

## Bộ ba “mở” cải thiện độ tin cậy của khoa học xã hội

Hồ Mạnh Tùng (*dịch*)

*Khoa học & Phát triển*

11/02/2018 07:10

*KH&PT* (11/2/2018; <http://khoa hocphattrien.vn/khoa-hoc/bo-ba-mo-cai-thien-do-tin-cay-cua-khoa-hoc-xa-hoi/20180206031637843p1c160.htm>) -- TS Vương Quân Hoàng (Đại học Thành Tây, Hà Nội và ULB, Bỉ) đề xuất trên *Scientific Data Updates* việc kết hợp bộ ba yếu tố dữ liệu mở, phản biện mở và đối thoại cộng đồng mở như một giải pháp đáp ứng đòi hỏi của xã hội đối với độ tin cậy của các kết quả nghiên cứu.

Thế giới ngày càng nhận thức rõ thách thức từ các kết quả nghiên cứu thiếu độ tin cậy cần thiết như tính tái lập kết quả nghiên cứu và điều này ảnh hưởng nghiêm trọng tới niềm tin của xã hội vào ngành khoa học xã hội. Một số kết quả khoa học lừng lẫy đã bị lật tẩy bản chất ngụy tạo dữ liệu hoặc không có cơ sở học thuật chặt chẽ.

Nhiều nhà nghiên cứu đã bỏ công điều tra mức độ tin cậy của các khám phá trong khoa học xã hội và nhân văn. Một ví dụ điển hình là Tỷ số cảm xúc tích cực tối thiểu tới hạn (Critical Minimum Positivity Ratio) 2.9013 của Fredrickson và Losada (2005), được tuyên bố là “cho phép dự đoán mức độ phát triển khả dĩ của một cá nhân thông qua tỷ lệ cảm xúc tích cực so với tiêu cực theo bước tiến thời gian của người đó”.

Con số huyền bí này từng có ảnh hưởng lớn và được ngưỡng mộ gọi là “hằng số” tâm lí cho đến khi nghiên cứu của Brown, Sokal và Friedman (2013) chỉ ra, đó chẳng qua là một con số thiếu cơ sở, tùy tiện và vô nghĩa.

Để cứu vãn sự tổn thất niềm tin của xã hội đối với độ tin cậy của các kết quả nghiên cứu trong khoa học xã hội, tôi đề xuất kết hợp bộ ba yếu tố dữ liệu mở, phản biện mở và đối thoại cộng đồng mở như một lựa chọn khả thi.



*Những nền tảng ấn bản phẩm mở như F1000Research cho phép các bài báo hoàn chỉnh được xuất bản chỉ trong vòng một tuần, trước cả khi quá trình phản biện bắt đầu.*

*Ảnh: [blog.f1000.com](http://blog.f1000.com)*

### **Dữ liệu mở**

Hiện nay vẫn tồn tại những cản trở về tâm lý và tài chính cho việc mở dữ liệu. Mặc dù thoạt nghe thì có vẻ đơn giản, nhưng việc theo đuổi chính sách dữ liệu mở lại rất phức tạp và không dễ đạt được, do đó “việc đưa ra những cú hích để hoạt động khoa học trở nên mở hơn đòi hỏi sự nỗ lực và hợp tác của tất cả các bên liên quan” (Nosek et al 2015). Thật không may là cộng đồng khoa học chưa thật sự nghiêm túc xem xét khả năng này (Begley and Ellis 2012), và như vậy đã bỏ qua cơ hội để dữ liệu có giá trị của họ được thử thách đăng trường tồn với thời gian.

Với sự ra đời của những hệ thống lưu trữ dữ liệu mở hoạt động ngày một hiệu quả hơn như Harvard Dataverse, Open Science Framework, Mendeley, UK Data Archive, chưa bao giờ việc chia sẻ dữ liệu cho cộng đồng sử dụng và tái lập các kết quả khoa học lại dễ dàng như hiện nay.

Những nền tảng ấn bản phẩm mở như *F1000Research* thậm chí giúp lưu trữ đồng thời nhiều bộ dữ liệu (có liên quan về logic khoa học) một cách an toàn trên chính hệ thống của họ mà không nhờ đến bên cung cấp dịch vụ thứ ba.

Tạp chí *Scientific Data* còn tiến xa hơn khi loại bỏ những rào cản gây hạn chế việc sử dụng dữ liệu mở cho mục tiêu thương mại bằng cách lên tiếng ủng hộ việc tạo ra giá trị kinh tế từ dữ liệu thông qua những công ty thương mại (Scientific Data Editorial, “Open for Business”).

Một lợi ích đối với những tác giả chia sẻ dữ liệu cùng với bài báo khoa học của mình là dữ liệu của họ được kiểm tra lại bởi những học giả khác, thậm chí còn dẫn tới những khám phá mới trong khoa học. Họ có cơ hội được đồng nghiệp xác nhận độ tin cậy của dữ liệu và độ chắc chắn của quá trình khoa học – những phẩm chất đóng vai trò quan trọng trong quá trình tích tụ cải thiện để làm giảm sự bất định (Aarts et al 2015). Đây chính là nền tảng cho khả năng tự hiệu chỉnh của khoa học.



