

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Open Science è una necessità, non una noia burocratica

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1792138> since 2021-06-25T18:28:03Z

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

Open Science è una necessità, non una noia burocratica

Se c'è un aspetto positivo di questa pandemia causata dal Coronavirus è aver reso evidente a tutti la necessità inderogabile di avere accesso immediato ai risultati e ai dati della ricerca scientifica. Triste constatarlo, ma i meccanismi attuali della comunicazione scientifica fanno ben altri interessi che non quelli del progresso della scienza.

La parola d'ordine è "Publish the evidence and embrace open research" (come si legge su [Nature](#) il 17 marzo) Sempre [Nature](#) (paradossalmente, nella stessa pagina, si legge l'invito "Subscribe") aveva già invitato a febbraio: "continue a condividere, restate aperti". Gli esperti scientifici degli Stati Uniti e di altri 11 paesi hanno firmato a inizio marzo 2020 [un appello](#) per la condivisione dei dati e dei risultati non solo in formato pdf – ovvero quello che normalmente si ottiene in cambio di costosi abbonamenti – ma in formato leggibile alle macchine, per permettere all'Intelligenza Artificiale di aiutare in una situazione di emergenza. Gli scienziati stanno condividendo tutto rapidamente (come risulta da questo [articolo](#) del 1 febbraio 2020), e altrettanto rapidamente la conoscenza sul virus progredisce grazie alla condivisione dei dati quali le sequenze del genoma del virus stesso (come si legge in questo altro pezzo del [Washington Post](#)).

Ma questo viene sottolineato ed evidenziato perché non è ciò che accade normalmente nella comunicazione scientifica. Anzi, tutti rilevano l'eccezionalità della risposta alla crisi eccezionale scatenata dal Coronavirus come una deviazione, una novità rispetto alla norma.

E qual è la norma? La norma è che il sistema attuale di comunicazione basato sulle riviste scientifiche pone barriere alla conoscenza, invece di favorirne la condivisione. Quando un ricercatore termina il suo esperimento o la sua ricerca, la propone per la pubblicazione a una rivista scientifica. L'articolo, dopo aver passato il processo di revisione (peer review) a cura di due revisori anonimi scelti fra gli esperti della materia, viene pubblicato sulla rivista. Ora, né gli autori né i revisori ottengono un compenso economico per il loro lavoro. Ciò che si aspettano dalla pubblicazione è il prestigio, sono le citazioni che possono far crescere i loro indici ai fini della valutazione della ricerca. I criteri di valutazione in vigore premiano infatti le pubblicazioni su riviste con alto indice di impatto.

Il problema è che l'ente che ha finanziato la ricerca, se vuole leggere i risultati di ciò che ha già contribuito a pagare con i propri fondi, deve "ricomprare" la rivista attraverso l'abbonamento. Aggiungiamo un tassello: sia i soldi che hanno finanziato la ricerca sia i soldi usati per ricomprare i risultati sotto forma di abbonamento alle riviste sono, nel caso delle Università, soldi pubblici.

Di fatto, quindi, noi oggi stiamo spendendo fondi pubblici per chiudere i risultati delle ricerche (a loro volta finanziate con i fondi pubblici) dietro abbonamenti il cui costo è [cresciuto del 521%](#) fra il 1986 e il 2015, mentre i bilanci delle biblioteche si sono ridotti, con conseguenti tagli, per cui i ricercatori possono leggere ed essere letti di meno. Aggiungiamo un altro tassello ancora: chi non ha abbonamento non può leggere i risultati della ricerca: medici, professionisti, piccole e medie imprese, start up, nessuno può permettersi di pagare abbonamenti di migliaia di dollari per una singola rivista – e ne vengono pubblicate circa 25.000.

Ma la ricerca è un processo incrementale, e si costruisce sui risultati di altri. Se non si ha accesso a questi risultati e ai dati, come si può progredire? Non si rischia di reinventare ogni giorno la ruota, con ulteriore spreco e duplicazione di finanziamenti? E come si può fare innovazione o semplicemente curare bene il proprio paziente se non si ha accesso alle ricerche recenti?

