

# Xếp hạng đại học và chất lượng giáo dục

Nguyễn Tấn Đại, Vũ Thị Phương Anh\*

## 1. Mở đầu

Trong những năm gần đây, xu hướng xếp hạng các trường đại học trên thế giới đã thu hút được nhiều sự chú ý của giới khoa bảng cũng như các nhà chức trách ở khắp các nước, đặc biệt là ở châu Á. Các cấp lãnh đạo Việt Nam cũng có chủ trương là phải nhanh chóng đưa một số trường đại học của Việt Nam đạt tầm khu vực và quốc tế.

Để đạt được mục tiêu này, trước hết cần phải có những hiểu biết sâu rộng về các lợi ích và rủi ro của việc xếp hạng cũng như các phương pháp đã và đang được áp dụng trên thế giới trong việc xây dựng lộ trình nâng cao vị trí xếp hạng của một trường đại học. Trước đây, một số hội thảo khoa học trong nước đã thu hút được sự tham gia của nhiều tác giả có kinh nghiệm làm việc trong lĩnh vực khoa học giáo dục hoặc đánh giá chất lượng giáo dục với nhiều bài tham luận khá sâu về một số bảng xếp hạng trên thế giới. Trên các tờ báo trong nước cũng đã có nhiều bài đăng với nội dung tương tự. Tuy nhiên, vẫn chưa có một nghiên cứu nào trong nước đánh giá một cách hoàn chỉnh các bảng xếp hạng phổ biến, gắn kết với bản chất của vấn đề đánh giá chất lượng giáo dục đại học (bao gồm cả giảng dạy, đào tạo và nghiên cứu).

Trong bài này, chúng tôi mô tả một số bảng xếp hạng theo hai xu hướng phổ biến là xếp hạng tổng thể toàn bộ trường đại học và xếp hạng theo đặc thù chuyên ngành, đồng thời cập nhật những thông tin mới nhất về sự thay đổi, điều chỉnh của các bảng xếp hạng này. Qua các ý kiến tổng hợp từ những cuộc thảo luận đa chiều trong giới học thuật trên thế giới, chúng tôi hệ thống hoá các vấn đề chính cần lưu tâm trong phương pháp tiên hành xếp hạng đại học cũng như mối liên hệ giữa xếp hạng và chất lượng giáo dục đại học. Trên cơ sở đó, các khuyến cáo của giới học thuật quốc tế trong việc sử dụng các bảng xếp hạng sẽ được gắn vào thực tiễn tại Việt Nam để đánh giá triển vọng tham gia các bảng xếp hạng quốc tế của các trường đại học trong nước, nhưng với mục tiêu cần được xác định rõ ràng và xuyên suốt: xây dựng và đảm bảo chất lượng giáo dục đại học.

## 2. Các xu hướng xếp hạng phổ biến

### 2.1. Xếp hạng tổng thể toàn bộ trường đại học

Các bảng xếp hạng dạng này thường xem đối tượng xếp hạng là toàn bộ trường đại học, bất kể những đặc thù về chuyên ngành, lĩnh vực hoạt động hay cơ cấu tổ chức. Trong những nỗ lực xếp hạng cấp trường ở phạm vi quốc tế đầu tiên, có thể kể đến bảng xếp hạng của *Asiaweek* từ năm 1997. Bảng xếp hạng này dựa trên các nhóm tiêu chí về uy tín học thuật do chính các trường bầu chọn lẫn nhau, tính cạnh tranh đối với sinh viên, nguồn lực giảng viên, thành tựu nghiên cứu và nguồn lực tài chính của nhà trường. Hầu hết các nguồn dữ liệu thu thập được chủ yếu là do các trường đại học cung cấp khi họ chấp nhận tham gia cuộc khảo sát của *Asiaweek*, chỉ trừ số liệu trích dẫn khoa học là lấy từ một nguồn độc lập. Sau bốn năm thực hiện, bảng xếp hạng này đã tạo ra nhiều luồng dư luận khác nhau, kể cả ủng hộ lẫn phản đối. Cuối cùng, sau ấn bản thứ tư năm 2000, *Asiaweek* đã quyết định ngưng xếp hạng vì có quá nhiều trường đại học trong khu vực rút lui, không hợp tác nữa (Salmi, Saroyan, 2007).

Tuy nhiên, phải đến năm 2003 thì mới thực sự bắt đầu “bùng nổ” cuộc đua giữa các bảng xếp hạng đại học ở quy mô toàn cầu, kể từ khi “Bảng xếp hạng Học thuật các trường Đại học trên Thế giới” (*Academic Ranking of World Universities – ARWU*) được Trung tâm Đại học Đẳng cấp Quốc tế và Viện Giáo dục Đại học (thuộc Đại học Giao thông Thượng Hải, Trung Quốc) công bố. Mục đích ban đầu của bảng xếp hạng này là nhằm xác định vị trí của các trường đại học hàng đầu Trung Quốc trên

---

\* Trung tâm Khảo thí và Đánh giá Chất lượng Đào tạo, Đại học Quốc gia TP. HCM.

toàn cầu, nhưng sau đó đã vượt lên trở thành một trong các bảng xếp hạng có tầm ảnh hưởng lớn nhất và được sử dụng nhiều nhất trên thế giới trong những năm sau đó. Bảng xếp hạng này dùng một số tiêu chí đánh giá kết quả đầu ra, đặc biệt là trong nghiên cứu khoa học. Cụ thể, ARWU có bốn tiêu chí và sáu chỉ số với trọng số khác nhau, liên quan đến số giải thưởng nghiên cứu khoa học lớn (*Nobel và Fields*), số bài báo đăng trên hai tạp chí *Nature* và *Science*, số bài báo được lập chỉ mục và số nhà nghiên cứu được trích dẫn nhiều dựa trên cơ sở dữ liệu *Science Citation Index-expanded* (SCIE) và *Social Science Citation Index* (SSCI) của *Thomson Reuters* (tiền thân là *Institute for Scientific Information – ISI*). Với mỗi chỉ số, trường đạt điểm cao nhất được tính điểm 100, và các trường khác sẽ được tính theo tỉ lệ phần trăm của điểm cao nhất; điểm của mỗi chỉ số được quy theo trọng số để tính thành điểm tổng cộng sau cùng của mỗi trường, và trường cao nhất cũng được tính điểm 100, còn các trường khác theo tỉ lệ so với trường cao nhất; điểm số cuối cùng được xếp theo thứ tự giảm dần, và vị trí của mỗi trường cũng chính là số trường khác đứng trên mình trong bảng xếp hạng (Liu, Cheng, Liu, 2005).

