



SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

SISSA Digital Library

Open Science, Fair Science: garantire la trasparenza della scienza attraverso l'organizzazione della pratica quotidiana in laboratorio

This is a pre print version of the following article:

Original

Open Science, Fair Science: garantire la trasparenza della scienza attraverso l'organizzazione della pratica quotidiana in laboratorio / Crepaldi, D.. - In: GIORNALE ITALIANO DI PSICOLOGIA. - ISSN 0390-5349. - 2018:4(2018), pp. 747-764.

Availability:

This version is available at: 20.500.11767/105380 since: 2019-12-12T01:02:58Z

Publisher:

Published

DOI:10.1421/92987

Terms of use:

openAccess

Testo definito dall'ateneo relativo alle clausole di concessione d'uso

Publisher copyright

(Article begins on next page)

Open Science, Fair Science:
Garantire la trasparenza della scienza attraverso l'organizzazione della pratica
quotidiana in laboratorio¹

Davide Crepaldi²
Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste

Novembre 2018

Abstract

Experimental Psychology and Cognitive Neuroscience are reflecting more and more on their research practices, under the spur of the so-called “replication crisis”. Many tools have been made available to researchers lately, such as diverse statistical approaches and pre-registration, or have gained an unprecedented momentum. In this paper I will argue that these tools are extremely useful and should be used widely, but aren't all-purpose solutions to our problems. A more universal remedy is transparency, intended as a bid to make available to the community all research material (e.g., stimuli, data, analysis code) at some point of the project, as appropriate for each specific experiment. I will also argue that this practice can only be widely adopted (and is most effective) if the lab develops a standard research pipeline, where researchers work from Day 1 *as if* all their material would be under the public eye at all times. I will describe the pipeline that I've implemented in my lab to realize this principle, its strengths and weaknesses, and some tools that we may use to put that pipeline into practice on a daily basis. Finally, I will argue that establishing such a transparent pipeline isn't only ethic duty to our community, but also an effective way to make our research more visible and, eventually, give ourselves a better prospect for a shiny career.

¹ Questo articolo è ora pubblicato presso il Giornale Italiano di Psicologia (volume 45/4, Dicembre 2018, pagine 747-764). Questa versione è un pre-print non processato dall'editore, pubblicato a mero scopo scientifico e con l'obiettivo di incoraggiare il dibattito sui temi della trasparenza e della fair science.

² Potete contattare l'autore via mail usando il seguente indirizzo, davide.crepaldi@sissa.it; dare un'occhiata al sito del suo laboratorio presso questo URL, <https://rlac.sissa.it/>; oppure seguirlo su Twitter e Facebook, <https://twitter.com/crepaldidavide>, <https://twitter.com/CrepaldiLab>, <https://www.facebook.com/LanguageLearningReading/>.

Keywords: replicability, open science, research pipeline, data sharing, open everything.

La Psicologia Sperimentale e le Neuroscienze Cognitive stanno acquisendo sempre più consapevolezza dei limiti di alcune delle proprie pratiche di ricerca consolidate. Termini come “p hacking”, “file drawer effect”, “HARKing” erano praticamente sconosciuti alla gran parte dei ricercatori fino a pochi anni fa (pur se già ben presenti nella letteratura, (Begg & Berlin, 1988; Kerr, 1998; Rosenthal, 1979)). Oggi sono ormai entrati nel vocabolario corrente della disciplina. Nel complesso, è sempre più chiaro che il modo in cui facciamo ricerca nel campo delle scienze cognitive e comportamentali **non garantisce a sufficienza la riproducibilità dei nostri risultati** (Open Science Collaboration, 2015).

La comunità scientifica ha ovviamente reagito a questa sempre maggiore presa di coscienza, e lo ha fatto attraverso molti strumenti diversi: nuovi approcci statistici (Cumming & Calin-Jageman, 2016; Etz & Vandekerckhove, 2018), la pre-registrazione degli esperimenti, la conduzione della *peer review* sull’idea e il protocollo sperimentale prima che i risultati siano stati ottenuti (*registered reports*; (Chambers, 2013)). Ciascuna di queste iniziative affronta un aspetto specifico del problema della replicabilità, e sta dando il suo contributo verso una scienza comportamentale di migliore qualità. In questo articolo non entrerò nel dettaglio di questi problemi e delle loro (spesso ancora nascenti) soluzioni; per questo, vi rimando agli altri articoli che compongono questo numero speciale del GIP³. Mi concentrerò invece sui risvolti più pratici della questione: come è possibile, nel complesso, lavorare ogni giorno per produrre una scienza più affidabile? Come possiamo modificare le mille piccole azioni che ciascuno di noi mette in pratica ogni giorno in laboratorio? E’ possibile immaginare **una pipeline di ricerca** complessiva che riassume e riunisca tutte le pratiche virtuose che sono necessarie per garantire una buona affidabilità ai nostri risultati? Con il paragone che ho usato aprendo il mio intervento alla giornata di studi che ha dato origine a questa colletanea (che, vi ricordo, potete facilmente trovare on line, ad esempio [qui](#) e [qui](#)), cercherò di fare la parte della buona massaia, raccontandovi il modo in cui nel mio laboratorio cerchiamo di trasformare in pratica quotidiana le indicazioni ricevute dai grandi cuochi.

³ [Qui](#), [qui](#), [qui](#) e [qui](#).

Dall'esterno all'interno: ciascuno di noi deve contribuire a una scienza migliore.

Penso che sia anzitutto molto utile portare l'attenzione sulla **propria pratica di ricerca quotidiana**. Studiare a fondo gli aspetti statistici e metodologici del problema della replicabilità è ovviamente molto importante; ma se poi questo studio non si traduce in un'attenta riflessione sul modo in cui facciamo ricerca tutti i giorni, difficilmente otterremo un cambiamento duraturo e profondo.

