



**DOTTORATO IN COMUNICAZIONE, RICERCA, INNOVAZIONE - XXX CICLO**

*Curriculum in Metodologia delle Scienze Sociali*

Coordinatore: Prof. Enzo Campelli

**Il cambiamento nel modo di produzione della conoscenza.  
Un'esplorazione degli effetti delle infrastrutture di ricerca  
nel lavoro degli scienziati sociali**

Tesi di  
Andrea Orazio Spinello

Tutor  
Prof. Maurizio Bonolis  
Dott.ssa Emanuela Reale



## *Ringraziamenti*

Desidero ringraziare vivamente i miei tutor, il *prof. Maurizio Bonolis* e la *dott.ssa Emanuela Reale*, per la dedizione con cui mi hanno seguito durante il percorso dottorale e per la stima che hanno da subito riposto nei miei confronti; la *dott.ssa Emilia Primeri* per i preziosi consigli e per la pazienza nel supportarmi; la *dott.ssa Serena Fabrizio* e il *dott. Antonio Zinilli*, colleghi presso l'IRCrES-CNR, che hanno incoraggiato il mio lavoro di ricerca dal primo giorno; i dirigenti di ricerca, professori, ricercatori e assegnisti che si sono resi disponibili a rilasciare le interviste che hanno costituito il materiale empirico utilizzato per questo studio.

E poi, un grazie di cuore all'amico *Salvatore Mertoli* per avermi sostenuto attivamente durante il lavoro di ricerca; a *Rafał Matuszewski*, che mi ha trasmesso entusiasmo e fiducia ad ogni fase del mio progetto; a *Marco Condorelli*, grazie al cui incitamento è in parte dovuta la mia decisione di intraprendere il percorso dottorale. Un ringraziamento speciale va anche ad alcuni amici che, seppur non vicini fisicamente, mai mi hanno fatto mancare supporto, stima e affetto, e che sono e saranno corresponsabili, in piccola parte e in modo diverso, del raggiungimento di ogni mio obiettivo: *Simona Coco, Santi Marcello Russo, Andrea Sessa, Paolo Tirrò*.

Infine, un grazie infinito ai miei genitori, *Giuseppe e Maria*, che mi hanno sempre trasmesso e mi continuano a trasmettere il loro amore, compreso quello per lo studio. A loro è dedicato questo lavoro.

# INDICE

<b>INTRODUZIONE</b>	<b>6</b>
<b>CAPITOLO 1</b>	
<b>LA PRODUZIONE DI CONOSCENZA SCIENTIFICA: APPROCCI DI STUDIO, INNOVAZIONI E TENDENZE</b>	<b>11</b>
<b>1.1 Rassegna della letteratura principale sugli studi sulla scienza</b>	<b>11</b>
1.1.1 La riflessione sulle rivoluzioni scientifiche	13
1.1.2 Dall’approccio istituzionale alla virata sui contenuti dell’attività di ricerca	14
1.1.3 L’etnografia in laboratorio e la scienza in azione	17
1.1.4 I mutamenti della scienza contemporanea	19
<b>1.2 Gli avanzamenti delle ICT e il lavoro di ricerca</b>	<b>21</b>
1.2.1 Dal computer all’avvento del Web 2.0	22
1.2.2 Parametri strutturali e determinanti dell’adozione di ICT	27
<b>1.3 La nuova scienza: aperta, collaborativa e basata sul dato</b>	<b>30</b>
1.3.1 La reinvenzione della scoperta scientifica	32
1.3.2 I dati aperti tra promozione e barriere	35
<b>CAPITOLO 2</b>	<b>40</b>
<b>LE INFRASTRUTTURE DI RICERCA NELLE SCIENZE SOCIALI</b>	<b>40</b>
<b>2.1 Infrastruttura di ricerca: un concetto moderno, ma non nuovo</b>	<b>40</b>
2.1.1 Una costellazione di referenti empirici	41
<b>2.2 Calibrazione dei referenti per il campo delle scienze sociali</b>	<b>43</b>
<b>2.3 Le infrastrutture derivate da progetti specifici (i <i>project-driven database</i>)</b>	<b>48</b>
2.3.1 Sfide per il futuro: accessibilità, manutenzione, stabilità	49
<b>2.4 Infrastrutture di ricerca stabili e aperte per le scienze sociali</b>	<b>52</b>
2.4.1 Le infrastrutture di interesse europeo	53
2.4.2 Infrastrutture finanziate dai programmi quadro europei – il caso RISIS	56
2.4.3 Esempi di infrastrutture di caratura internazionale: OCSE-PISA e OCSE-PIAAC	57

## CAPITOLO 3

### **APPROCCIO METODOLOGICO E STRATEGIA DI RICERCA 59**

#### **3.1 Lo studio di caso e l'obiettivo esplorativo 59**

3.1.1 Domanda di ricerca e ipotesi 61

3.1.2 Generalizzazione analitica 63

#### **3.2 Il disegno della ricerca 64**

3.2.1 Scelta del contesto di riferimento e individuazione del caso 64

3.2.2 Selezione e numerosità delle sub-unità 66

#### **3.3 La tecnica di raccolta dati: l'intervista semi-strutturata 70**

3.3.1 Vantaggi e caveat della socializzazione preliminare al *setting* di ricerca 72

#### **3.4 Strategia d'analisi 74**

## CAPITOLO 4

### **ANALISI DEL MATERIALE EMPIRICO: LE INFRASTRUTTURE DI RICERCA NEL LAVORO DEGLI SCIENZIATI SOCIALI 77**

#### **4.1 Le infrastrutture di ricerca nelle scienze sociali: in cerca di una definizione 77**

#### **4.2 Tipi di infrastrutture utilizzate dagli intervistati 82**

4.2.1 Riflessi sulla produzione scientifica generale 86

#### **4.3 Gli effetti dell'utilizzo delle infrastrutture sull'attività di ricerca 88**

4.3.1 Cambiamento tecnologico e accesso alle risorse informative 88

4.3.2 Nuove impostazioni del disegno della ricerca 93

4.3.3 Collaborazioni a livello locale, nazionale o internazionale 99

4.3.4 Propensione all'interdisciplinarietà 105

#### **4.4 Valutazioni sulle ricadute dell'utilizzo di infrastrutture sui prodotti di ricerca: aspetti positivi e criticità 110**

4.4.1 Aspetti qualitativi relativi agli *output* sviluppati da infrastruttura 110

4.4.2 Criticità nell'utilizzo delle infrastrutture precostituite 115

#### **4.5 Caratteristiche desiderabili delle infrastrutture 119**

### **CONCLUSIONI 123**

### **RIFERIMENTI BIBLIOGRAFICI 133**

### **ALLEGATO: PROTOCOLLO DI INTERVISTA 144**

## INTRODUZIONE

La natura dei processi di ricerca scientifica è entrata in una fase di transizione verso una modalità di produzione della conoscenza (la *knowledge production mode 2* profetizzata da Gibbons *et al.* nel 1994) caratterizzata, tra le altre cose, (i) dall'emersione di un ampio spettro di siti nei quali la conoscenza si produce e circola stimolando le interazioni tra ricercatori dapprima limitate da impedimenti fisici e tecnici; (ii) dalla valorizzazione dell'orientamento alla transdisciplinarietà che si è contrapposto ad assetti settoriali contraddistinti da rigidità e vincoli gerarchici; (iii) dall'allontanamento da modelli di scienza "pura" e fondata sulla teoresi che ha aperto la strada a disegni di ricerca innovativi, anche in virtù dei consistenti sviluppi delle innovazioni tecnologiche trasferite al campo scientifico che hanno radicalmente cambiato il modo di trattare e analizzare i dati. Da una decina di anni, inoltre, complice l'evoluzione tecnica della rete Internet, il rapporto tra tecnologie di informazione e comunicazione e mondo scientifico ha assunto nuove peculiari connotazioni: i ricercatori utilizzano le risorse *online* non solo per consultare le pubblicazioni, ma anche per comunicare e scambiare dati e informazioni; grandi basi di dati condivise a livello nazionale e internazionale sono state sviluppate allo scopo di raccogliere e disseminare dati in un contesto di comparazione continua di analisi e risultati e di produzione comune di standardizzazioni, nomenclature e classificazioni. Parallelamente, la progressiva istituzionalizzazione del dato di ricerca come "bene comune" ha conferito rilevanza al concetto di "apertura" (*open science*) nei termini di un'accessibilità al dato orientata verso platee sempre più ampie di ricercatori e verso la società (*cf.* Boulton *et al.*, 2012). Alla luce dell'organizzazione di una pratica di ricerca sempre più basata su *team* che operano su piattaforme sociotecniche aperte e condivise, alcuni autori hanno delineato i contorni di un "nuovo percorso per la scienza" fondato sulla condivisione della ricchezza informativa, sull'intensità del dato (Hey *et al.*, 2009) e su una maggiore capacità di "fare rete" per la risoluzione delle questioni scientifiche (Nielsen, 2011, tr. it. 2012).

In questo contesto, lo sforzo collettivo delle comunità scientifiche si riversa nella produzione e nello sfruttamento condiviso di dati osservazionali e sperimentali resi disponibili tramite risorse fisiche o virtuali, quali le *infrastrutture di ricerca*, accessibili secondo termini di utilizzo variabili, la cui costituzione è spesso supportata dall'azione delle agenzie di finanziamento o dalle istanze dei decisori politici. Se un primo referente della

locuzione si può far risalire alla biblioteca alessandrina, quale primordiale abilitatore al processo di produzione della conoscenza, ad oggi essa si attaglia ad una molteplicità di oggetti – dai laboratori, agli acceleratori di particelle, ai *database* di ricerca – che supportano i processi della ricerca scientifica in tutti i campi disciplinari (*cf.* European Commission, 2008, p. 10; Ciula *et al.*, 2011). Quando ci si riferisce al dominio delle scienze sociali, l'insieme dei referenti risulta tuttavia ridotto, e circoscritto alla categoria delle *research information infrastructure*, che include le raccolte di materiale accademico, gli archivi di dati e i *database* di ricerca strutturati (German Council of Science and Humanities, 2011). In particolare, al livello delle basi di dati, prescindendo dalla produzione di statistiche ufficiali nazionali e internazionali, si distinguono a) le *collezioni di dati project-driven*, legate a specifici progetti di ricerca europei, nazionali o generati nell'ambito di progetti d'ateneo, b) le infrastrutture caratterizzate da aggiornamento costante e ampia apertura alle comunità scientifiche, come quelle di *interesse europeo*, costituite nell'ambito di quadri legali di tipo europeo e afferenti allo *European Strategy Forum on Research Infrastructures* (ESFRI; *cf.* ESFRI, 2016) o di *caratura internazionale*, quali quelle derivanti indagini periodiche promosse da organizzazioni internazionali (ad esempio i database PISA e PIAAC di OCSE).

Proprio l'irruzione sulla scena scientifica delle infrastrutture di ricerca nel campo delle scienze sociali, siano esse stabilite nell'ambito di quadri legali internazionali o sorte nell'ambito di progetti di ricerca specifici, costituisce il nucleo d'interesse del presente lavoro. Il loro sfruttamento suggerisce infatti molteplici potenziali implicazioni rispetto ai processi di ricerca, tra cui una potenziata disponibilità di accesso ai risultati e ai dati di ricerca e una più accentuata possibilità di soddisfare nuove domande di ricerca a prescindere dalla produzione di dati originali (altri possibili effetti sono presentati da Farago, 2014). Sullo sfondo della nuova scienza aperta e collaborativa, la diffusione di tali risorse appare dunque predisporre ad un cambiamento nel modo di produzione della conoscenza.

Un'indagine estensiva sui cambiamenti che intervengono nel lavoro dei ricercatori per effetto dello sfruttamento e della condivisione di infrastrutture di ricerca è stata implementata ad oggi solo in uno studio del 2010 (*cf.* Hüsing *et al.*, 2010), che tra l'altro evidenziava dei limiti di rappresentatività relativamente agli utenti delle infrastrutture proprio nel campo delle scienze sociali. In generale, l'interesse alle dinamiche di coproduzione di nuovi prodotti scientifici e si è finora sostanzialmente diretto al campo delle c.d. *scienze dure* producendo già un discreto numero di studi (*cf.* Hine, 2006; Bietz e Lee, 2009; Leonelli, 2012; Binz-

Scharf *et al.*, 2014); al contrario, il tema dell'eventuale modifica dei parametri su cui si fonda l'organizzazione del lavoro nelle *scienze non dure*, e in particolare nel campo delle scienze sociali, è stato pochissimo affrontato (*cf.* Barkuus e Brown, 2012; Antonijevic *et al.*, 2013). Il presente lavoro ambisce dunque a posizionarsi in questo panorama contraddistinto dalla carenza di approfondimenti sugli effetti dello sfruttamento delle infrastrutture sul lavoro degli scienziati sociali, in ciò applicando un approccio che si propone di far luce sui processi che strutturano le azioni e le relazioni dei ricercatori che le utilizzano, e che consideri la produzione di conoscenza come il risultato dell'attivazione di meccanismi decisionali che presiedono al processo di ricerca, come quelli che pertengono alla costituzione delle strategie di ricerca e alla scelta delle fonti di dati e indicatori, o quelli relativi alle dinamiche di tipo aggregativo, quali la formazione di strutture interattive tra scienziati.

L'obiettivo della ricerca è primariamente *esplorativo* e si è tradotto nel controllo di una serie di ipotesi operative da testare ulteriormente in studi successivi. La *domanda di ricerca* si è focalizzata sul cambiamento indotto dall'utilizzo delle infrastrutture di ricerca nei processi che presiedono alla produzione di nuova conoscenza negli studi sociologici, con particolare riferimento agli *effetti sul disegno di ricerca*, sulle *dinamiche collaborative*, sulla *tendenza all'interdisciplinarietà*, approfondendo inoltre quale sia il *giudizio sulle ricadute sui prodotti della ricerca derivati dallo sfruttamento di tali risorse*. Su questa base, nella fase di disegno della ricerca, sono state sviluppate alcune ipotesi preliminari che facessero da guida al lavoro, indirizzate al riscontro di: (i) un'influenza delle infrastrutture verso l'attivazione di processi collaborativi che possano portare anche alla formazione di nuovi legami tra ricercatori; (ii) una spinta, favorita dalla disponibilità (o creazione) di tali risorse, verso la formazione di domande di ricerca orientate alla produzione di studi interdisciplinari; (iii) un riconoscimento, da parte dei ricercatori, di una serie di ripercussioni vantaggiose che investono la progettazione e la produzione di *output* derivanti dal loro utilizzo. Al fine di esplorare tali dimensioni, è stato creato un protocollo di intervista semi-strutturata che è stato somministrato a ricercatori appartenenti al dominio degli studi sociologici di diversa età, genere e posizione accademica. Nel corso delle interviste, è stata prevista l'inclusione di una parte "narrativa" che avesse ad oggetto il processo di formazione di uno o più *output* di ricerca generati dallo sfruttamento di infrastrutture, con lo scopo di ricostruire retrospettivamente le modalità peculiari di realizzazione. Tale scelta metodologica si è ritenuta utile in particolare ai fini di far emergere quale fosse, dal punto di vista dei ricercatori, il vantaggio cognitivo e

organizzativo (e l'insieme di vincoli connessi) derivato dall'utilizzo delle infrastrutture di ricerca per la produzione di nuova conoscenza scientifica.

Nel delimitare il contesto di studio alla comunità dei sociologi, si è deciso di implementare uno *studio di caso* riferito al maggior polo universitario italiano – Sapienza Università di Roma – e al maggior ente di ricerca pubblico italiano – il Consiglio Nazionale delle Ricerche (CNR). Le interviste sono state somministrate a 15 tra dirigenti di ricerca, professori, ricercatori, assegnisti di ricerca di tali istituzioni. Sono stati inoltre intervistati tre “testimoni privilegiati”, il cui esame del curriculum ha rivelato una particolare attitudine all'utilizzo di infrastrutture di ricerca, in modo che le loro esperienze potessero costituire il *benchmark* di riferimento su pratiche avanzate. Il materiale testuale è stato analizzato attraverso l'applicazione dei metodi dell'*analisi tematica* (Braun e Clarke, 2006). L'analisi è stata implementata attraverso la produzione di una sinossi descrittiva, suddivisa per temi, basata sulle testimonianze degli intervistati, e infine completata da un'interpretazione complessiva rispetto alle dimensioni esaminate che si propone di aprire interrogativi e sollecitare al dibattito e alla riflessione sul tema.

Dal punto di vista della struttura, il lavoro si articola in quattro capitoli e una sezione finale dedicata alle conclusioni.

- Il *primo capitolo* offrirà una panoramica degli studi più rilevanti nell'ambito dei *Science and Technology Studies* (STS) che hanno inteso esplorare le dinamiche di produzione della conoscenza scientifica, dalla teorizzazione classica di T. S. Kuhn (1962, tr. it. 1969), agli “studi di laboratorio” (tra gli altri, Latour e Woolgar, 1979; Knorr-Cetina, 1982), fino ai contributi più recenti di Gibbons *et al.* (1994) e Ziman (1998). Verranno poi presentati gli effetti degli avanzamenti delle tecnologie di informazione e comunicazione nell'ambito della pratica scientifica, con una particolare attenzione nei riguardi dei processi di adozione e di recepimento di innovazioni ed evoluzioni tecniche nei contesti di ricerca. Infine, verranno chiariti alcuni concetti emergenti negli ultimi anni riguardo il cambiamento nelle modalità di produzione e circolazione della conoscenza scientifica, quali *open science* (*cfr.* Boulton *et al.*, 2012), *networking science* (Nielsen, 2011, tr. it. 2012), “quarto paradigma che sostiene la scoperta scientifica” (Hey *et al.*, 2009).

- L'indagine non ha potuto naturalmente prescindere dalla delimitazione dell'ambito di ricerca, così il *secondo capitolo* sarà caratterizzato dal tentativo di delimitare i confini semantici della locuzione *infrastruttura di ricerca*. Una particolare attenzione verrà dedicata alla calibrazione dei referenti del concetto quando riferito alle *scienze sociali*. Sarà introdotta la distinzione tra infrastrutture derivanti da progetto specifico (*project-driven*) e quelle di tipo "istituzionalizzato", stabili e aperte. Riguardo questo ultimo tipo, un breve approfondimento sarà dedicato alle infrastrutture per le scienze sociali di interesse europeo sviluppate sotto l'ombrello dello *European Strategy Forum on Research Infrastructures* (ESFRI), a un esempio di infrastruttura di tipo distribuito finanziata attraverso il Settimo Programma Quadro europeo (EU-FP7) e a due infrastrutture di rilevanza internazionale promosse dall'Organizzazione per la cooperazione e lo sviluppo economico (OCSE).
- L'analisi della letteratura e la disamina della documentazione sulle infrastrutture di ricerca presentate nel primo e nel secondo capitolo sono state trattate come *background* teorico che ha permesso di identificare e raffinare le dimensioni rilevanti d'indagine da inserire nel disegno della ricerca. Il *terzo capitolo* tratterà in maniera accurata l'approccio metodologico utilizzato per questo studio, esplicitandone l'obiettivo esplorativo e l'istanza di generalizzazione analitica (*cf.* Yin, 2009). Saranno inoltre presentate compiutamente la domanda di ricerca, le ipotesi operative, la tecnica utilizzata per la raccolta dei dati e la strategia d'analisi.
- Il *quarto capitolo* descriverà i risultati dell'analisi del materiale empirico, esplorando come i ricercatori sociali declinano il concetto di infrastruttura di ricerca, rilevando gli effetti dell'utilizzo di tali risorse in termini di nuove impostazioni dei disegni di ricerca, collaborazioni e orientamento all'interdisciplinarietà, ed esaminando la loro prospettiva valutativa sui vantaggi (e sulle eventuali problematiche) derivati dall'utilizzo delle infrastrutture per la produzione di nuova conoscenza scientifica.
- L'ultima parte del lavoro, dedicata alle *conclusioni*, riguarderà la discussione dei risultati alla luce della domanda di ricerca e delle ipotesi, e mirerà a rappresentare quanto emerso dallo studio in modo che possa prestarsi a ulteriori controlli attraverso future repliche della ricerca (ad esempio, nuovi studi di caso), suggerendo inoltre approfondimenti rispetto alle dimensioni emerse ed analizzate.

## CAPITOLO 1

### **La produzione di conoscenza scientifica: approcci di studio, innovazioni e tendenze**

#### **1.1 Rassegna della letteratura principale sugli studi sulla scienza**

Le modalità di produzione della conoscenza scientifica costituiscono un oggetto d'indagine peculiare all'interno del filone dei *Science Studies*. Entro tale termine è condensato un insieme di pubblicazioni e ricerche di diversa matrice – *in primis* sociologica, ma anche antropologica, filosofica, storica – accomunate dal comune orientamento verso tematiche inerenti la scienza e la tecnologia: per questo motivo, il campo disciplinare è definito, in modo alternativo, come *Science and Technology Studies* (STS). Al suo interno afferiscono numerose scuole di pensiero tra cui le più note sono rappresentate dal movimento accademico denominato *Sociology of Scientific Knowledge* (SSK) e dalla corrente dei sostenitori dell'*Action-Network Theory* (ANT) (cfr. Daston, 2009).

Sebbene le prime riflessioni d'impronta sociologica sulla scienza abbiano iniziato a svilupparsi a partire dalla fine degli anni Trenta con Robert K. Merton<sup>1</sup>, fu solo dopo la Seconda Guerra Mondiale, e in particolare a partire dagli anni Sessanta, che si costituì un campo scientifico identificabile di studi sulla scienza, per effetto dell'azione catalizzatrice e della risonanza che suscitò l'opera di Thomas S. Kuhn sulla struttura delle rivoluzioni scientifiche (cfr. Kuhn, 1962, tr. it. 1969). L'interesse verso le nuove tematiche fu inoltre alimentato dall'esperienza del recente conflitto bellico che contribuì da un lato a dare risalto al ruolo rilevante svolto da alcuni scienziati e gruppi di ricerca e dall'altro a fortificare la convinzione che l'azione scientifica e lo sviluppo della tecnologia potessero essere considerati elementi determinanti per il funzionamento sia del sistema politico che di quello economico (Rudolph, 2014). Da quel momento, la letteratura di studi sulla scienza si è popolata di contributi che hanno utilizzato approcci diversificati, sulla scorta della considerazione che l'insieme dei processi e meccanismi che muovono la scienza insistono su più livelli analitici. Come riportato da Heimericks e Leydesdorff (2012), a questo proposito

---

<sup>1</sup> Il lavoro che è generalmente considerato l'atto fondativo degli studi sociologici sulla Scienza, è costituito dalla tesi di dottorato di Robert K. Merton su *Storia, tecnologia e società nell'Inghilterra del secolo XVII* del 1938 (cfr. Bucchi, 2004).

si è soliti distinguere tra: (i) studi classici di macro-livello sulla produzione della conoscenza, che si concentrano sugli sviluppi paradigmatici delle discipline, (ii) studi di micro-livello che coprono il contesto della scoperta e si fondano sul livello del laboratorio e sull'esperienza di ricerca concreta, (iii) studi più recenti che pongono l'attenzione sul contesto di applicazione dei risultati scientifici e riconoscono l'importanza degli ambienti socioeconomici in cui i processi di produzione sono immersi.

Inquadrabile appieno nell'ambito della prima tipologia è il lavoro sopracitato di Kuhn, attraverso il quale l'autore ha proceduto alla ricerca della proprietà astratte dei *pattern* dinamici che segnano l'evoluzione dei campi disciplinari, adottando un'impostazione riconducibile alla Storia della Scienza rivisitata sotto un profilo sociologico. La seconda declinazione si sviluppò dagli anni Settanta in reazione a quello che fino ad allora costituiva l'approccio classico agli studi sulla scienza, ossia l'orientamento struttural-funzionalista della *Sociology of Science* mertoniana. Se in Merton l'oggetto privilegiato d'indagine è coinciso col versante istituzionale dei processi di produzione della conoscenza, in ciò focalizzandosi sugli aspetti organizzativi e funzionali legati al sottosistema che regola la produzione e la trasmissione del sapere nel più ampio sistema sociale, soprattutto con Bloor (*cfr.* Bloor, 1981) emersero parallelamente sia un atteggiamento di critica nei confronti di tale visione che una svolta verso un approccio che avesse ad oggetto i contenuti concreti dell'attività scientifica evidenziando un interesse alla comprensione di fattori, come le ideologie politiche o gli elementi culturali, che contribuiscono a spiegare le direzioni in cui la scienza si muove. Partendo dall'assunto che la conoscenza scientifica sia il risultato di meccanismi innescati al contempo da variabili esogene, presenti a livello culturale e sociale, ed endogene, riscontrabili, nelle pratiche quotidiane dei ricercatori e spesso derivanti da specifiche culture di laboratorio, gli *Science Studies* si sono dunque sempre più concentrati per tutti gli anni Ottanta e l'inizio dei Novanta sul microlivello dei processi di ricerca. Particolare enfasi è stata dedicata ai processi sociali che stimolano la crescita della conoscenza e la strutturano attraverso le azioni degli scienziati e le relazioni nell'ambito delle comunità scientifiche (*cfr.* Latour e Woolgar, 1979; Knorr-Cetina, 1982) e alla formazione di *network* di ricerca come segno della formazione di alleanze dinamiche nel contesto scientifico (*cfr.* Callon, 1991). Più di recente, sono stati prodotti contributi a cavallo tra la Sociologia della Scienza e la teoria neo-istituzionalista che hanno spostato l'interesse verso le dinamiche del rapporto tra scienza e società. Tali lavori hanno delineato le peculiarità dei mutamenti delle forme di produzione

della scienza contemporanea in relazione alle caratteristiche degli ambienti socioculturali che la influenzano (Gibbons *et al.*, 1994; Nowotny *et al.*, 2001; Ziman, 2000, tr. it. 2002).

### *1.1.1 La riflessione sulle rivoluzioni scientifiche*

La teorizzazione sulle “rivoluzioni scientifiche”<sup>2</sup> di Thomas S. Kuhn (1962, tr. it. 1969) ha rappresentato una riflessione di ampio respiro sulla dinamica evolutiva dei campi disciplinari. Lo studioso statunitense riconfigurò il modo di pensare all’evoluzione delle discipline scientifiche, non più identificandolo col procedere lineare e cumulativo che gli era generalmente attribuito, ma caratterizzandolo attraverso l’alternanza tra fasi di “scienza normale” e momenti rivoluzionari segnati da slittamenti paradigmatici. Questi ultimi si impongono come momenti traumatici di discontinuità col passato sotto molteplici aspetti, soprattutto in termini di visione teorica generale sottesa alle speculazioni ed elaborazioni specifiche delle discipline scientifiche.

Il concetto di paradigma in Kuhn ha un’estensione molto ampia e incorpora la visione pre-teorica, la prassi di ricerca e gli strumenti di studio condivisi da una comunità scientifica e si traduce, nella concreta attività scientifica, nell’individuazione e scelta dei fatti rilevanti da studiare, nella formulazione di ipotesi entro cui collocare la spiegazione dei fenomeni e nella modulazione e nella scelta delle tecniche di ricerca empirica appropriate (*cf.* Corbetta, 2003). Il paradigma costituisce il modello sulla base del quale gli scienziati legittimano il loro modo di fare scienza e, ruotando la prospettiva, si riferisce parimenti alle regole che guidano le *routine* del lavoro scientifico in tutte le fasi del disegno di ricerca. In periodi di “scienza normale”, il grado di consenso verso il complesso di teorie, metodi e criteri è massimo e questa situazione fa sì che essi siano esenti da qualsivoglia forma di verifica e rappresentino le coordinate guida dell’attività sperimentale. Di conseguenza, durante queste fasi, lo sviluppo scientifico procede effettivamente in maniera lineare e cumulativa. Allorquando la comunità riscontra delle anomalie che rimettono in discussione tecniche e credenze consolidate, avviene una “trasformazione della struttura concettuale attraverso la quale gli scienziati guardano al mondo” (Kuhn, 1962, tr. it. 1969, p. 131) e si concretizza “un cambiamento dei problemi da proporre all’indagine scientifica e dei criteri con i quali la professione stabiliva cosa si sarebbe dovuto considerare come un problema ammissibile o

---

<sup>2</sup> Espressione declinata al plurale per poterla distinguere dalla "rivoluzione scientifica" sviluppatasi tra la fine del Rinascimento e il Seicento.

come una soluzione legittima ad esso” (*ibidem*, p. 25). Si apre così un periodo di “scienza straordinaria” nel quale la comunità si interroga su quali dei nuovi paradigmi accettare, finché un numero sufficiente di scienziati converrà su un quadro di riferimento convincente che risulterà prevalente.

Secondo Kuhn, alcune discipline, come le scienze sociali, non facendo riferimento ad un’unica struttura concettuale condivisa, si troverebbero ad uno stato pre-paradigmatico, caratterizzato cioè dall’esistenza di molte scuole differenti in competizione tra di loro e dall’assenza dell’immagine fondamentale che una disciplina ha del suo oggetto (*ibidem*). Al contrario, discipline afferenti alle scienze naturali hanno conosciuto slittamenti paradigmatici che hanno modificato la concezione pre-teorica della disciplina e ne hanno fondato una nuova (ad esempio, l’irruzione della visione quantistica nella fisica ottica, che ha sostituito le concezioni newtoniane e le teorie basate sul comportamento ondulatorio delle particelle)<sup>3</sup>.

### 1.1.2 Dall’approccio istituzionale alla virata sui contenuti dell’attività di ricerca

I lavori di sociologia della scienza di Robert K. Merton, già a partire dalla sua tesi di dottorato del 1938, presentavano un’impostazione lontana dalla prospettiva evolutiva che contraddistinse il lavoro di Kuhn. Il sociologo statunitense ha sviluppato un approccio prettamente istituzionale che si è tradotto in una concezione della scienza in quanto entità sorretta da un *ethos* che implicasse un rifiuto costante da parte degli scienziati all’applicazione di norme utilitaristiche nel loro lavoro. Tale *ethos* si traduceva nei quattro imperativi istituzionali, caratterizzanti la struttura normativa della scienza. Essi erano rappresentati dall’*universalismo* (la validità di una teoria scientifica è giudicata indipendentemente da chi ne sia l’autore), dal *comunitarismo* (ogni scoperta diviene parte di un corpus scientifico che diventa bene comune per la comunità di ricerca), nel *disinteresse* (la finalità ultima dello scienziato è il progresso scientifico) e nello *scetticismo organizzato* (è necessaria la valutazione critica di qualsiasi risultato fino all’ottenimento di prove sulla sua

---

<sup>3</sup> Anticipatorie di molte delle tesi di Kuhn, furono le argomentazioni di Ludwik Fleck (1935, tr.it. 1983). In particolare, il concetto di paradigma è molto vicino a quello di *denkstil* (stile di pensiero), ossia “ciò che disciplina il pensiero scientifico, lo orienta, mette a sua disposizione risorse cognitive e nello stesso tempo pone limiti alla sua attività cognitiva e alle sue pratiche” (*ibidem*). Esso influenza qualsiasi atto di osservazione e diventa nucleo dei cosiddetti collettivi di pensiero, cioè la “comunità degli uomini che hanno fra di loro un contatto intellettuale e che si scambiano le idee influenzandosi reciprocamente” (*ibidem*). Quella scientifica è solo una delle molteplici cerchie a cui uno scienziato può partecipare (si pensi alla classe sociale o alla fede religiosa) e l’interconnessione e la combinazione di tali sfere d’appartenenza determina un condizionamento sociale dell’attività di conoscenza.

validità). Questi principi si prestavano ad essere interpretati come guide che dovrebbero muovere le attività in seno al sottosistema scientifico, garantendone il funzionamento (*cf.* Bucchi, 2004).

