1. **CÁC CHỈ SỐ ĐÁNH GIÁ XẾP HẠNG CỦA WEBOMETRIC**

Webometrics xếp hạng website trường đại học không chỉ dựa vào quy mô, số lần truy cập, hình thức thiết kế website mà còn căn cứ vào việc công bố điện tử các kết quả nghiên cứu, các tài liệu và thông tin khoa học thông qua 4 chỉ số: Chỉ số nhận diện, Chỉ số qui mô, Chỉ số “file giàu”, Chỉ số thư tịch khoa học. Cả 4 chỉ số này đều được đo lường dựa trên trang web của trường đại học.

**1. Chỉ số nhận diện (V – Visibility, 50%)**

Chỉ số nhận diện (V) được tính toán bằng tích của căn bậc hai số các đường liên kết từ bên ngoài (refering backlinks) liên kết đến các kết nối bên trong trên một tên miền giáo dục và số lượng miền (refering domains) có nguồn gốc từ backlinks đó [1].



Webometrics lấy dữ liệu cho việc xác định xác định chỉ số V dựa vào số liệu của hai nhà cung cấp quan trọng nhất đó là Majestic SEO, Ahrefs. Các liên kết đó sẽ giúp công nhận uy tín của Trường, hiệu quả đào tạo, giá trị của thông tin và tính hữu ích của các dịch vụ của Trường khi được giới thiệu trong các trang web theo các tiêu chí của hàng triệu biên tập viên website trên toàn thế giới. Vì vậy, không những sự phổ biến của liên kết quan trọng, mà sự đa dạng của liên kết thậm chí còn quan trọng hơn.

Như vậy, chỉ số V đại diện cho khía cạnh ảnh hưởng giáo dục và danh tiếng của trường Đại học. Ưu điểm của phương pháp này là đưa ra ảnh hưởng của các trường ĐH với nhau, các trường ĐH trong nước và ngoài nước, các tổ chức giáo dục, các tổ chức xã hội trên cơ sở đồ thị web đã được xây dựng bởi công cụ tìm kiếm Majestic SEO, Ahrefs.

**2. Nhóm các chỉ số hoạt động khoa học (Activity, 50%)**

**2.1. Chỉ số Qui mô (S – Size, 1/3)**

Chỉ số qui mô website (S) thể hiện mức độ hiện diện của website, được tính thông qua tổng số trang được lưu trữ trong một tên miền (bao gồm cả tên miền con) đại học khi được chỉ mục bởi công cụ tìm kiếm Google. Google chấp nhận cả các trang web động, trang web tĩnh và các rich file khác. Vì vậy không thể có một sự hiện diện mạnh (chỉ số V lớn) mà không có sự đóng góp của tất cả mọi người trong đơn vị bởi các đối thủ hàng đầu luôn sãn sàng có thế xuất bản hàng triệu trang web.

**2.2. Chỉ số file giàu (R- Rich file, 1/3)**

Chỉ số phong phú dữ liệu (R) thể hiện tính mở của Trường trên website, được tính thông qua số lượng các loại file Adobe Acrobat (.pdf), Adobe file Script (.ps), Microsoft Word (.doc, .docx) và Microsoft Powerpoint (.ppt, .pptx) được đăng tải trên trang web của đơn vị.

Mục tiêu là để xem xét các ấn phẩm (đại diện cho kết quả nghiên cứu, tài liệu học tập..) mới công bố gần đây của Trường.

**2.3. Chỉ số thư tịch khoa học (Sc- Scholar, 1/3)**

Chỉ số thư tịch khoa học (Sc) được tính theo số lượng các thư tịch khoa học (academic records) gồm số lượng các bài báo, luận án, luận văn, công trình nghiên cứu khoa học, ấn phẩm khoa học, trích dẫn trên các tên miền học thuật (academic domain). Chỉ số này cùng với chỉ số R đại diện cho kết quả nghiên cứu của từng trường.

Chỉ số Sc được tính toán thông qua công cụ tìm kiếm Google Scholar. Với cú pháp tìm kiếm dạng: <tên miền>, ví dụ: [www.ntu.edu.vn](http://www.ntu.edu.vn) hoặc [ntu.edu.vn](http://www.ntu.edu.vn).

Để được chỉ mục trong Google scholar, chúng ta phải thỏa mãn các tiêu chí sau:

**2.3.1. Tiêu chí nhà xuất bản (Publisher type)**

a**.** *Đối với tác giả độc lập***:** Upload bài viết lên website (ví dụ: [www.example.edu/~professor/jpdr2009.pdf](http://www.example.edu/~professor/jpdr2009.pdf)), sau đó tạo một liên kết đến file mới upload trên một trang web công bố ấn phẩm độc lập (ví dụ: [www.example.edu/~professor/publications.html](http://www.example.edu/~professor/publications.html)). Toàn văn bài báo lưu ở dạng file PDF, các phần trong nội dung phải được bố trí và định dạng theo chuẩn một bài báo khoa học.

