

La scienza ai tempi dei Social Network

Gli archivi di dati
condivisi alla base della
collaborazione
scientifica on-line

Autore: Matteo Massicci
Relatore: Alessio Cimarelli



Scuola Superiore di Studi Avanzati di Trieste
Master in Comunicazione della Scienza "Franco Prattico"
Anno accademico 2014/2015

Sommario

Abstract	4
Introduzione	5
Obiettivi	16
Capitolo 1	18
Scienza 2.0	22
Le piattaforme della Scienza 2.0	24
Il Polymath Project	28
La storia del progetto Genoma	33
Intervista a Edoardo Boncinelli	37
I servizi di open access dell’Agenzia Spaziale Europea e il caso Survey Lab	40
Intervista a Pier Bargellini (progetto Sentinelles dell’ESA)	43
Intervista a Silvia Scifoni (progetto Edi.Modi)	45
Il caso di Ebola Research Database	46
Intervista a Ilaria Capua	54
A ogni problema la giusta competenza: il caso di Mathesia	57
Capitolo 2	64
I progetti collaborativi sviluppati online	67
Serendipità vs serendipità pianificata	71
Evoluzione dei dati sperimentali, interpretazione e collettivizzazione del lavoro scientifico	77
La valutazione della produzione collaborativa online	82
I criteri per la contribuzione sui social media e l’amplificazione dell’intelligenza collettiva	86
Conclusioni	90
Scienza post-accademica e Scienza 2.0, risposte necessarie all’evoluzione del lavoro di ricerca	90
Le caratteristiche della Scienza 2.0: intelligenza collettiva, riorganizzazione dell’attenzione e serendipità pianificata	92

Infrastruttura e metodologia della scienza contemporanea: il Web 2.0 e i suoi servizi	_ 97
Note	_____ 100
Bibliografia	_____ 102
Sitografia	_____ 105
Elenco Figure	_____ 106

Abstract

Come cambia la conoscenza scientifica con l'evolversi della tecnologia? L'uso degli strumenti digitali influisce enormemente sia sulla capacità di osservare e studiare in dettaglio la complessità dei fenomeni naturali, sia sulla possibilità di interpretare e dare un senso a questa analisi, altrimenti troppo dispendiosa a causa della mole di dati generati. Nel tempo il lavoro di interpretazione della vasta informazione derivante dalla ricerca scientifica di base si è dunque rivelato troppo problematico da affrontare in assenza di strumenti specifici che lo supportassero. È proprio questo il ruolo delle banche dati digitali e dei repository online, strumenti destinati ad accogliere i dati dell'osservazione sperimentale. Ma è davvero solo questa la loro funzione? I database sono diventati oggetti indispensabili all'interno della prassi scientifica: l'informazione che sono in grado di veicolare e la rete di contatti tra scienziati che generano li rendono imprescindibili all'interno dei grandi progetti scientifici o delle infrastrutture di ricerca. I database assolvono a molte delle funzionalità di un Social Network: gli utenti sono sia fruitori che creatori di contenuti e all'interno di essi c'è uno scambio quasi istantaneo di quegli stessi contenuti tra ricercatori di tutto il mondo. Attraverso l'analisi di esempi esplicativi tratti da diversi campi di indagine scientifica, il seguente elaborato esplorerà la tesi secondo cui i servizi di repository online ricoprono un ruolo centrale all'interno dell'odierna pratica scientifica, soprattutto per quelle discipline che richiedono l'intersezione di competenze differenti, e dimostrerà come la Scienza 2.0, e cioè il sistema di produzione della conoscenza scientifica supportato dall'utilizzo dei servizi del Web 2.0, non sia solo una mera speculazione sulle possibili trasformazioni a cui andrà in contro la ricerca scientifica nel prossimo futuro, ma una metodologia già diffusa e pienamente attestata.

Introduzione

Al fine di descrivere le trasformazioni subite dai rapporti e dalle gerarchie all'interno dei gruppi sociali a seguito della comparsa dei nuovi media, nel 2001 i sociologi canadesi Barry Wellman e Anabel Quan-Haase coniarono il termine "Hyperconnectivity". La parola fa riferimento al contemporaneo utilizzo di molteplici strumenti e sistemi di comunicazione: mail, strumenti di messaggistica istantanea, telefono, il tradizionale contatto frontale e, oggi, i servizi d'informazione forniti dal Web 2.0 come i Social Network. Secondo Wellman, la pervasività del Web nelle nostre vite e nel campo della comunicazione dimostrerebbe come le dinamiche e i processi che contraddistinguono i rapporti interpersonali si siano spostati da un ristretto piano fisico, determinato dalla contiguità di un individuo alla propria comunità, a uno esteso e pienamente interconnesso, il Web, in cui ogni soggetto è in grado di creare il proprio network sociale personalizzato e virtuale senza limitazioni di tipo geografico o temporale¹. Ma qual è il ruolo svolto dalla *iperconnettività* garantita dal Web, e soprattutto dai social network digitali, nella scienza contemporanea? Che posizione occupano i database e i repository condivisi in questi network di uomini e servizi automatici? Quali sono le difficoltà che, all'interno dell'impresa scientifica, l'utilizzo di questi strumenti consente di risolvere, e quali problematiche invece solleva?

Una prima risposta a queste domande prende le mosse dall'assunto secondo cui due persone impegnate a risolvere un problema hanno maggiori probabilità di avere successo di quante ne abbia un singolo individuo. Per cui dedurre che il Web, una vasta rete in grado di collegare miliardi di teste, è meglio di due, dieci o,

¹ 'Physical Place and Cyber Place: The Rise of Networked Individualism'. International Journal of Urban and Regional Research. (2011)

meglio ancora, di un intero palazzo di cervelli pensanti. Ma non ci troviamo di fronte a una trasformazione di tipo unicamente quantitativo. Come ha affermato il filosofo Silvano Tagliagambe, la rete rappresenta e caratterizza, allo stesso tempo, sia la forma del sapere odierno che il nuovo paradigma esplicativo della conoscenza: “dal paradigma ad albero, fondato sull’idea di radice, e quindi di base, di fondamento, che poi si sviluppava in verticale - pensiamo a tutta l’importanza che ha avuto nella matematica, nella fisica, dall’inizio del secolo fino agli anni ‘50, il problema dei fondamenti, il problema delle basi - si è arrivati al paradigma a Rete (orizzontale, n.d.r.) attraverso cui la conoscenza si profonde. Oggi la metafora esplicativa della conoscenza è la Rete, in cui non esiste una base, non esiste un centro, ma in cui importanti sono certamente i nodi della Rete e ancor più importanti sono le maglie, i link, le interconnessioni tra questi nodi. Quindi diciamo che la stessa Rete è diventata una nuova metafora esplicativa del proprio sapere, sostituendo la metafora dell’albero”². Della stessa idea sembra anche essere il filosofo americano del Web David Weinberger, uno degli autori del Cluetrain Manifesto³, documento del 1999 recentemente aggiornato in cui, attraverso 95 tesi, si analizza l’impatto di Internet sui mercati e sulle organizzazioni e in cui si suggeriscono i cambiamenti da adottare in risposta alle trasformazioni sociali ed economiche promosse dal Web. Per Weinberger “una rete di conoscenze genera una conoscenza più grande della conoscenza di ciascuna persona connessa”: l’entità più intelligente in una stanza in grado di rendere fattiva una perfetta connessione tra i suoi occupanti risulterebbe perciò la stanza stessa. La conoscenza si manifesterebbe così come proprietà emergente del Web⁴.

² ‘Rete, paradigma della conoscenza’. La Repubblica (9 marzo 1998)

³ <http://cluetrain.com/>

⁴ ‘Il disaccordo che accresce la conoscenza’, Il Sole 24 Ore (4 novembre 2012)

Le riflessioni dei due filosofi in merito alla struttura della conoscenza nell'era digitale sono in buona misura legate alla natura di Internet: in quanto rete di più calcolatori e di utenti interconnessi tra loro, il Web può essere infatti considerato in tutto e per tutto come il più grande sistema complesso mai realizzato dall'uomo. In fisica, per sistema complesso si intende un sistema in cui le singole parti sono interessate da interazioni locali, di breve raggio d'azione, che provocano cambiamenti nella struttura complessiva⁵. Da ciò consegue che le proprietà emergenti di un sistema complesso non possono essere valutate sommando i contributi apportati dalle singole parti e che la conoscenza scaturita dal Web crescerà all'aumento del numero e della qualità delle connessioni tra gli utenti della rete.

Il passaggio da un paradigma della conoscenza di tipo verticale a uno orizzontale e supportato dal Web non è però privo di difficoltà. Nonostante la conoscenza precedente alla comparsa di Internet non fosse in grado di rendere conto dell'inclusività e della dinamicità dell'informazione garantite dalla struttura della rete digitale, essa era tuttavia in grado di generare un sapere stabile e organico, ossia validato da fondamenti concettuali, come ricordato da Tagliagambe, e dalle competenze degli esperti. Al contrario, l'informazione veicolata dal Web rispecchia la natura complessa di questo medium, presentandosi spesso disorganica, frammentata e priva delle necessarie conferme. A tal proposito, sono state molte le voci autorevoli che nel recente passato hanno espresso le loro perplessità nei confronti di Internet e della sua funzione di mezzo di supporto privilegiato per la conoscenza. Una su tutte è quella dello scrittore americano Nicholas Carr che, nel suo saggio "Internet ci rende stupidi?"⁶ del 2011, tenta di dimostrare come la frammentata disposizione delle informazioni presenti nel Web sia di ostacolo allo

⁵ https://it.wikipedia.org/wiki/Sistema_complesso

⁶ 'Internet ci rende stupidi?', Raffaello Cortina Editore, (2011)

sviluppo di una conoscenza stabile, poiché inficerebbe la capacità di focalizzare l'attenzione sull'argomento delle nostre riflessioni. A questa complessa distribuzione dei contenuti online si affianca inoltre la dispersione del sapere prodotto dagli utenti della rete. Per rendere davvero effettiva la conoscenza di cui parla Weinberger è infatti necessario realizzare connessioni volte a raggiungere uno scopo comune e capaci di amplificare l'intelligenza inespressa di Internet, rappresentata dalle competenze latenti che nel Web dimorano⁷.

Anche - e soprattutto - in ambito scientifico l'immensa mole di dati generati dagli strumenti digitali e veicolata dal Web ha spinto i ricercatori a interrogarsi su vantaggi e svantaggi dell'utilizzo di queste tecnologie all'interno dell'impresa scientifica. Secondo quanto sostenuto da Tony Hey, Stewart Tansley e Kristin Tolle nel libro 'The Fourth Paradigm. Data-Intensive Scientific Discovery'⁸ ("Il Quarto Paradigma. La scoperta scientifica fondata su grandi volumi di dati"), la scienza odierna si troverebbe infatti alle porte di una svolta epocale. La rivoluzione informatica mette oggi a disposizione della ricerca una base empirica senza precedenti. Per gli autori di 'The Fourth Paradigm' il salto di quantità nella produzione di informazione è tale da costituire di per sé un salto di qualità. L'idea è che l'informazione raccolta in quantità mastodontica possa trasformarsi facilmente, quasi automaticamente, in nuova conoscenza. Anche se, a prima vista, potrebbe sembrare che con la possibilità di avere a disposizione una grande mole di dati si possa realizzare il progetto induttivista di Francis Bacon⁹, secondo cui la scienza dovrebbe partire solo dall'analisi dei dati su cui costruire le teorie, la gestione dell'enorme quantità di informazioni scientifiche presenti sul Web

⁷ 'La stanza intelligente. La conoscenza come proprietà della rete', Codice Edizioni. (2011)

⁸ Tony Hey, Stewart Tansley, Kristin Tolle 'The Fourth Paradigm: Data-intensive Scientific Discovery', (2009)

⁹ Pietro Greco, 'Il quarto paradigma', Micron n. 23, (dicembre 2012)

orientata alla creazione di nuova conoscenza pone serie criticità: insiemi così estesi di dati includono informazioni di tipo differente e ogni progetto di ricerca ha necessità di usare solo informazioni utili ai propri obiettivi. Quindi ciascuno di questi insiemi sarebbe pienamente efficace se fosse suddiviso e reso disponibile ad altri gruppi ed elaboratori. Tutto ciò implica che alla fine di queste fasi di predisposizioni, invio e lavorazione dei dati in ingresso, i risultati debbano poi essere a loro volta riassociati, probabilmente sulla base di una teoria, in modo da costituire un quadro di riferimento omogeneo da cui trarre considerazioni su andamento ed esito dell'esperimento. In altre parole, bisogna garantire che i dati possano essere usati più volte e interpretati di volta in volta in base al contesto di analisi. Quindi a un aumento del dettaglio e della quantità delle osservazioni nella ricerca odierna corrispondono una crescita del tempo necessario alla loro analisi, con il rischio di trascurare informazioni rilevanti su ciò che è stato acquisito, e una difficoltà nel gestire i dati in maniera tale che questi ultimi rendano possibile la riproducibilità degli esperimenti che li hanno prodotti e lo scambio di informazioni tra team scientifici differenti, che non sempre condividono metodologie e linguaggi di lavoro comuni. A ben vedere, le difficoltà risultanti dall'utilizzo degli strumenti digitali e del Web nell'impresa scientifica possono essere accomunate a quelle emerse dall'analisi della rete come supporto privilegiato per la conoscenza in generale, e cioè frammentazione dei contenuti e delle competenze in grado di interpretarli.

Ma se il Web, concepito come paradigma della conoscenza, e gli strumenti digitali prestano il fianco all'insorgenza di simili difficoltà, Internet fornisce anche servizi di aggregazione e di collaborazione sociale in grado di superare le problematiche della dispersione dell'informazione e delle competenze. Gli strumenti in questione sono repository e database, in quanto capaci di favorire la costituzione di reti sociali online e la collaborazione tra i professionisti della scienza, e costituiscono il soggetto di studio del seguente elaborato. Secondo i due esperti di Social Media

marketing Andreas Kaplan e Michael Haenlein¹⁰, i Social Media possono essere definiti come “un gruppo di applicazioni basate sul Web e costruite sui paradigmi (tecnologici ed ideologici) del Web 2.0 che permettono lo scambio e la creazione di contenuti generati dagli utenti”. Sotto questa definizione rientrano sei tipologie differenti di servizi online:

- blog e microblog, come Twitter;
- siti di social networking, come Facebook;
- mondi virtuali di gioco, come World of Warcraft;
- mondi virtuali sociali, come SecondLife;
- progetti collaborativi, come Wikipedia;
- content communities - community che condividono materiale multimediale -, come Youtube.

Nel prosieguo della tesi si dimostrerà che l'utilizzo di questi strumenti per la produzione di nuovo sapere, e in particolare dei servizi online maggiormente adatti a supportare e a favorire i progetti collaborativi, è in grado di realizzare l'aggregazione dei contenuti e delle competenze presenti online. Pur non essendo ancora una pratica codificata, l'utilizzo dei servizi sociali del Web 2.0 all'interno dell'impresa scientifica non può essere considerato una novità. Lo spostamento di buona parte delle pratiche scientifiche, almeno quelle che riguardano la sfera sociale del lavoro dei ricercatori, dai laboratori e dal mondo accademico al cyberspazio ha spinto molti osservatori a descrivere quello che sembra l'avvento di una Scienza 2.0, con tutto quello che consegue in termini di ridefinizione del processo di produzione e di valutazione della scoperta scientifica. Ciò non deve però sorprendere. Le ragioni delle trasformazioni in corso trovano la loro giustificazione nello stretto legame esistente tra il sistema scienza e i Social

¹⁰ 'Users of the world, unite! The challenges and opportunities of social media', Business Horizons n. 53. (2010)

Network. Infatti, al di là dell'evidenza storica che vede la nascita di Internet legata a esigenze prettamente scientifiche, in un primo momento "come strumento tecnologico in grado di collegare le comunità scientifiche diffuse sul territorio statunitense" e successivamente come mezzo di "condivisione ipertestuale di testi digitali presso il CERN di Ginevra"¹¹, il Web e - nello specifico - i Social Network condividono con la scienza elementi e obiettivi di tipo sia culturale che sociologico. Dal punto di vista culturale l'etica delle reti digitali e dell'informatica si pone infatti in continuità con quella del ricercatore, la quale per Robert K. Merton si contraddistingue sulla base di cinque principi: comunitarismo, universalismo, disinteresse, originalità e scetticismo organizzato. Senza voler entrare ora nel dettaglio del pensiero mertoniano, a cui sarà riservata un'adeguata sezione del seguente lavoro, è indubbio che le prime due norme, che tentano di descrivere come i risultati dell'attività scientifica siano sempre prodotti sul presupposto che entreranno a far parte della conoscenza comune e che nessuna persona può essere esclusa, in assenza di criteri rilevanti, dal processo di ampliamento della cultura scientifica, abbiano avuto un ruolo fondamentale nell'avvio del movimento Open Access di libera circolazione dei paper scientifici online. Oltre ai fattori culturali, impresa scientifica e Social Network sembrano inoltre operare sulla base delle stesse modalità, o finalità - a seconda dei casi -, sociologiche. Se la scienza può essere definita come "progetto collettivo organizzato in comunità (le comunità scientifiche), caratteristica a cui i sociologi della scienza fanno riferimento a volte con il termine collettivizzazione (Ziman, 2000) e altre ancora con il concetto dei college invisibili (Crane, 1972), è evidente che la capacità dei Social Network di creare - nonostante le eventuali distanze fisiche che possono separare i ricercatori - comunità accomunate da obiettivi condivisi sia una discriminante fondamentale

¹¹ Davide Bennato, 'Comunità connesse. L'impatto dei social media sulla professione del ricercatore' in 'Scienza Connessa. Rete, media e social Network', a cura di Sveva Avveduto, Gangemi Editore. (2015)

nel comprendere le ragioni della diffusione di questi strumenti all'interno dell'impresa scientifica. Alla possibilità di operare secondo il modello della collettivizzazione della collaborazione scientifica è inoltre fortemente associata l'idea, legata alla norma mertoniana del comunismo, che il sistema scienza sia un sistema che comunica" (Bennato, 2015). È infatti possibile affermare che la norma del comunismo attiva un elaborato sistema di comunicazione, all'interno del quale i risultati della ricerca non hanno valore scientifico se non vengono registrati, riferiti, diffusi, condivisi e infine trasformati in proprietà comune. Tentando di analizzare i processi di trasformazione in atto all'interno della scienza e vista l'importanza attribuita all'aspetto comunicativo - che per compiersi non può esaurirsi all'interno della ristretta cerchia degli scienziati ma deve estendersi al rapporto con l'intera società -, John Ziman, nel libro 'La vera Scienza'¹², arriva a sostenere che Internet, in quanto mezzo inclusivo per eccellenza e ambiente privilegiato per lo stoccaggio della conoscenza, sia diventato vero "motivo sociologico" della nuova scienza contemporanea. Quindi, considerando la scienza sia come processo collaborativo che sistema di comunicazione è facile comprendere il motivo per cui i servizi di collaborazione online trovino sempre più spazio nella vita professionale degli scienziati. A testimonianza di questo fenomeno possono essere riportati i numeri della crescita di ResearchGate, indubbiamente il più diffuso Social Network dedicato alla scienza. Dal 2008, anno di fondazione del servizio, a oggi, la comunità di ResearchGate¹³ è infatti passata da 10 mila membri a 10 milioni membri e le pubblicazioni scientifiche che la piattaforma raccoglie sono più di 100 milioni.

Abbiamo visto come il fenomeno della diffusione dei Social Media nella prassi scientifica è motivato sia da ragioni di tipo epistemologico e metodologico che da

¹² 'La vera scienza', Edizioni Dedalo. (2002)

¹³ <https://www.researchgate.net/press>

evidenze culturali e sociologiche. Ma cosa contraddistingue la Scienza 2.0, e quali conseguenze ha sul sistema scienza e sul modo di fare ricerca? Una prima considerazione in merito ai cambiamenti indotti dalla comparsa di questi strumenti all'interno della prassi scientifica riguarda l'interpretazione della mole di dati sperimentali oggi a disposizione. L'attività di data mining, e cioè la selezione, valutazione e modellazione di grandi quantità di dati al fine di estrarne una conoscenza - che rappresenta oggi giorno una cospicua parte del lavoro dei ricercatori - non potrebbe essere condotta in maniera proficua ed efficace senza strumenti online in grado di raccogliere e rendere intellegibile, disponibile e condivisibile tutta questa informazione. Assolvendo per loro stessa costituzione a simili funzioni, i servizi di archiviazione e condivisione dei dati, il cuore delle risorse del Web 2.0, sono perciò diventati elementi imprescindibili delle infrastrutture di ricerca. Secondo quanto sostenuto del fisico e studioso del Web Michael Nielsen¹⁴, il vantaggio apportato dall'utilizzo dei Social Network nella quotidianità del lavoro scientifico è il risultato della capacità di questi strumenti di amplificare l'intelligenza collettiva della rete. Concentrando infatti le conoscenze degli utenti - in questo caso degli scienziati -, in ambienti digitali destinati alla ricerca delle soluzioni per specifici problemi, i Social Media sono in grado di far emergere i meccanismi di riorganizzazione delle competenze e di serendipità pianificata - concetto che verrà chiarito nel secondo capitolo - che consentirebbero di diminuire il tempo necessario al processo di scoperta e di favorire la produzione di nuova conoscenza scientifica, sopperendo alla dispersione delle conoscenze e delle informazioni di cui si è discusso nei primi paragrafi. I meccanismi individuati da Nielsen sembrano fornire la loro maggiore incisività nell'ambito di quei settori scientifici che, a causa della complessità dei fenomeni studiati, sono a forte propensione transdisciplinare e che sembrano quindi maggiormente interessati dal

¹⁴ 'Le nuove vie della scoperta scientifica', Einaudi. (2012)

fenomeno della collettivizzazione del lavoro. Tra queste possiamo annoverare tutte quelle discipline che rientrano nelle due più ampie categorie di scienze ambientali e della vita. Tali campi di ricerca, non potendo sempre utilizzare un approccio riduzionista per il superamento di difficoltà teorico pratiche, necessitano infatti di strumenti in grado di creare reti stabili tra comunità scientifiche differenti capaci di valutare da vari punti di vista i dati sperimentali e di trovare soluzioni innovative o di aprire la strada ad approcci metodologici non ancora percorsi. Il concetto di transdisciplinarietà, così come concepito da Jean Piaget¹⁵, prende proprio atto delle trasformazioni indotte nel modo di fare scienza dalla complessità dei fenomeni studiati e dalla collettivizzazione del lavoro che ne deriva, e prospetta per la Scienza un futuro in cui la risoluzione dei problemi venga svolta da “un sistema totale senza confini stabili tra le discipline stesse”. Da queste considerazioni risulta chiaro che se la rete è, per Ziman, “motivo sociologico” della scienza contemporanea, i Social Network sono diventati ragione metodologica per la Scienza 2.0.

La seguente tesi si pone quindi come obiettivo quello di approfondire il ruolo svolta da database e repository, i servizi alla base delle risorse del Web 2.0, nella pratica scientifica odierna, facendo emergere come l’interazione tra i professionisti della scienza veicolata da questi strumenti sia centrale all’interno degli aspetti sociali che sottendono alla produzione di nuova conoscenza e verificando che il loro impiego si traduce in un più fondamentale supporto all’interpretazione dei dati sperimentali e all’individuazione di nuovi paradigmi o di nuovi filoni di ricerca. Tale indagine sarà condotta attraverso l’analisi di varie esperienze di collaborazione online di tipo scientifico. I casi studio affrontati consentiranno di delineare le caratteristiche di ciò che è stata definita intelligenza collettiva della

¹⁵ Fabio Marzocca, ‘Il nuovo approccio scientifico verso la transdisciplinarietà’, Edizioni Mythos. (ottobre 2014) http://ciret-transdisciplinarity.org/biblio/biblio_pdf/eBook_Transdisciplinarity.pdf

rete. Simili considerazioni consentiranno inoltre di proporre un'accezione forte del termine Scienza 2.0, che vede nell'utilizzo degli strumenti digitali di collaborazione un presupposto metodologico imprescindibile per la piena realizzazione della transdisciplinarietà a cui dovrebbe aspirare la scienza contemporanea.

