

ĐẠI HỌC BÁCH KHOA THÀNH PHỐ HỒ CHÍ MINH
KHOA KHOA HỌC VÀ KỸ THUẬT MÁY TÍNH



BÁO CÁO LUẬN VĂN TỐT NGHIỆP

Trực quan dữ liệu với thư viện C9js

Hội đồng:

GVHD: TS. Lương Thế Nhân

GVHD: TS. Huỳnh Tường Nguyên

GVPB: ThS. Trần Giang Sơn

Sinh viên thực hiện:

Phạm Thành Công - 51200399

Đỗ Đăng Thanh Huy - 51201337

Thành phố Hồ Chí Minh, 05/2016

Mục lục

Danh sách hình vẽ	5
Danh sách mã nguồn	6
1 Giới thiệu đề tài	7
1.1 Bối cảnh	7
1.1.1 Trực quan hoá dữ liệu	7
1.1.2 Ứng dụng trên nền tảng Web	8
1.2 Lí do chọn đề tài	9
1.3 Mục tiêu đề tài	10
2 Kiến thức nền tảng và Công nghệ	12
2.1 Công trình liên quan	12
2.1.1 Công cụ hỗ trợ trực quan dữ liệu phân tán Gapminder	12
2.1.2 Global Health Atlas - WHO	14
2.1.3 Thư viện đồ thị C3.js	17
2.1.4 Ứng dụng tạo bản đồ online Mapme	20
2.2 Kiến thức nền tảng	24
2.2.1 JavaScript	24
2.2.2 D3.js	26

2.2.3	OpenLayers	27
3	Giới thiệu thư viện C9js	30
3.1	Mô tả bài toán	30
3.2	Hướng dẫn sử dụng	30
3.2.1	Cài đặt (Setup)	31
3.3	Các ví dụ minh hoạ	33
3.4	Ứng dụng thực tế	33
4	Quá trình thực hiện	34
4.1	Kiến trúc xây dựng	34
4.2	Hiện thực thư viện	34
5	Kết luận	35
5.1	Đánh giá	35
5.2	Hướng phát triển	35
5.3	Kết luận	35
	Tài liệu tham khảo	36

Danh sách hình vẽ

1.1	Thông số đánh giá nhận được từ GitHub	11
2.1	Đồ thị trực quan dữ liệu Gapminder World	13
2.2	Biểu đồ tương tác về mật độ bác sĩ của WHO	15
2.3	Tương tác khi click vào một hàng trong bảng dữ liệu	16
2.4	Tương tác khi hover lên Map	17
2.5	Thư viện đồ thị C3js	18
2.6	Thư viện đồ thị C3js	18
2.7	Thư viện đồ thị C3js	19
2.8	Thư viện đồ thị C3js	20
2.9	Thư viện đồ thị C3js	21
2.10	Ứng dụng tạo bản đồ online Mapme	21
2.11	Ứng dụng tạo bản đồ online Mapme	22
2.12	Ứng dụng tạo bản đồ online Mapme	23
3.1	Các tính năng chính của C9js trên trang chủ c9js.me	31
3.2	Menu hướng dẫn trên trang chủ c9js.me	31
3.3	Trang hướng dẫn trên Website c9js.me	32

Danh sách Mã nguồn

1	Tải mã nguồn thông qua npmcdn	32
2	Tải mã nguồn thông qua unpkg	33
3	Tải mã nguồn <i>d3.js</i> trước tiên	33
4	Tải mã nguồn <i>OpenLayers 3</i> nếu muốn sử dụng chức năng Bản đồ	33

1 Giới thiệu đề tài

1.1 Bối cảnh

1.1.1 Trực quan hoá dữ liệu

Khái niệm

Trực quan dữ liệu (Data visualization), được xem như hướng đi trong tương lai của việc giao tiếp và truyền thông thông qua hình ảnh. Nó liên quan đến các công trình và nghiên cứu nhằm biểu diễn dữ liệu một cách trực quan, hay nói cách khác, được hiểu là "thông tin được trực quan hoá dưới dạng các lược đồ, bao gồm các thuộc tính và thông số đại diện cho các đơn vị thông tin^[?]]".

Lịch sử hình thành

Bắt đầu từ thế kỷ thứ 2 với việc sắp xếp dữ liệu vào các cột và các hàng, phát triển thành biểu diễn dưới dạng các đơn vị định lượng ban đầu trong thế kỷ 17^[?]]. Theo Hiệp hội Thiết kế Tương tác (The Interaction Design Foundation), nhà triết học và toán học người Pháp René Descartes đã đưa ra những nghiên cứu nền tảng đầu tiên làm bước đệm cho một kỹ sư người Scotland tên là William Playfair. Descartes đã phát triển một hệ trục tọa độ hai chiều để hiển thị các giá trị, vào những năm cuối thế kỷ 18, Playfair đã nhìn thấy tiềm năng trong việc thể hiện đồ họa của dữ liệu định lượng^[?]].

Trong nửa sau của thế kỷ 20, Jacques Bertin sử dụng đồ thị định lượng để đại diện cho thông tin "trực giác, rõ ràng, chính xác và hiệu quả".^[?]] John Tukey và đáng chú ý hơn là Edward Tufte đã thúc đẩy các giới hạn của trực quan dữ liệu. Tukey đưa ra cách tiếp cận mới của ông theo hướng thống kê: Phân tích dữ liệu thăm dò, và Tufte đã xuất bản cuốn sách "The Visual Display of Quantitative Information". Các đóng góp trên đã mở ra con đường rõ ràng hơn cho các công nghệ trực quan.

Tiềm năng của Trực quan dữ liệu

Ngày nay trong lĩnh vực kinh doanh hiệu quả (Business Intelligence) không có gì có thể mang chúng ta lại gần hơn với những tương tác thông minh hơn là Trực quan dữ liệu. Nhưng điều này chỉ xảy ra khi chúng ta thực sự hiểu và dùng nó một cách đúng đắn. Để đạt được điều đó, chúng ta phải thực sự hành động và vứt bỏ những quan niệm chưa đúng về Trực quan dữ liệu.

Trực quan dữ liệu ngày càng đóng vai trò quan trọng trong mọi lĩnh vực đời sống. Như sử dụng trong nghiên cứu, trong những công việc liên quan tới xử lý dữ liệu, giao dịch bởi người dùng phổ thông, và được áp dụng với tỷ lệ tăng cao trong lực lượng lao động trí óc, đặc biệt đối với các nhà phân tích.

Với sự phát triển của công nghệ cùng với sự tiến triển của dữ liệu trực quan; bắt đầu với hình ảnh vẽ tay và chuyển biến thành các ứng dụng kỹ thuật - bao gồm thiết kế tương tác cho đến các phần mềm.^[?] Các chương trình như SAS, SOFA, R, Minitab, và nhiều hơn nữa khởi đầu cho việc trực quan dữ liệu trong lĩnh vực thống kê. Các ứng dụng khác, tập trung nhiều hơn vào cá nhân hoá, các ngôn ngữ lập trình như D3, Python và JavaScript giúp cho việc hiện thực trực quan khả thi và dễ dàng hơn.

1.1.2 Ứng dụng trên nền tảng Web

Sự tăng trưởng trong khối lượng dữ liệu trong các dự án với nhu cầu chia sẻ mang tính toàn cầu ngày càng gia tăng, đòi hỏi sự đáp ứng bởi các công nghệ trên nền tảng Web để phân tích, xử lý, tương tác và hiển thị dữ liệu ngày càng lớn.

Các công nghệ trên nền Web hiện đại như HTML5, CSS3 và các thư viện JavaScript mạnh mẽ mang đến những triển vọng về trực quan trên các trình duyệt hiện tại. Hiện tại đã có hàng chục, thậm chí hàng trăm các loại thư viện, framework và ứng dụng phục vụ cho nhu cầu trực quan tùy từng mục đích cụ thể của cá nhân, doanh nghiệp, các công trình nghiên cứu,...

1.2 Lí do chọn đề tài

Đam mê trong các công nghệ mới trên nền tảng Web

Ngày nay,^[?]] Web ngày càng phát triển và trở thành một thế lực riêng, một đế chế riêng bao gồm các *Web pages* và *Web apps*, xây dựng nên từ các video, hình ảnh hay các nội dung tương tác. Điều mà người dùng phổ thông không để ý đến đó là các công nghệ Web và các trình duyệt đã đóng góp rất lớn trong quá trình trên.

Qua thời gian, các công nghệ Web tiến hoá và ngày càng đem lại nhiều trải nghiệm mới lạ cho các lập trình viên, nhằm đáp ứng khả năng sáng tạo và tạo mọi điều kiện để họ thể hiện khả năng thông qua nó.

