

Manifestul de la Leiden: un simptom și un îndemn (The Leiden manifesto: a symptom and an impulse)

RADU SILAGHI-DUMITRESCU

Babeș-Bolyai University, str. M. Kogălniceanu nr. 1, Cluj-Napoca 400028, Romania

The Leiden manifesto, recently published by Diana Hicks, Paul Wouters, Ludo Waltman, Sarah de Rijcke, and Ismael Rafolsin in the journal *Nature*, proposes ten „commandments” regarding the use of bibliometrics in evaluating scientific performance. Commented here is the current state of the scientific performance evaluation systems in Romania with respect to the principles laid out in the Leiden Manifesto – pointing out previous efforts complying with its „commandments” and arguing that such trends should be further and more widely supported by the Romanian academic community.

Keywords:

Diana Hicks (profesor de politici publice la Georgia Institute of Technology, Atlanta, Georgia, SUA), Paul Wouters (profesor de scientometrie și director al Centre for Science and Technology Studies, Leiden University, Olanda), Ludo Waltman (cercetător la aceeași universitate), Sarah de Rijcke (lector la aceeași universitate) și Ismael Rafols (cercetător în politica științei la Consiliul Național al Cercetării din Spania și la Universitatea Politehnică din Valencia) au publicat, recent, ca o concluzie a unei conferințe desfășurate la Leiden și pe fundalul unei comunități foarte active în domeniu, un articol în revista *Nature*, intitulat „Manifestul de la Leiden” („The Leiden Manifesto for research metrics”) [1]. Acesta propune zece principii de urmat în evaluarea cercetării științifice, toate menite a optimiza utilizarea (evitând abuzarea, dar și subestimarea) metodelor scientometrice.

Manifestul începe cu un preambul în care se subliniază rolul în creștere pe care îl joacă datele numerice în evaluări, pornind de la pionieratul bazei de date ISI dinainte de era internetului, trecând apoi la sistemele online create de Thomson Reuters (2002) și apoi de principalii competitori (Elsevier și Google, 2004), la sistemele cu algoritmi mult mai complecși precum InCites (Thomson-Reuters) și Sci-Val (Scopus), și numeroasele alte variante (dintre care noi am menționa inițiativa românească Epistemo a lui Răzvan Florian), unele bazate inclusiv pe medii sociale. În toate acele eforturi, două noțiuni par să transpară ca printre cele mai populare, uneori în sine și alteori ca punct de plecare pentru algoritmi mai evaluați: factorul de impact (din 1995 încoace) și indicele h (propus în 2005).

Autorii subliniază, în calitate de experți în scientometrie, în științe sociale și în administrarea activității de cercetare, un profund dezacord cu numeroase cazuri de aplicare greșită a indicatorilor scientometrici. Ei citează ca astfel de exemple negative:

1. clasamente internaționale ale universităților, precum Shanghai – ARWU sau Times Higher Education, bazate pe criterii incomplete. La această critică, mediul universitar românesc va rezona semnificativ. Universități de prim rang, precum „Babeș-Bolyai” și-au fixat, la un moment sau altul, ținte strategice legate de astfel de clasamente – ca de exemplu intrarea între primele 500 de universități din lume din clasamentul ARWU/Shanghai. Societatea rezonază și ea la astfel de clasamente, dar există și critici ai acestei atitudini, în sens apropiat celei din Manifest [2];

2. folosirea indicelui h drept criteriu la angajare;

3. folosirea unor „standarde minimale” în privința indicelui h și a numărului de articole în reviste de „mare impact” pentru a lua decizii asupra promovării pe post. La prima vedere, această critică poate pune metodologiile CNATDCU direct sub reflector. Destui membri ai sistemului academic românesc privesc lucrurile în mod simplist: îndeplinirea criteriilor CNATDCU atrage angajarea sau promovarea pe post (într-adevăr, „promovare”; conceptul nu există în lege, deși el există peste tot în lume – și până când va fi recunoscut ca atare este oricum aplicat la noi *de facto*, la scară aproape universală, în forme distorsionate și alambicate). Este, în fapt, rolul comisiilor de concurs sau de

examinare să adauge componenta calitativă, de „peer-review” (deci tocmai cum cere și Manifestul) – pentru că, până la urmă, CNATDCU nu oferă titluri, și nici nu pretinde că simpla îndeplinire a criteriilor „minimale” ar trebui să fie un pașaport „gata vizat” pentru obținerea unui titlu;

4. abuzarea (după principiul cunoscut în anumite medii din România drept „fără număr”) în privința afișării unor astfel de date (lucrări în reviste de „mare impact”, *h*-uri deosebite) în CV-urile cercetătorilor – mai ales, în domeniul biomedical, domeniu favorizat față de altele prin numărul mai mare de reviste, ceea ce duce la numere mai mari de citări, de unde și factori de impact mai mari și indici *h* mai mari.