Riflettere sulla nostra pratica di ricerca quotidiana enfatizza la **nostra responsabilità personale**. Leggere articoli di altri, in particolare su argomenti metodologici e statistici di cui spesso non siamo esperti, lascia sempre un po' un senso di alterità; un "altro" mi sta dicendo dove la comunità sbaglia, un "altro" si prenderà cura di sistemare le cose. Le riviste scientifiche cambieranno le loro linee guida, i revisori staranno più attenti, i miei collaboratori più giovani e più sensibili alle novità mi chiederanno di fare le cose in un modo nuovo: il motore del cambiamento è sempre qualcun altro. E invece no: nessun cambiamento è davvero duraturo se non viene dall'interno, da noi stessi, da uno sguardo critico sul modo in cui io faccio le cose, non qualcun altro. Parlando di questi temi ho spesso una sensazione simile a quando si parla di cambiamento climatico, o degli aspetti meno virtuosi del nostro sistema economico: si ragiona come se stessimo parlando di altro, come se il cambiamento climatico o il sistema economico fossero degli agenti esterni, altro da noi, indipendenti dalle nostre azioni. Raramente realizziamo che invece noi siamo il cambiamento climatico, quando prendiamo l'auto per fare un tragitto di 1km che potremmo tranquillamente fare a piedi; noi siamo il sistema economico un po' perverso, quando compriamo un articolo su una piattaforma on-line sapendo che il nostro risparmio esce dalle tasche dei dipendenti sottopagati e costretti a lavorare in condizioni terribili. Allo stesso modo, non possiamo delegare alle riviste scientifiche, ai *reviewers* o ai nostri colleghi la fatica del cambiamento, o nessun cambiamento ci sarà: noi stessi, io, ciascuno di noi deve riconsiderare il modo in cui produce ricerca.

Un altro elemento importante è la molteplicità degli aspetti che determinano l'affidabilità della nostra scienza: costruirne una migliore non è questione di introdurre **una** novità o cambiare **un** aspetto del nostro modo di lavorare, è **l'intera catena di**

azioni che ci portano dalla concezione dell'idea sperimentale fino alla pubblicazione dell'articolo che deve essere riconsiderata. Anche se suona molto radicale, questa affermazione non implica nulla di rivoluzionario: ci pensiamo probabilmente poco, ma ciascuno di noi ha, nei fatti, una pratica consolidata di ricerca intesa come un modo "tipico" di affrontare i vari passaggi che ciascun esperimento richiede. Non si tratta quindi di inventarsi qualcosa del nulla e "re-imparare" il nostro lavoro da zero, ma di riflettere a fondo su quanto già facciamo tutti i giorni. Il nodo cruciale qui è che è difficile innestare un comportamento virtuoso se non si rende virtuosa l'intera catena. Molte delle pratiche problematiche, infatti, si nascondono negli anfratti della nostra attività di ricerca, nelle tante piccole decisioni quotidiane che siamo chiamati a prendere: facciamo prima uno studio pilota, o partiamo direttamente con l'esperimento vero e proprio? Ci accontentiamo della stima dell'*effect size* che abbiamo fatto un po' a occhio, o cerchiamo di ottenere una stima migliore? Riportiamo nell'articolo anche quell'analisi che non è venuta benissimo, oppure teniamo soltanto il quadretto idilliaco dipinto dalle analisi standard, di cui difficilmente i *reviewers* si lamenteranno e che girano a meraviglia? Quasi tutti i ricercatori sanno ormai che esiste una risposta giusta a queste domande. Ma il punto è proprio questo: anche sapendo la risposta giusta, non è facile comportarsi come la teoria richiederebbe quando il giorno della decisione arriva, e dopo aver fatto le cose come si sono sempre fatte (tendenzialmente maluccio, ahinoi), ci troviamo a decidere se buttare a mare settimane di lavoro per evitare il rischio di p-hacking, oppure se indebolire il paper agli occhi dei *reviewer* riportando tutte le analisi, e non solo quelle che hanno funzionato bene. Senza costruirsi una *pipeline* di ricerca appropriata sin dall'inizio, prima di trovarci ad affrontare queste decisioni difficili in momenti in cui potremmo essere sotto forti pressioni in direzione opposta, difficilmente riusciremo davvero a radicare nella quotidianità le buone pratiche che quasi tutti ormai conosciamo.

Dalla molteplicità all'unità: il principio cardine è la trasparenza

Stabilire una pipeline di lavoro, tuttavia, crea inevitabilmente rigidità; e questo non si sposa bene con la complessità della questione replicabilità, né più in generale con la complessità della ricerca scientifica. Un esempio molto chiaro di questa complessità è la

molteplicità di iniziative che la comunità ha messo in piedi per ovviare al problema della replicabilità. Si è lavorato sugli aspetti più prettamente statistici (Cumming & Calin-Jageman, 2016), oppure sulle modalità di revisione degli articoli (Chambers, 2013), oppure ancora, come stiamo facendo qui, sugli aspetti più generali del modo in cui facciamo ricerca. Ciascuno di questi fronti è importante: ma nessuno di essi è, da solo, la risposta al problema. Un *registered report* è perfetto per alcune domande di ricerca; ma non è adatto ad altre, per esempio quando la ricerca sia principalmente di natura esplorativa e/o ci si stia muovendo in un campo piuttosto inesplorato, teoricamente o metodologicamente (dove, per esempio, è quasi impossibile stabilire una stima dell'*effect size*). Gli approcci Bayesiani alla modellazione dei dati sono incredibilmente informativi, particolarmente in certi casi; ma non per questo il vecchio approccio basato sul *Null Hypothesis Significance Testing (NHST)* è interamente da rottamare. Qualsiasi pipeline che sposi uno di questi approcci sarebbe necessariamente limitata, privandosi della possibilità di sfruttare i pregi dell'approccio opposto: e come tutti gli strumenti, anche i nostri arnesi di ricerca non sono di per sé né buoni né cattivi, dipende tutto dall'uso che se ne fa.

Come, dunque, adottare una pipeline di ricerca standard, con il suo scudo di protezione quotidiano alle cattive pratiche, mantenendo però la flessibilità metodologica necessaria per affrontare domande scientifiche di natura diversa? La soluzione che propongo in questo articolo è legata a un'unica parola chiave: **trasparenza**. L'idea centrale è che tutto quello che produciamo in laboratorio (stimoli, codice per presentare l'esperimento e raccogliere i dati, codice di analisi dei dati, ecc.) viene messo a disposizione della comunità. Così facendo, non dobbiamo necessariamente sottoscrivere a vita un certo strumento o approccio di ricerca: quando lo riterremo opportuno, pre-registreremo, altre volte no; qualche volta analizzeremo i dati con un semplice *t test*, altre volte li modelleremo con dei modelli misti o dei GAM, mantenendo così la nostra piena flessibilità metodologica, e con essa le molte virtù dei diversi approcci di ricerca e/o di analisi dei dati. Tuttavia, saremo protetti dalla tentazione di scorciatoie inappropriate dall'occhio pubblico: mettendo a disposizione del lettore tutte le informazioni di cui siamo a disposizione noi stessi, sappiamo che le nostre scelte dovranno sempre essere ben giustificate e ponderate, nonché ovviamente libere da quelle pratiche di ricerca che minano la solidità dei nostri risultati.