Nhóm đề tài gồm các thành viên từng làm việc với nhiều dự án liên quan đến Web. Đặc biệt đều có hứng thú và đam mê với khả năng của các xu hướng phát triển hiện tại, cũng như muốn khám phá khả năng bản thân trong thế giới luôn luôn đổi mới không ngừng của Web, đó là một trong những lí do nhóm chọn đề tài này.

Đóng góp vào sự phát triển của trực quan dữ liệu

^[?]] Có thể nói, với sự phát triển khoa học và công nghệ như vũ bão vào những năm cuối thập niên này, những sản phẩm công nghệ mới phát triển rầm rộ đã đem lại nhiều tiện ích cho cuộc sống, nó được ứng dụng rộng rãi ở nhiều lĩnh vực; đặc biệt là lĩnh vực thông tin truyền thông. Hệ thống thông tin điện tử, trực tuyến, các Website của các tổ chức, đơn vị, doanh nghiệp được phát triển mạnh mẽ, góp phần tăng cường mối quan hệ, giao lưu, hợp tác phát triển ở nhiều lĩnh vực, nhất là lĩnh vực kinh tế, văn hóa xã hội, khoa học công nghệ, y tế, giáo dục, giải trí, con người có trong tay nhiều công cụ để chia sẻ thông tin của mình qua blog, website, diễn đàn, qua các trang mạng xã hội như Facebook, Youtube, Twitter.

Trong thời đại bùng nổ thông tin hiện nay, được tiếp sức bởi các loại công nghệ tiên các phương tiện thông tin hiện đại đã giúp cho thế giới xích lại gần nhau một cách nhanh chóng hơn. Đặc biệt, công nghệ truyền thông Internet đang phát triển như vũ bão, tạo nên những siêu lộ thông tin có dung lượng lớn và tốc độ cao, chuyển tải mọi thông tin một cách nhanh

chóng.

Trong thời đại mà thông tin được coi như một loại tài sản, việc kết hợp trực quan dữ liệu với Web có thể được xem như một xu thế mới, một hướng đi đầy tiềm năng trong tương lai. Với mong muốn đóng góp vào quá trình này, đề tài *Trực quan dữ liệu với C9js* có thể xem như một viên gạch trong bức tường đang xây dựng của trực quan dữ liệu.

1.3 Mục tiêu đề tài

Rút ngắn thời gian cho các lập trình viên

Mang trực quan dữ liệu trở thành một ứng dụng trên nền Web vừa dễ vừa khó. Dễ ở việc các công nghệ Web rất đa dạng, phong phú, thậm chí là bao la khi đáp ứng các nhu cầu của trực quan dữ liệu. Nhưng đó cũng là điểm hạn chế, khi có quá nhiều thư viện, nền tảng, công nghệ lập trình viên cần nắm rõ và kết hợp nếu muốn hiện thực ý tưởng thành một ứng dụng mang tính thực tế trên Web.

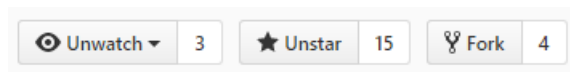
Bản thân nhóm cũng là các lập trình viên thích khám phá và nghiên cứu các công nghệ mới và phong phú trong thế giới Web. Mục tiêu thư viện C9js đề ra là giúp các lập trình viên khác khi sử dụng C9js sẽ cảm thấy dễ dàng nắm bắt, không đòi hỏi quá nhiều nỗ lực trong việc tìm hiểu hay xây dựng một ví dụ mẫu. Mặt khác, giảm thiểu tối đa thời gian cần thiết để đưa trực quan dữ liệu lên Web của mình.

Hướng đến cộng đồng mã nguồn mở

C9js hiện tại được triển khai là một mã nguồn mở (open-source) tại GitHub^[?]. Các công nghệ được sử dụng trong C9js hoàn toàn là mã nguồn mở, nên nhóm quyết định đóng góp vào cộng đồng một thư viện với khả năng mang trực quan dữ liệu trở thành một công cụ dễ dàng triển khai trên Web.

Về kiến trúc, cũng như tài liệu (documentation), C9js được thể hiện rất rõ ràng với mong muốn ngay từ ban đầu là nhận được sự đóng góp cũng như phát triển lâu dài từ các lập trình viên khác trong cộng đồng mã nguồn mở GitHub. Hiện tại, repository^[?] của C9js trên

GitHub đã nhận được sự quan tâm và đánh giá từ cộng đồng. Các thông số cụ thể (tính đến thời điểm bài báo cáo này) như hình 1.



Hình 1.1: Thông số đánh giá nhận được từ GitHub

2 Kiến thức nền tảng và Công nghệ

2.1 Công trình liên quan

2.1.1 Công cụ hỗ trợ trực quan dữ liệu phân tán Gapminder

Giới thiệu

Gapminder^[?]] được thành lập tại Stockholm bởi Ola Rosling, Anna Rosling Rönnlund và Hans Rosling vào ngày 25 tháng 2 năm 2005.

Gapminder là một tổ chức phi lợi nhuận nhằm thúc đẩy sự phát triển bền vững toàn cầu, và là một thành tựu đánh dấu mục tiêu phát triển mang tầm thế kỷ của tổ chức Liên Hiệp Quốc (United Nations). Với mục tiêu đề ra nhằm gia tăng khả năng tìm hiểu, phân tích và đánh giá các thống kê, dữ liệu liên quan đến các vấn đề xã hội, tài chính và phát triển môi trường ở địa phương, quốc gia, và tầm thế giới.

Gapminder hoạt động dựa trên nền tảng được xác định trước, xây dựng các dịch vụ bởi ban quản trị, đôi khi hợp tác với các dự án tại các trường đại học, các tổ chức quốc tế, cơ quan Nhà nước cũng như các tổ chức phi chính phủ.

Mục tiêu

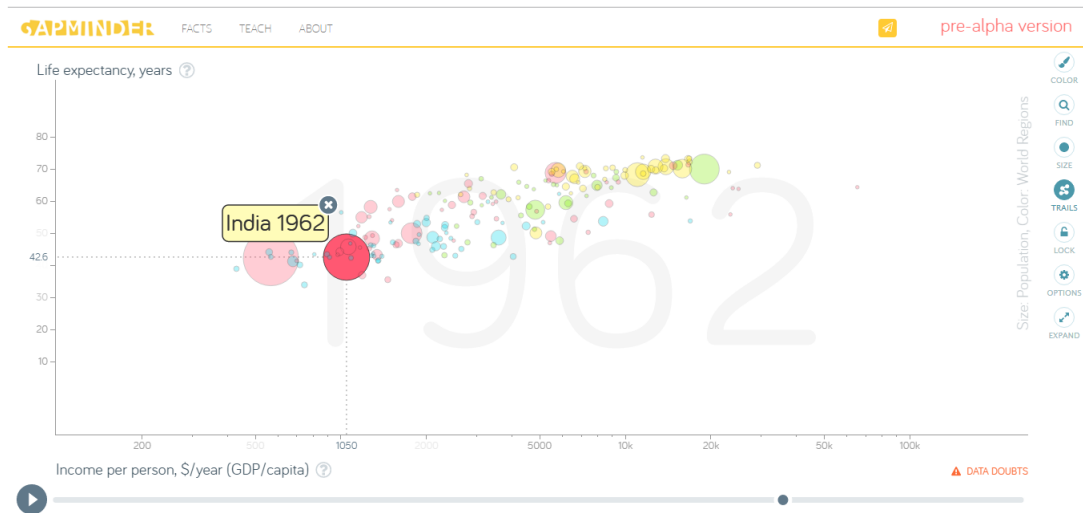
Các hoạt động ban đầu là theo đuổi sự phát triển của ứng dụng Trendalyzer (tháng 3 năm 2006). Trendalyzer là một phần mềm biểu diễn chuỗi thời gian thống kê bằng cách chuyển đổi những con số khô khan, nhàm chán thành các đồ thị chuyển động thú vị, và đặc biệt là khả năng tương tác. Trendalyzer còn được biết đến dưới tên gọi Gapminder World, xây dựng trên nền tảng web biểu diễn dữ liệu thống kê theo dòng thời gian cho tất cả các quốc gia trên thế giới.

Gapminder World^[?]] cung cấp các video, bài thuyết trình bằng Flash và các đồ thị dưới dạng PDF thể hiện xu hướng phát triển mang tính toàn cầu.

Kể từ năm 2006 Gapminder cung cấp một lượng lớn các video thể hiện thực tế dưới góc

nhìn toàn cầu.

Từ năm 2008 Gapminder còn cung cấp các biểu đồ con (sub-graph)^[?] thể hiện xu hướng, sự khác biệt giữa các tiểu bang của Mỹ, các tỉnh thành của Trung Quốc, hay các đô thị đặc biệt so sánh số liệu thống kê từ các thành phố ở Trung Quốc, Mỹ, Ấn Độ với các quốc gia khác trên thế giới. Từ Gapminder World, chỉ bằng một cú click chuột, các dữ liệu, thống kê liên quan nằm sau các đô thị được thể hiện trên các biểu đồ con.



