

Bibliyometri: Araştırma Ölçevleri için Leiden Manifestosu

The Leiden Manifesto for research metrics, Nature, 520, 429-431 (23 April 2015) doi:10.1038/520429a
<http://www.leidenmanifesto.org/>

Diana Hicks^a, Paul Wouters^b, Ludo Waltman^c, Sarah de Rijcke^d & Ismael Rafols^e.

a Diana Hicks Georgia Teknoloji Enstitüsü'nde (Atlanta, ABD) kamu politikası profesörüdür.

b Paul Wouters bilimetre profesörü ve yöneticidir.

c Ludo Waltman araştırmacıdır.

d Sarah de Rijcke Leiden Üniversitesi Bilim ve Teknoloji Çalışmaları Merkezi'nde (Hollanda) yardımcı doçenttir.

e Ismael Rafols İspanya Ulusal Araştırma Konseyi ve İspanya Valencia Polytechnic Üniversitesi'nde bilim politikası araştırmacıdır.

22 Nisan 2015

Bu on ilkeyi araştırma değerlendirmesine rehberlik etmesi için kullanın. Diana Hicks, Paul Wouters ve meslektaşlarını önerin.

Konu terimleri: Kariyer, Araştırma yönetimi, Yayıncılık

Bilim yönetiminde veri giderek daha fazla kullanılmaktadır. İsmarlama olarak hakemler tarafından başlatılan ve sıradanlaşan araştırma değerlendirmeleri günümüzde artık ölçümlere (metrics) dayanmaktadır¹. Sorun şu ki; günümüzde değerlendirme işi, muhakemeden ziyade veriler tarafından yönlendirilmektedir. Ölçümler hızla artmaktadır: genellikle iyi niyetli girişimler sonucu, her zaman geniş bilgiye dayanmamaktadır ve çoğunlukla kötü bir şekilde uygulanmaktadır. Değerlendirme; iyi uygulama, yorumlama bilgisi veya önerisi olmayan kuruluşlar tarafından giderek daha fazla uygulandığı için sistemi geliştirmek üzere tasarlanan araçların sisteme zarar vermesi riskiyle karşı karşıyayız.

2000 yılından önce, Institute for Scientific Information-ISI (Bilimsel Bilgi Enstitüsü) tarafından CD-ROM ortamında yayınlanan ve profesyonel analizler için uzmanlar tarafından kullanılan Science Citation Index-SCI (Bilim Atıf İndeksi) vardı. 2002 yılında, Thomson Reuters entegre bir web platformu oluşturdu ve Web of Science, kısacası WoS veri tabanını geniş çapta erişilebilir kıldı. Bundan sonra WoS'a rakip diğer atıf dizinleri türemeye başladı: 2004 yılında Elsevier tarafından tanıtılan Scopus'un ardından aynı yıl Google Akademik'in beta sürümü piyasaya sürüldü. Kurumsal araştırma verimliliğini ve etkisini rahatça karşılaştırmak amacıyla (WoS verilerini kullanarak) InCites ile (Scopus verilerini kullanarak) SciVal gibi bazı web tabanlı araçlar ile Google Akademik'i kullanarak bireysel atıf profillerini analiz etmeye yarayan yazılımlar geliştirildi ("Yayınla ya da yok ol" anlamına gelen "Publish or Perish", 2007'de yayınlandı).

Bireysel araştırmacıların atıf sayımını popüler hale getiren h-İndeksi 2005 yılında Kaliforniya San Diego Üniversitesi'nde fizikçi olarak görev yapan Jorge Hirsch tarafından önerildi. Dergi etki faktörüne olan ilgi 1995 yılından sonra istikrarlı bir şekilde arttı (bkz. Şekil: 'Etki faktörü endişesi').

Son zamanlarda sosyal kullanım ve çevrimiçi yorumlarla ilgili ölçümler ivme kazanmıştır- 2002 yılında F1000Prime, 2008 yılında Mendeley ve 2011 yılında Nature Publishing Group'un sahibi

Macmillan Bilim ve Eğitim (MacMillan Science and Education) tarafından desteklenen Altmetric.com kurulmuştur.

Bilimetre uzmanları, sosyal bilimciler ve araştırma yöneticileri olarak bizler bilimsel performansın değerlendirilmesinde, göstergelerin yaygın olarak yanlış kullanımını artan bir endişeyle gözlemledik. Çeşitli örneklerden birkaçı şu şekildedir: Dünya genelinde, üniversitelerin Şangay Sıralaması ve Times Yüksek Öğretim listesi gibi uluslararası sıralamalardaki konumlarına bağlılıkları saplantı düzeyine ulaşmıştır; bizim görüşümüze göre bu tür listeler, yanlış veriler ve keyfi göstergelere dayanmaktadır.