Nel primo capitolo verrà proposto un quadro generale sulle ragioni e sulla vastità della diffusione dei servizi sociali del Web 2.0 all'interno della prassi scientifica stabilendo criteri tramite cui classificare le varie tipologie di questi strumenti. Verranno inoltre presentate esperienze collaborative sviluppate attraverso Social Network o altri servizi online in grado di implementare la connessione e la collaborazione tra scienziati al fine di facilitare la risoluzione di problemi e di promuovere la comparsa di nuovi approcci metodologici nella prassi scientifica o nell'applicazione pratica: a partire dalla descrizione di uno dei primissimi progetti scientifici della storia di Internet ad aver utilizzato la capacità dei Social Media di collettivizzare il processo scientifico, il progetto Polymath; per passare all'esposizione di esperienze di condivisione aperta dei dati scientifici nel settore biochimico (GenBank ed Ebola Research Database), supportate dall'intervista alla virologa Ilaria Capua, e in quello del *remote sensing*, (European Space Agency Earth Observation Data), con la testimonianza di alcuni responsabili di piattaforme di collaborazione e di un'azienda, "Survey lab", che a partire dall'accesso ai dati satellitari di osservazione della Terra è stata in grado di sviluppare un sistema innovativo di monitoraggio della stabilità degli edifici e delle infrastrutture; fino ad arrivare all'analisi di una piattaforma di crowdsourcing con finalità scientifiche e destinata a offrire soluzioni tecniche al settore privato, Mathesia.

Nel secondo capitolo verranno esaminate, senza tralasciare le difficoltà, le caratteristiche che consentono alle esperienze introdotte di far emergere i due meccanismi della serendipità pianificata e dell'aggregazione delle competenze, come descritti da Michael Nielsen, in grado di implementare e amplificare l'intelligenza collettiva degli scienziati presenti in rete.

La terza parte dell'elaborato sarà invece rivolta all'esposizione delle conclusioni derivanti dall'analisi condotta nel secondo capitolo e alla presentazione di quella che verrà definita come *strategia sociale* della Scienza 2.0.

Obiettivi

Attingendo alla letteratura di riferimento, delineeremo le specificità della scienza contemporanea: motiveremo, attraverso i fenomeni della specializzazione delle competenze e della crescita del numero e del dettaglio dei dati sperimentali, la collettivizzazione odierna del lavoro di ricerca e daremo conto di due accezioni di Scienza 2.0, proponendone una terza con implicazioni maggiori sul tipo di prassi alla base della produzione odierna della conoscenza scientifica e in grado di realizzare il necessario ricorso a collaborazioni sempre più estese tra i professionisti della scienza. La validità di questa tesi verrà corroborata grazie all'analisi di esperienze che fondano il loro successo proprio sull'utilizzo degli strumenti di collaborazione del Web 2.0 e a interviste a scienziati. Mostriamo infine che i Social Network si dimostrano particolarmente adatti alla prassi scientifica poiché sono in grado di amplificare l'intelligenza collettiva inespresa della Rete attraverso i due meccanismi di serendipità pianificata e di riorganizzazione dell'attenzione delle competenze presenti su Internet.

Capitolo 1

La somma delle conoscenze umane supera la capacità di ogni singolo individuo; e anche di migliaia di individui. Con la distruzione della nostra costruzione sociale, la scienza verrà spezzettata in milioni di parti. Gli individui conosceranno poco meno che una sfaccettatura di tutto ciò che c'è da sapere. Da soli saranno indifesi e inutili. Tali frammenti insignificanti di conoscenza non saranno trasmessi e si disperderanno attraverso le generazioni. Se però prepariamo un gigantesco sommario del sapere, esso non andrà mai perduto. Le generazioni successive costruiranno sopra queste basi, senza doverle riscoprire. (Isaac Asimov, "Trilogia Galattica")

Quando Isaac Asimov, alla fine degli anni '40, mette in bocca queste parole a Hari Seldon, personaggio centrale nell'evoluzione delle vicende riguardanti la nascita della Fondazione Galattica, lo fa non solo in virtù della sua maestria come scrittore di fantascienza ma anche perché, da scienziato, è perfettamente consapevole di quanto la comunicazione e la condivisione delle scoperte giochino un ruolo essenziale all'interno dell'impresa scientifica e per il progresso della scienza tutta. Lo scrittore di origini sovietiche, prima di essere considerato a giusto merito il padre indiscusso della fantascienza moderna, fu infatti biochimico e divulgatore di riferimento per un'intera generazione. Dopo oltre 70 anni, gli strumenti di condivisione della conoscenza scientifica si sono evoluti al di là delle più rosee aspettative di Asimov, tanto che oggi ci si interroga sulla validità dell'uso dei Social Media a supporto della ricerca. A dimostrazione delle trasformazioni succedutesi dai tempi di Asimov, basti pensare che lo scrittore, nel libro sopracitato, considera lo sfruttamento del binomio *biblioteca* ed *enciclopedia* come unico modello possibile di comunicazione e condivisione della conoscenza scientifica. Nonostante le differenze e le distanze tra la tecnologia conosciuta da

Asimov e quella oggi a disposizione, i principi e le prassi sociali della pratica scientifica sono rimasti per lo più immutati. Nella Trilogia Galattica, infatti, la soluzione dello psicostoriografo Hari Seldon alla possibile stagnazione e dispersione della cultura scientifica conseguente alla disgregazione dell'Impero Galattico viene trovata nella creazione di una Fondazione in grado di accogliere il maggior numero possibile di scienziati e di mettere questi ultimi nella condizione di proseguire nei rispettivi studi e di pubblicare, attraverso un'enciclopedia galattica, i propri lavori.

La collettivizzazione, così come viene definita da John Ziman, è una prassi naturale dell'impresa scientifica, l'attività umana più innovativa di tutte, ma è una tendenza che si è fatta sempre più evidente con l'avanzamento del progresso tecnologico¹⁶.

La messa a punto di strumenti sempre più sofisticati spinge la ricerca verso modalità di azione sempre più collettive.

La stessa tendenza si manifesta chiaramente nella crescente quantità di saggi scientifici prodotti da due o più autori. Il lavoro di gruppo, la comunicazione in rete ed altre modalità di collaborazione tra ricercatori specializzati non sono semplici mode alimentate dal buon esito dell'attuale considerazione elettronica globale. Essi sono le conseguenze sociali dell'accumulo di conoscenze e tecniche. La scienza è progredita a un livello in cui i suoi problemi ancora irrisolti non possono essere risolti da singoli che operano autonomamente.

Questa analisi si pone ancora una volta in continuità con le parole di Seldon/Asimov: la quantità delle conoscenze scientifiche è ormai divenuta tale da non poter permettere ai ricercatori di lavorare in maniera autonoma. Ma anche se Ziman non ritiene che la comunicazione in rete e di conseguenza i Social Network sia il motivo principale della tendenza alla collettivizzazione del lavoro scientifico, è chiaro che questi strumenti, vista la loro capacità di trascendere lo spazio fisico e

¹⁶ 'La vera scienza', Edizioni Dedalo. (2002)

di collegare le persone, siano strumenti funzionali all'implementazione di una simile prassi.

Il fenomeno della collettivizzazione non può però prescindere dal principio secondo cui il sistema scienza è un sistema che deve comunicare e rendere sempre disponibile i risultati ottenuti. La collaborazione tra scienziati deve infatti essere supportata da un'ampia base condivisa e accessibile di conoscenza. Tale aspetto è una delle tante conseguenze scaturite dalle norme che contraddistinguono la pratica scientifica descritte da Robert Merton¹⁷. Le ricerche di Merton riguardavano soprattutto gli aspetti organizzativi e funzionali della scienza in quanto istituzione capace di autoregolamentarsi. Quali valori e norme di condotta, si chiede Merton, garantiscono il funzionamento della scienza? La sua risposta a questa domanda si incentra attorno a quattro imperativi istituzionali¹⁸.

- A. Universalismo: asserzioni o risultati scientifici vengono giudicati indipendentemente da caratteristiche inerenti al soggetto che li ha formulati quali la classe, la razza, la religione.
- B. Comunitarismo: i risultati e le scoperte non sono proprietà del singolo ricercatore, ma patrimonio della comunità scientifica e della società nel suo complesso. Lo scienziato non ottiene riconoscimento per la propria attività se non rendendola pubblica e mettendola quindi a disposizione degli altri.
- C. Disinteresse: ogni ricercatore persegue l'obiettivo primario del progresso della conoscenza, ottenendo indirettamente il riconoscimento individuale.
- D. Scetticismo organizzato: ogni ricercatore deve essere pronto a valutare in modo critico qualunque risultato, inclusi i propri, sospendendo il giudizio definitivo fino all'ottenimento delle necessarie prove.

¹⁷ Merton Robert, K. Elinor, G. Barber, 'Viaggi e avventure della serendipity', Il Mulino. (2002)

¹⁸ Massimiano Bucchi, 'Sociologia della scienza', Nuova Informazione Bibliografica n. 3. (2004)

Nell'enunciare questi principi, Merton sottolinea a più riprese come essi vadano considerati validi dal punto di vista istituzionale e non dal punto di vista delle motivazioni individuali di ciascuno scienziato. La norma che ha maggiori implicazioni per l'analisi che si sta svolgendo è ovviamente quella del comunitarismo. Questo precetto, infatti, attiva un elaborato sistema di comunicazione, all'interno del quale i risultati di ricerca non hanno valore scientifico se non vengono registrati, riferiti, diffusi, condivisi e infine trasformati in proprietà comune tramite la loro pubblicazione formale. La conoscenza scientifica deve perciò essere pubblicabile. Da ciò consegue che gli strumenti di stoccaggio e comunicazione della conoscenza assumono un ruolo fondamentale all'interno della pratica scientifica. Tali considerazioni e i modelli contemporanei di comunicazione della scienza rendono ragione dell'interesse che gli strumenti online per la collaborazione suscitano all'interno della comunità scientifica. Un esempio interessante di classificazione delle forme di comunicazione scientifica è quello del modello elementare che prevede sei componenti fondamentali della comunicazione¹⁹.

- A. *Comunicazione intraepistemica* che avviene tra il ricercatore e i propri pari.
- B. *Comunicazione trans-epistemica* che si dipana tra ricercatori in ambiti diversi.
- C. *Comunicazione di "rete"* che si svolge con soggetti dell'indotto della ricerca.
- D. *Comunicazione sociale* che è il momento di dialogo con la società civile.
- E. *Comunicazione politica* i cui soggetti coinvolti sono i politici, legislatori e decisori.
- F. *Comunicazione generale* che si rivolge al pubblico nel suo complesso.

Questo e tutti gli altri modelli di classificazione scientifica hanno due punti in comune: il ricercatore che dialoga con i colleghi del proprio settore di ricerca e di

¹⁹ Bennato (2015)

altri settori, spesso appartenenti a discipline molto differenti. L'aspetto intradisciplinare e transdisciplinare della comunicazione assume un ruolo rilevante ai fini del seguente lavoro, in quanto l'impiego degli strumenti sociali in queste due particolari categorie di dialogo tra scienziati si dimostra un fattore necessario e integrante della metodologia scientifica, soprattutto all'interno di quelle discipline votate a un approccio fortemente pluralistico come le scienze della vita e quelle dell'ambiente, così come discipline umanistiche quali la storia e la sociologia.

Scienza 2.0

Abbiamo visto come l'integrazione dei Social Network e dei Social Media nella quotidiana pratica scientifica sia coerente sia con la tendenza alla collettivizzazione del lavoro scientifico, sia con le norme mertoniane che regolano l'atteggiamento istituzionale degli scienziati e, di conseguenza, l'aspetto comunicativo che contraddistingue la scienza.

Coscienti della diffusione degli strumenti sociali online di tipo scientifico, gli osservatori e i sociologi della scienza hanno coniato il termine *Scienza 2.0* per cercare di cogliere i cambiamenti occorsi nel lavoro di ricercatori a partire dalla comparsa degli strumenti e delle logiche tipiche del Web 2.0. Nonostante questo termine non sia stato ancora codificato nella letteratura specialistica e non sia possibile rintracciare il momento in cui abbia avuto inizio il dibattito in merito al suo significato, è comunque possibile descrivere due diverse accezioni con cui questa etichetta viene utilizzata: una più edulcorata, che presuppone che Social Network e Social Media svolgano un ruolo solo superficiale all'interno dell'impresa scientifica - supportando e favorendo unicamente la partecipazione,

collaborazione e comunicazione degli scienziati - e una più forte, che vede nell'utilizzo degli strumenti partecipativi del Web 2.0 una nuova via alla metodologia scientifica in virtù della quantità di dati che questi ultimi consentono di veicolare²⁰.

Secondo l'accezione debole, per scienza 2.0 si intende l'uso dei social media e delle tecnologie del Web partecipativo per aiutare a semplificare il lavoro del ricercatore scientifico. Entrano in questa categoria le riflessioni sull'uso del blog da parte del ricercatore, lo sviluppo di servizi di Social Network per la creazione di reti di collaborazione scientifica o l'uso di piattaforme progettate per aumentare la circolazione delle attività scientifiche.

Per quanto riguarda l'accezione forte, essa sostiene che l'ingresso dei media partecipativi del Web 2.0 non solo ha un impatto notevolissimo sulla pratica della ricerca scientifica, per quanto riguarda relazioni e rapporti, ma in alcuni casi potrebbe modificare lo stesso metodo scientifico a causa dell'impressionante quantità di dati resi disponibili da diversi servizi, portando così a sostituire il processo di ipotesi / verifica dei dati con le tecniche di data mining o analisi massiva dei dati.

A queste due concezioni della Scienza 2.0 ne verrà affiancata una terza. Tale proposta implicherà che l'impatto forte dei media digitali, quello che attiene a una vera e propria trasformazione del processo scientifico, sia determinato proprio dalla capacità di questi strumenti di implementare l'intelligenza collettiva, e cioè quei comportamenti intelligenti che emergono in maniera non coordinata dall'interazione di un notevole numero di persone, attraverso tecnologie opportunamente progettate della rete grazie alla loro funzione di aggregatori e di facilitatori della partecipazione degli scienziati. Una proposta che ritiene quindi inseparabili le concezioni analizzate. Verrà inoltre dimostrato che la via presentata a una Scienza 2.0 è già una realtà e una concezione indispensabile all'interno delle infrastrutture di ricerca che fanno capo a quelle discipline scientifiche che si

²⁰ Maness J, 'Library 2.0 Theory: Web 2.0 and its implications for libraries', *Webology* vol.3, n. 2. (2006);

occupano di fenomeni complessi, la cui valutazione e interpretazione dipende da una forte partecipazione di comunità scientifiche differenti.

Le piattaforme della Scienza 2.0

È difficile stilare una classifica delle tipologie di piattaforme di ispirazione scientifica. Ciò dipende dal fatto che, il più delle volte, i servizi scientifici del Web 2.0 mettono a disposizione risorse molteplici, non sempre riconducibili a una sola funzione. Nuove applicazioni, strategie commerciali, uso di mashup rendono complicata questa classificazione. È tuttavia possibile orientarsi in questo panorama effettuando una valutazione sulla funzionalità generale che le piattaforme svolgono, ovvero rispetto al modello “comunicativo / relazionale” che esse esprimono e non ai specifici settori disciplinari a cui fanno riferimento.

La categoria maggiormente rappresentata tra le piattaforme della Scienza 2.0 è senza dubbio quella dei Social Network. Le ragioni di questa presenza dipendono dalla capacità di questi strumenti di creare reti di collaborazione tra ricercatori di una stessa disciplina, o discipline differenti, che lavorano sugli stessi problemi. L'esempio più famoso di Social Network pensato per aiutare e sviluppare la collaborazione tra ricercatori è ResearchGate²¹, progettato per consentire le relazioni fra ricercatori appartenenti a tutte le categorie senza nessuna restrizione.

Una seconda categoria di applicazioni online è quella che si basa sulle logiche di videosharing o photosharing. “Come suggerisce il loro nome, sono piattaforme pensate per la condivisione di video e/o fotografie, come - rispettivamente - Youtube o Flickr” (Bennato, 2015). Le piattaforme di questo genere progettate per i

²¹ <https://www.researchgate.net/>

ricercatori hanno come scopo quello della condivisione di materiali audiovisivi e fotografici o la condivisione di importanti risultati di studi e ricerche rappresentati in maniera grafica. In questo specifico settore, uno degli esempi più rappresentativi è la piattaforma DnaTube²². Con una comunità di oltre 80 mila membri, appartenenti per lo più al mondo degli enti di ricerca, delle università, delle scuole, questo sito si rivolge - a differenza di quanto suggerisce il nome - a studiosi di diversi settori, con particolare attenzione ai biologi, perché è questo il target di riferimento, e agli studenti universitari, in quanto il portale è utilizzato come supporto rivolto all'educazione da parte delle università.

Oltre a quelle appena presentate, che sono esperienze pensate come versione specializzata di servizi che possono essere considerati generalisti, esistono anche tipologie di strumenti appositamente progettati per il lavoro del ricercatore. Tra di essi possiamo annoverare i papers manager, piattaforme utilizzate per gestire la letteratura scientifica di riferimento per determinate categorie scientifiche, semplificandone l'organizzazione (*personal bibliographic management*), il riferimento bibliografico o le citazioni (*reference/citation management software*). È chiaro che, a parte strumenti strettamente mirati come Google Scholar, simili applicazioni possono essere rintracciati all'interno di piattaforme più generaliste quali ResearchGate. Uno strumento che può però essere associato a questa categoria è Mendeley²³: un software che si installa sul computer e automaticamente si collega anche a un Social Network. Il software installato archivia e classifica in maniera semi-automatica gli articoli scientifici presenti in una particolare cartella del computer dell'utente, estraendone automaticamente anche i riferimenti bibliografici per semplificare la scrittura degli articoli scientifici. Ha anche una serie di strumenti per lavorare direttamente sui file, spesso pdf, delle riviste

²² <http://www.dnatube.com/>

²³ <https://www.mendeley.com/>

scientifiche che possono essere annotati, commentati e condivisi con membri del gruppo di ricerca. Il Social Network a cui è connesso invece, svolge una funzione di connettività basata su articoli e interessi di ricerca. In pratica a partire dalla propria collezione di paper scientifici suggerisce altri paper simili, colleghi che stanno lavorando su simili temi e avere accesso ovunque alle proprie annotazioni svolte sugli articoli.

Continuando nella descrizione delle piattaforme progettate per i ricercatori, troviamo quelle per la gestione della ricerca (*research management*), programmate per l'amministrazione digitale dei flussi di lavoro di un esperimento o di un progetto di ricerca. Come emerso in precedenza, anche in questo caso, e a prova della tesi che si vuole dimostrare, sono molto rappresentate le scienze della vita quale la biochimica. "Lablife²⁴ è una piattaforma che contiene al suo interno diversi tool molto utili per il ricercatore come la possibilità di gestire il database dei propri reagenti, appunti sui protocolli di ricerca, un inventario dei prodotti da acquistare e - ovviamente - un archivio dei dati sperimentali ottenuti." (Bennato, 2015)

Un'ultima categoria di servizi della Scienza 2.0 è quella che riguarda il supporto condiviso al processamento dei dati (*cloud computing*). Per far fronte all'aumento dei dati sperimentali gli scienziati sono costretti a munirsi di notevoli risorse hardware e software in grado di supportare la computazione. Tuttavia, simili strumenti sono spesso software proprietari che devono essere implementati su calcolatori adeguati o che devono essere scritti di volta in volta per far fronte alle esigenze di elaborazione specifiche e rappresentano generalmente un'ingente voce nell'ambito della spesa per la ricerca. Nel tentativo di ridurre i costi e facilitare il lavoro di data mining, molti enti e istituti di ricerca pubblici hanno sviluppato portali che offrono risorse computazionali online in condivisione. All'interno di questi ambienti digitali, comunemente dotati anche di servizi di repository, gli

²⁴ <https://www.lablife.org>

utenti possono lavorare sui dati senza la necessità di doverli materialmente scaricare, evitando di gravare su infrastrutture locali di archiviazione e computazione. In molti casi questi servizi rendono inoltre completamente visibili e disponibili le attività di tutti gli utenti e i risultati ottenuti. Un esempio di questa categoria di risorse è rappresentato da INDIGO²⁵ (INtegrating Distributed data Infrastructures for Global ExpLOitation). INDIGO è un progetto finanziato nell'ambito del programma europeo Horizon 2020 della Commissione Europea e coordinato a livello continentale dall'Istituto Nazionale di Fisica Nucleare (INFN) e si comporrà di un'infrastruttura di calcolo distribuita e condivisa. Obiettivo del progetto è quello di rispondere allo stesso tempo alle esigenze di calcolo, elaborazione e archiviazione dati di ricercatori di discipline molto diverse, attraverso l'utilizzo comune di funzionalità avanzate fornite dalla piattaforma. INDIGO si rivolge in particolar modo ai grandi centri di elaborazione dati e alle applicazioni scientifiche sviluppate da esperimenti che sentono maggiormente la necessità di fare rete, sviluppare e utilizzare servizi innovativi e accedere alle risorse messe a disposizione in cloud. Questa stessa logica sottende ai servizi commerciali di clouding forniti da Amazon, Google o Microsoft.

La classificazione appena compiuta dei servizi scientifici del Web 2.0 deve però tener conto di alcune considerazioni sull'etica e sulle logiche della rete che fanno parte del DNA di tutti gli strumenti presentati. L'open access e lo scientific blogging - componenti, queste sì, ampiamente studiate e codificate dalla letteratura sui new media scientifici - devono infatti essere considerati come momenti imprescindibili della progettazione e dell'utilizzo dei Social Network per il lavoro scientifico. Senza queste componenti sarebbe impossibile pensare di far emergere l'intelligenza collettiva inespressa della Rete. L'idea di affidare la risoluzione dei problemi scientifici alle competenze presenti sul Web fa parte della concezione

²⁵ <https://www.cnaf.infn.it/indigo/>

della scienza 2.0 e della strategia partecipativa di social networking per la risoluzione dei problemi scientifici, in parte mutuati - o comunque coevi - ai progetti di crowdsourcing che si svolgono sulla Rete: un esempio su tutti è quello che riguarda il fenomeno della Citizen Science.

Nei prossimi paragrafi approfondiremo esperienze significative di collaborazioni condotte su servizi tecnologicamente e temporalmente antecedenti a quelli descritti, ma comunque in grado di supportare le logiche della Scienza 2.0, tentando di evidenziare gli effetti risultanti dalla comparsa spontanea delle strategie partecipative che caratterizzano invece i servizi scientifici del Web 2.0. La selezione dei casi di studio che andremo ad affrontare è stata quindi in parte basata su esperienze già istituzionalizzate ma in grado di mostrare, sia in maniera voluta che come conseguenza indiretta del loro utilizzo, una precisa propensione allo sfruttamento dell'intelligenza collettiva della Rete. Vedremo inoltre come nel campo delle scienze biologiche e ambientali la metodologia scientifica non possa prescindere da tali strumenti - di cui Social Media e Social Network sono una naturale evoluzione - che oggi costituiscono in tutto e per tutto una parte indispensabile delle infrastrutture di ricerca e che sono spesso la ragione della comparsa di soluzioni inattese a problemi teorici o pratici di natura scientifica.

Il Polymath Project

Il primo Polymath Project nasce su iniziativa del matematico medaglia Fields Tim Gowers. Nel gennaio del 2009 Gowers decide di usare il suo blog per sottoporre ai suoi colleghi matematici un esperimento sociale molto insolito. Scelse un importante problema matematico complesso e ancora irrisolto che, scrisse, gli "sarebbe tanto piaciuto risolvere". Ma invece di occuparsi in prima persona della

risoluzione del problema, decise di affrontarlo pubblicamente, usando il blog per postare idee e progressi. Il fattore degno di nota di tutta la vicenda fu però che Gowers invitò tutti i suoi lettori a confrontarsi con lui e a esprimere le loro idee.

La proposta che dà il via a questa sperimentazione, che verrà successivamente chiamata Polymath1 dalla comunità Polymath, è quella di trovare una nuova prova combinatoria del teorema di Hales-Jewett sulla densità. L'assunto di questo teorema è di facile comprensione. Per capire cosa dice il teorema ci avvaliamo della griglia di nove caselle (tre per tre) disegnata sotto.

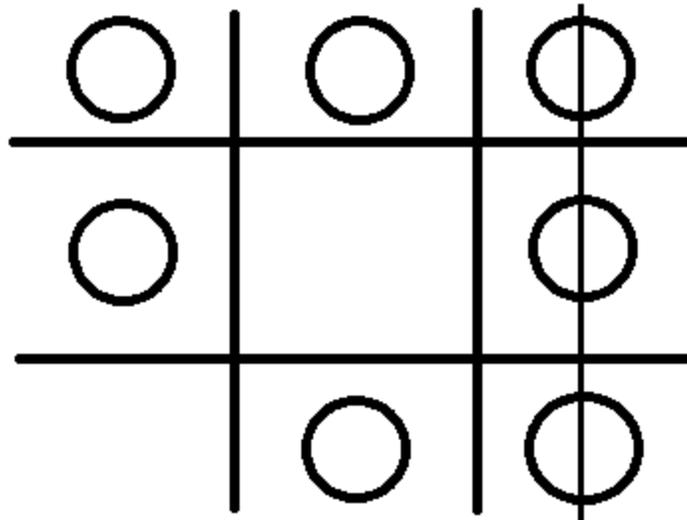


Fig. 1

I sette punti disegnati in questa maniera all'interno delle caselle scelte consentono di tracciare una linea retta che attraversa tre di quei punti. Al contrario, la configurazione della griglia sottostante non ha allineamenti: non è possibile tracciare una linea retta che attraversi tre punti.