Hình 2.1: Đồ thị trực quan dữ liệu Gapminder World

Như ở hình 2, chúng ta có thể thấy với 1 lượng dữ liệu lớn (ở đây là độ tuổi kỳ vọng ở các quốc gia trên thế giới), có thể được thể hiện trực quan bởi đồ thị bong bóng (bubble chart). Qua kích thước các bong bóng, có thể thấy sự tương quan khác biệt giữa các quốc gia như thế nào. Hay khi trở vào 1 bong bóng, chúng ta thấy được số liệu chi tiết ở quốc gia đó, như trong hình là độ tuổi kỳ vọng ở Ấn Độ, vào năm 1962 là 42.6. Ngoài ra ở phía dưới đồ thị, có 1 thanh thể hiện dòng thời gian (timeline), khi thanh thời gian trượt đi, kích thước các quả bóng thay đổi theo như số liệu tương ứng ở các năm của các quốc gia.

Phương pháp

Gapminder sử dụng D3.js, một thư viện JavaScript phổ biến trong việc thao tác với các thành phần của một trang web (Document Object Model) dựa trên dữ liệu.

D3 cho phép thể hiện dữ liệu thông qua HTML, SVG và CSS. D3 còn là một tiêu chuẩn web cung cấp hầu hết các khả năng mà một trình duyệt hiện đại mang lại, mà không cần phải tự viết lại hay dựa trên một nền tảng khuôn khổ. Kết hợp các thành phần trực quan mạnh mẽ và hướng tiếp cận dữ liệu để thao tác trên DOM.

Đánh giá

Với mục tiêu luôn cập nhật các công cụ cũng như dữ liệu, nhằm đảm bảo tính xác thực, và đúng đắn của ứng dụng, Gapminder World đảm bảo nguồn dữ liệu chính xác nhất và cập nhật nhất. Dữ liệu đề tài sẽ sử dụng từ Gapminder để đảm bảo nguồn thông tin luôn xác thực và thể hiện đúng đắn nhất các xu hướng phát triển toàn cầu.

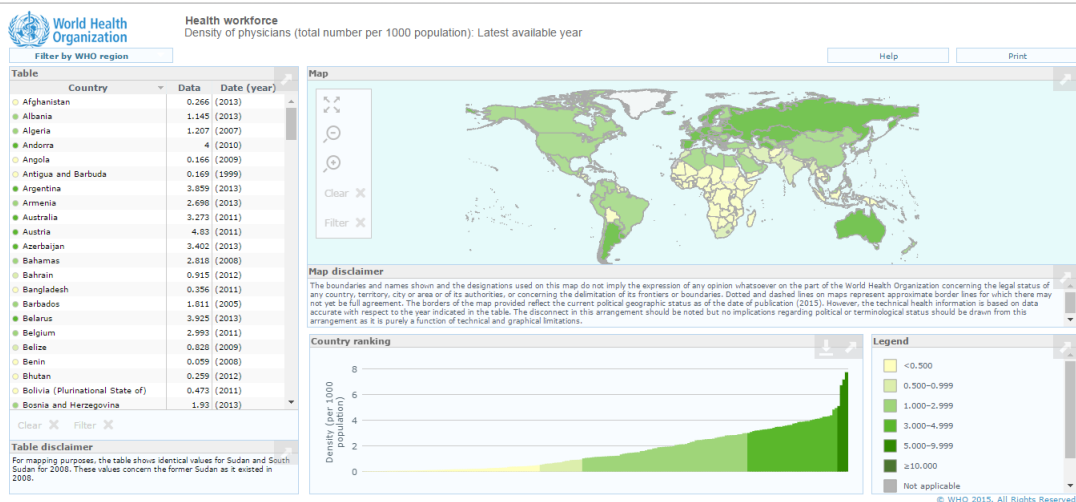
Ngoài ra, cách sử dụng các biểu đồ con để hiển thị dữ liệu liên quan của một đối tượng được tương tác giúp người dùng dễ dàng nắm bắt được chính xác các thông tin muốn truyền đạt, tăng tính thân thiện cũng như trải nghiệm người dùng. Áp dụng tính hiệu quả trên, đề tài đề xuất sử dụng cách thể hiện trải nghiệm người dùng (UI/UX) kế thừa các đặc tính của Gapminder World.

Về thư viện D3.js được sử dụng trong Gapminder World, đây là một thư viện mạnh mẽ tiếp cận dữ liệu dựa trên các thao tác với DOM. D3.js cho phép tái sử dụng các thành phần và trình nhúng (plugin) một cách đơn giản và hiệu quả. Đề tài đề xuất sử dụng D3.js làm thư viện chính để hiện thực thư viện tương tác giữa các đồ thị.

2.1.2 Global Health Atlas - WHO

Giới thiệu

Trong một môi trường điện tử, WHO's Communicable Disease Global Atlas được dùng cho việc phân tích và so sánh dữ liệu chuẩn, thống kê các bệnh truyền nhiễm ở các quốc gia, vùng lãnh thổ, và toàn thế giới. Việc phân tích và biểu diễn dữ liệu ngày càng được hỗ trợ tốt hơn thông qua các thông tin về dân số, điều kiện kinh tế xã hội, và các yếu tố môi trường. Với việc làm đó, biểu đồ Atlas này có thể cho biết được những yếu tố gây ra bệnh truyền nhiễm (2007)^[7].



Hình 2.2: Biểu đồ tương tác về mật độ bác sĩ của WHO

Mục tiêu

Ban đầu, với mục đích cung cấp những dữ liệu, báo cáo về các bệnh truyền nhiễm chính (sốt rét, HIV/AIDS,..), thì ngày nay Global Atlas của WHO cũng dùng để phân tích, biểu diễn, so sánh dữ liệu từ những vấn đề liên quan đến y tế khác như: mật độ bác sĩ, số lượng người chết do tai nạn giao thông, tuổi thọ trung bình,...

Phương pháp

Trực quan dữ liệu thành nhiều dạng biểu đồ khác nhau, các biểu đồ có thể tương tác với nhau giúp cho việc nhìn nhận, phân tích, đánh giá thông tin một cách chính xác, đầy đủ.

Ứng dụng^[?] bao gồm các thành phần:

- *Bảng dữ liệu*: Được lấy từ database của WHO.
- *Bản đồ (Map)*: Dùng thư viện Google Maps, kết hợp với bảng dữ liệu để tạo ra biểu đồ.
- *Biểu đồ cột*: Dựa vào bảng dữ liệu để tạo ra biểu đồ.
- *Chú thích*: Phân vùng dữ liệu.

Tính năng: Ứng dụng hỗ trợ tương tác giữa các biểu đồ, bảng dữ liệu

- Khi click vào một hàng trong bảng dữ liệu (hình 4), bên map sẽ phóng to đến vùng tương ứng, bên biểu đồ cột sẽ nổi bật cột tương ứng. Các tương tác sẽ được thể hiện tương tự khi click vào các biểu đồ.



Hình 2.3: Tương tác khi click vào một hàng trong bảng dữ liệu

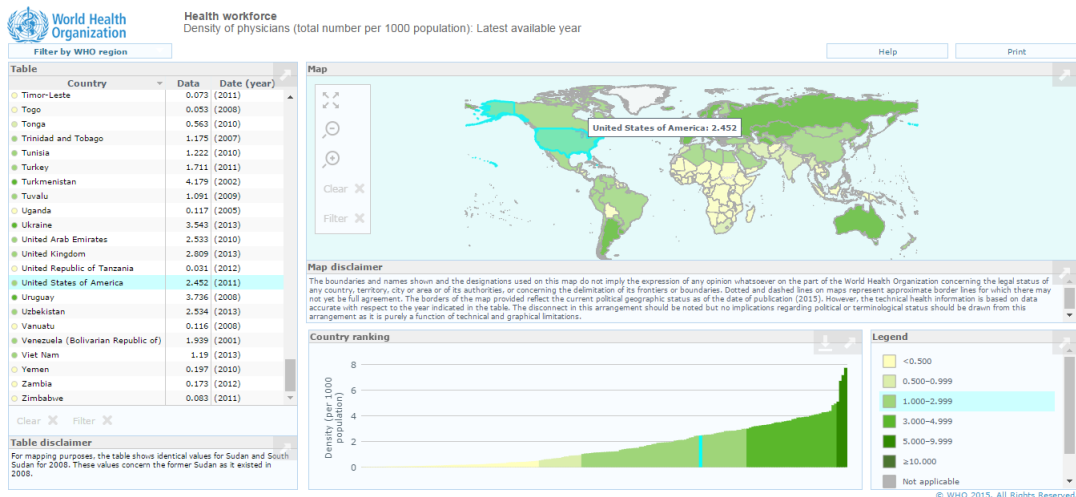
- Khi hover lên một hàng trong bảng dữ liệu (hình 5), bên các biểu đồ cũng sẽ được làm nổi bật vùng tương ứng như khi click, nhưng với 1 màu sắc khác (bên map sẽ không phóng to đến vùng tương ứng).

