Xây dựng phương pháp tìm kiếm tốt hơn nhờ đánh giá chất lượng tìm kiếm.

## Giới thiệu

Các công cụ tìm kiếm thường được phát triển bằng cách sử dụng các bộ test case có tính thực tế, việc này cho phép người phát triển đánh giá một cách tương đối chính xác hiệu quả của các cách tiếp cận khác nhau. Hội nghị Text Retrieval Conference (TREC) của Viện tiêu chuẩn và công nghệ Mỹ (NIST) đã thành lập lên một nền tảng cơ bản để đánh giá chất lượng của các kết quả được tiếm kiếm.

Một điều hiển nhiên, chúng ta thường tìm kiếm bằng ngôn ngữ mẹ đẻ của mình, nhưng các công cụ tìm kiếm trên Web rất phức tạp, và nội dung cũng rất rộng lớn. Dữ liệu cũng ngày một phong phú hơn, cùng với sự phát triển đó, chúng ta tìm kiếm trên mạng với ngày một nhiều lý do (có thể là giải trí, công việc, tìm một bức ảnh, một bộ phim hay một đoạn ghi âm,…) Trong nhiều trường hợp, công nghệ vẫn còn hạn chế, chưa thể giải quyết ngay được. Và các nhà khoa học đang phải đối mặt với các câu hỏi lớn: Làm sao để cải tiến công nghệ tìm kiếm? Và làm sao để biết được phải làm gì và vì sao?

Việc đánh giá cẩn thận và chi tiết hiệu suất của các máy tìm kiếm dựa trên một nền tảng, một tập hợp các test case mang tính thực tế với sự tham gia của một cộng đồng nghiên cứu rộng lớn và đa dạng đã được chứng minh là rất quan trọng. Thông qua dự án TREC, Viện tiêu chuẩn và công nghệ Mỹ (NIST) thành lập một công cụ để đánh giá và thúc đẩy sự phát triển của công nghệ tìm kiếm trong hơn 4 thế kỉ qua.

## Sự ra đời của TREC

Thuật toán tìm kiếm thường được phát triển bằng cách so sánh các phương pháp khác nhau dựa trên một tiêu chuẩn khắt khe được gọi là “Test collections”- các bộ test

Bộ test đầu tiên là kết quả của một loạt các thí nghiệm liên quan đến ngôn ngữ chỉ mục tại trường cao đẳng hàng không Cranfield trong những năm 1960

Bộ test Cranfield bao gồm 1 tập đoạn trích, bản tóm tắt của các bài báo về hàng không, và một tập các truy vấn đối với những bản tóm tắt này và kết quả chuẩn cho phản hồi đúng mỗi truy vấn. Mặc dù rất nhỏ so với tiêu chuẩn hiện nay. Bộ test Cranfield đã lập nên nền tảng mới bằng cách tạo ra một công cụ đo lường, chia sẻ đầu tiên cho các hệ thống thông tin.

Một số nhóm nghiên cứu khác cũng đã bắt đầu thực hiện theo các phương pháp thử nghiệm được giới thiệu bởi bộ test Cranfield, tạo ra một tập các bộ test khác được sử dụng vào những năm 1970, 1980. Nhưng vào năm 1990 đã xuất hiện sự những mâu thuẫn ngày càng lớn trong phương pháp này. Trong tình trạng bất mãn này, NIST đã được yêu cầu để xây dựng một bộ test lớn để sử dụng trong việc đánh giá công nghệ tìm.

## Sự đánh giá của cộng đồng

Cùng với sự bắt đầu của dự án TREC rất nhiều ngờ vực phát sinh liên quan đến hệ thống thống kê. Liệu một hệ thống được phát triển trong phòng thí nghiệm có thể hoạt động hiệu quả trên một khối lượng dữ liệu lớn trong thực tế?

Tuy vậy, TREC đã cho thấy, không chỉ những công cụ tìm kiếm vào đầu những năm 1990 mới có thể làm việc trên quy mô dữ liệu lớn, mà chính TREC cũng đã được cải thiện đáng kể. Tính hiệu quả của TREC đã được kiểm nghiệm không chỉ trong phòng thí nghiệm mà còn được sử dụng thường xuyên vào những bộ dữ liệu lớn hơn nhiều những bộ dữ liệu được cho là lớn vào những năm 1992.

## Sự đa dạng hóa

Mặc dù mục đích ban đầu của TREC chỉ đơn giản là xây dựng một hoặc 2 bộ test lớn và trả lời những câu hỏi liên quan đến phương pháp tổng hợp, nó nhanh chóng cho thấy rằng ad hoc task có thể chiếm lĩnh 1 phạm vi lớn. Ngày nay, mỗi TREC có 7 hoặc 8 tracks luôn phiên thay đổi nhau để giữ cho TREC không bị lỗi thời để hỗ trợ cho cộng đồng mới một cách hiệu quả.

## Lọc

Ad hoc task và routing task là 2 nhiệm vụ trong những năm đầu tiên của dự án TREC.

Trong khi việc đánh giá routing task cho phép quá trình sử lý tất cả các tài liệu trong bộ sưu tập một cách đồng loạt thì sự đánh giá filtering task đòi hỏi hệ thống phải xử lý các tài liệu được thu thập đồng loạt và thích ứng với mô hình người dùng trực tuyến.

Filtering track cho thấy được sự khó khăn khi thực hiện filtering track. Trong filtering track, cũng như trong cuộc sống, hầu như chỉ có một rất phần nhỏ các tài liệu liên quan được tìm thấy trong hàng ngàn tài liệu.

## Hướng phát triển trong tương lai

Hoạt động của TREC cũng có nhiều lợi ích không phải về mặt kinh tế. TREC giúp nâng cao chất lượng đào tạo ở bậc đại học, giúp các sinh viên trong việc học tập, nghiên cứu,… Nghiên cứu đã cho thấy TREC ảnh hưởng 1/3 của sự cải tiến hơn 200% trong các sản phẩm tìm kiếm web đã được quan sát thấy từ năm 1999- 2009

Mặc dù thành công này, còn nhiều việc phải làm. Máy tính vẫn không thể thực sự hiểu nội dung dữ liệu mà con người sử dụng trong khi dữ liệu đang phát triển ngày một lớn hơn.

Nguồn tham khảo http://www.infoq.com/articles/building-better-search-engines-by-measuring-search-quality