Dall'astratto al concreto: la vita quotidiana nel [Language, Learning and Reading lab in SISSA](#)⁴

Uno dei vantaggi della semplicità dell'idea chiave descritta sopra è che essa rende molto semplice anche l'implementazione pratica della pipeline. Di fatto, questa si riduce al **deposito di una serie di documenti chiave nel processo di ricerca in un repository** che sia intanto **condiviso all'interno del gruppo di ricerca**, e poi possa **diventare pubblico in modo semplice e veloce**. I documenti chiave rappresentano il punto finale delle diverse fasi di cui un esperimento si compone:

1. un brevissimo testo (tipicamente non più di 8-10 righe) che riassume l'idea sperimentale e gli elementi fondamentali dello studio (il paradigma sperimentale, il numero di partecipanti richiesti, i possibili risultati e la loro interpretazione teorica, ecc.);
2. Il set di stimoli per l'esperimento;
3. Lo script per la presentazione degli stimoli ai soggetti e la relativa raccolta dati;
4. I dati: in forma grezza, così come raccolti durante il testing (tipicamente); oppure in forma aggregata (quando la natura dei dati lo renda più appropriato, come nel caso di dati neurofisiologici);
5. Il codice per l'analisi dei dati;
6. Il manoscritto.

La condivisione di questi documenti, prima di tutto interna, porta tutta una serie di benefici che, prima ancora di rendere la ricerca più affidabile, **semplificano la routine quotidiana dei ricercatori**. E questo, si badi bene, è un ingrediente fondamentale della ricetta: se vogliamo che una pipeline venga adottata davvero come pratica di lavoro quotidiana, deve anche imporre un carico di lavoro gestibile ai ricercatori; meglio ancora se addirittura semplifica loro la vita.

Immaginate uno scenario concreto, che nel mio lab (ma sono certo anche in molti altri) è particolarmente frequente. Uno studente scrive la seguente mail al proprio *principal investigator* (PI):

⁴ Potete trovare una descrizione più completa della nostra pipeline di ricerca, e più in generale di come cerchiamo di lavorare nel mio lab alla SISSA, nella nostra [Lab Userguide](#).

Ciao Davide,

Ho fatto le analisi che abbiamo immaginato insieme due settimane fa: torna tutto, yeah! L'interazione tra lunghezza e predicibilità della parola in effetti è piuttosto solida, e resiste alla rimozione degli outlier (come abbiamo discusso durante il meeting). Che ne pensi?

Letizia

Ora, il PI —che tra la riunione di cui parla Letizia e il momento in cui legge la mail avrà fatto 8 riunioni di Dipartimento, elaborato circa 500 e-mail⁵, fatto da *reviewer* a 3 articoli, revisionato 2 articoli di altri studenti, discusso di altri 10 esperimenti con altri membri del lab, rivisto 2 poster che i suoi studenti presenteranno a una conferenza la settimana successiva, preparato e condotto 4 lezioni, e lavorato un paio di mezze giornate al suo prossimo *grant proposal*— penserà: “Analisi? Oddio, di quali analisi stiamo parlando? E perché predicibilità, non era parte del nostro design? O mamma, ma non è che forse mi ricordo male e alla fine avevamo deciso di studiare anche questa variabile nell'esperimento? O forse mi sto confondendo con l'esperimento di Giacomo?”. Tipicamente, il PI risponderrebbe alla mail con tutte queste domande, magari un po' camuffate così da non fare troppo la figura del vecchio amnesico; al che seguirebbero probabilmente ancora due o tre mail chiarificatorie, prima finalmente di arrivare al dunque. Ecco, ora immaginate che questo PI possa invece semplicemente andare nel repository concordato⁶, leggere il documento che descrive l'esperimento in 8 righe, aprire lo script con le analisi e farle correre: risolverebbe i suoi dubbi in non più di 10–15 minuti (ché, come tutti gli amnesici lievi, la sua memoria migliora moltissimo con i *cue* giusti), e sarebbe in grado di rispondere alla studentessa con la sua opinione senza alcun giro di e-mail ulteriore. Miglior scienza e miglior supervisione, il tutto condito da un gran risparmio di tempo.

Altro scenario classico: finalmente arrivano delle *review* positive su un articolo che abbiamo già provato a pubblicare tre volte senza successo, e il cui esperimento è stato dunque effettuato tre anni prima. Uno dei revisori chiede come mai i gradi di libertà

⁵ Io ne ricevo 35 e ne scrivo 15 al giorno in media, che moltiplicati per 10 giorni lavorativi fanno quel numero lì.

⁶ Nel lab chiamiamo questo *repository* “The Fount of All Wisdom”, o semplicemente “The Fount”, per noi amici.