Nei lavori del sociologo americano è emersa dunque una visione peculiare della scienza come oggetto d'analisi: essa infatti era connotata in modo duplice, sia come entità autonoma, dotata di un insieme di metodi e attività, che come sottosistema valoriale e normativo all'interno del sistema societale (*cf. ibidem*)<sup>4</sup>. Al centro della riflessione si ponevano dunque le caratteristiche esteriori del sottosistema scientifico, del funzionamento delle istituzioni scientifiche e delle loro proprietà e l'attenzione agli aspetti organizzativi e strutturali costituivano ciò che per Merton era il focus della disciplina e la pietra di paragone delle speculazioni. Nella visione struttural-funzionalistica mertoniana fu dato particolare risalto da un lato alla conflittualità dei valori (ad esempio, l'ideale di accesso comune all'informazione è spesso tradito dalla tendenza alla riservatezza), dall'altro alle deviazioni dalle prescrizioni normative. A quest'ultimo proposito, Merton rilevò un fenomeno, il cosiddetto *effetto San Matteo*, per cui le probabilità che contributi scientifici acquisiscano visibilità nella comunità di riferimento sono tanto più alte quanto più vengano presentati da scienziati di alto profilo piuttosto che da giovani ricercatori. Tale tendenza, reputata disfunzionale a livello del singolo ricercatore, è in realtà funzionale al mantenimento del sistema, sulla scorta del fatto di poter dirigere l'attenzione della comunità verso scoperte che stenterebbero ad essere accettate e di operare un'attività di selezione nel gran numero di contributi che il sistema produce (Merton, 1968).

A partire dagli anni Settanta, e in particolare con la pubblicazione nel 1974 del libro di Barnes *Scientific Knowledge and Sociological theory*, si proponeva sostanzialmente l'abbandono di una visione per cui la spiegazione sociologica dovesse riguardare qualsiasi forma di deviazione dalle caratteristiche funzionali del sottosistema scientifico<sup>5</sup> e fu così che

---

<sup>4</sup> “Storia, tecnologia e società nell’Inghilterra del secolo XVII” (Merton, 1938, tr. it. 1975) verteva sul rapporto tra sviluppo della Scienza come istituzione della società e la diffusione dei valori religiosi legati al protestantesimo. Per Merton, la spinta alla pratica scientifica nella società britannica del Seicento si traduceva spesso in lavori senza alcuna finalità pratica e trovava giustificazione semplicemente nell’impegno allo studio empirico della natura in quanto rivelatore della grandezza divina. Tale fattore di stimolo derivato dalla dottrina religiosa calvinista caratterizzata da un’ideale di esaltazione del lavoro come via di salvezza per l’individuo, concorreva ad istituzionalizzare la scienza nella società, conferendole “*social stability*” ed “*institutional integrity*” attraverso una codificazione sociale del ruolo dello scienziato.

<sup>5</sup> Ciononostante l’impostazione di Merton resiste nell’attività definitoria della Scienza, se è vero che Luciano Gallino (1995) ha definito la Scienza, dal punto di vista della sociologia come “sottosistema d’azione sociale” all’interno del sistema socioculturale “che forma con i sistemi comunitario, politico ed economico il primo livello di differenziazione d’una società moderna”.

l'approccio classico mertoniano trovò un contraltare in due correnti accademiche generalmente indicate come la Scuola di Edinburgo e la Scuola di Bath, che contraddistinsero l'emergere della *Sociology of Scientific Knowledge* (SSK) all'interno dei *Scienze and Technology Studies* (STS). L'approccio della Scuola di Edinburgo, spesso identificato con le teorizzazioni del filosofo David Bloor e il suo "programma forte" (cfr. Bloor, 1981), si proponeva di indirizzare la strategia degli studiosi ad un'analisi dei contenuti concreti della produzione scientifica anziché far stazionare gli studi ad un livello d'astrazione superiore. Ispirandosi a principi metodologici di causalità dei processi conoscitivi, imparzialità, simmetricità e riflessività, e forti dell'assunto secondo cui la creazione e la cessazione di qualsiasi tipo di conoscenza debbano essere spiegate in termini sociali, Bloor e colleghi (cfr. Bloor, 1976) profetizzarono un intervento della Sociologia nello studio della conoscenza scientifica da implementarsi in quelle situazioni in cui il corso automatico della razionalità e del progresso scientifico venisse ostacolato dalla presenza di fattori sociali, quali la religione, la politica o la cultura (Bucchi, 2004). Tale approccio doveva essere applicato sia nei riguardi della conoscenza considerata reale, ossia basata sui fatti, sia quella gravata da parzialità, errore o che voglia essere ingannevole. Sull'altro fronte, la Scuola di Bath, col suo programma empirico del relativismo, ha indirizzato il proprio interesse verso l'analisi della cosiddetta "flessibilità interpretativa" dei risultati di ricerca, cioè la possibilità che questi si prestino a più di una interpretazione, e ai meccanismi attraverso cui viene raggiunta la risoluzione di una controversia scientifica. L'attenzione di questi studiosi è focalizzata su quei "punti critici di contingenza" nel processo della ricerca che scaturiscono allorché determinati eventi, come ad esempio la rilevazione del dato o l'interpretazione dei risultati, avrebbero potuto verificarsi in maniera diversa se non si fosse raggiunta una mediazione su quale fosse il modo definito corretto di rilevare il dato o quale fosse l'interpretazione definita più affidabile dei risultati. Nel caso citato, le definizioni di correttezza e di affidabilità sono addebitabili, secondo la visione relativista, a criteri sociali solitamente rappresentati dall'inserimento dei ricercatori che li sostengono nelle reti scientifiche più rilevanti, nonché influenzati dalla loro nazionalità degli stessi, dalla loro reputazione e da quella delle istituzioni ospitanti (Collins, 1992).

L'assunto che le circostanze sociali, le convinzioni dei ricercatori e le relazioni tra gli stessi siano fattori rilevanti nella creazione di risultati scientifici tanto quanto i fenomeni naturali studiati ha aperto la strada a sempre più massicci studi etnografici sul campo da parte

dei sociologi della conoscenza, esaltando la svolta paradigmatica verso lo studio delle dinamiche micro-sociologiche inerenti ai processi scientifici (*cf.* Latour e Woolgar, 1979; Knorr-Cetina, 1981).

### *1.1.3 L'etnografia in laboratorio e la scienza in azione*

L'approccio basato sul microlivello dei processi di produzione della conoscenza ha trovato negli studi etnografici sulle dinamiche interne a specifici gruppi di ricerca una declinazione intesa a privilegiare la dimensione relazionale e negoziale dell'attività scientifica. Per gli studiosi afferenti al filone noto all'interno degli STS come "studi di laboratorio", la sede dell'attività scientifica, vista come sito d'azione, costituisce un'opportunità per indagare le dinamiche di produzione del sapere poiché consente di penetrare all'interno della "scatola nera" delle contingenze legate ai processi di ricerca. Le conversazioni che avvengono durante gli esperimenti, gli scambi di pareri, gli appunti, le stesure provvisorie di *paper* si pongono in questo senso quali fonti informative privilegiate per l'analisi dei meccanismi. Tuttavia, si evidenzia come il "sito di ricerca" non costituisca un'unità isolata entro la quale descrivere e circoscrivere il processo. Le situazioni e le contingenze osservate in laboratorio devono tenere in considerazione il carattere di mediazione, che trascende i singoli gruppi ed alimenta "catene di relazione transepistemiche" che coinvolgono non solo i ricercatori, ma le agenzie che li finanziano, i fornitori di strumenti e materiali e i *policymaker* (Knorr-Cetina, 1982).

Latour e Woolgar sono considerati pionieri di questo filone di studi. Nel classico *Laboratory life* (Latour e Woolgar, 1979), essi posero sotto la lente d'ingrandimento le attività di un gruppo di ricerca presso il Salk Institute di La Jolla in California. La ricerca della conoscenza che avviene in laboratorio venne riconfigurata da questi autori nei termini di catene decisionali costituite da negoziazioni basate su aggiustamenti progressivi che trovano compimento nella pubblicazione di articoli sulle riviste scientifiche. La difesa dei risultati scientifici, e prima ancora la concettualizzazione degli oggetti di studio e le forme di presentazione dei dati, sono il frutto di strategie discorsive che si muovono entro le argomentazioni proprie dei campi disciplinari, ma fanno riferimento anche ad "argomenti non epistemici". Il *paper* scientifico infatti "dimentica molto di ciò che è accaduto in laboratorio" e lo ricostruisce razionalmente e selettivamente, costituendo una forma letteraria propria, distante dallo stile di ragionamento informale che caratterizza il lavoro concreto interno ai

gruppi. Sulle riviste il processo di ricerca viene infatti presentato attraverso un modello discendente “a cascata” in cui tutte le mosse del ricercatore discendono da specifici obiettivi fissati in partenza, celando fattori quali casualità annesse al processo e disponibilità contingente di risorse.

Uno sviluppo e completamento degli studi di laboratorio fu costituita dalla *actor-network theory* (ANT), un approccio che concettualizza l’ambiente sociale di ricerca in una prospettiva di alleanze dinamiche. La ANT è stata definita come un “metodo crudo” per imparare dagli attori senza imporre su di loro una definizione *a priori* sulle loro capacità di costruzione del mondo (Latour, 1999). Partendo dall’assunto che la realtà è un fenomeno emergente e relazionale, gli studiosi intendono i *network* di ricerca quale segno della formazione di alleanze dinamiche nel contesto scientifico. Gli attori che vi sono inseriti sono trattati da mediatori che rendono visibile il movimento dell’azione sociale. L’inclusione in un *network* comporta l’allineamento ad un programma d’azione comune, a una sorta di copione che delega ruoli e traiettorie nel processo di ricerca. Questi “copioni” possono essere delegati anche a macchine, tecnologie o altre componenti materiali, artefatti che divengono essi stessi attori mobilitando competenze e relazioni con agenti umani. La ANT dunque elimina la distinzione tra tecnologia, attori non umani e umani, considerandoli come nodi correlati in reti sociotecniche in costante evoluzione, e postulando la loro inseparabilità nelle attività concrete. Le alleanze dinamiche producono il risultato di istituzionalizzare *fatti scientifici*. Un enunciato o un risultato di ricerca non può infatti penetrare nel patrimonio comune, acquisendo lo status di *fatto*, se rinuncia all’appoggio e alla cooperazione di tutta una serie di “alleati”, non solo all’interno ma anche fuori del laboratorio<sup>6</sup>. Come rileva Bucchi (2004), secondo i sostenitori dell’approccio ANT, una citazione in bibliografia, un collega col quale si collabora, un’apparecchiatura che consente di attuare un particolare studio, un committente disposto ad investire in una ricerca, un’osservazione confermativa di una tesi, un gruppo di potenziali utenti per un’innovazione tecnologica, sono tutti attori della rete di sostegno che alimenta quel processo che traduce un insieme di risultati sperimentali e asserzioni in una scatola nera: un fatto scientifico o un prodotto tecnologico.

Negli anni Novanta, gli STS subiscono quella che Woolgar definisce una “svolta tecnica” e, facendo tesoro di molte delle acquisizioni teoriche e metodologiche elaborate nelle ricerche

---

<sup>6</sup> Contrariamente alla proposta della Scuola di Edimburgo, la tesi è dunque quella di considerare il fatto scientifico non come un punto di partenza a cui imputare l’influenza di alcuni fattori sociali, ma come il risultato del determinarsi una complessa rete di alleanze che consentono al ricercatore di “difenderlo” (Bucchi, 2004).

sulla vita dei laboratori scientifici, cominciano a puntare l'attenzione sulla progettazione e l'uso e la produzione delle tecnologie dell'informazione (*cf.* Star, 1995). L'esperienza etnografica, che ha disvelato i meccanismi interni al mondo della ricerca scientifica si sarebbe rivelata utile all'emersione di tematiche quali la standardizzazione, la selezione, la conservazione degli strumenti e dei metodi di produzione del sapere (Star e Bowker, 2007, *cf.* anche Clarke e Fujimura, 1992).

#### *1.1.4 I mutamenti della scienza contemporanea*

Più recentemente gli studi sulle dinamiche generali di produzione della conoscenza scientifica hanno spostato la loro attenzione verso i mutamenti indotti sul modo di fare di ricerca dal contesto socio-economico in cui essi sono immersi, applicando una prospettiva evolutiva in riferimento ai cambiamenti che investono la società moderna dal punto di vista politico, tecnologico e culturale.

La tesi centrale di Gibbons *et al.* (1994) in *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies* fu che la natura del processo di ricerca si stava trasformando radicalmente per effetto di alcune tendenze sulla cui effettività gran parte degli scienziati convergeva: (i) una crescente eterodirezione delle priorità di ricerca, per cui a livello europeo e nazionale con la creazione dei Programmi Quadro e di programmi di ricerca tematici si stava tentando di far incontrare la capacità di produrre ricerca con la soddisfazione di bisogni economici e sociali; (ii) la commercializzazione della ricerca, come conseguenza della riduzione di fondi pubblici, che provocava da una parte una crescente ricerca di finanziamenti esterni e dall'altra alimenta una maggiore consapevolezza da parte delle università del valore della proprietà intellettuale dei propri prodotti; (iii) la crescente enfasi sull'aspetto di *management* della ricerca indirizzato alla valutazione della sua efficienza e al miglioramento della sua qualità. Sulla scorta di questi fattori, gli autori descrivevano dunque un vero e proprio cambiamento paradigmatico, per dirla à la Kuhn, tra una *knowledge production mode 1* e una *knowledge production mode 2* – la prima caratterizzata primariamente dall'egemonia della produzione teoretica o sperimentale, da un assetto disciplinare rigido, tassonomico e gerarchico basato sull'accademia e su grandi laboratori centralizzati e la seconda contraddistinta da reti di conoscenza distribuita, transdisciplinarietà, orientamenti applicativi, riflessività e da gradi di responsabilità (*accountability*) da gestire su fronti molteplici. La prima delle nuove

caratteristiche si riferisce all'ampio spettro di siti nei quali la conoscenza è prodotta e circola: i progressi delle tecnologie di informazione e comunicazione hanno favorito e stimolato le interazioni tra comunità di ricerca, che erano dapprima limitate da impedimenti fisici e tecnici. Il fattore di transdisciplinarietà si esprime invece con la mobilitazione di una gamma di prospettive e metodologie pratiche da orchestrare e far comunicare per risolvere problemi di ricerca che possono avere risvolti in molti campi del sapere. Connesso a quanto appena descritto, l'orientamento applicativo costituisce un allontanamento dai modelli di scienza "pura" e fondata sulla teoresi per far posto al trasferimento tecnologico e a disegni di ricerca in cui lo sviluppo di metodologie e la disseminazione dei risultati siano tarati rispetto al problema scientifico da risolvere. La quarta caratteristica è costituita dalla riflessività, per cui l'ambiente influenza la scelta delle domande di ricerca e degli utilizzi degli utenti. Infine, l'ultima novità rilevata dagli autori è la sottoposizione della pratica scientifica a principi di *accountability* nell'ottica di un effettivo collegamento tra risultati raggiunti e allocazione delle risorse: ciò ha richiesto nuove forme di controllo di qualità rispondenti a definizioni multiple al fine di facilitare i processi selettivi dei *policymaker* e delle agenzie di finanziamento e tenere conto di eventuali impatti economici e risvolti etici. Il netto discrimine individuato tra *knowledge production mode 1* e *knowledge production mode 2* è stato ribadito ed ampliato qualche anno più tardi da tre dei sei autori della sopracitata pubblicazione con *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty* (Nowotny *et al.*, 2001), lavoro in cui si pone in evidenza come il cambiamento nel modo scientifico si innesti in un più generale cambiamento della società che diventa essa stessa *mode 2*. La penetrazione di tendenze generali nel sub-sistema scientifico ha sancito una confusione tra scienza e società tale che esse costituiscono ormai un unico sistema integrato, caratterizzato dal costituirsi di un'agorà pubblica grazie a cui si alimenta una "distribuzione sociale dell'*expertise*". La negoziazione delle soluzioni a problemi scientifici non è più delegata alla comunità scientifica ma trova spazio in nuovi contesti aperti anche a non esperti.

Le argomentazioni di Gibbons *et al.* sono state messe in discussione da alcuni autori che hanno preferito l'adozione di uno schema interpretativo che propone un bilanciamento tra gli elementi delle due tipologie, che continuano a co-esistere e co-evolversi (*cf.* Martin, 2003). Sulla stessa scia della constatazione di profondi mutamenti che investono il mondo scientifico parallelamente a quelli della società si pone invece Ziman con la sua definizione di scienza "post-accademica". Il cambio di marcia è argomentato sia dal punto di vista pratico che da

quello epistemologico e parte da lontano, dalla fine della Seconda Guerra Mondiale, quando gli scienziati abbandonano le proprie “torri d’avorio” e condividono le decisioni rilevanti in merito al proprio lavoro insieme ad altri gruppi sociali: politici, burocrati, manager delle industrie, militari, movimenti di opinione e organizzazioni non governative, società nel suo complesso (Ziman, 1998). Si costituisce dunque una scienza “collettivizzata” in cui all’individualismo proprio dello stile accademico, si sostituisce un’azione collettiva di produzione della ricerca che attinge a risorse economiche e tecnologiche e cerca un più ampio consenso alla propria attività. Allo stesso tempo, fanno breccia nuove prospettive di interrelazione tra oggetti di ricerca che travalicano gli ambiti disciplinari e danno vita a nuove metodologie.

## **1.2 Gli avanzamenti delle ICT e il lavoro di ricerca**

La diffusione su scala planetaria delle tecnologie di informazione e comunicazione (ICT) ha prodotto un radicale cambiamento nella vita quotidiana e nel lavoro degli individui. Da un lato autori come McLuhan (1964) hanno rilevato come i componenti elettronici e le tecnologie di base generino una forte influenza sul modo di pensare delle persone; dall’altra studiosi quali Friedmann (1978) hanno evidenziato le implicazioni seguite all’introduzione dei processi di automazione nell’organizzazione del lavoro.

In questo contesto, le attività legate alla produzione di conoscenza scientifica non si sono limitate ad essere passivamente investite dai processi di cambiamento tecnologico ma, per certi versi, hanno incarnato l’esemplificazione più emblematica di nuove dinamiche non solo organizzative, ma anche sociali e culturali. Se dapprima l’introduzione del computer in ambiente scientifico ha innescato dinamiche tali da rivoluzionare il modo stesso di immaginare e svolgere l’attività di ricerca in tutti i campi disciplinari, imponendosi come strumento di lavoro essenziale, grazie alle caratteristiche di velocità di risposta, precisione, capacità di sistematizzazione e consistenza logica (National Academy of Sciences, 1989), in anni recenti l’implementazione delle nuove infrastrutture dell’informazione, dalla nascita di Internet in poi, ha posto il “sistema scienza” in posizione privilegiata nello sfruttamento delle potenzialità derivanti dal progresso delle reti internazionali di comunicazione (Star e Bowker, 2007).

### 1.2.1 Dal computer all'avvento del Web 2.0

I primi calcolatori ad uso di ricerca, introdotti dalla fine degli Quaranta<sup>7</sup>, furono in un primo tempo concepiti come rimpiazzati delle vecchie tecnologie di calcolo, ed erano utilizzati dai ricercatori quasi solo al fine di svolgere operazioni matematiche in sostituzione di carta e penna, abachi e calcolatori meccanici. Solo una frazione relativamente ridotta di studiosi poteva avervi accesso, essendo la loro disponibilità gravata da alti costi di acquisto e di gestione e dalla necessità di costante supporto tecnico. Nei primi anni dalla loro introduzione, il beneficio sul lavoro di ricerca era sbilanciato sul piano dell'analisi quantitativa piuttosto che sull'organizzazione generale del lavoro di ricerca: le funzioni di calcolo divenivano affidabili e veloci, ma le attività quotidiane dei ricercatori, nelle fasi antecedenti e posteriori all'elaborazione del dato, non risentivano nel complesso di mutamenti rilevanti<sup>8</sup>.

Dall'inizio degli anni Ottanta, per effetto del progresso legato delle tecnologie di semiconduzione e del conseguente incremento delle velocità di elaborazione, e a seguito della riduzione dei costi di gestione, l'uso del computer in ambito scientifico è andato diffondendosi capillarmente e i suoi effetti sulle attività quotidiane dei ricercatori sono divenuti sempre più rilevanti e pervasivi. Un'indagine svolta alla fine degli anni Ottanta dal Panel on Information Technology and the Conduct of Research della National Academy of Sciences americana (1989) già segnalava consistenti vantaggi concentrati su tre dimensioni: (i) la raccolta e l'analisi delle informazioni; (ii) le comunicazioni e la collaborazione tra ricercatori e (iii) la conservazione e il recupero dei dati. Il primo aspetto faceva riferimento alla crescita esponenziale di dati reperibili dal ricercatore e alla sempre maggiore disponibilità di pacchetti di *software* specifici per l'attività di ricerca, che promettevano di offrire una grande varietà di opzioni di analisi e grazie ai quali diveniva possibile processare milioni di righe. Il computer ha iniziato a consentire lo sviluppo di analisi simulate, che sostituivano parzialmente l'esperimento nelle situazioni in cui non vi fosse disponibilità di dati, e a permettere la modellizzazione di sistemi dinamici complessi, peraltro costringendo in un certo senso i ricercatori a sviluppare un'attività teorica più precisa al fine di ottenere un valido

---

<sup>7</sup> È del 1946 l'introduzione di ENIAC, il primo supercomputer che poteva svolgere fino a 5000 addizioni al secondo.

<sup>8</sup> Per comprendere l'impatto che fin da subito l'utilizzo del computer ha generato sulla spinta agli studi quantitativi, si può citare il caso emblematico degli articoli nel campo delle scienze sociali: si stima che alla fine della Seconda Guerra Mondiale nel 50% circa dei lavori pubblicati su *American Sociological Review* e *American Journal of Sociology* non vi fosse traccia di alcuna analisi quantitativa, mentre trent'anni più tardi, tale percentuale scendeva drasticamente al 13%" (Gigerenzer, 2001).

quadro di implicazioni rispetto alle loro idee (Gigerenzer, 2001). Il secondo fattore includeva l'insieme delle attività di comunicazione svolte per mezzo del terminale ed atte a facilitare ed articolare il processo scientifico: grazie alla posta elettronica legata a Internet, che andava parzialmente a sostituire la posta tradizionale e il contatto telefonico, e al *word processing* i ricercatori potevano scambiarsi e condividere le informazioni in modo rapido e istantaneo e, al contempo, creare nuovi *network* collaborativi o alimentare gli esistenti. La comunicazione interna tra ricercatori iniziava a sfidare le barriere geografiche e poteva finalmente avvenire in maniera quasi istantanea, non limitandosi agli scambi verbali, ma estendendosi al passaggio di *file* testuali o *set* di dati. Il terzo punto riguarda la modalità di gestione e recupero elettronici delle informazioni, che ha comportato un vantaggio economico nella fase di conservazione del dato e un vantaggio logistico, allorché le consultazioni in digitale non hanno più reso necessario lo spostamento fisico. La digitalizzazione del processo di ricerca ha permesso di risparmiare un gran numero di ore-lavoro e consentito la conservazione di grandi quantità di dati e risultati che potevano essere predisposti a riutilizzi e controlli successivi.

Da metà degli anni Novanta, assecondando il crescente fenomeno di alfabetizzazione informatica, il computer ha smesso in modo definitivo il ruolo di mero calcolatore per insediarsi stabilmente in molteplici fasi del processo di ricerca (Arcila-Calderón *et al.*, 2015), dalla fase concettuale – con la ricerca della letteratura su *database* elettronici sempre più strutturati – alla stesura di un articolo scientifico su versioni di *word processor* maggiormente performanti. Uno sfruttamento più incisivo della rete Internet ha mutato le tradizionali norme del comportamento scientifico introducendo nuovi modi di implementare i disegni di ricerca, soprattutto in riferimento all'acquisizione del dato, e moltiplicando nuove modalità comunicative disponibili (forum, chat, videoconferenze). In un contesto globale nel quale la Rete ha iniziato a rappresentare lo spazio più comune nel quale le persone interagiscono e collaborano per creare, condividere e disseminare conoscenza (*cf.* Giles, 2005), l'utilizzo delle risorse legate al *web* ha prodotto, in forme diverse, un impatto consistente sulle attività di ricerca scientifica, innescando nuove consuetudini sia dal punto di vista istituzionale che organizzativo e proponendo nuovi modelli di *routine* che si configurano come evoluzioni di quelli emersi a seguito della prima ondata delle novità in campo ICT (Borgman, 2007). Già nel 2003, la dichiarazione di Berlino sull'accesso aperto alla letteratura scientifica, affermava che “*Internet ha radicalmente modificato le realtà pratiche ed economiche della*

*distribuzione del sapere scientifico e del patrimonio culturale*”<sup>9</sup>. Gallino (2003) ha rilevato un cambio di passo rispetto alle modalità tradizionali di lavoro in laboratorio che risulta evidente soprattutto nella trasformazione di certi atti quotidiani: azioni come la consultazione di libri o riviste in formato cartaceo o l’apertura di classificatori nelle biblioteche diventano via via più rare allorché Internet e il World Wide Web sono adoperati in maniera intensiva per svolgere ricerca e veicolare i contenuti da essi prodotti: “[il ricercatore scientifico] va direttamente in Rete, nella quale sa di trovare in abbondanza i documenti, gli articoli, i dati, i risultati di altri esperimenti, [...] si metterà in contatto con altri colleghi sparsi nel mondo, al fine di ottenere suggerimenti e commenti, [...] quando considererà maturi i risultati della ricerca, li diffonderà in qualche minuto, attraverso la Rete, a tutte le riviste specializzate, i media, i laboratori alleati e concorrenti”.

Uno dei primi cambiamenti rilevanti indotti per effetto della diffusione massiccia di Internet nel lavoro di ricerca è stato rappresentato dalla possibilità di accesso ad articoli e riviste in formato elettronico (Kennan, 2008). La distribuzione in Rete di pubblicazioni scientifiche diventava disponibile in diversi formati e si basava su insiemi di strumenti adattati sulle esigenze di specifiche comunità di ricerca e di servizi forniti da editori o biblioteche digitali. Tra le risorse fruibili a questo proposito è possibile, tra le altre, annoverare *database* citazionali come *Web of Science* di Thomson Reuter e "SciVerse" Scopus di Elsevier, ai quali è possibile accedere previa sottoscrizione di un abbonamento. Entrambi gli archivi sono tuttora ampiamente utilizzati allo scopo di attingere a documenti, tesi, libri, *abstract*, articoli di case editrici accademiche, ordini professionali e studi più rilevanti nel campo della ricerca accademica mondiale e per dare contezza delle loro ricadute scientifiche<sup>10</sup>. Inoltre, un’ulteriore innovazione è stata costituita dall’apertura dei primi archivi digitali su larga scala che rendono possibile l’accesso ai microdati di ricerca. Tuttavia tale innegabile avanzamento sul lato del *data mining*, con la possibilità di interrogare grandi *database* elettronici *online*, e il potenziamento dell’attività di condivisione tra ricercatori è stato controbilanciato da potenziali rischi alla proprietà intellettuale per via di attacchi informatici fraudolenti o da attività di plagio degli articoli scientifici (*cfr.* Davison, 2003).

---

<sup>9</sup> Dichiarazione di Berlino, promossa dai partecipanti alla Conference on Open Access to Knowledge in the Sciences and Humanities, del 22 ottobre 2003.

<sup>10</sup> Le due risorse, pur utilizzando analoghi principi propri delle analisi bibliometriche, si differenziano per caratteristiche e contenuti. Sono inoltre da segnalare motori di ricerca “speciali” come Google Scholar che affiancano ai tradizionali strumenti per la ricerca bibliografica anche funzionalità di tipo bibliometrico.

A partire dalla metà degli anni duemila il rapporto tra ICT e scienza ha iniziato ad assumere forme peculiari allorché un nuovo modo di concepire la Rete ha prodotto nel lavoro dei ricercatori implicazioni inedite legate allo sfruttamento delle potenzialità e delle caratteristiche del nuovo ambiente di comunicazione virtuale, il *Web 2.0*<sup>11</sup> (Ponte e Simon, 2011). La nuova visione alla base della piattaforma, che già da una decina d'anni aveva fortemente supportato il processo scientifico, si sostanzia principalmente nella generazione di dati e contenuti operata dagli stessi utenti e si caratterizza per un comune sforzo collaborativo per la loro condivisione, in ciò rimodulando il ruolo degli utilizzatori da semplici fruitori dell'informazione verso una triplice riqualificazione a produttori, editori e consumatori dell'informazione digitale<sup>12</sup>. I ricercatori hanno potuto dunque utilizzare l'ambiente *online* non solo al fine di consultare le pubblicazioni ma anche per comunicare, distribuire e condividere informazioni e dati attraverso un più ampio spettro di modalità, strumenti e piattaforme. La proliferazione di grandi *database online*, continuamente alimentati da una molteplicità di studiosi dislocati su tutto il pianeta, ha generato una ricchezza informativa straordinaria che viene utilizzata per condurre studi altrimenti fuori dalla portata del singolo ricercatore<sup>13</sup>. Un'ampia gamma di strumenti *online* è iniziata ad emergere e ad essere sfruttata dai ricercatori per la loro attività, favorendo l'emersione di nuovi modelli di organizzazione dell'attività di ricerca basati su un impegno attivo dei ricercatori, siano essi coinvolti in veste di autori di prodotti scientifici da diffondere, di lettori o di *reviewer*. In primo luogo, ci si riferisce ai siti di *social networking*, come *ResearchGate* o *Academia.edu*, che hanno lo scopo di collegare tra loro ricercatori con interessi comuni e mettere a disposizione propri prodotti di ricerca. Questi servizi forniscono un'esperienza simile a quella offerta da *Facebook* o *LinkedIn*, avendo però come target la comunità di ricerca: gli iscritti a questi servizi creano profili personali e sono incoraggiati a elencare le

---

<sup>11</sup> La nozione di *Web 2.0* è inclusiva di una molteplicità di elementi che descrivono da una parte l'introduzione di nuove pratiche sociali di produzione e disseminazione dell'informazione da parte di individui o gruppi, dall'altra una riconfigurazione dal punto di vista tecnico degli ambienti digitali su cui queste pratiche si innestano.

<sup>12</sup> A far da supporto a questo nuovo contesto di pratiche comunicative vi è stata la rapida diffusione di nuovi servizi e applicazioni il cui sviluppo ha consentito di concepire l'ambiente-Rete come una piattaforma in grado di generare e riprodurre contenuti attraverso sistemi di *network* che alimentano sia il loro ciclo di produzione/consumo che l'interazione tra utenti (Anderson, 2007). Tra le più famose applicazioni e i più rilevanti servizi di questa nuova "generazione" di Rete è possibile annoverare Skype, gli ambienti wiki, blog and microblog (come Twitter), *social network* quali Facebook.

<sup>13</sup> È il caso, ad esempio, di *GenBank*, l'enorme archivio di dati genetici gestito dal *National Center for Biotechnology Information* degli Stati Uniti, uno strumento *online* pubblicamente accessibile, che sistematizza il lavoro di migliaia di biologi che hanno caricato i propri dati di laboratorio in modo che altri scienziati possano utilizzarli in vista del progresso scientifico nella materia.

loro pubblicazioni e altre attività accademiche, caricare copie di manoscritti di cui sono autori, e costruire legami con gli altri studiosi. Tuttavia, come riportato da Rinaldi (2014), tali risorse non vengono solitamente utilizzate per condividere lavori o dati sperimentali ma sono perlopiù concepiti come mezzo per accrescere la visibilità pubblica di scienziati, in particolare a beneficio dei finanziatori della ricerca, e come piattaforme dove sia possibile svolgere un costruttivo dibattito scientifico. Inoltre, altre “pratiche 2.0” hanno iniziato ad affacciarsi, come quella dell’apertura dei *blog* scientifici creati principalmente per rivendicare la priorità dei creatori degli stessi su temi di ricerca specifici e stimolare il dialogo sulla ricerca intrapresa, rappresentando allo stesso tempo un modo di impegnarsi nella divulgazione scientifica attraverso la relazione con un pubblico che altrimenti avrebbe difficilmente interagito con le pubblicazioni scientifiche tradizionali (Wilkins, 2008).