Một năm sau khi ARWU được công bố lần đầu, một bảng xếp hạng khác của *Times Higher Education Supplement* (THES) – đổi thành *Times Higher Education* (THE) từ 2008 – cũng ra đời, qua sự hợp tác với hãng *Quacquarelli Symonds* (QS). Bảng xếp hạng THES/THE-QS này được duy trì từ 2004 đến 2009 với cơ sở lí luận xếp hạng dựa vào “bốn cột trụ” tạo nên giá trị của một trường đại học hàng đầu thế giới: chất lượng nghiên cứu cao; chất lượng giảng dạy cao; tỉ lệ sinh viên tốt nghiệp có việc làm cao; và một “quan điểm quốc tế” hay mức độ quốc tế hoá (Baty, 2009). Từ cuối năm 2009, do bất đồng quan điểm, mối quan hệ hợp tác giữa THE và QS đã chấm dứt, để từ đó THE chuyển sang hợp tác với *Thomson Reuters* để xây dựng một bảng xếp hạng mới chặt chẽ và minh bạch hơn. Các hoạt động cốt lõi của một trường đại học được chia thành bốn nhóm tiêu chí, bao gồm: nghiên cứu khoa học; hoạt động kinh tế và sáng tạo; tính đa dạng quốc tế và nhóm chỉ số khoa bảng. Nếu như trong các bảng xếp hạng cũ, chỉ có sáu chỉ số với tỉ lệ rất cao dành cho kết quả khảo sát danh tiếng (*reputation survey*), thì bảng xếp hạng mới dự kiến sẽ ưu tiên nhiều hơn cho các chỉ số hoạt động nghiên cứu, với thông tin lấy từ các cơ sở dữ liệu xuất bản khoa học của *Thomson Reuters*, và đặc biệt là giảm mức độ phụ thuộc vào một chỉ số riêng rẽ để tăng giá trị của các nhóm chỉ số đa tiêu chí, liên quan mật thiết đến chức năng và sứ mạng của các trường đại học. Qua đó, họ hi vọng có thể tạo ra được một bảng xếp hạng chặt chẽ, cân đối và minh bạch hơn (Baty, 2010).

Sau khi THE thay đổi quan điểm xếp hạng, hãng QS cũng chọn đối tác mới là *US News* để đưa ra bảng xếp hạng “Các trường đại học tốt nhất thế giới” (*World's Best Universities*) từ năm 2009. Và từ chỗ chỉ xếp hạng các trường đại học ở Hoa Kỳ, *US News* đã bắt đầu vươn ra phạm vi toàn cầu. Các tiêu chí và trọng số của bảng xếp hạng này dựa vào cả đánh giá từ bên ngoài và bên trong của trường đại học, về cơ bản vẫn giữ nguyên như khi QS hợp tác với THE, bao gồm kết quả khảo sát bình duyệt học thuật (*academic peer-review*) và khảo sát chủ lao động, tỉ lệ sinh viên:giảng viên, số lượng bài báo được trích dẫn, số lượng sinh viên và giảng viên quốc tế. Ngoài việc đưa ra danh sách 400 trường được điểm cao nhất thế giới trong năm, bảng xếp hạng này còn có các bảng phụ theo châu lục, bao gồm Châu Á và Trung Đông, *Australia* và *New Zealand*, Anh và châu Âu, và *Canada* (Morse, 2010).

Cùng xu hướng xếp hạng cấp trường dựa vào các nguồn dữ liệu học thuật, từ năm 2004, Phòng thí nghiệm Đo lường Mạng thông tin điện tử (*Cybermetrics Lab*) ở Tây Ban Nha đã phát triển một sáng kiến xếp hạng các trường đại học về mức độ sử dụng và phổ biến thông tin học thuật của mình thông qua mạng Internet: “Bảng xếp hạng đại học thế giới trên Mạng toàn cầu” (*Webometrics Ranking of World's Universities*). Về cơ sở lí luận của bảng xếp hạng này, Aguillo và cộng sự (2006) cho rằng Mạng học thuật toàn cầu (*academic web*) có khả năng đóng góp to lớn giúp nền giáo dục đại học hoàn thành được ba sứ mạng của mình: giảng dạy, nghiên cứu và chuyển giao công nghệ. Mạng học thuật toàn cầu với các *website* đại học và ấn bản khoa học là nguồn cung cấp và trao đổi thông tin khoa học nhanh chóng, tiện lợi và ít tốn kém hơn rất nhiều so với phương thức phổ biến thông tin truyền thống là sách báo in trên giấy. Các tiêu chí được xem xét trên *website* chính thức của trường đại học được xếp hạng là: số lượng trang mạng cung cấp thông tin về tổ chức, quy trình và hoạt động của nhà trường; số tập tin thuộc các định dạng giàu thông tin (được xem là nguồn cung cấp nội dung học thuật cho sinh viên và giảng viên) gồm PDF, PPT, DOC và PS; số bài báo khoa học được lập chỉ mục trong

*Google Scholar*, bộ máy tìm kiếm chuyên sâu từ các cơ sở dữ liệu học thuật; và tính khả kiến (*visibility*) trên Internet thông qua số đường liên kết từ các *website* bên ngoài dẫn vào. Điều cần lưu ý ở đây là các tác giả của phương pháp này không chỉ dựa vào yếu tố công nghệ mà còn dựa cả vào các yếu tố khoa học, với nhiều nghiên cứu so sánh thống kê mối liên hệ với các chỉ số trích dẫn của *ISI/Thomson* hay với kết quả xếp hạng *ARWU*. Nhiều kết quả nghiên cứu đã chứng tỏ rằng xếp hạng theo các tiêu chí thông tin trên Mạng toàn cầu là một cách có hiệu quả nhất định để tiếp cận và xếp hạng các trường đại học ở các nước đang phát triển hoặc không sử dụng tiếng Anh. Tương tự, các trường thiên về kỹ thuật và công nghệ sẽ được đánh giá tốt hơn so với các bảng xếp hạng khác dựa vào tiêu chí số lượng trích dẫn tham khảo vốn thiên về y sinh và khoa học tự nhiên (Aguillo *et al.*, 2006, 2008).