totali di una certa analisi non corrispondano alla somma dei datapoints raccolti (così come descritti nei metodi) meno uno. Panico totale, dopo dieci minuti di imprechi e la visione autoconsolatoria del [fantastico video sul famigerato Reviewer 3](#), si capisce. Lo studente che aveva fatto l'esperimento ora lavora in Lapponia per un progetto sullo scioglimento dei ghiacci in Artide. Avendo altro da fare, risponde 5 giorni dopo l'e-mail del PI, che nel frattempo ha esplorato l'intero archivio di posta elettronica e l'intero hard-disk nella speranza di trovare una copia di un qualche file che potesse vagamente ricordargli l'esperimento in causa. Essendo lo studente giovane e dotato di una memoria molto migliore del PI⁷, risponde che si ricorda benissimo dell'esperimento; ma che i dati stavano su un computer che ha rottamato ormai due anni prima, e che quindi deve controllare nel suo archivio. Altri giorni passano, e un sacco di ore di lavoro vengono sprecate per recuperare dei files che non abbiamo immagazzinato appropriatamente in un archivio *ab origine*. Anche ammettendo che alla fine si risolve l'arcano e il PI possa scrivere tutto felice che i 21 gradi di libertà mancanti vengono da una *response box* malfunzionante che aveva generato 21 tempi di reazione negativi, è ovvio che una pipeline di ricerca appropriata dovrebbe evitare questi inconvenienti, e con essi le enormi perdite di tempo che ne derivano. Bene, se il PI avesse potuto andare nel *repository* condiviso dell'esperimento, avrebbe probabilmente risolto il problema in mezz'ora grazie ai dati originali e allo script di analisi che lui e il suo studente avevano providenzialmente immagazzinato lì tre anni prima. (O, meglio ancora, il reviewer non avrebbe nemmeno sollevato il punto, potendo andare lui stesso a controllare nei dati e nello script di analisi, se essi fossero stati anche resi pubblici al momento della spedizione dell'articolo alla rivista.)

I due casi appena raccontati fanno riferimento all'organizzazione interna del gruppo di ricerca; ma le dinamiche non sono diverse da quelle che si innescano quando si vuole rendere pubblico il proprio lavoro. Pensate al caso in cui un ricercatore decida, appunto, di pubblicare i propri dati dopo averli trattati come tipicamente avviene nelle *pipeline* tradizionali, con i file che vengono elaborati in locale, sul proprio computer, da ciascun membro del gruppo di ricerca. Anche qui, la ricostruzione dei vari passaggi e la preparazione del dataset *ex post* richiede un sacco di lavoro; e questo ovviamente inficia

⁷ Chiaramente qui sto parlando di me.

la potenziale qualità della nostra ricerca (ciò che è difficile e richiede molto tempo genera inevitabilmente errori) complicando la vita a chi cerca di essere trasparente. Riprendendo un tweet di Tal Yarkoni, professore alla University of Texas at Austin, riassumerei i precedenti tre paragrafi in un principio che chiamerei **too-lazy-too-busy**: abbiamo troppe cose da fare, e siamo qualche volta anche troppo pigri, per poter pensare di fare la stessa cosa più volte. Se vogliamo rendere pubblici i prodotti della nostra ricerca, come consiglio a tutti caldamente di fare, dobbiamo costruire una *pipeline* che sia pensata per questo sin dall'inizio, così che non ci siano aggravii di lavoro in fase di "confezionamento", e anzi, il nostro lavoro sia reso migliore e più semplice sin dall'inizio, anche quando le nostre cose non sono ancora sotto l'occhio pubblico.

Gli strumenti

E' ovviamente necessario disporre degli strumenti giusti per realizzare la condivisione di cui abbiamo parlato nei paragrafi precedenti; e in questo periodo storico ne abbiamo a disposizione in abbondanza (si pensi, ad esempio, a Google Drive, Dropbox, iCloud, i vari sistemi Git, the [Open Science Framework](#), ecc.). La difficoltà sta, casomai, proprio nel decidere in questa abbondanza: essa fa sì che ciascun membro del lab arrivi nel gruppo con le sue preferenze, sviluppate in anni di utilizzo di un certo strumento, e magari anche con piani di abbonamento che sta già pagando. Portare tutti a usare lo stesso sistema non è semplice; né, forse, la scelta migliore.

Rivelo subito la mia politica: il lab paga un account per alcuni servizi (al momento, soltanto il Git server) e chiede ai membri di usare quei servizi nel momento in cui fossero necessari. In generale, però, per ciascun progetto specifico discutiamo quale sia il tipo di *repository* migliore, anche sulla base della familiarità che i membri più giovani della squadra hanno con le varie opzioni sul tavolo. Questo sistema richiede ovviamente un po' di flessibilità da parte mia, trovandomi a lavorare con strumenti diversi su diversi esperimenti. Inoltre, ha il difetto di non permettere un unico *repository* per tutti i progetti del lab, il che rende più difficile a un utente esterno trovare i nostri lavori. Però ha il vantaggio di non costringere gli studenti e i post-doc a familiarizzarsi con un sistema completamente nuovo per loro, il che richiederebbe ovviamente un po' di tempo e implicherebbe un certo rischio di commettere errori. La ragione fondamentale

di questa politica è che saranno poi proprio gli studenti e i post-doc a lavorare tutti i giorni con quel sistema, per cui mi pare giusto dare la priorità al loro comfort piuttosto che al mio, o a quello dell'utente esterno.

In questi anni di esperienza con la pipeline del mio laboratorio, ho visto che ci sono sostanzialmente **tre tipi di repository/servizi on-line** tra cui scegliere per gestire il flusso e i prodotti della propria ricerca. Un primo modello è probabilmente quello più semplice, rappresentato prototipicamente da **Google Drive**. Esso offre uno spazio di archiviazione on-line, con un sistema di controllo di accesso molto semplice da usare, ma, al tempo stesso, sufficiente nella grande maggioranza delle situazioni. La facilità d'uso è senz'altro il pregio maggiore di questi tipi di sistemi, a cui si aggiunge il fatto, specifico per Google Drive, che molte istituzioni di ricerca stanno esternalizzando il proprio servizio di posta elettronica all'azienda di Mountain View, il che garantisce uno spazio di archiviazione teoricamente infinito a ciascun utente senza imporre costi aggiuntivi al laboratorio.

A questo si aggiunge forse il maggior pregio del sistema Google Drive, l'**editing condiviso on-line**: questo permette di lavorare a più mani su qualsiasi documento, senza il classico problema di dover necessariamente "serializzare" il lavoro tra i membri del gruppo. E il sistema di editing on-line di Google è del tutto comparabile ai classici strumenti *Office* in termini di capacità e ricchezza di strumenti: ci si trova tutto quel che serve nello scrivere un documento insieme, dalla possibilità di inserire commenti, al tracciamento dei cambiamenti, a un vero e proprio controllo di versione (si può sempre viaggiare sulla "linea del tempo" e tornare a versioni precedenti).