Se da un lato è ormai usuale per gli scienziati sottomettere i propri lavori alle riviste per via digitale o pubblicarli autonomamente sui *blog*, esporre il proprio curriculum vitae sui siti di *social networking*, ricercare su Internet i lavori più rilevanti, comunicare a distanza con i propri collaboratori e iscriversi alle conferenze sfruttando le nuove tecnologie (Mahapatra, 2010), Paccagnella (2012) segnala come, nonostante la retorica del *web* partecipativo si riveli perlopiù ottimistica quando vengono esaltate le sue potenzialità applicate ai processi di produzione della conoscenza scientifica, alcune voci critiche abbiano sollevato dubbi sull’effettiva realizzazione di tale partecipazione<sup>14</sup>. Tra i ricercatori è permanentemente in atto un processo di *social learning* che caratterizza l’utilizzo del *Web 2.0* per le loro attività, non ancora compiuto del tutto (i termini della diffusione verranno chiariti nel par. 1.2.2), ma destinato a maturare nel breve/medio termine (Procter *et al.*, 2010). Ad ogni modo, fattori legati al *Web 2.0* come la diffusione di strumenti di comunicazione tecnologicamente avanzati e l’alimentazione di un panorama informativo sempre più ricco e condiviso hanno *de facto* prodotto l’emersione di nuove e più efficaci forme di collaborazione nella ricerca, da una parte allentando la pressione all’innalzamento dei tassi di produttività sempre più richiesti a fini valutativi, dall’altra alimentando il trasferimento di conoscenza all’interno della comunità afferente al *field* disciplinare e verso gli *stakeholder* (*ibidem*).

---

<sup>14</sup> Citando la “legge di Pareto” sulle disuguaglianze distributive della ricchezza economica, lo studioso rileva come “la gran parte dei contenuti generati dagli utenti sia prodotto in realtà da una ristretta minoranza di utenti particolarmente attivi” (*ibidem*).

### 1.2.2 Parametri strutturali e determinanti dell'adozione di ICT

L'adozione di nuove pratiche presuppone negoziazioni e processi di scoperta (Williams *et al.* 2005) in un contesto fluido dove nuove pratiche si innestano nel tempo su nuovi supporti: da una parte i potenziali utenti “combattono” per scoprire, esplorare e sfruttare le nuove capacità tecnologiche adattandole ai loro fini e contesti, dall'altra gli sviluppatori cercano di comprendere gli utenti e gli usi emergenti, generando un processo di sperimentazione potenzialmente senza fine. In riferimento all'introduzione di innovazioni tecnologiche nel contesto lavorativo, non vi è in letteratura un chiaro consenso sull'identificazione delle determinanti di adozione di ICT. Venkatesh *et al.* (2003) hanno provato a uniformare la vasta produzione di *framework* teorici per la comprensione e la modellizzazione di tali processualità (*Diffusion of Innovations, Theory of reasoned action, Technological acceptance model, Planned Behavior Theory, PC use model, Social cognitive theory*), rilevando come ogni costrutto fosse fortemente legato al concetto di *performance expectancy*, un indicatore composito, i cui elementi costitutivi sono riscontrabili nei concetti di utilità percepita, motivazione estrinseca, vantaggio relativo, *job-fit* e aspettative di risultato. Nel modello creato dagli autori, conosciuto come *Unified theory of acceptance and use of technology* (UTAUT), la *performance expectancy* è moderata da età e genere (l'effetto risulta più forte per giovani e per uomini) ed è definita come “il grado a cui un individuo crede che l'utilizzo del sistema lo aiuterà a ottenere un guadagno nella performance di lavoro” (*ibidem*).

Data la loro attuale diffusione, negli ultimi anni si è riscontrato un crescente interesse verso lo studio delle dinamiche d'adozione di strumenti *Web 2.0* per differenti scopi riguardanti il lavoro di ricerca. A livello globale, la diffusione delle più recenti risorse per la diffusione scientifica e l'interesse verso ICT per la ricerca non risulta essere simile in tutti i campi disciplinari (Dutton e Meyer, 2008). È emerso infatti che i ricercatori nel campo delle *hard science* utilizzano più estensivamente le nuove opportunità digitali mentre, nel campo delle scienze sociali, l'adozione è meno diffusa e più eterogenea, anche se negli ultimi anni un insieme di tecniche, strumenti e dinamiche sta trovando spazio all'interno delle “*digital humanities*” (*cfr.* tra gli altri Dutton, 2010). Procter *et al.* (2010) hanno riscontrato una significativa correlazione tra il tasso di adozione dei *tool 2.0* e le caratteristiche anagrafiche, quali età e genere, e l'inquadramento del ricercatore, concludendo che i ricercatori maschi più anziani e con livelli accademici più alti sono più inclini all'uso del *Web 2.0* per ricerca.

Inaspettatamente, lo studio non ha riscontrato che i giovani e i cosiddetti *nativi digitali* fossero gli utenti più estensivi degli strumenti *social* in ambiente accademico. Tuttavia, tale rilievo non è stato sempre supportato dalle ricerche successive. Se da una parte i risultati del report *Researchers of tomorrow* (JISC e British Library, 2012) hanno sottolineato come la nuova generazione di ricercatori effettivamente non utilizzi le tecnologie digitali al loro pieno potenziale, dall'altra i dati di Arcila-Calderón *et al.* (2015) mostrano di contro come la diffusione scientifica tramite ICT stia diventando sempre più importante per i giovani ricercatori. Entrando nel merito degli specifici servizi offerti dalla nuova tecnologia di Rete, uno studio di Pearce (2010) ha mostrato che l'età può essere considerata una determinante all'adozione di strumenti orientati alla diffusione scientifica, mentre ha riscontrato una correlazione positiva tra il genere e il tasso di adozione e conoscenza di altri strumenti avanzati più orientati all'analisi dei dati e alla loro conservazione.

Procter *et al.* (2010) e Ponte e Simon (2011), rispettivamente in uno studio focalizzato sui ricercatori inglesi e in un lavoro condotto attraverso una *survey* a livello europeo, hanno evidenziato come, a cinque e sei anni dalla coniazione del termine *Web 2.0* (*cfr.* O'Reilly, 2005), un numero sempre maggiore di studiosi si affidasse o esprimesse favorevolmente l'intento di affidarsi in futuro alle nuove pratiche di comunicazione *online* non solo per promuovere o diffondere un *paper* o un articolo, ma anche per coordinare lavori basati su collaborazioni o condividere artefatti digitali di ricerca. Tuttavia, le due ricerche erano concordi nel mostrare un certo grado di "conservatorismo" nell'adozione delle nuove tecnologie digitali. Ancora molti ricercatori infatti dichiaravano di affidarsi a meccanismi tradizionali di scambio delle informazioni che abbiano dimostrato di funzionare in passato e che fossero radicati in sistemi di valutazione e ricompensa istituzionali e professionali da tempo consolidati, in ciò seguendo la tendenza ad utilizzare strumenti già sperimentati piuttosto che compiere un investimento rischioso su novità inesplorate. Ciò che è emerso è che la proliferazione di risorse e la costante frammentazione dell'offerta di strumenti e piattaforme si è esposta paradossalmente a inattese resistenze dei potenziali *user* sia per via della difficoltà di tenere traccia degli sviluppi delle nuove opportunità e sia a causa della carenza di informazione o preparazione per valutare i benefici per particolari tipi di attività. Per questo motivo, in particolare Procter *et al.* (2010) sottolineano che solo servizi *Web 2.0* costruiti incrementalmente su pratiche esistenti, con caratteristiche di genericità, intuitività, dal facile utilizzo e a costo zero, hanno conosciuto una più rapida ascesa rispetto a strumenti

più specializzati e settoriali che si propongano di esplorare nuove modalità di comunicazione scientifica *online*. Ad ogni modo, laddove i vantaggi di adozione siano stati percepiti come sufficientemente alti e i costi d'investimento favorevolmente bassi per motivare l'adozione di una comunità, le nuove risorse riescono comunque a raggiungere masse di utenti generando externalità di *network*<sup>15</sup>, portando le stesse ad essere pervasivamente adottate da particolari comunità. Gli studiosi individuavano nel ruolo giocato dai gruppi di ricerca locali e dagli intermediari della conoscenza, all'interno dei dipartimenti o dei *network*, le determinanti chiave per stimolare, attraverso modalità formali o informali, l'adozione di servizi e strumenti del *Web 2.0*. Sulla stessa scia, Arcila-Calderón *et al.* (2015), applicando il modello UTAUT in contesti di ricerca, seppur con qualche variazione, hanno sostanzialmente confermato che la *performance expectancy* – intesa in questo caso come complesso di aspettative riguardanti il futuro della produttività scientifica e della generazione di nuove scoperte – fosse un importante predittore sull'utilizzo di ICT in ambiente scientifico; pertanto, i ricercatori che credono che le nuove tecnologie possano aiutarli nella produzione di nuova conoscenza le adotteranno con maggiore probabilità. Tuttavia, sia rispetto all'utilizzo di ICT per la disseminazione dei risultati che per l'analisi dei dati, il più forte predittore del modello è stato identificato nella collaborazione scientifica.

Nonostante una tendenza all'utilizzo delle nuove tecnologie nella comunicazione, Scheliga (2015) mostra tuttavia come i ricercatori utilizzino canali multipli di comunicazione per le loro comunicazioni di ricerca con i propri coautori, non rintracciando segni per cui la comunicazione digitale induca a un declino di quella *face-to-face*, che rimane una pratica densa di implicazioni sociali per i soggetti e largamente istituzionalizzata (*cf.* anche Borgman, 2007). Al contrario, l'autrice ha rilevato, in uno studio sui ricercatori tedeschi, una correlazione positiva tra la comunicazione mediata digitalmente e gli incontri *face-to-face*. Questi ultimi vengono ancora considerati molto rilevanti ai primi stadi del disegno di ricerca allorché vengano prese importanti decisioni sui passi successivi; negli altri stadi la comunicazione digitale dimostra di assumere un valore maggiore in quanto veloce ed economica e, fattore non di secondo piano, poiché consente di tenere traccia delle decisioni prese e delle conversazioni intraprese. Altro fattore di freno è rappresentato dai casi di diffusione di informazione non sottoposta ai sistemi di controllo della qualità scientifica, quali

---

<sup>15</sup> Tale condizione si ottiene allorché il beneficio di ogni utente si innalza all'aumentare del numero degli utenti.

la *peer-review*<sup>16</sup>. I ricercatori tendono a manifestare resistenze a far uso di nuove forme di comunicazione scientifica che esulino dai *network* più prestigiosi perché non ripongono fiducia nelle risorse che non siano state sottoposte a un controllo da parte di pari (Camussone *et al.* 2011, scoperta coerente rispetto allo studio di Ware e Monkman, 2008).

### 1.3 La nuova scienza: aperta, collaborativa e basata sul dato

Sulla scorta del potenziale per la ricerca che gli ultimi avanzamenti delle tecnologie promettono di supportare, la pubblicazione da parte della Royal Society del rapporto *Science as an open enterprise* (Boulton *et al.*, 2012) è apparsa avvalorare la tesi di quanti (*cf.* tra gli altri Gray [Hey *et al.*, 2009] e Nielsen [2011, tr. it. 2012], di cui si approfondirà il pensiero nel par. 1.3.1) hanno scorto nei primi anni del nuovo millennio un nuovo importante cambio di paradigma nelle dinamiche di produzione della conoscenza scientifica<sup>17</sup>. Le sei linee d'azione suggerite dal rapporto agli scienziati e alle istituzioni di ricerca andavano nella direzione di un importante cambiamento organizzativo, pratico e culturale nel lavoro dei ricercatori che fosse in grado di sfruttare le potenzialità legate alle evoluzioni ICT e, in particolare, al dilagare delle opportunità legate al *Web 2.0*. Venivano paventate: la necessità di una maggiore apertura degli scienziati, sia in termini di incentivo alle loro collaborazioni che al rapporto con la società; un crescente riconoscimento del valore della raccolta, analisi e comunicazione del dato scientifico; l'utilizzo di standard comuni per la condivisione dell'informazione in modo da renderla largamente fruibile; la pubblicazione dei dati in formati riutilizzabili; la formazione di *expertise* nella gestione dei dati digitali; la creazione di strumenti appropriati per l'analisi di grandi quantità di dati.

Alla luce della proliferazione e del crescente affidamento a risorse *online* e dell'organizzazione delle pratiche scientifiche basata su *team* di collaborazione che operano su infrastrutture di ricerca (*cf.* cap. 2), il rapporto ha delineato i contorni di un "nuovo

---

<sup>16</sup> Procedura costituita da un insieme di pratiche eterogenee e non standardizzate (frutto di un'evoluzione delle tecniche adottate dalla comunità autonoma ed indipendente degli studiosi), attraverso la quale un ricercatore sottopone un proprio lavoro al giudizio di altri ricercatori, i cosiddetti "pari", che ne stabiliscono la validità (Baccini, 2010).

<sup>17</sup> Il rapporto tra le peculiarità distintive del nuovo stadio tecnologico e modo di intendere e realizzare la pratica scientifica sono esemplificati dalla definizione di nuova scienza della Leibnitz Association (2014) secondo cui "*Science 2.0 deals with the investigation of new fields for research and development, originating from the application of new participative and collaborative internet technologies in all phases of research.*".

percorso per la scienza" basato sulla ricchezza e la condivisione del panorama informativo, esplicitando le linee guida per gestire la fase di transizione verso una nuova modalità di ricerca basata sull'intensità del dato (*data-intensive science*) e sulla maggiore capacità di fare rete per la risoluzione delle questioni scientifiche (*networked science*). Il carattere *data-intensive* designa una modalità di ricerca centrata sul valore del dato, in cui non solo le ipotesi vengono testate attraverso la raccolta e analisi di dati, ma sempre più spesso gli *output* di ricerca derivano dalla combinazione e l'estrazione di insiemi di dati già disponibili e accessibili (Anderson, 2008). Le collaborazioni e il *networking*, resi possibili dalle nuove tecnologie, sottolineano l'evoluzione della pratica scientifica nei termini di "una grande collaborazione nella quale gli scienziati condividono e costruiscono un enorme bene comune dell'informazione" (Nielsen, 2011, tr. it. 2012). Tale nuova configurazione dell'attività di ricerca genera, come in un ciclo senza fine, un'esplosione della circolazione di *output* non solo costituiti da nuova letteratura, ma anche *terabyte* di micro-dati scientifici sui quali sono basate le pubblicazioni originali, la cui disponibilità d'accesso rende possibile la replicabilità degli studi, la validazione dei risultati e l'estensione delle ricerche (Gorgolewski *et al.* 2013). Su queste basi, prende forma un processo scientifico sempre più "aperto", appropriatamente denominato *Open Science* (Boulton *et al.*, 2012), un "termine ombrello" che:

(i) indica la tendenza a rendere la ricerca scientifica e la diffusione dei dati e dei risultati accessibili a qualunque livello, dai cittadini agli scienziati amatoriali fino ai professionisti del settore, caratterizzandosi per un ampio sforzo collaborativo ad ogni fase del processo di produzione di nuova conoscenza;

(ii) comprende pratiche quali la pubblicazione aperta delle ricerche, iniziative a favore dell'*open access*, incoraggiando gli scienziati a praticare la cosiddetta *open notebook science*, ossia la divulgazione dei dati "grezzi" raccolti durante le fasi della ricerca, e rendendo più facile la pubblicazione e la diffusione delle informazioni scientifiche;

(iii) non coincide unicamente con l'attitudine da parte dei *performer* della ricerca a comportamenti di condivisione stimolati dalla rivoluzione informatica nelle pratiche scientifiche e dall'utilizzo delle tecnologie *Web 2.0*, ma coinvolge anche organizzazioni di ricerca, enti finanziatori e il pubblico generale nel dominio della scienza e il loro approccio verso l'infrastruttura tecnologica su cui poggia l'attività di produzione della conoscenza scientifica (Gagliardi *et al.*, 2014);

(iv) non si configura come una novità assoluta nel panorama scientifico<sup>18</sup>, ma piuttosto come il risultato di un processo evolutivo a lungo termine nella conduzione della scienza che negli anni recenti ha trovato una nuova e maggiore enfasi per via della nuova capacità infrastrutturale legata alle tecnologie di informazione. Dalla forma più tradizionale di organizzazione della scienza a quella attuale, il processo scientifico si è storicamente fondato sulla diade concatenata “creazione-disseminazione” della conoscenza: la prima guidata da curiosità, ingenuità ed etica, la seconda validata da pari scettici e diffusa dai mezzi di comunicazione disponibili (*ibidem*), per cui l’aumento esponenziale del bagaglio conoscitivo degli ultimi secoli appare connesso all’evoluzione delle pratiche di “apertura” legate al secondo termine della diade, di cui il progresso ICT è stato un potente abilitatore.

### *1.3.1 La reinvenzione della scoperta scientifica*

I termini in cui cambierà il modo di produrre conoscenza, espandendo la gamma di questioni risolvibili dalla mente umana ed introducendo nuovi modi di intendere l’attività scientifica nel suo complesso sono stati declinati in vario modo, convergendo però su un punto: se già da oggi le comunità scientifiche godono dell’opportunità di poter sfruttare nuovi metodi e strumenti informatici che consentono a singoli ricercatori e *team* di ricerca di avere a portata di clic volumi di osservazioni non paragonabili col passato, nel prossimo futuro potenzialmente tutta la letteratura e i relativi dati sperimentali potrebbero essere disponibili sulla Rete, gestibili grazie agli sviluppi delle tecnologie ICT e predisposti ad un accesso globale (Borgman, 2012; *cf.* Walker *et al.*, 2011). Si prevede l’emersione della cosiddetta Rete dei Dati (*cf.* Nielsen, 2011, tr. it. 2012), indicando con questo termine l’insieme di tutti i dati già esistenti sulla rete Internet e dei nuovi dati che gli utenti immetteranno o condivideranno nel tempo sul *web*: essa viene considerata come unico aggregato di informazioni in continuo accrescimento, interrogabile da ogni parte del mondo, sistematizzato, a seconda degli argomenti trattati, in archivi differenziati ma potenzialmente armonizzabili rispetto ad alcuni oggetti di riferimento. Tale Rete non è da considerarsi solo un supporto per veicolare la diffusione dei prodotti scientifici, ma anche una base per favorire la conduzione di collaborazioni di ricerca ad ampio raggio.

---

<sup>18</sup> Come rilevato da David (2004), a partire dalla Rivoluzione Scientifica, gli scienziati hanno via via riacquisito il regime di segretezza nel lavoro scientifico, a condizione del principio di priorità di attribuzione della scoperta, in ciò esaltando il valore che un approccio collaborativo potrebbe apportare alla scienza nel suo complesso.

Gray ha sostenuto l'emergere del cosiddetto "quarto paradigma" che sostiene la scoperta scientifica fondato appunto sul *data deluge* (diluvio dei dati, *cfr.* Hey *et al.* 2009). L'informatico americano osserva come il primo paradigma scientifico fosse fondato sulla descrizione dei fenomeni naturali attraverso processi di deduzione empirica e un atteggiamento scientifico caratterizzato dalla "passività" del ricercatore rispetto al fenomeno studiato. Il progressivo utilizzo di modelli e generalizzazioni e lo sviluppo dell'attività teoretica hanno rappresentato il passaggio a un secondo paradigma basato sulla ricerca attiva delle spiegazioni ai fenomeni. La combinazione dei due paradigmi "classici", sperimentale e teorico, ha consentito un notevole incremento delle conoscenze negli ultimi quattrocento anni. Il recente avvento della rivoluzione digitale e della scienza computazionale ha tuttavia prodotto un riassetto nel modo di produrre conoscenza, consentendo la simulazione di fenomeni complessi: in tal mondo, la ricerca scientifica, specialmente nel campo delle scienze sperimentali, ha potuto distaccarsi dall'osservazione empirica e proiettarsi sulla riproduzione di un mondo virtuale realizzato in analogia a quello reale. A questa terza fase basata sulla possibilità della simulazione succede, nell'ultimo decennio, la scienza fondata sull'ICT, che rappresenta al contempo una fusione e un superamento dei tre paradigmi precedenti. Si è aperta un'era che nasce dal cambiamento e dallo sviluppo delle tecnologie di informazione e comunicazione, e in particolare del World Wide Web, che hanno generato e soddisfatto le aspettative relative a disponibilità di materiali di ricerca 24 ore al giorno in un'ampia varietà di formati. La nuova scienza si fonda sul "diluvio di dati" consentito dal proliferare degli archivi digitali, e si caratterizza per la riunione di teoria, esperimenti e simulazioni, nel luogo ideale dove le tecnologie informatiche "incontrano" lo scienziato, consentendo la condivisione e l'apertura della produzione scientifica. I nuovi risultati scientifici, profetizzava Gray, deriveranno sempre più dall'immersione nell'oceano sconfinato di dati ottenuti e gestiti grazie a nuovi strumenti tecnologici e risorse digitali per la ricerca. Secondo la visione di Gray, l'interrogazione tramite *software*, la memorizzazione digitale e l'analisi statistica avanzata completeranno il processo ideale che consentirà agli scienziati del nuovo millennio di giungere alla nuova conoscenza. Inoltre, l'aumento del volume informativo andrà a costituire la prima fase verso l'aggregazione di grandi masse di informazioni (*big data*), dando vita a strategie di ricerca che aprono la possibilità a studi su larga scala.

Nielsen (2011, tr. it. 2012), nel presentare i caratteri della scienza del prossimo futuro, ha posto maggiormente l'accento sull'aspetto collaborativo dei processi di produzione della

conoscenza come chiave per comprendere il cambio di direzione, rilevando come per la scienza contemporanea i processi di *problem solving* scientifico si fondino su un nuovo percorso caratterizzato sempre di più su modalità collaborative che sfruttano la varietà cognitiva e l'insieme di informazioni e competenze disponibili su scala globale attraverso la Rete. La Rete di Dati e gli archivi *online* consentono un aumento considerevole delle dimensioni delle collaborazioni, favorendo l'abbandono della tradizionale visione statica della divisione del lavoro scientifico e l'emergere di un modello dinamico caratterizzato dalla comparsa sulla scena di una sorta di "intelligenza collettiva", un *network* potenziato che interviene nella risoluzione delle domande di ricerca<sup>19</sup>. Le modalità fluide di organizzazione si rivelano per Nielsen fondamentali ai fini del superamento degli ostacoli cognitivi e delle *impasse* che possono intervenire nel lavoro individuale. Esse sono attivate dalle nuove possibilità di accesso comune al dato scientifico, e si basano sull'idea che più grande sia un gruppo coinvolto nell'approccio a un problema scientifico, maggiore sia la quantità di competenza latente complessiva disponibile per la risoluzione di domande di ricerca. Premessa ineludibile per l'implementazione effettiva di tale modello di collaborazione, fondato sulla "modularizzazione", è la condivisione di un bagaglio di conoscenze tecniche e di una prassi metodologica condivisa<sup>20</sup>. La collaborazione così delineata diventa autostimolante e si caratterizza per dinamiche di "serendipità pianificata"<sup>21</sup>, che permettono di risolvere più rapidamente molti problemi privi di soluzione, canalizzando le questioni aperte verso soggetti maggiormente qualificati a risolverle. A livello di laboratorio, osserva Nielsen, si tende ad una ridefinizione della relazione tra osservazione e analisi nel lavoro scientifico: mentre tali fasi del disegno della ricerca erano solitamente operate dal medesimo ricercatore o gruppo di ricerca, in alcuni campi scientifici si assiste alla tendenza alla scissione

---

<sup>19</sup> La visione statica si caratterizza per una certa rigidità nei ruoli e lascia inutilizzate molte micro-competenze, stimolando solo a livello del singolo la componente creativa legata alla fase della scoperta. L'autore cita a modello di divisione statica, seppur di successo, il lavoro che nel 1983 portò alla scoperta del bosone Z presso il CERN di Ginevra. Si trattò di un'impresa scientifica in cui tutti gli scienziati coinvolti avevano un ruolo prestabilito, dall'ingegnere addetto al raffreddamento del fascio di particelle agli statistici che elaboravano i risultati finali.

<sup>20</sup> Per Nielsen, la scienza è particolarmente adatta al modello di collaborazione basato sull'intelligenza collettiva. Infatti, la maggior parte dei campi scientifici possiede metodi e strumenti sulla cui applicabilità vi è un sostanziale accordo nella comunità scientifica di riferimento: è il caso della matematica, per cui le regole del discorso matematico non vengono messe in discussione. Più complicato è immaginare tale condivisione di paradigma nell'ambito economico, la cui comunità di ricercatori discute su importanti questioni e non concorda unitariamente sulla metodologia di ricerca.

<sup>21</sup> L'autore mutua il termine da Jon Udell e non ne fornisce una definizione precisa. Ad ogni modo, lo intende come capacità dell'"intelligenza collettiva" di risolvere un problema scientifico in modo da un lato inatteso (significato letterale di "serendipità"), ma dall'altro indotto dallo sfruttamento di una situazione globale di decentralizzazione del sapere e delle competenze (che risiedono in tante menti diverse) che gli strumenti *online* attivano attraverso la creazione di "architetture dell'attenzione" (piattaforme adeguate a catturare l'attenzione degli esperti) dirette verso ogni persona competente nella risoluzione del problema (*cf.* Nielsen, 2011, tr. it. 2012, pp. 34-38).

di competenze. Tale tendenza è testimoniata dall'emersione di nuovi campi disciplinari specifici come la biologia computazionale – che è popolata da studiosi che possiedono sia le competenze del programmatore che la capacità interpretativa del biologo – o l'astroinformatica e le *digital humanities* (Hey *et al.*, 2009). A livello globale, grazie alle risorse *online*, l'apporto di grandi quantità di contributi diversificati e modulari forma una cosiddetta “massa critica conversazionale”, ossia una collaborazione creativa senza confini tale che le idee del singolo inducano altre persone, anche cittadini comuni, a farsi venire altre idee e così via. Perché si compia la “modularizzazione” della collaborazione e l'accesso quanto più ampio alla varietà delle competenze disponibili, occorre per Nielsen abbattere le barriere alla partecipazione, portando avanti l'idea di una scienza fondata sulla condivisione dei dati e delle conoscenze, ancora non completamente accettata nel mondo scientifico (*cfr.* par. 1.3.2).

La *networking science* profetizzata da Nielsen va oltre la metafora del cervello collettivo e inerisce alla istituzionalizzazione di una visione del bene comune dell'informazione, sempre più spesso reso alla portata dei cittadini in termini di accessibilità, in ciò concorrendo a ridefinire il rapporto tra mondo scientifico e resto della società anche in riferimento all'aumento e alla diversificazione del numero degli attori che producono o collaborano alla formazione di prodotti scientifici. Per mezzo della Rete, anche *audience* “esterne” non specializzate dovranno avere la possibilità di accedere facilmente all'informazione scientifica e addirittura potenzialmente intervenire nel processo di ricerca. L'accesso sempre più globale ai dati di ricerca ha già permesso la partecipazione dei cittadini allo sforzo cooperativo per l'avanzamento della conoscenza, al punto che alcuni esperimenti scientifici aperti a un pubblico non scientifico, come Galaxy Zoo e Polymath, hanno costituito casi di successo per una nuova organizzazione del lavoro scientifico che permetta ampi margini di manovra nel tipo di problemi che possono essere affrontati (Franzoni e Sauer mann, 2014).

### *1.3.2 I dati aperti tra promozione e barriere*

L'Open Knowledge Foundation (2014) rileva che i dati di tipo scientifico costituiscono solo una di otto tipologie di possibili dati “aperti” (vi sono anche i dati culturali, finanziari,

statistici, meteorologici, ambientali e di trasporto)<sup>22</sup>. Il dibattito sulla loro apertura, sia in ambito scientifico che nella retorica degli *stakeholder*, è caratterizzato, secondo Pasquetto *et al.* (2015), da otto dimensioni ricorrenti – destinatari; aspetti legali; modalità di disseminazione; standard di accesso; gestione e conservazione; scala di condivisione; riferimenti materiali; benefici connessi. In prima battuta, secondo gli autori, emerge spesso una vaghezza di fondo nell’attività definitoria in quanto è spesso posta un’uguaglianza non sufficientemente specificata con un termine altrettanto ambiguo come *data sharing*. Tale fenomeno è identificato con la generica attività di rilascio e condivisione di dati di ricerca per uso altrui (con questo termine indicando i soggetti che non hanno partecipato alla loro produzione), che lascia imprecisato un riferimento definitivo ai beneficiari dell’apertura, demandando tale specificazione di volta in volta attraverso l’individuazione di determinati requisiti. Essi possono coincidere col cittadino comune fino ad essere identificato con particolari comunità scientifiche o istituzioni di ricerca. Il raggio di potenziale condivisione è variabile rispetto alla prospettiva di “riutilizzo” del dato adottata dai produttori dello stesso e dai fruitori. Essa è prevista dai proprietari dei *database* nella strategia di gestione, conservazione, cura dei *database*, attività queste che si presentano come integrate e inscindibili nella politica di conservazione. Tuttavia, è da segnalare come nella caratterizzazione di “apertura” definita dall’OCSE (2007), all’interno dei *Principi e linee guida per l’accesso ai dati di ricerca*, la platea del dato scientifico “aperto” corrisponda con la comunità scientifica internazionale: “*access on equal terms for the international research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination*”.

Ciò su cui la maggior parte delle definizioni converge è tuttavia l’aspetto della disponibilità dal punto di vista legale e tecnico. Sotto l’aspetto legale, i produttori di dati aperti possono fornire accesso all’esterno preservando la propria proprietà intellettuale e vincolando in vari modi le attività di elaborazione. Le soluzioni più utilizzate si esprimono con l’adozione di una licenza *Open Data Commons*, la quale prevede che l’utilizzo del dato comporti l’obbligo di citazione della fonte, e attraverso l’inclusione di aspetti di *privacy* e restrizioni di carattere etico. Esse sono specificate attraverso una documentazione puntuale

---

<sup>22</sup> Quando si citano nel dibattito scientifico esempi di apertura si fa generalmente riferimento ai dati collezionati attraverso fondi pubblici, gli esempi generalmente riportati riguardano *repository* e archivi (quali GenBank, Protein Data Bank, Sloan Digital Sky Survey), *network* di dati armonizzati (come ad esempio World Data Centers, Global Biodiversity Information Facility; NASA Distributed Active Archive Centers), osservatori virtuali (tra cui International Virtual Observatory Alliance, Digital Earth) e *repository* istituzionali (come PubMedCentral, arXiv).

che descrive anche la metodologia di raccolta, le fonti e gli scopi per i quali sono stati collezionati. L'accesso tecnico ai dati è invece direttamente legato alle modalità di disseminazione, che possono assumere varie forme dallo scambio privato, alla pubblicazione *online* su siti personali o istituzionali, alla sistemazione in archivi o *repository*, alla creazione di strutture relazionali più complesse sotto forma di *database* pubblici. Il contatto diretto con i dati avviene generalmente attraverso soluzioni tecnologiche digitali quali *Digital Object Identifiers*, basato sullo standard del World Wide Web, e *Linked Open Data*, attraverso cui è possibile la generazione di “fogli eseguibili” per estrarre dati e riprodurre analisi. Nonostante gli standard di accesso siano simili più o meno per la maggior parte degli insiemi di dati “aperti”, un sistema di coordinamento di scambio internazionale di dati non è ancora realizzato e le pratiche di condivisione di dati di ricerca variano per formato, modalità di gestione, campi disciplinari.

La pratica di condivisione di dati scientifici produce potenzialmente un doppio valore, sia naturalmente sul fronte dei fruitori esterni dei dati che dal lato dei proprietari degli stessi. Se infatti la conseguenza più ovvia è rappresentata dalla possibilità per un ricercatore di avere accesso a masse d'informazione collezionate da colleghi afferenti allo stesso dominio, altrimenti non disponibili o reperibili a fronte di una grande quantità di lavoro, è parimenti da considerare per coloro che mettono a disposizione le loro raccolte una potenziale apertura a collaborazioni trans-disciplinari talvolta necessarie a rispondere alle grandi sfide che contraddistinguono la ricerca contemporanea (Boulton *et al.*, 2012; Borgman, 2012; Hey *et al.*, 2009). Inoltre, come afferma Nielsen (2011, tr. it. 2012), l'accesso aperto ai dati scientifici consente potenzialmente ai ricercatori di stabilire domande di ricerca che sarebbero difficilmente emerse in assenza dell'accesso stesso e costituisce anche un deterrente alla frode, che incoraggia la pubblicazione di ricerche di alta qualità.