b**.** *Đối với Kho tài nguyên số Đại học***:** Lưu trữ, công bố các bài viết sử dụng các phần mềm mới nhất như Dsspace, Eprints, Digital Commmon (ví dụ: <http://repository.ntu.edu.vn>).

c**.** *Nhà xuất bản tạp chí:* Để lưu trữ số lượng nhỏ các bài báo, Google khuyến nghị sử dụng các dịch vụ lưu trữ như Atypon, Highwire Press or MetaPres. Nếu số lượng lớn, Google khuyến nghị sử dụng dich vụ lưu trữ như JSTOR hoặc SciELO. Hoặc sử dụng Open Journal System nếu có năng lực tự quản trị website và các ấn phẩm của mình (ví dụ <http://pkp.sfu.ca/ojs> )

**2.3.2. Tiêu chí nội dung website**

**-** *Về việc upload các bài báo học thuật***:**  Để thỏa mãn điều kiện cho Google scholar đánh chỉ mục tài liệu của mình, nội dung của website đa số phải là các ấn phẩm học thuật, bao gồm các bài báo, báo cáo, bản thảo, các tóm tắt bài báo và các tài liệu tương tự. Google Scholar không đánh chỉ mục các đánh giá sách, các bài xã luận, tin tức, báo tạp chí. Nếu file muốn được đánh chỉ mục có kích thước lớn hơn 5MB, Google yêu cầu phải upload lên Google Books. Các nội dung học thuật được upload lên Google Books mạc nhiên sẽ thỏa mãn các điều kiện chỉ mục hóa tài liệu của Google Scholar.

**-** *Về việc sử dụng Tóm tắt (abstracts)***:** Bản tóm tắt đầy đủ của bài viết (hoặc toàn bộ bài viết) phải dễ dàng nhìn thấy và dễ truy cập. Người xem không cần phải thực hiện bất kỳ thao tác nào (như đăng nhập, đóng cửa sổ quảng cáo, cuộn chuột) để truy cập và đọc toàn bộ tóm tắt.

**2.3.3. Tiêu chí về định dạng và cấu trúc**

*a. Định dạng file*

Đề bộ tìm kiếm thông tin của Google hoặt động một cách hiệu quả, file ấn phẩm phải ở định dạng HTML hoặc PDF. Các file PDF phải có văn bản bên trong sao cho trình đọc PDF có thể tìm kiếm một từ khóa nhất định.

*b. Cấu trúc trang web*

Website cần phải được cấu trúc hóa theo một khuôn mẫu nhất định, sao cho bộ tìm kiếm thông tin có thể dễ dàng xác định địa chỉ URL của tất cả các ấn phẩm học thuật của Trường. Khuôn mẫu đó phải đảm bảo việc đi tới mỗi bài báo phải không quá 10 liên kết HTML.

Nếu website có một số lượng nhỏ các bài báo, cần một trang HTML độc lập cung cấp một danh sách các liên kết đến các bài báo đó, trang này nên đơn giản và dễ truy cập. Chú ý rằng, file bài báo phải ở định dạng PDF và nội dung ở dạng toàn văn.

Nếu website có số lượng hàng ngàn bài báo, cách tốt nhất là cung cấp cho bộ tìm kiếm thông tin của Google một danh sách đầy đủ các bài báo được sắp xếp theo ngày công bố. Các giá trị sắp xếp khác như tên tác giả hoặc từ khóa làm mất nhiều thời gian tìm kiếm của Google, và vì vậy dẫn tời thứ hạng thấp trong kết quả tìm kiếm của Google

Nếu website có hàng trăm ngàn bài báo, cần thêm một trang liệt kê tất cả các bài báo từ 2 tuần qua, đông thời sắp xếp theo thứ tự ngày công bố. Làm được điều này, các paper của bạn sẽ được tái thu thập thông tin một cách thường xuyên hơn, vì vậy đảm bảo số lượng lớn các ấn phẩm đều được chỉ mục trong Google Schoolar

*c. Khả năng tiếp cận người dùng và Bộ tìm kiếm thông tin*

Bộ tìm kiếm thông tin sẽ duyệt các trang web theo chu kỳ để tạo chỉ mục mới hoặc cập nhật nội dung, và vì vậy các bài báo của bạn phải liên tục online. Trong trường hợp trang web đang tạm thời bị lỗi, phải thoogn báo cho bộ tìm kiếm bằng việc sử dụng mã HTTP 5xx.

Trường hợp ấn phẩm được chuyển sang một địa chỉ mới, thì cần phải điều hướng lại chúng, sử dụng mã HTTP 301, việc điều hướng phải trỏ tới URL mới của ấn phẩm.

*d. Điều tiết thu tập thông tin*

Tệp robot.txt trên website có tính năng điều tiết việc tìm kiếm thông tin. Vì vậy, cần phải đảm bảo rằng tệp này đã không được cấu hình để ngăn cấm bộ tìm kiếm thông tin truy cập các ấn phẩm của bạn, đồng thời cần cấu hình cấm bộ tìm kiếm truy cập các nội dung rộng và được tạo ra một cách tự động, bởi vi điều này là mất nhiều thời gian cho bộ tìm kiếm duyệt trang của bạn.