Credo ci siano, però, ancora **due limiti fondamentali** che rendono impossibile eleggere Google Drive a unico strumento di lavoro (oltre ad alcuni problemi di carattere etico/economico di cui parlerò sotto). Quando si lavora a un codice invece che a un documento o a un foglio di lavoro, il sistema di editing on-line non è disponibile; questo significa che è necessario scaricare il file, lavorarci sul proprio computer e poi ricaricare il risultato finale sul Drive. Così facendo, si perde l'immediatezza d'uso garantita dal lavoro on-line, e anche la chiarezza che deriva dal poter racchiudere tutti i nostri commenti, le nostre proposte di modifica e le nostre discussioni in un unico file.

Esistono [software che permettono di sincronizzare il Drive su una o più cartelle locali](#),

ma la mia esperienza con questi sistemi non è del tutto positiva: ci sono spesso lag, anche significativi, tra il Drive on-line e le cartelle sincronizzate sul proprio computer, ad esempio.

Il secondo limite fondamentale è la gestione automatica delle citazioni, che rende non semplicissimo usare il Drive come strumento per scrivere il manoscritto da spedire poi alle riviste per la pubblicazione. Sono emersi negli ultimi anni strumenti che affrontano in qualche modo questo problema, incluso [Paperpile](#), che sto usando ora per redigere questo documento. I risultati sono senz'altro buoni, ma non ancora comparabili a quelli che si ottengono scrivendo in Latex, ad esempio, o utilizzando i più classici *citation manager* come [Mendeley](#), EndNote o [JabRef](#).

Il secondo tipo di strumento è esemplificato dal sistema [Git](#). Questo software è stato sviluppato specificamente per garantire un efficace controllo di versione, ed è quindi senza dubbio il sistema più completo e sicuro per lavorare a più mani su uno stesso progetto. L'architettura è pensata strutturalmente per permettere a ciascun utente di lavorare in locale, e dunque permette di operare sul sistema in modo molto efficace anche quando si è off-line (al contrario di Google Drive, che è invece piuttosto scomodo da gestire senza una connessione Internet). Proprio perché il sistema è fondato sul lavoro in locale, Git ha un altro grande vantaggio rispetto al Drive: si condivide con tutti (*commit*, nel gergo del software) solo quando si è pronti a farlo, mentre si è liberi di esplorare quanto si vuole sulla propria macchina senza che questo rischi di disturbare il lavoro dei collaboratori⁸. Anche il sistema di condivisione e controllo è certamente migliore di quello di Google Drive: si pensi ai *forks* o ai *pull request*⁹, che non esistono, oppure sono molto meno potenti, nel mondo Google.

Ma niente di buono viene a questo mondo senza un po' di fatica, e la indubitabile maggiore potenza di Git richiede tipicamente un po' di apprendimento da parte nostra e dei nostri collaboratori: il sistema è davvero semplice da usare una volta che lo si è imparato, ma impone un modo di ragionare diverso rispetto al classico approccio dei

⁸ Ovviamente questo è possibile anche usando Google Drive, ma richiede che i file condivisi vengano scaricati, e poi ricaricati in un secondo momento; lavorando on-line, qualsiasi modifica diventa immediatamente visibile a tutti.

⁹ Sono consapevole che questi termini potrebbero essere incomprensibili a chi non è familiare con il sistema Git, ma spiegare nel dettaglio come questo funziona esula dagli scopi di questo articolo. Rimando il lettore interessati ai moltissimi video e guide on-line sul tema, ad esempio [qui](#) e [qui](#).

software più popolari. Insomma, il mio messaggio è: vale certamente la pena imparare Git, ma bisogna armarsi di un po' di pazienza. Non iniziate a usarlo a 4 giorni da una deadline!

L'ultimo tipo di risorsa on-line che usiamo in lab si rende necessaria per lavorare collaborativamente alla scrittura del manoscritto. Su questo fronte, **Latex** domina incontrastato: nessun altro sistema ha la sua potenza e permette di concentrarsi così bene sul contenuto e la struttura di quanto si sta scrivendo. Poi, certo, anche qui la qualità complessiva del sistema si paga con un po' di complessità: c'è una curva di apprendimento da affrontare all'inizio. Il lavoro collaborativo non è inoltre semplicissimo in Latex; ma è qui che intervengono le risorse on-line cui facevo riferimento qualche riga sopra. Esistono infatti alcune piattaforme (ad esempio, [Overleaf](#)) che permettono di tenere traccia dei cambiamenti, conservare uno storico del documento in lavorazione e inserire commenti (che poi possono svilupparsi in vere e proprie conversazioni, come in Google Drive); tutto questo in Latex, conservando quindi tutti i principali vantaggi di questo sistema, come l'universalità (rispetto al sistema operativo, per esempio), la strepitosa gestione automatica della bibliografia, la semplicità di integrazione di formule matematiche e la formattazione semplice, che poi può essere adattata a qualsiasi formato sia richiesto dall'una o l'altra rivista con una semplice riga di comando. Queste piattaforme on-line hanno anche il vantaggio (notevole) di gestire il sistema Latex nel suo complesso al proprio interno: nessun bisogno, quindi, di installare e mantenere il software sul proprio computer.

Due note importanti, in chiusura di questa sezione. Anzitutto, gli aspetti etico-economici cui ho fatto riferimento sopra. Molte delle risorse che ho menzionato richiedono un pagamento; e anche quelle che non lo fanno, sono comunque offerte da aziende profit, il cui fine ultimo non è servire l'umanità, ma fare soldi. Nel momento in cui il meritorio scopo di semplificare il lavoro della gente dovesse entrare in contrasto con il loro interesse economico, non c'è alcuna garanzia che queste aziende continuino a garantire i servizi di cui ho parlato sopra. E siccome su questi servizi noi staremmo basando la nostra scienza (oltre a immagazzinare qualche quintale di dati su cui abbiamo una responsabilità legale), la questione è piuttosto delicata. Si potrebbe

ovviare al problema implementando tutti **questi servizi in-house, sui server della nostra istituzione**. Ci sono sistemi liberamente disponibili sul web che permetterebbero questa soluzione; la mia personale preferenza va a [Sandstorm](#), ma sono certo che esistono molte valide alternative che non conosco. Il problema tipico è che questo ovviamente richiede un lavoro non banale da parte dei nostri sistemi IT, che, come risaputo, non sono facili da convincere: in presenza di risorse limitate, il supporto “casalingo” di un’infrastruttura come quella descritta sopra chiaramente fa fatica a raggiungere un livello di priorità sufficientemente alto. Anche su questo versante, è necessario un cambiamento culturale, tanto nascosto ai riflettori quanto fondamentale: senza una buona infrastruttura, nessun buon sistema può funzionare.