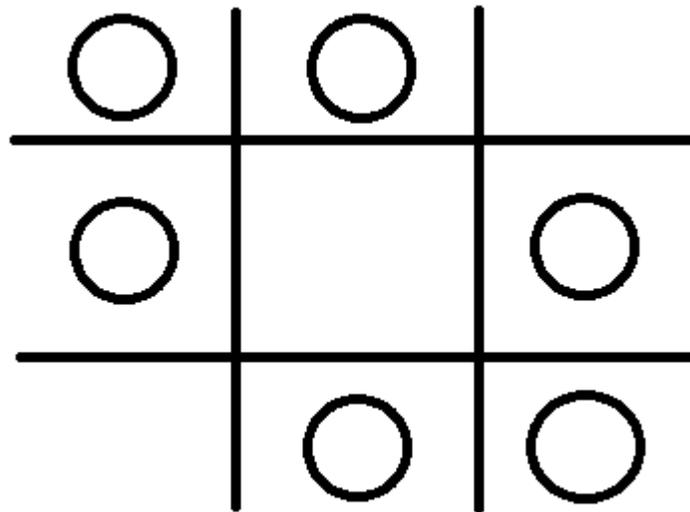


Fig. 2

Facendo delle prove è semplice notare che questo è il massimo numero di punti che consente una configurazione senza allineamenti. Disegnando sette punti nella griglia è quindi sempre possibile tracciare una linea retta che ne tocchi tre. Con uno sforzo maggiore, possiamo immaginare di estendere la griglia in un numero arbitrario di dimensioni spaziali. A questo punto possiamo fare la seguente domanda: considerando n come una dimensione qualsivoglia, quale sarà il numero massimo di punti che consentono una configurazione priva di allineamenti? Il teorema di Hales-Jewett ci dà delle indicazioni per rispondere a questa domanda. In particolare, una conseguenza del teorema è che quando il numero n di dimensioni è molto alto, il numero di punti della maggiore configurazione priva di allineamenti è solo una frazione minuscola del numero totale di posizioni sulla griglia.

Il Polymath Project ebbe una partenza lenta. Sette ore dopo che Gowers aveva aperto il suo blog di discussione matematica, ancora non c'era un solo commento. Poi un matematico della University of British Columbia, Jozsef Solymosi, postò un commento per suggerire una variante che era più semplice. Quindici minuti più tardi un professore di liceo dell'Arizona, Jason Dyer, postò un suo contributo. Poco

dopo il matematico dell'Ucla (University of California, Los Angeles) Terence Tao aggiunse un commento. Nei successivi trentasette giorni ventisette persone scrissero ottocento commenti matematici, per un totale di oltre 170 mila parole.

Mentre il progetto prendeva forma, emersero due filoni principali di ricerca. Il primo thread, condotto sul blog di Gowers, avrebbe continuato con l'obiettivo originale di trovare una dimostrazione combinatoria al teorema. Il secondo, sviluppato attraverso i commenti sul blog di Terence Tao, si concentrava su un aspetto differente del problema.

Dopo sette settimane, Gowers annunciò sul suo blog che il problema era stato "probabilmente risolto"²⁶, anche se i lavori proseguirono per entrambi i filoni di indagine fino al maggio 2009, circa tre mesi dopo l'annuncio.

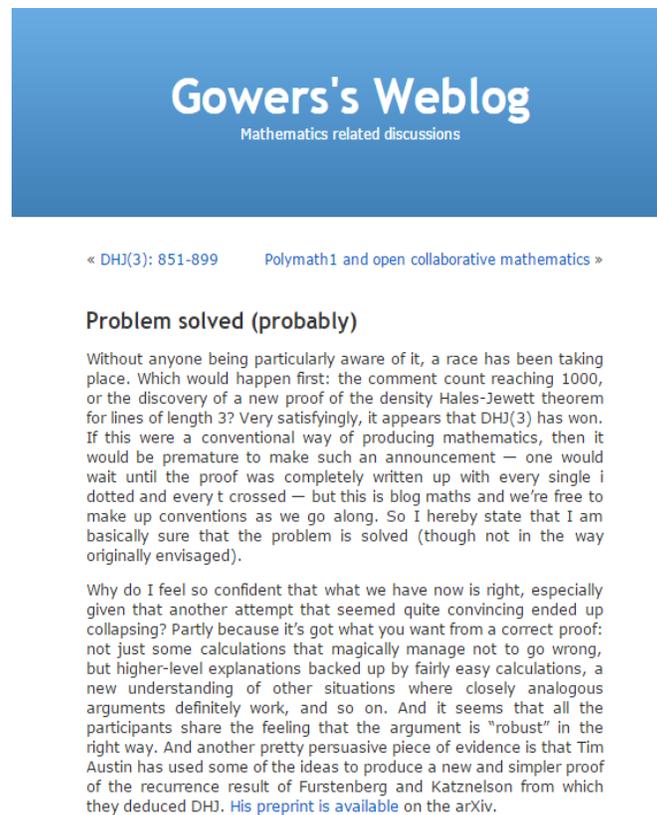


Fig. 3

²⁶ <https://gowers.wordpress.com/2009/03/10/problem-solved-probably/>

Il 9 gennaio 2010 Gowers, attraverso il suo blog, lancia la sua proposta per il quinto Polymath²⁷: risolvere il problema della discrepanza di Erdős, di cui verrà fornita una descrizione nel secondo capitolo. La proposta di Gowers, insieme ad altre tre ipotesi di ricerca, venne sottoposta al giudizio della comunità ed eletta come la più adatta da affrontare per mezzo della collaborazione online. Il lavoro sul Polymath5 continuerà fino al 2012, quando viene definitivamente abbandonato senza aver raggiunto nessun risultato. Nonostante l'apparente insuccesso, il progetto ha fatto emergere le idee e l'approccio metodologico che hanno consentito a Terence Tao di risolvere successivamente il problema. Pur essendo firmatario del paper finale, Tao ha ampiamente riconosciuto il contributo fornito dal Polymath al processo risolutivo da lui poi sviluppato.

The argument uses three ingredients. The first is a Fourier-analytic reduction, obtained as part of the Polymath5 project on this problem, which reduces the problem to the case when f is replaced by a (stochastic) completely multiplicative function g . The second is a logarithmically averaged version of the Elliott conjecture, established recently by the author, which effectively reduces to the case when g usually pretends to be a modulated Dirichlet character. The final ingredient is (an extension of) a further argument obtained by the Polymath5 project which shows unbounded discrepancy in this case.²⁸

Soltanto un altro progetto Polymath (Polymath8) è stato risolto e parecchi altri sono stati "abbandonati con discrezione", scrive Tim Gowers. Nel frattempo, altre proposte sono emerse e hanno dato vita a nuove collaborazioni. Il successo di Terence Tao ha spinto la comunità Polymath a lanciare una nuova sfida su un'altra congettura di Erdős: il sistema Erdős Rado Delta, detto anche il problema del girasole.

²⁷ <https://gowers.wordpress.com/2010/01/09/erds-discrepancy-problem-continued/>

²⁸ <https://arxiv.org/abs/1509.05363>

La storia del progetto Genoma

Per avere un'idea sul tipo e sulla quantità di dati con cui la genetica ha giornalmente a che fare, ripercorriamo velocemente le tappe del progetto Genoma Umano e della genetica contemporanea, fornendo anche qualche dato. A questa titanica impresa, infatti, hanno contribuito e continuano a contribuire scienziati con le formazioni più diverse: ricercatori provenienti dalla matematica e dalla fisica, dalla chimica, dalla medicina. Nell'aprile 1996 cento diversi laboratori situati negli Stati Uniti, in Europa, in Canada e in Giappone, avevano completato la prima sequenza di un genoma eucariotico, quello del lievito di birra *Saccharomyces Cerevisiae*, il cui DNA è formato da 12 milioni di basi. Nel dicembre 1998 il Sanger Centre, un centro di ricerca britannico sostenuto finanziariamente dal Wellcome Trust, un'istituzione filantropica che opera in campo medico, portò a termine, dopo quindici anni di sforzi, il lavoro di sequenziamento dei 97 milioni di basi del verme *C. Elegans*, rivelando che nel genoma di questo organismo erano presenti 19.099 geni, il 30% dei quali simili ai geni umani ma, nella maggior parte dei casi, mai osservati prima di allora.

Nel 1993 il progetto di sequenziamento del genoma portato avanti dai National Institutes of Health americani faceva affidamento su sei principali centri di sequenziamento che operavano negli Stati Uniti, ma all'inizio del 1998 questi centri sequenziavano al massimo 30 milioni di paia di basi l'anno, meno della metà di quelle necessarie a completare la sequenza del genoma umano entro la data inizialmente stabilita (2005).

Verso la fine della primavera del 1998, grazie all'introduzione di nuovi sequenziatori (PRISM ABI), il progetto pubblico americano di sequenziamento del genoma subì una forte accelerazione. Nel 1999 fu completato il sequenziamento del genoma della *Drosophila Melanogaster*, il moscerino della frutta su cui la genetica si

esercitava da lungo tempo, che possiede 120 milioni di basi. Un'impresa che dimostrò l'efficacia della nuova tecnologia.

Il 26 giugno 2000 il presidente degli Stati Uniti Bill Clinton organizzò una conferenza stampa per annunciare il completamento di un abbozzo del genoma umano, da cui risultava che quest'ultimo conteneva tra i 30 e i 40 mila geni e 3,1 miliardi di basi circa. Nell'aprile 2003, il progetto pubblico di studio americano del genoma annunciò il completamento dell'intera sequenza del genoma umano.²⁹

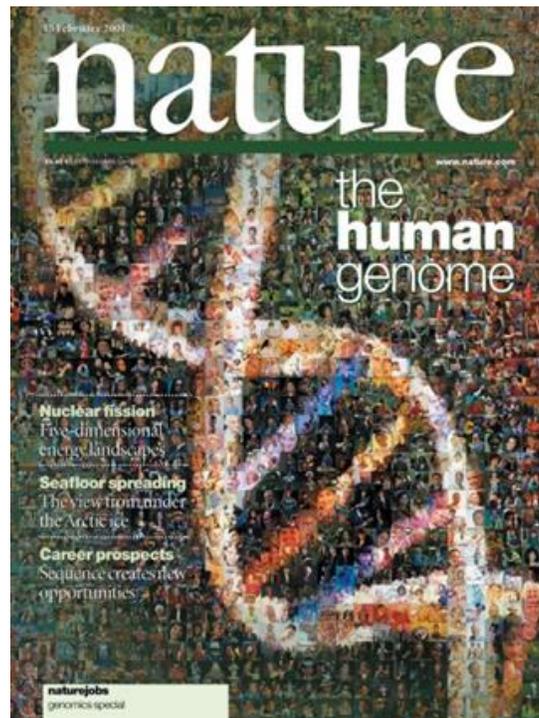


Fig. 4

Meno nota, ma forse persino più importante, è la HapMap³⁰ (abbreviazione di Haplotype Map, mappa degli aplotipi), che descrive come e dove gli esseri umani diversi differiscono nel loro codice genetico. Il progetto HapMap inizia nel 2002 e, studiando 270 individui di 4 razze diverse, mappa 1 milione di polimorfismi

²⁹ [http://www.treccani.it/enciclopedia/genoma-il-progetto-genoma-umano_\(Enciclopedia-della-Scienza-e-della-Tecnica\)/](http://www.treccani.it/enciclopedia/genoma-il-progetto-genoma-umano_(Enciclopedia-della-Scienza-e-della-Tecnica)/)

³⁰ http://www.ncbi.nlm.nih.gov/variation/news/NCBI_retiring_HapMap/

genetici. Nel 2010 HapMap ne aveva già recensiti oltre 3 milioni, in 11 popolazioni diverse a livello mondiale. Lo scopo di HapMap è quello di aiutare la comprensione delle basi genetiche di molte malattie comuni.

Un altro grande progetto di sequenziamento dei dati genetici è ENCODE³¹, consorzio di ricerca pubblico lanciato dall'Istituto di ricerca nazionale del genoma umano degli Stati Uniti (NHGRI) nel settembre 2003. L'obiettivo è di trovare tutti gli elementi funzionali presenti nel genoma umano, il quale rappresenta uno dei progetti più ambiziosi dopo il completamento del progetto Genoma Umano. Già nel 2012 ENCODE aveva acquisito una quantità di dati pari a 15 terabyte (1 terabyte equivale a 1024 gigabyte).

Grazie all'utilizzo di strumenti di sequenziamento sempre più potenti, che hanno moltiplicato le loro prestazioni di cinque volte tra il 2008 e il 2013, oggi l'ammontare dei dati a nostra disposizione riguardanti il DNA è dell'ordine dei petabyte, cioè di miliardi di byte. La grande quantità di dati proveniente dalla genetica viene caricata su service centralizzati come GenBank³², il deposito di informazioni genetiche gestito dal National Center for Biotechnology Information degli Stati Uniti. GenBank integra tutte le informazioni genetiche in un unico database online pubblicamente accessibile, un archivio del lavoro di migliaia di biologi. Secondo Nielsen:

Sono così tante informazioni che sarebbe pressoché impossibile analizzarle a mano. Per fortuna chiunque può scaricare liberamente la mappa genetica, e poi analizzarla tramite algoritmi, magari scoprendo verità insospettite sul genoma umano. [...] Questo è di fatto quel che rende possibili gli studi che associano certi geni a certe malattie. I ricercatori che svolgono quegli studi individuano prima di tutto un ampio gruppo di persone affette da una certa malattia e un gruppo di controllo sano, e poi usano la mappa genetica umana per scovare delle correlazioni tra l'incidenza della malattia e le differenze genetiche tra i due gruppi. Questo metodo viene usato in

³¹ <https://genome.ucsc.edu/ENCODE/>

³² <http://www.ncbi.nlm.nih.gov/genbank/>

tutta la scienza. [...] stiamo creando collettivamente una mappa del mondo intero, integrando il lavoro di centinaia di migliaia di scienziati. (Nielsen, 2012)

Eseguendo una ricerca sul database di letteratura biomedica PubMed³³, uno dei primi esempi di successo di piattaforma open access che consente un'ampia circolazione della letteratura scientifica di diverse discipline, è possibile mostrare quanti paper, a partire dal 1982, anno di creazione di GenBank, abbiano fatto ricorso alle sequenze genetiche presenti in questa banca digitale. Il numero complessivo delle ricerche in ambito genetico e biomedico che hanno condotto a nuove scoperte grazie all'utilizzo dei dati contenuti in GenBank ammonta a 227.740³⁴. Un'ulteriore analisi sulla quantità degli studi pubblicati annualmente dimostra poi la sensibile progressione con cui il numero è cresciuto negli anni, in corrispondenza della crescente mole di sequenze genetiche, calcolata nel grafico sottostante in termini di bytes caricati, presente su GenBank.

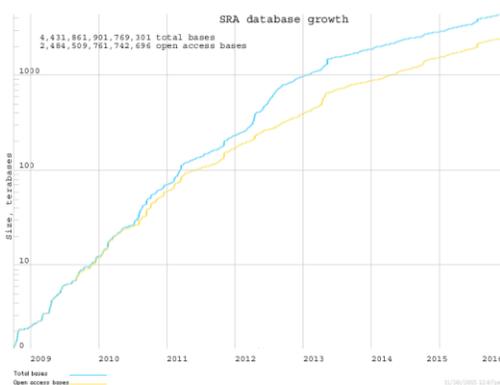


Fig. 5

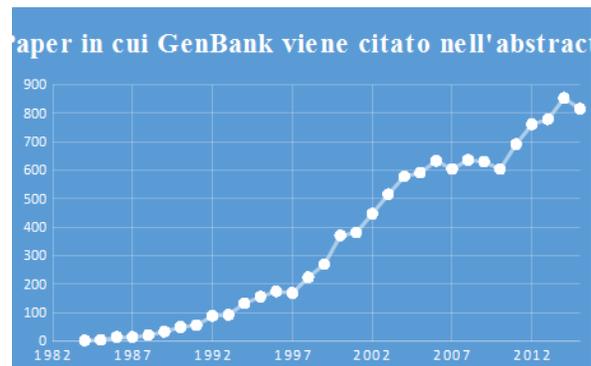


Fig. 6

Se il Polymath Project rappresenta un cambiamento nel modo in cui gli scienziati collaborano per creare conoscenza, GenBank e gli studi genetici illustrano un

³³ <http://www.ncbi.nlm.nih.gov/pubmed>

³⁴ Numero complessivo di paper presenti su PubMed che a giugno 2015 esibivano la parola GenBank nell'abstract. La ricerca è stata condotta sottoponendo una query e gli opportuni filtri al motore di ricerca del portale.

cambiamento nel modo in cui gli scienziati trovano un significato nei dati scaturiti dall'osservazione sperimentale. Entrambe le tipologie di risorse online sono quindi state fatte proprie dagli scienziati e accolte come strumenti di lavoro capaci di aiutare la collaborazione e la partecipazione a progetti di ricerca. Tuttavia, l'esperienza di GenBank e delle piattaforme dedicate alla gestione delle referenze e della ricerca in ambito biologico sono esemplificative della prassi di lavoro della genetica. Questa disciplina, sin dai suoi primi passi, è stata infatti fortemente votata alla collettivizzazione del lavoro e a una necessaria transdisciplinarietà dettata dalla varietà e della quantità dei dati sperimentali di riferimento. Tali esigenze rendono conto della forte componente digitale che caratterizza la genetica contemporanea e fanno di questa disciplina l'esempio perfetto dell'integrazione delle logiche del Web 2.0 nella pratica scientifica contemporanea.

Intervista a Edoardo Boncinelli

Edoardo Boncinelli è uno dei più noti biologi italiani. Nonostante l'ambito di studi a cui ha dedicato la sua intera carriera da ricercatore, la genetica, Edoardo Boncinelli è un fisico di formazione, a dimostrazione dell'interdisciplinarietà che contraddistingue questa disciplina.

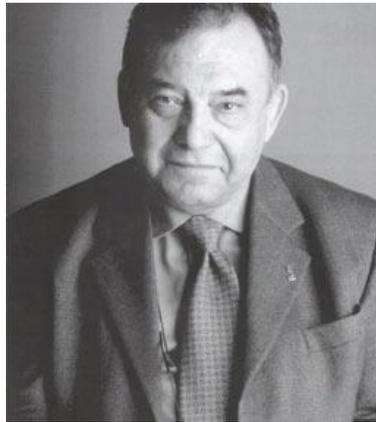


Fig. 7

Laureato in Fisica a Firenze con una tesi sperimentale di elettronica quantistica, dal 1968 al 1992 ha ininterrottamente svolto ricerca nel campo della genetica presso l'Istituto internazionale di genetica e biofisica del CNR (Consiglio Nazionale delle Ricerche). È stato direttore della Scuola Internazionale Superiore di Studi Avanzati di Trieste (SISSA), del laboratorio di biologia molecolare dello sviluppo presso l'Istituto scientifico universitario San Raffaele e direttore di ricerca presso il Centro per lo studio della farmacologia cellulare e molecolare del CNR di Milano. È attualmente docente presso la Facoltà di Filosofia dell'Università Vita-Salute San Raffaele di Milano. Nel 1985, al CNR di Napoli, ha scoperto, insieme al suo team, i geni omeotici, responsabili delle prime fasi di sviluppo dell'organismo umano. Nel 1996 scopre l'alterazione genetica responsabile della schizoencefalia e nel 2000, insieme ad Antonello Mallamaci, del Dipartimento biotecnologie del San Raffaele di Milano e con il contributo di Telethon, il gene della corteccia cerebrale che regola il pensiero astratto.

Considerata la sua lunga carriera nell'ambito della biologia molecolare e i successi da lei ottenuti, cosa pensa del rapporto tra strumenti informatici digitali e genetica.

"La storia della genetica contemporanea è legata a doppio filo con quella dell'informatica, nascono infatti entrambe nel secondo dopoguerra e procedono parallele per molto tempo.

Non è difficile comprendere le motivazioni di questa commistione: la complessità degli oggetti di studio della genetica necessita di una tecnologia in grado di implementare algoritmi per descrivere modelli affidabili sul comportamento dei geni. Da ciò deriva anche la condivisione di termini tra le due discipline. I primi grandi genetisti hanno infatti trovato appropriato mutuare parole quali modifica e trascrizione proprio dal vicino mondo informatico.”

Come pensa sia cambiato nel tempo questo rapporto?

“L’evoluzione degli strumenti sperimentali, diventati sempre più potenti, ci mette oggi di fronte a una impressionante quantità di dati. Abbiamo infatti a disposizione biblioteche digitali complete, o quasi, di sequenze di genomi, trascrittomi - i vari RNA di trascrizione del codice genetico -, proteomi - le proteine che sostengono la vita biologica - e persino di connettomi - il vasto insieme di connessioni che si dipanano all’interno del nostro cervello - . Una tale valanga di dati impone ai ricercatori di spostare l’attenzione dal laboratorio all’interpretazione di quegli stessi dati. In maniera del tutto ipotetica, potremmo affermare che al giorno d’oggi i genetisti potrebbero tranquillamente appendere i camici bianchi e svolgere solo attività di data mining davanti ai loro computer. Quello che però manca è la capacità di trovare un significato o correlazioni rilevanti all’interno di questa moltitudine di dati.”

In che modo, secondo lei, si può sopperire a questa mancanza?

“Ciò che si sta cercando di fare è arrivare a uno sviluppo di una bioinformatica matura, in cui le risorse computazionali e algoritmiche siano in grado di darci risposte chiare sulla natura dei dati sperimentali. Tuttavia, questa prospettiva potrebbe anche rivelarsi del tutto infondata e inapplicabile. Quello che abbiamo ottenuto finora è comunque molto e i servizi di biblioteca 2.0 hanno sostenuto i successi ottenuti, diventando una parte necessaria ed essenziale delle infrastrutture di ricerca nel campo della biologia molecolare. Una seconda soluzione è quindi quella di continuare su questa strada cercando di allargare la collaborazioni tra ricercatori e comunità scientifiche differenti ponendoli di fronte a progetti e sfide che richiedano la partecipazione di tutti.”

I servizi di open access dell'Agencia Spaziale Europea e il caso Survey Lab

Fondata nel 1975, L'Agencia Spaziale Europea (ESA, European Space Agency) è stata creata in ambito europeo con l'obiettivo di aprire nuove strade all'esplorazione spaziale, sviluppare tecnologie avanzate ed edificare un'industria competitiva in grado di emergere sul piano mondiale. Attualmente, sono venti i paesi europei che fanno parte di questo ente. La strategia spaziale dell'ESA si articola attorno a quattro principali assi: lo sviluppo delle conoscenze scientifiche, il miglioramento della qualità della vita sulla Terra, il successo della cooperazione europea, la promozione dell'industria europea.

Le attività principali dell'ESA riguardano programmi scientifici e di esplorazione spaziale, osservazione della Terra, telecomunicazioni, navigazione satellitare, trasporto spaziale e lanciatori, Stazione Spaziale Internazionale³⁵. Nell'ambito dello studio e del controllo ambientale, negli ultimi trent'anni l'ESA ha ideato e sviluppato una vasta quantità di missioni di telerilevamento tutt'ora operanti, da ultimo - ma non per importanza - il progetto Copernicus³⁶, che consta di una flotta di 30 satelliti equipaggiati con differenti strumenti di osservazione della Terra chiamati Sentinel. Tre i satelliti ad oggi in orbita: Sentinel-1A, Sentinel-1B e Sentinel-2. Oltre a fornire informazioni per il controllo e la gestione dei disastri ambientali, uno degli obiettivi primari del programma Copernicus e dell'Agencia Spaziale Europea è quello di sviluppare sistemi integrati di acquisizione, interpretazione e condivisione della grande quantità di dati forniti dal remote

³⁵

http://www.esteri.it/mae/it/politica_estera/economia/cooperaz_econom/agenziaspazialeeuropea.html

³⁶ <http://copernicus.eu/>

sensing (Ground Segment), allo scopo di favorire e supportare il progresso scientifico e tecnologico in campo ambientale e non solo. In questo specifico settore, l'ESA offre una serie di servizi online, quali banche dati, risorse di processamento e portali di community, che mirano alla creazione di reti di collaborazione tra scienziati e all'agevolazione del lavoro di ricerca. Tali caratteristiche rendono l'operato dell'agenzia estremamente interessante ai fini del seguente elaborato.