Si ritiene che un primo passo significativo e ufficiale verso una politica scientifica orientata alla condivisione di dati su larga scala sia stato rappresentato da un'intesa collettiva ratificata nel 1996 presso le Isole Bermuda dai più importanti laboratori coinvolti nella ricerca sul genoma umano (*ibidem*). Il cosiddetto “Accordo delle Bermuda” riportava che “*tutte le informazioni relative alla sequenza del genoma umano devono essere disponibili gratuitamente e accessibili al pubblico, in modo da incoraggiare la ricerca e lo sviluppo e da massimizzare i benefici sociali di questa ricerca*”. A seguito di esso, tutti i dati genetici umani sarebbero stati resi accessibili *online* e la pratica di condivisione costituì requisito

essenziale per lavorare sul genoma umano. Nielsen (*ibidem*) riporta che, nel tempo, si sono succeduti vari tentativi di estendere lo spirito dell'accordo ad altre comunità scientifiche. Un *endorsement* dai massimi vertici della politica mondiale al nuovo modo di produrre conoscenza attraverso la collaborazione e la condivisione arrivò il 14 marzo 2000 quando il presidente USA Bill Clinton e il primo ministro britannico Tony Blair rilasciarono una dichiarazione congiunta in cui elogiavano i principi contenuti nell'accordo e invitavano gli scienziati di ogni parte del mondo ad adottarne di analoghi. Anche singole *research funding organisation* o giornali specializzati contribuiscono ad attivare *policy* miranti a favorire la condivisione di dati. Ad esempio, National Science Foundation ha annunciato nel 2010 che tutte le future proposte di finanziamento avrebbero richiesto un piano di due pagine, soggetto a *peer-review*, riguardante la gestione dei dati e rispondente ai requisiti di cui sopra.

Ad ogni modo, nonostante i numerosi stimoli a livello internazionale e la promozione delle agenzie di finanziamento, l'apertura del processo scientifico è diventata un'aspirazione che ha avuto negli anni differenti fortune. Sebbene la tendenza generale assecondi l'enfasi sulla necessità di condivisione dei dati scientifici, molta scienza ad oggi continua a non essere accessibile oltre la comunità di riferimento, le attività di rilascio e l'interesse allo sfruttamento degli stessi appaiono concentrati in pochi campi specifici, e la disponibilità di dati empirici appare ancora troppo frammentata, in ciò costituendo un approccio da "era della stampa" più che da "era digitale" (Boulton *et al.*, 2012). Borgman (2012) ha rilevato come nel complesso poca ricerca sia circolata effettivamente al di fuori dei *team* scientifici che l'hanno prodotta e un numero non elevato di richieste è stato riscontrato per questi dati. Inoltre, pare non essersi ancora instillata del tutto nella cultura di ricerca l'idea che l'apertura del dato sia un bene per la scienza. Al contrario, persiste da parte degli scienziati la tendenza a conservare un vantaggio competitivo sui colleghi. L'autrice parla di un vero e proprio dilemma (*conundrum*) che ruota attorno alla condivisione di dati scientifici, inerente al tema del bilanciamento tra vantaggi e svantaggi dell'apertura all'esterno del lavoro scientifico. Condividendo le loro ricerche, i ricercatori temono di perdere *expertise*, poiché – nonostante la crescente tendenza alle collaborazioni – gli stessi competono per finanziamenti, lavori e pubblicazioni. Inoltre, i dati sono più probabilmente condivisi quando esistono politiche che ricompensano chi li produce. Parallelamente, esistono limiti pratici alla condivisione di dati. Talvolta, i dati non esistono in forme trasferibili per ragioni etiche o epistemologiche, oppure – come evidenzia ancora Borgman (*ibidem*) – il tentativo di sfruttare i dati in contesti diversi da quelli della loro

produzione ha messo in luce che un effettivo riutilizzo è spesso molto difficoltoso per i singoli scienziati, poiché i vari *dataset* possono utilizzare protocolli differenti e sfruttare differenti interfacce, per cui gli scienziati impiegherebbero molto tempo per capire come accedere agli stessi, selezionarli, armonizzarli e in generale, prepararli all'utilizzo voluto.

## CAPITOLO 2

### Le infrastrutture di ricerca nelle scienze sociali

#### 2.1 Infrastruttura di ricerca: un concetto moderno, ma non nuovo

La locuzione *infrastruttura di ricerca* appare con sempre maggiore frequenza sia nelle parole dei ricercatori, quando presentano un lavoro la cui realizzazione è stata consentita o facilitata dall'utilizzo di particolari risorse informative o strumenti di ricerca resi disponibili dal progresso tecnologico, sia nella retorica legata alle politiche della ricerca, quando si sottolineano le nuove potenzialità legate all'innovazione in campo scientifico. Sebbene l'interesse verso queste risorse si sia intensificato solo di recente in conseguenza degli investimenti indirizzati alle nuove *facility* fisiche e digitali per la ricerca in Europa (Ciula *et al.*, 2011)<sup>23</sup>, esse in realtà rappresentano, sotto varie forme, un filo conduttore del modo di produrre conoscenza da più di duemila anni a questa parte.

La prima infrastruttura realizzata al servizio dell'avanzamento della conoscenza è generalmente individuata nel *Mouseion* eretto ad Alessandria d'Egitto nel 307 a.C. su iniziativa di Tolomeo I (*ibidem*). Si trattava di un ampio edificio che ambiva alla raccolta del sapere del tempo, rappresentando al contempo un centro culturale, un'università e una biblioteca. Questo “centro d'informazione” *ante litteram* ha ispirato nel tempo iniziative tese a strutturare la conoscenza in un sistema organizzato e consultabile, che hanno enfatizzato l'importanza della raccolta, della conservazione e della diffusione del sapere<sup>24</sup> e hanno reso la pratica scientifica profondamente ancorata al precetto della conservazione della documentazione (Latour, 1987). Seguendo l'esempio della biblioteca alessandrina, sin dall'inizio della scienza moderna, alcuni peculiari “supporti di ricerca” sono stati creati e percepiti dagli scienziati nella forma di *abilitatori* del processo di produzione di conoscenza. Lauer (2014) nota che Bacone, nel suo *Advancement of Learning* del 1605, menziona due istituzioni, che presentano tali funzioni di ausilio e incentivo: le biblioteche e le collezioni di

---

<sup>23</sup> Cfr. par. 2.4. Già nel 2010, le infrastrutture di ricerca nell'Unione Europea aggregavano investimenti nell'ordine di 100 miliardi di euro, si stima che venissero utilizzate ogni anno da circa 50000 ricercatori all'anno e che producessero un impatto scientifico complessivo misurabile tra i 3000 e i 6000 prodotti di ricerca da esse derivate (Rizzuto, 2016).

<sup>24</sup> Esempi in tal senso vanno dalla pubblicazione delle prime enciclopedie su larga scala alla creazione delle riviste scientifiche in Europa. La diffusione di queste risorse ha spinto allo sviluppo di nuovi strumenti di ricerca, come ad esempio indici, bibliografie, dizionari biografici, che si sono imposti come supporti insostituibili per l'accesso alla conoscenza.

dati. Tali risorse sono state, e vengono ancora oggi, sfruttate in maniera intensiva dai ricercatori di tutto il mondo per la produzione dei loro lavori.

Dalla fine della Seconda Guerra Mondiale, il progresso tecnologico legato alla strumentazione scientifica e all'evoluzione dei supporti informatici ha trasformato la natura dei processi di produzione, sistematizzazione, condivisione e diffusione dei dati di ricerca, connotando il ciclo di alimentazione del sapere in modo inedito (*cf.* par. 1.2). In particolare, la fase di raccolta di dati e informazioni si rivela oggi notevolmente facilitata dal formato digitale che rende più rilevabile, accessibile e maneggiabile il materiale da consultare o da lavorare. Così, agli “abilitatori di ricerca” fisici, che a loro volta si sono raffinati e adattati alle nuove esigenze degli scienziati, si sono aggiunti quelli virtuali o elettronici, formando un'ampia gamma di risorse, che va sotto il nome di *infrastrutture di ricerca*, che costituiscono nel loro insieme, per alcuni osservatori, la vera spina dorsale della nuova scienza (Farago, 2014).

### *2.1.1 Una costellazione di referenti empirici*

È opportuno da subito rilevare che non esiste in letteratura una singola definizione accettata e condivisa di infrastruttura di ricerca (Renschler *et al.*, 2013). Farago (2014) ha osservato che all'interno dei report, saggi e articoli pubblicati sul tema negli ultimi due decenni è stata prodotta una costellazione di referenti relativi a caratteristiche umane, tecniche, operative e organizzative. La ragione di tale “caos definitorio” ha origine nel fatto che lo stesso concetto di *infrastruttura* è di difficile delimitazione. Generalmente, le infrastrutture si connotano quale sub-strato grazie al quale si abilitano e si sviluppano le attività sociali della vita quotidiana e tale dimensione di *background* le rende invisibili e date per scontate (Star, 1991); presentano inoltre una composizione complessa poiché incorporano componenti umane, organizzative e tecnologiche, formando sistemi dinamici e relazionali di tipo socio-tecnico (Star e Ruhleder 1996; Bowker *et al.* 2010). In particolare, nell'ambito dei processi di produzione della conoscenza, le infrastrutture – secondo il filone di studi afferente ai *Science and Technology Studies* – sono immerse nelle pratiche, si evolvono dinamicamente sulla base delle tecnologie e mantengono un insieme di conoscenza specifica in modi complessi e interconnessi (*cf.* Monteiro e Hanseth, 1996; Edwards, 2010). Per questi motivi, sono percepite dai ricercatori in vario modo: per alcuni rappresentano un *quadro tecnico e operativo* che permette loro di collaborare e condividere dati e risultati di ricerca; per altri si

sostanziano nel *contenuto informativo a cui è offerto un accesso*, senza considerare le strutture intorno ad esso; e per altri ancora rappresentano entrambe le cose (Farago, 2014).

A mettere ordine nel panorama definitorio sulle infrastrutture di ricerca è stato il livello politico, a livello nazionale ed europeo, che ha individuato una serie di referenti validi per tutti i campi scientifici, orientando il modo di intendere queste risorse abilitatrici di conoscenza secondo un approccio prettamente normativo. A livello italiano, il Programma Triennale della Ricerca 2010-2012 ha designato con tale termine, riprendendo la definizione riportata nel *Community legal framework for a European Research Infrastructure* (ERIC), “*gli impianti, le risorse e i servizi utilizzati dalla comunità scientifica per compiere ricerche di alto livello nel loro settore*”. La definizione comprende genericamente “*grandi impianti o strumenti di ricerca scientifici; risorse di conoscenze come collezioni, archivi o informazioni scientifiche strutturate; le strutture basate sulle TIC come le reti di tipo GRID, il materiale informatico, il software e le comunicazioni; qualsiasi altro mezzo essenziale per raggiungere l'eccellenza nelle azioni di R&S*”<sup>25</sup>. La definizione di *infrastruttura di ricerca di tipo europeo o internazionale* prevede alcuni requisiti ulteriori tra i quali l'accesso aperto, cioè l'utilizzo da parte di ricercatori di qualsiasi organismo di ricerca o economico selezionati sulla base del merito scientifico di una proposta di ricerca, e la manutenzione periodica. In questo senso, il progetto *Mapping of the European Research Infrastructure Landscape* (MERIL), finanziato nell'ambito del Settimo Programma Quadro europeo e sostenuto dalle organizzazioni afferenti alla European Science Foundation, il cui obiettivo è coinciso con la realizzazione di un portale per la mappatura delle infrastrutture rilevanti in campo europeo, riconosce come infrastrutture di ricerca di tipo europeo quelle *facility* o piattaforme che forniscono la comunità scientifica con risorse e servizi per condurre una ricerca di alto livello nei rispettivi campi [...] che “*a) offrano prestazioni scientifiche e tecnologiche di alta qualità, che dovrebbero essere riconosciute come di rilevanza europea; b) permettano l'accesso ad utenti scientifici provenienti dall'Europa e oltre, attraverso un processo di selezione trasparente*

---

<sup>25</sup> Cfr. Programma Triennale della Ricerca 2010-2012: glossario, disponibile all'indirizzo web <http://www.miur.it/UserFiles/3248.pdf> (ultimo accesso: agosto 2017). Nel *Community legal framework for a European Research Infrastructure* (ERIC) con infrastruttura di ricerca “*si intendono facilities, risorse e servizi connessi utilizzati dalla comunità scientifica per condurre ricerca di alto livello nei rispettivi campi e copre grandi attrezzature scientifiche o gruppi di strumenti; risorse basate sulla conoscenza quali collezioni, archivi o informazioni scientifiche strutturate; infrastrutture ICT come reti di calcolatori (Grid computing), software e tecnologie di comunicazione, o qualsiasi altra entità di natura unica essenziale per raggiungere l'eccellenza nella ricerca*” (traduzione mia, European Commission, 2008, p. 10).

sulla base di eccellenza; c) abbiano una gestione stabile ed efficace" (traduzione mia, European Science Foundation, 2013, pp. 9-10; cfr. anche MERIL project, 2011).

Tutte le concezioni prodotte a livello europeo o nazionale convergono comunque nell'identificare un criterio che combina posizionamento geografico e discrimine tra fisicità e virtualità: le infrastrutture di ricerca sono di tipo localizzato se rappresentano risorse posizionali rispetto ad una singola unità fisica (*single-sited*); di tipo distribuito (*distributed*) se si sostanziano in reti di risorse multicentriche (si veda ad esempio European Commission [2008] 467); sono inoltre di tipo virtuale (*e-infrastructure*) se si riferiscono a servizi forniti per via elettronica.

Combinando le varie definizioni disponibili, Ciula *et al.*, (2011) hanno individuato quattro gruppi principali di referenti designabili col termine infrastrutture di ricerca (2011):

- infrastrutture fisiche, come collezioni di oggetti, installazioni, laboratori e strumenti fisici;
- *e-infrastrutture*, ossia reti informatiche, GRID o *cloud computing* che si sviluppano su varie istituzioni;
- infrastrutture di dati digitali, che comprendono archivi, collezioni e *repository*;
- meta-infrastrutture, che consentono un accesso unico verso infrastrutture indipendenti, con differenti formati e strutture collegate utilizzando metadati.

## 2.2 Calibrazione dei referenti per il campo delle scienze sociali

L'articolazione del concetto di infrastruttura di ricerca presente nelle definizioni "istituzionali" appare caratterizzata da un *alto livello di genericità* poiché fa riferimento a elementi riconducibili al complesso del panorama scientifico. La delimitazione del campo semantico dovrebbe essere invece calibrata rispetto ai singoli domini ed essere sufficientemente estensiva da includere tutte le infrastrutture di ricerca utilizzate dalle comunità di studiosi afferenti ai vari campi disciplinari, e al contempo abbastanza restrittiva nel tenere conto delle speciali dinamiche e degli aspetti che dovrebbero essere considerati per ciascun *field* (Renschler, *et al.*, 2013).

Così, parlando delle *scienze sociali*, non sarà di certo possibile annoverare i grandi impianti, come ad esempio acceleratori di particelle, telescopi, navi di ricerca, satelliti per il telerilevamento e così via. Allo stesso modo, le infrastrutture *hardware* come le infrastrutture GRID per il calcolo distribuito, sfruttate per l'elaborazione di grandi quantità di dati, mediante l'uso di una vasta quantità di risorse, o *mainframe* ad alte prestazioni e alta capacità, hanno utilizzi di gran lunga più orientati rispetto alle esigenze delle scienze naturali. Al contrario, dovranno essere considerate le *research information infrastructure* (German Council of Science and Humanities, 2011), ossia quelle risorse create attraverso l'acquisizione e la raccolta dei dati come le collezioni di informazioni strutturate o le banche dati digitali e i servizi ad esse connessi.

Il progetto MERIL individua, tra i vari referenti di infrastrutture di ricerca, quelle utilizzate nelle *scienze sociali*, specificando che tali risorse non sono di esclusiva pertinenza di tale dominio scientifico, ma possono essere sfruttate, a seconda del loro contenuto, anche da altri domini: si tratta delle collezioni di dati organizzate in *data archives, repositories and collections; national and international statistical facilities and research data service facilities; registers and survey-led studies/databases*<sup>26</sup>. Soffermandoci sull'ultima categoria, si dovrà tenere conto che, a dispetto di una crescente presenza di collezioni di dati istituzionalizzate a livello europeo, incentrate su ampia apertura alla comunità dei ricercatori, aggiornamento e interoperabilità, vi è nelle scienze sociali una persistenza di infrastrutture sviluppate in relativa indipendenza nell'ambito di specifici progetti di ricerca (Dutton e Meyer, 2009; German Council of Science and Humanities, 2011). Per questo motivo, i *database* di ricerca possono essere ulteriormente caratterizzati ad un secondo livello di classificazione attraverso il duplice criterio dell'obiettivo di costituzione e del grado di apertura in *database project-driven*, derivati da progetti specifici, e *infrastrutture di interesse europeo* (ESFRI, 2016), ormai affermate come principali elementi della competitività all'interno dello Spazio Europeo della ricerca, o di *caratura internazionale*.

#### *Data archives, repositories and collections*

Un tipico esempio di oggetti inclusi in questa categoria è rappresentato dagli *archivi istituzionali*, biblioteche e servizi digitali che contengono materiale accademico prodotto e

---

<sup>26</sup> La classificazione è tratta da <https://portal.meril.eu> (ultimo accesso: marzo 2017), ma presentata in forma raggruppata per una migliore presentazione.

gestito all'interno di un'università o ente di ricerca specifico<sup>27</sup>. Già da quasi due decenni, le istituzioni scientifiche, stanno progettando, costituendo e organizzando tali infrastrutture allo scopo di collezionare e classificare in modo sistematico documenti e lavori accademici prodotti dai propri afferenti. Tale attività si connota come un orientamento istituzionale nella gestione delle risorse accademiche intrapreso in funzione dei processi sociali che sottendono al loro sfruttamento (*cf.* Reale *et al.*, 2011). La costituzione di *repository* dal carattere cumulativo, persistente e aperto sta conoscendo uno sviluppo significativo, come testimonia il numero in continua ascesa di università ed enti di ricerca che, nel mondo, hanno aperto archivi istituzionali. Si conta che oggi oltre 1300 archivi attivi contribuiscano ad alimentare i processi di diffusione dei risultati delle attività di ricerca e, conseguentemente, generino un impatto sui nuovi studi in tutti i campi del sapere scientifico. Tra essi possono essere individuati gli archivi c.d. “disciplinari”, come quelli per le discipline umanistiche e sociali, ossia a piattaforme condivise tra più enti che interagiscono nel deposito di materiale di una stessa disciplina o di un argomento specifico (Kennan, 2008).

Nell'ambito dello sforzo di costruzione di infrastrutture condivise e transnazionali a livello europeo, più di 70 archivi istituzionali nazionali, aderiscono al progetto DRIVER, un archivio istituzionale di deposito dei prodotti di ricerca europea, che contiene ogni tipo di documentazione scientifica: articoli, rapporti tecnici, *working paper*, *pre-print*, dati e metadati, pubblicati nella *European Research Area*. Nel contesto italiano, si segnalano ad oggi AMS ACTA, l'archivio istituzionale dell'Università degli Studi di Bologna Alma Mater Studiorum, e PLEIADI (Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali) una piattaforma nazionale sviluppata dal CILEA e dal CASPUR per l'accesso centralizzato alla letteratura scientifica depositata negli archivi gestiti da università ed enti di ricerca. Queste due risorse costituiscono due esempi della tendenza a raccogliere in maniera sistematica i dati della produzione scientifica che si sta comunque diffondendo negli ultimi anni anche sul territorio italiano.

---

<sup>27</sup> La Conferenza dei Rettori delle Università Italiane (CRUI) ha definito l'archivio istituzionale è come: “[...] uno strumento concreto a disposizione della comunità scientifica, un'infrastruttura informativa e comunicativa che raccoglie in un unico luogo tutta la produzione scientifica di un ateneo, alla quale viene conferita autorevolezza, assicurandone nel contempo la persistenza in rete e la conservazione a lungo termine. [...] Gli archivi istituzionali possono essere considerati degli indicatori tangibili della qualità di una istituzione accademica, la sua estensione naturale in quanto volani della ricerca primaria – potenzialmente la componente più importante nell'evoluzione dei nuovi modelli di comunicazione scientifica” (CRUI, 2009, p.6).

### *National and international statistical facilities and research data service facilities*

Nel caso delle agenzie statistiche nazionali (a livello italiano si pensi all'ISTAT), l'iniziativa di raccolta e sistematizzazione delle informazioni scaturisce da una committenza statale, che risponde a finalità politiche di monitoraggio e valutazione. I dati sono messi a disposizione dell'utenza generale attraverso siti *web* ufficiali in forma tabulare, consentendo tuttavia solamente un livello d'analisi aggregato (per classi, tipologie organizzazioni, livelli territoriali, ecc.). L'accessibilità al microdato è regolata da norme e convenzioni che riguardano le modalità di acquisizione del dato stesso e rispondono a specifiche condizioni d'accesso (OCSE, 2013). A livello europeo, l'esigenza di comprendere come le politiche nazionali e quelle d'integrazione operino all'interno dell'Unione Europea, attraverso l'individuazione di eventuali punti di forza e di debolezza, è stata per anni soddisfatta attraverso l'azione dell'agenzia statistica europea, EUROSTAT, che coordina la produzione dei censimenti nazionali e la produzione di dati amministrativi. Sin dagli anni Sessanta, la ex Comunità Europea ha provveduto a sviluppare collezioni di microdati per misurare il progresso sociale e monitorare le dinamiche di integrazione (si pensi all'Indagine Europea sulla Forza Lavoro o all'indagine EU-SILC su Reddito e Condizioni di vita). Assimilabili alle statistiche ufficiali degli istituti statistici sono i *data service* delle organizzazioni nazionali, come Censis, o internazionali, come la World Bank o la World Health Organisation (*ibidem*).

### *Registers and survey-led studies/databases*

Grazie ai consistenti sviluppi delle piattaforme informatiche e dei *software* digitali, i *database* elettronici rappresentano oggi un elemento chiave nella strategia di gestione delle informazioni in tutto il ciclo di una ricerca: assistono la raccolta dei dati fornendo modi personalizzati di imputazione, abilitano l'integrazione con altri insiemi di dati; permettono la gestione di *corpus* eterogenei di materiale quantitativo e qualitativo; forniscono modalità flessibili per la ricerca e l'elaborazione dei dati (*cfr.* Martinez-Uribe e Patrick, 2011). Questo tipo di infrastrutture di ricerca rappresenta una risorsa particolarmente importante per le scienze sociali, ma naturalmente manifesta una sua rilevanza anche per altri domini, come le scienze della vita e ambientali, le scienze fisiche e ingegneria, le scienze mediche e così via. Il potenziale offerto dall'avanzamento delle tecnologie dell'informazione e della comunicazione è stato riconosciuto con grande tempestività dai ricercatori delle *hard science* che hanno sfruttato le enormi possibilità offerte dalle banche dati digitali condivisibili da

remoto. Di contro, le scienze umane e sociali sono state più lente nel cogliere le nuove opportunità legate all'era digitale<sup>28</sup>.

Per quanto riguarda le *hard sciences*, grandi database condivisi sono stati sviluppati allo scopo di raccogliere e disseminare dati in un contesto di comparazione continua di analisi e risultati e di produzione comune di standardizzazioni, nomenclature e classificazioni (*cfr.* lo studio sull'attività di ricerca in ambito biologico e clinico di Leonelli [2012]). Inoltre, lo sviluppo di strutture relazionali per dati e informazioni ha contribuito alla fase d'analisi facilitando la produzione di report e automatizzando procedure di selezione e calcolo. Proprio l'ingresso sulla scena di nuove basi di dati digitali negli ultimi 15 anni, pensate non solo come strutture di sistematizzazione ma anche come strutture di condivisione della conoscenza ha consentito, secondo Hine (2006) la soddisfazione di domande di ricerca cui singoli ricercatori, alle prese con risorse limitate e tempi compressi, avrebbero fatto fatica a rispondere. Uno studio di Bietz e Lee (2009) si è focalizzato sulle implicazioni della produzione e dell'uso di *database* per il lavoro collaborativo nel campo della metagenetica, evidenziando come tali prodotti siano descrivibili quali *boundary negotiated artifact* (secondo la concettualizzazione di Lee, 2007), ossia infrastrutture negoziali costruite su collaborazioni già esistenti che legano diverse comunità di pratiche, raccogliendo, organizzando, condividendo idee e travalicando gli abituali confini di laboratorio attraverso la comune soddisfazione di bisogni informativi. Dagiral e Peerbaye (2016) hanno mostrato come un *database* digitale, originariamente progettato per condividere le informazioni può rappresentare un oggetto stabile e rappresentativo per alcuni ricercatori, ma anche problematico e potenzialmente utile per ulteriori rinegoziazioni per altri. Sul lato della coproduzione della conoscenza, i *database* stanno portando a mutamenti nella pratica di ricerca, specie in riferimento all'ambiente spaziale dove essa prende forma. Binz-Schaft *et al.* (2014), analizzando il campo della biologia molecolare e cellulare, hanno rilevato come l'introduzione di basi di dati condivise non abbia tuttavia spezzato gli aspetti sociali della generazione di conoscenza scientifica, tra i quali la forza dei legami di attaccamenti generati da interazioni *face-to-face*: tale aspetto, insieme alla prossimità spaziale, continua, secondo gli autori, ad essere predittivo dei *pattern* di collaborazione. Sulla stessa scia, Hine (2006) ha sostenuto che la disponibilità di nuovi

---

<sup>28</sup> Sebbene già nel 1949 sia possibile individuare una data simbolica per la prima idea infrastruttura umanistica digitale, ossia l'*Index Thomisticus* immaginato e realizzato nei successivi 30 anni da Roberto Busa, consistente nella raccolta delle opere complete di San Tommaso (Lauer, 2014).

archivi informativi condivisi ha mostrato finora effetti contraddittori: se da una parte ha stimolato l'ingresso di nuovi partecipanti alle attività di ricerca su specifiche tematiche, dall'altra ha contribuito a supportare ulteriori collaborazioni tra ricercatori che avevano già stabilito legami preferenziali.

Passando al campo delle scienze non dure, non sono molti gli studi che hanno trattato il tema della condivisione dei *database* di ricerca (tra le poche ricerche *cfr.* lo studio sulla condivisione di basi di dati da parte di una comunità di storici presentato in Antonijevic *et al.*, 2013). Per quanto riguarda le dinamiche collaborative legate alla condivisione di basi di dati in particolare nelle scienze sociali, Barkuus e Brown (2012), hanno ribadito l'importanza delle collaborazioni a piccola scala e lunga durata per il lavoro di ricerca, in contrasto con le collaborazioni a larga scala e vita breve che, secondo gli autori, sono il *target* delle infrastrutture digitali condivise, in ciò contrastando la previsione di Atkins *et al.* (2010), secondo cui la ricerca sociale avrebbe intrapreso un sempre più crescente percorso di trasformazione verso strutture globali di collaborazione col concorso di risorse computerizzate a larghissima scala.

### **2.3 Le infrastrutture derivate da progetti specifici (i *project-driven database*)**

Nell'ambito delle scienze sociali, è molto comune la creazione di collezioni di dati, di tipo più o meno focalizzato, create su misura per rispondere a domande di ricerca specifiche (*project-driven*) nell'ambito di un progetto di ricerca e destinati, almeno inizialmente, ad uso dei creatori primari. Dutton e Meyer (2009) hanno rilevato che vi è una proliferazione di *database* isolati nelle scienze sociali e che molti scienziati sociali hanno costruito i propri *set* di dati spesso in modi idiosincratici, per soddisfare le loro particolari esigenze e poiché nessun altro strumento e *set* di dati disponibile è ritenuto adeguato per soddisfare queste esigenze. Nel panorama nazionale e internazionale si riscontra una considerevole frammentazione, poiché le infrastrutture sorgono isolatamente in risposta alla domanda locale e non coordinate con altre basate su tematiche analoghe. I ricercatori realizzano collezioni di dati empirici per obiettivi di ricerca precisi, sulla spinta di finanziamenti pubblici provenienti dagli enti di pertinenza o da fonti esterne. In questo caso, l'accesso all'infrastruttura può inizialmente essere limitato al gruppo di ricerca coinvolto nello specifico progetto. La disponibilità all'esterno del dato può dipendere da specifici accordi e condizioni eventualmente stipulati

con le fonti del dato stesso, tuttavia è più probabile che ne venga promossa la diffusione attraverso meccanismi come l'obbligo imposto dalle agenzie di finanziamento nazionali di assicurare la condivisione dei dati (OCSE, 2013). L'orientamento verso l'apertura di un'infrastruttura di ricerca e lo scopo della stessa non possono essere tuttavia stabiliti a priori e una volta per tutte. Un'infrastruttura può anche essere stabilita nell'ambito di uno speciale progetto di ricerca, con relativamente pochi utenti, e solo successivamente espandere in un'infrastruttura aperta a una comunità più ampia (German Council of Science and Humanities, 2011). Ciononostante, la condizione di esclusività dei possessori del dato può generare una ritrosia da parte dei singoli e dei gruppi a disseminare alla comunità scientifica l'informazione sulla disponibilità dei microdati, anche laddove essi possano essere resi disponibili. Queste infrastrutture *project-driven* offrono infatti ai primi utilizzatori opportunità di guadagnare in reputazione attraverso il trasferimento dei primi risultati dei progetti nelle proprie pubblicazioni (OCSE, 2013).

Oltre che rispetto alle caratteristiche d'accesso, tali collezioni possono differenziarsi rispetto alla possibilità di mantenimento: la loro creazione e il loro sfruttamento può essere limitato nel tempo o si può prevedere un loro aggiornamento più o meno permanente. Nel momento in cui le infrastrutture di ricerca non riescano più a soddisfare le aspettative ad esse connesse o quando perdono la loro connessione rispetto allo sviluppo di metodi di ricerca, il loro sfruttamento viene terminato nell'ambito del contesto di creazione. Tuttavia, molte infrastrutture sorte all'interno di progetti di ricerca limitati nel tempo, si sono sviluppate nel corso degli anni in istituzioni più stabili, non solo per soddisfare le esigenze di ricerca esistenti al momento della loro creazione, ma anche per crearne di nuove, incorporando sempre di più altri gruppi di utenti attraverso innovazioni metodologiche e l'incorporazione di nuovi argomenti di ricerca (German Council of Science and Humanities, 2011).

### 2.3.1 Sfide per il futuro: accessibilità, manutenzione, stabilità

Ciò che viene da più parti auspicato è che i dati contenuti nelle infrastrutture *project-driven* possano essere sempre più aperti, condivisi e armonizzati consentendo alle scienze sociali di abbandonare il modello fondato sui *database* isolati e di abbracciarne uno incentrato su infrastrutture che contribuiscano alla fornitura scientifica di base (*basic scientific supply*), offrendo le stesse condizioni di partenza ad ogni membro di una comunità scientifica in termini di accesso a dati e informazioni, senza restrizioni, e il cui scopo è quello di consentire

un miglioramento delle condizioni generali di lavoro nella ricerca e per la didattica<sup>29</sup>. Seguendo Renschler *et al.* (2013) e Farago (2014), le infrastrutture di ricerca per le scienze sociali dovrebbero rappresentare “*istituzioni durevoli, strumenti tecnici e piattaforme, e/o servizi che sono stabiliti per supportare e migliorare la ricerca come risorsa “bene pubblico” per la comunità delle scienze sociali*”<sup>30</sup>. Da questa definizione, l’autore individua cinque caratteristiche chiave desiderabili per fronteggiare il futuro nella ricerca sociale: i) fornire risorse per la ricerca sotto forma di bene non competitivo e disponibile a tutti; ii) essere orientate all’utente, vale a dire corrispondere ai bisogni dei ricercatori, sia per quanto riguarda i dati che conservano, sia per l’*expertise* metodologica che ha contribuito alla loro realizzazione (il trasferimento della quale può essere attivato attraverso attività di *training* legate alle stesse infrastrutture); iii) essere durevoli a lungo termine attraverso l’ancoraggio a politiche pubbliche e a risorse continue e stabili; iv) garantire adattabilità alle esigenze della comunità scientifica guadagnando e mantenendo il supporto degli *stakeholder*; v) essere legate ai requisiti del metodo scientifico, offrendo accesso aperto e trasparente ai dati e potenziando le opportunità per testare ipotesi e replicazioni attraverso la creazione di standard e la codificazione di pratiche e strumenti. Un accesso facile e aperto ai dati sulle scienze sociali, dovunque siano locati, permetterebbe ai ricercatori di scambiare esperienze e seguire agende comuni, promuovendo progetti transnazionali e generando *partnership* sia a livello di istituzioni che di team di ricerca.