Una possibile soluzione al problema di non delegare ad aziende profit l’infrastruttura di ricerca e al contempo non oberare i sistemi IT della nostra singola istituzione è che **la comunità scientifica si costruisca l’infrastruttura essa stessa**, in modo che più ricercatori, anche da istituzioni diverse, possano utilizzarla insieme: questo è esattamente quello che ha generato [Open Science Framework](#). OSF è diventato ormai lo standard per la pre-registrazione degli esperimenti; ma non tutti sanno che la piattaforma offre molto più di questo, incluso una gran parte dei sistemi di cui ho parlato sopra. In OSF è possibile, tra le altre cose, co-editare documenti (con qualche funzionalità in meno rispetto a Google, ma nel complesso con una buonissima *user experience*), mantenerne uno storico, organizzare il progetto in una struttura di sub-progetti collegati, e impostare una pagina *wiki* per aiutare l’utente (collaboratore o semplice lettore esterno, nel caso in cui il progetto venga a un certo punto reso pubblico) a navigare nel progetto. Insomma, OSF è ormai un’infrastruttura in grado di aiutare gli utenti a **gestire progetti di ricerca**, oltre che a pre-registrarli e pubblicarli.

La seconda nota è forse ovvia, ma non per questo meno importante. La pipeline di ricerca che vi ho raccontato, così come l’opinione che mi sono fatto sugli strumenti che ho usato in questi anni per realizzarla, riflettono la **mia esperienza personale**; il mio gusto, la mia sensibilità, il mio approccio alla ricerca. **Sono una soluzione possibile; non LA soluzione**. Vi incoraggio moltissimo a ragionare sulle falle del sistema che vi sto proponendo (si veda il prossimo paragrafo, tra l’altro), oltre che sulle sue virtù; e vi incoraggio a esplorare altre possibili *pipeline* e/o altri strumenti per realizzarle.

Incoraggio anche i PI all'ascolto: stimolate l'esplorazione e la critica nei vostri collaboratori più giovani, anche (anzi, soprattutto!) quando questo significa mettere in discussione il vostro modo di lavorare. Più menti brillanti sono sempre meglio di una sola mente brillante; e spesso le menti dei nostri giovani collaboratori sono più brillanti delle nostre. E' molto importante che la "scoperta" della relativa inaffidabilità della nostra scienza non diventi un motivo per irregimentarla, per costringere tutti in un unico vestito, trasformando così un atto di creatività, qualche volta visionario, in niente più che un'arida e ripetitiva catena di operazioni degna del fordismo anni '70. Dobbiamo certamente eradicare le brutte abitudini; ma rimangono moltissime le buone abitudini che ciascuno di noi può decidere di adottare. *Your science, your way* (purché la tua strada non preveda p-hacking, *of course*; e meglio ancora se è interamente trasparente).

No pain, no gain: i problemi che emergono nel cambiamento

Spero di avervi convinto, a questo punto dell'articolo, (i) dell'importanza di costruire una pipeline di ricerca per il vostro laboratorio e il vostro lavoro; e (ii) che questa pipeline, se sufficientemente trasparente e aperta, può proteggerci dalle cattive pratiche di ricerca che minano la solidità della nostra scienza, senza per questo limitare gli strumenti (buoni) che compongono la nostra cassetta degli attrezzi. Spero davvero di avervi convinto, perché dopo pagine di sperticate lodi ora mi tocca rivelarvi anche il **lato oscuro della storia**, ahimé. Scherzi a parte, penso sia importante condividere anche le difficoltà del percorso, un po' perché non vi scoraggiate troppo quando ci proverete anche voi, un po' nello spirito del finale del paragrafo precedente: è anche dalla condivisione delle difficoltà che possono nascere idee nuove, o nuovi modi di perseguire gli stessi obiettivi.

Direi che le vere difficoltà nell'implementazione del sistema di ricerca descritto in questo articolo sono sostanzialmente due, in parte connesse tra loro. La prima, e forse quella più fondamentale, deriva dal fatto che la pipeline impone una cura del dettaglio nel corso dello svolgimento della ricerca che effettivamente **rallenta il processo**. Invece di accontentarsi della confusione totale della classica cartella "stimoli" sul nostro computer ("primaBozza", "primaBozzaDavide", "primaBozzaDavideLetizia", "primaBozza2", "versioneQuasiFinale", "versioneQuasiFinale2", ecc.), la pipeline ci

richiede di fare ordine da subito e generare un unico file che sia completo e, cosa non banale, comprensibile a un utente esterno. Questo richiede tempo, un tempo che gli studenti o i post-doc, nella loro attività standard, non aspetterebbero prima di correre in laboratorio a testare. (Notate come questa procedura aiuti, peraltro, a fare emergere eventuali errori in una fase in cui è ancora possibile rimediare ad essi. Questo aiuta la qualità della nostra ricerca in modo diretto, come è ovvio; ma anche in modo indiretto, riducendo la probabilità che ci si trovi poi con dei dati già raccolti, ma compromessi da eventuali errori, in una posizione dove la pressione a pubblicare potrebbe portare a fare scelte sbagliate anche il ricercatore meglio intenzionato.)

Un altro esempio classico è lo script di analisi dei dati. Anch'esso diventa molto velocemente un ammasso di linee di codice incomprensibile per tutti tranne che per noi; presi dall'eruzione di idee e dalla *serendipity* che spesso caratterizza la modellazione dei dati, il nostro percorso logico si perde velocemente in un coacervo di comandi. Senza la pressione imposta dal dover rendere pubblico quel codice, la maggioranza dei ricercatori non ci tornerebbe più sopra una volta ottenuta la serie di risultati che andranno a costituire lo scheletro del proprio articolo. Organizzarlo appropriatamente, invece, forza ciascuno di noi a ripensare bene a quanto fatto, a distinguere il necessario dal superfluo, a chiarire la struttura logica che ha determinato il nostro percorso di esplorazione dei dati: tutto questo ci protegge sia, banalmente, dagli errori che, meno banalmente, dalle pratiche di ricerca che sappiamo essere nefaste per l'affidabilità dei nostri risultati. Ma sistemare un codice di analisi e commentarlo appropriatamente può richiedere giorni, se non settimane. Ancora una volta, sono più settimane che ci separano dall'agognato momento della *submission* e, speriamo, della pubblicazione del nostro lavoro.