*Với sự ra đời của những hệ thống lưu trữ dữ liệu mở như Harvard Dataverse, Open Science Framework, Mendeley, UK Data Archive, chưa bao giờ việc chia sẻ dữ liệu cho cộng đồng sử dụng và tái lập các kết quả khoa học lại dễ dàng như hiện nay.*

*Ảnh: Internet*

Trong kinh nghiệm của tôi, có những lúc quá trình tái sử dụng dữ liệu có thể diễn ra rất nhanh. Khi xuất bản một bài báo về dữ liệu trên tạp chí *Data in Brief*, với bộ dữ liệu mở được đưa lên Mendeley Data (Vuong & Nguyen, 2016), ban đầu tôi không dám chắc dữ liệu đó đóng góp được gì cho cộng đồng khoa học.

Thật bất ngờ, một nhà nghiên cứu tại Đại học Pittsburgh đã sử dụng chúng để hoàn thành một công bố khoa học có chất lượng cho tạp chí *International Journal of Equity in Health* (Nguyen, 2017). Mọi chuyện đã xảy ra nhanh như thế nào? Bài báo của tôi được xuất bản vào 30-9-2016, và bản thảo của Nguyen được nộp cho tạp chí vào ngày 15-11-2016!

Xuất bản dữ liệu mà sau đó được tái sử dụng là một trải nghiệm hết sức đặc biệt. Những câu hỏi bản thảo về giá trị của dữ liệu được giải đáp không chỉ một lần, với quá trình phản biện và biên tập, mà là hai lần và có lẽ là hơn nữa, với những công bố khoa học tiếp theo có sử dụng bộ dữ liệu đã được đóng góp công khai.

### **Phản biện mở**

Trong khi khuyến khích tính minh bạch và công khai của khoa học hướng tới làm khoa học trở nên tốt hơn cho tất cả, thì khái niệm phản biện mở hoàn toàn đã dần trở dậy như một giải pháp đáng khích lệ.

Khái niệm này có thể còn lạ lẫm đối với những nhà nghiên cứu trong khoa học xã hội, nơi mà quá trình phản biện kín hai chiều vẫn đang thống trị; nhưng trên lý thuyết, những thế mạnh của nó đã được chứng minh.

Bằng chứng từ Walsh, Rooney, Appleby và Wilkinson (2000) ủng hộ phản biện mở vì những phản biện có chữ ký chính danh được đánh giá có chất lượng cao hơn, phong thái lịch sự và tận tâm hơn. Hơn nữa, cũng không thể lờ đi giá trị xã hội của phản biện mở vì “sự tiến bộ trong chính sách của tạp chí có thể giúp những giá trị đó [giá trị nghiên cứu] trở nên hiển hiện hơn trong thực tiễn hằng ngày, và cuối cùng, sẽ làm gia tăng niềm tin của cộng đồng vào khoa học cũng như khiến cho chính khoa học tiến bộ hơn,” (Nosek et al 2015).

Những nhà tài trợ nặng ký đã sớm nhìn thấy giá trị của quá trình phản biện minh bạch, sự đánh giá của quần chúng, và khả năng của hệ thống nhằm đối phó với những nhược điểm đáng kể của quá trình phản biện nghiên cứu liên ngành như Bammer (2016) đã nêu ra.

Đó chính là lý do những tạp chí trẻ như *F1000Research* nhận được ủng hộ mạnh mẽ không lâu sau khi ra mắt (Van Noorden, 2013), từ Quỹ Wellcome Trust vào năm 2016 (Butler 2016) và Quỹ Bill & Melinda Gates vào năm 2017 (Butler 2017). Như Begley và Ellis (2012) đã chỉ ra, khoa học xã hội ngày nay cần một hệ thống phản biện mở để tạo điều kiện cho quá trình khám phá minh bạch mang lại lợi ích xã hội.

## **Đối thoại cộng đồng mở**

Chúng ta đang nói đến một khái niệm mở rộng của thuật ngữ “đối thoại” bao gồm những thảo luận mang tính kỹ thuật của giới chuyên gia về phương pháp luận khoa học và mã máy tính (Eglen et al 2017) và các quá trình trao đổi nghiên cứu sẵn sàng được đưa ra cho cộng đồng đánh giá, phê bình, tái sử dụng hoặc mở rộng (Nosek et al 2015).

Hệ thống phản biện cộng đồng PubPeer đã thúc đẩy nhu cầu đối thoại cộng đồng mở đi xa hơn, mặc dù việc đánh dấu một bài báo khoa học là “có vấn đề” trên trang này đang bị nhiều người coi (một cách không chính xác) là sự đe dọa.