## **2.2. Xếp hạng theo đặc thù chuyên ngành**

Dù xu hướng xem cả trường đại học như là một chủ thể xếp hạng chung đang phổ biến mạnh mẽ, nhưng đó không phải là cách tiếp cận duy nhất trong đánh giá và xếp hạng đại học. Từ lâu, một số nơi đã đi sâu phân tích các chi tiết ở những bậc tổ chức thấp hơn trong các trường đại học, thường là khoa hay bộ môn, gọi chung là chương trình, để từ đó có thể đưa ra bảng xếp hạng có xét các yếu tố đặc thù theo lĩnh vực chuyên ngành. Có rất nhiều bảng xếp hạng dạng này, đặc biệt là trong lĩnh vực kinh tế - thương mại, tuy nhiên đáng kể nhất là các bảng xếp hạng do Trung tâm Phát triển Giáo dục Đại học (*Centrum für Hochschulentwicklung – CHE*) và Văn phòng Hợp tác Đại học Đức (*Deutsche Akademische Austauschdienst – DAAD*) phối hợp thực hiện (Usher, Savino, 2006).

Về phương pháp luận, CHE cho rằng nhóm đối tượng sử dụng bảng xếp hạng trước tiên là học sinh tốt nghiệp trung học. Cái họ cần là lựa chọn một chương trình chuyên ngành hơn là một trường đại học, do đó bảng xếp hạng nên chú trọng từng chương trình một thay vì tổng thể cả trường. Hơn nữa, thực tế cho thấy rằng ngay bên trong một trường đại học thường có sự khác biệt khá lớn giữa chương trình này và chương trình khác, và một đánh giá trung bình tổng thể về nhà trường không có ý nghĩa gì đối với sinh viên của một chương trình cụ thể. Hơn nữa, không có cơ sở lý thuyết hay kinh nghiệm thực tiễn nào chứng minh được tính hợp lý của việc xếp hạng dựa trên một giá trị tổng hợp từ các chỉ số định lượng riêng lẻ. Thay vào đó, chú trọng tính đặc thù cá nhân của nhóm đối tượng sử dụng chính là sinh viên tương lai, với những mục đích và mong đợi khác nhau, CHE chọn cách cho phép người dùng tự kết hợp các chỉ số khác nhau để thiết lập ra bảng xếp hạng theo ý muốn của mình. Điểm đáng lưu ý nhất của CHE, đó chính là quy trình xây dựng bộ tiêu chí xếp hạng, với một hệ thống chỉ số bao quát được nhiều khía cạnh khác nhau khi muốn đánh giá một chương trình đào tạo theo góc nhìn của các đối tượng sử dụng bảng xếp hạng. Các chỉ số ấy phản ánh nhiều loại thông tin từ tổng quát đến chi tiết, theo quan điểm tiếp cận đa chiều của CHE, với hai đặc trưng chính: các chỉ số lấy từ nhiều nguồn dữ liệu khác nhau và thuộc về cả hiện thực khách quan lẫn đánh giá chủ quan. Ngoài ra, CHE còn duy trì một cơ chế đảm bảo chất lượng: chuyển dữ liệu thu thập được từ các trường đại học và chương trình được xếp hạng về cho chính họ để kiểm tra, bổ sung, chỉnh sửa các sai sót nếu có; đồng thời hàng năm đều tổ chức một hội đồng tư vấn cho các chuyên ngành được xem xét nhằm kiểm tra, đánh giá các chỉ số và phương pháp được áp dụng cũng như tính xác thực của kết quả thu được (Federkeil, 2002; Van Dyke, 2005).

Cùng quan điểm với CHE về vai trò của đặc thù chuyên ngành trong xếp hạng đại học, nhưng với cách thu thập và xử lý dữ liệu khác, Van Raan đã cùng các cộng sự của mình tại Trung tâm Nghiên cứu Khoa học và Công nghệ (CWTS), Đại học Leiden (Hà Lan), phát triển “Bảng xếp hạng khoa học” (*Ranking of Science*) dựa vào các số liệu thống kê trích dẫn từ *Web of Science* (Thomson Reuters). Bảng xếp hạng này bao gồm khoảng 1.000 trường đại học có trên 700 bài báo trong *Web of Science* mỗi năm. Là nơi chuyên nghiên cứu về “đo lường thông tin thư mục” (*bibliometrics*), CWTS thiết lập một hệ thống các chỉ số khác với *ARWU* dù cả hai cùng sử dụng nguồn dữ liệu tương tự nhau (của Thomson Reuters, tiền thân là *ISI*). Đặc trưng quan trọng nhất của bảng xếp hạng này là “chỉ số đỉnh cao” (*crown indicator*): số lần trích dẫn quy theo mức độ tác động trên bình diện quốc tế trong chuyên ngành (*CPP/FCSm*), một khái niệm được nhóm nghiên cứu của CWTS phát triển trong những năm gần đây dùng để đánh giá mức độ tác động của các công trình nghiên cứu được công bố trên các tạp

chí khoa học quốc tế, tùy theo đặc thù của mỗi lĩnh vực chuyên ngành (Van Raan, 2003; Opthof, Leydesdorff, 2010).

Ngay cả những bảng xếp hạng theo xu hướng tổng thể cũng tỏ ra quan tâm nhiều hơn đến xu hướng xem xét các đặc thù chuyên ngành này. Như ARWU đã bắt đầu bổ sung các bảng xếp hạng trường đại học hàng đầu ở năm lĩnh vực nghiên cứu chính (khoa học tự nhiên và toán học; kỹ thuật, công nghệ và khoa học máy tính; khoa học về sự sống và nông nghiệp; y học lâm sàng và dược học; khoa học xã hội) kể từ năm 2007, và theo năm chuyên ngành sâu (toán học; vật lý; hoá học; khoa học máy tính; kinh tế & thương mại) kể từ năm 2009. Hay như mới đây, bảng xếp hạng của QS phối hợp với *US News* cũng có phân chia theo năm nhóm ngành: nghệ thuật và nhân văn; kỹ thuật, công nghệ và tin học; khoa học sự sống và y sinh; khoa học tự nhiên và vật lý; khoa học xã hội.