In tutto questo nessuno, nemmeno i ricercatori che scrivono articoli all'interno delle nostre università, si chiede quanti milioni di euro si spendano per gli abbonamenti. Globalmente, nel 2016, sono [stati spesi 7,6 miliardi di dollari](#), il che significa che se dividiamo questa cifra per il numero di articoli pubblicati ogni anno [stiamo pagando dai 3800 ai 5000 dollari](#) (a seconda della banca dati utilizzata che conta 1,5 o 2 milioni di articoli) per ogni articolo, generando guadagni netti per gli editori commerciali che [sfiorano il 38%](#) (quando,

per fare un raffronto, il guadagno netto di un colosso come Google [si attesta sul 25%](#)). Ricordiamo che gli abbonamenti vengono pagati da tutti gli enti di ricerca nel mondo ogni anno, crescono ogni anno di circa il 3-5%, e tutti paghiamo per la stessa rivista (ovvero, non è che se l'Università di Torino si abbona a Nature, l'Università di Milano può leggere Nature: deve a sua volta pagare lo stesso abbonamento). Chi non ha abbonamento ed è interessato a un articolo, quando cerca di scaricarlo vede comparire un pop-up che chiede dai 30 ai 40 dollari per acquistare il diritto a leggere un solo pdf. Come può essere sostenibile tutto ciò non solo nei paesi a scarso reddito, ma anche per i laboratori di ricerca non affiliati a un ateneo o per le unità di ricerca e sviluppo delle aziende? O per i medici e i professionisti? Se questo non fosse vero, se davvero come sostengono gli editori commerciali tutti quelli cui serve hanno accesso alle risorse, il [sito pirata Sci-Hub](#) non esisterebbe. Sci-Hub è stato creato anni fa da una giovane ricercatrice – un po' hacker, diciamo – che, stufo di scontrarsi con la comparsa del famoso pop-up di richiesta dei 30 dollari ad articoli semplicemente è andata a prendere i pdf dove li ha trovati, ovvero siti di università o di editori. SciHub è un sito illegale, ma ha oltre [6 milioni di download](#) al mese da ogni parte del mondo, inclusi i paesi ben finanziati e con un potenziale accesso a tutto ciò che serve per far ricerca [un suggerimento: esiste [Unpaywall](#), che è un modo legale per accedere ai testi cui non si ha abbonamento]. Il Rettore emerito dell'Università di Liegi, [Bernard Rentier](#), notava a questo proposito: “il solo fatto che condividere la conoscenza sia diventato illegale deve farci riflettere su quanto assurdo sia diventato questo sistema”. Ora, quando esistevano solo riviste cartacee è evidente che non c'erano alternative, ma oggi, grazie al web, ci sono tutti gli strumenti e i servizi per voltare pagina, se solo lo si volesse. Il sistema resta in piedi solo perché i criteri di valutazione della ricerca continuano a perpetuare la logica perversa del prestigio della rivista in cui si pubblica.

Ma questa logica ha generato [comportamenti adattivi](#) (si pubblica solo ricerca “di moda” per essere pubblicati nelle riviste più prestigiose, si enfatizzano i risultati, si suddivide una ricerca in più articoli per gonfiare la propria produzione scientifica...) o, peggio, si falsificano e addirittura a volte si fabbricano i dati. Questo fenomeno si traduce spesso nella ritrattazione (il ritiro dalla pubblicazione): un'occhiata al blog scientifico [Retraction Watch](#) è sconcertante, perché ad esempio se si tolgono le 96 ritrattazioni di Joachim Boldt dalle metanalisi, il rischio di danni ai reni o di morte per assunzione della molecola da lui studiata aumenta esponenzialmente. E ciò che più sconcerta è la [correlazione diretta fra il numero di ritrattazioni](#) e il fattore di impatto della rivista: ovvero, più è alto il prestigio della rivista più sono alte le ritrattazioni, perché se devo fare di tutto pur di pubblicare su certe riviste, faccio letteralmente di tutto, incluso falsificare o fabbricare i dati, per cui paradossalmente [le riviste più “prestigiose” pubblicano la peggiore scienza](#), e quello che viene pubblicato oggi è la [“selezione naturale” della peggior scienza](#). Non a caso sta crescendo il numero di riviste che richiedono di pubblicare anche i dati che supportano l'articolo, per trasparenza. E non a caso l'Open Science consiste in primo luogo nel dare accesso ai dati della ricerca, non solo per il riuso ma anche per la verificabilità e riproducibilità.