Come per quanto già messo in luce nel paragrafo relativo alla classificazione delle piattaforme sociali del Web 2.0, anche le risorse online messe a disposizione dall'ESA hanno subito nel tempo un'evoluzione in linea con la tendenza al mash up dei servizi. La nuova fase di sviluppo del Ground Segment dell'Agenzia prevede infatti l'implementazione di siti con funzionalità sociali sempre più variegata e accessibili. Uno degli esempi del nuovo corso dello sviluppo dei servizi collaborativi online di ESA è rappresentato dalla Thematic Exploitation Platform (TEP), sito rivolto allo sfruttamento dei dati satellitari di osservazione della Terra. Il progetto TEP raccoglie cinque differenti community impegnate su tematiche differenti: Coastal, Forestry, Hydrology, Geohazards, Polar and Urban. Scopo principale del sito è quello di avvicinare gli utenti e le comunità interessate ai dati e agli strumenti per la loro interpretazione offrendo un ambiente digitale che consenta le seguenti attività:

- accesso a dati rilevanti di osservazione della Terra e non solo;
- accesso a un network modulare, a risorse computazionali e a ambienti virtuali in grado di ospitare gli algoritmi di processamento degli utenti;
- accesso a una piattaforma che metta gli utenti nella condizione di integrare, testare, eseguire e gestire applicazioni senza la complessità di costruire e mantenere una propria infrastruttura, e di fornire l'accesso ai servizi della piattaforma standard e alle funzioni più rilevanti: strumenti di

collaborazione, data mining e applicazioni di visualizzazione, strumenti di sviluppo, strumenti di comunicazione (Social Network) e documentazione, di contabilità e strumenti di reporting per gestire l'utilizzo delle risorse;

- accesso a repository di applicazioni o store, che forniscono applicazioni da utilizzare per elaborazioni avanzate e rilevanti³⁷.



Fig. 8

La varietà degli strumenti di collaborazione online appena elencata ha sicuramente consentito all'Agenzia Spaziale Europea di allargare la rete di ricercatori, di istituzioni scientifiche e di aziende private in grado di accedere ai dati provenienti dall'osservazione satellitare della Terra, promuovendo lo sfruttamento dell'informazione prodotta dal remote sensing per la realizzazione di soluzioni o di applicazioni tecnologiche innovative, anche in contesti apparentemente lontani dal settore spazio, come la sicurezza alimentare e quella delle infrastrutture. A tal proposito, risulta esemplificativo il caso di Survey Lab, spin off dell'Università "La Sapienza" di Roma. Grazie allo sfruttamento dei dati acquisiti da una particolare categoria di satelliti ed erogati dall'Agenzia Spaziale Europea attraverso le piattaforme descritte, l'azienda è stata in grado di realizzare EDI.MODI - anagramma di EDifice MOonitoring DIspacement -, algoritmo di monitoraggio della stabilità degli edifici e delle infrastrutture civili. EDI.MODI utilizza dati di

³⁷ <https://tep.eo.esa.int/>

osservazione della terra ottenuti da satelliti equipaggiati con interferometri SAR (Synthetic Aperture Radar), integrati con dati acquisiti da sensori posizionati sul territorio e alimentati all'interno di modelli numerici, per fornire analisi avanzate sulle deformazioni e i fenomeni evolutivi di danneggiamento.

Intervista a Pier Bargellini (progetto Sentinelles dell'ESA)

Pier Bargellini è un ingegnere aerospaziale italiano. Dopo la laurea presso l'Università di Pisa, entra nel 1997 nell'Agenzia Spaziale Europea. Per 15 anni Bargellini ha lavorato nell'European Space Operations Center (ESOC) di Darmstadt, in Germania. Durante questo periodo è stato coinvolto nello sviluppo e nella gestione di una serie di missioni di osservazione della Terra. Dal 1997 al 2004 è stato responsabile per le operazioni della piattaforma Envisat. È stato poi nominato Spacecraft Operations Manager, ruolo che ha ricoperto per diverse missioni relative all'osservazione della terra, tra cui il lancio e la prima fase delle Sentinelles. Nel 2012 si è trasferito al centro dell'ESA per l'osservazione della Terra (ESRIN) in Italia, dove è stato nominato capo della Copernicus Space Component (CSC) Mission Management and Ground Segment Development Division.

Dottor Bargellini, quali sono i servizi che il ground segment dell'ESA mette a disposizione della comunità scientifica?

“Distinguerei la ground segment tradizionale dall'evoluzione che stiamo mettendo in piedi. Il ground segment tradizionale, il modello utilizzato fino a oggi e che riguarda tutte le agenzie spaziali, offre un servizio principale che è quello della disseminazione dei contenuti. Quindi si acquisiscono dati dai satelliti, si processano, si archiviano, si procede a un controllo della qualità e poi vengono messi a disposizione degli utenti. Nel passato, anche se

la fornitura dei dati era comunque gratuita, i ricercatori e le aziende interessate dovevano compilare una domanda giustificando le motivazioni della richiesta dei dati. Tutto questo processo veniva inoltre tracciato in modo da rendere conto delle spese necessarie alla disseminazione di questi prodotti. Negli ultimi cinque anni questo meccanismo in ESA è stato dismesso. Oggi gli utenti possono registrarsi autonomamente con un loro account e, senza compilare nessun form, hanno subito a disposizione i dati. Questo è il principale servizio che mettiamo a disposizione, ma che rimane ancora una funzione di stampo tradizionale. Nel recente passato abbiamo però iniziato a sviluppare e sperimentare servizi che rientrano nella categoria che chiamiamo exploitation platform. Questo genere di infrastrutture consistono in un ambiente di lavoro digitale che comprende i dati, la macchine virtuali con cui processarli e una serie di tool, che possono essere di visualizzazione, analytics, data ming. Ciò consente all'utente di lavorare interamente all'interno dell'environment, di condividere eventualmente i risultati grazie al servizio di community e anche di riprodurre i risultati già ottenuti senza doverli scaricare. È facile capire che questa serie di servizi consente di far risparmiare ai ricercatori tempo e denaro."

Qual è la natura e la quantità di dati che mettete a disposizione?

"Con il progetto Copernicus e l'inizio dell'attività delle prime Sentinel, la quantità dei dati che riceviamo giornalmente è nell'ordine dei terabyte. Inoltre, a seconda del range di missioni, abbiamo immagini di tipo ottico di varie risoluzioni: low resolution, che corrispondono a una definizione pari a un'acquisizione ottenuta da 1 chilometro di distanza, high resolution, da circa 10 metri, e very high resolution, sotto i 2 metri. Queste tipologie si accompagnano poi a varie bande spettrali, come l'infrarosso. Abbiamo strumenti e di conseguenza immagini radar, anche in questo caso con due differenti tipi di risoluzione. Infine dati provenienti da missioni dedicate al rilevamento della composizione atmosferica."

Che genere di categorie professionali utilizzano i vostri dati?

"Negli ultimi anni abbiamo osservato che, oltre ai ricercatori ovviamente più vicini a ESA, ci sono molti scienziati, anche nel settore commerciale e molto lontani dall'ambito spaziale, che utilizzano i nostri dati per lo sviluppo di applicazioni. Esempi molto recenti riguardano

il mondo delle compagnie assicurative e quelle dell'oil and gas, fino ad arrivare ad aziende come Google e Amazon, che mettono a disposizione i nostri dati sui loro portali sia a uso commerciale sia per finalità di ricerca. Abbiamo quindi applicazioni in ogni genere di dominio, come il trasporto marittimo, la security in ambito civile e il settore energetico."

Intervista a Silvia Scifoni (progetto Edi.Modi)

Silvia Scifoni è assegnista di ricerca presso il Dipartimento di Ingegneria Civile, Edile e Ambientale dell'Università "La Sapienza" di Roma. Laureata in Ingegneria Ambientale, ha conseguito il dottorato di ricerca in geofisica. È socio di Survey Lab e responsabile del settore ricerca e sviluppo dell'azienda.

Dottoressa Scifoni, cos'è esattamente Edi.Modi e che tipo di tecnologia sfrutta?

Edi.Modi si occupa del monitoraggio degli spostamenti degli edifici in area urbana e di grandi infrastrutture strategiche, come ad esempio autostrade e dighe. Utilizza la tecnica DinSAR, cioè l'interferometria radar satellitare, per estrarre informazione dai dati SAR di osservazione della Terra. Questa metodologia di intervento non prevede quindi l'installazione di target sugli edifici, permette di osservare spostamenti delle strutture nel lungo termine e di valutare il loro comportamento andando ad analizzare e comparando i dati di archivio. Questa tecnologia non è stata inventata dalla Survey Lab, ma la nostra società ha l'obiettivo, attraverso il progetto Edi.Modi, di creare dei servizi a valore aggiunto che possano essere utilizzabili e comprensibili anche ai non esperti del settore. Infatti la tecnologia DinSAR viene attualmente utilizzata soprattutto in ambito universitario e di ricerca. La Survey Lab, attraverso il progetto Edi.Modi, cerca di realizzare servizi a livelli progressivamente più complessi adatti a diversi tipi di utenti: dal singolo cittadino fino ai tecnici e professionisti che operano sul territorio."

In che modo siete arrivati all'idea di utilizzare i dati satellitari per il controllo della stabilità delle infrastrutture?

Il personale della nostra azienda proviene nella sua quasi interezza dall'ambito universitario. Pur avendo competenze e percorsi multidisciplinari, abbiamo tutti esperienza con questa tipologia di dati, che vengono spesso utilizzati per scopi di ricerca accademica. Dal 2008, dopo la fondazione della Survey Lab, abbiamo iniziato a pensare di rendere accessibili le possibilità che questa tecnologia offre anche ai non esperti del settore. Questo soprattutto perché, stando a stretto contatto con la realtà dell'ingegneria civile, ci siamo resi conto che la soluzione di alcune delle problematiche che un tecnico che opera sul territorio si trova ad affrontare potrebbero essere supportate dall'utilizzo di dati di osservazione della Terra.

Che ruolo hanno avuto e hanno le piattaforme open access e i servizi online messi a disposizione nello sviluppo e realizzazione del vostro strumento?

Oltre ai dati satellitari SAR, che sono attualmente open access su portali come ERS, ENVISAT e SENTINEL, verranno utilizzate ulteriori piattaforme open access per la base cartografica, tipo Open Street Map e altre, per acquisire dati informativi sul territorio analizzato. Infatti, oltre ai dati DinSAR, il servizio si avvale di diverse informazioni accessorie direttamente disponibili sul territorio: alcune servono per creare i livelli di servizio, mentre altre servono per andare ad analizzare in dettaglio le aree critiche individuate dai dati DinSAR. Inoltre si prevede l'utilizzo di tool open source per lo sviluppo del Webgis, sistemi informativi geografici, e di applicativi utili al servizio.

Il caso di Ebola Research Database

Nel corso del 2014 un'epidemia di Ebola colpì l'Africa occidentale. Iniziata in Guinea nel febbraio dello stesso anno, la malattia si diffuse successivamente in Liberia, Sierra Leone e Nigeria. Ad oggi, con 28.646 casi segnalati e 11.323 morti,

questa epidemia di Ebola è considerata da virologi ed epidemiologi come la più grave mai avvenuta. A causa della veloce diffusione e dell'estrema aggressività della malattia, molte organizzazioni e istituzioni scientifiche, tra cui la Comunità economica degli Stati dell'Africa occidentale, i centri statunitensi per il controllo delle malattie, lo United Kingdom Collaborative on Development Science, la Commissione Europea e ovviamente la World Health Organization, impegnarono risorse e attivarono progetti - alcuni dei quali ancora in corso - al fine di contrastare l'epidemia e la sua futura ricomparsa.

I primi segnali della presenza della malattia furono rilevati nel dicembre del 2013 nella Regione boschiva della Guinea sud-orientale, vicino al confine con la Liberia e la Sierra Leone. I casi sospetti furono registrati nelle prefetture di Conakry, Guéckédougou, Macenta e Dabola. A seguito del rapporto sulla situazione in Guinea da parte dell'Istituto Pasteur di Lione, il 1 aprile l'Organizzazione Mondiale della Sanità lanciò l'allerta per l'epidemia, confermando che le morti per febbre emorragica avvenute nel paese erano da attribuirsi a Ebola, e più precisamente al ceppo noto con il nome di *Zaire Ebolavirus*. Un primo rapporto suggeriva che si trattasse di un nuovo ceppo di virus Ebola, ma quest'ipotesi fu smentita da studi successivi che confermarono il rapporto dell'istituto virologico francese. Un caso sospetto fu ricoverato in ospedale il 28 marzo 2014. Il 31 marzo, i Centers for Disease Control degli Stati Uniti inviarono un team di cinque persone per supportare le autorità delle nazioni colpite dal virus e l'Organizzazione Mondiale della Sanità nel lavoro di controllo del contagio e gestione dell'epidemia. A questo stesso scopo, a settembre dello stesso anno, gli Stati Uniti avrebbero inviato un contingente di 3 mila soldati.

Intorno al 23 maggio, l'epidemia si propagò fino alla capitale della Guinea, Conakry, una città di circa due milioni di abitanti. In Liberia, la malattia si ripresentò verso la fine di marzo, e già entro la metà di aprile il Ministero della Salute e della previdenza sociale della Liberia (MoHSW) registrò possibili casi nelle

contee di Margibi e Montserrado. In Sierra Leone, Mali e Ghana furono individuati alcuni casi sospetti entro la metà di aprile. Gueckedou era l'unico punto caldo dove furono ancora segnalati casi di trasmissione e decessi. Dal 23 al 27 maggio 2014 furono riportati diversi nuovi casi clinici, interessanti tre distretti già precedentemente colpiti (Guéckédou, Macenta e Conakry), quattro nuovi distretti (Boffa, Téliimélé, Boke e Dubreka) e un nuovo paese (la Sierra Leone). A metà giugno furono segnalati i primi casi a Monrovia, in Liberia. Il direttore delle operazioni di Medici Senza Frontiere, Bart Janssens, a fine giugno descrisse la situazione come "senza precedenti, totalmente fuori controllo e la situazione può solo peggiorare".

Ai primi di luglio, in Liberia furono segnalate 107 infezioni (52 vennero confermate in laboratorio) e almeno 65 morti causati da Ebola. Dalla metà di aprile erano stati segnalati solo quattro morti al di fuori della contea di Lofa, mentre a metà giugno il Ministero della Salute e del Welfare liberiano annunciò ulteriori sette morti solo nella contea di Montserrado. L'epidemia progredì rapidamente nella Sierra Leone. I primi casi furono riscontrati il 25 maggio nel distretto di Kailahun, vicino al confine con Guéckédou, in Guinea³⁸.

Il primo caso a Freetown, la capitale della Sierra Leone, fu riportato verso la fine di luglio. Il primo caso in Nigeria fu invece segnalato dall'OMS il 25 luglio 2014 e riguardava un cittadino liberiano, un quarantenne naturalizzato americano di nome Patrick Sawyer, che era sbarcato in aereo nella città di Lagos e che morì successivamente a causa della malattia. L'ospedale dove l'uomo era stato in cura fu chiuso e messo in quarantena; gli operatori sanitari che lo avevano in cura furono a loro volta isolati nel tentativo di fermare la diffusione del virus³⁹.

³⁸

https://it.wikipedia.org/wiki/Epidemia_di febbre_emorragica_di_Ebola_in_Africa_Occidentale_del_2014

³⁹ <http://www.who.int/csr/disease/ebola/en/>

Il 22 ottobre 2014 l'Organizzazione Mondiale della Sanità ha dichiarato concluse le epidemie di virus Ebola in Nigeria e in Senegal e il 9 maggio 2015 in Africa Occidentale con un comunicato⁴⁰.

L'epidemia di Ebola fece la sua comparsa anche negli Stati Uniti e in Europa, per lo più paesi di provenienza dei molti operatori sanitari rimpatriati che avevano prestato servizio nelle zone del massimo contagio. A fine 2014 era infatti giunto a 22 il conteggio dei casi di Ebola riscontrati fuori dall'Africa Occidentale. Negli Usa furono 10 i pazienti trattati per il virus, il primo dei quali, Kent Brantly, era un medico missionario rimpatriato dalla Liberia e guarito anche grazie al siero sperimentale Zmapp. Come lui, altri due medici e un'infermiera, che avevano contratto il virus in Africa, e un cameraman della NBC guarirono grazie all'utilizzo del trattamento sperimentale, lo stesso che aveva permesso di curare le due infermiere di Dallas venute in contatto con il "paziente zero" statunitense, il cittadino liberiano Thomas Eric Duncan, invece morto. Stessa sorte per Martin Salia, un medico statunitense contagiato in Sierra Leone e rimpatriato il 15 novembre 2014. In Spagna morirono entrambi i missionari rimpatriati, uno dei quali fu anche il primo caso di Ebola fuori dai confini africani, mentre Teresa Romero, l'infermiera infettatasi dopo aver curato uno dei due, guarì. Gli altri casi di operatori sanitari rimpatriati e curati con successo dopo aver contratto la malattia in Africa riguardarono Francia, ancora la Spagna e la Norvegia. Sempre in Francia un operatore dell'Unicef ricoverato a Parigi ai primi di novembre dopo il contagio in Sierra Leone fu dimesso dopo quasi un mese. Furono invece tre i pazienti curati in Germania: due medici, entrambi guariti, e un operatore dell'Onu invece morto. In Svizzera il 21 novembre fu ricoverato un medico cubano che riuscì a guarire. Fabrizio Pulvirenti, il medico italiano di Emergency che contrasse la

⁴⁰ <http://www.who.int/mediacentre/news/statements/2015/liberia-ends-ebola/en/>

malattia in Sierra Leone, fu invece trasferito a novembre 2014 all'Istituto Spallanzani di Roma e dimesso a inizio gennaio 2015.⁴¹

La cronistoria della diffusione di Ebola proposta e i dati dell'OMS aggiornati al 22 novembre 2015 dimostrano la velocità di diffusione del virus e il suo alto tasso di mortalità.

Table 1: Confirmed, probable, and suspected EVD cases in Guinea, Liberia, and Sierra Leone

Country	Case definition	Cumulative cases	Cases in past 21 days	Cumulative deaths
Guinea	Confirmed	3351	0	2083
	Probable	453	*	453
	Suspected	0	*	‡
	Total	3804	0	2536
Liberia [§]	Confirmed	3151	-	‡
	Probable	1879	-	‡
	Suspected	5636	-	‡
	Total	10 666	-	4806
Liberia**	Confirmed	9	3	2
	Probable	*	*	‡
	Suspected	‡	*	‡
	Total	9	3	2
Sierra Leone [†]	Confirmed	8704	0	3589
	Probable	287	*	208
	Suspected	5131	*	158
	Total	14 122	0	3955
Total	Confirmed	15 215	3	‡
	Probable	2619	*	‡
	Suspected	10 767	*	‡
	Total	28 601	3	11 299

Fig. 9

Dopo l'aprile del 2014, nel tentativo di opporsi all'escalation del contagio del virus, gli scienziati, hanno iniziato a unire gli sforzi scegliendo la strada della collaborazione. Una delle iniziative realizzate per cercare di velocizzare la ricerca

⁴¹ http://www.adnkronos.com/salute/sanita/2015/01/02/dall-africa-agli-usa-storia-dei-contagi-del-virus-maledetto_CSRkgQcq27tQJXRSScpeoM.html

di soluzioni in grado di contrastare Ebola è stata l'attivazione dell'Ebola Research Database⁴² (ERD). Sviluppata dalla United Kingdom Collaborative on Development Science, un'organizzazione che si pone come obiettivo quello di condividere conoscenze al fine di promuovere lo sviluppo internazionale, questa risorsa ha tutt'ora lo scopo di riunire le informazioni in merito alle ricerche che riguardano il virus Ebola. Nello specifico, l'Ebola Research Database contiene notizie su tutte le attuali ricerche accademiche rilevanti per contrastare Ebola e un'ampia banca dati di sequenze genetiche. Lo scambio di contenuti che il database consente è pensato per supportare l'allineamento tra i finanziatori della ricerca, al fine di ridurre un'indebita duplicazione e contribuire allo sviluppo di attività complementari in cui è richiesto un maggiore sforzo.

Invece di proseguire indipendentemente nel proprio lavoro, gli enti presenti su ERD, creando una vasta rete di connessioni di dati, hanno perciò accettato di condividere volontariamente informazioni sensibili, con la consapevolezza che una simile gestione della conoscenza prodotta avrebbe consentito di fronteggiare efficacemente e velocemente l'emergenza pandemica rappresentata da Ebola.

Ma qual è il modo migliore per fronteggiare la diffusione di un virus aggressivo come Ebola? Apparentemente, la migliore strategia è arrivare alla produzione di un vaccino nel minor tempo possibile. Nel periodo di maggiore crisi dell'epidemia, ricercatori di tutto il mondo incrementarono gli sforzi per cercare di raggiungere questo obiettivo. A questo proposito, il 3 agosto 2015 la rivista Lancet annunciò che la terza fase di un trial clinico su un vaccino per Ebola condotto in Guinea e finanziato dall'Organizzazione Mondiale della Sanità aveva ottenuto risultati convincenti⁴³. Nonostante ciò, prima che si arrivasse all'individuazione di una

⁴² <http://www.ukcds.org.uk/resources/ebola-research-database>

⁴³

http://www.repubblica.it/salute/ricerca/2015/07/31/news/ebola_oms_entusiasmanti_i_primi_risultati_del_vaccino_-120185068/

cura, la comunità scientifica, considerato il complicato quadro igienico sanitario che caratterizzava e continua a caratterizzare la zona dell'ultima epidemia di Ebola, aveva vagliato strade alternative con la finalità di comprendere l'efficacia di strategie differenti per contrastare la diffusione. Quindi, in mancanza di un vaccino, i biologi avevano bisogno di uno strumento grazie al quale riorganizzare l'attenzione nei confronti di quegli aspetti secondari, ma comunque implicati nel fenomeno del rapido contagio di Ebola in Africa, su cui fosse stato possibile intervenire. Sono state proprio queste le ragioni che hanno portato all'attivazione di un supporto di aggregazione delle informazioni prodotte dalla ricerca.

La creazione dell'Ebola Research Database ha perciò consentito di coordinare il lavoro della ricerca, esplorando contromisure alternative alla diffusione del virus. Visitando il sito di UK CDS e accedendo alla pagina riservata all'ERD è possibile consultare la vasta gamma di possibilità esplorate dagli studi condotti su Ebola. Analizzeremo qui solo le più significative, riportando i relativi abstract dei paper pubblicati.

- L'epidemia di Ebola in Africa ha fatto emergere i rischi derivanti dalla facilità con cui, in un mondo sempre più globalizzato e fisicamente connesso, le persone si spostano e entrano in contatto con comunità differenti. Le abitudini che contraddistinguono gli spostamenti odierni degli individui aumentano quindi esponenzialmente le possibilità di contagio - Ente: Wellcome Trust/ DFID - "Titolo della ricerca: Predicting the geographic spread of Ebola virus disease in West Africa" - Link: <http://www.elrha.org/map-location/oxford-uni-predicting-geographic-spread-ebola/>
- Una delle ipotesi più fondate sull'inizio della diffusione di Ebola in Guinea afferma che il caso zero sia stato il risultato della trasmissione del virus da un animale infetto all'uomo. Quindi gli animali hanno giocato un ruolo

importante sulle dimensioni del contagio in quanto vettori della malattia. - Ente: Federal Ministry of Education and Research (Germania) - Titolo della ricerca: "Project 9a: Investigating a possible secondary reservoir in animals in West Africa" - Link:

http://www.rki.de/DE/Content/InfAZ/E/Ebola/Ebola_node.html

- Vista l'aggressività di Ebola e la velocità del decorso dei sintomi a esso associati, fare affidamento su sistemi diagnostici rapidi ed economici potrebbe tradursi nella possibilità di implementare in maniera tempestiva le cure adeguate e in maggiori probabilità di sopravvivenza per i pazienti. - Ente: Commissione Europea - Titolo della ricerca: "FILODIAG" - Link: http://ec.europa.eu/research/press/2015/pdf/imi_ebola_project_overview_january_2015.pdf
- Individuare velocemente nuovi casi di contagio può risultare di fondamentale importanza per comprendere le dimensioni della diffusione della malattia e di conseguenza per adottare le giuste contromisure. A questo scopo, l'analisi dei dati generati dalla telefonia e dai mobile device possono fornire una valida risorsa. - Ente: Federal Ministry of Education and Research (Germania) - Titolo della ricerca: "Ebola surveillance with mobile real-time data transmission in Nigeria" - Link: http://www.dzif.de/en/news_press/news_press_releases/view/detail/artikel/ebokon_strengthening_ebola_research/

Queste sono solo alcune delle quasi 60 tra ricerche e iniziative dall'Ebola Research Database. La conoscenza prodotta grazie a questi studi ha quindi consentito e consentirà di implementare attività tramite cui controllare la diffusione di Ebola e di nuove epidemie che potrebbero presentarsi in futuro.