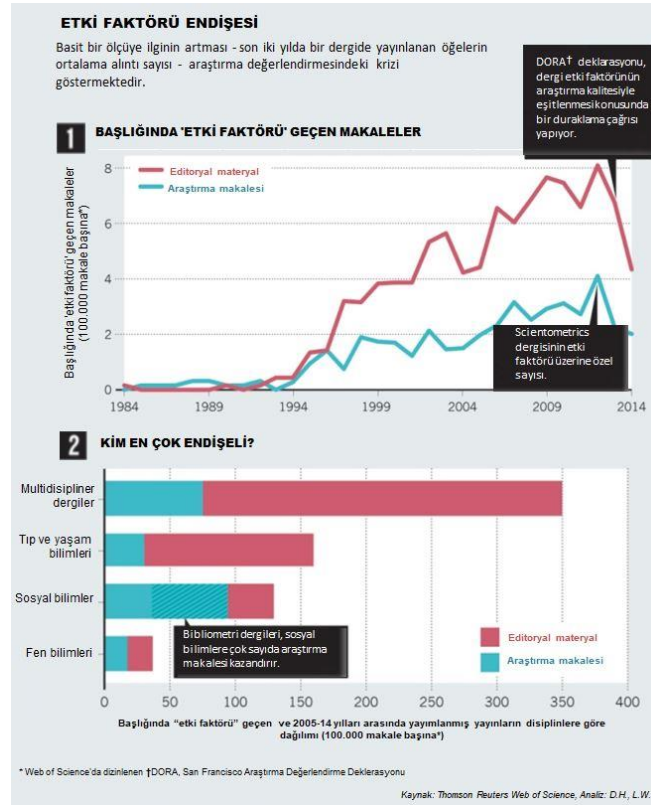
Bazı işverenler, adaylardan h-indeksi değerlerini talep etmektedir. Birçok üniversite terfi kararlarında eşik h-indeksi değerlerini ve etki faktörü yüksek dergilerdeki makalelerin sayısını temel almaktadır. Özellikle biyotıp alanında araştırmacılar özgeçmişlerinde bu puanlarla övünme fırsatı bulmuşlardır. Her yerde, danışmanlar doktora öğrencilerinden hazır olmadıkları halde yüksek etki faktörlü dergilerde yayın yapmalarını ve dış finansman sağlamalarını talep etmektedir.

İskandinav ülkeleri ve Çin'de, bazı üniversiteler sayısal bir değere bağlı olarak araştırma fonu veya teşvik ayırmaktadırlar. Örneğin; 'performans kaynaklarını' ayırmak için bireysel etki puanları hesaplanmakta veya araştırmacılara etki faktörü 15'ten yüksek olan bir dergide yayın için teşvik verilmektedir (ref. 2).

Birçok durumda, araştırmacılar ve değerlendiriciler hâlâ dengeli karar verebilmektedirler. Yine de araştırma ölçevlerinin kötüye kullanımı, görmezden gelinemeyecek kadar yaygınlaşmıştır.

Bu nedenle, konferanstan (19. International Conference on Science and Technology Indicators /19. Uluslararası Bilim ve Teknoloji Göstergeleri Konferansı) sonra ismi kesinleşen Leiden Manifestosu'nu sunmaktayız (bkz. <http://sti2014.cwts.nl>). Manifestoda yer alan 10 ilke bilimetre alanında çalışan uzmanlar için yeni değildir fakat düzenleme eksikliğinden hiçbirimizin bu ilkeleri ezbere anlatması mümkün olmayacaktır. ISI'nın kurucusu Eugene Garfield gibi alanda önde gelen kişiler, bu ilkelerin bazılarını belirtmişlerdir^{3,4}. Ancak, değerlendiriciler ilgili metodolojide uzman olmayan üniversite yöneticilerine rapor verirken bu alanın öncüleri orada yer almamaktadır. Bir değerlendirme ile (bilimsel) rekabete dahil olmak amacıyla literatür araştırması yapan bilim insanları, aradıkları materyalin erişim sorunu yaşadıkları ve az tanınan (bilinmeyen) dergilere dağılmış olduğunu görürler.

Burada ölçev tabanlı araştırma değerlendirmesinde en iyi uygulamanın özünü sunmaktayız. Böylece araştırmacılar değerlendiriciyi ve değerlendirme yapan kişiler de göstergeleri hesaba katabilirler.