Ma ci sono altri aspetti critici in questo sistema: proprio la riproducibilità, per le pratiche scorrette di cui sopra, è in crisi, in [alcune discipline per oltre il 50%](#): ma se la scienza non è riproducibile, si può ancora definire scienza? Crescono poi a dismisura le [autocitazioni - +179%](#), sempre per ragioni di valutazione, in Italia siamo diventati un caso di studio internazionale-, ma anche questo cos'ha a che fare con una sana prassi scientifica? I tempi medi di pubblicazione, in questo sistema, vanno dai [9 ai 18 mesi](#). Pensiamo all'esempio del Coronavirus: che senso avrebbe leggere un risultato fra quasi un anno? Una giovane ricercatrice, [Paola Masuzzo](#), lo scorso anno festeggiava l'uscita di un articolo a firma di suoi amici su una prestigiosa rivista dopo due anni e mezzo di revisione, felicitandosi per la loro carriera ma chiedendosi allo stesso tempo cosa ci fosse da festeggiare in termini di avanzamento della conoscenza, dopo un periodo così lungo.

Serve un nuovo modello: testi e dati e l'intero processo della ricerca aperto... in una parola, Open Science

L'emergenza Coronavirus ha fatto esplodere le contraddizioni di questo [sistema disfunzionale](#). Alcune [biblioteche](#) e [reti di biblioteche](#) stanno chiedendo agli editori di aprire i testi relativi all'infezione. Esiste [COVID-19 Open Research Dataset](#), una pagina in cui oltre ai dati e agli articoli già Open alcuni editori

commerciali hanno reso aperte (ma solo fino a quando l'emergenza durerà; lo stesso avevano fatto all'indomani dell'incidente di Fukushima sugli effetti delle radiazioni) le loro pubblicazioni sul virus.

Ma resta da chiedersi: perché solo durante un'emergenza? E perché solo i risultati e i dati sul Coronavirus? Solo perché ci tocca da vicino? E la ricerca sul cancro allora? E quella sull'Alzheimer? O sul riscaldamento globale e la crisi climatica?

Quello che serve è un radicale cambio di rotta dell'intero sistema di comunicazione scientifica, per riportare la scienza alla sua natura di "grande conversazione" e di collaborazione, invece che competizione.

Questo è ciò che si intende per Open Science, ovvero un nuovo modo aperto e partecipativo di fare ricerca, diffonderla e valutarla. Uno degli obiettivi è aumentare la trasparenza e la riproducibilità in ogni passo del processo di ricerca, attraverso la condivisione più tempestiva possibile di dati, pubblicazioni, software, strumenti, protocolli. In altre parole, se ci pensiamo, Open Science è solo la scienza, fatta bene.

Esistono strumenti per "aprire" ognuno dei passi della ricerca, come si vede in questo "arcobaleno":

Come puoi rendere Open ogni passo della ricerca...

- aggiungendo misure di impatto alternative, es. altmetrics
- comunicando sui social media, es. Twitter
- condividendo poster e presentazioni, es. su FigShare
- utilizzando licenze aperte, es. Creative Commons BY
- depositando in archivi o pubblicando su riviste Open
- provando la open peer review, es. PubPeer o F1000
- condividendo preprints, es. su OSF, arXiv o bioRxiv
- con formati leggibili dalle macchine, es. Jupyter o CoCalc
- con la scrittura collaborativa, es. Overleaf o Authorea
- condividendo protocolli e workflow, es. su Protocols.io
- condividendo note di laboratorio, es. OpenNotebookScience
- condividendo software, es. su GitHub con licenza GNU/MIT
- condividendo i dati, es. su Dryad, Zenodo o DataVerse
- pre-registrando esperimenti, es. su OSF o AsPredicted
- commentando pagine web, es. su Hypothes.is o Pund.it
- usando bibliografie condivise, es. su Zotero
- condividendo progetti di ricerca, es. su RIO Journal