Đánh giá

Dữ liệu được WHO thu thập trên toàn thế giới, cập nhật theo thời gian, đảm bảo tính chính xác của ứng dụng.

Bên cạnh đó, việc trực quan thành nhiều loại biểu đồ, thể hiện được nhiều cái nhìn khác nhau của dữ liệu giúp đánh giá vấn đề 1 cách đúng đắn, đầy đủ.

Một tính năng khác đó là việc tương tác giữa các đồ thị giúp người dùng có nhiều cái nhìn từ nguồn dữ liệu, hiểu được thông tin rút trích từ các đồ thị, tăng tính thân thiện,



Hình 2.4: Tương tác khi hover lên Map

trải nghiệm của người dùng.

Do đó, đề tài đề xuất sử dụng nhiều đồ thị để biểu diễn cho nguồn dữ liệu đầu vào, và sự tương tác giữa các đồ thị.

2.1.3 Thư viện đồ thị C3.js

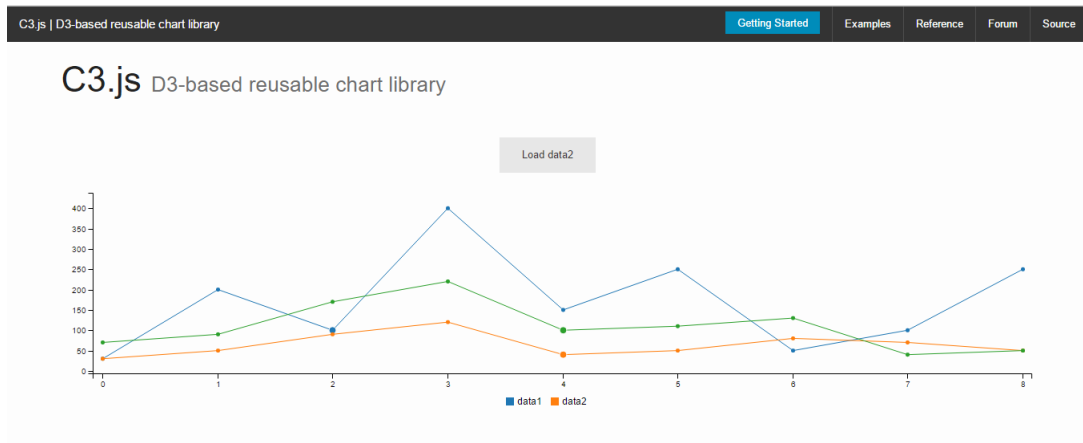
Giới thiệu

C3.js là một thư viện mã nguồn mở^[7] dựa trên thư viện D3.js, nhằm cho phép người dùng có cái nhìn sâu hơn về các ứng dụng web tích hợp đồ thị.

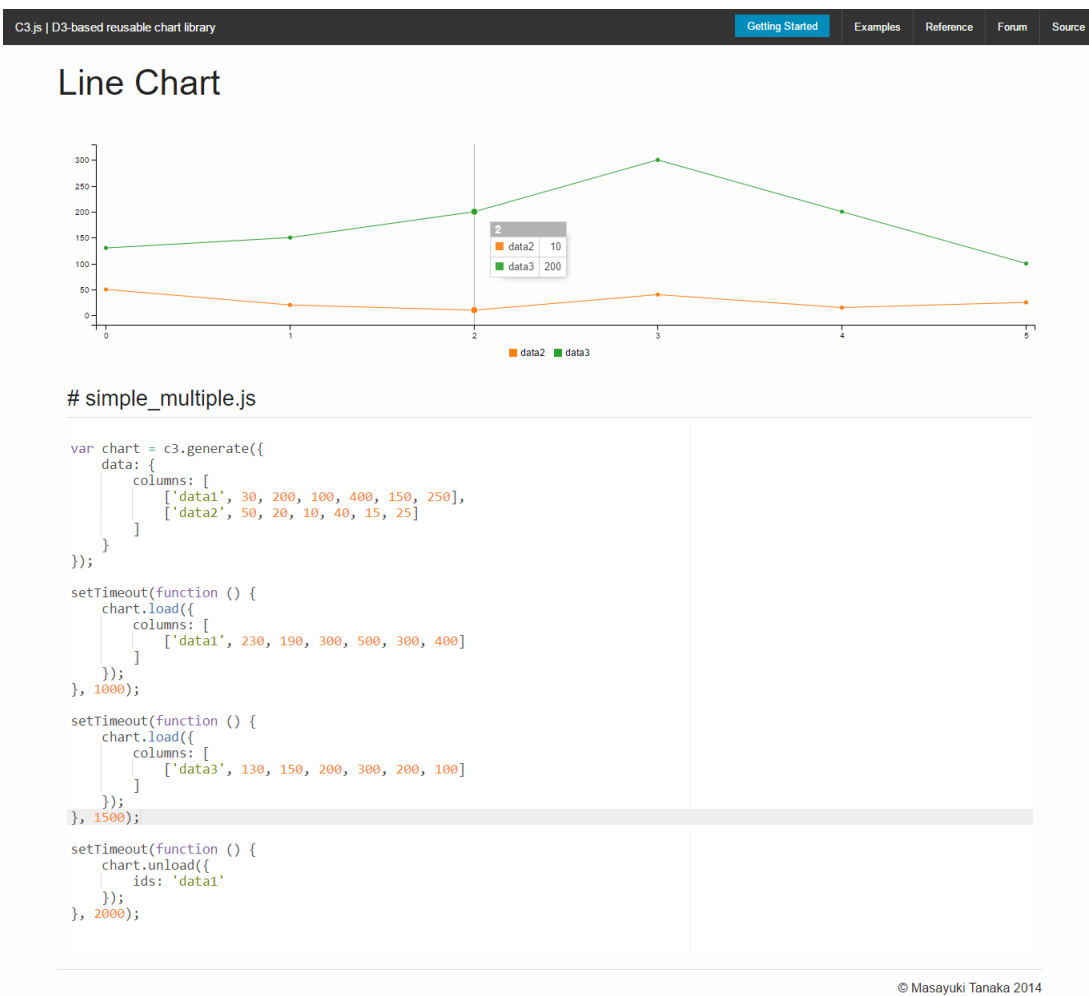
Toàn bộ mã nguồn của thư viện nằm ở: <https://github.com/masayuki0812/c3>

Mục tiêu

- *Sự thoải mái (Comfortable)*: C3 giúp người dùng có thể dễ dàng tạo một đồ thị dựa trên D3 bằng cách gộp chung các đoạn code cần thiết để cấu trúc một đồ thị hoàn chỉnh. Qua đó, người dùng không cần phải viết các đoạn code D3 rườm rà như trước nữa.



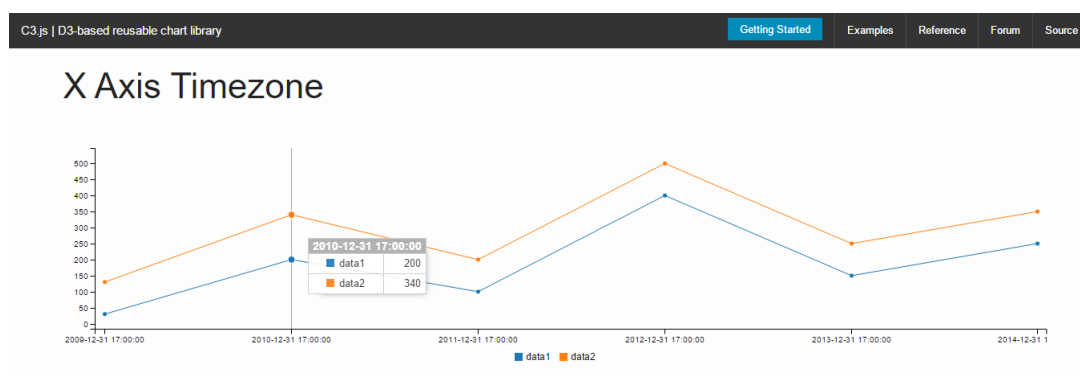
Hình 2.5: Thư viện đồ thị C3js



Hình 2.6: Thư viện đồ thị C3js

[Hình 6] Thông qua một số câu lệnh định nghĩa dữ liệu, C3 khởi tạo một đồ thị cho phép tương tác một cách trực quan thông qua hover, click,... các thành phần của đồ thị.