La presenza di infrastrutture così connotate eserciterebbe, secondo alcuni osservatori, un profondo effetto sul modo in cui la ricerca nelle scienze sociali è organizzata e condotta a livello nazionale e internazionale offrendo immediatamente due vantaggi visibili: efficienza nei costi – dal momento che i dati vengono prodotti una sola volta e riutilizzati da molti ricercatori – e promozione di una competizione bilanciata e proponendo allo stesso tempo diverse sfide, dalla tensione tra accesso ai dati e confidenzialità, alla capacità di manutenzione, al superamento della frammentazione e al finanziamento (Ciula *et al.*, 2011). *Survey* coordinate permetterebbero una maggiore diffusione di studi comparativi e, se

---

<sup>29</sup> Il cambio di obiettivo e l’ampliamento del grado di apertura tra database *research-driven* e *basic scientific supply* ricalca la classificazione fluida applicata dal German Council of Science and Humanities (2011, p. 21) per qualificare le infrastrutture di ricerca nelle scienze umane e sociali.

<sup>30</sup> Tale definizione è da annoverare tra i risultati della conferenza *Facing the Future: European Research Infrastructure for Humanities and Social Sciences*, svoltasi a Berlino il 21-22 novembre 2013, promossa dal Social and Cultural Innovation Strategy Working Group of ESFR (SCI-SWG) e dal German Federal Ministry of Education and Research (BMBF).

associate a programmi multi-tematici di lunga durata, incoraggerebbero l'interdisciplinarietà conducendo verso avanzamenti e convergenze in termini di metodologie e procedure. L'accesso a crescenti volumi di dati aggiornati e validati, declinati su una moltitudine di settori tematici, stimolerebbe al dialogo interno alle comunità di ricerca, rafforzando i processi di *network building*. Oltre a rafforzare i *network* collaborativi, questo avrebbe l'effetto di creare tecniche, vocabolari e metodologie comuni. Queste infrastrutture possono infine contribuire all'innovazione metodologica e ad avanzamenti rispetto a come i dati vengono usati e raccolti e giocano un importante ruolo nella disseminazione di *know-how* tecnici (Farago, 2014).

L'unico studio empirico su vasta scala delle ricadute sulle comunità scientifiche dello sfruttamento di infrastrutture di ricerca *aperte e stabili* consiste in un report per la Commissione Europea prodotto da Hüsing *et al.* (2010) con l'obiettivo di descrivere gli sviluppi organizzativi, collaborativi e tecnologici connessi a un campione di infrastrutture di ricerca<sup>31</sup> afferenti a *tutti i settori disciplinari*, comprese le scienze sociali. La ricerca, che risale al 2010, è stata basata su un questionario somministrato agli utenti delle infrastrutture e ai gestori delle stesse, e ha rilevato un ampio accordo sull'impatto sostanzialmente positivo di tali risorse sul lavoro scientifico, evidenziando però effetti meno sostanziali per il campo delle scienze sociali<sup>32</sup>. I principali vantaggi segnalati dagli utenti delle infrastrutture hanno riguardato la velocità di ricerca dei dati e benefici concreti sul lavoro di analisi degli stessi; altrettanto sottolineata è stata la capacità di lavorare su nuovi problemi che i ricercatori non avrebbero potuto affrontare con le risorse disponibili in precedenza. Sono emersi con una frequenza lievemente minore effetti positivi sulla produttività, sul fronte dei costi e sulla qualità degli *output*. L'importanza attribuita allo sfruttamento di infrastrutture di ricerca è risultata essere variabile rispetto al campo disciplinare e ha ricalcato le diverse intensità e modi di utilizzo delle infrastrutture: in questo senso, gli astronomi e gli scienziati della terra sono risultati essere i più influenzati positivamente; gli scienziati sociali e i fisici hanno invece manifestato reazioni un po' più tiepide. Il report ha analizzato, tra le altre cose, anche l'effetto delle infrastrutture nei processi collaborativi. Gli intervistati, in tre quarti dei casi, hanno

---

<sup>31</sup> Caratterizzandole in questo caso come *e-infrastructure*, ma includendo in questa categoria anche archivi e collezioni di dati, a ulteriore testimonianza della confusione definitoria.

<sup>32</sup> Per le scienze sociali sono stati intervistati gli utenti delle infrastrutture CLARIN (Common Language Resources and Technology Infrastructure), creata per sostenere la condivisione, l'uso e la sostenibilità dei dati linguistici e degli strumenti per la ricerca nelle scienze umane e sociali DARIAH (Digital Research Infrastructure for the Arts and Humanities) che mette a disposizione testi, ricerche e dati sul patrimonio culturale.

manifestato accordo con le affermazioni “collaboro di più” e “collaboro più ampiamente in senso geografico”. Un minore grado di accordo ha invece ricevuto l'affermazione di una maggiore collaborazione con colleghi di altri settori della scienza secondo l'etichetta di "collaborazione interdisciplinare". Sono emerse, inoltre, differenze tra campi accademici riguardo la concretizzazione delle collaborazioni, ma gli autori non le hanno ritenute significative a causa del basso numero di risposte ricevute per alcuni ambiti disciplinari. In generale, una minore influenza sulla collaborazione si è riscontrata nei soggetti che hanno qualificato il proprio campo come “un campo in cui la collaborazione è ancora la modalità dominante del lavoro, ma meno che in altri campi”, e tra questi un buon numero di soggetti afferisce alle scienze umane e sociali; al contrario si rilevava una maggiore influenza sui soggetti per cui nel proprio campo “la collaborazione è considerata essenziale per il raggiungimento dei progressi in questi settori e il lavoro è più spesso basato su collaborazioni su larga scala”, e tra questi il maggior numero di soggetti afferisce a “Scienze della Terra e altre scienze naturali”, Fisica, Biologia, Medicina (Hüsing *et al.*, 2010). I risultati della ricerca, con riferimento all'ambito delle scienze sociali sono tuttavia da prendere con cautela, a causa di una scarsa rappresentatività nel campione di studiosi afferenti a questo campo disciplinare (solo 13, equivalenti al 3,4% degli intervistati) e del fatto che le infrastrutture selezionate per le scienze sociali fossero fondate su temi molto specifici (*cf.* nota 32).

## **2.4 Infrastrutture di ricerca stabili e aperte per le scienze sociali**

La creazione di infrastrutture di ricerca per le scienze sociali disponibili su vasta scala, coordinate e aggiornate, come quelle che stanno sorgendo in ambito europeo e internazionale, sta preparando il campo a un cambiamento delle pratiche di ricerca nella direzione di modelli di produzione di dati su larga scala prodotti e condivisi in maniera comune. In particolare, una disponibilità di nuove infrastrutture pan-europee, caratterizzata da un'accessibilità potenziata a masse di dati su tematiche variegata, promette di massimizzare il potenziale per la ricerca insito nelle nuove tecnologie digitali, supportandola e abilitandola verso le grandi sfide globali (*grand challenge*) e di generare hanno un crescente impatto sul piano sociale ed economico, favorendo i metodi di integrazione delle conoscenze e di organizzazione del lavoro tecnico-scientifico (ESFRI, 2016).

In questo paragrafo saranno brevemente presentati alcuni esempi di infrastrutture che stanno intervenendo o promettono di intervenire su tali dinamiche.

#### 2.4.1 *Le infrastrutture di interesse europeo*

Il tema delle infrastrutture di ricerca ha assunto da qualche anno una particolare rilevanza a livello delle istituzioni nazionali ed europee, sia con riferimento all'azione delle agenzie di finanziamento che alle istanze dei decisori politici. Come riportato da Lossau (2012), dall'inizio degli anni Duemila si contano centinaia di finanziamenti a progetti legati a costruzione, estensione, apertura o aggiornamento di infrastrutture di ricerca. Già nel 2001 era emerso, nell'ambito del Consiglio dei Ministri europei, l'interesse alla creazione di un approccio europeo sulle infrastrutture di ricerca nella *European Research Area*, sostanziato in un invito alla Commissione e agli Stati membri dell'UE all'instaurazione di nuovi accordi per sostenere le politiche relative su questo tema: l'obiettivo di fondo era appunto di istituire un sistema integrato di infrastrutture di interesse pan-europeo che fossero in grado di fornire collezioni di dati su larga scala, coordinate, armonizzate e interdisciplinari e che rendessero possibili analisi su molteplici tematiche.

La nascita dell'*European Strategy Forum on Research Infrastructures* (ESFRI) è stata molto significativa in questo contesto, poiché ha contribuito al quadro di strutturazione dello Spazio Europeo della Ricerca, e accresciuto tra i ricercatori la consapevolezza che vi è un'interdipendenza tra di loro sia in riferimento ad una concezione unitaria e collaborativa della scienza sia in riferimento alle risorse e alla condivisione dei finanziamenti (NORFACE, 2009). ESFRI è stato fondato nel 2002 e supporta un approccio strategico mirato alla definizione delle politiche in materia di infrastrutture di ricerca in Europa relativamente a tutte le discipline scientifiche, facilitando iniziative multilaterali che consentano un loro migliore utilizzo e siano tese al loro sviluppo. Il Forum comprende nel suo *board* delegati nominati dai ministeri titolari del settore Ricerca dei Paesi membri e associati. La maggior parte degli Stati Membri ha elaborato la propria *roadmap* nazionale allo scopo di individuare necessità e strategie di valorizzazione delle proprie infrastrutture di dimensione nazionale aprendole a schemi di internazionalizzazione e individuando le priorità di partecipazione alle

infrastrutture selezionate dalla *Roadmap ESFRI*<sup>33</sup>. ESFRI individua le infrastrutture di ricerca di interesse pan-europeo relative a tutte le aree scientifiche e include *progetti* e *landmark*. I *progetti* sono infrastrutture selezionate sulla base dell'eccellenza scientifica e della maturità e sono inclusi nella *roadmap* ESFRI, al fine di sottolineare la loro importanza strategica per il sistema di infrastrutture di ricerca europee e sostenere la loro attuazione tempestiva (ESFRI, 2016). I *landmark* sono infrastrutture già sviluppate e ormai affermate come principali elementi della competitività all'interno dello Spazio Europeo della ricerca. Esse "hanno bisogno di un supporto continuo per il completamento, il funzionamento e l'aggiornamento in linea con la gestione ottimale e il massimo ritorno sugli investimenti" (*ibidem*). Tra i *landmark* di ESFRI vi sono quelle infrastrutture che fanno capo allo *European Research Infrastructure Consortium* (ERIC), un quadro legale per favorire la costruzione, garantire l'operabilità e assicurare la sostenibilità a lungo termine per le infrastrutture di ricerca nuove o esistenti (esempi in questo senso sono la ESS - *European Social Survey* e la SHARE - *Survey of Health, Ageing and Retirement in Europe*).

Tra i *landmark* ESFRI per la tematica *Social and Cultural Innovation* spicca il *Consortium of European Social Science Data Archives* (CESSDA), un'entità legale permanente che consorzia un *network* archivi di dati distribuiti in tutta Europa con l'obiettivo di promuovere i risultati della ricerca relativa alle scienze sociali e di sostenere la cooperazione nazionale e internazionale in ambito scientifico. I suoi servizi propongono l'accesso a numerose infrastrutture di livello europeo che presentano *cross-sectional survey* come quelle facenti capo all'Eurobarometro della Commissione Europea<sup>34</sup>, la *European Social Survey* (ESS-ERIC)<sup>35</sup>, la *Survey of Health, Ageing and Retirement in Europe* (SHARE-ERIC)<sup>36</sup>, la *European Values Survey* (EVS)<sup>37</sup> e la *World Values Survey* (WVS)<sup>38</sup>.

- L'Eurobarometro è stato istituito nel 1974 per fornire dati a livello europeo sulle tendenze sociali, i valori e per testare l'opinione pubblica su varie tematiche riguardanti, tra le altre cose, la situazione politica, l'economia e le istituzioni

---

<sup>33</sup> Ne è un esempio la *Roadmap Italiana delle Infrastrutture di Ricerca di interesse Pan-Europeo*. Altri Paesi, come la Germania, si sono mossi verso l'individuazione di *best practice* come le *Recommendation* prodotte dal German Scientific Council (*Wissenschaftsrat*).

<sup>34</sup> <http://ec.europa.eu/COMMFrontOffice/PublicOpinion/>

<sup>35</sup> <http://www.europeansocialsurvey.org/>

<sup>36</sup> <http://www.share-project.org/>

<sup>37</sup> <http://www.europeanvaluesstudy.eu>

<sup>38</sup> <http://www.worldvaluessurvey.org/wvs.jsp>

europee. Ai dati quantitativi si affiancano studi qualitativi che investigano in profondità su motivazioni, percezioni e reazioni di determinati gruppi sociali su determinate materie. L'obiettivo, sin dalla sua fondazione, è quello di fornire materiale su cui tarare le nuove politiche europee e valutare quelle esistenti. Le *survey* di Eurobarometro prevedono circa un migliaio di interviste *face-to-face* svolte in ogni Paese.

- La ESS è un'indagine a direzione accademica, svolta ogni due anni dal 2001, progettata per fornire informazioni su atteggiamenti, credenze e comportamenti della popolazione europea. Questa risorsa si segnala per l'attenzione ai problemi metodologici e relativi alla comparabilità dei dati nazionali. Nelle edizioni più recenti ha approfondito temi come l'immigrazione e le disuguaglianze rispetto al diritto alla salute; nelle prossime, una particolare attenzione verrà posta sulle grandi sfide della modernità come il cambiamento climatico e il futuro del *welfare state*.
- SHARE è un *database* multidisciplinare contenente microdati su salute, status socio-economico, sociale e familiare della popolazione *over 50* in 20 Paesi europei più Israele. Il suo scopo primario è quello di documentare e interpretare gli effetti dell'innalzamento della durata di vita per gli individui e per la società europea.
- EVS e WVS sono indagini progettate al fine di consentire una comparazione "cross-nazionale" e "cross-culturale" focalizzata su valori e norme su una vasta gamma di argomenti e per monitorare i cambiamenti nei valori e negli atteggiamenti in tutto il mondo. Gli argomenti trattati comprendono la percezione della vita, la famiglia, il lavoro, i valori tradizionali, la disponibilità finanziaria personale, la religione e la morale, l'economia, la politica e la società, l'ambiente, allocazione delle risorse, le questioni sociali contemporanee.

Le infrastrutture afferenti a CESSDA hanno permesso modalità di ricerca di tipo comparativo a livello europeo e fornito l'input per un gran numero di pubblicazioni scientifiche. Dal 2003 al 2014 si contano 2705 pubblicazioni tra articoli su rivista, libri o capitoli di libro e atti di conferenze basate su dati ESS, con un aumento costante del numero di pubblicazioni annuali (dalle 210 del 2008, alle 313 del 2010 alle 416 del 2014) (Fitzgerald, 2016). Il medesimo *trend* in ascesa è rilevabile nel numero di pubblicazioni basate sui dati SHARE: dalle circa 400 del 2009, alle quasi 800 del 2012 a più di 1200 nel 2015 (ESFRI,

2016). In generale, le infrastrutture di ricerca di interesse europeo per le scienze sociali stanno dimostrando di disseminare nuova conoscenza verso ampie platee di ricercatori, consentendo lo sviluppo di nuove abilità e istruendo giovani leve all'uso dei dati (*ibidem*).

#### 2.4.2 Infrastrutture finanziate dai programmi quadro europei – il caso RISIS

La Commissione Europea si è posta quale organo finanziatore di infrastrutture di ricerca, integrando gli sforzi messi in atto dai Paesi membri, attraverso lo stanziamento di ben 1715 milioni nell'ambito del Settimo Programma Quadro europeo 2007-2014 (Lossau, 2012)<sup>39</sup>. La Commissione definisce queste risorse quali centri strategici di eccellenza per la ricerca che facilitano la collaborazione interdisciplinare e istituzionale e la formazione dei ricercatori, nonché quali abilitatori di *partnership* pubblico-privato nel settore della ricerca. Anche nel programma di Horizon2020 le infrastrutture della ricerca sono state inserite nella priorità tematica 1 (*Excellence science*) che promuove, tra le altre cose, lo sviluppo di nuove infrastrutture di ricerca di livello mondiale, l'ottimizzazione dell'uso delle strutture nazionali attraverso integrazione, interoperabilità e apertura a tutti i ricercatori europei.

Tra le infrastrutture finanziate nell'ambito del Settimo Programma Quadro europeo<sup>40</sup>, catalogate dalla Commissione Europea nel campo delle *Social Sciences and Humanities* (European Commission, 2015), è interessante prendere ad esempio RISIS – *Research infrastructures for research and innovation policy studies*, finanziata con circa 5 milioni di Euro in quattro anni<sup>41</sup> e sviluppata da una *partnership* di scienziati che fanno capo agli studi su Scienza e Innovazione<sup>42</sup>. Tale comunità ha recentemente fatto leva sullo sfruttamento comune di infrastrutture informative digitali, esplicitando il valore aggiunto derivato dall'utilizzo di basi di dati condivise che supportino in particolare nuove domande relative alle *policy* legate a ricerca e innovazione (RISIS project, 2013). Il nuovo approccio, ideato al fine fronteggiare le nuove sfide cognitive, si lega al passaggio dalla prima generazione di

---

<sup>39</sup> Ossia il 42% del finanziamento destinato al programma specifico “Capacities” (*cf.* anche [https://ec.europa.eu/research/fp7/index\\_en.cfm?pg=capacities](https://ec.europa.eu/research/fp7/index_en.cfm?pg=capacities), ultimo accesso: agosto 2017).

<sup>40</sup> Asse Infrastructures, nel programma specifico “Capacities”.

<sup>41</sup> *Cfr.* <http://risis.eu>

<sup>42</sup> La comunità epistemica di Scienza e Innovazione produce studi che mirano a comprendere le dinamiche della conoscenza, studiano i processi di innovazione e le condizioni che stimolano lo sviluppo scientifico, sottolineando l'importanza delle interazioni tra gli attori (produttori e utenti della ricerca), le dinamiche di collaborazione e competizione e i processi di apprendimento istituzionali (RISIS Project, 2013).

indicatori di tipo *input/output* (basate sulla misurazione delle risorse investite nel settore R&I) alla seconda generazione di indicatori posizionali che catturano i processi collocati tra *input* e *output* facendo emergere la complessità delle strategie degli attori coinvolti nelle dinamiche che coinvolgono scienza e innovazione (Lepori *et al.*, 2008). Laddove i nuovi tipi di indicatori non erano rappresentati nei *database* statistici o bibliometrici a fruizione pubblica, un crescente numero di nuove basi di dati (alcune delle quali supportate dal finanziamento della Commissione Europea) si è sviluppato non organicamente a livello locale attorno a dimensioni centrali delle dinamiche della conoscenza. RISIS si è posto l'obiettivo di integrare in forme accessibili tali *database* proprietari, formando un'infrastruttura armonizzata, stabilizzando e mantenendo queste risorse, rendendole inoltre accessibili ai ricercatori europei. Il progetto incorpora ad oggi 13 basi di dati, di cui alcune già esistenti prima dell'inizio del progetto e altre sorte durante la fase di implementazione dello stesso. I temi di interesse sono multidisciplinari: i) le dinamiche del settore della ricerca pubblica; ii) i percorsi di carriera di coloro che hanno conseguito il titolo di dottore di ricerca; iii) l'innovazione nelle imprese, iv) l'integrazione europea e v) le politiche su ricerca e innovazione. Gli obiettivi finali di RISIS consistono, tra gli altri, nel lavoro di armonizzazione dei *database* sulla base di alcune variabili chiave per assicurare integrazione e complementarità nell'affrontare domande di ricerca interdisciplinari e lo sviluppo di un importante numero di progetti di ricerca sviluppati per effetto dello sfruttamento di uno o più *database* dell'infrastruttura distribuita. Le singole infrastrutture vengono aggiornate periodicamente e il progetto finanzia corsi per introdurre al loro utilizzo. L'accesso è garantito attraverso visite *on-site* o via Internet e si basa su procedure di selezione e accreditamento basati sul merito scientifico di progetti di ricerca presentati al *board* del progetto.

#### 2.4.3 Esempi di infrastrutture di caratura internazionale: OCSE-PISA e OCSE-PIAAC

*Survey* condotte su larga scala possono essere promosse da organizzazioni internazionali e rese accessibili alle comunità di ricercatori attraverso la pubblicazione di microdati su infrastrutture di ricerca stabili e aggiornate periodicamente. In questi casi, i dati vengono sin da principio predisposti alla condivisione e al riutilizzo.

Rappresentativo di questo fenomeno è lo sforzo di raccolta di organizzazioni quali l'OCSE (Organizzazione per la cooperazione e lo sviluppo economico), che mira alla

formazione di collezioni di dati comparabili e rispondenti a specifici standard che concorrono al monitoraggio delle politiche nazionali. Già da alcuni anni proprio OCSE ha realizzato, in collaborazione con i Paesi membri, alcune indagini comparate internazionali come quelle su competenze ed apprendimento di giovani e adulti, al fine di mostrare il panorama in mutamento nei sistemi educativi e formativi. Le due *survey*, OCSE-PISA e OCSE-PIAAC, costituiscono due esempi di infrastrutture di caratura internazionale accessibili dai ricercatori:

- L'indagine PISA (*Programme for International Student Assessment*)<sup>43</sup> è implementata al fine di valutare la preparazione degli studenti ad affrontare la vita adulta, rilevando le competenze in matematica, scienze, lettura e in ambito finanziario e raccogliendo informazioni di contesto sulle pratiche educative nei paesi partecipanti. Si tratta di un programma di ricerca che fornisce dati rilevanti per le politiche e le pratiche dell'istruzione e permette di monitorare nel tempo i risultati dei processi di acquisizione di conoscenze e abilità attraverso la comparazione di contesti nazionali diversi e in contesti demografici differenziati all'interno dello stesso Paese.
- Complementare all'indagine PISA, è la PIAAC (*Programme for the International Assessment of Adult Competencies*)<sup>44</sup> che ha lo scopo di conoscere abilità fondamentali della popolazione adulta compresa tra i 16 e i 65 anni, in particolare sulla lettura, sulle abilità logico-matematiche e sulle competenze collegate alle tecnologie dell'informazione e comunicazione. L'indagine viene ripetuta nel tempo, per avere informazioni e dati aggiornati e informa su come le persone fanno uso delle competenze non solo nella loro vita personale, ma anche durante la loro attività lavorativa. I dati PIAAC assumono rilevanza sotto vari aspetti, sia per migliorare i percorsi di istruzione e formazione sia per capire come accrescere e arricchire le competenze degli adulti.

---

<sup>43</sup> <http://www.oecd.org/pisa/>

<sup>44</sup> <http://www.oecd.org/skills/piaac/>

## CAPITOLO 3

### Approccio metodologico e strategia di ricerca

#### 3.1 Lo studio di caso e l'obiettivo esplorativo

Il sempre maggiore ricorso, da parte dei ricercatori, a tecnologie informatiche e risorse digitali (*cfr.* par. 1.2) suggerisce implicazioni dirette nell'attività di ricerca e induce a interrogarsi su quale sia il suo effetto rispetto ai processi che generano la produzione di conoscenza. Diversi osservatori hanno rilevato un mutamento nella pratica scientifica nella direzione di modalità collaborative, aperte e basate sul dato, sostenuto dall'aumento del volume di dati presente in Rete e dalla creazione di piattaforme di ricerca ad accesso libero, o quantomeno condiviso tra più ricercatori (*cfr.* par. 1.3). In questo contesto, lo sfruttamento delle *infrastrutture di ricerca* nell'ambito di comunità scientifiche specifiche emerge come uno dei fattori di cambiamento più interessanti da approfondire.

Un'indagine, riferita a tutti i campi disciplinari, sui cambiamenti che intervengono su varie dimensioni del lavoro dei ricercatori per effetto dell'utilizzo di infrastrutture di ricerca è stata implementata ad oggi solo in uno studio un po' datato, voluto dalla Commissione Europea (*cfr.* Hüsing *et al.*, 2010). Tale studio, i cui risultati sono stati riassunti nel par. 2.3.1, soffriva di limiti di rappresentatività relativamente agli utenti delle infrastrutture proprio nel campo delle scienze sociali. Inoltre, mentre nel campo delle scienze dure (*hard sciences*) esiste un discreto numero di studi sui meccanismi che regolano la coproduzione di conoscenza e sui cambiamenti che intervengono nelle pratiche laboratoriali per effetto della condivisione di infrastrutture di ricerca (*cfr.* Hine, 2006; Bietz and Lee, 2009; Leonelli, 2012; Binz-Scharf *et al.*, 2014), l'indagine su come la disponibilità di basi di dati e l'irrompere di tecnologie virtuali abbiano modificato i parametri su cui si fondano l'organizzazione del lavoro e le modalità di aggregazione degli scienziati in specifici campi di scienze non dure (*soft sciences*) è stata poco affrontata (*cfr.* Barkuus e Brown, 2012, Antonijevic *et al.*, 2013). Il presente lavoro mira dunque a colmare tale *gap* conoscitivo, con riferimento all'ambito delle scienze sociali interrogandosi sul cambiamento che l'utilizzo di basi di dati e informazioni, e il loro eventuale sfruttamento condiviso, sono in grado di generare nei processi di produzione della conoscenza che presiedono alla creazione di *output* di ricerca.

L'approfondimento di un aspetto legato alle attività di ricerca fino a questo momento non adeguatamente esplorato in letteratura ha imposto una riflessione accurata su quale fosse l'approccio metodologicamente più affidabile da implementare, che consentisse di gettare luce sul fenomeno e possibilmente aprire la strada ad ulteriori futuri approfondimenti sullo stesso. A questo proposito, è apparso opportuno orientare l'indagine verso il contesto di una comunità scientifica *specificata*, raccogliendo testimonianze dirette delle esperienze e delle opinioni dei soggetti attivi entro tale gruppo di riferimento. Si è deciso dunque di delimitare il contesto di studio focalizzandosi su un'unità d'analisi definita dal punto di vista *disciplinare* – la comunità dei sociologi –, inquadrando il presente lavoro nell'alveo della ricerca qualitativa operata attraverso uno *studio di caso* applicato al maggior polo universitario italiano – Sapienza Università di Roma – e al maggior ente di ricerca pubblico italiano – il Consiglio Nazionale delle Ricerche (CNR).

Si è soliti definire un'indagine empirica quale *studio di caso* se incorpora caratteristiche quali a) l'investigazione in profondità di un fenomeno contemporaneo, o di un particolare aspetto di un fenomeno, attuata attraverso l'osservazione o la raccolta di informazioni all'interno del suo contesto naturale, poiché si ritiene che non sia possibile giungere ad una sua comprensione prescindendo dalla rilevazione di elementi contestuali, e b) la focalizzazione del ricercatore sulla descrizione e sulla spiegazione dei processi sociali che coinvolgono le persone, i gruppi o le organizzazioni (a seconda dell'unità di analisi) immerse nel contesto sotto studio (Yin, 2009; Swarnborn, 2010). È possibile individuare tre tipi riferibili a questa strategia di ricerca, distinti rispetto all'obiettivo finale. Vi sono infatti casi di studio *esplorativi*, definiti così se sono utili a definire le domande e le ipotesi di ricerca per uno studio successivo o per verificare la fattibilità di una procedura di ricerca; casi di studio *descrittivi*, che mirano semplicemente ad una descrizione completa del fenomeno nel suo contesto; casi di studio *esplicativi*, se sono implementati per testare teorie di tipo causale attraverso la combinazione di evidenze sia qualitative che quantitative (Yin, 2009). In accordo con la tipologia presentata, lo studio proposto incontra le caratteristiche del primo tipo, in quanto esso viene utilizzato come banco di prova non solo per controllare la validità di certe assunzioni discendenti dalla domanda di ricerca iniziale, ma anche per testare i metodi e le tecniche utilizzati, che potenzialmente possono essere estese ad altri casi.

### 3.1.1 Domanda di ricerca e ipotesi

Swarnborn (2010) ha suggerito che, nell'ambito di uno studio di caso, la domanda di ricerca dovrebbe essere definita nel modo più ampio possibile, in modo da essere raffinata o ridefinita nel corso dell'implementazione della ricerca, con l'acquisizione e l'analisi dei primi dati, e predisposta ad eventuali integrazioni con l'emergere di aspetti o elementi inattesi connessi al fenomeno. Poiché l'obiettivo di studio è stato identificato nell'esplorazione di come le nuove opportunità di accesso, creazione e condivisione di infrastrutture di ricerca si traducano nell'attività concreta degli scienziati sociali<sup>45</sup>, in via preliminare si è costituita una domanda di ricerca di ampia portata, sintetizzabile in questi termini:

*Qual è il cambiamento indotto dall'utilizzo delle infrastrutture di ricerca sui processi che presiedono alla produzione di nuova conoscenza negli studi sociologici?*

In particolare, gli aspetti su cui si è inteso focalizzarsi sono stati indirizzati all'approfondimento delle seguenti questioni:

- *Sub-1*: Come muta il processo di costruzione della conoscenza quando i ricercatori utilizzano infrastrutture di ricerca, con particolare riferimento al disegno di ricerca, alle dinamiche aggregative, alla tendenza all'interdisciplinarietà della ricerca?

- *Sub-2*: Quali sono le ripercussioni percepite dai ricercatori sull'*output* di ricerca conseguenti al passaggio da modalità tradizionali di accesso alle informazioni (biblioteca, *workshop*, informazioni non sistematizzate presenti sul *web*) a modalità aperte e/o condivise basate su creazione o sfruttamento di infrastrutture di ricerca condivise?

La domanda di ricerca ha preparato all'identificazione di alcune dimensioni da esplorare, riguardanti i processi che concorrono ad alimentare la produzione di nuova conoscenza scientifica. Tali dimensioni, per le quali si è supposto che l'utilizzo delle infrastrutture potesse avere un'influenza, sono state riprodotte all'interno di tre ipotesi operative di lavoro.

Le ipotesi operative a partire dalle quali lo studio si è mosso sono:

*H1: Potenziamento delle capacità di collaborazione e networking.* Le infrastrutture di ricerca possono essere pensate come strumenti di conoscenza condivisa che sostengono il

---

<sup>45</sup> Si tenga conto della restrizione del perimetro definitivo di infrastruttura di ricerca in riferimento al campo delle scienze sociali, così come evidenziato nel par. 2.2.

*networking* tra i ricercatori, promettendo di supportare le reti di collaborazione e offrendo nuove possibilità che travalicano i confini geografici dell'organizzazione spaziale del lavoro scientifico. L'ipotesi è che tali risorse consentano la partecipazione a nuovi *network* di ricerca e/o il consolidamento di quelli già esistenti e la formazione di nuove collaborazioni tra ricercatori.

• *H2: Maggiore incentivo alla ricerca interdisciplinare.* La disponibilità di infrastrutture che potenzialmente ospitano dati e indicatori su una molteplicità di temi offre la possibilità ai ricercatori di ampliare il panorama informativo relativo a sub-settori specifici, consentendo di approfondire tematiche di interesse da prospettive differenti. L'ipotesi è che risorse informative di tal genere possano incidere sulla formazione di domande di ricerca orientate alla produzione di studi interdisciplinari.

• *H3: Produzione di output di ricerca investiti da vantaggio qualitativo.* Eventuali cambiamenti nelle modalità operative del lavoro dei ricercatori per effetto dello sfruttamento delle infrastrutture digitali di ricerca possono ripercuotersi su alcune dimensioni che riguardano l'*output* prodotto in termini di migliore adeguamento agli obiettivi cognitivi, copertura campionaria, accuratezza del dato, robustezza dello stesso, maggiore capacità di generalizzazione, opportunità di comparazione. L'ipotesi è che l'utilizzo delle infrastrutture produca una serie di vantaggi qualitativi che i ricercatori riconoscono quali punti di forza degli *output* prodotti.