Il caso appena analizzato mette in evidenza un altro aspetto a supporto della nostra proposta di una terza accezione al termine Scienza 2.0 che sostiene che

l'utilizzo dei servizi per la collaborazione del Web 2.0 da parte dei ricercatori costituisca già una parte importante della metodologica scientifica odierna. La componente in questione è quella temporale. Infatti, sfruttando l'intelligenza collettiva emergente dalla collaborazione digitale tra scienziati, la Scienza 2.0 è in grado di rendere conto a problematiche complesse in tempi contenuti, senza rinunciare all'indispensabile fase di verifica delle teorie.

Intervista a Ilaria Capua

Al fine di supportare l'ipotesi secondo la quale i database digitali siano il risultato di un compromesso sociale e che svolgono una funzione volta a promuovere l'interconnessione tra dati e ricercatori, abbiamo intervistato Ilaria Capua, virologa di fama mondiale e direttrice del Dipartimento di Scienze Biomediche Comparete dell'Istituto Zooprofilattico Sperimentale delle Venezie di Legnaro (PD). Nel 2000 la dottoressa Capua ha sviluppato la strategia DIVA (Differentiating Vaccinated from Infected Animals), la prima strategia di vaccinazione contro l'influenza aviaria con un test che è in grado di svelare se gli anticorpi presenti in un soggetto sono indotti dal vaccino o da infezione. DIVA è adesso tra le strategie raccomandate da organizzazioni internazionali come la FAO e dalla Unione Europea per combattere l'Influenza aviaria su scala globale.



Fig. 10

Nel 2006, con la sua decisione di sfidare il sistema e depositare la sequenza genetica del primo ceppo africano di influenza H5N1 in GenBank e non in un database ad accesso limitato, diede inizio a un dibattito internazionale sulla trasparenza dei dati che ha cambiato i meccanismi internazionali alla base dei piani pre-pandemici. La sua iniziativa è stata ripresa dalla stampa internazionale tra cui il Wall Street Journal, il New York Times e il Washington Post oltre che dalla stampa scientifica tra cui Science e Nature. Oggi, grazie anche al suo gesto, a distanza di 5 anni, l'OMS, la FAO e l'OIE promuovono e sostengono meccanismi di condivisione più efficienti, la trasparenza dei dati e un approccio interdisciplinare per migliorare la preparazione a eventi pandemici.

Quali sono state le ragioni che nel 2006 l'hanno spinta a rendere pubblica la sequenza genetica del virus H5N1, meglio noto come virus dell'influenza aviaria?

“Questa decisione, volendo ribaltare in un certo senso il messaggio che è arrivato al pubblico, l'ho presa anche per ragioni professionalmente egoistiche. All'epoca, infatti, le sequenze dei virus influenzali non erano disponibili, e quindi io e il mio team, che ricevevamo finanziamenti per effettuare studi che prevedevano la comparazione dei virus da noi isolati su territorio italiano, non avevamo accesso alle sequenze degli altri paesi. Gli studi che fanno i virologi, in particolare quelli di filogenesi, si basano però sulle

comparazioni, un po' come succede per l'analisi delle sequenze genetiche familiari. Per fare un esempio, tracciare il percorso che ha portato una macchina americana, con tutte le componenti prodotte in quel paese, al suo proprietario, che potrebbe essere italiano, è un procedimento semplice. Se invece pensiamo a una macchina assemblata con componenti provenienti da tutte le parti del mondo, ricostruirne la storia risulta praticamente impossibile. Nel 2006 noi avevamo le sequenze genetiche dei virus, ma non potevamo svolgere un lavoro accurato a causa dell'assenza di condivisione delle informazioni con i nostri colleghi. Ciò voleva dire realizzare delle mappature e degli studi incompleti. Quindi, quando ho avuto in mano il virus che tutto il mondo avrebbe voluto guardare, e cioè l'H5N1 africano, ho pensato di dare l'esempio, mostrando che tutta la nostra comunità aveva molto più da guadagnare accettando di condividere i dati della ricerca. L'altro aspetto importante di tutta questa vicenda è che io sono stata molto fortunata perché questa mia decisione è arrivata nel momento giusto. In quel periodo la nostra comunità era infatti pronta a recepire il messaggio che volevo lanciare e a realizzare un cambiamento nella gestione di questa tipologia di dati."

In che modo i database a libero accesso e la pratica della condivisione delle informazioni tra scienziati contribuiscono al lavoro dei virologi? Questi due strumenti possono effettivamente aiutare gli scienziati a trovare, in tempi rapidi, contromisure a emergenze sanitarie come Ebola?

"Grazie alla condivisione e alla disponibilità dei dati garantita dalle risorse digitali e dal Web, gli scienziati sono agevolati nel compiere il loro lavoro, anche nel caso di crisi pandemiche, come dimostra l'Ebola Research Database. È vero però che le cose non sono così semplici come sembrano. Per una persona esterna al nostro ambiente, potrebbe sembrare automatico pensare che condividendo e conoscendo tutte le informazioni si possa lavorare meglio. Il problema è che quello che qui si propone è un vero cambio di paradigma e di impostazione nella cultura scientifica odierna. Gli scienziati con più anni di esperienza, coloro che non sono nati con gli attuali strumenti che il Web mette a disposizione, sono molto restii ad abbandonare codificati sistemi di produzione della conoscenza e di

attribuzione del merito. Quindi ciò che serve è quel cambio di mentalità verso una scienza condivisa di cui le nuove generazioni di ricercatori sono portatori, in quanto hanno sempre vissuto e fatto scienza attraverso e con le strumentazioni digitali. Bisogna perciò gestire di fatto un cambio di paradigma tenendo sempre bene a mente che tutte le opinioni sono rispettabili e degne di essere ascoltate, ma pensando che ciò che bisogna fare è andare avanti.”

A ogni problema la giusta competenza: il caso di Mathesia

Attraverso le esperienze presentate nei paragrafi precedenti, abbiamo messo in evidenza come anche i servizi del Web non direttamente destinati a favorire la collaborazione, ma che consentono l'estesa interconnessione dei complessi di dati scaturiti dall'osservazione sperimentale e la creazione di reti partecipative di scienziati, abbiano la capacità di amplificare l'intelligenza collettiva appartenuta agli esperti presenti sul Web e di accelerare il processo “creativo” di risoluzione di difficoltà teorico pratiche.

Nel campo economico e aziendale, il modello di business che consiste nell'affidare la progettazione, la realizzazione o lo sviluppo di un progetto, oggetto o idea a un insieme indefinito di persone non organizzate precedentemente è chiamato crowdsourcing⁴⁴. Questo processo viene favorito dagli strumenti che il Web mette a disposizione. Solitamente il sistema delle open call viene reso disponibile attraverso dei portali presenti in Rete. Il crowdsourcing si fonda sugli stessi meccanismi in gioco nelle collaborazioni online descritte nei paragrafi precedenti

⁴⁴ <https://it.wikipedia.org/wiki/Crowdsourcing>

ed è uno dei capisaldi della progettazione dei servizi del Web 2.0. Grazie a questo modo di fare business, le esigenze delle aziende e dei privati - che si manifestano solitamente come difficoltà tecniche - vengono proposte alla rete e alle competenze che appartengono agli utenti del Web con lo scopo di ottenere risposte. A questo proposito, analizzeremo ora una piattaforma online, Mathesia⁴⁵, che, usufruendo della riorganizzazione dell'attenzione favorita dal crowdsourcing, mette a disposizione delle aziende le competenze dei professionisti della matematica con l'obiettivo di trovare soluzioni con tempistiche prestabilite.

Mathesia è una piattaforma di crowdsourcing dedicata alla matematica applicata. Sul sito di Mathesia i problemi industriali sono trasformati in sfide di carattere matematico proposte a una rete internazionale di ricercatori, professionisti e società specializzate.



Fig. 11

La piattaforma è stata sviluppata da Moxoff, spin-off del Politecnico di Milano, e Yottacle, startup incubata all'interno di Polihub. Lanciata a dicembre 2014, Mathesia è una piattaforma di crowdsourcing 2.0 con oltre 2000 iscritti tra ricercatori, studenti, consulenti, aziende e start-up.

Attraverso un wizard, un sistema che consente di tradurre i problemi industriali in linguaggio matematico, Mathesia raccoglie le esigenze delle aziende - che sulla

⁴⁵ <http://www.mathesia.com/projects/>

base della difficoltà del problema proposto definiscono la natura della ricompensa per la sua risoluzione - e le pone all'attenzione della comunità dei ricercatori che aderiscono a Mathesia. "Il meccanismo sembra, a una prima impressione, il cuore di un videogioco evoluto. I protagonisti sono due squadre di provenienza pressoché opposta. I *brainies* sono gli esperti di matematica. I *pitchers* sono le aziende che lanciano le sfide e che vogliono migliorare la qualità dei loro prodotti o dei loro processi. Che vogliono concretamente fare innovazione."⁴⁶

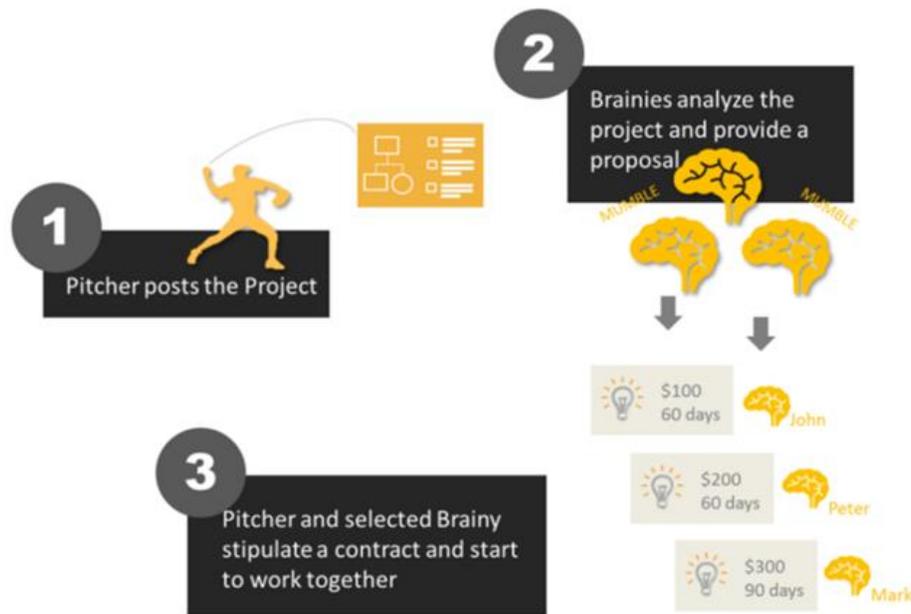


Fig. 12

Il sito di Mathesia si compone di due parti: la parte *projects*, ovvero la piattaforma di crowdsourcing vera e propria, dove le aziende che devono sviluppare nuove idee o risolvere problemi industriali, postano il loro progetto passando attraverso il wizard; e la parte *community*, una sorta di Social Network, con bacheca, blog, eventi, news, forum, dedicato ai brainies, ossia ricercatori, esperti, consulenti di matematica (quelli a cui vengono lanciate le sfide di crowdsourcing). Un altro

⁴⁶ Fonte: Ufficio Stampa Mathesia

aspetto differenziante di Mathesia rispetto ad altre piattaforme è che la sua rete è altamente specializzata (ricercatori, professori, PhD, dipendenti di spinoff e aziende di consulenza) e reputation-based.

La nostra è una piattaforma di disintermediazione che mette in contatto il mondo dell'impresa e i centri dove si fa ricerca - spiega l'amministratore delegato di Mathesia Ottavio Crivaro. Il pitcher può registrarsi gratuitamente e iniziare a vivere Mathesia. Nel momento in cui vuole lanciare una sfida, viene guidato nel descrivere in termini matematici il suo problema e i suoi obiettivi. E qui che scatta la competizione tra i brainies. La piattaforma controlla i tempi e il modo in cui le parti si confrontano. Intanto si assiste a una vera e propria gara in cui i brainies si sfidano a colpi di algoritmi e formule matematiche e raccontano come risolvere il problema. Si arriva sul traguardo formalizzando una proposta concreta e con indicazioni di carattere tecnico ed economico. È il pitcher alla fine a decretare il vincitore. È come avere un consulente sotto casa anche se insegna all'università di Atlanta. È come avere un reparto di specialisti e ricercatori che attinge idee da tutto il mondo e che fa a gara per proporre la soluzione migliore.

A oggi, sono otto le sfide portate a termine dai brainies di Mathesia, e tutte riguardano aspetti peculiari e molto tecnici delle procedure di ottimizzazione, controllo e analisi dei processi produttivi, eccone un paio.

- Azienda: THUSIA - Settore di riferimento: Sicurezza alimentare e analisi dei rischi - Una delle maggiori sfide per le imprese operanti nel campo della produzione e vendita al dettaglio di generi alimentari è sviluppare servizi finalizzati a mantenere e migliorare gli standard di igiene, sicurezza e qualità dei prodotti attraverso le informazioni raccolte dall'analisi delle procedure esistenti e dei sistemi di gestione. Per aumentare in modo adeguato la sicurezza dei loro prodotti le aziende sono perciò costrette a dotarsi di strumenti adatti al controllo e alla previsione dei rischi. Questo obiettivo viene generalmente perseguito attraverso lo sviluppo di strumenti software basati su algoritmi statistici per elaborare i dati provenienti da prove sperimentali e di laboratorio sui prodotti alimentari e da registri di

controllo e di tracciabilità, al fine di trovare correlazioni e fare previsioni sulla loro sicurezza e il rischio di previsione. Gli algoritmi utilizzati devono anche considerare correlazioni incrociate tra variabili, e fornire la visualizzazione intuitiva dei risultati. - Link:

http://www.mathesia.com/projects/business_projects/

- Azienda: SIEMENS VAI MT - Settore di riferimento: Siderurgia - Nel settore siderurgico la realizzazione di un prodotto viene ottenuta attraverso fasi successive di deformazione di un materiale per mezzo di differenti tipi di strumenti. Tali strumenti sono però soggetti a usura e hanno una vita limitata, che dipende dalla quantità del materiale lavorato, e devono essere periodicamente riparati. Una volta raggiunto il numero massimo di riparazione gli strumenti devono essere sostituiti e poi rigenerati. Queste procedure richiedono ovviamente tempo, perché per ogni sostituzione gli impianti devono essere arrestati, e gravano sui programmi produttivi delle aziende. Per ottimizzare tutto il processo è quindi necessario calcolare il numero minimo di arresti che servono per raggiungere un obiettivo di produzione e il numero minimo di strumenti che devono essere sostituiti su un periodo di tempo stabilito. - Link: <http://www.mathesia.com/projects/>

Il primo progetto descritto ha coinvolto 15 brainies, mentre il secondo 31. In merito alla sfida lanciata da SIEMENS VAI MT, Emanuele Osto, responsabile di ricerca e sviluppo dell'azienda, ha spiegato il perché sono ricorsi a Mathesia.

Ci siamo trovati di fronte alla necessità di realizzare modelli sofisticati di simulazione per cui erano necessarie skill matematiche che non avevamo internamente e quale modo migliore che confrontarsi con specialisti di tutto il mondo in un modo così agevole? Avevamo già sperimentato piattaforme di crowdsourcing per l'editing e formattazione di documenti, abbiamo scelto Mathesia perché consente di accedere a un consistente bacino di risorse, nel loro complesso molto variegato sia in termini di skills che di capacità di assorbire incarichi last minute senza una pianificazione dedicata.

Il progetto proposto da Primetals era invece un problema di ottimizzazione⁴⁷.

Un laminatoio riduce la sezione del materiale in passaggi successivi attraverso due rulli (cilindri di laminazione) sulla cui circonferenza sono ricavati dei canali (calibri). Ogni dimensione di prodotto finale che si vuole realizzare richiede l'utilizzo di un certo numero di passaggi e di calibri. Poiché per passare da un set-up di impianto all'altro occorre del tempo, è necessario poter ottimizzare numero e distribuzione dei calibri sui cilindri sia per limitare i tempi improduttivi di set-up impianto sia per ridurre il parco cilindri. La sfida era raggiungere una buona produttività al minor costo di attrezzaggio e di immobilizzazione in un parco cilindri. Il progetto è andato bene: data la complessità dell'ottimizzazione e considerato il budget limitato che abbiamo deciso di investire, abbiamo ottenuto un software che è in grado di restituire una soluzione in tempi ragionevoli per problemi di complessità medio bassa. Non converge velocemente per un problema di complessità reale, ma direi che comunque il risultato ottenuto è soddisfacente e adatto a studi di fattibilità veloci o proiezioni di costo come analisi per i nostri clienti, sempre in ambiti molto semplificati.

Il successo di Mathesia nel fornire soluzioni a problemi di matematica applicata complessi dipende perciò dalla capacità di sfruttare l'aggregazione delle competenze. La riorganizzazione dell'attenzione degli esperti sui problemi proposti dalle aziende è inoltre agevolata dal ruolo svolto dal wizard, che rende fattibile la connessione tra mondi e linguaggi differenti. Le dichiarazioni di Osto dimostrano inoltre come anche le grandi aziende e compagnie possano trarre beneficio dalla vasta disponibilità di competenze garantita da piattaforme come Mathesia. Esternalizzare attraverso le risorse online il lavoro di risoluzione dei problemi riguardanti la gestione dei processi produttivi consente infatti di diminuire i tempi e di contenere i costi.

Sebbene la piattaforma descritta non incentivi un approccio collaborativo tra i ricercatori per la risoluzione dei problemi, poiché i brainies di Mathesia si sfidano in vista di una ricompensa, tutti i contributi forniti sono a disposizione degli utenti

⁴⁷ <http://aecanalisiecalcolo.it/pubblicazioni/aec/70/crowdsourcing-e-partecipazione-collettiva/>

che sottoscrivono la proposta. Ciò determina un confronto creativo che stimola i ricercatori a fornire soluzioni migliori. Mathesia è quindi in grado di associare a ogni problema gli individui con le competenze necessarie affinché si giunga a una soluzione.

Quello che emerge dall'analisi di questa esperienza di crowdsourcing, una delle modalità primarie della progettazione dei servizi del Web 2.0 dedicati al mondo della scienza, è di fatto la sua concretezza e la sua capacità di sfruttare l'intelligenza collettiva della rete. A differenza degli altri casi analizzati finora, il lavoro della vasta rete di ricercatori presenti su Mathesia offre infatti risposte immediatamente disponibili e applicabili. Ciò è in parte dovuto al fatto che, per quanto tecnica e specialistica, la natura dei problemi affrontati esula dal campo della ricerca di base e dal contesto di lavoro dei professionisti della scienza. Fermo restando che carriera e compensi dei ricercatori sono in larga misura ancora legati alla "Scienza Accademica", termine con cui John Ziman descrive il sistema scientifico in senso classico e a cui contrappone i nuovi modelli della scienza post accademica (Ziman, 2002), senza i vincoli di valutazione previsti dal meccanismo di peer review e grazie all'incentivo rappresentato dalla ricompensa economica, i ricercatori sono più inclini a prestare il loro contributo e a mettere a disposizione la loro esperienza nella risoluzione dei problemi. Nonostante Mathesia riguardi esclusivamente esperienze di scienza applicata e il fattore economico costituisca un forte incentivo alla partecipazione delle sfide proposte agli scienziati, il portale offre una verifica importante all'idea di Scienza 2.0 che si intende qui verificare. Sono presenti alcune delle componenti in grado di far emergere l'intelligenza collettiva della rete e di giustificare la modifica della prassi del lavoro del ricercatore apportata dall'utilizzo dei Social Network: concentrazione delle competenze appartenute ai brainies su un singolo problema e conseguente riduzione dei tempi di risoluzione delle sfide presentate dalle aziende.

Capitolo 2

Polymath5 era dedicato al problema della discrepanza di Erdős. Attivato nel 2010, ha ricevuto contributi aggiuntivi nel 2012, senza raggiungere una soluzione. Il problema è stato risolto da Terry Tao usando alcune osservazioni del Polymath Project combinate con importanti recenti sviluppi nella teoria analitica dei numeri.

Con queste parole, il 22 settembre 2015, Gil Kalay, professore di matematica alla Hebrew University di Gerusalemme e a Yale, annunciava sul blog del Polymath Project, di cui è uno degli amministratori, che quella che sembrava ormai destinata a rimanere un'indimostrabile congettura della teoria analitica dei numeri, nota con il nome di problema della discrepanza di Erdős, era stata finalmente risolta. Artefice di questo importante successo Terence Tao - Terry per Kalay -, vincitore della medaglia Fields nel 2006. Tao è noto nel mondo scientifico per essere un vero prodigio. A nove anni è stato il più giovane studente della John Hopkins University a ottenere il titolo di eccezionale talento. A 11, 12, 13 anni è diventato il più giovane vincitore di una medaglia di bronzo, di argento e d'oro alle Olimpiadi della matematica. A 16 anni ha conseguito il master, a 21 i dottorato, a 24 ha ottenuto la cattedra e a 31 ha vinto la medaglia Fields⁴⁸.

Il post di Kalay mette in evidenza come la risoluzione della congettura sia stata resa possibile da alcune osservazioni e spunti emersi all'interno del quinto Polymath Project. Ma procediamo con ordine. Paul Erdős, uno dei matematici più prolifici di tutti i tempi, oltre agli essenziali contributi forniti in diversi campi della matematica, è famoso per aver posto un gran numero di problemi. Uno dei più famosi è senza dubbio quello della discrepanza. La congettura, che ha una

⁴⁸ Piergiorgio Odifreddi, 'Il Tao della matematica', Le Scienze. (dicembre 2015)

enunciazione elementare, riguarda la proprietà addizionale dei numeri primi. Considerando una qualsiasi successione infinita di +1 e -1, potremmo immaginare serie del genere:

+1-1+1+1+1-1+1+1+1...

oppure

-1-1-1-1-1-1-1-1-1...

o ancora

+1-1+1-1+1-1+1-1+1...

Se vogliamo formalizzare il problema in modo matematico, in pratica abbiamo una funzione f , definita sui numeri naturali, che può assumere solo i valori +1 e -1, per esempio:

$f(1)=+1, f(2)=-1, f(3)=-1, f(4)=+1, \dots$

Ora consideriamo la somma dei suoi primi 100 valori. Il risultato deve essere compreso tra +100 e -100 (al massimo abbiamo tutti +1 o tutti -1) e anzi, è ragionevole pensare che, se i valori di f sono distribuiti a caso, il risultato sarà molto più piccolo di 100, perché alcuni dei +1 si cancellano con i -1. In realtà il numero 100 non ha niente di speciale: se decidiamo di sommare i primi N valori, il risultato non supera N in valore assoluto (ossia non può essere maggiore di N né minore di $-N$). Fra tutte le funzioni possibili, ve ne sono due per cui il risultato in valore assoluto vale esattamente N (quella che assume solo il valore +1 e quella che assume solo il valore -1). D'altra parte, se f assume alternativamente i valori +1 e -1, la somma dei primi N valori vale 0 se N è pari, e 1 se N è dispari.

È possibile costruire funzioni f per cui si hanno comportamenti intermedi. Se prendiamo la funzione periodica di periodo 3 che assume ciclicamente i valori +1,

+1 e -1, la somma dei primi N valori vale circa $N / 3$ (il valore esatto dipende dal resto della divisione di N per 3). Mediante piccole variazioni di questo esempio, possiamo costruire funzioni che danno risultati simili. Erdős però non si è accontentato di studiare solamente il comportamento della somma dei primi N valori, ma ha considerato quello che succede facendo, più in generale, la somma dei valori della funzione f calcolata nei primi N multipli di un certo numero naturale q , cioè $q, 2q, 3q, \dots, Nq$. Naturalmente, anche in questo caso il valore assoluto delle somme in esame non può superare N , come detto sopra, perché ci sono solo N addendi.