- *Khả năng tùy chỉnh (Customizable)*: C3 cung cấp các class tương ứng cho mỗi đối tượng trong đồ thị khi khởi tạo. Qua đó, người dùng có thể định nghĩa các style theo ý muốn theo class cũng như mở rộng cấu trúc các thành phần của đồ thị (trục dữ liệu, hình dạng hiển thị, kiểu đồ thị,..) một cách trực tiếp.



Hình 2.7: Thư viện đồ thị C3js

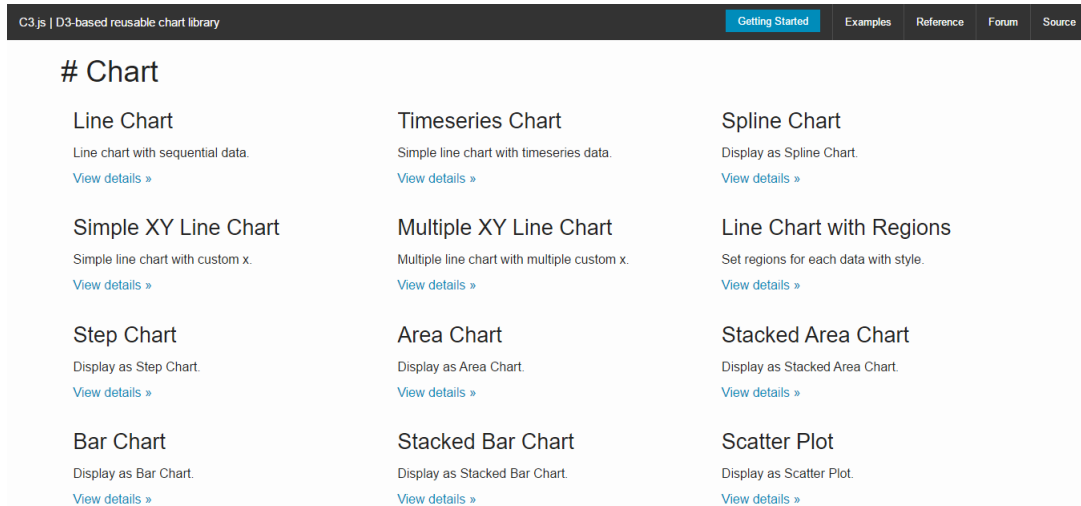
(Hình 7) Trục hoành của đồ thị được tùy chỉnh theo thời gian hiện tại với định dạng hoàn toàn tùy thuộc người sử dụng.

- *Khả năng kiểm soát (Controllable)*: C3 cung cấp đa dạng các APIs và các hàm callback để truy xuất trạng thái của đồ thị. Bằng cách đó, người dùng có thể cập nhật lại đồ thị thậm chí sau khi đồ thị đã được thể hiện. [Hình 8] Thể hiện một số thành phần được hỗ trợ bởi thư viện C3js.

Phương pháp

Như đã trình bày ở trên, C3js sử dụng thư viện D3js nhằm tạo các đồ thị có khả năng tái sử dụng, cũng như giúp người dùng đào sâu hơn vào khái niệm trực quan dữ liệu thông qua đồ thị trên nền ứng dụng web.

Ngoài ra dựa theo cấu trúc mã nguồn mà C3js chia sẻ, có thể thấy cấu trúc tổ chức của C3js



Hình 2.8: Thư viện đồ thị C3js

theo hướng hướng đối tượng, được tổ chức theo các class (trục, dữ liệu, đồ thị,..) và các phương thức, thuộc tính đi kèm class.

Đánh giá

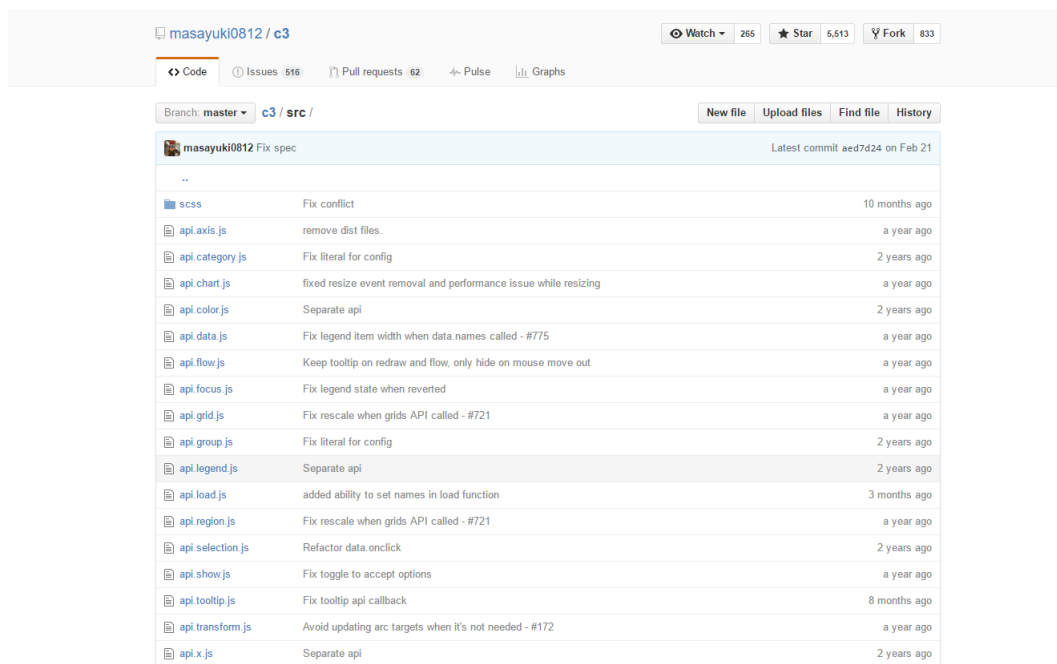
Cách tổ chức và hiện thực theo hướng hướng đối tượng như trên giúp tiết kiệm số dòng mã, tận dụng khả năng tái sử dụng, đồng thời giúp việc mở rộng về sau rất hiệu quả.

Đề tài đề xuất sử dụng cách hiện thực API theo hướng như thư viện C3js.

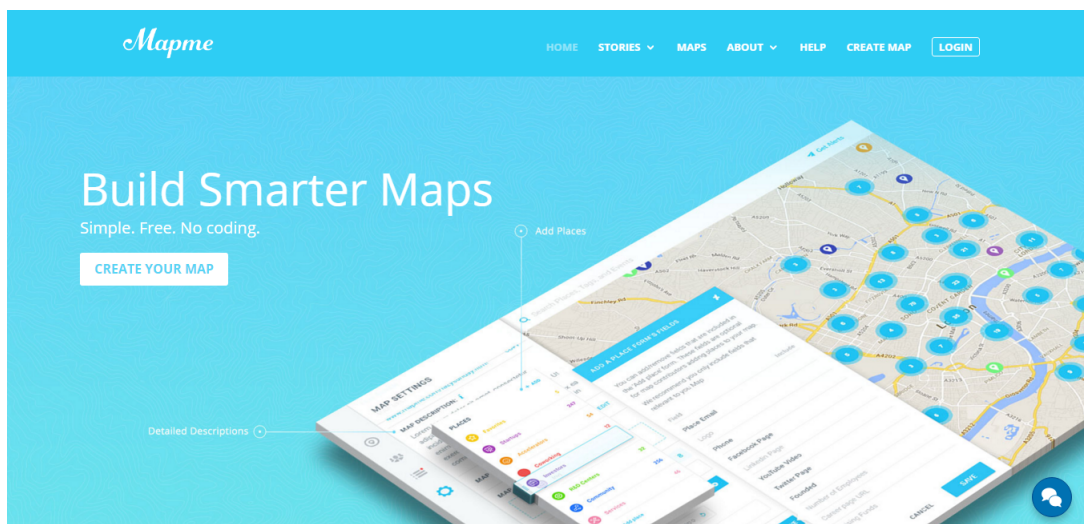
2.1.4 Ứng dụng tạo bản đồ online Mapme

Giới thiệu

Mapme^{[?]1} là một ứng dụng cho phép người dùng tạo bản đồ, dữ liệu được xử lý trực tiếp từ các file Excel đầu vào. Bản đồ tạo ra có các marker tương ứng với dữ liệu, có khả năng tương tác trực tiếp thông qua các sự kiện định nghĩa trước. Ngoài ra, bản đồ được tạo ra còn có thể được nhúng vào các trang web khác thông qua một đoạn mã nhúng. Đặc biệt toàn bộ tương tác từ những bước khởi tạo bản đồ đầu tiên, đến lúc xuất thành quả, đều không yêu cầu người dùng phải viết một dòng mã nào.