Le ipotesi operative appena presentate sono state sviluppate all'inizio del processo di ricerca, sulla base di una *review* della letteratura sul tema (capp. 1-2), e sono state formalizzate per essere testate attraverso il confronto con i risultati ottenuti attraverso l'analisi empirica. Si è supposto che il loro controllo potesse favorire una comprensione globale del fenomeno investigato, ma allo stesso tempo si è guardato con interesse alla possibile emersione di *nuovi elementi inattesi* riferibili a dimensioni non prese in considerazione in fase preliminare. I risultati derivati dallo specifico studio di caso potranno prestarsi ad ulteriori controlli da attuare attraverso future replicazioni della ricerca su contesti differenti (nuovi studi di caso), in ciò perseguendo un'istanza di generalizzazione *analitica* (Yin, 2009). Tale concetto fa riferimento al carattere di estendibilità dello studio, i cui termini verranno brevemente esplicitati nel prossimo sub-paragrafo.

### 3.1.2 Generalizzazione analitica

Le indagini basate su studi di caso si caratterizzano per l'applicazione a contesti specifici e per il numero molto limitato di unità osservate (o dalle quali si acquisiscono informazioni). Tali fattori pongono parecchi problemi per quanto riguarda l'estensione dei risultati ottenuti a unità *diverse* da quelle studiate, e pertanto influenzano il grado di generalizzabilità della ricerca. Solitamente il concetto di generalizzabilità è dunque utilizzato in riferimento a studi quantitativi e inerisce al grado con cui i risultati ottenuti attraverso un campione – selezionato attraverso metodi probabilistici – siano applicabili ad una popolazione di riferimento o ad altri campioni<sup>46</sup>.

Tuttavia, come rilevato da Yin (2009), esistono dei buoni motivi per implementare uno studio di caso anche al fine di perseguire obiettivi di generalizzazione. La strategia di ricerca dello studio di caso contribuisce infatti alla generazione e all'alimentazione di nuova teoria sulla tematica studiata, che si presta ad essere reiteratamente validata o rigettata da ulteriori ricerche, anche attraverso il riutilizzo di costrutti e di ipotesi delineati all'interno del precipuo disegno progettuale da cui è scaturita. A questo proposito, alla classica generalizzazione *di tipo statistico* – per cui è possibile inferire, con un certo grado di fiducia, i risultati ottenuti dall'analisi di dati empirici raccolti su un gruppo di unità alla popolazione da cui esse sono state tratte – l'autore oppone una generalizzazione di tipo *analitico*, subordinata alla logica della *replicazione*, per cui l'estensibilità è legata alla possibilità di riproposizione di controlli empirici su una teoria elaborata anche a partire da uno studio di caso<sup>47</sup>. In altre parole, la singola ricerca basata su studio di caso, se considerata atomisticamente, dà conto di come il fenomeno sotto esame opera in un contesto specifico e con riferimento al caso studiato, ma, allo stesso tempo, analogamente a quanto accade con gli esperimenti, nel contesto generale del complesso di ricerche intraprese su una determinata tematica, può produrre una nuova proposta analitica complessiva (riconfermando o revisionando la precedente), che si costituisce come fattore abilitante per l'intrapresa di un processo reiterativo di controllo e comparazione o generare nuove ipotesi di lavoro da controllare ulteriormente (Falk e Guenther, 2006; Yin, 2009). Nel caso in cui tutte le repliche dello stesso studio su contesti

---

<sup>46</sup> Il concetto di generalizzabilità è stato legato anche termini quali “trasferibilità” e “validità esterna” di una ricerca (cfr. Tashakkori e Teddlie, 2003).

<sup>47</sup> Yin (2009, p. 15): “Case studies (...) are generalizable to theoretical propositions, not to populations or universes (...) your goal will be to expand and generalize theories (analytical generalizability) and not to enumerate frequencies (statistical generalization)”.

diversi restituiscano i medesimi risultati, la teoria risulterà essere confermata e sarà possibile asserire la sua generalizzazione oltre il singolo caso, fino ad una replica confutativa. Se, al contrario, dall'analisi di nuovi casi emergessero risultati non collimanti ai precedenti, le ipotesi su cui poggia la teoria potranno essere riviste e a loro volta ritestate su un altro caso o un insieme di casi (attraverso disegni a caso singolo o a casi multipli). Più è alto il numero di repliche, più valida sarà la teoria (Ishak e Bakar, 2014).

Per quanto detto, questa ricerca si propone di esplorare il tema prescelto a partire da un insieme di ipotesi operative, da un lato prestandosi alla produzione di risultati legati alla specificità del caso, da controllare in studi successivi attraverso nuove ipotesi di lavoro, dall'altro fornendo una proposta metodologica per studiare il fenomeno, testata attraverso l'applicazione di tecniche e strategie precise. Relativamente all'ultimo punto, Yin (2009) ha rimarcato l'importanza della validità e dell'affidabilità del disegno di ricerca anche nel campo della ricerca basata su studi di caso, dal momento che la generazione di risultati, seppure convincenti, se scissa dal rigore legato alla scelta di disegni e metodi, non è condizione sufficiente perché uno studio di caso sia valido. Ad esempio, la scelta della singola unità informativa può giocare un ruolo fondamentale ai fini dei risultati ottenuti, in termini di appropriatezza e affidabilità. Inoltre, la modalità di acquisizione del dato (documenti, archivi, interviste, questionari, osservazioni dirette) può, in ogni caso, costituirsi quale punto di forza o di debolezza rispetto al fenomeno studiato (*ibidem*, *cfr.* par. 3.3).

## **3.2 Il disegno della ricerca**

### *3.2.1 Scelta del contesto di riferimento e individuazione del caso*

La ricerca ha optato per una modalità di studio di caso *singolo* di tipo *embedded*, caratterizzata dalla rilevazione di informazioni su più sub-unità riferibili al medesimo contesto di riferimento.

Yin (2009) opera una distinzione basica tra disegni di ricerca a caso *singolo* o a caso *multiplo*, a seconda se i casi sotto esame appartengano a uno o molteplici contesti di riferimento, intesi come scenari di caratterizzazione ove gli attori operano e i processi sotto studio si svolgono. Il presente lavoro si riferisce ad un unico contesto, delimitandolo dal punto di vista disciplinare: la comunità degli scienziati e dei ricercatori che svolgono attività di

ricerca *in campo sociologico*. L'esplorazione nell'ambito di un settore disciplinare specifico ha richiesto che i partecipanti alla ricerca appartenessero tutti al medesimo ambito.

Quando lo studio di caso singolo coinvolge una o più sub-unità d'analisi, la letteratura distingue tra studio di caso di tipo *olistico* e di tipo *embedded (ibidem)*. Nel primo tipo di disegno, il ricercatore è interessato a investigare la natura globale di una singola unità considerandola nella sua unica ed irripetibile complessità; nel secondo tipo di disegno, l'analisi coinvolge unità informative multiple, qualsivoglia sia il modo con cui vengano selezionate, poiché si ritiene che possano fornire un materiale informativo ricco ed eterogeneo, utile per un'analisi estensiva. A questo proposito, si è scelto di perseguire la seconda modalità selezionando e coinvolgendo diverse sub-unità, afferenti al maggiore ateneo italiano – Sapienza Università di Roma – e al maggiore Ente di Ricerca Pubblico italiano – il Consiglio Nazionale delle Ricerche (CNR), che dunque costituiscono il caso studiato in questa ricerca. La diversità delle sub-unità in termini di inquadramento in differenti strutture è da considerarsi ininfluenza, in quanto il lavoro non persegue obiettivi comparativi tra unità di ricerca, bensì aspira a riprodurre l'eterogeneità di esperienze e percezioni rispetto a due tra i principali poli che sviluppano ricerca sociologica in Italia. È altresì opportuno puntualizzare che il caso non è stato individuato perché rappresentativo di particolari criticità, né per caratteristiche di unicità o rappresentatività statistica, bensì in quanto sufficientemente rivelativo rispetto agli obiettivi cognitivi dello studio (tale caratteristica connota ulteriormente il presente studio di caso come *intrinseco* [Stake, 1995])<sup>48</sup>.

La figura in basso chiarisce il posizionamento della ricerca nell'ambito delle tipologie di studi di caso, così come identificate dalla quadripartizione di Yin (2009).

---

<sup>48</sup> In altre parole, si è ritenuto che l'applicazione della domanda di ricerca su tale contesto potesse fornire al conduttore dello studio le risposte o le evidenze che si è proposto di ricercare.

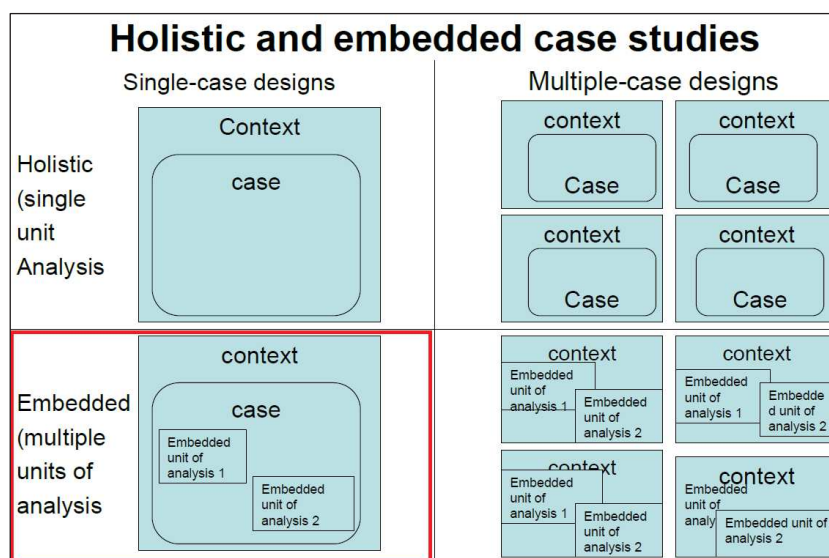


Figura 1.  
Tipologia di casi di studio di Yin (2009) e posizionamento del presente lavoro<sup>49</sup>

### 3.2.2 Selezione e numerosità delle sub-unità

La base empirica dello studio di caso è stata formata da ricercatori che operano presso:

- i 2 dipartimenti sociologici di Sapienza Università di Roma:
  - il Dipartimento di Comunicazione e Ricerca Sociale (CORIS) e
  - il Dipartimento di Scienze Sociali ed Economiche (DiSSE)
- i 2 principali istituti di ricerca sociale del Consiglio Nazionale delle Ricerche:
  - l'Istituto di Ricerca sulla Crescita Economica Sostenibile (IRCrES) e
  - l'Istituto di Ricerche sulla Popolazione e le Politiche Sociali (IRPPS).

Il CORIS, risultato della fusione dei “Dipartimenti di Ricerca Sociale e Metodologia Sociologica” e di “Sociologia e Comunicazione”, promuove e coordina le attività di didattica e di ricerca nel campo delle varie aree della sociologia, delle scienze umane e dei *media studies*, producendo studi sia in ambito nazionale che internazionale<sup>50</sup>. Il DiSSE, risultato della fusione tra il precedente “Dipartimento di Scienze Sociali” e il “Dipartimento di Analisi Economiche e Sociali”, offre alla comunità scientifica un contributo interdisciplinare basato sull’apporto di sociologi, economisti, storici, psicologi, giuristi, antropologi, statistici,

<sup>49</sup> La figura è stata tratta da una presentazione di Skogerbø (2011).

<sup>50</sup> <http://www.coris.uniroma1.it/>

adottando prospettive di ricerca orientate all'applicazione empirica<sup>51</sup>. L'IRCrES è un istituto di ricerca che presenta tra le principali aree tematiche di ricerca l'organizzazione e sostenibilità dei grandi sistemi delle società contemporanee, lo studio specifico sulle politiche della scienza e della tecnologia e l'analisi socio-economica delle tematiche ambientali, ed è composto da uno staff con competenze su materie sociologiche e materie economiche<sup>52</sup>. L'IRPPS è infine un istituto interdisciplinare che svolge studi su tematiche demografiche e migratorie, sui sistemi di welfare e sulle politiche sociali, sui rapporti tra scienza e società, su creazione, accesso e diffusione della conoscenza e delle tecnologie dell'informazione ed integra un insieme di analisi, di studi teorici ed empirici, favorendo la collaborazione tra diverse frontiere disciplinari<sup>53</sup>.

Le quattro strutture di ricerca sopracitate hanno tutte sede nella città di Roma, per cui le evidenze che emergeranno dallo studio saranno da considerarsi espressione di una *local community* caratterizzata territorialmente: tale circostanza incontra tuttavia il carattere esplorativo dello studio di caso, che non richiede che i partecipanti alla ricerca rappresentino la loro popolazione (la comunità dei sociologi italiani) ma che riportino informazioni in grado di approfondire la comprensione del fenomeno sotto esame. Nonostante l'eliminazione dell'inferenza statistica, si è inteso comunque formare il campione tenendo conto di fattori quali genere, età, posizione accademica/ruolo nell'ente di ricerca. Si è ritenuto infatti che uno studio accurato di come le nuove opportunità di accesso e condivisione, relative alle infrastrutture di ricerca, si traducano nell'attività concreta dei ricercatori dovesse tener conto sia del livello gerarchico occupato sia del passaggio generazionale tra soggetti che abbiano conosciuto modalità diverse di diffusione della conoscenza – antecedenti all'avvento del *web* (con dinamiche tradizionali di circolazione delle informazioni), caratterizzate dalla disponibilità sul *web* di informazioni non ancora sistematizzate, basate sull'accesso o creazione di risorse condivise e/o aperte.

La selezione delle singole unità all'interno delle strutture individuate si è dunque ispirata a una duplice esigenza: i. coinvolgere soggetti con caratteristiche che incontrassero l'obiettivo dello studio, ossia avessero capacità di fornire informazioni rilevanti sul tema, e che riproducessero un'eterogeneità di prospettive esperienziali; ii. assicurare un numero di

---

<sup>51</sup> <https://web.uniroma1.it/disse/>

<sup>52</sup> <http://www.ircres.cnr.it/index.php/it/>

<sup>53</sup> <http://www.irpps.cnr.it/it>

partecipanti sufficiente a spiegare il fenomeno coerentemente alla domanda di ricerca formulata.

- i. Il successo di un'investigazione basata sulla raccolta di informazioni da unità informative specifiche dipende in larga parte dalla loro "qualità": dovrebbero essere infatti selezionate in modo da garantire la raccolta di informazioni pertinenti, estensive, complete e dettagliate. Due delle domande che il conduttore dello studio deve porsi sono "chi possiede le informazioni rilevanti?" e poi "quali persone che possiedono l'informazione sono più capaci di fornire un *accounting* più accurato?" (Gorden, 1975). Specialmente se la ricerca intende servirsi della "conoscenza speciale" delle unità rispetto alla comprensione di una situazione o un processo sociale, il processo di selezione deve considerare elementi che siano quanto più indicativi della rispondenza delle unità alle esigenze del ricercatore (Glaser e Laudel, 2009b).

Nel caso della presente ricerca, si è inteso selezionare scienziati e ricercatori a livelli di carriera – ricercatori junior/assegnisti di ricerca, professori associati/ricercatori di enti di ricerca pubblico, professori ordinari/dirigenti di ricerca –, i cui *curriculum vitae*, disponibili sui siti Internet delle strutture di afferenza, presentassero elementi che facessero desumere una propensione all'utilizzo o alla creazione di infrastrutture di ricerca, o che la indicassero esplicitamente<sup>54</sup>. A tale scopo si è posta particolare attenzione su elementi di carriera quali pubblicazioni, reportistica, atti di conferenza, partecipazione a progetti di ricerca nazionali e internazionali. L'eventuale presenza, insieme a soggetti che abbiano utilizzato infrastrutture di ricerca a diversi livelli, di altri che non lo abbiano fatto è stato considerato come fattore di possibile arricchimento per lo studio. Ad ogni modo, si è ritenuto opportuno prevedere preliminarmente la presenza sicura di soggetti che presentassero un'esperienza più avanzata rispetto alla pratica di sfruttamento di tali risorse. Per questo motivo, è stata prevista l'inclusione di tre "testimoni privilegiati", il cui esame del *curriculum* ha rivelato una particolare attitudine all'uso di infrastrutture, che potesse costituire il *benchmark* di riferimento su pratiche avanzate: uno è stato individuato nell'ambito di una struttura che costituisce il caso di questa ricerca, gli altri due afferiscono invece a strutture estere e sono stati scelti in

---

<sup>54</sup> La conoscenza del campo sociologico del conduttore dello studio ha reso meno difficoltosa tale valutazione, sebbene non sia stato possibile stabilire *a priori* l'esatto livello di *expertise* degli intervistati relativamente allo sfruttamento di queste risorse.

quanto portatori di un'esperienza particolarmente avanzata di sfruttamento e condivisione di infrastrutture maturata nell'ambito del progetto europeo RISIS, finanziato dal Settimo Programma Quadro. Tale progetto, al quale sia Sapienza Università di Roma che Consiglio Nazionale delle Ricerche partecipano come *partner*, ha acquisito particolare rilevanza per il suo obiettivo di integrare in forme accessibili 13 *database* proprietari, formando un'infrastruttura armonizzata di tipo distribuito rivolta alle comunità scientifiche di sociologi ed economisti (*cf.* par. 2.4.2).

- ii. Quanto alla numerosità delle unità, i ricercatori qualitativi, specialmente negli studi di caso, sono soliti selezionare gradualmente i loro casi, non determinandone in anticipo la quantità, ma procrastinando tale scelta a fasi successive dell'implementazione della ricerca (Glesne e Peshkin, 1992).

Il campionamento di questo studio è stato dunque concepito come un processo iterativo e sequenziale in cui l'aumentare della mole di dati e una prima riflessione analitica sul materiale ottenuto hanno guidato la partecipazione di nuovi soggetti. Era stato previsto, in fase preliminare, un numero provvisorio di 10 unità su cui raccogliere le informazioni di interesse, includendo sia rappresentanti dell'Accademia sia di enti pubblici di ricerca in una proporzione di 7:3. Nel corso della ricerca, il conduttore dello studio ha continuato a raccogliere informazioni su nuove unità, finché ha ritenuto che l'ammontare delle stesse avesse raggiunto il punto di saturazione (*saturation point*), per cui la raccolta di nuovi dati non avrebbe concorso all'incremento della ricchezza informativa dei risultati (Ishkar e Bakar, 2014). Per questo motivo, il numero di unità è prima salito a 13, con l'aggiunta dei 3 testimoni privilegiati, e infine si è stabilizzato a 18, per effetto dell'ulteriore aggiunta di 5 soggetti appartenenti al contesto di riferimento.

Per questi motivi, il campione finale ha incluso:

- 15 soggetti operanti negli istituti e nei dipartimenti presentati più sopra<sup>55</sup>:
  - 4 tra ricercatori *junior* e assegnisti di ricerca  
(2 uomini, 2 donne);

---

<sup>55</sup> Dodici dei quindici soggetti operano nell'ambito dei due dipartimenti di Sapienza Università di Roma, mentre i rimanenti tre operano negli istituti del Consiglio Nazionale delle Ricerche. La proporzione a largo vantaggio dell'istituzione universitaria riflette una stima della distribuzione di ricercatori sociali nelle università e negli enti pubblici di ricerca di Roma.

- 8 tra professori associati e ricercatori di enti pubblici di ricerca (4 uomini, 4 donne);
- 3 tra professori ordinari e dirigenti di ricerca di enti pubblici (1 uomo, 2 donne).
- 3 testimoni privilegiati, portatori di esperienze di tipo “avanzato”:
  - 1 ricercatore afferente all’IRCrES CNR;
  - 1 professore dell’Università della Svizzera Italiana, afferente al progetto RISIS;
  - 1 ricercatore dell’Austrian Institute of Technology, afferente al progetto RISIS.

### 3.3 La tecnica di raccolta dati: l’intervista semi-strutturata

Le interviste rappresentano una delle fonti privilegiate nell’applicazione degli studi di caso (Yin [2009] le annovera tra le più utilizzate insieme a questionari, osservazioni dirette, e lo studio di artefatti fisici). Esse possono assumere la forma di: *interviste con domanda aperta*, in cui il contenuto non è prestabilito ma può variare da intervista a intervista; *interviste semi-strutturate*, in cui il contenuto delle domande è prestabilito, ma non lo è la forma, poiché l’intervistatore dispone di una traccia che riporta gli argomenti che deve toccare nel corso dell’intervista; *interviste strutturate*, caratterizzate da un insieme fisso e ordinato di domande aperte che vengono sottoposte a tutti gli intervistati nella stessa formulazione e nella stessa sequenza (Corbetta, 1999)<sup>56</sup>.

La strategia di ricerca implementata al fine di raccogliere i dati per questo studio di caso si è basata su *interviste semi-strutturate*. Per garantire che le medesime linee base di indagine fossero seguite nel contesto del processo comunicativo con i soggetti intervistati, è stata creata una traccia d’intervista (vedi allegato) in cui i temi da esplorare sono stati operazionalizzati attraverso domande a risposta aperta. La presenza di un tale supporto ha garantito una focalizzazione della comunicazione sulle dimensioni d’interesse individuate

---

<sup>56</sup> In genere, la ricerca qualitativa evita l’utilizzo di metodi standardizzati di intervista poiché chiudono la porta all’opportunità di capire cosa gli intervistati pensano realmente e si prestano a interpretazioni soggettive dei significati (gli intervistati attribuiscono certi significati alle domande poste e rispondono sulla base di interpretazioni soggettive). Le modalità più flessibili di intervista, al contrario, si caratterizzano per un alto grado di sensitività rispetto al contesto, sulla base dell’assioma che la ricerca sociale scientifica si debba adattare alle regole di comunicazione quotidiana che esistono preliminarmente al processo di ricerca, permettendo all’intervistato di esprimere liberamente le sue strutture di rilevanza.

preliminarmente (influenze dello sfruttamento di infrastrutture di ricerca nel disegno della ricerca, nella stipula di collaborazioni, nell'approfondimento di temi interdisciplinari), consentendo tuttavia al ricercatore una certa libertà nell'utilizzo del *wording* e nella gestione dell'ordine delle domande. Agli intervistati non è dunque stata posta una rigida ed immutabile batteria, quanto piuttosto un insieme di input precostituiti, lasciando alla "sensibilità ermeneutica" (*cfr.* Montesperelli, 2008) del rilevatore e al flusso argomentativo delle risposte, la possibilità di introdurre domande non previste. Questo modo di affrontare la situazione di intervista ha consentito di affrontare le questioni chiave con tutti gli intervistati, offrendo flessibilità nel determinare quando fosse appropriato sondare certe materie più in profondità o acquisire nuove informazioni, traendo vantaggio da particolari opportunità contingenti (Patton, 2002).

Trinczek (2009) ha rilevato che nelle *thematically focused expert interview* la caratterizzazione argomentativo-discorsiva e l'innesto di modalità narrative rappresentassero le forme tecniche più appropriate da utilizzare. Lo stesso autore ha tuttavia sottolineato come l'operazionalizzazione degli *item* nell'atto dell'intervista debba essere ad ogni modo concepita in riferimento all'adeguatezza rispetto agli argomenti da trattare. Per questa ricerca si è optato per una struttura domanda-risposta in apertura e in chiusura, per la raccolta di definizioni e opinioni, e per una forma maggiormente narrativo-discorsiva per la parte centrale. In particolare, è stato previsto che la parte "narrativa" avesse ad oggetto uno o più prodotti di ricerca (pubblicazioni, report, atti di conferenza) generati dallo sfruttamento di infrastrutture di ricerca, con l'obiettivo di ricostruire retrospettivamente il processo di ricerca che ha concorso alla loro formazione, in modo da far emergere modalità peculiari di realizzazione in riferimento alle infrastrutture utilizzate. Dopo aver raccolto l'informazione sul numero complessivo di eventi di utilizzo di infrastrutture sperimentati si è proceduto ad individuare alcuni *output* di ricerca da essi derivati e a stimolare il processo di ricostruzione delle fasi della ricerca, dalla progettazione, alle modalità di eventuale lavoro collaborativo, alla presentazione dei risultati. L'identificazione degli *output* ha richiesto uno sforzo mnemonico da parte dell'intervistato, tuttavia – nei casi di *impasse* – l'analisi curriculare

preliminare, operata in fase di selezione, ha consentito all'intervistatore di suggerire possibili prodotti di ricerca legati all'utilizzo di infrastrutture<sup>57</sup>.

### 3.3.1 Vantaggi e caveat della socializzazione preliminare al setting di ricerca

Le interviste effettuate nell'ambito del presente studio sono da considerarsi *scientifically informed interview* ("interviste scientificamente consapevoli", *traduzione mia*) (Laudel e Glaser, 2007), cioè interviste operate da un sociologo che è addentro al *field* scientifico, e quindi al contesto di riferimento, sotto studio, e che quindi può comprenderne le dinamiche con maggiore facilità rispetto a un rilevatore esterno.

Seguendo Weber (1922, trad. it. 1962), se la dimensione interpretativa dell'azione sociale è rappresentata dalla spiegazione causale del suo corso e delle sue conseguenze, il sociologo che voglia giungere alla creazione di modelli esplicativi dovrebbe fare il più possibile proprio il quadro di riferimento della cultura studiata, assimilando conoscenza relativa ad aspetti legati a fattori di materialità dei comportamenti attuati dai soggetti osservati. Al contrario, l'impossibilità di partenza di praticare questo approccio "empatico" nella lettura delle azioni, data dal trovarsi all'esterno della cultura studiata, può essere considerato un forte dazio potenzialmente dirimente per la buona riuscita della ricerca. In questo senso, il sociologo che si trovi a studiare il *setting* della disciplina cui afferisce, ponendosi come conduttore di studio "consapevole" rispetto al *field* da analizzare<sup>58</sup> lo pone in una condizione privilegiata nel reperimento dei dati utili alla sua ricerca, dalla facilità del contatto con i soggetti da intervistare all'interpretazione dei dati ottenuti.

Che la socializzazione o un'informazione preliminare rispetto al *setting* di ricerca, al contrario dell'adozione di approcci prettamente *naive* ("ingenui")<sup>59</sup>, rappresenti un

---

<sup>57</sup> In letteratura è stato talvolta utilizzato un metodo alternativo per espletare lo stimolo narrativo verso ricercatori che dovessero ricostruire i propri percorsi di ricerca, consistente in visualizzazioni grafiche rispetto al *network* di pubblicazioni e temi di ricerca affrontati (*cf.* Glaser e Laudel, 2009a).

<sup>58</sup> Cioè che abbia acquisito una comprensione scientifica del campo di studio sotto esame.

<sup>59</sup> L'approccio *naive* utilizzato nello studio di Latour e Woolgar (1979, *cf.* par. 1.1.3) si basava sulla cosiddetta "osservazione ingenua", per cui il sociologo-osservatore rimaneva di fatto estraneo all'ambito di studio. I due studiosi hanno asserito che il pericolo di "diventare nativi", acquisendo ulteriori elementi di cultura appartenenti alla realtà scientifica studiata, poteva distorcerne in qualche modo la comprensione. Lynch (1993) ha sollevato due obiezioni: a) che le stesse descrizioni di Latour e Woolgar provano che non è possibile mantenere l'approccio "ingenuo", sia da un punto di vista temporale (durante le fasi di osservazione si acquisisce comunque una certa conoscenza della realtà studiata), sia dal punto di vista della comprensione di procedure tecniche osservate, altrimenti non possibile da parte di soggetti completamente estranei al campo di studio; b) che tale approccio può fortemente limitare la comprensione, in particolare inducendo gli osservatori a selezionare solo quegli eventi o azioni che appaiano loro neutralmente intellegibili.

prerequisito desiderabile nell'ambito di ricerche aventi a tema la comprensione del lavoro scientifico è convinzione prevalente negli studi etnografici e negli studi di caso nell'ambito dei *science studies* (cfr. Laudel e Glaser, 2007)<sup>60</sup>. Sebbene un rilevatore “esterno” possa in qualche modo acquisire le competenze dei “nativi” del *setting scientifico* sotto studio, permane la circostanza che egli può raggiungere una comprensione del loro campo di studio non tale da divenire così competente da comprendere (o al limite performare egli stesso) la ricerca che osserva. Essere un “nativo” del *setting* scientifico consente invece di colmare il distacco culturale tra conduttore dello studio e soggetto studiato, mettendo il primo in condizione di pervenire ad una più profonda conoscenza della pratica scientifica sotto esame (cfr. lo studio dei radio-astronomi Mulkay ed Edge [1976] sul campo della radio-astronomia).

Un intervistatore consapevole del *frame* di riferimento o “nativo” può meglio sfuggire a situazioni di “abbaglio comunicativo” (Briggs, 1986), per cui nelle interviste qualitative si incorre talvolta nel rischio di sottomettere domande non pienamente comprensibili, ricevere risposte non richieste o nel mancare la comprensione delle stesse. Peraltro, quando gli scienziati parlano dei loro processi di ricerca tendono a farlo in maniera tecnica riferendo il contenuto epistemico del loro lavoro (problemi di ricerca, oggetti, metodi di sperimentazione, strumenti): nel caso dell'intervistatore “nativo” la comprensione risulta molto facilitata, data la sua immersione nelle dinamiche del *field*, mentre l'intervistatore “esterno” dovrà in qualche modo adattarsi ad un livello scientifico raggiungibile, che dovrà negoziare con l'intervistato. Il dimostrarsi un “addetto ai lavori” e mostrare competenza genera inoltre una maggiore probabilità di essere “presi sul serio” e ottenere risposte affidabili (Laudel e Glaser, 2007). Tale approccio non è però completamente scevro da possibili implicazioni negative. Ad esempio, Mulkay ed Edge (1976) hanno notato che gli “intervistatori consapevoli” possono mettere in atto scambi interessati di giudizi e opinioni, che potrebbero distorcere i punti di vista, riproducendo false e incomplete assunzioni come se fossero prodotte dal gruppo che si studia (Laudel e Glaser, 2007). Tuttavia, nel caso della presente ricerca, il pericolo di orientare o indirizzare l'intervistato si è attestato su livelli molto bassi, dato che

---

<sup>60</sup> Sulla base del bagaglio esperienziale scaturito dalla moltitudine di ricerche aventi a tema la comprensione del lavoro scientifico (per citarne alcuni, Mulkay ed Edge [1976], Latour e Woolgar [1979], Laudel [2001]), Laudel e Glaser (2007), hanno passato in rassegna i diversi modi di affrontare il problema di come pervenire, attraverso ricerche qualitative, ad una comprensione scientificamente fondata del lavoro degli scienziati, rilevando che un modello di intervista “scientificamente consapevole” a scienziati, sia inserita nell'impianto di una ricerca etnografica sia nell'ambito di metodologie *stand-alone*, fosse emerso quale modalità più appropriata per la comprensione della pratica scientifica.

gli intervistati sono tutti operanti nel campo delle scienze sociali e si suppone siano soliti riconoscere e annullare tali influenze nel loro stesso lavoro di ricerca.

Ad ogni modo, anche nel caso di un'intervista "scientificamente consapevole", è stato necessario mettere in pratica le usuali accortezze da applicare in ogni situazione di intervista (*cf.* Seidman, 1991), declinate in riferimento alla particolare unità d'analisi. Data la peculiarità del rapporto tra ricercatore "conduttore di studio" e ricercatori "oggetto di studio", è stato indispensabile che sia la preparazione sia il momento dell'intervista fossero gestiti in maniera particolarmente attenta. In particolare, è stato tenuto presente che i rispondenti potessero non esprimersi in maniera esplicita su certi argomenti assumendo che fossero già di conoscenza dell'intervistatore "informato". In questo caso, la soluzione al problema è stata rappresentata da pratiche di *probing* in profondità attuate durante il dialogo (come suggerito da Laudel e Glaser, 2007). Inoltre, sebbene il protocollo di intervista fosse unico per tutti i 18 partecipanti, l'investigazione del contenuto della ricerca del singolo rispondente, la conoscenza di problemi, strategie e logiche di ricerca scientifica con cui lo scienziato ha avuto a che fare hanno aiutato a gestire la situazione di intervista, facilitando il processo comunicativo.