La domanda di Erdős è questa: esiste una funzione f per cui tutte queste somme, al variare di N e q tra tutti i numeri interi positivi, sono limitate (in valore assoluto) da una certa costante C ? La costante C potrebbe essere molto grande. Il punto non è quanto sia grande C , ma il fatto che abbia un valore finito. Insomma, la domanda di Erdős è se questo tipo di successione finisce per andare all'infinito (ossia alla fine prevalgono i +1 o -1), oppure si riesce a *mischiare* i valori talmente bene che, anche prendendo quelle sottosuccessioni speciali, i +1 e -1 si annullano parzialmente tra di loro, per cui la loro somma rimane finita.⁴⁹

Nonostante la sua forma apparentemente intuitiva, la congettura di Erdős ha importanti implicazioni in diversi ambiti e ha resistito per anni ai tentativi di risoluzione da parte dei matematici, i quali, nel tempo, hanno tentato di verificare l'ipotesi del matematico ungherese anche attraverso l'adozione di metodi di dimostrazione di tipo computazionale, senza ottenere risultati convincenti.

⁴⁹ <http://maddmaths.simai.eu/divulgazione/focus/terence-tao-erdos/>

I progetti collaborativi sviluppati online

Possiamo ora rivolgere l'attenzione al vero protagonista di questa storia – e attore principale per l'economia del seguente elaborato –, il Polymath Project. Prima della dimostrazione fornita da Tao, la discrepanza di Erdős era solo una delle tante questioni ancora irrisolte della matematica. Nel gennaio del 2009 Tim Gowers, medaglia Fields nel 1998, affida a un post pubblicato sul suo blog (Gowers's Weblog) il compito di rendere pubbliche alcune considerazioni su un modello alternativo di produzione della conoscenza matematica, un modello fondato sull'estesa collaborazione tra scienziati garantita dai Social Media.

Con l'articolo 'Is massively collaborative mathematics possible?'⁵⁰, il matematico inglese riflette sulla possibilità di rendere più sociale il lavoro dei matematici e si spinge a tratteggiare le linee guida di quello che diventerà il Polymath Project, un forum online in grado di raccogliere i contributi di una vasta comunità di individui, non necessariamente professionisti della matematica, con competenze differenti e di favorire la discussione su un determinato problema, diminuendo il tempo necessario per la sua risoluzione.

It seems to me that, at least in theory, a different model could work: different, that is, from the usual model of people working in isolation or collaborating with one or two others. Suppose one had a forum (in the non-technical sense, but quite possibly in the technical sense as well) for the online discussion of a particular problem. The idea would be that anybody who had anything whatsoever to say about the problem could chip in.

Nel 2009 quella di Gowers non rappresenta però una voce isolata. A diciotto anni dalla sua comparsa, che non a caso nasce in un grande centro di ricerca con lo scopo di rendere più efficiente la trasmissione e l'elaborazione delle informazioni,

⁵⁰ <https://gowers.wordpress.com/2009/01/27/is-massively-collaborative-mathematics-possible/>

il World Wide Web è infatti una tecnologia ormai matura, le cui ricadute hanno già determinato il fenomeno dello spostamento del dibattito specialistico sulle nuove scoperte scientifiche dalla carta stampata ai blog gestiti dai singoli ricercatori. Nel 2010 la studiosa di media e comunicazione Janelle Ward, ricercatrice presso l'Erasmus University di Rotterdam, descriveva così questo processo di trasformazione della comunicazione accademica.

Un numero crescente di studiosi utilizza strumenti del Web 2.0 come blog e i wiki per condividere le proprie scoperte, discutere di nuovi sviluppi e trovare nuove strade di ricerca collaborativa [...] Nonostante esistano ancora numerosi ostacoli e dubbi da superare, il Web 2.0 rappresenta molto più che una nuova tecnologia. Il frequente utilizzo che gli stessi studiosi ne fanno, potrebbe permettere di arrivare alla soluzione di molti problemi già esistenti e di dare nuove prospettive alla ricerca nel campo della comunicazione.⁵¹

Jill Walker dell'Università di Bergen, in Norvegia, ha indicato tre tipi di blog dedicati alla ricerca, notando che alcuni si allontanano più degli altri dalle forme di pubblicazione accademica tradizionali. L'evoluzione dell'utilizzo della seconda tipologia di blog descritti è proprio quella che ha consentito a Gowers di proporre l'esperimento sociale successivamente trasformatosi nel Polymath Project.

- I blog del primo tipo sono indirizzati a quello che la ricercatrice chiama "pubblico di intellettuali" e sono prevalentemente forum per il dibattito sulle varie dottrine politiche, su femminismo, analisi di mercato e così via.
- I blog del secondo tipo sono utilizzati come registri di ricerca e servono a "raccolgere i risultati a cui una determinata ricerca ha portato e le idee che si potrebbero seguire". Questa tipologia di blog riprende la metodologia di ricerca tradizionale, assomigliando al libretto degli appunti di un sociologo

⁵¹ http://it.masternewmedia.org/2010/03/17/scrittura_collaborativa_online_in_che_mod_o_i.htm

o al taccuino su cui sono appuntati i risultati degli esperimenti di un laboratorio scientifico.

- Il terzo tipo comprende “blog pseudonimi sulla vita universitaria” che spesso “danno voce a chi si rifiuta di riverire l’esperienza della torre d’avorio”. In questi blog, i ricercatori tendono a non concentrarsi sul loro lavoro, ma su aspetti della loro vita all’interno del mondo universitario.⁵²

Forte del ruolo di promotrice della rivoluzione digitale e consapevole delle potenzialità dei nuovi media, la scienza si interroga perciò sulla validità di una nuova metodologia di produzione della conoscenza che tenga conto della connettività su larga scala garantita dal Web e dai Social Network digitali. La lucida disamina contenuta in ‘Is massively collaborative mathematics possible?’ è quindi una semplice tappa del processo di trasformazione della scienza - o almeno della comunicazione scientifica - messo in atto con l’avvento di Internet. Secondo quanto sostenuto in un altrettanto fondamentale post di Michael Nielsen (“Doing science online”)⁵³, da cui Glowers prende le mosse per la sua analisi, i vantaggi prodotti da una collaborazione estesa sviluppata attraverso blog o wiki riguardano aspetti differenti della metodologia di lavoro prescritta da un sistema di impresa di tipo scientifico. Ciò comporta che il modello alternativo di lavoro supportato dai Social Media proposto dai due ricercatori non implica lo stravolgimento della pratica scientifica, né dal punto di vista della diffusione pubblica dei suoi contenuti né da quello che riguarda le metodologie di lavoro adottate dai ricercatori, ma ne rappresenta solo una naturale evoluzione sia tecnologia che sociale. Lo stesso genere di trasformazioni prodotte dall’introduzione nella scienza passata di nuovi sistemi e tecnologie di comunicazione. Secondo Pietro Greco è proprio grazie

⁵² Axel Bruns, ‘Blogging from inside the ivory tower’. (2006)

⁵³ <http://michaelnielsen.org/blog/doing-science-online/>

all'avvento del libro che nel Seicento le potenzialità di creatività scientifica latenti diventano attuali⁵⁴. Per Lucio Russo, fisico e storico della scienza, la scienza in senso moderno è nata e si è affermata in epoca ellenistica⁵⁵. Prima di declinare in epoca romana, si può infatti notare non solo che è proprio in quell'epoca che la cultura scritta dei Greci si affranca da quella orale con l'affermazione di uno strumento di comunicazione chiamato libro, ma anche che gli uomini di scienza ellenistici sono soliti scrivere libri in cui non c'è commistione con la religione, l'arte (la tecnica) o la filosofia e che comunicano a tutti in maniera trasparente un sapere tendenzialmente universale. Sempre Greco, in merito al contributo del libro e dell'invenzione della stampa a caratteri mobili nelle trasformazioni che caratterizzarono la rivoluzione scientifica del Seicento:

[...] il mezzo della comunicazione modifica anche la qualità della comunicazione. In particolare, il fatto che l'ars naturaliter scribendi, la tecnica naturale di scrivere, quella a mano, si modifica in ars artificialiter scribendi, in riproduzione a stampa di un manoscritto, determina delle modificazioni profonde nei contenuti e che non sono state ancora compiutamente indagate dagli statuti della comunicazione letteraria. Tra gli statuti letterari modificati vi sono, naturalmente, anche quelli della comunicazione scientifica. [...] L'importanza del libro nel '600 è tale che esso diviene una metafora del mondo. La natura è un libro scritto in caratteri geometrici, dice Galileo. La natura è un codice, anzi un codex vivus, ribadisce Campanella. Quanto a Francis Bacon, non ha dubbi: l'invenzione della stampa a caratteri mobili è, con la bussola e la polvere da sparo, la maggiore e la più significativa tra le innovazioni tecnologiche che rendono quella attuale, il Seicento, un'epoca di progresso. Anzi, come rileva ancora Tommaso Campanella nella Città del Sole, quelle tre invenzioni coincidono con un'improvvisa e inedita accelerazione della storia.

Quindi, a dispetto della diffidenza con cui parte della comunità scientifica accolse il sistema della pubblicazione delle scoperte e con cui viene accolta oggi la collaborazione estesa veicolata dalle risorse della Rete, l'apertura dell'impresa

⁵⁴ 'L'idea pericolosa di Galileo: storia della comunicazione della scienza nel Seicento', Utet Università. (2009)

⁵⁵ 'La rivoluzione dimenticata', Feltrinelli. (1998)

scientifica a una crescente condivisione del sapere prodotto è un tema ricorrente nella storia della scienza e consente di incrementare l'espressione delle potenzialità di creatività scientifica latenti. I cambiamenti indotti dall'avvento delle risorse online, secondo quanto sostenuto da Nielsen, pur non avendo avuto un immediato vantaggio, stanno confermando, proprio come avvenuto durante la rivoluzione scientifica, la loro inevitabilità e il loro fondamentale contributo nel lungo termine.

Galileo could not imagine a world in which it made sense for him to freely share a discovery like the rings of Saturn, rather than hoarding it for himself. Certainly, he couldn't share the discovery in a journal article, for the journal system was not invented until more than 20 years after Galileo died. Even then, journals took decades to establish themselves as a legitimate means of sharing scientific discoveries, and many early scientists looked upon journals with some suspicion. The parallel to the suspicion many scientists have of online media today is striking. [...] Most scientists, like Galileo, can't conceive of a world in which it makes sense to share all that information, in which sharing information on blogs, wikis, and their descendants is viewed as being (potentially, at least) an important contribution to science [...] Over the short term, things will only change slowly. We are collectively very invested in the current system. But over the long run, a massive change is, in my opinion, inevitable. The advantages of change are simply too great.

Serendipità vs serendipità pianificata

Ma quali sono i benefici derivanti dall'utilizzo delle risorse online nel problem solving e nella produzione di nuova conoscenza scientifica? Quali gli svantaggi? E soprattutto, a quali criteri dovrebbe sottostare un tipo di scienza collaborativa sviluppata attraverso le risorse del Web 2.0? Gowers individua tre potenziali e ovvie prerogative dell'approccio collaborativo alla risoluzione dei problemi matematici. La prima coinvolge quell'ambito della ricerca e degli studi scientifici

che hanno a che fare con il caso. A volte, sostiene infatti il matematico inglese, per adottare un'adeguata strategia di risoluzione, è necessario che la fortuna metta di fronte allo scienziato la giusta idea in grado di indirizzarlo verso la migliore strategia di risoluzione. Un elevato numero di individui avrebbe perciò maggiori probabilità di imbattersi nell'ipotesi adatta di quanto ne abbia un ricercatore impegnato a lavorare per conto proprio o con una ristretta cerchia di collaboratori su una questione di rilevanza scientifica.

Sometimes luck is needed to have the idea that solves a problem. If lots of people think about a problem, then just on probabilistic grounds there is more chance that one of them will have that bit of luck. (Nielsen, 2012)

Michael Nielsen, nel libro 'Le Nuove vie della scoperta scientifica'⁵⁶, in cui tenta di sistematizzare le riflessioni sull'utilizzo del Web in ambito scientifico contenute nel suo blog, definisce questa proprietà emergente dalla condivisione online del lavoro con il termine "serendipità pianificata". Il mondo scientifico riconosce da tempo l'importanza della serendipità, ovvero la predisposizione a fare scoperte inaspettate mentre si sta cercando altro, per lo sviluppo e la pianificazione del lavoro dei ricercatori. Si deve a Robert K. Merton l'introduzione del concetto all'interno di quel quadro di norme sociali non scritte che contraddistinguono la prassi scientifica contemporanea. Se la serendipità mertoniana è però inerente all'esperienza personale del singolo scienziato, il fenomeno descritto da Nielsen scaturisce invece dalla capacità dei media digitali - e in particolare dei social media - di agevolare le connessioni casuali tra individui in possesso di microcompetenze differenti.

Per Social Media, o media sociali, lo ricordiamo, intendiamo tutte quelle tecnologie o pratiche online che consentono di condividere esperienze, opinioni, contenuti e informazioni. I media sociali possono assumere forme diverse tra cui testo,

⁵⁶ 'Le nuove vie della scoperta scientifica', Einaudi. (2012)

immagini, audio e video. I siti che appartengono a questa categoria si servono di strumenti quali bacheche elettroniche, forum, podcast, bookmark, comunità, wiki, weblog e molto altro ancora. Il termine “sociale” indica un mezzo di comunicazione che consente e incoraggia la partecipazione attiva dei fruitori alla costruzione, aggiornamento o valutazione dei contenuti. Gli utenti sono invitati a commentare, votare, categorizzare i contenuti presenti su un sito e a condividerli con altri. La socialità favorisce la costituzione di gruppi sociali con gusti, interessi e obiettivi comuni.

A differenza dei media di massa, in cui il contenuto è il frutto di un sistema di produzione industriale di tipo verticale, i media sociali producono il contenuto attraverso un sistema di co-costruzione di tipo orizzontale, in cui gli utenti non sono destinatari passivi, ma hanno la possibilità di interagire e di creare nuovi messaggi. Le caratteristiche che definiscono un medium come sociale sono la presenza di un contenuto digitale, una rete neutrale di computer, una massa critica di utenti. Il contenuto digitale consente di automatizzare molti passaggi necessari alla produzione e distribuzione di oggetti sociali. La rete mette in connessione creatori, distributori e manipolatori di informazioni in modo neutrale, ovvero consente la comunicazione tra i vari nodi, a prescindere dal contenuto trasmesso. Gli utenti devono rappresentare una massa critica, cioè superare la soglia quantitativa minima necessaria a produrre un mutamento qualitativo. I media sociali per essere tali devono essere di massa, nel senso che devono essere condivisi da una moltitudine di individui non omogenei.

I Social Network digitali, grazie alle caratteristiche appena descritte, consentono di amplificare l'intelligenza inespressa della rete, rappresentata dal sapere appartenuto ai singoli utenti, attraverso la “riorganizzazione dell'attenzione” nei confronti di un unico contenuto digitale. Accentrando l'interesse e le competenze sopra un'unica questione o problema, gli individui, interconnessi per mezzo dei servizi sociali online, hanno la possibilità di contribuire collettivamente per

ottenere un obiettivo comune, come sviluppare nuova conoscenza o risolvere difficoltà teoriche.

Alla base di questo quadro generale c'è il fatto che collettivamente sappiamo molto di più anche degli individui più brillanti. Secoli fa era forse possibile per una persona dotata - un Aristotele o un'Ipazia o un Leonardo - superare tutti gli altri in molti ambiti del sapere. Oggi non è più così, perché lo scibile umano si è ampliato moltissimo. Il sapere è stato decentralizzato e risiede in tante menti diverse, persino nelle aree specializzate. (Nielsen, 2012)

La natura connettiva dei Social Media può perciò risultare di capitale importanza nel contrastare il fenomeno centrifugo della frammentazione del sapere risultante dalla suddivisione del lavoro scientifico in ambiti sempre più specialistici. Questo perché la concentrazione delle microcompetenze degli individui per mezzo dei grandi progetti collaborativi online consente la proliferazione e la condivisione estesa delle idee, aumentando le probabilità di far emergere l'intuizione che, se sviluppata correttamente, avrà le maggiori speranze di superare un problema teorico. Una ricerca del 2006 della Harvard Business School, in collaborazione con la Copenhagen Business School, dal titolo 'The Value of Openness in Scientific Problem Solving'⁵⁷ che tentava di verificare la capacità di risolvere problemi ancora aperti da parte di gruppi di esperti al variare di condizioni quali la condivisione delle informazioni, le tipologie di competenze appartenute agli individui coinvolti nella risoluzione e il numero di contributori, ha dimostrato che può essere vantaggioso riunire differenti risolutori di problemi e incoraggiarli a collaborare. Lo studio indica infatti che gruppi composti da persone in possesso di conoscenze differenti hanno maggiori possibilità di avere successo di quante ne abbiano gruppi altamente qualificati, ma composti da individui con competenze omogenee. La ricerca ha inoltre confermato che l'inclusione casuale di diversi

⁵⁷ Karim R. Lakhani, Lars Bo Jeppesen, Peter A. Lohse e Jill A. Panetta. 'The Value of Openness in Scientific Problem Solving'. (2006)

collaboratori nella soluzione dei problemi può aumentare le prestazioni di calcolo. Secondo Karim R. Lakhani, Lars Bo Jeppesen, Peter A. Lohse e Jill A. Panetta, autori dello studio, è ragionevole pensare che un approccio al problem solving che contempi e favorisca la trasparenza, il libero accesso all'informazione e la collaborazione in tutto il processo scientifico di risoluzione dei problemi ha le potenzialità per incrementare i tassi di risoluzione. Questa efficacia derivante dall'apertura dei processi risolutivi implica inoltre una sistematica inclusione di diverse prospettive metodologiche ed euristiche all'interno del problem solving di tipo scientifico.

È proprio questa caratteristica ad aver reso Mathesia un'esperienza di successo. La capacità di questa piattaforma di fornire soluzioni a problemi di matematica applicata complessi dipende infatti dalla capacità di sfruttare il meccanismo dell'aggregazione delle competenze. La riorganizzazione dell'attenzione degli esperti sui problemi proposti dalle aziende è inoltre agevolata dal ruolo svolto dal wizard, che rende fattibile la connessione tra mondi e linguaggi differenti. Mathesia rappresenta quindi un esempio dell'applicazione del meccanismo descritto, secondo cui un gruppo di esperti in possesso di conoscenze disomogenee impegnati a risolvere un problema aumenta la probabilità di ottenere una soluzione.

Anziché confidare sulle capacità personali di un singolo scienziato o di un ristretto team di ricercatori, che potrebbero spendere anni prima di imbattersi in un'idea vincente - eventualità descritta dal concetto di serendipità fortuita di matrice mertoniana - il Web, riunendo l'apporto di molte menti connesse tra loro, mette a disposizione degli scienziati una vasta gamma di contributi potenzialmente adatti a superare le difficoltà che potrebbero presentarsi durante la risoluzione di un problema. In definitiva, i network digitali svolgono la funzione di porre la giusta persona, quella in possesso dell'adeguata microcompetenza, di fronte al giusto problema, nel giusto momento.

La serendipità pianificata scaturita dalle collaborazioni creative di cui parla Nielsen, che mutua il termine dal mondo della programmazione e dell'open source, è dunque la caratteristica che rende gli strumenti online particolarmente adatti al problem solving e alla pratica scientifica tutta. Questa capacità dei Social Media si pone in contrasto con quello che Anabel Quan-Haase e Kim Martin, in uno studio del 2011, chiamano invece "perdita di serendipità"⁵⁸. Secondo i due sociologi canadesi, i processi di ricerca non lineari e non focalizzati che sottendono alla capacità di far emergere il fenomeno della serendipità ricoprono un ruolo centrale nella realizzazione delle scoperte scientifiche. Al contrario, i motori di ricerca e i contenuti digitali di cui fanno uso oggi gli studiosi, che permettono un rapido reperimento delle informazioni, non consentirebbero di passare in rassegna le conoscenze in maniera approfondita, diminuendo così le probabilità di imbattersi fortuitamente in qualche idea. Nonostante ciò, conclude la ricerca, servizi Web come i Social Media sono in grado di generare ambienti digitali in cui sono favorite le connessioni e le condivisioni casuali tra individui.

L'autore di "Le nuove vie della scoperta scientifica" chiarisce che, come in una reazione a catena, maggiore è l'estensione della collaborazione, maggiore sarà la possibilità di creare connessioni tra individui con competenze non omogenee e lo scambio casuale dei contenuti all'interno dei Social Media. Grazie all'estensione dei gruppi collaborativi, il meccanismo della serendipità pianificata diventerebbe perciò autosufficiente.

Nella collaborazione creativa succede qualcosa di simile (reazione a catena n.d.r.). Quando cerchiamo di risolvere da soli un difficile problema creativo quasi tutte le nostre idee sono vicoli ciechi. Ma in una buona collaborazione creativa, alcune delle nostre idee - che da soli non riusciremmo a sviluppare appieno - stimolano qualcun altro a farsi venire altre idee. Queste a loro volta stimolano ancora altre persone a

⁵⁸ Anabel Quan-Haase and Kim Martin, 'Seeking knowledge: An exploratory study of the role of social networks in the adoption of E-books by historians'. (2011)

farsi venire ancora altre idee. E così via. Idealmente raggiungiamo una specie di massa critica conversazionale, dove la collaborazione diventa autostimolante e noi traiamo il beneficio reciproco della connessione fortuita più e più volte. È proprio questa transizione che la serendipità pianificata permette di fare e che avviene quando la collaborazione aumenta di dimensione, incrementando il numero e la varietà dei partecipanti, e dunque incrementando la possibilità che un'idea stimoli un'altra, nuova idea.

Evoluzione dei dati sperimentali, interpretazione e collettivizzazione del lavoro scientifico

Prima di avere un quadro esaustivo delle condizioni che hanno reso possibile la nascita del Polymath Project, è necessario riflettere anche sull'evoluzione della conoscenza prodotta in seno ai vari ambiti di ricerca e sul modo in cui essa è oggi prodotta, poiché le trasformazioni subite dai dati sperimentali raccolti in seno all'attività scientifica e le loro conseguenti ricadute sulla prassi giornaliera dei ricercatori rappresentano le ragioni del ricorso a una Scienza 2.0 e un'ulteriore prova di quanto Web e Social Network possano oggi incidere sulla metodologia della ricerca.

Negli ultimi cinquant'anni la scienza ha visto un progressivo incremento delle proprie aree di indagine sperimentale.

È giusto descrivere una specializzazione scientifica come area di problemi, poiché si tratta non tanto di un sistema conoscitivo quanto di un programma di ricerca. In linea di principio, essa include gli obiettivi attuali di un insieme di progetti genericamente connessi, che determinano caratteristiche in comune. Ma se un tale programma persiste per qualche decennio, allora diviene un'istituzione culturale, tramandata di generazione in generazione come tradizione epistemica, tecnica e comune.

Come sostiene John Ziman⁵⁹, la suddivisione degli ambiti della ricerca, oltre a far emergere il fenomeno della specializzazione delle competenze, ha reso possibile la moltiplicazione di linguaggi e metodologie che fossero conformi ai nuovi contesti sperimentali. Di certo, questo processo ha contribuito a quella sorta di dispersione delle competenze a cui fa riferimento Nielsen nella sua analisi. Nonostante ciò, la scienza, nelle rispettive aree che le sono ormai proprie, ha proseguito nel tempo a produrre conoscenza grazie all'interpretazione dei dati scaturiti dall'indagine sperimentale. La rivoluzione digitale e informatica hanno messo inoltre a disposizione degli scienziati nuovi e potentissimi strumenti tramite cui osservare i fenomeni naturali. La prassi scientifica, grazie all'utilizzo di simili strumenti, vive un'epoca d'oro. La quantità e la risoluzioni dei dati è tale da suggerire a molti, come abbiamo già indicato nell'introduzione di questo lavoro, l'avvento di un quarto paradigma per la scienza: il salto di quantità nella produzione di informazione è tale da costituire di per sé un salto di qualità.