5. succesul pe care cercetătorii prea puțin experimentați îl au în obținerea de finanțare doar pe criteriul că au avut, chiar ca subordonați/doctoranzi, lucrări de „mare impact”. Aici, România, ca parte a unor necesare eforturi de reformă și de recunoaștere diferențiată a meritelor pe criterii calitative, a îmbrățișat, uneori, această viziune greșită. În diverse proceduri de evaluare, calitatea de prim-autor, rezervată, de multe ori, *executantului*, care a efectuat cea mai mare parte a muncii pentru acel articol, a fost pusă pe picior de egalitate cu calitatea de autor corespondent, rezervată, de obicei, autorului de concepție, care a gândit și dirijat experimentele și, adesea, a asigurat și fondurile ca *lider* al unui grup de cercetători;

6. alocarea de fonduri sau premii doar pe criterii numerice simpliste, precum factorul de impact individual, sau numărul de lucrări publicate în reviste cu un factor de impact peste o anumită valoare. Aici, România are exemple asupra cărora poate reflecta: sistemul de „premiere a rezultatelor cercetării” al UEFISCDI pare a se înscrie, perfect, în vizorul acestei critici din Manifestul de la Leiden. Dacă ea este asumată, corolarul ar fi un sistem de redirecționare a unor astfel de fonduri către granturi și burse acordate prin „peer-review” (*informat* însă neapărat de date scientometrice, după cum se va argumenta și mai jos).

În acest context, autorii Manifestului propun zece principii de urmat

1. „**Evaluările cantitative trebuie să susțină evaluări calitative efectuate de experți în domeniu**”. Astfel, datele scientometrice pot fi foarte utile în evitarea unor prejudecăți locale sau accidentale, inclusiv sau, mai ales, atunci când se fac comparații între colegi. Este esențial, însă, ca evaluatorii să își păstreze dreptul și responsabilitatea

deciziei, în calitate de experți, și să nu existe dubii că acea decizie ar fi fost rodul rece al unui algoritm numeric inerent incomplet. Pentru România, am citat mai sus exemplul criteriilor numerice ale CNATDCU și al modului în care ele sunt, potrivit regulamentului, doar unul dintre argumentele pe care comisiile de concurs le iau în calcul. Din acest punct de vedere, Manifestul de la Leiden poate fi privit drept un apel către comisii de a insista pe dreptul și responsabilitatea, pe care le au, de a servi drept mai mult decât un oficiu secretarial, care verifică și ordonează „numărul de...”. Exercițiul recent de evaluare propus de către CNFIS,[3] ca și versiunile precedente ale algoritmilor de finanțare a universităților, sunt contraexemple pentru acest punct: ele propun a colecta date numerice, fără a adăuga componenta de evaluare de către experți. Prin comparație, sistemul de clasificare calitativă a universităților, conform art. 193 din Legea Educației Naționale nr. 1/2011 și HG 789/2011 (centrate pe educație, de educație și cercetare științifică, de educație și creație artistică și de cercetare avansată și educație),[4] pare mai apropiat de principiul 1 al manifestului de la Leiden.

2. „**În măsurarea performanței, natura misiunii de cercetare a instituției/grupului/cercetătorului trebuie să fie punctul de referință**”. Trebuie, așadar, identificat ce se dorește de la entitatea evaluată, care sunt obiectivele care i-au fost menite, și abia apoi să se aleagă, pe acele baze, indicatorii de performanță. Acești indicatori nu au cum să fie universalii atât pentru cercetarea fundamentală de la frontiera cunoașterii, cât și pentru cea aplicativă cu impact social imediat. În unele cazuri, se poate vorbi de principiile excelenței academice, în altele - despre efectul asupra unor politici, asupra industriei sau asupra publicului. În mediul academic românesc aceste argumente nu sunt noi. Din păcate, ele au stat și la baza unor false argumente. Astfel, atunci când cu justificare, se respinge utilizarea factorilor de impact, sau chiar a bazelor ISI și Scopus pentru anumite domenii, se propun la schimb uneori paliative fără substanță. În arii, în care multe reviste *din afara* ISI sunt într-adevăr cele mai de calitate din domeniu, există tendința de a apela, în aceeași măsură, la aproape orice altă revistă. Acolo unde cărțile ar prima în fața revistelor, există tendința de a trece prea departe în planul secundar calitatea acelei cărți. În domenii în care cercetarea aplicată și interacțiunea cu mediul privat sunt propuse ca indicatori, se înregistrează brevete de invenție niciodată aplicate și se înființează entități private (firme, ONG-uri, asociații) al căror singur scop este accesarea de fonduri destinate cercetării – mimând doar caracterul de inițiativă privată reală pe piața comercială sau socială.