### **3.4 Strategia d'analisi**

Se da una parte i lavori di tipo descrittivo si pongono l'obiettivo di esplorare un fenomeno o "dare voce" all'esperienza e alle opinioni dei soggetti studiati (in particolare se del fenomeno in esame o del gruppo in questione si conosce ben poco), l'analisi interpretativa penetra maggiormente la superficie del dato, provando a comprendere come e perché particolari processi si siano generati, operando costruzioni di senso che vadano oltre il significato letterale del dato raccolto e argomentando una sorta di teorizzazione generale (Clarke e Braun, 2013). Tuttavia, non occorre operare una totale disgiunzione tra le due modalità. L'analisi descrittiva infatti non ha un valore che si limita alla sola riproduzione di resoconti descrittivi, bensì fornisce un supporto per la lettura e il collegamento dei dati, consentendo la produzione di un quadro concettuale propedeutico a letture ulteriori e più penetranti dei fenomeni. Nel caso della presente ricerca, l'obiettivo esplorativo generale si è riverberato anche nella fase d'analisi, poiché il ricercatore, cercando di sondare alcune

dimensioni legate al fenomeno sotto studio, ha mirato ad accendere i riflettori su una particolare realtà, riproducendone le dinamiche e i processi attraverso le parole dei soggetti intervistati.

Questo lavoro ha inteso utilizzare per l'analisi effettiva del materiale empirico il metodo dell'*analisi tematica*. Sebbene fosse utilizzata con frequenza nell'ambito di ricerche qualitative, tale procedura ha acquisito visibilità autonoma nell'ambito delle scienze sociali solo recentemente (Braun e Clarke, 2006). L'analisi tematica consiste in una procedura per identificare e analizzare temi che emergono da un insieme di dati testuali. Non prescrive metodi specifici di raccolta dei dati, posizioni teoretiche, quadri ontologici ed epistemologici e ha dunque nella flessibilità di adattamento al disegno di ricerca il suo maggior punto di forza. I temi possono essere identificati in modalità *bottom-up* guidata dai dati (*data-driven*), sulla base di un approccio induttivo; oppure possono essere identificati con una modalità maggiormente *top-down*, per cui il ricercatore usa i dati per esplorare particolari idee teoretiche o portare queste a influenzare l'analisi che viene condotta. Tali modalità di identificazione sono spesso combinate in un'unica analisi e in questo non si è distaccata la presente ricerca, in cui la fase analitica ha preso le mosse dal controllo delle ipotesi teoriche, ma ha lasciato spazio all'emersione di aspetti e dimensioni impreviste.

La procedura d'analisi implementata è sintetizzabile in tre fasi:

- i. Il primo fondamentale *step* verso questo tipo di analisi è stato ovviamente costituito dalla trascrizione integrale del contenuto delle interviste. Questa fase ha visto il costituirsi di una mole cospicua di materiale empirico che ha riprodotto la varietà di esperienze e di punti di vista riguardo la tematica di interesse della ricerca. Una delle caratteristiche connesse al metodo degli studi di caso, ossia la flessibilità di compiere aggiustamenti e deviazioni durante il processo di raccolta dei dati sugli aspetti da indagare (Yin, 2009), ha consentito la costituzione di una quantità di dati testuali sufficiente a fornire un panorama informativo quanto più completo rispetto al fenomeno in esame.
- ii. Un secondo passaggio è stato rappresentato dalla procedura di *coding*, cioè dal processo di identificazione degli aspetti collegati alla domanda e alle ipotesi di ricerca, attraverso l'applicazione di etichette testuali a segmenti di dati. A questo scopo, sono stati identificati *data-derived code*, che hanno riflesso contenuti

semantici dei dati e *researcher-derived code* che derivano dall'interpretazione concettuale e teoretica dei dati da parte del ricercatore. Successivamente, si è proceduto alla produzione di una mappa tematica che presentasse anche una possibile relazione tra i temi emersi. Tale azione si è rivelata propedeutica alla definizione finale dei temi e alla scelta degli estratti da utilizzare per illustrare le differenti sfaccettature rispetto alle dimensioni selezionate.

- iii. Infine, l'analisi è stata completata con la produzione di una narrativa basata su tali estratti, interpolata da citazioni letterali, che desse evidenza al lettore di ciò che è emerso dall'implementazione delle interviste per ciascun tema identificato. Al fine di non appesantire il testo, il riferimento ai soggetti citati avverrà tramite l'indicazione di un codice (R1 nel caso di ricercatori *junior* e assegnisti di ricerca; R2 nel caso di professori associati e ricercatori di enti pubblici di ricerca; R3 nel caso di professori ordinari ed dirigenti di ricerca di enti pubblici), con la giustapposizione delle lettere U (uomo) e D (donna) a seconda del genere dell'intervistato. Non si è ritenuto di applicare codici diversi nel caso di afferenza a università o enti pubblici di ricerca poiché nel corso della ricerca non si è riscontrata una significativa differenza nelle risposte che potesse essere riferita all'istituzione di provenienza. Infine, dato il numero ridotto di testimoni privilegiati, al fine di preservarne l'anonimità, non si è assegnato alcun codice per differenziarli.

I risultati derivati dal presente lavoro di ricerca potranno costituire la base per applicazioni di studio successive rispetto alla tematica di oggetto di studio.

## CAPITOLO 4

### Analisi del materiale empirico:

#### le infrastrutture di ricerca nel lavoro degli scienziati sociali

#### 4.1 Le infrastrutture di ricerca nelle scienze sociali: in cerca di una definizione

Il lavoro di analisi del materiale empirico scaturito dalle interviste ha inteso dapprima esplorare i confini del perimetro semantico entro cui gli scienziati sociali collocano il concetto di *infrastruttura di ricerca*, in che termini lo percepiscono e a quali referenti empirici lo attagliano. Nell'ambito della somministrazione delle interviste non sono state infatti esplicitate definizioni operative preliminari del concetto di infrastruttura di ricerca per mantenere aperto il carattere esplorativo dell'indagine. Come già rilevato, la locuzione è associata in letteratura ad una costellazione di elementi relativi a caratteristiche umane, tecniche, operative e organizzative (Renschler *et al.*, 2013)<sup>61</sup>. Tale molteplicità definatoria si è in effetti riflessa nella varietà delle risposte alla domanda d'esordio "*Se le fosse chiesto di dare una definizione di infrastruttura di ricerca, in che termini ne parlerebbe?*". La ragione principale di questo "caos semantico" è ascrivibile alla circostanza per cui ciò che è percepito come infrastruttura di ricerca si declina in maniera estremamente mutevole in rapporto a diversi fattori, tra cui le visioni personali dei ricercatori sull'assetto del lavoro scientifico e la variabilità di esperienze sviluppate e obiettivi conseguiti attraverso il suo utilizzo. Così, nel corso delle varie interviste, la definizione di infrastruttura di ricerca è andata modellandosi su specifiche caratteristiche descrittive come quelle relative al posizionamento e dematerializzazione ("*risorse delocalizzate e virtuali*"), alla conservazione di contenuti informativi di tipo misto ("*database numerici e testuali*"), all'inclusione nel sistema della ricerca ("*elementi del sistema sociotecnico della ricerca*"), o funzionali (es. "*piattaforme per reperire informazione e veicolare la collaborazione*", "*complessi digitali e fisici d'ausilio al lavoro dello scienziato*", "*strumenti di connessione tra ricercatori*", "*complessi di competenze per l'analisi avanzata dei dati*"), in pochi casi fornendo visioni che contemplassero insieme più aspetti, se non dopo la sottoposizione di stimoli e *probe* da parte

---

<sup>61</sup> Cfr. par. 2.1.1.

dell'intervistatore. In un caso le infrastrutture di ricerca sono state associate alla rete Internet *tout court*, che con la sua evoluzione tecnica consente oggi una maggiore facilità di diffusione delle informazioni e dei dati utili per affrontare il lavoro di ricerca.

Specialmente gli scienziati che vantano una più lunga carriera hanno espresso qualche esitazione rispetto alla possibilità di circoscrizione del concetto, situandosi lungo uno spettro che va dall'espressione di una supposta distanza dello stesso rispetto alle scienze sociali (*“le infrastrutture di ricerca per solito si abbinano alle scienze hard per cui uno pensa all'anello del CERN o al genoma umano, ecc. Nelle scienze sociali è un po' particolare la definizione... – R3D)* all'implicazione sottesa di una visione peculiare del lavoro di ricerca, lontana dalla propria (*“il concetto di infrastruttura di ricerca non l'ho mai utilizzato perché è un concetto che, in qualche modo, tende a rappresentare il lavoro di ricerca nel campo delle scienze sistemiche, e delle scienze sociali all'interno delle scienze sistemiche, come il ricercatore da una parte e l'infrastruttura dall'altra [...] sono aduso a pensare a me stesso in un sistema interattivo in cui la controparte è l'infrastruttura che però comprende anche gli altri utenti dell'infrastruttura e il team che lavora sull'infrastruttura” – R3U).*

In riferimento a quest'ultimo aspetto, in cinque interviste è emersa una definizione per così dire “sistemica” dell'infrastruttura di ricerca, sganciata da referenti empirici unitari e coerente con la visione sociotecnica della scienza elaborata nell'ambito dei *Science and Technology Studies*<sup>62</sup>. Paradigmatica in questo senso è l'immagine dell'infrastruttura come “oggetto frontiera”, che assume un significato per la funzione di tenere insieme oggetti e soggetti: laboratori, macchine e addetti alla ricerca (*“le infrastrutture non sono mai soltanto autoconsistenti ma hanno bisogno di oggetti intermedi che tra loro possono consentire collegamenti anche con altre infrastrutture [...]; per funzionare [l'infrastruttura] ha bisogno di tutti i network di expertise che gli stanno intorno. Non riesco a guardare all'infrastruttura come dimensione autonoma che è capace di produrre” – R2D).* Tali “oggetti di collegamento” possono consistere ad esempio nelle basi di dati, considerate come elementi di caratterizzazione su cui l'infrastruttura si incardina (*“il database può essere il mattone di un'infrastruttura, il database da solo è inerte. [...] Quindi è un pezzo fondamentale di un'infrastruttura più complessa, di insieme più complesso, insomma...” – R1U).* Dall'intervista ad un giovane ricercatore universitario, è emerso un quadro riassuntivo di cosa sia un'infrastruttura di ricerca dal punto di vista “sistemico”. Essa è rappresentata da tre livelli

---

<sup>62</sup> Cfr. par. 1.1.3.

a generalità discendente: un livello “organizzativo” che è costituito dall'apparato burocratico che supporta il ricercatore sul piano fattuale della progettazione e nell'implementazione dei progetti e coincide con l'istituzione universitaria dal punto di vista economico, giuridico, logistico; un livello c.d. dell' “*actor-net*”, costituito dalla rete di attori che circonda il ricercatore e con i quali egli condivide le attività di tipo scientifico; infine, un livello pratico, legato alla possibilità di sviluppo di modalità metodologiche e operative e coincidente con interfacce fisiche quali computer, biblioteche, programmi di analisi dei dati, e con piattaforme informative quali i *database*<sup>63</sup>. L'infrastruttura si pone dunque, in questi termini, come l'insieme delle condizioni necessarie per la costituzione di un ambiente di lavoro che consenta al ricercatore di operare e di condividere l'attività di ricerca con altri attori, al cui interno si stagliano degli oggetti come i *database* che, costituendo “*l'elemento tangibile più imponente dell'infrastruttura, anche perché senza quello non si arriverebbe poi a dei risultati scientifici e alla conseguente produzione scientifica*” (**R1D**), sono percepiti come mezzi abilitanti di processi e collegamenti, o come veri e propri costrutti sociali distinguibili all'interno della rete di relazioni e conoscenze.

Al netto di tale visione di tipo diffuso, in cui l'infrastruttura sfugge all'individuazione di referenti empirici univoci, la maggior parte degli intervistati rileva invece una coincidenza tra infrastrutture di ricerca e oggetti specifici, risorse identificabili e autoconsistenti. Esse coincidono con una pluralità di servizi di tipo diverso, di disponibilità perlopiù informative, che aggregano competenze e dati in forma fisica o virtuale: “*archivi istituzionali*”, “*database*”, “*banche dati sistematizzate e non*”, “*strutture di dati di ricerche classiche*”, “*laboratori di ricerca*”, “*archivi cartacei*”, “*archivi di dati amministrativi*”, “*archivi di prodotti della ricerca*”, raccolte di “*lavori che sono ancora in progress, cioè non definitivi e non pubblicati nel senso editoriale del termine*”. Se la forma dell'infrastruttura di ricerca prevalente nel passato era quella fisica (“*fino a qualche tempo fa avrei definito l'infrastruttura di ricerca come di fatto una “struttura di ricerca”, come un laboratorio o un centro in cui erano disponibili alcune attrezzature, dei macchinari*” – **R2D**), l'innovazione tecnologica e lo *switch* digitale hanno reso tali risorse sempre più virtuali, andando a costituire in parte “*l'evoluzione della biblioteca e in parte l'evoluzione di andare a cercarsi i dati qua e là*”

---

<sup>63</sup> Il giovane ricercatore intervistato ha esplicitato una visione del *database* come infrastruttura di livello micro che acquisisce senso all'interno di un disegno infrastrutturale più ampio: “*è ovvio che il database di per sé è un'infrastruttura, cioè quello in primis...nel più semplice dei casi è una matrice casi per variabili e già quello è un disegno architettonico di tipo infrastrutturale...quindi lo faccio rientrare lì, però lo racchiudo in una rete più ampia*” (**RIU**).

dovunque, con grande difficoltà, e la speranza di trovarli in qualche modo già preorganizzati e più user-friendly e facili da individuare” (R3D). Due caratteristiche delle “nuove infrastrutture”, che appaiono rilevanti nelle combinazioni definitorie, sono la possibilità di interoperabilità (“Credo che sia un sistema che mette in rete più risorse tecnologiche al fine di produrre conoscenza, con un utilizzo di tecnologie avanzate, e ricorso anche a big data e a fonti di dati, che prima non erano accessibili” – R2U), e l’impossibilità di separazione con componenti fisiche e hardware (“le infrastrutture di ricerca secondo me sono un’unione tra la capacità di accedere a database digitali, [...] come per esempio derivanti da survey ripetute nel tempo oppure da grossi database internazionali con survey fatte a livello internazionale e comparate, insieme alla struttura fisica con spazi, postazioni e tutte le apparecchiature necessarie per accedere a questi database, sostanzialmente” – R1U). Si identifica dunque un passaggio da infrastrutture “tradizionali” a infrastrutture “moderne” che ha consentito di svolgere il lavoro di ricerca in modo diverso rispetto al passato (“per me [queste] sono le infrastrutture: cioè quello che mi permette di accedere molto velocemente a documenti, a materiali a letteratura che altrimenti ci metterei una vita a cercarla” – R3D). Tale implicazione si traduce in una ricorrente caratterizzazione rispetto alla facilitazione nell’attività di ricerca (“tutto ciò che possa essere d’ausilio al ricercatore per poter svolgere attività di ricerca o per esempio anche per poter confrontare i risultati della propria attività di ricerca” – R1D). Una parola in particolare, il termine inglese “scaffolding”<sup>64</sup>, è emersa da un’intervista ad un professore associato e sintetizza efficacemente la visione dell’infrastruttura come elemento di sostegno alla ricerca, “un modo di organizzare la conoscenza e renderla fruibile, disponibile, utilizzabile, condivisibile” (R2D).

Il sostegno garantito dalle infrastrutture di ricerca si sostanzia nell’accesso a fonti “di seconda mano”, ossia non prodotte direttamente dal ricercatore, ma alle quali può accedere per supportare i propri percorsi di ricerca *in fieri* (ma anche per costruirli in una fase d’inizio del disegno della ricerca). Allo stesso modo, può coincidere con lo sfruttamento di fonti derivate da raccolta originale sviluppata attraverso lavoro collaborativo, come ad esempio nell’ambito di un progetto al quale il ricercatore partecipa in quanto responsabile dell’infrastruttura o in quanto partner. La tipologia di contenuto di tali fonti è percepita nella sua potenziale diversificazione: “conservano al loro interno una serie di informazioni che

---

<sup>64</sup> Il termine identifica le impalcature di sostegno nell’ambito della costruzione degli edifici.

*possono essere di carattere sia qualitativo che quantitativo; quindi, sia dati che metadati, che anche informazioni documentali, quindi documenti, prodotti scientifici” – R1D).*

Un *trait d’union* che è ricorso più volte tra le varie definizioni è rappresentato dal fatto che le infrastrutture di ricerca possono fungere non solo da fonti per i dati di ricerca, ma assurgere anche a veicoli di relazione (“[...] *dovrebbero permettere anche un contatto vero e proprio tra i ricercatori, cioè non solo quindi un livello di scambio di informazioni intese come dati, ma anche uno scambio di informazioni inteso come dialogo su prodotti o su risultati della ricerca*” - R1D). L’idea dell’infrastruttura come risorsa realizzatrice di istanze di condivisione e di abilitazione di reti potenziali di collegamento, fra ricercatori e fra diversi lavori di diversi ricercatori, è riconosciuta da un professore associato come l’attualizzazione di un’idea, quella della Scienza come impresa comune, che è tipica della Sociologia della scienza e soprattutto di Merton, “*vale a dire l’idea delle diverse forme in cui un [prodotto della ricerca]... cioè qualche cosa su cui stiamo lavorando [è reso pubblico] non nella sua forma tradizionale e standardizzata, all’interno ad esempio di una collocazione editoriale. In questo senso, l’idea di un’infrastruttura di ricerca è qualcosa che non solo attraversa le comunità scientifiche, [...], ma è in grado di collegare tra di loro sottoinsiemi e sottostrutture delle medesime comunità scientifiche che rischiano di rimanere isolate da un lavoro collettivo, quale dovrebbe essere la scienza*” (R2U). Tale immagine che riassume reperibilità del dato, sforzo collaborativo e possibilità di lavorare da remoto richiama, come rilevato da un professore associato, le caratteristiche del *cloud*: “*cioè [le infrastrutture di ricerca sono] spazio di conservazione, condivisione e lavoro allo stesso tempo*” (R2U), non scatole chiuse pronte ad essere usate e messe da parte, ma risorse disponibili e comuni che rappresentano configurazioni in allestimento continuo.

I testimoni privilegiati, portatori di esperienze di tipo “avanzato”, in quanto coinvolti nella creazione di infrastrutture a livello nazionale ed europeo, hanno fornito definizioni che si attagliano maggiormente a una visione delle infrastrutture di ricerca in quanto oggetti specifici, come i *database*. Particolarmente suggestiva è l’immagine del “*punto di cristallizzazione*” della conoscenza che ne veicola lo scambio, fornita da uno dei tre testimoni: per questo ricercatore l’infrastruttura si qualifica primariamente come ambiente *abilitante* di tipo fisico (come un laboratorio per attuare simulazioni) o digitale (un *database* di ricerca) che promuove e coltiva informazioni e risultati di esperimenti o studi e si pone come base per la discussione scientifica e le collaborazioni. Di uguale avviso è un altro testimone: “*le*

*infrastrutture rappresentano per il ricercatore la possibilità di accrescere prima di tutto la sua capacità di conoscenza e poi le proprie relazioni all'interno della comunità di riferimento” e possono coincidere con “sistemi informativi di vario tipo”, non necessariamente digitali, che contengano non solo “dati numerici, ma [...] pure dati di tipo qualitativo”. Il terzo testimone opta per una definizione più sostanziale, puntando sulla declinazione del concetto entro il settore delle scienze sociali: “le infrastrutture di ricerca nel mio settore sono banche dati sostanzialmente”, risorse poliedriche rispetto a molteplici usi di ricerca che implicano una professionalizzazione della gestione e possiedono “caratteristiche di accessibilità tecnica, di standardizzazione che sono diverse o superiori ai dati che uno può avere sul computer”.*

## **4.2 Tipi di infrastrutture utilizzate dagli intervistati**

Il processo di selezione degli intervistati, oltre a prevedere l’inclusione di ricercatori sociali giunti a stadi diversi della propria carriera (ricercatori junior / assegnisti di ricerca, professori associati / ricercatori di enti di ricerca pubblici, professori ordinari / dirigenti di ricerca), con età differenti e con una rappresentazione di genere proporzionata, ha tenuto conto dei relativi *curriculum vitae*, disponibili sui siti Internet delle strutture di afferenza, al fine di ricercare elementi tali da far desumere una propensione all’utilizzo o alla creazione di infrastrutture di ricerca. A tale scopo, si è posta particolare attenzione su elementi di carriera quali pubblicazioni, reportistica, atti di conferenza, partecipazione a progetti di ricerca nazionali e internazionali. Durante la fase iniziale dell’intervista, si è dunque cercato conferma alle indicazioni emerse da tale analisi curriculare attraverso la domanda-filtro “*Lei usa o ha usato infrastrutture di ricerca nella sua attività di ricerca?*”.

Alla domanda “*Lei usa o ha usato infrastrutture di ricerca nella sua attività di ricerca?*” tutti gli intervistati hanno risposto indicando l’utilizzo in carriera di almeno un’infrastruttura di ricerca (generalmente, il riferimento è stato a 2 o 3 infrastrutture per intervista). Un punto interessante è emerso relativamente alla “consapevolezza” di aver utilizzato o creato un’infrastruttura di ricerca. Appare a questo proposito particolarmente interessante riportare tre testimonianze che fanno evincere come il legame tra il concetto di infrastruttura di ricerca e l’oggetto al quale esso si riferisce non sia talvolta così automatico: un’intervistata, dopo una

riflessione posteriore alla domanda-filtro, ha rilevato di essersi resa conto solo in un secondo tempo che “*di fatto quella a cui avevo accesso poteva essere definita un'infrastruttura di ricerca, quindi un insieme di informazioni, dati, anche strumenti che mi consentivano di sviluppare l'attività di ricerca che avevo in mente [...] L'ho usata senza saperlo*” (R2D); una seconda intervistata ha manifestato qualche titubanza nel rispondere positivamente alla domanda-filtro: “*Sì...adesso parlandone mi è venuto in mente [...] Infatti non so come non mi è venuto in mente prima...Su Eurobarometro ci lavoro sistematicamente, soprattutto su questo discorso della partecipazione civica*” (R3D); infine, un professore ordinario, dopo aver risposto negativamente alla domanda-filtro, ha successivamente citato l'utilizzo di dati tratti da *database* internazionali della World Bank per uno studio metodologico.

Quanto ai tipi di infrastrutture utilizzate, più della metà degli intervistati ha indicato l'utilizzo, per almeno uno studio (a fine di sfruttamento di microdati, di costruzione di quadro di riferimento o di pura consultazione), di *infrastrutture di ricerca di interesse europeo o caratura internazionale* (cfr. par. 2.4.1), il cui scopo primario è la fornitura di dati di ricerca per utenti terzi rispetto ai creatori dell'infrastruttura. Tra le infrastrutture di interesse europeo, l'infrastruttura ESFRI Eurobarometro è risultata essere la più utilizzata, mentre la European Social Survey e la European Value Survey sono state citate solo una volta ciascuna; tra quelle di caratura internazionale, due infrastrutture derivanti da indagini OCSE, quali OCSE-PISA e OCSE-PIAAC, sono state utilizzate da due intervistati. L'utilizzo di tali risorse si è tuttavia configurato come un'opzione minoritaria se si considera che in 11 occasioni, gli intervistati hanno fatto riferimento allo sfruttamento di dati da uno o più *database creati “su misura”* da una *partnership* di ricercatori per rispondere a domande di ricerca specifiche (*project-driven*) nell'ambito di un progetto di ricerca e destinati, almeno inizialmente, ad uso dei creatori primari. Durante le interviste si sono approfonditi in particolare 4 progetti che hanno generato infrastrutture di ricerca che avevano dimensione europea (ad es. finanziati dal Settimo Programma Quadro), 2 dimensione nazionale (due PRIN) e 5 sviluppati all'interno di dipartimento. Quanto al sottoinsieme delle infrastrutture ascrivibili agli archivi di dati forniti attraverso piattaforme di consultazione da istituzioni riconosciute a livello nazionale, vi è stato un riferimento ad esse in 3 occasioni: a questo proposito sono state citate le banche dati Cineca, Censis, Invalsi. A livello internazionale l'utilizzo di dati World Bank è stato

riscontrato 2 volte, mentre l'accesso allo specialistico *database* eCORDA<sup>65</sup> è stato citato una volta. L'utilizzo di *database* legati alla produzione di statistiche ufficiali fornite a livello nazionale da Istat, a livello europeo da EUROSTAT è ricorso rispettivamente 5 e 3 volte, spesso a fini di alimentazione di infrastrutture *project-driven*, e quando non legato a progetti è sempre rimasto comunque escluso dalle narrative sugli *output* di ricerca. Infine, archivi di dati amministrativi sono stati sfruttati in due occasioni, con riferimento agli archivi Sapienza, ma la loro rilavorazione ha generato nuove infrastrutture nell'ambito di 2 dei 5 progetti dipartimentali citati sopra. In un solo caso si è fatto riferimento a un'infrastruttura, disponibile a pagamento e messa a disposizione da un'azienda privata. Si noti che naturalmente tra categorie sopracitate vi è *sovrapposizione*, nel senso che, ad esempio, chi ha utilizzato un'infrastruttura di interesse europeo ha talvolta riferito di aver utilizzato durante la sua carriera anche un'infrastruttura nata da progetto. Sulle infrastrutture appartenenti a queste due categorie si è basata la quasi totalità delle narrative sugli *output* di ricerca (*cf.* par. 3.3).

Gli intervistati che hanno fatto riferimento allo sfruttamento di infrastrutture nate da progetto, quasi sempre hanno rivestito il duplice ruolo di utente e co-creatore (*“a questa infrastruttura ho partecipato non solo come utente/utilizzatrice ma anche come responsabile della sua costruzione nella fase iniziale”* – R2D), partecipando alla progettazione e all'alimentazione della risorsa. Nel caso dell'infrastruttura nata da un PRIN che ha coinvolto 4 università italiane, un professore associato ha dichiarato che la costruzione *ex-novo* dell'infrastruttura si poneva come elemento indispensabile per portare avanti i propri interessi di ricerca, in assenza di altre infrastrutture a livello europeo che potessero soddisfare il bisogno informativo: *“La mia predilezione è per la produzione di nuove infrastrutture perché di solito mi muovo con degli interrogativi di ricerca rispetto ai quali non esistono già delle infrastrutture preesistenti che possano dare delle risposte a quegli interrogativi”* (R2U). Nel caso dell'utilizzo di infrastrutture di tipo europeo, i ricercatori che ne hanno fruito hanno agito nel solo ruolo di utilizzatori, con l'eccezione di un ricercatore *junior* che ha dichiarato di aver in aggiunta fornito un contributo all'alimentazione della risorsa, in quanto presente nel *board* dell'indagine ad essa connessa (*“facciamo i trial, proviamo i vari item e vediamo come funzionano”* – R1U). Infine, coloro che hanno dichiarato di non aver utilizzato infrastrutture di interesse europeo hanno evidenziato da un lato la poca informazione su questo tipo di

---

<sup>65</sup> eCORDA (COMmon REsearch DATA warehouse) contiene dati sulle proposte e sui progetti finanziati nell'ambito dei Programmi Quadro europei FP6, FP7 e H2020.

risorse, dall'altro la mancanza di interesse verso tali risorse informative (*“Non c'è una ragione, nel senso che non è l'esito di un processo decisionale per cui mi sono detto uso questi e non uso altri. Evidentemente per motivi non voluti sono rimasti fuori dal mio modo di lavorare. [...] È proprio mancata l'occasione, nel senso che li conosco poco”* (R2U).

Le caratteristiche d'accesso alle infrastrutture utilizzate hanno riflesso un'accentuata variabilità. Le infrastrutture di tipo europeo utilizzate e le banche dati degli istituti di statistica nazionali o europee sono ad accesso libero o ad accesso subordinato a registrazione gratuita. Solo per l'infrastruttura eCORDA è stato necessario ricevere un'autorizzazione, posteriore all'invio di un progetto di ricerca, che consentisse di giustificare la richiesta di accesso. Le infrastrutture generate da progetto citate dagli intervistati sono invece tutte ad accesso limitato alla *partnership* che ha dato vita al progetto. Nel caso dell'infrastruttura generata dal PRIN, la remora principale riguardo l'apertura all'esterno è derivata dal fatto che *“l'infrastruttura esistente andrebbe ampliata a livello di apertura geografica e andrebbe aggiornata rispetto a una wave successiva della stessa indagine e in quel caso sarebbe un'infrastruttura che avrebbe senso aprire”*. Nel caso delle infrastrutture nate da progetto europeo, l'apertura è condizionata dalle regole imposte dal committente finanziatore e, se avviene, è sempre posteriore ad un periodo di “embargo”: *“Dunque al momento no [non è prevista un'apertura], perché come lei sa la Commissione Europea è proprietaria di tutto quello che uno fa e quindi al di là di Bruxelles non esce. Noi ci auguriamo che sia “sì” e che al momento in cui il progetto sarà finito e questo sarà uno dei risultati del progetto, questa possa essere condivisa e seguire la tendenza generale all'open access”* (R3D). Sfruttando questa circostanza i ricercatori hanno talvolta potuto godere di una posizione di vantaggio nello sfruttamento dei dati nella struttura di competizione con gli altri ricercatori non coinvolti nel progetto.

L'orientamento generale degli intervistati si è rivelato favorevole all'*open access* e quindi alla divulgazione dei dati derivanti da proprie ricerche e da propri progetti (*“questo è comunismo mertoniano”*. *Se la scienza è un'impresa aperta e democratica, chiunque da qualunque parte del mondo può mettere il naso nelle cose che faccio io, sia le infrastrutture di ricerca sia i risultati”* – R3U), sebbene in nessuna occasione tale orientamento si sia tradotto nei fatti nelle loro narrative. In controtendenza, una ricercatrice ha fatto tuttavia presente che a suo avviso per le scienze sociali non c'è ancora la propensione a *“ragionare in termini di infrastruttura di ricerca che presuppone la condivisione, la sistematizzazione, la messa a disposizione del dato con criteri di accesso e livelli di libertà d'accesso che possono*

*essere diversi” (R2D). Sulla stessa scia, una professoressa associata ha denunciato lo scarso interesse alla condivisione del dato da parte di alcuni colleghi, citando ad esempio di ciò il mancato conferimento dei dati in un’infrastruttura da lei creata: “Sono stata promotrice di un altro progetto che si chiamava \*\*\*, che aveva l’obiettivo di mettere in condivisione risorse, soprattutto dati, software, esperienze di ricerca sull’uso di metodi misti, quantitativi e qualitativi insieme. Anche questo è stato un progetto che è morto sul nascere nel senso che...non abbiamo potuto creare questo repository proprio perché le persone che erano nel progetto non avevano interesse ad alimentarlo. [...] l’idea di condivisione è anche di utilizzare il proprio tempo per alimentare una risorsa da condividere, non è un’idea che trova successo in tante persone. Perché molti ricercatori sono abituati ad avere un obiettivo individuale da perseguire senza perdere tempo in altre cose” (R2D). Una leva che può incentivare i ricercatori alla condivisione del dato e al mantenimento dell’infrastruttura potrebbe essere quella monetaria, che lega l’infrastruttura a una progettualità definita e a un finanziamento costante: “Spesso questo avviene e si fa perché c’è un progetto di una certa portata e rilevanza che lo consente, [...] quindi è su come strutturare un po’ anche il finanziamento della ricerca in questi campi che andrebbe considerato per dare un corso a questo sviluppo”.*

#### *4.2.1 Riflessi sulla produzione scientifica generale*

I testimoni privilegiati di questo studio hanno dichiarato che l’impatto della disponibilità di infrastrutture di ricerca sulla loro produzione scientifica è stato molto forte. Uno di loro si è spinto a dichiarare che *“una metà della mia produzione scientifica è legata ad attività infrastrutturali. L’impatto è molto forte, sì”*. Un altro testimone ha addirittura fondato, insieme ad alcuni colleghi, una rivista su cui è stata riversata la mole di lavoro scaturito dallo sfruttamento di un’infrastruttura da lui creata (*“Sì, noi avevamo una rivista che si chiamava \*\*\*. Abbiamo riversato quel tipo di esperienza in modi classici, cioè attraverso delle pubblicazioni”*).