Insomma, il sistema, come diciamo noi psicologi, è a ricompensa ritardata, e questo non è facile da digerire in un mondo accademico dove lo *storytelling* dominante dice che un paper in più o in meno al momento giusto può decidere della nostra carriera¹⁰.

La seconda difficoltà è che **il sistema funziona solo se tutto il gruppo di lavoro lo adotta**, e anzi, si convince appieno che quello sia il modo giusto di fare ricerca. Una volta immaginata l'architettura, dunque, il lavoro è solo all'inizio: convincere i

¹⁰ Questo *storytelling* è molto fuorviante, in verità; ma questa è un'altra storia.

collaboratori a fare lo sforzo aggiuntivo di cui sopra non è per nulla semplice, e tuttavia indispensabile. Questo vale ovviamente per i ricercatori all'inizio della loro carriera, che chiaramente non possono imporre al proprio gruppo e al proprio supervisore un certo modo di lavorare. Ma vale anche per chi invece è già alla guida del proprio gruppo: imporre non è mai una buona idea; e, in ogni caso, non funziona. Nel mio laboratorio il lavoro di coinvolgimento del gruppo è ancora molto incompleto. Alcuni dei miei studenti e post-doc hanno adottato il sistema qui descritto in modo molto naturale e, per esempio, mantengono il *repository* di progetto in splendida forma. Alcuni altri sono invece meno convinti e, anche se magari non sollevano esplicitamente le loro perplessità, non seguono la *pipeline* come a me piacerebbe. Tutto ciò è nell'ordine delle cose; e in fondo, chi deve gestire la ricerca nella quotidianità, e dunque pagare il costo di uno standard di ricerca più alto, sono prima di tutto loro. Io posso indicare loro una strada, e mi sforzo ogni giorno di trovare buoni argomenti per convincerli che quella è la strada giusta. Ma credo che nessun PI possa (né debba) costringere i propri colleghi più giovani sulla stessa strada dove andrebbe lui.

Etica e (legittimo) interesse personale

La sezione precedente ci porta dritti alla domanda con cui vorrei concludere l'articolo. Seguire il metodo di ricerca che ho proposto qui, e più in generale cercare di conformarsi a standard di ricerca migliori, è soltanto un debito di onestà nei confronti del resto dell'umanità, oppure **conviene anche al singolo ricercatore?** Anche qui, svelo subito la mia risposta. Fare ricerca bene è prima di tutto un dovere etico, e chi non lo fa commette un'ingiustizia, ed è dunque colpevole, nei confronti dei suoi colleghi e della comunità *at large*. Tuttavia, fare ricerca bene conviene anche al singolo ricercatore, nel senso che ne asseconda il legittimo interesse personale e migliora le sue prospettive di carriera.

Che sia un dovere morale fare del nostro meglio per garantire l'affidabilità dei nostri risultati scientifici dovrebbe essere del tutto ovvio. Qui mi limiterò a sottolineare un aspetto che noi scienziati spesso dimentichiamo. I soldi che pagano il nostro stipendio e i nostri esperimenti vengono dalle tasche dei contribuenti, e dunque, come minimo, dobbiamo loro la conoscenza più solida e affidabile che possiamo generare.

Prendere qualche scorciatoia, sapendo che di conseguenza i risultati dei nostri esperimenti potrebbero non indicare la direzione giusta, è un modo piuttosto clamoroso di tradire il nostro contratto con la comunità.

Approfitto di questo passaggio sul nostro “contratto” con la comunità non accademica per evidenziare un altro aspetto in cui la *pipeline* proposta in questo articolo onora al meglio i finanziamenti che riceviamo dalle tasche dei contribuenti. Rendere pubblico tutto quanto produciamo con il nostro lavoro (dati, stimoli, pezzi di codice, ecc.) riduce la probabilità che qualcun altro faccia il nostro stesso lavoro una seconda volta, o che noi facciamo una seconda volta il lavoro che qualche nostro collega ha già fatto. A maggior ragione essendo pagati dalla stessa fonte, questo non ha davvero alcun senso: e pensate invece a quante volte capita, con il “solito modo” di fare le cose! Addirittura, non è raro che tra colleghi (e addirittura all’interno dello stesso lab, qualche volta) ci si neghi la condivisione, con il pensiero, implicito o esplicito: “con tutta la fatica che ho fatto!”. Ma proprio per questo si dovrebbe condividere: pensate a quante altre cose interessanti potrebbe fare il nostro collega (o potremo fare noi, la prossima volta) con il tempo e lo sforzo che risparmierebbe (risparmieremmo). Mettiamoci nei panni di chi ci paga: quanto saremmo contenti di pagare due volte per ottenere due volte la stessa identica cosa?

Proprio l’esempio del paragrafo precedente ci introduce a una nuova, più che legittima, domanda: perché mai dovremmo “regalare” il nostro lavoro a un collega, che magari tra qualche mese sarà il nostro concorrente per ottenere il posto che tanto vorremmo? (Per inciso, il solo fatto che questa domanda suoni legittima ci dice quanto il nostro sistema attuale sia assurdo.) Qui la parola chiave diventa **riconoscimento**. Se il nostro collega, come ovviamente dovrebbe fare, riconoscerà il nostro contributo, noi avremo ottenuto una ricompensa nella moneta più preziosa nel mondo accademico: il riconoscimento dei pari. Se ci saremo preoccupati di pubblicare più ufficialmente la risorsa che stiamo condividendo (alcune autorevoli riviste ormai pubblicano senza problemi set di stimoli e di dati¹¹), quel riconoscimento prenderà addirittura la forma di una vera e propria citazione. Più conveniente di così!

¹¹ Tra i vari esempi virtuosi, cito qui il neonato [Journal of Cognition](#), pubblicato e curato dalla European Society for Cognitive Science (ESCoP).

