về Web ngữ nghĩa, gồm:

* DOAP: Dự án mô tả thông tin về các dự án phần mềm nguồn mở.

* FOAF: Dự án mô tả thông tin về con người và các mối quan hệ xã hội.

* SIOC: Dự án mô tả thông tin về các trang cộng đồng trực tuyến các blog, bulletin board,... và sử dụng thông tin này để kết nối các địa chỉ này với nhau.

* Linking Open Data về Web ngữ nghĩa: là dự án có mục đích là làm cho các nguồn dữ liệu mở khác nhau khả dụng trên Web ở dạng RDF và thiết lập các liên kết RDF giữa các hạng mục dữ liệu từ các nguồn dữ liệu khác nhau.

* Planetrdf Blog aggregator tập hợp các blog của một số nhà phát triển Semantic Webtích cực trên toàn thế giới.

3.4.3. Khuyến nghị về chính sách phát triển và ứng dụng triển khai

Xét về khía cạnh quản lý, định hướng phát triển và ứng dụng triển khai, công nghệ web 3.0 và Schematic Web hiện nay là một xu hướng phát triển tiếp theo của công nghệ web; là một xu hướng công nghệ được phát triển nhằm đáp ứng các nhu cầu ngày càng gia tăng trong xã hội về các hình thức trao đổi thông tin, tìm kiếm và chia sẻ thông tin, phục vụ nhu cầu cá nhân, cộng đồng trong mọi mặt hoạt động của đời sống kinh tế xã hội của một quốc gia. Tuy nhiên cũng cần phải nhận thức rằng, giữa việc đáp ứng nhu cầu thực tế và khả năng đáp ứng của công nghệ và các điều kiện đặc thù của Việt Nam còn là một thách thức rất lớn. Nó đòi hỏi phải có những chính sách phát triển và ứng dụng triển khai theo cách tiếp cận hợp lý và một lộ trình trình. Một số định hướng tiếp cần phù hợp có thể là:

- ***Về nghiên cứu phát triển:*** Cần khuyến khích phát triển đội ngũ cán bộ nghiên cứu chuyên sâu và chuyên 'ngiệp về công nghệ Web 3.0 và Semantic Web.

- ***Về mặt đào tạo:*** Tiến hành đưa công nghệ Web 3.0 và Semantic Web như là môn học trong quá trình đào tạo,

giảng dạy ở các trường đại học, và các chương trình nghiên cứu cao học.

- ***Phát triển cơ sở hạ tầng mạng truyền tải thông tin:***

Đây là điều kiện quan trọng, làm cơ sở hạ tầng để ứng dụng công nghệ Web 3.0 và Semantic Web trong tương lai.

- ***Xây dựng và chuẩn hóa các loại hình cơ sở dữ liệu tập trung và phân tán, xây dựng chính sách chia sẻ dữ liệu:***

Có thể nói, các cơ sở dữ liệu là tiền đề và cũng là các đối tượng hướng tới của công nghệ Web 3.0 và Semantic Web trong việc ứng dụng và triển khai các loại hình dịch vụ.

- ***Cần có những định hướng và lộ trình hợp lý về ứng dụng triển khai Web 3.0 và Semantic Web tại Việt Nam:***

Việc xây dựng các định hướng và lộ trình ứng dụng công nghệ Web 3.0 và Schematic Web cần bám sát xu hướng của thế giới về mức độ phát triển, độ chín muồi công nghệ này để có thể ứng dụng triển khai trên thực tế.

Về mặt định hướng chung, việc ứng dụng công nghệ Web 3.0 và Semantic Web cần được đề cập đến và bao hàm trong định hướng phát triển ứng dụng các công nghệ thông tin tiên tiến, hiện đại theo "*Định hướng Chiến lược phát triển Công nghệ thông tin và Truyền thông Việt Nam giai đoạn 2011 – 2020*".

2.3. Kết luận chương

Nội dung của chương 3 ngoài việc khái quát tình hình triển khai tiếp cận Web 3.0 trên thế giới cũng đã đánh giá được tình hình phát triển Web 2.0 tại Việt nam bước đầu đã đạt những thắng lợi. Tuy nhiên việc sử dụng các trang Web trong nước ít hơn nhiều so với những trang nước ngoài do các nguyên nhân như giao diện không thân thiện, độ bảo mật chưa cao và tốc độ chậm là các lý do chính. Từ những đánh giá đó chúng ta cần phải hội tụ cả điều kiện cần và đủ đồng thời giải quyết các khó khăn như xây dựng dữ liệu theo cấu trúc chuẩn,

xây dựng các ontology đại diện cho các vùng miền dữ liệu theo từng lĩnh vực...

Phần nội dung quan trọng nhất của chương và cũng là của luận văn này là đưa ra được khuyến nghị triển khai áp dụng tại Việt nam cùng với các định hướng và khuyến nghị cụ thể bao gồm định hướng các hoạt động nghiên cứu phát triển, khuyến nghị đối với việc đào tạo tìm hiểu thông tin về các công nghệ Semantic Web3.0 và định hướng và chính sách phát triển và ứng dụng triển khai tại Việt nam. .

KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TIẾP THEO

1. Kết luận

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. Luận văn đã từng bước đi từ lý thuyết từ kiến trúc các lớp, đến xu hướng phát triển các công nghệ Web 3.0 để từ đó phân tích đánh giá hiện trạng tình hình, các khó khăn, vướng mắc cần giải quyết khi triển khai trong nước. Với các khuyến nghị, định hướng rõ ràng đi kèm một lộ trình triển khai hợp lý có thể hy vọng rằng tương lai thực sự của Web3.0 tại Việt nam sẽ ngày càng tốt đẹp và vượt xa hơn nhiều những gì đã được dự đoán.

2. Hướng nghiên cứu tiếp theo

Học viên sẽ tiếp tục nghiên cứu, cập nhật các công cụ phát triển hỗ trợ của Web 3.0, triển khai thử nghiệm với mục tiêu đưa Web 3.0 vào thực tế cuộc sống theo đúng định hướng chung phát triển ứng dụng các Công nghệ Thông tin tiên tiến và hiện đại.