Tuttavia, questo aspetto apre il problema dell'interpretazione delle esperienze sperimentali effettuate. Il processamento dei dati generati è ormai tale da richiedere strumenti algoritmici di elaborazione sempre più potenti, una quantità di tempo superiore rispetto al passato e lo sforzo condiviso di un ingente numero di scienziati. Come se non bastasse, l'interpretazione di questi differirà al variare sia dell'ambito di osservazione che, ovviamente, a quello di applicazione dell'interpretazione stessa, incrementando la distanza e una certa incomunicabilità tra le varie discipline scientifiche. È infatti impossibile supporre che un fatto empirico osservato tramite gli strumenti sperimentali possa essere esente dal contesto nel quale è stato osservato. Il contesto contiene sempre caratteristiche sia soggettive che oggettive, di solito strettamente correlate. Uno strumento sofisticato, del resto, incarna molti concetti teorici ma si tratta solo di elaborazioni ed

⁵⁹ 'La vera scienza', Edizioni Dedalo. (2002)

estensioni delle teorie necessarie a un osservatore preparato per vedere ciò che è rilevante dal punto di vista scientifico nella sua personale esperienza del mondo. Pertanto, pur non potendo negare il sostanziale ruolo delle risorse informatiche e digitali nel perseguimento della nuova conoscenza, tali strumenti sono allo stesso tempo responsabili dell'accresciuta difficoltà nell'attività di comprensione e interpretazione dei dati. La dispersione e le articolate specificità teoriche con cui si presentano oggi le osservazioni sperimentali prodotte all'interno dei vari ambiti di ricerca non si prestano quindi a quella apertura interdisciplinare che da sempre caratterizza la scienza e che consente l'individuazione e la prolifica stimolazione di nuove ipotesi.

David Weinberger ha coniato il termine *sovraccarico cognitivo* per descrivere il fenomeno della crisi della conoscenza risultante dall'impossibilità di interpretare la sovrabbondanza di dati che siamo oggi in grado di generare. Secondo Weinberger, mentre nel passato la quantità dell'informazione prodotta era tale da poter essere gestita attraverso i classici canali della comunicazione, che applicavano filtri interpretativi modellati sul medium cartaceo al fine di estrapolare la conoscenza utile contenuta nei dati, oggi la mole e la natura articolata delle informazioni hanno bisogno di essere veicolate da un tipo di infrastruttura più dinamica e interconnessa per avere la speranza di essere visti, interpretati, e di potersi levare al grado di conoscenza pratica.

[...] la strategia fondamentale che abbiamo usato per comprendere un mondo che va ben oltre le capacità del nostro cervello è consistita nel filtrarlo, vagliarlo e infine ridurlo a qualcosa di più gestibile. Abbiamo tenuto sotto controllo la manichetta delle informazioni riducendone il flusso. Ci siamo riusciti con un elaborato sistema di filtri editoriali che hanno impedito che la maggior parte di ciò che veniva scritto fosse pubblicato; con un elaborato sistema di filtri curatoriali che hanno evitato che la maggior parte di ciò che veniva pubblicato fosse custodito nelle nostre biblioteche e librerie locali; e con un elaborato sistema di filtri professionali grazie ai quali molti di noi non erano tenuti a conoscere la maggior parte di ciò che superava gli altri filtri. [...] La carta resterà con noi ancora a lungo, sebbene la situazione sia

chiaramente a favore del nuovo medium digitale e interconnesso. [...] il collegamento della conoscenza - il networking, la messa in rete o in circolo - sta in realtà cambiando la nostra più antica e fondamentale strategia del sapere.⁶⁰

Ma se il Web si presenta come unico medium capace di incarnare questo nuovo e caotico sapere, non è altrettanto capace di restituirci una conoscenza strutturata e lineare: l'inclusività di Internet ha di fatto reso disorganica e frammentata sia l'informazione che la competenza appartenuta dagli utenti della rete, rendendo difficile la ricerca di significato nel mare magnum dei dati. Per far fronte all'aumento incontrollato dell'informazione scientifica e alla dispersione di conoscenza ed expertise, oltre ad augurarsi che il progresso tecnologico ci metta a disposizione strumenti algoritmici di elaborazione più potenti, è perciò importante implementare strategie collaborative e di condivisione online complementari, che possono essere definite di tipo sociale, in grado di rendere più efficiente l'interpretazione dei dati e l'estrapolazione di nuova conoscenza da essi. Tali strategie dovrebbero però essere in grado di far emergere i due meccanismi della "riorganizzazione delle microcompetenze" e della "serendipità pianificata" che abbiamo visto comparire all'interno dell'analisi delle esperienze scientifiche sviluppate attraverso i Social Media. È ancora una volta Weinberger a definire la complementarietà funzionale tra strumenti algoritmici e strumenti sociali online di interpretazioni dei dati

Di conseguenza abbiamo rapidamente sviluppato una serie di tecnologie in grado di aiutarci (a contrastare il sovraccarico cognitivo n.d.r.). Ricadono in due categorie - algoritmica e sociale - anche se la maggior parte degli strumenti di cui oggi disponiamo le contempla entrambe. Le tecniche algoritmiche sfruttano le enormi memorie e la grande potenza di calcolo dei computer per manipolare vorticosi ammassi di dati, al fine di trovare delle risposte. Gli strumenti sociali ci aiutano a individuare gli elementi interessanti usando come guida le scelte dei nostri amici.

⁶⁰ 'La stanza Intelligente', Einaudi. (2012)

La strategia sociale per l'estrapolazione di significato dalla vasta quantità di dati generati dall'osservazione sperimentale non è però prerogativa dei soli Social Media. Come abbiamo osservato nel primo capitolo con la descrizione di supporti come GenBank ed Ebola Research Database, anche i servizi online di archiviazione possono contribuire a far emergere i due meccanismi della "riorganizzazione delle competenze" e della "serendipità pianificata". Queste piattaforme favoriscono infatti la connessione di dati specifici con gruppi di ricerca in possesso delle competenze necessarie a interpretarli. Pur non collaborando classicamente tra di loro, i vari centri di ricerca dispersi su tutto il pianeta, accomunati dall'utilizzo di un archivio condiviso di gestione dei dati sperimentali, lavorano di fatto a un unico obiettivo, estrapolare conoscenza dall'informazione contenuta all'interno di una vasta mole di sequenze genetiche o trovare soluzioni per contrastare la diffusione di una pandemia. Ciò non sarebbe ovviamente possibile se tali archivi non fossero immersi in un ambiente interconnesso e se non ci fosse un libero accesso e una libera condivisione dei contenuti attraverso un unico database, oltre ovviamente a una metodologia e a modelli di base comuni, soprattutto per quanto riguarda le informazioni che descrivono i dati stessi (i metadati). Fattore imprescindibile dell'implementazione della strategia sociale mediante l'utilizzo degli archivi online è quindi quello di rendere liberamente accessibili le informazioni presenti sul Web e passare a un modello di pubblicazione online di tipo open access, e cioè una modalità che consenta accesso libero e senza restrizione ai dati sperimentali primari.

L'analisi della trasformazione dei dati scientifici e della conseguente evoluzione degli strumenti e del metodo di lavoro nell'ambito della ricerca appena proposta ci consente di dare conto del passaggio, descritto nel primo capitolo, da un sistema di tipo classico e accademico, pienamente aderente alle norme mertoniane, al modello "post-accademico" di produzione della conoscenza scientifica proposto da Ziman e

caratterizzato dal fenomeno della collettivizzazione della ricerca e della sua transdisciplinarietà. Secondo il fisico e sociologo neozelandese, anche quando non è impegnata a occuparsi di problemi prettamente pratici, la scienza post-academica si contraddistingue infatti per la sovrapposizione delle diverse scienze, le quali vengono a compenetrarsi lungo innumerevoli canali cognitivi e tecnici.

Quando la ricerca è orientata verso problemi più ampi, sia che questi sorgano in contesti applicati o di base, essa dovrà necessariamente rapportarsi alle attività comuni provenienti da diverse discipline

Se collettivizzazione e transdisciplinarietà della ricerca sono le risposte sociali indotte da un accumulo di dati, conoscenze e tecniche, l'assunzione della Rete e dei servizi di collaborazione online sia come ragioni sociali - vista la centralità della comunicazione all'interno del sistema scienza - che come principi metodologici della scienza post-academica, di cui la scienza 2.0 non è che altro una manifestazione, risulta quindi fondata.

La valutazione della produzione collaborativa online

Torniamo a parlare della nascita del Polymath Project e di quali fattori Tim Glowens - e di concerto Nielsen - ritiene possano ostacolare la nascita di un progetto collaborativo di tipo scientifico sviluppato attraverso un Social Network. Abbiamo già fatto riferimento al fatto che i due scienziati ritengono il passaggio da un sistema del riconoscimento del valore della scoperta scientifica incentrato sul classico modello della pubblicazione a uno basato sulla collaborazione estesa sviluppata sui Social Media sia una trasformazione necessaria e che comporterà enormi vantaggi per l'impresa scientifica.

Nell'analisi dei vantaggi derivanti da una scrittura collaborativa e creativa sviluppata attraverso i wiki, Janelle Ward osserva che

[...] il processo redazionale universitario tradizionale non è affiancabile a questo modo di raccolta di informazioni che va dal basso verso l'alto e che difficilmente supererà il dogma accademico del diritto d'autore. Ma l'effetto migliore, tuttavia, è che tale collaborazione crea una nuova forma di revisione paritaria tra studiosi internazionali e quindi allarga il bagaglio di conoscenze ed esperienze a disposizione. È però chiaro che la scrittura collaborativa online è ancora in fase di rodaggio e ha bisogno di tempo per acquisire stabilità e legittimità. Tuttavia i critici più conservatori ruotano ancora attorno al concetto di diritto d'autore come unico mezzo per dare credito, nel modo più giusto, a chi scrive gli articoli. Ecco ciò che dice chi si schiera contro l'utilizzo dei wiki.

Il metodo accademico di revisione della ricerca, oltre a essere universalmente accettato, dà la certezza che vengano riconosciuti i giusti meriti per il lavoro svolto, meriti che si tradurranno in prestigio e in possibilità di carriera per lo scienziato o il gruppo di scienziati responsabili di una scoperta. Vista la grande mole di contributi e di persone presenti sui Social Media, ciò che sembra mancare ai grandi progetti collaborativi proposti da Gowers è proprio questa facoltà di distinguere a chi dovrebbe essere riconosciuto il merito della risoluzione di un problema e di mostrare quali vantaggi, in termini di credito scientifico, potrebbero derivare dal contribuire alla risoluzione dello stesso.

La prossima domanda ovvia è questa. Perché qualcuno dovrebbe accettare di condividere le proprie idee? Sicuramente lavoriamo su problemi al fine di essere in grado di pubblicare le soluzioni e ottenere credito per loro. E se la collaborazione massiva portasse a una buona idea? Non c'è il pericolo che qualcuno riesca a utilizzare l'idea per risolvere il problema in fretta e di pubblicarlo individualmente? È qui che sta la bellezza di blog, wiki, forum ecc.: sono completamente pubblici, come tutta la loro storia. Per vedere l'effetto che ciò potrebbe avere, immaginate che un problema sia stato risolto nel corso della collaborazione attraverso i commenti su un post del blog. Supponiamo che il blog sia stato piuttosto attivo e che il post abbia sempre avuto diversi commenti interessanti. E supponiamo che tu abbia avuto un'idea che pensi sia quella giusta. Invece della solita reazione di aver paura di

condividerla perché non vuoi che qualcun altro ti batta sul tempo nella soluzione, dovresti aver paura di non condividerla perché qualcuno ti potrebbe battere sul tempo nella pubblicazione di quella particolare idea. E se il problema fosse alla fine risolto, e pubblicato sotto un qualche pseudonimo, "come Polymath", questo potrebbe spiegare, con una nota in calce o un link al blog, il metodo di risoluzione del problema, quindi chiunque potrebbe andare sul blog e guardare tutti i commenti. E lì troverebbe la vostra idea e saprebbe esattamente in che modo essa abbia contribuito. Ci potrebbero essere delle argomentazioni sul fatto che alcune idee hanno dimostrato di essere più importanti per la soluzione, ma almeno tutte le prove sarebbero lì, a disposizione di chiunque voglia guardarle. È vero, potrebbe essere molto difficile descrivere sul tuo CV: "ho avuto un'idea che si è rivelata essenziale per la soluzione del problema Polymath". Ma se hai dato un contributo significativo a diversi progetti di collaborazione di questo tipo, allora si potrebbe anche cominciare a guadagnare una reputazione tra le persone che leggono i blog matematici, e ciò potrebbe iniziare a contare qualcosa (anche se non conta più di tanto ora, questa modalità di risoluzione è destinata a diventare sempre più importante). E potrebbe allora non essere così difficile immaginare come tutto questo possa essere segnalato sul tuo CV: si potrebbe iniziare a vederti come un coautore, con l'ulteriore vantaggio che la gente potrebbe scoprire esattamente ciò a cui hai contribuito. E per quanto riguarda la persona che cerca di tagliare la corda quando il progetto non è ancora concluso? Beh, potrebbe succedere, ma tutti saprebbero quello che hai fatto. L'arbitro del paper, si spera, potrebbe dire, per esempio: "ehm, non dovresti essere accreditato dal Polymath per il tuo cruciale Lemma 13?" E questa sarebbe una cosa piuttosto imbarazzante da dover fare.

Gowers sostiene dunque che l'accREDITAMENTO scientifico derivante dalla proficua collaborazione degli utenti ai progetti sviluppati online possa essere garantito dal fatto che i Social Media rendono di pubblico dominio e costantemente disponibili tutti i passaggi che contraddistinguono la soluzione di un problema. La comunità degli utenti coinvolti nel progetto si fa inoltre garante della bontà dei contributi del singolo individuo, decidendo se le idee proposte abbiano o meno lo status per essere messe in evidenza una volta che la collaborazione abbia raggiunto i propri obiettivi.

Nonostante la fiducia di Gowers nei confronti del nuovo sistema di revisione paritaria, la comunità scientifica sembra restia ad abbandonare il concetto di diritto

d'autore, unico criterio a oggi esistente per determinare la qualità e il merito da attribuire al lavoro accademico degli scienziati. Il metodo di valutazione proposto da Gowers sembra infatti porre un'eccessiva fiducia nella possibilità che all'interno della collaborazione online di tipo scientifico gli autori di contributi di maggiore importanza siano disposti a condividere la paternità di un risultato; e di conseguenza nel fatto che gli utenti responsabili di contributi di minore spessore, ma che potrebbero - con ulteriori riflessioni - arrivare a una risoluzione definitiva per conto proprio, siano spinti a fare la stessa cosa. In sintesi, il problema si traduce, al di là delle migliori intenzione del Polymath Project e dei suoi ideatori, nell'attestazione del merito della scoperta.

Le problematiche appena emerse spingono a una considerazione sulla natura della Scienza 2.0. Secondo quanto sostenuto da Davide Bennato nel libro 'Scienza Connessa: Rete, Media e Social Network', i modelli alternativi di produzione del sapere scientifico, come quello rappresentato dalla Scienza 2.0, sono in realtà modelli che devono confrontarsi anche con metodi alternativi di legittimazione del sapere scientifico. Considerati quindi i grandi risultati ottenuti dalla scienza in termini di avanzamento culturale e tecnologico, il passaggio a un differente sistema scientifico potrebbe risultare del tutto infondato e privo di effettivi benefici, se non addirittura dannoso. Tuttavia, la Scienza 2.0, almeno nell'accezione che vuole essere qui dimostrata, non rifiuta i meccanismi classici di revisione e legittimazione, ma trova la sua migliore descrizione in quel nuovo approccio alla collaborazione visibile nelle esperienze descritte nel primo capitolo. Una collaborazione che, a dispetto delle indicazioni di Gowers, emerge dall'utilizzo condiviso degli strumenti online di collaborazione e dalla necessità di far fronte alla quantità e alla complessità dei dati scientifici, indirizzando i ricercatori verso soluzioni e applicazioni innovative. Questo è il motivo del successo di piattaforme come Mathesia e TEP, che si contraddistinguono per un certo grado di concretezza

e nell'indipendenza degli utenti di poter sfruttare a loro esclusivo vantaggio i risultati dei processi collaborativi online.

In breve, il supporto degli scienziati veicolato dalle risorse online e dal meccanismo del crowdsourcing al mondo delle aziende e delle imprese sembra essere facilitato da questi fattori e offre vantaggi riguardevoli. Il successo delle Tematic Explication Platform ha invece a che vedere sia con la quantità dei dati analizzati sia con la natura dei tool messi a disposizione. La possibilità di utilizzare le risorse condivise di processamento e di mettere a disposizione i risultati consente agli scienziati di collaborare all'ampliamento del portale mantenendo l'autonomia di sfruttare quegli stessi risultati per lo sviluppo di applicazioni.

I criteri per la contribuzione sui social media e l'amplificazione dell'intelligenza collettiva

Terminata la descrizione dei vantaggi derivanti dall'amplificazione delle collaborazioni creative implementate attraverso i Social Network digitali, ci resta da analizzare quali siano i criteri che i partecipanti ai progetti collaborativi debbano rispettare per rendere efficace questo genere di esperienze e per amplificare l'intelligenza in loro possesso. Le regole della collaborazione online si riferanno ovviamente alle caratteristiche degli strumenti utilizzati.

Come abbiamo avuto modo di osservare, Nielsen ritiene che il successo dei progetti collaborativi sviluppati online risieda nella capacità di amplificare l'intelligenza collettiva inespresa appartenuta agli utenti della rete. Nel libro 'The Wisdom of Crowds'⁶¹, il giornalista americano James Surowiecki osserva che "un

⁶¹ 'The Wisdom of Crowds', Anchor Book. (2004)

gran numero di persone è più intelligente di un élite costituita da pochi, non importa quanto brillanti; sono più bravi nel risolvere problemi, nell'incentivare le innovazioni, nel prendere decisioni e persino nel prevedere il futuro". A questo proposito, Nielsen individua tre criteri in grado di rendere produttiva la contribuzione degli utenti ai progetti collaborativi online.

- Rendere modulare la collaborazione, cioè fare in modo di scindere il compito generale in sotto compiti più piccoli che possono essere affrontati in maniera indipendente o semi-indipendente. Ciò facilita il coinvolgimento di nuove persone munite di nuove competenze, e di conseguenza aumenta la varietà delle competenze disponibili. Spesso la modularità è difficile da realizzarsi ed esige un impegno consapevole e indefesso da parte dei partecipanti.

Questo argomento ha a che vedere con la scelta del problema da affrontare, perché non tutte le questioni hanno la possibilità di essere risolte attribuendo ai collaboratori i sotto compiti risultanti dalla scissione dello stesso. Ma, una volta fatta la scelta, se l'argomento soddisfa questi requisiti, il guadagno in termini di ottimizzazione del tempo e delle competenze potrebbe risultare ingente. Anche Gowers sottolinea l'importanza della modularità⁶².

Persone diverse hanno caratteristiche differenti quando si tratta di ricerca. Alcuni sono adatti a proporre idee, altri a criticarle, altri a venire a capo dei dettagli, altri a rispiegare le idee in un linguaggio differente, altri a formulare problemi diversi ma correlati, altri a fare un passo indietro e passare da una situazione di stallo e confusione di idee a una situazione più chiara e coerente, e così via. Un progetto di grande collaborazione renderebbe possibile la specializzazione delle persone. Ad esempio, se si è interessati a un problema e si ha un'idea solo abbozzata, ma si è meno interessati o portati a testare questa stessa idea, allora si può solo suggerirla e sperare che gli altri la troveranno abbastanza interessanti per valutarla.

⁶² <https://gowers.wordpress.com/2009/01/27/is-massively-collaborative-mathematics-possible/>

Il secondo criterio indicato da Nielsen riguarda invece la tipologia del contributo che deve essere proposto.

- Incoraggiare i piccoli contributi, sempre per ridurre le barriere alla partecipazione e ampliare la varietà delle competenze disponibili.

I contributi non possono essere espressi in un linguaggio troppo specialistico. Questo perché i partecipanti, in quanto esperti in possesso di conoscenze teoriche non omogenee, devono avere il modo di comprendere l'idea contenuta nello stesso. In caso contrario, il processo della serendipità pianificata non potrebbe essere attivato.

I contributi non possono essere troppo lunghi. Ciò sarebbe infatti di ostacolo alla riorganizzazione dell'attenzione e all'ampliamento del bacino delle competenze. La forma lunga non si addice inoltre al mezzo Web che, come è già stato fatto notare, è adatto a veicolare contenuti ipertestuali e meno costretti dai vincoli fisici imposti dalla carta. Anche in questo caso il pensiero di Gowers è perfettamente in linea con quello di Nielsen.

[...] l'etica del forum - in qualsiasi forma essa sia presa - si manifesterà nel fatto che i commenti saranno per lo più brevi. In altre parole, ciò che non dovrebbe essere fatto, almeno se ci si vuole attenere allo spirito della collaborazione online, è passare un mese con il pensiero interamente rivolto su un problema e poi tornare indietro e scrivere dieci pagine su di esso. Piuttosto, si dovrebbe contribuire con idee, anche nel caso non siano pienamente sviluppate o addirittura sbagliate.

Il terzo principio che dovrebbe consentire di amplificare l'intelligenza collaborativa riguarda invece la tipologia di infrastruttura che dovrebbe supportare l'utilizzo dei Social Network come mezzo per produrre nuova conoscenza scientifica.

- Sviluppare un bene comune dell'informazione ricco e ben strutturato, così che si possa costruire sul lavoro precedente. Più è facile riutilizzare contenuti preesistenti, più in fretta crescerà l'offerta dell'informazione.

Senza l'istantanea disponibilità dell'informazione e l'illimitato accesso alla conoscenza garantite dal Web e dalle sue risorse, i progetti collaborativi non avrebbero quindi speranze di risultare vincenti e di attestarsi come nuova prassi condivisa che conduce alla scoperta. Ciò che Nielsen si auspica è perciò la piena realizzazione del processo di open access dei contenuti online.

Se si valutano nel loro insieme le indicazioni di Nielsen sui comportamenti e gli aspetti della contribuzione online, è possibile notare come la pratica scientifica sia particolarmente adatta all'intelligenza collettiva: la scienza è portatrice di una metodologia che ha molto in comune con i criteri appena elencati. Al di là della condizione riguardante l'estensione dei contenuti che dovrebbero essere prodotti dagli utenti, norma che sembra adattarsi maggiormente al tipo di contribuzione rappresentato dal Polymath Project, la ricerca procede infatti sia in maniera modulare, e cioè attraverso la scomposizione dei problemi in sotto problemi che possono essere affrontati in maniera indipendente, sia grazie alla strutturazione e la stratificazione successiva della conoscenza prodotta. Le esperienze di Ebola Research Database e di GenBank presentate dimostrano che le potenzialità dei Social Network sono tanto più rilevanti quanto maggiore è la capacità di accedere alle informazioni e alle conoscenze già esistenti, e conseguentemente di trovare e risolvere indipendentemente problemi ancora aperti che possano contribuire a migliorare la comprensione su questioni più generali.

Conclusioni

La nostra analisi dei servizi del Web 2.0 dedicati alla scienza e alla facilitazione della collaborazione online ci ha permesso di esplorare un'accezione alternativa del termine Scienza 2.0. Ne è emerso che la proposta qui discussa, oltre a essere supportata dai casi di studio esaminati nel primo capitolo, trova riscontro sia all'interno dell'evidenza - di matrice mertoniana - che la scienza è fondamentalmente un sistema che comunica, che nelle trasformazioni sociali e metodologiche subite dalla scienza contemporanea. Abbiamo inoltre cercato di valutare le caratteristiche che contraddistinguono la Scienza 2.0 e che rendono questo modello di produzione della conoscenza scientifica estremamente efficace e concreto. Ciò che abbiamo stabilito è che il ricorso alle risorse di collaborazione online è in grado di amplificare l'intelligenza della rete attraverso i due meccanismi della riorganizzazione delle competenze e della serendipità pianificata. Questi due aspetti emergenti dall'utilizzo dei Social Network scientifici rendono infatti possibile attribuire significato alla dispersione di contenuti e dei dati tipica sia dell'informazione veicolata da Internet che del lavoro di ricerca odierno.