Condividere dà visibilità al nostro lavoro anche al di là delle citazioni vere e proprie. Pensate a quanto possa rendervi popolari un pezzo di codice che fa in 30 secondi una cosa che, fatta a mano, richiede giorni di lavoro, ad esempio. Qualcosa di simile mi è capitato proprio di recente. Qualche anno fa, stanco del fatto che non riuscivo a rendere simili per una certa variabile due gruppi di stimoli che mi interessava confrontare (il famigerato *matching*, spauracchio di generazioni di psicolinguisti), ho scritto un codice che, dato un certo gruppo di parole A, pesca da un set di potenziali candidati X un sottogruppo B che mostra proprietà simili a quelle di A. Il codice può gestire diverse variabili contemporaneamente (ad esempio, può rendere equiparabili A e B sia per lunghezza che per frequenza d'uso), e permette al ricercatore di stabilire un peso relativo all'importanza delle diverse variabili (ad esempio, possiamo stabilire che la lunghezza conti il doppio della frequenza d'uso nella qualità del *matching*). Il codice è stato pubblico sul mio sito personale per anni senza entusiasmanti conseguenze. Qualche mese fa, però, una ricercatrice dell'Università di York ha twittato chiedendo ai suoi follower se conoscessero un codice del genere; e una seconda collega le ha risposto puntando a questo mio codice. Ne è nata una conversazione in cui sono intervenuti un'altra decina di colleghe e colleghi (che hanno poi svelato l'esistenza di altri software simili che io ovviamente non conoscevo, peraltro); e in quei due giorni i contatti sul mio sito sono più che raddoppiati. Questi contatti non sono ovviamente citazioni che potrò spendere alla prossima Valutazione della Qualità della Ricerca o al mio prossimo concorso; ma mi hanno messo sulla mappa mentale di molti miei colleghi, di cui avrò guadagnato la stima, se il mio codice sarà divenuto uno strumento utile alla loro ricerca.

Morale della favola: se si lavora con qualità e si condividono i risultati del proprio sforzo, forse impiegheremo più tempo a fare le cose, ma ci daremo la possibilità di guadagnare un capitale di stima tra i nostri colleghi che difficilmente potremmo ottenere diversamente. E badate: la **stima dei propri colleghi è la prima determinante della nostra carriera**, molto più fondamentale dei *papers* e delle citazioni, che sono utili proprio nel momento in cui ci permettono di guadagnare la stima di chi poi dovrà decidere se assumerci.

Mi riservo l'ultimo paragrafo per sfatare un altro dei miti che rende difficile la condivisione dei prodotti della ricerca, specialmente prima della pubblicazione dell'articolo relativo ad essi: il **timore dello scooping**, che qualcuno cioè ci rubi l'idea, faccia il nostro stesso esperimento, e poi pubblichi i risultati prima che riusciamo a farlo noi. Forse la parola "mito" che ho usato sopra è effettivamente un po' ingenerosa: la probabilità che lo scooping si verifichi non è ovviamente nulla, e alcuni colleghi giurano che a loro sia effettivamente capitato. Però credo concorderemo tutti che questa probabilità è comunque piuttosto bassa: lo *scooper* dovrebbe trovare il nostro materiale/poster/pre-print¹², e poi essere così veloce da preparare l'esperimento, condurlo, analizzare i dati, scrivere e pubblicare il suo articolo prima che lo facciamo noi. Con il vantaggio temporale che chi mette a disposizione il materiale ha per definizione, credo che una cosa del genere possa verificarsi davvero solo in casi del tutto eccezionali. In compenso, proprio la pubblicazione del materiale —sia appunto esso un poster, un pre-print, o un set di stimoli— costituisce una protezione contro lo *scooping*: nel momento in cui qualcuno riuscisse davvero a essere così tanto più veloce di noi e pubblicasse il suo articolo senza darci riconoscimento, avremmo proprio la pubblicazione stessa del materiale a dimostrare che la paternità intellettuale su quell'idea spetta a noi. E pensate invece a quanti effetti positivi potrebbe avere un *early sharing* dei nostri risultati! Se veicolata nel modo giusto (i social qui possono essere una risorsa molto preziosa), la pubblicazione di un pre-print, ad esempio, può davvero aumentare significativamente l'impatto di un articolo. Inoltre, può generare commenti preziosi da parte dei lettori, che miglioreranno l'articolo e addirittura, in qualche caso, aumenteranno le nostre chance di pubblicazione, ad esempio rendendoci consapevoli di problemi prima che li notino i reviewers, o suggerendoci analisi aggiuntive che chiariscano il messaggio teorico dell'articolo.

Nel complesso, mi pare dunque di poter dire che i potenziali benefici di una condivisione spinta del materiale di ricerca siano ben maggiori dei suoi rischi, soprattutto se pesati per la loro probabilità di occorrenza. **Essere giusti aiuta a realizzare le nostre legittime aspirazioni di carriera**, invece di contrastarle. In

¹² Il manoscritto così come spedito per la pubblicazione alla rivista prescelta, che noi tendenzialmente rendiamo pubblico al momento della *submission*.

questo senso, il concetto di *open science* può forse essere allargato a intendere una più generale *fair science*, che non soltanto garantisce la miglior qualità possibile per i nostri risultati sperimentali, ma rende anche più trasparente l'evoluzione delle carriere degli scienziati e porta verso una competizione più "generosa" e collaborativa rispetto a quella attuale, che spesso è oppositiva e platealmente anti-economica. Non capita tanto spesso che il mondo più giusto sia anche quello in cui è più bello vivere: dovremmo approfittarne, non credete?

Bibliografia

- Begg, C. B., & Berlin, J. A. (1988). Publication Bias: A Problem in Interpreting Medical Data. *Journal of the Royal Statistical Society. Series A*, 151(3), 419.
- Chambers, C. D. (2013). Registered Reports: A new publishing initiative at Cortex. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 49(3), 609–610.
- Cumming, G., & Calin-Jageman, R. (2016). *Introduction to the New Statistics: Estimation, Open Science, and Beyond*. Routledge.
- Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian Inference for Psychology. *Psychonomic Bulletin & Review*, 25(1), 5–34.
- Kerr, N. L. (1998). HARKing: hypothesizing after the results are known. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, 2(3), 196–217.
- Open Science Collaboration. (2015). PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641.