3. **„Protejarea excelenței în cercetarea de relevanță locală”**. Revistele care au la nivel internațional cel mai mare impact, tind să fie editate în limba engleză și de edituri din zona anglofonă – în special SUA. Aceasta poate aduce un accent exagerat pe problemele relevante pentru acea zonă culturală și socială. Se poate ajunge, atunci, la situații în care studiul unor probleme legate de cultura anglo-saxonă, sau al unor probleme de relevanță pentru regiuni îndepărtate, să primească *artefactual* mai multă recunoaștere pe criterii bibliometrice decât studii de impact real și imediat asupra propriei națiuni, propriei comunități, propriei regiuni. Desigur, aceasta nu înseamnă negarea excelenței din cercetarea de relevanță internațională; dimpotrivă, în Manifest se face un apel la construirea de metrici bazate pe literatura de specialitate „de mare impact” publicată în alte limbi decât cea engleză. În România, una dintre modalitățile de a rezolva astfel de probleme a fost „echivalarea”, după diverși algoritmi, a unor publicații locale cu unele de circulație internațională. De asemenea, a existat exercițiul evaluării editurilor și revistelor în categorii calitative (A, B, C...) de către CNCS.[5]

4. **„Procesul de analiză și colectare a datelor trebuie să fie deschis, transparent și simplu”**. Bazele de date trebuie construite după principii și mecanisme anunțate clar, evitându-se variantele care, sub justificarea secretului comercial sau a proprietății intelectuale, nu fac publici algoritmi și mecanismele, furnizând doar clasamente sau scoruri ce nu pot fi supuse verificării independente. Se poate merge chiar până la a cere simplitate în indicatori; deși, după cum se arată mai jos, excesul de simplitate este, de asemenea, periculos.

5. **„Cei evaluați trebuie să aibă ocazia de a verifica datele și analizele”**. Dacă etapa de verificare de către cei evaluați sau de către terți independenți poate costa timp și bani, construirea de baze de date de încredere, utile pentru evaluări, implică prea multă răspundere pentru a nu merita bugete pe măsură. Procedurile de evaluare din România fac, într-adevăr, adesea, uz de acest principiu. În fapt, *ele fac chiar abuz de el*. Cercetătorii din România nu doar că au posibilitatea de a verifica datele, ci, prea adesea, au răspunderea de a le *aduna ei înșiși*. Prea rar, fac autoritățile evaluatoare efortul de a pre-colecta informațiile publice disponibile din baze de date, și prea des, ceea ce s-a colectat o dată este în format complet incompatibil, din punct de vedere tehnic, cu ceea ce se colectează după câteva luni, pentru alte evaluări (de la cea a universităților, la cele interne departamentale, cele pentru diverse premii, granturi, reforme, concursuri etc).

6. **„Să se țină cont de variabilitatea între domenii în privința practicilor de publicare și de citare”**. Se recomandă, în acest context, să se pună la dispoziție o gamă de indicatori, dintre care fiecare domeniu să îl poată alege pe cel potrivit. Spre exemplu, în informatică este nevoie să se ia în considerare, cu precădere, articolele publicate în volume de conferință („proceedings”), în istorie și științe sociale trebuie luată în considerare, cu precădere, literatura de specialitate în limba națională respectivă (și cu precădere cărțile înaintea articolelor de revistă), în timp ce în alte domenii, lucrările de la conferințe și cărțile în limbi de mică circulație nu mai reprezintă contribuții științifice importante. De asemenea, dacă ne referim la reviste, rata de citare este foarte diferită între domenii, uneori și cu factori de câte un ordin de mărime. Sunt esențiali, atunci indicatorii normalizați la domeniu, urmărindu-se nu numere absolute ci, de exemplu, plasarea între primele x% reviste *din domeniu*. România aplică acest principiu, de câțiva ani, prin folosirea categoriilor roșu-galben-gri pentru reviste și prin utilizarea acestor indicatori pentru evaluări, cum este cea CNFIS, cum au fost standardele minimale impuse de către CNCS sau CNATDCU, dar și altele de natură mai locală.