Hình 2.9: Thư viện đồ thị C3js



Hình 2.10: Ứng dụng tạo bản đồ online Mapme

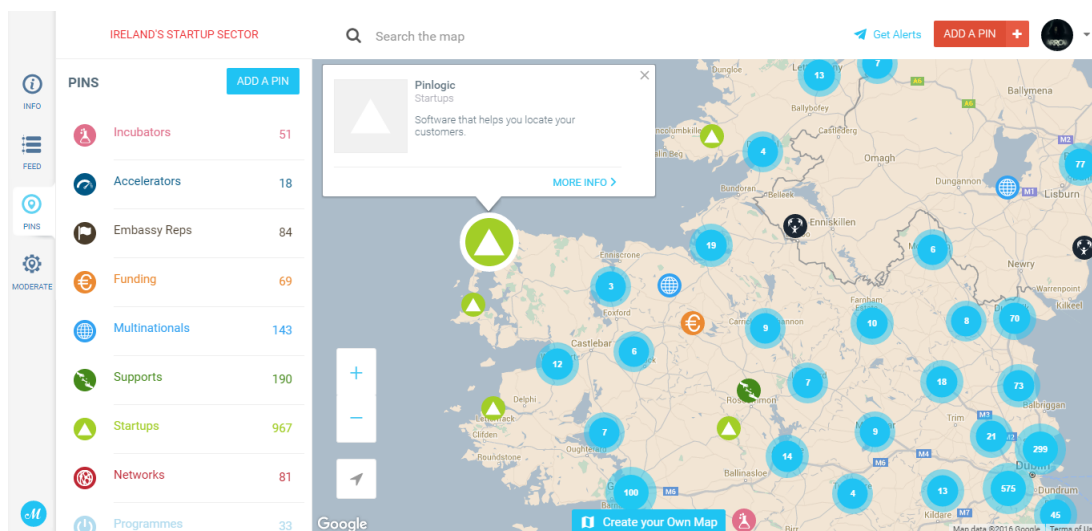
Mục tiêu

Đầu tiên nói đến các đối tượng mà Mapme hướng tới, đó là các tổ chức, chính phủ, các doanh nghiệp hay các tổ chức cộng đồng phi lợi nhuận.

Phục vụ trong các lĩnh vực khác nhau như giáo dục, các sự kiện, xuất bản hay truyền thụ cảm hứng trong đời sống hằng ngày.

Phương pháp

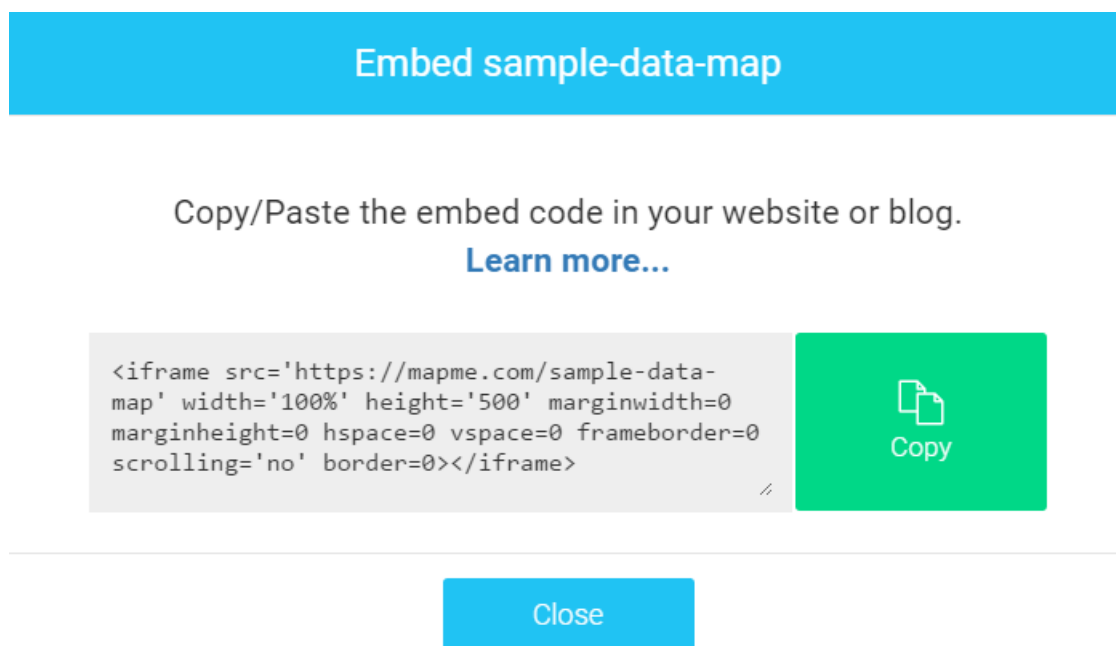
Sử dụng OpenStreetMap (viết tắt OSM)^[7], một dịch vụ bản đồ thế giới trực tuyến với nội dung mở, nhằm mục đích cung cấp dữ liệu địa lý do nhiều người cùng cộng tác với nhau trên hệ thống Wiki. Nó thường được gọi "Wikipedia của bản đồ". Dự án OSM được sáng lập năm 2004, chủ yếu để cạnh tranh với các công ty và cơ quan chính phủ cung cấp dữ liệu địa lý theo các điều khoản sử dụng được coi là quá chặt chẽ. OSM cho phép ứng dụng cập nhật dữ liệu địa lý trực tuyến một cách chính xác, luôn được cập nhật nhằm đảm bảo các dữ liệu luôn đi sát thực tế nhất.



Hình 2.11: Ứng dụng tạo bản đồ online Mapme

Như hình 11 có thể thấy, các thông tin về khởi nghiệp ở quốc gia Ireland được trực quan thông qua các biểu tượng (marker). Cùng với đó là các số liệu tương ứng được xử lý trực tiếp từ các file Excel đầu vào. Ngoài ra, người dùng còn có thể tương tác với các marker này

để thể hiện thông tin chi tiết ở các điểm chỉ định.



Hình 2.12: Ứng dụng tạo bản đồ online Mapme

Đặc biệt, Mapme còn hỗ trợ sinh mã nhúng [Hình 12], dưới dạng iframe, giúp người dùng nhúng trực tiếp bản đồ mình vừa tạo vào các website tùy mục đích.

Đánh giá

Đầu tiên là về cách sử dụng OpenStreetMap làm dịch vụ cung cấp bản đồ, đảm bảo ứng dụng luôn chính xác về các dữ liệu địa lý so với thực tế. Nhóm đề tài đề xuất sử dụng OpenStreetMap làm dịch vụ cung cấp bản đồ cho API mà nhóm đề tài thực hiện.

Ngoài ra, ứng dụng cho phép người dùng tương tác trực tiếp với các marker trên bản đồ, hiển thị các thông tin liên quan đến các địa điểm đó. Thay vì phải đọc các dữ liệu thô theo vùng miền, theo quốc gia, vùng lãnh thổ như trước, cách tương tác này giúp người dùng dễ dàng cập nhật thông tin cũng như có cái nhìn tổng quan, khả năng so sánh giữa các địa điểm khác nhau mà không phải chăm chú vào các con số nhàm chán như trước.

2.2 Kiến thức nền tảng

2.2.1 JavaScript

Giới thiệu

Javascript là một ngôn ngữ nhẹ (lightweight), thông dịch, hướng đối tượng với first-class function, và được biết đến nhất như là một ngôn ngữ kịch bản (scripting language) cho các trang Web, nhưng Javascript cũng có thể được sử dụng trong các môi trường non-browser như node.js hay Apache CouchDB. Javascript là một prototype-based, ngôn ngữ kịch bản đa mô hình lập trình (multi-paradigm), động (dynamic), hỗ trợ nhiều kiểu như hướng đối tượng, lập trình mệnh lệnh (imperative programming), và lập trình hàm (functional programming).

Chuẩn cho Javascript là ECMAScript. Phiên bản hiện nay là ECMAScript 7 được công bố vào tháng 6/2016. Tuy nhiên, cho đến hiện nay, các trình duyệt chỉ hỗ trợ chuẩn ECMAScript 5, sắp tới đây sẽ hỗ trợ cho ECMAScript 6. Nhóm sử dụng chuẩn ECMAScript 6 như là tiêu chuẩn để coding thư viện. Các tính năng mạnh mẽ của ECMAScript 6 sẽ được liệt kê ở những phần sau.

Tính năng

JavaScript nâng cao sự linh động và khả năng tương tác cho website bằng cách sử dụng các hiệu ứng của nó như thực hiện các phép tính, kiểm tra form, viết các trò chơi, bổ sung các hiệu ứng đặc biệt, tùy biến các chọn lựa đồ họa, tạo ra các mật khẩu bảo mật và hơn thế nữa.