Şekil: 'Etki faktörü endişesi'

10 ilke

1) Nicel değerlendirme nitel ve uzman değerlendirmesini desteklemelidir. Nicel ölçümler akran/hakem değerlendirmesinde önyargılı eğilimlere engel olabilir ve tedbirli olmayı olanaklı kılar. Bu, akran/hakem değerlendirmesini güçlendirmelidir; çünkü meslektaşları hakkında karar vermek, bir dizi ilgili bilgi olmaksızın zordur. Bununla birlikte, değerlendirme yapan kişiler karar vermeyi sayılara bırakma eğiliminde olmamalıdır. Göstergeler bilinçli karar yerine geçmemelidir. Herkes kendi değerlendirmesinden sorumludur.

2) Kurum, grup veya araştırmacının araştırma misyonuna karşı performansını ölçün. Başlangıçta programın hedefleri belirtilmeli ve performansı değerlendirmek için kullanılan göstergeler bu hedefler ile açıkça ilgili olmalıdır. Göstergelerin seçimi ve bunların kullanım şekilleri daha geniş sosyo-ekonomik ve kültürel bağlamda dikkate alınmalıdır. Bilim insanlarının araştırma misyonları çok çeşitlidir. Akademik bilginin sınırlarını ileriye taşıyan araştırmalar, toplumsal sorunlara çözüm getirmeye odaklanan araştırmalardan farklıdır. Araştırma, akademik mükemmellik fikirlerinden ziyade politik, endüstriyel ya da kamusal esaslara dayanarak yapılabilir. Tek değerlendirme modeli tüm şartlarda uygulanamaz.

3) Bölgesel araştırmalarda üstünlüğü koruyun. Dünyanın birçok yerinde, araştırma üstünlüğü İngilizce yayımla eşdeğer tutulur. Örneğin İspanya hukuku, İspanyol akademisyenlerin yüksek etki faktörüne sahip dergilerde yayın yapma isteğini beyan eder. Etki faktörü, ABD merkezli olan, hâlâ çoğunlukla WoS'ta dizinlenen İngilizce dergiler için hesaplanır. Bu önyargılar, özellikle araştırmacının daha bölgesel ve ulusal olduğu sosyal bilimler ve insani bilimler için daha çok

sorunludur. Örneğin, Sahra altı Afrika'da HIV epidemiyolojisi gibi birçok alanda ulusal veya bölgesel bir boyut vardır.

Bu çoğulculuk ve toplumsal alaka politikası, yüksek etkisi olan İngilizce dergilerin ilgisini çekebilecek türden makalelerin yayımlanması adına hasıraltı edilmektedir. WoS'ta yüksek oranda atıf yapılan İspanyol sosyologlar soyut modeller veya ABD verileri üzerinde çalışmışlardır. Yüksek etkiye sahip İspanyolca makalelerde sosyologların yerel iş hukuku, yaşlılar için aile sağlığı bakımı veya göçmen istihdamı gibi özgül konular göz ardı edilir⁵. İngilizce dışındaki yüksek nitelikli literatüre dayanan ölçevler, bölgesel araştırmalardaki üstünlüğü tanımlamaya ve ödüllendirmeye hizmet edecektir.

4) Veri toplama ve analitik süreçleri açık, şeffaf ve basit tutun. Değerlendirme için gerekli olan veri tabanlarının oluşturulması, araştırma tamamlanmadan önce kuralların açıkça tanımlanmasını gerektirir. Bu, birkaç on yıldır bibliyometrik değerlendirme metodolojisi inşa eden akademik ve ticari gruplar arasında yaygın bir uygulamadır. İlgili gruplar, hakemli literatürde yayımlanan protokolleri referans almıştır. Bu şeffaflık, araştırmaların daha dikkatli yapılmasına olanak sağlamıştır. Örneğin, 2010 yılında, gruplarımızdan biri tarafından kullanılan önemli bir göstergenin teknik özellikleri hususundaki kamuoyu tartışması (Hollanda'daki Leiden Üniversitesi Bilim ve Teknoloji Çalışmaları Merkezi) bu göstergenin hesaplanmasında bir revizyona gidilmesine olanak sağlamıştır⁶. Son ticari girişimler aynı standartlara tabi tutulmalıdır; hiç kimse bir kara kutu değerlendirme makinesini kabul etmemelidir.

Sadelik, bir göstergede erdemdir; çünkü şeffaflığı arttırır. Ancak basit ölçevler verileri çarpıtabilir (bkz. [ilke 7](#)). Değerlendirme yapan kişiler, araştırma sürecinin karmaşıklığı durumunda, basit göstergelerle dengeyi sağlamak için çabalamalıdır.