### 3. Những vấn đề đặt ra trong xếp hạng đại học

#### 3.1. Xếp hạng và vai trò, sứ mạng của trường đại học

Kể từ khi ra đời ở châu Âu từ thời Trung cổ, các trường đại học đã trở thành trung tâm của tiến trình phát triển xã hội loài người. Cho đến nay, giới nghiên cứu giáo dục đại học đã đồng thuận rằng đại học chính là nơi sáng tạo, lưu trữ và truyền bá tri thức của loài người. Ở từng thời điểm khác nhau, trong từng bối cảnh chính trị, kinh tế, xã hội khác nhau, ba chức năng này được phát triển trong các mô hình khác nhau, như là nơi “giảng dạy kiến thức phổ quát” (Newman, 1852), hay nơi kết hợp chặt chẽ giữa nghiên cứu và giảng dạy-đào tạo theo mô hình Humboldt (Husén, 1991), cho đến mô hình “vòng xoắn bộ ba” (*triple helix*) đại học-doanh nghiệp-chính phủ với ba sứ mạng căn bản của trường đại học là giảng dạy, nghiên cứu và chuyển giao tri thức cho xã hội (Etzkowitz, 2003). Đặc biệt, trong bối cảnh toàn cầu hoá hiện tại, các trường đại học còn phải chịu áp lực của quan hệ cung-cầu của thị trường và dần dần phải áp dụng các biện pháp hay mô hình tổ chức và quản lý của doanh nghiệp (Denman, 2005). Do đó, nếu muốn đo lường chất lượng của một trường đại học thì cần phải sử dụng nhiều phương pháp đánh giá hay xem xét nhiều mặt trong sứ mạng và lí do tồn tại của họ chứ không thể chỉ dựa trên một quan điểm tiếp cận duy nhất (Montesinos *et al.*, 2008).

Như đã phân tích ở trên, đại đa số các bảng xếp hạng đại học đều có những bộ chỉ số tập trung vào đánh giá hai nhóm năng lực nghiên cứu (thông qua số lượng bài báo khoa học công bố trên các tạp chí chuyên ngành, số lần được trích dẫn của các bài báo ấy, các giải thưởng nghiên cứu khoa học,...) và giảng dạy (thông qua các tiêu chí cho phép ước lượng hoặc phán đoán về kết quả đào tạo-giảng dạy như số lượng sinh viên, tỉ lệ sinh viên:giảng viên, kinh phí đầu tư vào hoạt động giảng dạy, tỉ lệ tốt nghiệp,...), còn lại, hầu như có rất ít những chỉ số xác đáng để đo lường “sứ mạng thứ ba” của trường đại học: chuyển giao công nghệ và tri thức cho xã hội.

Trong đó, dù là “điểm tựa” chủ lực của đa số các tiêu chí dùng trong các bảng xếp hạng đại học quốc tế cho đến nay, hoạt động nghiên cứu khoa học lại không xuất hiện trong định nghĩa một trường đại học trong thế kỉ XXI do Brian Denman (Đại học New England, Australia) đề xuất:

*“Đại học là một trường giảng dạy bậc cao phức tạp, chính thức được phép tổ chức đào tạo và cấp phát văn bằng ở trình độ cao trong ít nhất ba lĩnh vực hay chuyên ngành học tập.”*  
(Denman, 2005)

Theo ông, nghiên cứu khoa học đúng là hoạt động thiết yếu của tất cả các trường đại học, nhưng đó không phải là sứ mạng trước tiên của hầu hết các trường đại học hiện đang tồn tại trên thế giới. Khái niệm “ở trình độ cao” trong định nghĩa mà ông nêu ra đã bao hàm cả hai khái niệm “nghiên cứu” và “học tập chuyên sâu”.

Chính vì vậy, sự chú trọng quá mức đối với các tiêu chí kết quả nghiên cứu khoa học trong các bảng xếp hạng quốc tế – với giả định rằng “nghiên cứu giỏi thì sẽ dạy giỏi”, hay “nghiên cứu giỏi mới có thể dạy giỏi” – chưa hẳn đã là cách thích hợp để phản ánh được một cách khách quan tổng thể các hoạt động của một trường đại học cũng như là hiệu quả giảng dạy của họ.

### **3.2. Những điểm bất ổn trong phương pháp xếp hạng**

Không phủ nhận sức hút mạnh mẽ của các bảng xếp hạng đại học trong những năm gần đây, nhưng hầu hết các bảng xếp hạng này đều chưa có tính thuyết phục cao, nhất là đối với giới học thuật, bởi còn ẩn chứa nhiều điểm bất ổn trong phương pháp thực hiện của mình.

#### **3.2.1. Sai sót trong đo lường thông tin thư mục**

Cách đánh giá năng lực nghiên cứu của nhà khoa học dựa vào các thông số xuất bản và trích dẫn khoa học được gọi là phương pháp “đo lường thông tin thư mục” (*bibliometrics*). Mặc dù phải thừa nhận những ưu điểm vượt trội của các tạp chí chuyên ngành với hệ thống bình duyệt quốc tế, nhưng việc duyệt đăng, trích dẫn các bài báo khoa học cũng như thu thập dữ liệu vẫn có những yếu tố chủ quan hay khách quan dẫn đến kết quả thiếu chính xác, với tỉ lệ sai sót trong thống kê trích dẫn lên đến gần 1/3 (Moed, 2002; Van Raan, 2005).

Đó là chưa kể, hiện nay hầu như chưa có cơ sở khoa học chắc chắn nào cho phép khẳng định mối liên hệ giữa mức độ trích dẫn và chất lượng khoa học của một công trình đã công bố. Thậm chí, các cơ sở dữ liệu trích dẫn còn chưa xử lý rất ráo các trường hợp “trích dẫn âm” hay tự trích dẫn các công trình trước đó của mình; hoặc chưa phân biệt giữa các thể loại bài báo khoa học như bài báo nghiên cứu, bài báo tổng quan, bài viết trao đổi, thư ngắn, điểm sách,... vì mỗi loại bài đều có tầm quan trọng và mức độ ảnh hưởng khác nhau (Van Raan, 2005).