Cụ thể, JavaScript có thể được dùng để:

- *Tương tác với người dùng*: Chúng ta có thể viết mã để đáp lại các sự kiện. Các sự kiện này sẽ có thể phát sinh bởi người dùng: nhấp chuột hay được phát sinh từ hệ thống, định lại kích thước của trang ...
- *Thay đổi nội dung động*: Mã JavaScript có thể dùng để thay đổi nội dung và vị trí các phần tử một cách động trên một trang nhằm đáp lại sự tương tác với người dùng.

- *Kiểm tra tính hợp lệ dữ liệu*: Chúng ta có thể viết mã nhằm kiểm tra tính hợp lệ của dữ liệu do người dùng nhập vào trước khi nó được gửi lên Web server để xử lý.

ECMAScript 6

ECMAScript 6 hay gọi tắt là ES6 hay ES2015, đây là phiên bản nâng cấp, tập hợp các kỹ năng nâng cao của Javascript. ES6 ra đời năm 2015 nên cái tên ES2015 được lấy làm tên chính thức với nhiều tính năng mới học hỏi các ngôn ngữ cấp cao khác, hy vọng dần theo thời gian Javascript trở thành một ngôn ngữ lập trình hướng đối tượng.

Một số tính năng mạnh mẽ của ES6:

- *Classes*: Với việc hỗ trợ class, các cú pháp trông đẹp đẽ, đơn giản, dễ nhìn, rõ ràng hơn rất nhiều trong việc tạo ra đối tượng và giải quyết vấn đề thừa kế.
- *Modules*: Hỗ trợ import/export các giá trị (biến, hàm, ...) từ/đến các module.
- *Block Scoped*: Khai báo biến với từ khóa let để cho biết biến này chỉ tồn tại trong phạm vi block của nó.
- *Destructuring Assignment*: Dễ dàng khởi tạo các biến trong suốt các câu lệnh gán từ một mảng.
- *Arrow function*
- *Default Parameter Values*: Cho phép gán giá trị mặc định cho tham số.
- *Rest Parameter*: Kết hợp các argument còn lại thành một tham số của hàm.

Ứng dụng

Dựa trên những đặc điểm và khả năng mạnh mẽ trong hỗ trợ lập trình API của JavaScript đã kể trên, nhóm đề xuất sử dụng JavaScript làm ngôn ngữ chính để hiện thực API.

Ngoài ra, nhóm đề xuất sử dụng chuẩn ECMAScript 6 làm tiêu chuẩn cho các ràng buộc, quy tắc (coding conventions) khi hiện thực API.

2.2.2 D3.js

Giới thiệu

D3.js được viết bởi Mike Bostock, là 1 thư viện JavaScript dùng để thao tác với document dựa trên dữ liệu (data). D3 giúp mang dữ liệu đến cuộc sống bằng cách sử dụng HTML, SVG, và CSS. Sức mạnh của D3 trên chuẩn web cung cấp đầy đủ các tính năng của những trình duyệt hiện đại mà không buộc mình vào một khuôn khổ nhất định nào, kết hợp với các thành phần trực quan mạnh mẽ và một phương pháp hướng dữ liệu để thao tác DOM (Document Object Model).

D3 không phải là framework giải quyết mỗi vấn đề riêng biệt nào, thay vào đó D3 giải quyết điểm cốt lõi của vấn đề: sự thao tác hiệu quả các document dựa trên dữ liệu. Do đó, D3 có được sự linh hoạt cao, có thể sử dụng toàn bộ các tính năng của các tiêu chuẩn web như HTML, SVG, và CSS.

D3 cho phép nhúng bất kì kiểu dữ liệu nào vào 1 phần tử DOM, và sau đó có thể áp dụng CSS3, HTML, và/hoặc SVG lên dữ liệu này. Cuối cùng, D3.js có thể giúp tạo ra dữ liệu tương tác thông qua việc sử dụng những sự biến đổi (transformation), chuyển đổi (transition) hướng dữ liệu của D3.js.

Tính năng

Một số tính năng mạnh mẽ của D3:

- *Selections*: Tập hợp các node tùy ý gọi là selections, và D3 sử dụng phương pháp declarative để biến đổi selections này. Đặc biệt D3 cũng áp dụng các selectors được định nghĩa bởi W3C Selectors API, nhờ đó các phần tử sẽ được chọn lựa ra một cách dễ dàng.
- *Dynamic Properties*: Dễ dàng thay đổi các thuộc tính, style một các động thông qua các function trong D3.
- *Enter và Exit*: Sử dụng enter và exit các selections, người dùng có thể tạo ra nhiều

node mới cho các dữ liệu mới và xóa các node mà không cần thiết. Với việc sử dụng đúng những tác vụ cần thiết trên các node sẽ làm tăng hiệu suất và mang đến những hiệu ứng chuyển đổi tuyệt vời hơn.

- *Transformation*: D3 sử dụng các tiêu chuẩn web: HTML, SVG, và CSS. Do đó nếu trình duyệt giới thiệu một phiên bản mới, người dùng có thể sử dụng D3 ngay lập tức, không yêu cầu cập nhật. Ngoài ra giúp người dùng dễ dàng debug với sự trợ giúp của trình duyệt.
- *Transitions*: Interpolator của D3 hỗ trợ cả 2 kiểu primitive, ví dụ như là số và số nhúng trong string (font-size, dữ liệu về path, ...) và các giá trị phức tạp. Người dùng thậm chí có thể mở rộng interpolator để hỗ trợ cho các thuộc tính phức tạp và cấu trúc dữ liệu.

Ứng dụng

D3 được viết bằng JavaScript, và sử dụng kiểu lập trình hàm (functional programming) do đó có thể sử dụng lại code và thêm những hàm chức năng mà người dùng muốn. Điều đó có nghĩa là nó sẽ mạnh mẽ như thế nào là dựa vào người dùng muốn làm ra nó. Cách người dùng chọn để định dạng, thao tác, tương tác với dữ liệu là tùy thuộc ở họ.

D3 rất thích hợp dùng trong đề tài, nó giúp xây dựng framework trực quan dữ liệu mà người dùng muốn, dễ dàng sử dụng, do đó giúp tạo ra API một cách dễ dàng hơn, API dễ hiểu, dễ sử dụng.

2.2.3 OpenLayers

Giới thiệu

OpenLayers được tạo ra bởi MetaCarta sau hội nghị O'Reilly Where 2.0 (29-30/06/2005) và chính thức phát hành thành mã nguồn mở trước hội nghị O'Reilly Where 2.0 (13-14/06/2006) bởi MetaCarta Labs. Hai dự án khác về công cụ xây dựng bản đồ phát hành bởi MetaCarta là FeatureServer và TileCache. Từ tháng 11/2007, OpenLayers trở thành dự án mã nguồn mở chính thức của Hội đồng Không gian địa lý.

Sơ qua về OpenLayers, đây là một thư viện JavaScript mã nguồn mở (được chứng nhận bởi BSD License) nhằm phục vụ mục đích biểu diễn thông tin trên bản đồ đối với các trình duyệt Web, không phụ thuộc phía server. Nó cung cấp API chuyên về xây dựng các ứng dụng địa lý trên nền Web tương tự như Google Maps và Bing Maps, với một điểm khác biệt duy nhất, OpenLayers hoàn toàn miễn phí, được xây dựng và phát triển nên bởi cộng đồng phát triển phần mềm mã nguồn mở. Thư viện xây dựng trên Prototype JavaScript Framework.

Hiện nay, OpenLayers đã ra tới phiên bản OpenLayers 3. OpenLayers 3 được viết lại nhằm dễ hiểu hơn, hướng tới các tính năng của HTML5 và CSS3. Thư viện tiếp tục hỗ trợ trình chiếu, các giao thức tiêu chuẩn và các chức năng chỉnh sửa. Phiên bản mới này tập trung vào cải thiện hiệu năng, xây dựng đơn giản hơn, các thành phần trở nên trực quan hơn, API được cải tiến và còn nhiều hơn nữa.

Tính năng

Openlayers có 4 tính năng chính:

- *Thể hiện Lớp dưới dạng ô (Tiled Layers)*: Các lớp này được kéo về từ OpenStreetMap (OSM), Bing, MapBox, Stamen, MapQuest và bất cứ nguồn nào mà bạn có thể tìm thấy. Hỗ trợ đa dạng các Layer mà bạn có thể sử dụng trong ứng dụng của mình.
- *Thể hiện Lớp dưới dạng Vector (Vector Layers)*: Sinh ra dữ liệu dạng vector từ GeoJSON, TopoJSON, KML, GML và một số định dạng khác.
- *Tính đối ứng nhanh gọn (Fast & Mobile Ready)*: Hỗ trợ tùy chỉnh các thành phần một cách nhanh gọn và thuận tiện nhất. Cho phép xây dựng một bản tùy chỉnh gọn nhẹ gồm các thành phần mà lập trình viên cần.
- *Mài dũa các góc cạnh và dễ dàng tùy chỉnh (Cutting Edge & Easy to Customize)*: Bản đồ sinh ra bởi công nghệ WebGL, Canvas 2D, và từ những tính năng ưu việt nhất của HTML5. Dễ dàng tùy chỉnh style bằng CSS.