Bianca Kramer & Jeroen Bosman <https://101innovations.wordpress.com> DOI: 10.5281/zenodo.1147025 Traduzione: Elena Giglia DOI: 10.5281/zenodo.1195648

Va sottolineato che ognuno di questi passi può essere aperto anche subito, anche in presenza di criteri di valutazione obsoleti: Open Science e Open Access si possono fare sempre, non sono in antitesi con la valutazione della ricerca.

La scienza aperta è non solo più robusta e verificabile perché mette a disposizione di tutti i dati, le procedure, i protocolli, ma poiché garantisce accesso a tutti si traduce in un maggiore impatto anche sulla società. **Non a caso l'economista Jeffrey Sachs individua nella Open Science lo strumento principe per raggiungere i 17 Obiettivi sostenibili delle Nazioni Unite.**

Cosa succede in Europa

Per questo motivo, perché la scienza sia al servizio dell'innovazione e della crescita, in Europa da anni Open Access e Open Science sono diventati una priorità nell'agenda politica. La strategia e le principali tappe sono ben riassunte nell'[articolo](#) di Jean Claude Burgelman dal titolo assai significativo "rendere la scienza adatta al 21° secolo".

In Italia sono ancora praticamente sconosciuti, e son poche le realtà virtuose o i Rettori che abbiano una visione strategica o che hanno adottato politiche e pratiche di scienza aperta. In tempi molto recenti, da fine 2018, il Ministero della Ricerca ha iniziato a [impegnarsi su questi temi](#), anche se con il 2020 si è subito una brusca battuta di arresto.

In Europa a novembre 2018 è stata lanciata [EOSC, la European Open Science Cloud](#). EOSC sarà [un ambiente virtuale](#) in cui produttori di dati, produttori di servizi e innovatori si incontreranno. I suoi blocchi costitutivi saranno i dati FAIR, Findable, Accessible, Interoperable, Reusable, che renderanno molto più semplice e lineare il processo di ricerca e abilitano intelligenza artificiale o tecniche quali il text e data mining.

Va chiarito che FAIR non significa Open: “Accessible” significa solo che come ricercatore devo sapere in quale repository si trovano i dati e a quali condizioni posso accedere. Ci sono dati che possono essere perfettamente FAIR ma non saranno mai Open per la loro stessa natura (si pensi a interviste a rifugiati politici): con le parole dell’ex Commissario [Carlos Moedas](#), “as open as possible, as closed as necessary”.

Per supportare i ricercatori sui dati FAIR è necessaria una buona rete di [data stewards](#), esperti di dominio che acquisiscano anche competenze sui metadati e gli identificativi persistenti (Findable), sulla conservazione e i formati (Accessible), su ontologie e standard (Interoperable), su documentazione e licenze (Reusable).

Perché serve la Open Science

La scienza aperta è strategica. In primo luogo per la scienza stessa, per la crescita della conoscenza, perché le idee circolano prima e circolano di più, e la scienza risulta più solida grazie alla condivisione dei dati e dei metodi. È strategica per il benessere di tutti, perché permette di affrontare in modo efficace crisi come quella del virus Zika o dell’attuale Coronavirus, in cui tecniche come il data mining dimostrano tutta la loro utilità, purché possano lavorare su testi aperti. Ed è strategica anche per l’innovazione. Sapete qual è il migliore esempio di Open Science? Il protocollo http, che era stato inventato da Tim Berners-Lee al CERN di Ginevra come strumento interno di scambio di documenti di lavoro fra i gruppi di ricerca del CERN. Avrebbero potuto brevettarlo, avrebbero potuto tenerlo chiuso, hanno deciso di aprirlo a tutti. Open Science significa una scoperta scientifica che cambia la vita di tutti noi. Ovvero, come fa un [laboratorio di neuroscienze](#) canadese, mettere a disposizione di tutti i dati per trovare prima, insieme, la cura.