5) Değerlendirilenlerin veri ve analizi doğrulamasını sağlayın. Veri kalitesini sağlamak için, bibliyometrik çalışmalarda yer alan tüm araştırmacılar, sonuçlarının doğru bir şekilde tanımlandığını kontrol edebilmelidir. Değerlendirme süreçlerini yönlendiren ve yöneten herkes, kendi kendini doğrulama veya üçüncü kişi denetimi aracılığıyla veri doğruluğunu sağlamalıdır. Üniversiteler bunu kendi araştırma bilgi sistemlerinde uygulayabilirler ve bu tür sistem sağlayıcılarının seçiminde rehber kaynak olmalıdırlar. Doğru ve nitelikli veri zaman, maliyet ve süreç gerektirir. Bunun için bütçe ayrılmalıdır.

6) Yayın ve atıf uygulamalarında alanlara göre farklılıkların hesaplanması. En iyi uygulama, olası bir göstergeler grubunu seçmek ve disiplinlerin bunlar arasında seçim yapmasını olanaklı kılmaktır. Birkaç yıl önce, Avrupalı bir grup tarihçi, WoS tarafından dizinlenen dergilerde makale yayımlamaktan ziyade kitap yazdıkları için ulusal hakem değerlendirmesinde nispeten düşük bir puan almıştır. Bu tarihçiler bir psikoloji bölümünün parçası olmanın talihsizliğini yaşamışlardır. Tarihçiler ve sosyal bilimciler, yayın sayılarına kitap ve ulusal dilde üretilmiş literatürün, bilgisayar bilimcileri ise konferans bildirilerinin hesaplamaya dahil edilmesine ihtiyaç duyarlar.

Atıf oranları bilimsel alanlara göre farklılık gösterir. Matematik alanında en üst sıradaki dergilerin yaklaşık 3; hücre biyolojisinde ise üst sıradaki dergilerin yaklaşık 30 etki faktörü vardır. Normalleştirilmiş göstergelere ihtiyaç duyulmaktadır ve en sağlam normalleştirme yöntemi yüzdelikler bazındadır: her bir yayın, ait olduğu disiplinin atıf dağılımının yüzdelik

dilimleri temelinde değerlendirilir (örneğin en üst %1, %10 veya %20). Bir üniversitenin, yüzdelerle göstergelere dayanan sıralamadaki yerini, yüksek atıflı tek bir yayın belli belirsiz değiştirebilmektedir. Ancak, üniversiteyi atıf ortalamalarına göre oluşturulmuş bir sıralamada ortaldan en yukarı taşıyabilir⁷.

7) Bireysel araştırmacıların yayınlarının nitel değerlendirmesi hakkında temel değerlendirme. Yaşı ilerledikçe bir araştırmacı, yeni yayın çıkarmasa dahi, h-indeksi artar. H-indeksi alanlara göre değişir: yaşam bilimcileri 200'de zirveye çıkar; fizikçiler 100'de ve sosyal bilimciler de 20-30'da (ref. 8). Bu durum veri tabanından veri tabanına değişiklik gösterir: Bilgisayar bilimlerinde, WoS'ta 10 civarında bir h-indeksi olan ancak Google Scholar⁹'da 20-30 h-indeksi olan araştırmacılar yer almaktadır. Bir araştırmacının çalışmasını okumak ve incelemek bir sayıya güvenmekten çok daha uygundur. Çok sayıda araştırmacıyı karşılaştırırken bile, bireyin uzmanlığı, deneyimi, faaliyetleri ve etkisi hakkında daha fazla bilgi edinen bir yaklaşım her zaman en iyisidir.

8) Belirsizlikten ve hatalı kesinlikten kaçın. Bilim ve teknoloji göstergeleri, tereddüt ve kavramsal belirsizliğe eğilimlidir ve evrensel olarak kabul edilmeyen güçlü varsayımlar gerektirir. Örneğin, atıf hesaplamanın anlamı uzun süredir üzerinde tartışılan bir husustur. Bu nedenle, en iyi uygulamada daha sağlam ve çoğulcu bir tablo ortaya koymak için çoklu göstergeler kullanılır. Belirsizlik ve hata ölçülebilirse, örneğin hata çubukları (error bars) kullanılarak, bu bilginin yayınlanmış gösterge değerlerine eşlik etmesini sağlamalıdır. Bu durum mümkün değilse, gösterge üreticileri en azından hatalı kesinliği önlemelidirler. Örneğin, dergi etki faktörü yayınlanırken sayıların eşit olmasını önlemek için üç ondalık basamak halinde verilmektedir. Ayrıca, kavramsal belirsizlik ve atıf hesaplamanın rastgele değişkenliği göz önüne alındığında, çok küçük etki faktörü farklılıkları temelinde dergiler arasında ayırım yapmanın bir anlamı olmaz. Hatalı kesinlikten kaçın: sadece bir ondalık fark kazandırır.