Ứng dụng

Đề tài hướng tới xây dựng thư viện JavaScript cho phép tương tác giữa bản đồ và các đồ thị dựa trên dữ liệu đầu vào, cho nên OpenLayers, cụ thể OpenLayers 3 sẽ là công cụ đặc lực cho việc xây dựng thành phần bản đồ trong đề tài.

3 Giới thiệu thư viện C9js

3.1 Mô tả bài toán

Thông qua các công trình đã tìm hiểu, nhóm đề tài đánh giá và trích xuất ra các điểm hạn chế, các vấn đề mà các công trình đó chưa giải quyết được hoặc đã giải quyết nhưng đã lỗi thời cần nâng cấp để phù hợp với bối cảnh hiện tại của các công nghệ Web. Phần đánh giá, so sánh với các công trình liên quan sẽ được thể hiện chi tiết trong mục **Đánh giá**.

Bài toán mà nhóm đề tài đặt ra:

Hỗ trợ trực quan dữ liệu trên Web thông qua thư viện C9js.

Các vấn đề "thất nút" cần giải quyết:

Cũng là các mục tiêu và hướng phát triển **C9js** đã, đang và mong muốn đạt được:

- *Hoàn toàn là mã nguồn mở:* **C9js** là thư viện JavaScript mã nguồn mở, được triển khai thông qua GitHub ở địa chỉ <https://github.com/csethanhcong/C9js>.
- *Kiến trúc xây dựng theo tiêu chuẩn ES2015 (hay ES6):* Nghiên cứu cẩn trọng qua các mã nguồn mở có khả năng trực quan dữ liệu, việc chọn kiến trúc để xây dựng từ ban đầu là rất quan trọng, đặc biệt trong khả năng mở rộng nhằm đáp ứng nhu cầu tùy chỉnh hay đóng góp từ cộng đồng mã nguồn mở.
- *Bao gồm các tính năng mới, mạnh mẽ:* Các tính năng riêng biệt của C9js được thể hiện rõ trên trang chủ của thư viện, tại địa chỉ <http://c9js.me/>. Phần trình bày chi tiết sẽ được nêu ở phần sau

3.2 Hướng dẫn sử dụng

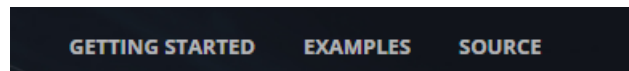
Với mong muốn đem lại trải nghiệm nhanh nhất, đơn giản nhất và ít tốn công sức nhất cho người dùng (ở đây là các lập trình viên), nhóm đề tài đã triển khai Website <http://c9js.me/>



Hình 3.1: Các tính năng chính của C9js trên trang chủ c9js.me

gồm các hướng dẫn từ cơ bản đến chuyên sâu, và gồm cả các ví dụ cụ thể nhất.

Qua Website c9js.me, có thể thấy mục **Getting Started** ở các nút ở trang chủ (landing page) hay thông qua menu ở đầu trang như hình 3.2



Hình 3.2: Menu hướng dẫn trên trang chủ c9js.me

Khi đó, màn hình hướng dẫn xuất hiện như hình 3.3

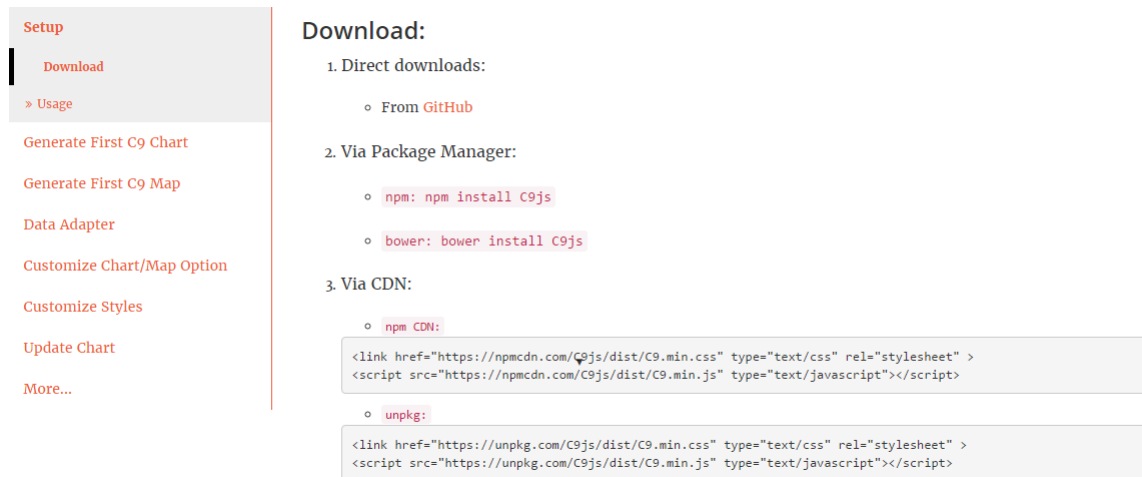
Như có thể thấy, menu giao diện bên trái thể hiện các mục hướng dẫn người dùng, bao gồm các bước cơ bản nhất để sử dụng và tùy biến các tính năng trong C9js theo ý muốn

3.2.1 Cài đặt (Setup)

Tải xuống (Download):

Hiện tại, nhằm đáp ứng tối đa lượng người dùng, C9js được triển khai qua nhiều kênh khác nhau

- *Tải trực tiếp:* Có thể tải trực tiếp ở địa chỉ <https://github.com/csethanhcong/>



Hình 3.3: Trang hướng dẫn trên Website c9js.me

C9js/releases/latest.

- *Các trình quản lý tập tin (Package Manager):* Hiện tại phổ biến nhất có 2 trình quản lý là **npm** và **bower**. Và C9js được phân phối qua cả 2 kênh này. Người dùng cài đặt thông qua một trong các lệnh:

- **npm:** npm install C9js
- **bower:** bower install C9js

- *Mạng phân phối nội dung (CDN):* Ngoài ra, người dùng còn có thể truy cập các mã nguồn cần thiết từ C9js thông qua hai dịch vụ CDN phổ biến là **npmcdn** và **unpkg**

- **npmcdn:**

```
1 <link href="https://npmcdn.com/C9js/dist/C9.min.css"
  type="text/css" rel="stylesheet" >
2 <script src="https://npmcdn.com/C9js/dist/C9.min.js"
  type="text/javascript"></script>
```

Mã nguồn 1: Tải mã nguồn thông qua **npmcdn**

- **unpkg:**

```
1 <link href="https://unpkg.com/C9js/dist/C9.min.css" type="text/css" rel="stylesheet" >
2 <script src="https://unpkg.com/C9js/dist/C9.min.js" type="text/javascript"></script>
```

Mã nguồn 2: Tải mã nguồn thông qua **unpkg**

Cách dùng (Usage):

Thư viện C9js hỗ trợ tạo biểu đồ dựa trên *d3.js* và bản đồ dựa trên *OpenLayers 3* nên người dùng cần tải các thư viện này trước.

Tùy theo mục đích sử dụng, nếu chỉ dùng chức năng tạo biểu đồ thì chỉ cần tải thư viện *d3.js*, tương tự đối với chức năng tạo bản đồ.

```
1 <link href="/path/to/C9.css" rel="stylesheet" type="text/css">
2 <!-- Load d3.js and C9.js -->
3 <script src="/path/to/d3.v3.min.js" charset="utf-8"></script>
4 <script src="/path/to/C9.min.js"></script>
```

Mã nguồn 3: Tải mã nguồn *d3.js* trước tiên

```
1 <link rel="stylesheet" href="https://openlayers.org/en/v3.19.1/css/ol.css" type="text/css">
2 <script src="https://openlayers.org/en/v3.19.1/build/ol.js" type="text/javascript"></script>
3 <script src="/path/to/C9.min.js"></script>
```

Mã nguồn 4: Tải mã nguồn *OpenLayers 3* nếu muốn sử dụng chức năng Bản đồ

3.3 Các ví dụ minh họa

3.4 Ứng dụng thực tế

4 Quá trình thực hiện

4.1 Kiến trúc xây dựng

4.2 Hiện thực thư viện

5 Kết luận

5.1 Đánh giá

5.2 Hướng phát triển

5.3 Kết luận

Tài liệu tham khảo

- [1] Albert Einstein. Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]. *Annalen der Physik*, 322(10):891–921, 1905.