Bên cạnh đó, có một dạng sai sót khác khá phổ biến trong đo lường thông tin thư mục, đó là sự nhầm lẫn tên đơn vị chủ quản. Khi công bố kết quả nghiên cứu, các nhà khoa học tự khai báo cơ quan chủ quản của mình, và với các nước không sử dụng tiếng Anh thì nguy cơ sai sót rất cao, do dịch tên trường không chính xác, hoặc dùng nhiều tên gọi khác nhau. Đặc biệt là ở các nước có sự phân biệt trong hệ thống tổ chức giảng dạy (trường đại học) và nghiên cứu (viện nghiên cứu) thì nhiều khi khó phân định được nơi nào nên được ưu tiên khai báo khi đề tài được thực hiện với sự hợp tác của cả hai bên, và sự sai sót có thể dẫn đến việc trường đại học bị “mất điểm”. Tương tự, có người chỉ nêu tên đơn vị thành viên nơi họ làm việc trực tiếp chứ không phải tên chính thức của trường đại học trong quan hệ với bên ngoài; hay có những trường đại học trong cùng thành phố trùng tên với nhau,... Tỉ lệ sai sót dạng này trong các dữ liệu đo lường thư mục là rất đáng kể (Van Raan, 2005; Guillemin, 2007).

#### **3.2.2. Bất bình đẳng trong thông tin trích dẫn**

Các bảng xếp hạng hiện nay hầu như đều sử dụng thông tin đo lường thư mục lấy từ các cơ sở dữ liệu lớn như ISI/*Web of Science*, Scopus,... Điều đó dẫn đến một xu hướng bất bình đẳng, trước tiên là về chuyên môn, vì các cơ sở dữ liệu này thường thiên về các lĩnh vực khoa học tự nhiên và y sinh; phần “thua thiệt” thuộc về các trường kĩ thuật và công nghệ, khoa học xã hội, nhân văn và nghệ thuật. Nhưng bất bình đẳng rõ ràng nhất là về ngôn ngữ: hầu hết các công trình xuất bản được lập chỉ mục trong các cơ sở dữ liệu ấy đều bằng tiếng Anh; trong khi bộ phận xuất bản khoa học bằng các thứ tiếng khác như Đức, Pháp, Nga, Tây Ban Nha, Bồ Đào Nha, Nhật, Hoa, v.v. cũng có đóng góp quan trọng cho nền khoa học toàn cầu (Ren, Zu, Wang, 2002; Van Raan, 2005).

Sự bất bình đẳng còn biểu hiện ở chỗ, có xu hướng “thiên vị” các tác giả Hoa Kỳ bởi họ luôn chiếm lĩnh hàng loạt vị trí cao nhất trong các bảng xếp hạng hàng năm và các nguồn dữ liệu trích dẫn. Việc lạm dụng các tiêu chí mang tính chất thuần túy tính đếm số lượng trích dẫn khoa học sẽ rất nguy hiểm trong việc đánh giá trình độ khoa học hiện tại của các trường, các nước và các châu lục. Thực tế là Châu Âu có sự thụt lùi so với Hoa Kỳ trong nghiên cứu khoa học, ở một số chừng mực nào đó, nhưng không đến nỗi chênh lệch quá lớn như trong một số bảng xếp hạng thời gian gần đây (Ren, Zu, Wang, 2002; Van Raan, 2005; Albatch, 2006).

#### **3.2.3. Các tiêu chí xếp hạng thiếu thuyết phục**

Có thể nói, đây là vấn đề lớn nhất của hầu hết các bảng xếp hạng đại học quốc tế hiện nay, trong đó

ARWU là đối tượng bị chỉ trích nhiều nhất. Theo nhiều nhà nghiên cứu, cách sử dụng các chỉ số và xác định trọng số trong việc xếp hạng các trường đại học của ARWU là không hợp lý, kết quả công bố là không lặp lại được, mức độ sai lệch so với các nguồn dữ liệu khác có thể lên rất cao. Hiệu ứng quy mô của trường đại học và của toàn hệ thống giáo dục có tác dụng quyết định đến kết quả xếp hạng của ARWU. Mặt khác, cả bộ bốn tiêu chí và sáu chỉ số xếp hạng của ARWU có mối liên hệ chặt chẽ với nhau, và sự thiếu độc lập giữa các tiêu chí này ảnh hưởng đến tính chính xác khách quan trong kết quả của bảng xếp hạng (Altbach, 2006; Enserink, 2007; Salmi, Saroyan, 2007).

Ngoài ra, có lẽ cũng là do sự khác biệt trong quan niệm đánh giá thể hiện qua các chỉ số và trọng số khác nhau, các bảng xếp hạng thường đưa ra một danh sách các trường được xếp theo thứ tự từ trên xuống dưới, tạo ra một cảm giác rất rõ ràng về khoảng cách chênh lệch trình độ giữa các trường, cho dù rằng trong thực tế khoảng cách ấy không đáng kể, hoặc kết quả thu được không có ý nghĩa thống kê. Đồng thời, giữa các bảng xếp hạng khác nhau trong cùng một năm, thứ tự của nhiều trường có thể khác biệt rất lớn, trong khi không có ai đánh giá được mức độ tin cậy của chính bản thân các bảng xếp hạng này với nhau (Altbach, 2006; Enserink, 2007).

Bên cạnh đó, một nhóm tiêu chí khác thường được dùng trong xếp hạng, đó là kết quả khảo sát ý kiến của các đối tượng khác nhau (giảng viên, sinh viên, lãnh đạo, doanh nghiệp sử dụng lao động) nhằm đo lường mức độ đánh giá của họ về các chương trình đào tạo hay trường đại học. Vấn đề là, khảo sát danh tiếng theo cách này tiềm ẩn nguy cơ mang lại kết quả thiếu khách quan, nhất là khi tự đánh giá về mình. Và ít nhiều, do không có khung chuẩn về chất lượng của người đánh giá được lựa chọn tham gia chương trình, cũng như do yếu tố tâm lý và nhiều nguyên nhân khác, phương pháp này thường mang lại lợi thế cho các trường lâu đời và thiết thời cho các trường và chương trình mới (Salmi, Saroyan, 2007; Van Dyke, 2008).