*e. Sách, tài liệu tham khảo và Phân tích cú pháp chỉ mục*

Google scholar nhận biết thông tin từ sách và tài liệu tham khảo sử dụng phần mềm tự động gọi là “Parsers – Bộ phân tích cú pháp”. Các bài báo có thể được chỉ mục một cách nghèo nàn nếu thông tin sách tham khảo không đúng hoặc không phù hợp với tài liệu tham khảo từ các bài báo khác. Vì vậy, cần thiết phải công bố bài báo với các siêu dữ liệu (meta-data) về thư mục và tài liệu tham khảo để bộ phân tích cú pháp có thể làm việc.

**2.3.4. Chỉ mục hóa tài liệu với thẻ meta**

Thông tin sách thao khảo của một bài báo phải được cung cấp trong các thẻ HTML meta. Các thẻ meta (ví dụ: <meta name="citation\_author" content=“Smith, John J.”>) là cụ thể và thường chỉ liên quan đến một trang web nhất định của một bài báo.

Google Scholar yêu cầu các trang web của bất kỳ bài báo nào cũng phải có các thẻ meta về tiêu đề (citation\_title), meta về các tác giả (citation\_author) và thẻ meta về ngày công bố (citation\_publication\_date).

Trang web của các bài báo hội nghị và hội thảo yêu cầu thẻ meta, vì vậy chúng có thể được dễ dạng được nhận diện khi nó được trích dẫn bởi bài viết khác. Các thẻ gồm article’s volume và issue number (citation\_volume và citation\_issue), cũng như trang đầu và cuối của bài báo được tham chiếu tới (citation\_firstpage và citation\_lastpage).

Đối với các báo cáo kỹ thuật, các đề tài, luận văn, luận án. Google yêu cầu các thẻ meta như: Cơ quan đơn vị của tác giả (citation\_dissertation\_institution, hoặc citation\_technical\_report\_institution), cũng như là số của báo cáo kỹ thuật đang được tham chiếu tới (citation\_technical\_report\_number). Các thẻ meta này giúp các báo cáo có thể được xác định khi được trích dẫn bởi các paper khác.

Vớt trang web chỉ trình bày mỗi phần abstract của một bài báo, Google yêu cầu chúng ta phải chỉ rõ vị trí của tài liệu full-text ở định dạng PDF sử dụng thẻ meta URL (citation\_pdf\_url), ở đó giá trị của thuộc tính content trong thẻ phải là địa chỉ URL tuyệt đối của tài liệu PDF (e.g. <meta name="citation\_pdf\_url" content=“URL”>). Chú ý rằng, URL của phần abstract và bài viết tương ứng của nó phải tồn tại dưới cùng thư mục con.

**2.3.5. Chỉ mục hóa tài liệu không dùng thẻ meta**

Nếu một bài viết không có các thẻ meta, nó phải cung cấp các thông tin về sách và tài liệu tham khảo ở định dạng văn bản đặc biệt.

Tiêu đề bài báo phải có kích thước font chữ tối thiều 24pt (đối với file PDFs), được đặt trong thẻ H1, hoặc H2 (đối với dạng HTML), hoặc được nhận diện với CSS class “citation\_title” (đối với file HTML sử dụng định dạng bằng CSS).

Sử dụng môt font cho toàn bộ tiêu đề bài báo, và tiêu đề đó phải có kích thước font lớn nhất so với bất kỳ văn bản nào của toàn bộ bài báo.

Tên tác giả, các đồng tác giả của một bài báo phải được in ngay trước hoặc sau tiêu đề, font chữ từ 16-23pt (PDFs) hoặc đặt trong thẻ H3 (HTMLs), hoặc trong CSS class “author” (đối với file HTML sử dụng định dạng bằng css). Sử dụng một font cho các tác giả. Font size của kho lưu trữ, tạp chí và các phân nhóm phải nhỏ hơn font size của tác giả.

Các tiêu đề, tác giả, kho lưu trữ và tạp chí toàn văn phải được viết ở dạng title-case (ví dụ. “Indexing Scholarly Publications”), trong khi subheadings phải sử dụng sentence-case (ví dụ “Indexing scholarly publications”).

Bạn bắt buộc phải chèn trích dẫn tài liệu tham khảo trên một dòng riêng, trong header và footer của trang đầu tiên (đối với file PDFs), hoặc ngay sau tiêu đề và tác giả (đối với HTMLs). Trích dẫn tài liệu tham khảo phải có một phiên bản của ấn phẩm công bố ở dạng bài báo toàn văn.

Các trích dẫn của các bài báo được công bố, phải được định dạng một cách rõ ràng (ví dụ. “J. Smith. Index., vol. 24, no. 8, pp. 873-876, September 2007”), tuy nhiên, nếu bài báo của bạn chưa được công bố, bạn phải thêm vào ngày tháng sẽ công bố trên một dòng riêng (ví dụ. September 30, 2007).