Thực ra, đây là một cộng đồng chuyên gia kỹ thuật có thể giúp chỉnh sửa các vấn đề, giải quyết sự non yếu về thống kê (nhằm giúp chất lượng chung của các bản thảo nghiên cứu trở nên tốt hơn), và phục vụ với tư cách là một người gác cửa về chất lượng cho những tạp chí khoa học. PubPeer đã thúc đẩy điều mà Eglen et al (2017) ủng hộ, “Chia sẻ phương pháp luận và mã máy tính.”

Trong khoa học xã hội, giá trị của việc chia sẻ nằm ở chỗ giúp cập nhật và xác minh “những kết quả nghiên cứu với tư cách sự thật ước định” trong một thế giới xác suất Bayesian, nơi mà “một biện luận khoa học quan trọng sai lầm không nhất thiết phải dẫn đến một kết luận sai lầm,” theo Định lý Gödel (Jaynes 2003).

Ngày nay, bảo vệ niềm tin vào khoa học xã hội có nghĩa là tham gia vào những đối thoại cộng đồng mở bởi điều này giúp đảm bảo tính chắc chắn và tăng tính hiệu quả của các phương pháp nghiên cứu. Sự kết hợp của dữ liệu mở, phản biện mở và đối thoại mở thoạt nghe thì đơn giản nhưng để thực thi bộ ba này trong thực tế không hề dễ dàng.

Tuy nhiên, nếu đồng ý với quan điểm của Begley và Ellis (2012), “quá trình khoa học đòi hỏi tiêu chuẩn cao nhất về chất lượng, đạo đức và sự chặt chẽ”, thì bộ ba vừa thảo luận ở trên xứng đáng được triển khai, bất chấp các thách thức. Quan trọng hơn, chúng có khả năng thành công và giúp tái lập độ tin cậy cần có của các nghiên cứu trong khoa học xã hội.

Vì thế, tôi nhận thấy rằng bộ ba dữ liệu mở, phản biện mở và đối thoại mở sẽ dần trở nên trở thành yêu cầu bắt buộc trên phương diện chính sách, dù nguồn tài trợ của nghiên cứu đến từ đâu.

*\*Bài viết được dịch với sự đồng ý của Tổng biên tập tạp chí Scientific Data Updates, TS Andrew Hufton.*

### **Tài liệu tham khảo:**

- Aarts AA et al (2015) Estimating the reproducibility of psychological science. *Science*; 349 (6251): aac4716, doi:10.1126/science.aac4716.
- Bammer G (2016) What constitutes appropriate peer review for interdisciplinary research? *Palgrave Communications*; 2: 16017, doi:10.1057/palcomms.2016.17.
- Begley CG and Ellis LM (2012) Drug development: Raise standards for preclinical cancer research. *Nature*; 483 (7391): 531–533, doi:10.1038/483531a.
- Brown NJL, Sokal AD, and Friedman HL (2013) The complex dynamics of wishful thinking: the critical positivity ratio. *American Psychologist*; 68 (9): 801–813, doi:10.1037/a0032850.
- Butler D (2016) Wellcome Trust launches open-access publishing venture. *Nature News*.
- Butler D (2017) Gates Foundation announces open-access publishing venture. *Nature*; 543 (7647): 599, doi:10.1038/nature.2017.21700.
- Eglen SJ et al (2017) Toward standard practices for sharing computer code and programs in neuroscience. *Nature Neuroscience*; 20 (6): 770-773, doi:10.1038/nn.4550.
- Fredrickson BL and Losada MF (2005) Positive affect and the complex dynamics of human flourishing. *American Psychologist*; 60 (7): 678-686, doi:10.1037/0003-066X.60.7.678.
- Jaynes ET (2003) *Probability Theory: The Logic of Science*. London, UK: Cambridge University Press.
- Nguyen HM (2017) Inequality in healthcare costs between residing and non-residing patients: evidence from Vietnam. *Int. Journal for Equity in Health*; 16 (1): 76, doi:10.1186/s12939-017-0581-3.
- Nosek BA et al (2015) Promoting an open research culture. *Science*; 348 (6242): 1422-1425, doi:10.1126/science.aab2374.
- Scientific Data Editorial (2017) Open for business. *Scientific Data*; 4: 170058, doi:10.1038/sdata.2017.58.
- Van Noorden R (2013) Company offers portable peer review. *Nature*; 494 (7436): 161, doi:10.1038/494161a.
- Vuong QH (2017) Survey data on Vietnamese propensity to attend periodic general health examinations. *Scientific Data*; 4: 170142, doi:10.1038/sdata.2017.142.

Vuong QH and Nguyen TK (2016) Data on Vietnamese patients' financial burdens and risk of destitution. *Data in Brief*; 9: 543-548, doi:10.1016/j.dib.2016.09.040.

Walsh E, Rooney M, Appleby L, and Wilkinson G (2000) Open peer review: a randomised controlled trial. *British Journal of Psychiatry*; 176 (1) 47-51; doi:10.1192/bjp.176.1.47.