#### **3.2.4. Đơn vị xếp hạng không hợp lý**

Vấn đề cuối cùng trong những điểm bất ổn của các bảng xếp hạng, đó là đơn vị so sánh khi xếp hạng. Như đã nói ở trên, thông thường không thể có một hình ảnh chất lượng chung về mọi thứ trong một trường đại học. Mỗi trường đều có những thế mạnh, sứ mạng, điều kiện tổ chức và kinh tế-xã hội riêng, sinh viên cũng học trong những chương trình riêng, do đó so sánh tất cả theo cùng một tiêu chí sẽ không đảm bảo đủ tính khách quan và công bằng (Federkeil, 2002; Altbach, 2006; Guillevin, 2007; Montesinos *et al.*, 2008).

Hơn nữa, hệ thống tổ chức giáo dục khác nhau giữa các nước cũng dẫn đến sự khác biệt trong nhiệm vụ xếp hạng. Ví dụ, nền giáo dục Hoa Kỳ hoạt động theo nhu cầu sử dụng dịch vụ giáo dục (*demand-driven*), việc xếp hạng thường đặt mục tiêu đáp ứng nhu cầu được cung cấp thông tin của người sử dụng dịch vụ giáo dục: họ tự do chọn lọc những thông tin nào thích hợp nhất với nhu cầu và mong đợi của mình. Trong khi đó, các trường đại học châu Âu lại hoạt động theo năng lực cung cấp dịch vụ giáo dục (*supply-driven*), việc xếp hạng được xem là cơ sở nền tảng để đánh giá chất lượng đào tạo và nghiên cứu ở bậc đại học. Chính vì vậy, khi xếp hạng toàn cầu, vượt qua khuôn khổ của một hệ thống giáo dục nhất định, rất khó để so sánh thật tốt về năng lực học thuật giữa các trường hoặc chuyên ngành nếu không lưu tâm đến sự khác biệt về hệ thống tổ chức; cũng như là, việc phân tích được sâu sắc các chỉ số năng lực (*performance indicator*) để đánh giá được hiệu quả hoạt động của một trường đại học không phải dễ dàng gì. Trong vấn đề này, cho đến nay vẫn chưa thấy được yếu tố thuyết phục của các bảng xếp hạng đã và đang tồn tại (Clarke, 2005; Jobbins, 2005; Kivinen, Hedman, 2008).

### **3.3. Xếp hạng đại học và chất lượng giáo dục**

Dù luôn khuyến cáo mọi người sử dụng một cách cẩn trọng, nhưng nhìn chung các bảng xếp hạng vẫn thường đem lại cho nhà chức trách, giới khoa bảng và công chúng một cảm giác vô hình về chất lượng nghiên cứu và/hoặc đào tạo của các trường đại học. Thế nhưng kết quả xếp hạng theo một vài tiêu chí nhất định đo đếm được ấy có nhất thiết phản ánh hay phản ánh đầy đủ chất lượng của một trường đại học hay không? Để trả lời câu hỏi này, cần phải quay trở lại với một câu hỏi căn bản: thế nào là chất

lượng giáo dục?

Đi tìm lời định nghĩa đó, Harvey và Green (1993) đã cho rằng chất lượng giáo dục liên quan mật thiết đến người dùng dịch vụ giáo dục cũng như bối cảnh tổ chức giáo dục. UNICEF (2000) còn nêu ra những điều kiện cụ thể hơn, không chỉ cho chất lượng giáo dục nói chung mà cả cho chất lượng giáo dục đại học, liên quan đến các mặt sức khoẻ và điều kiện sống của người học; môi trường học tập; nội dung, chương trình, phương tiện, tiến trình học tập; kết quả và thành tựu học tập gắn với đời sống xã hội,... Từ đó, có thể hiểu giáo dục là một hệ thống phức tạp gắn bó chặt chẽ với bối cảnh chính trị, văn hoá và kinh tế của mỗi quốc gia, đồng thời chịu cả các ảnh hưởng toàn cầu và quốc tế trong việc thảo luận về khái niệm chất lượng giáo dục. Có nghĩa là, các hệ thống xếp hạng hiện nay chưa thể đánh giá được chất lượng giáo dục đại học một cách tổng thể, nhất là khi vẫn còn chưa có sự đồng thuận cao về các yếu tố cấu thành nên chất lượng trong giáo dục đại học.

Rất thông thường, các bảng xếp hạng chỉ sử dụng một vài chỉ số năng lực nhất định để cho điểm các trường được xếp hạng. Các chỉ số này được chọn lọc có chủ đích, phản ánh quan niệm cụ thể của nhóm xây dựng bảng xếp hạng đối với chất lượng nghiên cứu và chất lượng giáo dục. Tuy nhiên, các tiêu chí xếp hạng được đo đếm với những mục đích riêng biệt đó dễ làm người ta quên rằng thực chất các chỉ số năng lực ấy thường là do chính phủ hay bản thân các trường công bố nhằm cho thấy mức độ hoàn thiện của một trường khi so với một bộ tiêu chuẩn đối sánh, hoặc chỉ đơn giản là phục vụ mục tiêu “minh bạch hoá” trong một tiến trình đảm bảo chất lượng giáo dục nói chung. (Usher, Savino, 2006).

Mặc khác, một số bảng xếp hạng áp dụng một trọng số quan trọng cho tỉ lệ giảng viên: sinh viên vì cho rằng đây là một yếu tố quan trọng quyết định (hoặc biểu hiện) chất lượng giảng dạy mà sinh viên có thể tiếp thụ được. Tuy vậy, trong giới học thuật vẫn chưa có bằng chứng rõ ràng về việc quy mô lớp học lớn hay nhỏ có tác động như thế nào đến kết quả học tập của sinh viên. Hơn nữa, khi chú trọng quan hệ số lượng giữa quy mô lớp học và thành tựu học thuật thì người ta thường quên một số yếu tố quan trọng khác, như cảm nhận về điều kiện làm việc của giáo viên, hay tinh thần trách nhiệm của họ khi giảng dạy (UNICEF, 2000).