Scienza post-accademica e Scienza 2.0, risposte necessarie all'evoluzione del lavoro di ricerca

Quale significato ha oggi il pensiero di Merton in merito alle norme sociali che dovrebbero contraddistinguere il lavoro degli scienziati? Risulta chiaro che il

valore delle regole CUDOS, sigla che raccoglie le quattro norme sociali del Comunitarismo, dell'Universalismo, del Disinteresse e dello Scetticismo Organizzato, sia tale solo in virtù di una concezione idealizzata della scienza, una visione che tende a escludere lo stretto rapporto esistente tra ricerca e resto della società civile: il sistema scientifico non si concretizza, e non lo ha mai fatto, al di fuori del proprio contesto storico e non è mai pienamente neutrale rispetto agli interessi e alle esigenze degli stakeholders. Tuttavia, i principi proposti da Merton continuano a incarnare una concezione limite che si fonda su prassi sociali che la scienza tende a rispettare. Una di queste riguarda certamente il fatto che la ricerca scientifica ha nella comunicazione uno dei principali fattori di progresso, aspetto questo riconducibile per l'appunto alla norma mertoniana del comunitarismo. Sotto questo punto di vista è quindi facile comprendere il motivo per cui studiosi della scienza come Ziman considerino il Web come ragione sociale della scienza contemporanea. Internet, che muove i suoi primissimi passi proprio all'interno della comunità scientifica per ragioni di condivisione della conoscenza, ha rappresentato una possibilità senza precedenti in termini di centralizzazione dei dati sperimentali, velocizzazione dello scambio dell'informazione e abbattimento delle distanze fisiche caratteristiche della collaborazione scientifica precedente all'invenzione del World Wide Web e dei Social Network e dei servizi del Web 2.0 scientifici. Ma tutto ciò non esaurisce il quadro di riferimento dentro il quale la scienza si muove oggigiorno. Uno degli aspetti che contraddistinguono la ricerca post-accademica è infatti la tendenza a rendere più diffusa possibile la rete di professionisti e di competenze che lavorano su uno stesso progetto. Questa evoluzione della ricerca verso forme sempre più collettive e transdisciplinari trova le sue ragioni all'interno delle trasformazioni subite dalla scienza stessa. Stratificazione delle conoscenze e progresso tecnologico hanno infatti determinato nel tempo una specializzazione delle competenze e un aumento esponenziale dei dati dell'osservazione sperimentale, fenomeni che hanno modificato drasticamente

il lavoro di ricerca: la quotidianità degli scienziati consiste oggi in lunghe procedure di sviluppo di algoritmi di computazione e in un'ancora più lunga attività di data mining, con l'ovvia conseguenza dell'aumento dei costi e della durata della ricerca stessa e della dispersione di preziosa informazione. Ecco il contesto in cui la Scienza 2.0, ossia un sistema scientifico che contempra la produzione di nuova conoscenza supportata dai servizi sociali del Web 2.0, non solo prende senso, ma si presenta, grazie all'analisi dei casi di studio presi in considerazione, nell'accezione radicale delineata nel primo capitolo: una realtà già ampiamente diffusa in cui i servizi online dedicati all'ottimizzazione della collaborazione fanno parte della necessaria metodologia in grado di realizzare collettivizzazione e transdisciplinarietà, soprattutto in quegli ambiti, quali le scienze naturali e quelle della vita, in cui convivono e si sovrappongono differenti discipline e oggetti di studio.

Le caratteristiche della Scienza 2.0: intelligenza collettiva, riorganizzazione dell'attenzione e serendipità pianificata

Giustificata l'ipotesi dell'utilizzo dei servizi scientifici della Rete come ragione metodologica della scienza contemporanea, nel secondo capitolo abbiamo descritto attraverso quali meccanismi essa si concretizzi nella quotidianità del lavoro dei ricercatori. La vera forza promotrice della Scienza 2.0 è la capacità di sfruttare e convogliare l'intelligenza collettiva appartenuta agli utenti della Rete al fine di utilizzarla per la risoluzione dei problemi o allo sviluppo di applicazioni innovative.

Il primo progetto del Polymath ha riscosso un discreto successo e il quinto ha consentito di ridurre la discrepanza di Erdős a un caso particolare del problema - risolto però solo grazie agli sforzi individuali di Terence Tao -, ma la maggior parte delle collaborazioni è stata presto abbandonata. Tuttavia, nel complesso, le riflessioni di Gowers e Nielsen in merito ai vantaggi derivanti dalla collaborazione estesa supportata dalle risorse online in ambito scientifico rimangono valide.

Ciò che caratterizza l'expertise scientifico della scienza contemporanea è la sua suddivisione in ambiti di competenza sempre più specifici e specialistici, perfino all'interno di uno stesso campo della ricerca. A questo fenomeno, risultato della compartimentazione del lavoro descritta nel capitolo precedente, ha fatto seguito una dispersione dell'informazione, risultante dalla natura complessa e inclusiva del nuovo paradigma del sapere, il Web. In un simile scenario, le tappe che contraddistinguono il processo della scoperta o della risoluzione di un problema scientifico complesso devono perciò essere integrate attraverso un sistema in grado di riorganizzare le microcompetenze diffuse e di concentrare l'attenzione degli esperti su uno specifico argomento, e i Social Network hanno le caratteristiche adatte a adempiere questa finalità. Aumentando il numero delle persone impegnate a dirimere una difficoltà teorica, aumentano anche le probabilità di trovare una soluzione. Ecco allora affiorare la seconda proprietà delle collaborazioni sviluppate online, e cioè quella che abbiamo definito serendipità pianificata: l'ampia interconnessione garantita dai Social Network consente che la giusta competenza, altrimenti latente all'interno del mare magnum del frammentato sapere odierno, possa incontrare il giusto problema, incrementando le possibilità di trovare la soluzione migliore. All'aumento del numero degli individui che partecipano alla collaborazione diminuiranno inoltre i tempi necessari per trovare la migliore strategia risolutiva.

Una delle problematiche riguardanti la realizzazione delle collaborazioni online individuate da Gowers concerne la difficoltà di pensare a un sistema di controllo in grado di regolare contributi e comportamento dei singoli partecipanti ai progetti. Secondo il matematico inglese, e come dimostra l'esperienza del Polymath Project, è la stessa comunità dei contributori nei progetti collaborativi a svolgere una simile valutazione. Anche se il Polymath è stato pensato e realizzato al fine di accogliere le idee di tutti, e non solo dei professionisti della scienza, i partecipanti con le maggiori competenze utilizzeranno infatti le loro conoscenze per giudicare i contenuti condivisi dagli altri utenti, attivando perciò un virtuoso meccanismo di autocontrollo.

Il secondo ostacolo a una piena attuazione della strategia di collaborazione sociale sviluppata attraverso i Social Network digitali riguarda invece l'attribuzione del merito della scoperta. Ed è proprio qui che riscontriamo le criticità che hanno impedito al Polymath Project di attestarsi come un modello per il problem solving di tipo scientifico. La soluzione proposta da Gowers per rendere conto del contributo di tutti i partecipanti a un progetto collaborativo è confidare sulla trasparenza garantita dagli stessi strumenti utilizzati per sviluppare la collaborazione. I Social Media danno infatti garanzia che tutti i passaggi, le idee e le intuizioni alla base di una scoperta siano consultabili da chiunque online.

La progressiva perdita di efficacia del Polymath è stata conseguenza diretta di questa metodologia di revisione paritaria della scoperta, non conforme al codificato sistema di valutazione e ricompensa della ricerca in vigore. All'interno della prassi scientifica odierna, i ricercatori non sono infatti incentivati a espandere il raggio delle collaborazioni poiché il prestigio e la carriera di uno scienziato sono commisurati alla quantità e qualità dei paper pubblicati. Perciò, se un ricercatore ha le capacità e le competenze necessarie per fare una scoperta di rilevanza scientifica, preferirà lavorare per conto proprio o con il suo ristretto team di collaboratori, anche se ciò vuol dire impiegare maggiori risorse e dilatare i tempi,

in modo da garantirsi finanziamenti e un ruolo di rilievo all'interno della comunità scientifica di appartenenza. Tuttavia, questa esperienza ha dimostrato in che modo sia possibile amplificare l'intelligenza inespressa degli utenti della rete al fine di ottenere nuova conoscenza scientifica.

Non è quindi possibile tralasciare le difficoltà derivanti dal ricorso alla Scienza 2.0 nell'accezione qui proposta, poiché un differente sistema di produzione della conoscenza deve di fatto fare i conti anche con un modello alternativo della sua valutazione. Tuttavia, come abbiamo avuto modo di osservare attraverso i casi studio e i servizi analizzati nel primo capitolo, l'utilizzo di strumenti di collaborazione scientifica del Web 2.0 - e in particolar modo delle risorse dedicate alla condivisione, al processamento dei dati e alla gestione dei contenuti - non sono in contraddizione e non presuppongono variazioni alla canonica revisione paritaria dei risultati della ricerca scientifica.

Il sistema di attribuzione dei meriti in una collaborazione creativa sviluppata online, come abbiamo appena evidenziato, è quindi la chiave di lettura della diffidenza con cui il mondo della scienza guarda a questo genere di esperienze di collaborazione estesa. Ma, al variare dei fattori problematici, l'utilizzo delle risorse online da parte del mondo scientifico mostra chiaramente le sue potenzialità. È questo il caso della piattaforma web Mathesia e delle altre esperienze discusse, prese in esame nel primo capitolo. Mathesia si ispira infatti alla pratica del crowdsourcing e mira a dare sì una risposta a difficoltà di tipo tecnico scientifico, ma che non hanno rilevanza in ambito scientifico, in quanto concernono lo sviluppo di conoscenza già acquisita. L'altra differenza riguarda il metodo della ricompensa previsto dal crowdsourcing: le open call, che nel caso di Mathesia sono proposte da aziende, hanno un valore economico prestabilito. Per quanto riguarda invece le similitudini, le piattaforme di crowdsourcing consentono la comparsa delle stesse proprietà che favoriscono l'amplificazione dell'intelligenza della rete, e cioè la riorganizzazione dell'attenzione e la serendipità pianificata.

Mathesia raccoglie e aggrega una vasta rete di competenze, e nello specifico di competenze matematiche, e le indirizza verso i problemi presentati. La chiarezza dei contributi auspicata da Glowers - criterio che il matematico inglese pone al fine di diminuire le distanze tra i differenti linguaggi relativi ai diversi campi di conoscenza - nel definire i principi delle collaborazioni viene inoltre simulato in Mathesia da uno specifico wizard in grado di tradurre le questioni dei richiedenti in linguaggio matematico. Grazie a questa piattaforma, questioni complesse di ottimizzazione e gestione dei processi produttivi trovano perciò una rapida soluzione, con beneficio reciproco per le aziende e per i ricercatori che rispondono alla call. Al di là degli evidenti risultati ottenuti da Mathesia, ciò che emerge dall'analisi di questo strumento online è che le risorse di aggregazione, favorendo l'interconnettività tra gli esperti, sono in grado di implementare entrambe le proprietà, riorganizzazione delle competenze e serendipità pianificata, che caratterizzano ciò che abbiamo definito strategia sociale per la ricerca di significato all'interno dei contenuti web e per la scoperta scientifica.

Quindi, uno degli aspetti che caratterizzano l'accezione di Scienza 2.0 presentata è che essa non presuppone un modello ad hoc per la valutazione delle conoscenze prodotta, poiché si fonda sull'assunzione secondo cui i servizi di collaborazione scientifica del Web 2.0 siano ragione metodologica della scienza contemporanea. Una simile ipotesi è giustificata dal fatto che i meccanismi di serendipità pianificata e riorganizzazione dell'attenzione che questi strumenti consentono di far emergere sono necessari per far fronte al fenomeno della dispersione delle competenze e dei dati tipici del lavoro di ricerca odierno, e si traduce perciò in un aspetto indispensabile della scienza contemporanea. È inoltre possibile dimostrare questa tesi attraverso il comportamento e i risultati indotti dall'utilizzo delle piattaforme esaminate nel primo capitolo. Oltre a Mathesia, le esperienze della Thematic Exploitation Platform dell'Agenzia Spaziale Europea, di Survey Lab e di Ebola Research Database consentono infatti di stabilire la capacità dei social

network scientifici di ottimizzare l'esplorazione di soluzioni innovative in tempi ristretti e di sviluppare applicazioni pratiche anche in ambiti che si allontanano da quello che genera i dati, mettendo in evidenza l'apporto metodologico di questi strumenti.

Infrastruttura e metodologia della scienza contemporanea: il Web 2.0 e i suoi servizi

Come abbiamo avuto modo di constatare grazie alla descrizione dei risultati e della tipologia di dati sperimentali che contraddistinguono la genetica contemporanea e la scienza tutta, senza il supporto di particolari risorse online e la condivisione delle informazioni, la biologia molecolare troverebbe numerose difficoltà nel produrre nuova conoscenza. Abbiamo perciò determinato che archivi digitali aperti e database open access non adempiono unicamente alla funzione di depositi del sapere, come accade nel caso dell'aggregazione fisica della conoscenza realizzata da un biblioteca, ma sono in grado, allo stesso modo dei Social Media, di implementare la strategia sociale per la ricerca di significato all'interno dell'immensa mole di dati supportata da Internet. Una simile capacità non potrebbe però manifestarsi se questi strumenti non fossero immersi in un ambiente ipertestuale che renda possibile la connessione dei dati e dell'informazione e l'immediata disponibilità di questi ultimi. Ciò rende l'open access e le risorse di collaborazione del Web 2.0 indispensabili per l'attuazione di quella nuova e necessaria metodologia scientifica che costituisce la Scienza 2.0 e rende inoltre conto di una delle tre regole a cui dovrebbero sottostare le collaborazioni online in

grado di amplificare l'intelligenza collettiva descritte da Nielsen, e cioè lo sviluppo di un patrimonio comune e strutturato dell'informazione.

Ma per quale ragione gli scienziati dovrebbero affidare una parte così importante del loro lavoro al Web, e soprattutto, perché dovrebbero pensare di propendere verso un sistema collaborativo della scoperta? Le motivazioni risiedono ancora una volta all'interno dell'esame fatto sul tipo di dati con cui la genetica contemporanea si confronta giornalmente. La natura dei dati che riguardano il nostro patrimonio genetico è infatti talmente articolata ed estesa da richiedere uno sforzo condiviso da parte di tutti i biologi. Privando questo campo della ricerca di risorse come GenBank, e cioè di grandi archivi pubblici di sequenze genetiche liberamente consultabili, la capacità di correlare i geni a gravi malattie o la possibilità di trovare delle cure tempestive per queste ultime potrebbe venire radicalmente compromessa dalle lungaggini con cui ogni singolo ricercatore dovrebbe fare i conti nel provvedere singolarmente alla produzione e all'interpretazione dei dati.

Tra i comportamenti che Nielsen individua per la contribuzione dei social network digitali, e quindi per l'implementazione delle collaborazioni estese, c'è la predisposizione di questi strumenti a gestire la suddivisione modulare del lavoro. La scienza sembra particolarmente adatta a far emergere l'intelligenza collettiva poiché la metodologia con cui essa affronta problemi complessi è infatti suddividere le questioni rilevanti in sottoproblemi differenti, a cui corrisponderà poi la rispettiva suddivisione dei compiti all'interno dei team di ricerca. L'analisi dell'esperienza dell'Ebola Research Database mostra chiaramente che gli strumenti online consentono di ottimizzare la coordinazione del lavoro degli scienziati accentrando la loro attenzione su quegli aspetti di un problema che necessitano di maggiore attenzione e accelerando il processo che porta a superare le difficoltà che si presentano. Nello specifico, il fattore temporale ha giocato un ruolo centrale nella realizzazione di misure volte a contrastare la veloce propagazione di Ebola in Africa.

In conclusione, se la Rete, a causa della crescita esponenziale dei dati della ricerca, della stratificazione e specializzazione delle conoscenze, è ormai diventata il paradigma e l'infrastruttura di riferimento per il sapere scientifico, le piattaforme e i servizi del Web 2.0 rappresentano gli strumenti indispensabili per riuscire a estrarre significato dall'osservazione sperimentale, favorire la collaborazione transdisciplinare della scienza contemporanea e sviluppare ipotesi e applicazioni innovative. Quindi la Scienza 2.0 non rappresenta un'ipotesi sulle possibili trasformazioni che il sistema di produzione della conoscenza scientifica potrebbe subire a causa dell'utilizzo dei Social Network o un semplice insieme di risorse in grado di rendere solo un contributo marginale all'impresa scientifica, ma è già una realtà concreta e una parte imprescindibile della prassi odierna dei ricercatori.

Note

Citizen Science: Sistematica raccolta e analisi di dati; sviluppo di tecnologia; verifica di fenomeni naturali; e la distribuzione di queste attività da parte dei ricercatori su una base costituita principalmente da non professionisti. [pag. 19]

Ebola: Il 9 giugno 2016 l'Organizzazione Mondiale della Sanità (WHO) ha dichiarato conclusa l'epidemia di Ebola anche in Guinea e in Liberia; <http://www.who.int/mediacentre/news/releases/2016/ebola-liberia/en/>
[pag. 34-35-36]

HapMap: A causa di problemi di sicurezza informatica, di una diminuzione nel suo utilizzo e del cambiamento dei protocolli di ricerca, il 16 giugno 2016 la risorsa è stata ritirata e sostituita dal progetto 1000 Genomes Project. [pag. 24]

Ilaria Capua: Il 5 luglio 2016 il tribunale di Verona ha proscioltto Ilaria Capua dalle accuse di associazione a delinquere finalizzata alla diffusione di epidemie mosse a suo carico nel 2006 e seguito della diffusione del genoma del virus H5N1.
[pag. 40]

Jean Piaget: (Neuchâtel, 9 agosto 1896 - Ginevra, 16 settembre 1980). Psicologo, biologo, pedagogista e filosofo svizzero. Fu il primo a introdurre il termine transdisciplinarietà durante una sua conferenza a un seminario internazionale in Francia nel 1970. Il termine, nell'uso che ne fa Piaget, fa riferimento al concetto di ecologia umana, o meglio la 'oikologia' umana (letteralmente ed etimologicamente il discorso sullo 'abitare' nel senso più ampio della parola), la quale procede con un approccio trasversale alle discipline e collaborativo nel suo tentativo di analisi e di comprensione della complessità e di soluzione ai problemi globali del nostro tempo. [pag. 9]

Mashup: Pratica di programmazione in cui si ottiene una nuova piattaforma mescolando tra loro diversi servizi Web. [pag. 16]

Mathesia: Le informazioni prive di riferimento contenute nel capitolo relativo a Mathesia sono state ottenute per mezzo di un'intervista a Ottavio Crivaro, amministratore delegato di Moxoff, e per gentile concessione di Francesca Gatti, responsabile comunicazione di Moxoff. [pag. 43-43]

Serendipità: Termine coniato da Orazio Walpole nel 1754 circa e ispirato alle avventure dei tre principi di Serendippo. La Serendipity indica la capacità di fare scoperte fortunate per caso. Il termine fu utilizzato con successo, nella medesima accezione, con anche da Merton nel tentativo di descrivere i fattori che contraddistinguono i processi della scoperta scientifica. [pag. 56]

Sistema Erdős Rado Delta (o problema del girasole): Questa congettura assume che un girasole di dimensione r è una faglia di insiemi A_1, A_2, \dots, A_r tali che ogni elemento che appartiene a più di uno degli insiemi appartiene a tutti gli altri. [pag. 23]

Bibliografia

- Barry Wellman, 'Physical Place and Cyber Place: The Rise of Networked Individualism'. *International Journal of Urban and Regional Research*. (2011)
- 'Rete, paradigma della conoscenza'. *La Repubblica*. (9 marzo 1998)
- Tony Hey, Stewart Tansley, Kristin Tolle, 'The Fourth Paradigm: Data-intensive Scientific Discovery'. (2009)
- Luca De Biase, 'Il disaccordo che accresce la conoscenza', *Il Sole 24 Ore*. (4 novembre 2012)
- Nicholas G. Carr, 'Internet ci rende stupidi?', Raffaello Cortina Editore. (2011)
- Michael Nielsen, 'La stanza intelligente. La conoscenza come proprietà della rete', Codice Edizioni. (2011)
- Tony Hey, Stewart Tansley, Kristin Tolle 'The Fourth Paradigm: Data-intensive Scientific Discovery'. (2009)
- Pietro Greco, 'Il quarto paradigma', *Micron* n. 23. (dicembre 2012)
- Michael Nielsen, 'Le nuove vie della scoperta scientifica, come l'intelligenza collettiva sta cambiando la scienza', Einaudi. (Torino, 2012)
- Andreas M. Kaplan, Michael Haenlein, 'Users of the world, unite! The challenges and opportunities of social media', *Business Horizons* n. 53. (2010)
- Davide Bennato, 'Comunità connesse. L'impatto dei social media sulla professione del ricercatore' in 'Scienza Connessa. Rete, media e Social Network', a cura di Sveva Avveduto, Gangemi Editore. (2015)

- Fabio Marzocca, 'Il nuovo approccio scientifico verso la transdisciplinarietà', Edizioni Mythos. (ottobre 2014) http://cirt-transdisciplinarity.org/biblio/biblio_pdf/eBook_Transdisciplinarity.pdf
- Massimiano Bucchi, 'Sociologia della scienza', Nuova Informazione Bibliografia n. 3. (2004)
- Maness J., 'Library 2.0 Theory: Web 2.0 and its implications for libraries', Webology vol.3, n.2. (2006)
- <https://gowers.wordpress.com/2009/03/10/problem-solved-probably/>
- <https://gowers.wordpress.com/2010/01/09/erds-discrepancy-problem-continued/>
- <https://gowers.wordpress.com/2009/01/27/is-massively-collaborative-mathematics-possible/>
- [http://www.treccani.it/enciclopedia/genoma-il-progetto-genoma-umano_\(Enciclopedia-della-Scienza-e-della-Tecnica\)/](http://www.treccani.it/enciclopedia/genoma-il-progetto-genoma-umano_(Enciclopedia-della-Scienza-e-della-Tecnica)/)
- https://it.wikipedia.org/wiki/Epidemia_di_febbre_emorragica_di_Ebola_in_Africa_Occidentale_del_2014
- http://www.esteri.it/mae/it/politica_estera/economia/cooperaz_economia/agenziaspazialeeuropea.html
- http://www.repubblica.it/salute/ricerca/2015/07/31/news/ebola_oms_entusiasmati_i_primi_risultati_del_vaccino_-120185068/
- http://www.adnkronos.com/salute/sanita/2015/01/02/dall-africa-agli-usa-storia-dei-contagi-del-virus-maledetto_CSRkgQcq27tQJXRSScpeoM.html
- <http://www.who.int/mediacentre/news/statements/2015/liberia-ends-ebola/en/>
- <https://it.wikipedia.org/wiki/Crowdsourcing>

- <http://aecanalysisicalcolo.it/pubblicazioni/aec/70/crowdsourcing-e-partecipazione-collettiva/>
- <https://arxiv.org/abs/1509.05363>
- <http://maddmaths.simai.eu/divulgazione/focus/terence-tao-erdos/>
- http://it.masternewmedia.org/2010/03/17/scrittura_collaborativa_online_in_che_modo_i.htm
- Axel Bruns, 'Bloggare da dentro la torre d'avorio'. (2006)
- <http://michaelnielsen.org/blog/doing-science-online/>
- Pietro Greco, 'L'idea pericolosa di Galileo: storia della comunicazione della scienza nel Seicento', Utet Università. (2009)
- Karim R. Lakhani, Lars Bo Jeppesen, Peter A. Lohse, Jill A. Panetta. 'The Value of Openness in Scientific Problem Solving'. (2006)
- Lucio Russo, 'La rivoluzione dimenticata', Feltrinelli. (1998)
- James Surowiecki, 'The Wisdom of Crowds', Anchor Book. (2004)
- Anabel Quan-Haase and Kim Martin, 'Seeking knowledge: An exploratory study of the role of social networks in the adoption of Ebooks by historians'. (2011)
- Karim R. Lakhani, Lars Bo Jeppesen, Peter A. Lohse e Jill A. Panetta. 'The Value of Openness in Scientific Problem Solving'. (2006)
- David Weinberger, 'La stanza intelligente, la conoscenza come proprietà della rete', Codice Edizioni. (Torino, 2012)
- Merton Robert, K. Elinor, G. Barber, 'Viaggi e avventure della Serendipity', Editore Il Mulino. (Bologna, 2002)
- John Ziman, 'La vera scienza, natura e modelli operativi della prassi scientifica', Edizioni Dedalo. (Bari, 2002)
- Piergiorgio Odifreddi, 'Il Tao della matematica', Le Scienze. (dicembre 2015)

Sitografia

- <http://www.treccani.it/>
- <https://www.wikipedia.org/>
- <http://polymathprojects.org/>
- <http://www.mathesia.com/>
- <https://gowers.wordpress.com/>
- <https://terrytao.wordpress.com/>
- <http://michaelnielsen.org/blog/>
- <http://www.ncbi.nlm.nih.gov/genbank/>
- <https://www.researchgate.net/>
- <https://www.lablife.org>
- <http://www.dnatube.com/>
- <https://www.mendeley.com/>
- <https://www.cnaf.infn.it/indigo/>
- <http://www.ukcds.org.uk/resources/ebola-research-database>
- <https://genome.ucsc.edu/ENCODE/>
- <http://www.ncbi.nlm.nih.gov/genbank/>
- <http://www.ncbi.nlm.nih.gov/pubmed>
- <http://copernicus.eu/>

Elenco Figure

Fig. 1	29
Fig. 2	30
Fig. 3	31
Fig. 4	34
Fig. 5	36
Fig. 6	36
Fig. 7	38
Fig. 8	42
Fig. 9	50
Fig. 10	55
Fig. 11	58
Fig. 12	59