7. **„Evaluările individuale trebuie bazate pe aprecierea calitativă a portofoliului cercetătorului”**. Spre exemplu, acordarea de calificative doar pe baza indicelui *h*, este greșită din mai multe puncte de vedere. Astfel, acest indice crește inevitabil în timp, odată cu acumularea citărilor - indiferent ce sau cât mai publică autorul. Prin urmare, clasarea după *h* va da, artefactual, întâietate cercetătorilor mai vârstnici, în fața unor colegi mai tineri, care s-ar putea să aibă realizări calitativ similare sau chiar ușor superioare. De asemenea, *h* variază dramatic între domenii: în timp ce valorile maxime tind către 200 în științele vieții sau 100 în fizică, ele ajung la cel mult 20-30 în științele sociale. De asemenea, *h* poate varia semnificativ între bazele tradiționale ca ISI și Scopus (care tind să dea rezultate similare una alteia) și alternative, precum Google. Autorii Manifestului argumentează, așadar, că *nu există alternativă serioasă la lecturarea și evaluarea rezultatelor și publicațiilor de către experți*.

8. **„Trebuie evitată falsa precizie numerică”**. Mulți indicatori scientometrici sunt bazați pe aproximări sau teorii general, *dar nu universal* acceptate. Ele au inerent marje de eroare, care ar trebui citate, sau, unde este posibil, prezentate chiar ca „bare de eroare”. Aceste marje trebuie recunoscute, ca atare, evitând situații de departajare după zecimale, acolo unde, de fapt, cifrele semnificative se opriseră cu mult înainte de a se

ajunge la virgulă. Spre exemplu, cele trei zecimale din valoarea factorului de impact nu au, neapărat, o semnificație universal acceptată. De asemenea, tendința, vizibilă în România, de a ordona cercetători după valori relativ mici ale indicilor h (și mici sunt în general aceste valori la noi față de standardele internaționale [6]), este de descurajat. Autorii Manifestului, în fapt, argumentează pentru luarea în considerare, la evaluări, a unor valori ale factorilor de impact rotunjite la o zecimală. De asemenea, este propusă folosirea, în paralel, de mai mulți indicatori, astfel încât, acolo unde unul riscă să inducă în eroare, altele să poată repune lucrurile în ordine.

9. **„Înțelegerea efectelor pe termen lung ale evaluării și indicatorilor”**. Indicatorii de performanță vor fi inerent luați în considerare de mulți cercetători, atunci când își stabilesc livrabilele și obiectivele profesionale. Un set de indicatori prea simplist duce, invariabil, la distorsionarea sistemului. Dacă se evaluează doar numărul de lucrări (nu și calitatea), se ajunge la scăderea numărului de lucrări de calitate superioară, în beneficiul celor de calitate mai slabă și se va extinde practica acordării de calitate de coautor pe motive administrative. Dacă se evaluează doar numărul de citări, ne vom autocita fără justificare; dacă se exclud autocitările, se ajunge la caste sau grupuri create în scopul citării reciproce. Dacă evaluăm numărul de pagini, se va mări, artificial, acest număr modificând fontul. România a văzut, în anii recentți, astfel de fenomene, în care, din cauza abuzului termenul BDI (baze de date internaționale) sau chiar ISI, sunt privite, de o parte a cercetătorilor, ca expresii peiorative, din cauza revistelor nou-introduse în acele categorii și care în sisteme de evaluare prea simpliste ajung, stând la coada clasamentelor bazate pe impact, să aibă statut egal cu revistele de impact *real* indexate în aceleași baze de date. Este, prin urmare, de urmărit un set de criterii suficient de diverse, încât să recompenseze calitatea fără a oferi posibilitatea unor ocolișuri ieftine către un calificativ bun. *Acea excelență care poate fi identificată doar din algoritmi obscuri sau înguști nu mai este excelență, ci obscurantism și îngustime.*

10. **„Indicatorii înșiși trebuie examinați și actualizați periodic”**. Unii indicatori își vor pierde, în timp, actualitatea, alții vor fi definiți pentru a acoperi tendințe, metodologii și tehnici mai noi. Se poate cita aici „numărul de accesări/citiri/descărcări de pe internet” al unui articol – un element metric ce nu ar fi avut sens acum 30 de ani, dar este, azi, luat în considerare de unele platforme online.