9) Değerlendirme ve göstergelerin sistemik* etkilerinin farkına varın. Göstergeler, kurdukları teşviklerle sistemi değiştirir. Bu etkiler beklenmelidir. Bu, bir dizi göstergenin daima tercih edilir olduğu anlamına gelir — tek bir tanesi oyun ve hedefin şaşmasına davetiyedir (bu durumda ölçüm hedef haline gelir). Örneğin, 1990'larda, Avustralya, bir enstitünün yayınladığı yayın sayısına dayanan bir formül kullanarak üniversite araştırmalarını finanse etmiştir. Üniversiteler, hakemli bir dergideki bir makalenin 'değerini' hesaplayabilirler; 2000 yılında araştırma finansmanı için 800 Avustralya doları (2000 yılında yaklaşık 480 ABD Doları) ayrılmıştır. Tahmin edilebileceği gibi, Avustralyalı araştırmacılar tarafından yayımlanan yayın sayısında artış olmasına rağmen, makaleler az atıf almış dergilerde yayımlandığı için makalelerin niteliğinin düşük olduğu izlenimini uyandırmıştır¹⁰.

10) Göstergeleri düzenli olarak dikkatlice inceleyin ve güncelleyin. Araştırmaların amaçları ve değerlendirme hedefleri ile araştırma sisteminin kendisi birlikte gelişmektedir. Yararlı olan ölçümler belli bir zaman sonra yetersiz olabilmekte ve yeni ölçümler ortaya çıkabilmektedir. Gösterge sistemleri gözden geçirilmeli ve hatta belki de değiştirilmelidir. 2010 yılında basit kurallar sistemini fark eden Avustralya, kaliteyi vurgulayan, daha karmaşık "Excellence in

* "Sözlük anlamı: Sayısal ve ekonomik konulara belli sistemler çerçevesinde bakan. Kaynak: Türk Dil Kurumu Güncel Türkçe Sözlük, erişim linki: <http://sozluk.gov.tr/>

Research for Australia initiative (ERA) / Avustralya için Araştırmada Mükemmellik inisiyatifi”ni başlatmıştır.

Sonraki adımlar

Araştırma değerlendirmesi bu on ilkeye uyulmasıyla, bilimin gelişmesinde ve toplumla etkileşiminde önemli bir rol oynayabilir. Araştırma ölçevleri, bireysel uzmanlık yoluyla toplanması veya anlaşılması zor olabilecek önemli bilgiler sağlayabilir. Ancak bu nicel bilginin bir araştırmadan hedefe dönüşmesine izin verilmemelidir.

En iyi kararlar, sağlam istatistiklerin, değerlendirilen araştırmanın amacı ve doğasına olan duyarlılıkla birleştirilmesiyle elde edilir. Hem nicel hem de nitel kanıtlara ihtiyaç vardır; her biri kendine göre objektiftir. Bilim ile ilgili karar verme, en yüksek kalitede veri ile donanmış yüksek kaliteli süreçlere dayalı olmalıdır.

Çeviri: Arş. Gör. Demet Işık, Ankara Üniversitesi Bilgi ve Belge Yönetimi Bölümü, Türkiye

Prof. Dr. Özlem Gökkurt Demirtel, Ankara Üniversitesi Bilgi ve Belge Yönetimi Bölümü, Türkiye

Kaynaklar

1. Wouters, P. in *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (eds Cronin, B. & Sugimoto, C.) 47–66 (MIT Press, 2014).
2. Shao, J. & Shen, H. *Learned Publ.* 24, 95–97 (2011).
3. Seglen, P. O. *Br. Med. J.* 314, 498–502 (1997).
4. Garfield, E. *J. Am. Med. Assoc.* 295, 90–93 (2006).
5. López Piñero, C. & Hicks, D. *Res. Eval.* 24, 78–89 (2015).
6. van Raan, A. F. J., van Leeuwen, T. N., Visser, M. S., van Eck, N. J. & Waltman, L. *J. Informetrics* 4, 431–435 (2010).
7. Waltman, L. et al. *J. Am. Soc. Inf. Sci. Technol.* 63, 2419–2432 (2012).
8. Hirsch, J. E. *Proc. Natl Acad. Sci. USA* 102, 16569–16572 (2005).
9. Bar-Ilan, J. *Scientometrics* 74, 257–271 (2008).
10. Butler, L. *Res. Policy* 32, 143–155 (2003).