Chi fa ricerca sta cercando di scardinare dal basso i meccanismi ormai obsoleti di una comunicazione scientifica ancora organizzata come quando esisteva solo la carta. Si stanno moltiplicando, per esempio, gli strumenti di scrittura collaborativa, e sta cambiando il concetto di “pubblicazione scientifica”: non si pubblicano più solo articoli ma anche dati, software, immagini, note di laboratorio. Sulla carta non si poteva fare altro, ma se ci pensiamo l’articolo è solo il “riassunto” dell’esperimento o della ricerca, mentre oggi possiamo condividere tutti i passaggi. Non solo, una volta la pubblicazione era il traguardo finale, l’articolo finalmente stampato su carta, oggi è invece solo l’inizio, perché se accompagnato dalle giuste licenze ogni lavoro può essere riusato. Sta cambiando il concetto di peer review, anche questa delegata alla fase post-pubblicazione, e intesa come giudizio da parte della comunità intera, oppure declinata come Open peer review, in cui i giudizi siano aperti e valutabili: una buona revisione richiede lungo tempo, e produce un “pezzo” di conoscenza, perché gli autori non dovrebbero essere valutati anche su questo? Stanno cambiando le riviste, che accolgono materiali diversi dagli articoli tradizionali, utilizzano la scrittura collaborativa, offrono servizi innovativi e spesso stanno al confine fra un repository e una rivista tradizionale, perché danno accesso anche ai pre-print. Stanno cambiando a livello europeo le norme sul copyright, che devono essere una tutela dei diritti ma non una protezione di interessi commerciali che si frappongano alla diffusione della scienza.

In uno scenario in rapida evoluzione, devono cambiare anche le regole con cui si valuta la ricerca, e in area Open si è alla ricerca di nuovi modi per misurare l’impatto reale di un lavoro. È appena uscito in Italia il nuovo Decreto sulla prossima VQR, valutazione nazionale della ricerca, che prevede che i prodotti scelti per la valutazione siano liberamente disponibili in un archivio Open Access. È un passo importante in termini di

trasparenza e riscontro sui risultati delle ricerche finanziate con fondi pubblici. È importante sottolineare che ciò che chiede il decreto è il deposito (tecnicamente, green road) in archivi Open di un lavoro ovunque esso sia stato pubblicato, in accordo con le politiche di copyright degli editori quali risultano nella banca dati [SHERPA-RoMEO](#). Il deposito è consentito nell'82% dei casi, è fattibile sempre e subito a costo zero, e consente di pubblicare sulle riviste di riferimento del settore, ottemperando ai criteri di valutazione in vigore. Il Decreto VQR non significa in alcun modo essere obbligati a pubblicare su riviste Open (cosiddetta "gold road"). Va sottolineato anche che solo il 27% delle oltre 14.000 riviste Open Access listate nella [Directory of Open Access Journals](#) richiede il pagamento delle cosiddette APC, Article Processing Charges o spese per la pubblicazione: ricordiamoci però che, a differenza degli abbonamenti, le APC vengono pagate una volta per tutte e non ogni anno, dalla sola istituzione dell'autore, e aprono il contenuto per tutti, per sempre.

Open Access e Open Science significano tutto questo. Non sono solo un ennesimo fardello burocratico imposto dai Regolamenti di Ateneo o dai bandi europei di Horizon 2020, che pure si basano sull'inattaccabile principio per cui la ricerca finanziata con fondi pubblici deve essere pubblicamente disponibile. Open Access e Open Science sono un'opportunità reale per una comunicazione scientifica più efficace, trasparente e utile all'innovazione, non più "confiscata" dietro barriere all'accesso insormontabili per molti.

E, come ci ricorda [Jean Claude Burgelmann](#), ci sono sì rischi a essere i primi, ma ci sono rischi maggiori a essere gli ultimi: una lezione che in Italia dovremmo imparare presto se non vogliamo essere esclusi dagli sviluppi (e dai finanziamenti) della European Open Science Cloud nei prossimi anni.