Autorii Manifestului conclud prin a argumenta că o evaluare de calitate trebuie să ia în considerare atât aspectele cantitative, cât și aspectele calitative. Aici rămâne rolul esențial al scientometriei de a încerca să

asigure statistici robuste și date de calitate, fără de care evaluările corecte și eficiente nu sunt posibile.

Așa cum s-a arătat mai sus, *Manifestul de la Leiden prezintă principii dintre care multe sunt deja aplicate în România, ca și în alte părți. Îl putem privi ca pe un simptom al lucrurilor bine făcute până acum, și un îndemn pentru a le continua.* Ca posibil studiu de caz, Universitatea „Babeș-Bolyai” a propus, în anii recentți, un model de evaluare, în care cercetătorii sunt plasați în patru clase de valoare, în baza impactului rezultatelor, descris în detaliu în altă parte.[7] Pe scurt, acest sistem face apel la criteriile scientometrice pentru *informarea* evaluatorilor (sugerându-se anumite calificative calitative pentru anumite numere de lucrări/citări/cărți, dar subliniind că este rolul evaluatorului să judece calitatea și greutatea contribuțiilor autorului – conform Principiului 1 de la Leiden), oferă criterii specifice diverselor domenii, conform Principiului 2 (se oferă posibilitatea atingerii de calificative maxime fie prin articole de impact, fie prin citări, fie prin cărți de impact, fie prin brevete, fie prin orice alte realizări de impact demonstrabil), se recompensează realizările cu specific local atâta vreme cât impactul lor în comunitatea ce le definește este demonstrat (Principiul 3), se folosesc numai baze de date accesibile public (Principiul 4), se oferă cercetătorului ocazia de a identifica singur itemii valoroși și de a contesta eventuala lor ignorare de către evaluator (Principiul 5), permite variabilitatea practicilor de publicare, conform Principiului 6 (folosește în privința revistelor clasamentul roșu-galben-gri al UEFISCDI, permite, dar nu limitează accesul la punctaje bune pe baza unor numere mari de citări sau lucrări, ia în considerare cărțile de impact, brevetele etc), insistă pe evaluarea calitativă conform Principiului 7 (permite acordarea de calificative maxime, chiar și acolo unde lipsesc aparent indicatorii scientometrici ai excelenței, recompensează, preferențial, calitatea de autor corespondent ca inițiator al inovării și cercetării), acordă un set restrâns de calificative calitative recunoscând necesitatea de a se evita departajările la virgulă (Principiul 8). Evident, că astfel de proceduri sunt inerent imperfecte, în liniile punctate de Principiile 9 și 10. Spre exemplu, ca răspuns la tratamentul preferențial acordat autorilor corespondenți, unele grupuri mai mari de cercetare tind să acorde acest statut acelor membri, care ar beneficia cel mai mult de el și nu celor care îl merită după standardele academice comune - așa cum, pentru motive similare, poate exista practica de a coopta un autor la o publicație, sau de a acorda un statut de prim autor etc. Evaluarea calitativă efectuată de experți din domeniu poate, în bună parte, identifica

astfel de derapaje; combaterea ar fi mai facilă, dacă s-ar realiza, mai clar, că ele pot fi considerate a fi elemente de fraudă academică.

Referințe

- [1] D. Hicks, P. Wouters, L. Waltman, S. de Rijcke, I. Rafols, *Bibliometrics: The Leiden Manifesto for research metrics*, *Nature* 520 (2015) 429-31.
- [2] http://adevarul.ro/educatie/universitar/interviu-acadioan-aurel-pop--rectorul-universitatii-babes-bolyai-trebuie-dispara-ideea-profesorul-estapan-1_5538f5cdcfbe376e357aa6fc/index.html.
- [3] <http://www.cnfis.ro/raportare-ic2015/>.
- [4] <http://chestionar.uefiscdi.ro/>.
- [5] <http://www.cnes-nrc.ro/>, .
- [6] R. Silaghi-Dumitrescu, A. Sabau, *Scientometric analysis of relative performance in a key University in Romania*, *Scientometrics* 99 (2014) 463-474.
- [7] R. Silaghi-Dumitrescu, submitted manuscript (2015).

Autor corespondent: rsilaghi@chem.ubbcluj.ro