Đồng thời, do bị ràng buộc bởi tính chất phải đo đếm được của các chỉ số năng lực, các bảng xếp hạng thường chỉ đánh giá trên kết quả đầu ra chứ không phải đầu vào. Tuy vậy, để đạt được chất lượng đầu ra thì nhất thiết phải quan tâm đến chất lượng đầu vào (Kivinen, Hedman, 2008). Ít nhất, với các trường muốn được xếp hạng cao, thì điều này có nghĩa rằng họ phải cải thiện chính các yếu tố đầu vào của mình để có thể đạt được kết quả tốt hơn. Một trường học với môi trường vật lí, tâm lí và dịch vụ có chất lượng chính là nền tảng cần thiết để đảm bảo chất lượng cho các hoạt động dạy-học và nghiên cứu của sinh viên cũng như giảng viên (UNICEF, 2000). Cho đến nay không có bảng xếp hạng nào đo lường được các yếu tố nền tảng cho chất lượng giáo dục nói chung và giáo dục đại học nói riêng này.

#### 4. Kết luận

Với các nhược điểm về phương pháp luận, những khác biệt về quan điểm tiếp cận, sự thiếu hụt những nguồn cung cấp dữ liệu khách quan, đầy đủ và đáng tin cậy, v.v. như đã phân tích ở trên, có thể thấy còn lâu mới có được một hay những bảng xếp hạng nào đó nhận được sự đồng thuận từ nhiều phía. Tuy nhiên, sự phát triển và tầm ảnh hưởng ngày càng rộng khắp của các bảng xếp hạng đại học trên thế giới là điều không cần phải bàn cãi. Vấn đề là, đối diện với sức lan toả nhanh chóng đó, hoặc là chúng ta nhất quyết bỏ qua, không bận tâm đến bất cứ bảng xếp hạng nào, hoặc là tìm cách phân tích để hiểu rõ phạm vi giới hạn của chúng nhằm có cách sử dụng thích hợp nhất (Altbach, 2006; Salmi, Saroyan, 2007).

Vậy thì cần phải làm gì trước trào lưu xếp hạng này? - Câu trả lời là, *cần phải xác định rõ ràng mục đích sử dụng các bảng xếp hạng để tránh rơi vào tình trạng lạm dụng hoặc diễn dịch sai lệch kết quả xếp hạng*. Salmi và Saroyan (2007) đã khuyến cáo rằng:

- Với các trường đại học, các bảng xếp hạng có thể giúp ích trong việc xây dựng kế hoạch chiến lược cũng như cải thiện chất lượng giảng dạy và nghiên cứu của nhà trường. Điều cần chú ý ở

đây, đó không phải là định ra một số điểm cần đạt được hay một thứ hạng cần vượt qua trong một bảng xếp hạng quốc tế, mà là phân tích các chỉ số cụ thể để hiểu rõ kết quả xếp hạng ấy, từ đó suy nghĩ về những mặt mạnh và yếu của mình, đề ra chiến lược hành động để chỉnh sửa, hoàn thiện và nâng cao dần chất lượng giảng dạy, đào tạo và nghiên cứu.

- Với nhà chức trách, chỉ nên sử dụng các bảng xếp hạng để khuyến khích văn hoá chất lượng. Ở các nước chưa có hệ thống đánh giá hay tiêu chuẩn hoá chất lượng giáo dục thì các bảng xếp hạng có thể là công cụ để đo lường chất lượng; còn ở phạm vi quốc tế thì đây có thể là công cụ kiểm soát chất lượng sinh viên nước ngoài khi không có một tổ chức chuyên trách đảm bảo chất lượng toàn cầu. Khi muốn xây dựng một hệ thống xếp hạng quốc gia, nhất thiết phải cho phép các trường cùng tham gia, xây dựng phương pháp, xác định mục tiêu chung, tạo ra tinh thần chia sẻ và trách nhiệm tập thể.
- Với sinh viên và gia đình, nhà tuyển dụng hay công chúng, chỉ nên xem các bảng xếp hạng là nguồn cung cấp thông tin và tạo kênh thảo luận đa chiều. Đặc biệt, cần chú trọng các bảng xếp hạng đa tiêu chí thay vì đồng quy mọi chỉ số về một tiêu chí duy nhất, vì chỉ có các bảng xếp hạng này mới cung cấp được thông tin xác hợp với chương trình đào tạo mà sinh viên cần tìm, hay về đối tượng và kỹ năng nghề nghiệp mà chủ doanh nghiệp cần tuyển dụng.

Và cuối cùng, xét trong mối quan hệ giữa xếp hạng và chất lượng giáo dục, Harvey và Green (1993) đã cho rằng sử dụng máy móc các chỉ số năng lực có nguy cơ dẫn đến tác dụng ngược đối với mục đích tạo ra chất lượng thực sự, vì mỗi thành phần tham gia vào hệ thống giáo dục đều có những mục tiêu, nhu cầu và cách thức định nghĩa không hoàn toàn trùng khớp với nhau về chất lượng. Chính vì vậy, thay vì bắt đầu với một cách định nghĩa chất lượng giáo dục duy nhất (thông qua một con số, một vị trí xếp hạng), thì xác định những tiêu chí mà các nhóm lợi ích khác nhau dùng để đánh giá chất lượng mới là cách khả dĩ để tìm ra một giải pháp linh hoạt, cho phép hiểu và thừa nhận các góc nhìn và quan niệm khác nhau thể hiện đúng đặc điểm của từng nhóm chia sẻ lợi ích đối với chất lượng giáo dục, tạo tiền đề xây dựng một bộ tiêu chí thích hợp để đánh giá chất lượng giáo dục.

Thiết nghĩ, với những kết luận và khuyến cáo của giới học thuật quốc tế như trên, các nhà chức trách và lãnh đạo nền giáo dục đại học Việt Nam trước khi đặt ra mục tiêu về vị trí xếp hạng cần phải quan tâm đến ý nghĩa thực sự của chất lượng giáo dục đối với đối với người học nói riêng và đối với xã hội nói chung, để có sự đầu tư hợp lý vào việc xây dựng một chiến lược và tiến trình nâng cao chất lượng của hệ thống đại học nước nhà; tiến trình ấy tuy có thể chậm, nhưng phải vững chắc, với tầm nhìn không chỉ giới hạn trong phạm vi năm năm, mười năm hay hai mươi năm, mà thậm chí phải còn dài hơn thế./.

### Tài liệu tham khảo

- [1] **Aguillo IF, Granadino B, Ortega JL, Prieto JA.** 2006. Scientific research activity and communication measured with cybermetrics indicators. *Journal of the American Society for Information Science and Technology*, 57(10): 1296-1302.
- [2] **Aguillo IF, Ortega JL, Fernandez M.** 2008. Webometric ranking of world universities: introduction, methodology, and future developments. *Higher Education in Europe*, 33(2-3): 233-244.
- [3] **Albath PG.** 2006. The dilemmas of ranking. *International Higher Education*, (42): 2-3.
- [4] **Baty P.** 2009. Rankings 09: Talking points [trực tuyến]. *Times Higher Education*, 8 October 2009. Địa chỉ truy cập: <<http://www.timeshighereducation.co.uk/story.asp?storycode=408562>> [tham khảo ngày 01/05/2010].
- [5] **Baty P.** 2010. THE unveils broad, rigorous new rankings methodology [trực tuyến]. *Times Higher Education*, 3 June 2010. Địa chỉ truy cập : <<http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=411907&c=1>> [tham khảo ngày 29/07/2010].
- [6] **Clark M.** 2005. Quality assessment lessons from Australia and New Zealand. *Higher Education in Europe*, 30(2): 183-197.



- [7] **Denman B.** 2005. Comment définir l'université du XXI<sup>e</sup> siècle ? *Politiques et gestion de l'enseignement supérieur*, 17(2): 9-30.
- [8] **Enserink M.** 2007. Who ranks the university rankers? *Science*, 317(5841): 1026-1028.
- [9] **Etzkowitz H.** 2003. Innovation in innovation: the triple helix of university-industry-government relations. *Social Science Information*, 42(3): 293-337.
- [10] **Federkeil G.** 2002. Quelques aspects de la méthodologie de classement – le classement du CHE des universités allemandes. *L'Enseignement supérieur en Europe*, 27(4): 389-398.
- [11] **Guillevin L.** 2007. Classement de Shanghai : valorisons nos universités en améliorant le classement des publications. *La presse médicale*, 36(12 I): 1709-1711.
- [12] **Harvey L, Green D.** 1993. Defining quality. *Assessment and Evaluation in Higher Education*, 18(1): 9-34.
- [13] **Husén T.** 1991. The idea of a university: changing roles, current crisis and future challenges. 2<sup>th</sup> UNESCO/NGO Collective Consultation on Higher Education, Paris, France, 8-11 April. UNESCO New Papers on Higher Education, ED.91/WS/11, 48 p.
- [14] **Jobbins D.** 2005. Moving to a global stage: A media view. *Higher Education in Europe*, 30(2): 137-145.
- [15] **Kivinen O, Hedman J.** 2008. World-wide university rankings: A Scandinavian approach. *Scientometrics*, 74(3): 391-408.
- [16] **Liu NC, Cheng Y, Liu L.** 2005. Academic ranking of world universities using scientometrics: A comment to the Fatal Attraction. Author's reply. *Scientometrics*, 64(1): 101-112.
- [17] **Moed HF.** 2002. The impact-factors debate: The ISI's uses and limits. *Nature*, 415(6873): 731-732.
- [18] **Montesinos P, Carot JM, Martinez JM, Mora F.** 2008. Third mission ranking for world class universities: beyond teaching and research. *Higher Education in Europe*, 33(2-3): 259-271.
- [19] **Morse R.** 2010. *World's Best Universities: the methodology* [trực tuyến]. U.S.News & World Report. Địa chỉ truy cập: <<http://www.usnews.com/articles/education/worlds-best-universities/2010/02/25/worlds-best-universities-the-methodology.html>> [tham khảo ngày 01/05/2010].
- [20] **Newman JH.** 1852. Discourse 2. Theology a branch of knowledge. In: *The idea of a university*. London: Longmans, Green, and Co., 1907, p. 19-42 [new impression]. Phiên bản trực tuyến: <<http://www.newmanreader.org/works/idea/index.html>> [tham khảo ngày 22/04/2009].
- [21] **Ophof T, Leydesdorff L.** 2010. Caveats for the journal and field normalizations in the CWTS (“Leiden”) evaluations of research performance. *Journal of Informetrics*, 4(3): 423-430.
- [22] **Ren SL, Zu GA, Wang HF.** 2002. Statistics hide impact of non-English journals. *Nature*, 415(6873): 732.
- [23] **Sadlak J.** 2006. Validity of university ranking and its ascending impact on higher education in Europe. *Bridges*, vol. 12, december 2006.
- [24] **Salmi J, Saroyan A.** 2007. Les palmarès d'universités comme moyens d'action : usages et abus. *Politiques et gestion de l'enseignement supérieur*, 19(2) : 33-74.
- [25] **United Nations Children's Fund (UNICEF).** 2000. *Defining quality in education*. Working Paper Series, Document No. UNICEF/PD/ED/00/02. Presented at the meeting of The International Working Group on Education Florence, Italy, June 2000.
- [26] **Usher A, Savino M.** 2006. *A world of difference: a global survey of university league tables*. Canadian Education Report Series. Toronto, ON: Educational Policy Institute.
- [27] **Van Dyke N.** 2005. Vingt ans de registres de résultats universitaires. *L'enseignement supérieur en Europe*, 30(2): 7-32.
- [28] **Van Dyke N.** 2008. Self- and peer-assessment disparities in university ranking schemes. *Higher Education in Europe*, 33(2-3): 285-293.
- [29] **Van Raan AF.** 2003. The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technikfolgenabschätzung, Theorie und Praxis* [Technology Assessment - Theory and Practice], 12(1): 20-29.
- [30] **Van Raan AF.** 2005. Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1): 133-143.