In generale, è emerso un impatto sempre crescente nel tempo, sebbene non affatto totalizzante, sulla produzione scientifica degli intervistati a tutti i livelli di carriera, spesso legata allo sfruttamento di infrastrutture derivanti da progetto, con sole due eccezioni di due professori ordinari che hanno ridimensionato il fenomeno: una donna (*“nel mio caso è una cosa abbastanza recente. Nel senso che non è che sono molti anni che ci lavoro su queste*

*cose” – R3D) e un uomo (“[ho usato le infrastrutture di ricerca] come quadro generale di riferimento su cui costruire un lavoro mio” – R3U). In particolare tra i ricercatori più giovani, l’impatto è stato indubbiamente percepito come rilevante, anche in ragione del fatto che la scarsità di risorse finanziarie per il settore della ricerca (più volte denunciata) si è tradotta nella necessità di sfruttare infrastrutture precostituite o create all’interno di progetti finanziati. Grazie al supporto delle infrastrutture sono stati derivati almeno due *output* di ricerca a testa tra articoli su riviste italiane o internazionali, paragrafi di libri e report.*

Tutti gli intervistati sono stati invitati a scegliere tra gli *output* prodotti da infrastrutture uno o più di uno su cui declinare gli effetti sul lavoro di ricerca che verranno esaminati nei prossimi paragrafi. Dalla narrativa sui prodotti selezionati è emerso che le infrastrutture derivate da progetto o da archivi istituzionali hanno sempre generato studi di tipo analitico che hanno sfruttato i microdati in esse contenuti e in un caso hanno prodotto un articolo metodologico sulla costruzione dell’infrastruttura<sup>66</sup>; mentre le infrastrutture di tipo europeo, oltre ad un utilizzo analitico dei dati, riscontrabile nel lavoro di soli 3 intervistati, sono spesso usate come supporto di sfondo per motivi di semplice consultazione (“*ci serviva come confronto anche per confermare la solidità del nostro dato, quindi dei risultati emersi. Quindi l’infrastruttura [Eurobarometro] è stata un supporto... però non solo per quanto riguarda il raffronto dei dati e la raccolta dei dati ma anche a livello metodologico per la definizione dell’attività di ricerca” – R1D*), test di procedure metodologiche (“*uso un’infrastruttura per testare una procedura, questa è una cosa che se vuole sperimentarla non serve a niente, però è pure una parte del lavoro del ricercatore. A me è capitato tante volte di progettare procedure e quindi non c’era il render conto del fenomeno” – R3U*) o combinando un interesse metodologico con un interesse più sostanziale per il dato (“*considerando la banca dati OCSE-PISA come un esempio di infrastruttura, quindi l’abbiamo analizzata dal punto di vista storico di come la banca è stata costruita e poi anche utilizzando i dati della banca dati*”. – R2D). Altro utilizzo delle infrastrutture di tipo europeo, slegato dall’utilizzo a fine analitico dei microdati, riguarda l’uso didattico, come emerso dalla testimonianza di un professore associato (“*Io, ad esempio, altri database di questi li ho usati molto di più per la didattica. Li ho usati quasi più per la didattica che per la ricerca” – R2U*).

---

<sup>66</sup> Un professore associato ha dichiarato che in sede di pubblicazioni si è mosso su un duplice binario: “*Da un lato ho prodotto risultati di ricerca e da un lato via via che poi ho svolto anche la ricerca sul \*\*\* ho presentato anche articoli internazionali sui mixed methods, su quest’approccio multilivello integrato [su cui è basata la metodologia di costruzione dell’infrastruttura]” (R2U).*

Rispetto alle infrastrutture derivate da progetto, i ricercatori hanno paventato uno sfruttamento ripetuto (*“Penso di sfruttarla ancora per la generazione di altri output; è un'infrastruttura alla quale periodicamente accedo sempre”*, **R3D**), sebbene solo in un caso abbiano dichiarato di aver prodotto più di un articolo legato alla medesima infrastruttura.

### **4.3 Gli effetti dell'utilizzo delle infrastrutture sull'attività di ricerca**

Uno degli interrogativi su cui questo studio ha inteso soffermarsi ha riguardato gli effetti nel processo di costruzione della conoscenza dovuti all'utilizzo di infrastrutture di ricerca da parte degli scienziati sociali, con particolare riferimento al disegno di ricerca, alle dinamiche aggregative, alla tendenza all'interdisciplinarietà della ricerca. Prima di affrontare separatamente queste tre dimensioni appare interessante approfondire un tema di ordine più generale, ossia come gli avanzamenti in ambito ICT e la parallela proliferazione di risorse informative *online* abbiano prodotto, in una prima valutazione degli intervistati, effetti nel modo di accesso all'informazione e al dato.

#### *4.3.1 Cambiamento tecnologico e accesso alle risorse informative*

Come già rilevato (*cf.* par. 1.2), in particolare nell'arco degli ultimi trent'anni, le attività legate alla produzione di conoscenza sono state investite da sempre più pressanti processi di cambiamento tecnologico che hanno provocato nel lavoro dei ricercatori alcuni effetti sia rispetto all'accesso alle risorse informative che in riferimento alle pratiche quotidiane di ricerca. Il rapporto tra tecnologie informatiche e scienza ha iniziato ad assumere forme peculiari allorché l'evoluzione della Rete ha prodotto implicazioni inedite inerenti allo sfruttamento delle potenzialità e delle caratteristiche dell'ambiente virtuale, come la possibilità di ricercare e condividere dati e prodotti di ricerca e di comunicare o scambiare dati in tempo reale a grandi distanze.

Considerando le differenti età degli intervistati e il diverso tipo di esperienza maturata da ognuno di essi, è stato possibile riconoscere il riferimento a tre modalità di circolazione della conoscenza, cronologicamente susseguenti, ma tuttora compresenti e non necessariamente sostitutivi: *modalità tradizionali di circolazione della conoscenza* (es. produzione di conoscenza basata su biblioteche, pubblicazioni cartacee, conferenze,

*workshop*); *web-based data* (reperimento di fonti e dati in Rete non sistematizzati); *disponibilità di infrastrutture di ricerca con dati e indicatori sistematizzati* (sfruttamento di risorse ad accesso libero o controllato su siti *web* o disponibili in *cloud* condivisi). Le tre tipologie di accesso all'informazione sono apparse convivere in distribuzioni variabili nell'attività dei ricercatori più anziani, mentre i ricercatori più giovani quasi propendono a tralasciare la prima modalità a favore delle più recenti.

Dalle parole, connotate da una certa nostalgia, di una delle intervistate con maggiore esperienza sembra di percepire nettamente un "prima" e un "dopo" nel modo di fare ricerca dovuto all'irruzione di Internet e alla disponibilità di risorse informative virtuali nel lavoro di ricerca: *"Io non sono una persona giovane e Internet è veramente stata una rivoluzione nel nostro lavoro, nel mio lavoro, perché, abituata ad andare in biblioteca – che devo dirle un po' rimpiango però –, avere la possibilità di fare tutto attraverso il computer per la mia generazione è una grande svolta epocale, decisamente grossa"* (R3D). Conseguenze concrete della "svolta epocale" sono rappresentate dalla progressiva perdita del rapporto con la carta, sebbene non del tutto scomparso (*"io sono cresciuta coi volumi dell'Istat, ce li ho ancora a casa, volumi con le serie storiche dell'Istat, in cui si lavorava sul cartaceo prevalentemente, rifacendosi le tabelle [...] noi copiavamo i dati materialmente perché non era possibile fare altrimenti"*) e dal superamento del trattamento manuale del dato (*"[...] lo dovevi copiare, il dato, oppure lo registravi"*)<sup>67</sup>. Al contrario, l'approccio moderno alla produzione scientifica si traduce nel lavoro dell'intervistata sempre più nello sfruttamento delle risorse bibliografiche digitali *online* che contengono i più svariati prodotti di ricerca (*"oggi, io attraverso Google Scholar e attraverso la sede dell'università io posso accedere ad una marea di riviste, di volumi"*) e delle infrastrutture di ricerca presenti in Rete o scambiate tra ricercatori, quali i *database* condivisi, che consentono una sistematizzazione precostituita del dato con conseguente abbreviazione dei tempi di ricerca (*"oggi tu hai dall'altra parte la possibilità di velocizzare e di avere di fronte a te il dato che al massimo sarà dell'anno prima,*

---

<sup>67</sup> L'abbondono di modalità analogiche di consultazione ed elaborazione del dato in favore dell'accesso a *database* e dell'analisi tramite *software* è emerso anche dalla testimonianza di una professoressa associata: *"i miei maestri al tempo mi facevano lavorare su fonti documentali cartacee anche nella costruzione di dati con matrici cartacee da costruire. Addirittura ricordo delle matrici forate sulle quali ho studiato e che provenivano dall'Istat e le forature corrispondevano alle stringhe di quei documenti a casi rilevati per una tipologia. Quindi ad esempio oggi se potessi rielaborare quei dati ad applicazioni per la lettura di quei dati io posso ritornare ad utilizzare quelle fonti facendolo però con strumenti che mi ottimizzano tempi e anche risultati output definitivi finali con l'applicazione tecnologica specifica. Quindi quello che il ricercatore tradizionale poteva fare semplicemente rileggendo il testo e catalogandolo a penna con alcuni criteri di selezione di concetti specifici, oggi questi software consentono di farlo con un ampio raggio di scelta"* (R2D).

*se sei in arretrato, con tutta una serie di aggiornamenti che ti vengono dati e subito sfruttabili...hai possibilità di scrivere soprattutto alle fonti internazionali”*). A questo proposito, un professore ordinario ha segnalato che, nella sua pur limitata esperienza nell’uso di grandi *database*, ha rilevato nel tempo un salto di livello sia sul fronte dell’apertura e proliferazione del dato, due fenomeni che si sono riverberati in una progressiva appropriazione comune del dato di ricerca da parte della comunità scientifica (“[Fino agli anni Novanta] *si creava quello che diceva Alain Touraine... delle persone proprietarie della conoscenza ed altri esclusi*”), sia una facilità d’uso delle banche dati non paragonabile rispetto al passato (“*i sistemi informativi sono diventati più elastici, e hanno ammesso quindi il fatto che uno possa fare delle queries più specifiche per ottenere dei dati filtrati in un certo modo [...] Voglio dire, i dati fino agli anni ‘70 erano utilizzabili secondo la logica di un antico pacchetto che era nato per la gestione dei magazzini, e come si gestivano i magazzini si gestivano i dati, in maniera non interattiva. Adesso la cosa è molto più flessibile per cui io posso chiedere (però mi serve il concorso dei responsabili dei dati) dei dati filtrati in maniera molto più articolata, dettagliata*” (R3U).

Sebbene l’utilizzo delle risorse *online* sia ormai divenuto attività prevalente nelle pratiche quotidiane di ricerca, un dirigente di ricerca ha evidenziato, in una prima valutazione nel corso dell’intervista, una relativa continuità nel tempo nel proprio modo di fare ricerca (sottintendendo la parte teoretica della ricerca, i metodi e le tecniche), sottolineando come le nuove disponibilità informatiche abbiano comportato unicamente mutamenti di *forma*: “*perché la ricerca è sempre la stessa, ma gli strumenti con cui la fai sono diversi per cui la sostanza non si modifica, ma si modifica infinitamente la forma, quindi la facilità d’accesso, la facilità di reperimento dei dati o dei libri. La forma si può modificare, si può avere un’accelerazione, però i parametri con cui fai ricerca quelli sono. Poi se ci metti una settimana perché devi andare in biblioteca fisicamente a cercare il libro oppure te lo trovi sul computer in due secondi [non cambia] però la sostanza della tipologia di ricerca, non si modifica*” (R3D). Sulla medesima scia si è posto un professore associato che, non rilevando cambiamenti rivoluzionari rispetto al proprio modo di fare ricerca imputabili alla proliferazione di risorse informative e all’evoluzione di Internet (“*non è un cambiamento drammatico, strutturale perché continuo a fare anche un lavoro che ho fatto di più e che continuerò a fare, più di tipo di analisi di letteratura. Il mio “prima” continuerà a sopravvivere secondo me*”), ha tuttavia notato possibilità di “*arricchimento e non di*

sostituzione” rispetto alla propria pratica di lavoro, nei termini di una maggiore propensione a comportamenti utilitaristici percepiti in termini negativi (“*si combinano opportunità ma anche dei cambiamenti in negativo, nel senso che è un modo per affrontare una fase di risorse scarse. L'accesso a infrastrutture che esistono indipendentemente da me è un modo per fare ricerca spendendo di meno*” – **R2U**). Una professoressa associata ha, al contrario, messo in evidenza due ordini di cambiamenti generali nella sua attività di ricerca; uno legato alla Rete quale mediatore d’accesso ad una quantità di dati impensabile in passato (per cui “*in qualche modo si impone per noi ricercatori un’opera di discernimento, tu devi sapere che cavolo di dati stai utilizzando*” – **R2D**), e uno legato ad un marcato *shift* di approccio alla ricerca reso possibile dalla disponibilità di infrastrutture di ricerca, col passaggio da studi qualitativi ad una propensione ad analisi quantitative: “*Io nasco come sociologa qualitativa, quindi interviste, etnografie, osservazioni, ecc. Ma negli ultimi dieci anni mi sono appassionata dei numeri, delle quantità, considerandoli, come dicevo prima, come degli effetti sociali rilevanti, quindi anche criticamente sostituiti di teorie, anche dal punto di vista delle grandi teorie educative [...] oggi io prendo OCSE-PISA per studiare la comparazione dei sistemi educativi*”.

I ricercatori più giovani hanno sottolineato come l’incipit della loro carriera sia stato da subito investito dalla rivoluzione nella soddisfazione del bisogno informativo che ha caratterizzato il settore della ricerca, grazie alla reperibilità *online* del dato e alla proliferazione delle infrastrutture di ricerca (in particolare uno di loro ha affermato “*la mia carriera è stata basata su questo [...] ottenere le stesse informazioni o informazioni abbastanza vicine a quello che era il bisogno informativo in maniera facile e veloce e con numeri che una volta non erano possibili*”)<sup>68</sup>. Tale circostanza ha, in un certo senso, controbilanciato la difficoltà dei giovani ricercatori a reperire dati primari dovuta all’indisponibilità di fondi di ricerca personali, sebbene la gratuità delle risorse informative *online* non sia scontata: “*Effettivamente una buona parte del lavoro è quella di trovare un’infrastruttura che ti permette di aver accesso ai dati acquisiti nell’ambito di programmi di ricerca europei oppure nelle università, dove ci sono e sono un po’ dappertutto. Il punto cruciale è in realtà nella disponibilità del dato perché non tutte le università rendono fruibile il dato o lo rendono disponibile gratuitamente. E questa è una forte limitazione perché un*

---

<sup>68</sup> Un’altra giovane ricercatrice ha confermato che l’importanza nell’utilizzo di infrastrutture di ricerca era stata promossa già durante gli anni dell’università: “*sono stata subito indirizzata nell’importanza di dover a mia volta utilizzare questo tipo di informazioni*” (**R1D**).

*giovane ricercatore, nel nostro caso, non ha fondi di ricerca suoi e quindi è difficile avere accesso... anche a buoni dataset o buone infrastrutture che però non sono gratis, sostanzialmente” (R1U). Il tema della scarsità delle risorse finanziarie destinato al settore della ricerca pubblica, e il suo legame col cambiamento del modo di lavorare dei ricercatori, è ripreso da un professore associato, che guarda favorevolmente alle opportunità offerte dal web: “Quindi ha cambiato il mio modo di lavorare? Assolutamente sì! [...] si ha la possibilità non solo di accorciare un pezzo di lavoro, ma anche di fare fronte ad una scarsità di risorse a cui tutti siamo obbligati. Quindi, siccome ho scarsità di risorse, siccome non potrò mai rifare una ricerca come hanno fatto quelli con quella disponibilità di personale tecnico, di collegamenti, di reti di ricerca, allora utilizzo questa strategia, quindi provo a prendere il meglio rispetto a quanto le nuove possibilità del web mi consentono di fare... perché tutto questo è possibile grazie al web...” (R2U).*

Un aspetto più volte evidenziato dagli intervistati ha riguardato il risparmio di tempo rispetto alla collezione dei materiali di supporto alla ricerca. Una ricercatrice ha, a questo proposito, sottolineato come *“l'utilizzo di un accesso unico dal mio computer, [...], senza più dovermi recare in biblioteca, mi ha consentito innanzitutto di risparmiare tempo”*, concordando tuttavia sul fatto che l'effettivo cambiamento che è possibile riconoscere nell'utilizzo di ICT si delinea rispetto alle opportunità *“che mi hanno dato l'utilizzo di questi strumenti per trovare e accedere alle informazioni e alla facilità e alla ricchezza di accesso a queste informazioni. Se ripenso al periodo del Dottorato in cui per la mia attività di tesi andavo ad analizzare la partecipazione italiana ai programmi quadro e tutta la mia attività iniziale di ricerca di desk ha riguardato la partecipazione delle università italiane e tutto questo è stato fatto analizzando delle schede di Cordis, [...] dall'altra parte sia per quanto riguarda \*\*\*, sia \*\*\*, io ho potuto accedere ai dati ed informazioni sulle partecipazioni, in termini di numerosità alla partecipazione, finanziamenti ottenuti, tipologia di ruolo... con un unico clic. Quindi il risparmio è stato notevole in termini di tempo” (R2D).*

Interessante la notazione di un testimone privilegiato che, sottolineando come la sua attività di ricerca si sia maggiormente declinata verso un approccio quantitativo parallelamente al suo sempre maggiore utilizzo di infrastrutture di ricerca (poiché per *“avere dei sample grandi significa che posso usare di più tecniche statistiche perché il numero di osservazioni è tale che un'analisi statistica funziona”*), ha rilevato un cambiamento nella sua capacità di gestione dei dati, dovuta alla presenza in remoto di infrastrutture di ricerca: *“il*

*mio modo di gestire i dati è estremamente cambiato. È diventato meno artigianale, nel senso che prima uno aveva il suo dataset, se lo teneva, se lo aggiornava, la copia era unica non c'era controllo del versioning. Adesso non c'è più: io uso l'infrastruttura. [...] Prima io avevo un file Excel coi dati delle università, poi ogni tanto aggiungevo un dato, ci pasticciavo e così via. I dati so che sono tenuti dall'infrastruttura, vado ogni volta ad estrarre i dati che mi servono, mi creo un database ad hoc. Quindi io ho trenta copie di \*\*\* per i diversi paper che sto scrivendo [...] Evito di fare data storage, evito il data handling e faccio delle estrazioni che servono esattamente per quella cosa".* Un altro testimone privilegiato ha infine evidenziato il minore utilizzo di fonti cartacee in favore del reperimento di fonti sulla Rete, che da una parte *"implica a volte l'uso di un controllo di qualità minore sulle fonti"*, ma dall'altra consente tempi di accesso più veloci (*"quindi hai molto più rapidamente e molto più ampio accesso ai dati e alle informazioni e ai diversi tipi di fonti"*).

#### *4.3.2 Nuove impostazioni del disegno della ricerca*

Il compimento di ogni sforzo di produzione di nuova conoscenza dipende da un complesso di azioni poste in essere dal ricercatore, tra le quali figurano la formulazione del problema d'indagine (o domanda di ricerca); la progettazione e costruzione della base empirica (per mezzo di collezione originale o sfruttando fonti secondarie); il trattamento, l'elaborazione e l'analisi dei dati acquisiti. Tali fasi, oggetto di decisioni puntuali del conduttore dell'indagine, vanno a costituire il cosiddetto "disegno della ricerca", e orientano l'implementazione del lavoro scientifico. Si è ritenuto dunque utile esplorare, nelle situazioni in cui i ricercatori sociali hanno utilizzato infrastrutture di ricerca, quali fossero le ripercussioni rispetto al percorso che ha guidato l'attuazione concreta dei loro studi, stimolando il riferimento a episodi tratti dalla loro esperienza diretta. Le testimonianze degli intervistati possono essere sintetizzate in alcuni punti generali tra i quali la propensione alla ricerca di fonti di dati pre-organizzati, la tendenza alla ridefinizione in *itinere* della domanda e degli obiettivi di ricerca e l'attenzione verso le procedure di trattamento del dato preliminari all'analisi.

Il primo dato generale, che è emerso in più di un'occasione dall'analisi delle interviste (in particolare a ricercatori che hanno fatto uso di infrastrutture già precostituite), riguarda un approccio alla progettazione della ricerca, maggiormente orientato alla verifica della disponibilità di fonti preorganizzate rispetto alla tematica da approfondire: *"Oggi qualunque*

ricercatore si ponga una domanda di ricerca, il momento dopo si chiede se ci sono già elementi che ti aiutano a costruire il quadro del fenomeno” (R2D). Un professore associato ha esplicitato, a questo proposito, un cambiamento nel tempo della preferenza del ricercatore rispetto allo sviluppo di ricerche che utilizzino fonti secondarie rispetto a studi basati su dati di prima mano: *“Fino a prima l’istinto, la tendenza naturale era quella... io ho un’idea, un tema che voglio trattare; la prima cosa che mi viene in mente è che comincio ad organizzare la raccolta di dati, la rilevazione originaria dei dati; invece, questa disponibilità orienta il percorso di riflessione su come impostare la ricerca in modo diverso, si esplora la strada dell’analisi secondaria con tutte le sue potenzialità e problemi”* (R2U). Si fa viva l’idea che lo sfruttamento di risorse informative pre-esistenti alimenti il processo di cumulatività dei risultati rispetto agli studi che le hanno sfruttate precedentemente: in questo senso le infrastrutture costituiscono il punto di partenza per un’impresa scientifica volta a scandagliare aspetti non ancora esplorati o la cui attualità è emersa solo successivamente (*“avendo a disposizione quest’infrastruttura, ho potuto utilizzare i dati della ricerca precedente come la base empirica a partire dalla quale articolare un’ipotesi di ricerca. Quindi...già l’ipotesi di ricerca che io avevo individuato si costituiva come un passo successivo rispetto alla ricerca, quindi una forma di cumulazione, cumulatività dei risultati...”* – R2U). Tali nuovi avanzamenti possono ritornare sull’infrastruttura in ottica migliorativa: *“Quindi l’approfondimento dei dati conduce a una revisione della forma con cui le domande e le informazioni vengono rilevate e quindi c’è questo circolo virtuoso tra la fonte statistica ufficiale che produce il dato e l’approfondimento in sede di ricerca accademica”* (R2D). In controtendenza, solo uno tra i ricercatori più esperti, un professore ordinario, ha fatto risaltare il proprio ancoraggio ad uno stile di ricerca “tradizionale” che la disponibilità e la facilità di accesso alle infrastrutture di ricerca già disponibili non hanno scalfito (*“Ho mantenuto il mio stile nel tempo perché bene o male è difficile, un po’ si tende alla ripetizione. Merton lo diceva chiaramente che passata una certa età, passati 45-50 anni, un po’ si ripete sé stessi [...] sono più legato quindi a uno stile di ricerca, se possiamo usare questo concetto, che non dico che sia giusto, probabilmente è un grosso limite, più legato alla ricerca primaria, legata all’originalità del progetto che non all’analisi secondaria”* – R3U). Ciò deriva dal fatto che ai

tempi della sua formazione “era considerato disdicevole o comunque non utile fare alcun tipo di ricerca che viene indotta dalla disponibilità di basi informatiche”<sup>69</sup>.

Nell’approfondimento sul rapporto tra infrastruttura di ricerca e formulazione del problema d’indagine<sup>70</sup>, sono emerse interessanti evidenze dall’analisi delle interviste. Al netto di domande di ricerca pregresse che orientano all’approccio all’infrastruttura e che costituiscono la modalità prevalente per la scelta della risorsa da utilizzare (o da creare in sede di progetto), la disponibilità (o la costituzione) della stessa ha stimolato, come si evince dalle esperienze di molti intervistati, un approccio induttivista tale per cui la presenza di determinati dati ed indicatori ha condotto alla formazione di domande di ricerca non previste, anche nel caso in cui esse siano scaturite nell’ambito di progetti con obiettivi specifici<sup>71</sup>. La disponibilità di dati sistematizzati e la possibilità di interrogazioni descrittive consente ai ricercatori di compiere ragionamenti e riflessioni “dentro le matrici” ancor prima di elaborare una qualche ipotesi di lavoro. Interessante, a questo proposito la narrativa di un giovane ricercatore rispetto a una domanda di ricerca inattesa, emersa dall’approccio all’infrastruttura europea OCSE-PIAAC: “OCSE-PIAAC è pensato fundamentalmente per la misurazione delle competenze della popolazione adulta, si intende popolazione tra i 16 e i 65 anni all’atto della rilevazione, divisa in gruppi di età e così via. Nel questionario retrospettivo io però ho anche il background familiare, la carriera scolastica, le condizioni di vita, ecc. ecc. Ad un certo punto, con un collega cominciamo a ragionare su un aspetto fondamentale delle disuguaglianze che in realtà in Italia è stato preso poco in considerazione, la genitorialità. Possiamo dire che diventare genitori in Italia è un nuovo elemento di disuguaglianza? cioè che ci possono essere dei fattori culturali, economici, congiunturali che discriminano tra le classi sociali? A quel punto, abbiamo preso un database pensato per tutt’altro ma che

---

<sup>69</sup> Il netto distacco generazionale si avverte nella convinzione di un abbassamento della qualità della ricerca dovuta al modo di approcciarsi alle nuove disponibilità informatiche: “la generazione nuova ha molta più dimestichezza nell’utilizzazione della banca dati, c’è nata dentro e ragiona anche in un “modo da navigare nella banca dati e creare modelli all’interno della banca dati”. Mentre prima il ragionamento era: “mi creo il modello, ho queste capacità di calcolo, in prospettiva posso farmi questa ricerca” e questa ricerca era di qualità diversa” (R3U).

<sup>70</sup> La domanda di ricerca di un’indagine prevede l’esplicazione degli interrogativi che hanno fatto emergere l’esigenza stessa di una ricerca e ai quali si intende fornire risposta.

<sup>71</sup> A quel punto il database ti esplose in mano nel senso positivo perché ti permette di formulare tantissime domande che prima non avresti fatto...” (R1U); “In entrambi i casi, sia per \*\*\* che per \*\*\*, [la domanda di ricerca] è stata legata alla disponibilità dell’infrastruttura” (R2D); “La presenza [dell’infrastruttura] ti può far nascere nuove domande di ricerca e quindi farti poi pensare ad una successiva analisi magari integrativa rispetto a quello che avevi fatto prima” (R1U); “ci si è arricchiti su quello che già ci si riprometteva di trovare [e] però appunto ci sono state aperture nuove che non avevamo previsto a livello di programmazione iniziale (R3D); “ti dà la possibilità di intervenire su cose che cambiano continuamente e che vengono costantemente aggiornate e che ti danno un quadro della situazione molto più ampio e poi ti permettono di mettere dentro delle cose a cui magari non avevi pensato” (R3D).

*includeva delle variabili di segmentazione tali da poter individuare il gruppo di “giovani adulti fuoriusciti da casa”, quello dei “giovani adulti fuoriusciti da casa con uno o più figli”, e così via incrociandolo con diverse altre variabili. [...] perché queste indagini poi hanno questa caratteristica di avere dei moduli, loro li chiamano “nested”, “annidati”, dei moduli dentro altri che ti permettono di andare ad approfondire delle tematiche specifiche” (R1U).*

Su questa scia, una professoressa si è spinta a qualificare l’atteggiamento del ricercatore rispetto al dato già disponibile, come mosso dalla voglia e dall’audacia di spingersi in ipotesi inedite e azzardate, utilizzando un tema ricorrente della tragedia greca e della letteratura greca, quello della *hybris* che indica "tracotanza" o "superbia": *“ha aumentato, lo dico a livello emozionale, la hybris del ricercatore. Il ricercatore potenzialmente sa che può accedere anche a strumenti che facilitano alcune operazioni o che ottimizzano i suoi tempi. Quindi nel disegno della ricerca, non nell'esito finale, ritengo che il ricercatore sia più spronato a spingersi in ipotesi di ricerca sulle quali prima andava più cauto. Perché sapeva che i tempi e le risorse a disposizione potevano essere non amichevoli, da un certo punto di vista. È un elemento positivo e negativo al contempo. Positivo ovviamente perché ci si può spingere a fare ipotesi dati che abbracciano più ambiti, aiutano nella multidisciplinarietà della ricerca. [...] Negativo, perché spesso avendo un’overdose di strumenti, di risorse potenzialmente disponibili – sono ancora nella fase del disegno della ricerca – non credo che ci sia – almeno io non l’ho sviluppata forse – la capacità di selezionarle in maniera funzionale per la ricerca” (R2D).*

Un punto significativo è costituito dal fatto che l’infrastruttura di ricerca, essendo un oggetto in divenire, aggiornabile, flessibile, lascia libertà creativa nel processo di ricerca consentendo continue ridefinizioni nel percorso di produzione di nuova conoscenza. Tale circostanza permette al ricercatore di tornare sulle sue decisioni, ad esempio riformulando gli obiettivi, sperimentando nuove metodologie o tecniche di analisi. Secondo un’intervistata, tale circostanza rende la ricerca sociale sempre più simile ai processi che connotano la ricerca valutativa: *“Il disegno della ricerca secondo me cambia nei termini progettuali che noi vediamo più simili che nella ricerca sociale nelle tradizioni della ricerca valutativa. [...] La ricerca valutativa cosa fa? Nel momento in cui si pone le domande della ricerca, apre un percorso continuo di ridefinizione. Secondo me, il fatto di avere delle infrastrutture di ricerca innovative che sono quindi costantemente implementabili dal punto di vista metodologico, fa sì che questo disegno della ricerca sia una fase sempre aperta. Tanto che io stessa – ecco*

*qual è l'elemento in più – torno sempre nuovamente a ridefinire il mio disegno della ricerca. Perché ho chiamato in causa la ricerca valutativa? Perché lì è una prassi consolidata; al contrario, nella ricerca sociale, è più difficile che la comunità scientifica accetti questo percorso continuo. Questo per me è un valore aggiunto invece” (R2D).*

Che l'infrastruttura sia di tipo precostituito o sia scaturita da progetti di ricerca nei quali i ricercatori sono stati coinvolti, il condizionamento della sua disponibilità sulla possibilità stessa di progettare ed effettuare lo studio è più volte ribadito: *“L'infrastruttura ha influito in maniera totale, nel senso che era quasi una condizione necessaria. Il disegno della ricerca è stato totalmente condizionato? Totalmente, totalmente. [...] buona parte, forse tutto no, dell'impianto di definizione del problema, ma anche di ricerca, è tutto scaturito dalla disponibilità dell'infrastruttura. Ah, visto che c'è questa cosa, facciamo questo, questo, questo: così è andata. Non ci sarebbe forse nemmeno venuto in mente, ecco” (R2U).*

In particolare, decisivi nella scelta di “disegnare” studi attraverso infrastrutture di ricerca già precostituite si rivelano fattori quali la convenienza economica e il risparmio di tempo, e si accompagnano a valutazioni di qualità dell'*output* (sull'aspetto qualitativo relativo ai prodotti della ricerca si tornerà in maniera estesa nel par. 4.4.1), come la possibilità di rappresentare il fenomeno di studio più esaurientemente rispetto ad una *survey* classica (*“La disponibilità, e la capacità soprattutto, di estrarre il dato in un certo tipo di modo e di un certo tipo di dato è stato fondamentale per riuscire a svolgere l'analisi di ricerca. L'alternativa sarebbe stata quella di realizzare una survey classica campionaria, quindi ridotta” – R1U*); mentre nei casi in cui i ricercatori abbiano sfruttato infrastrutture di ricerca nell'ambito di progetti nei quali sono stati coinvolti, il disegno della ricerca è coinciso, in larga parte, con la progettazione e l'alimentazione dell'infrastruttura stessa. Tale sovrapposizione ha impattato, oltre che sugli obiettivi generali di ricerca, sulle metodologie d'indagine e sugli strumenti di rilevazione delle informazioni che sarebbero andate a costituire il contenuto dell'infrastruttura. Un esempio paradigmatico è rappresentato dall'esperienza di un professore associato che riferisce di un PRIN del quale è stato coordinatore. La sua volontà di svolgere uno studio caratterizzato da un approccio integrato e multilivello rispetto al fenomeno di studio ha avuto una ricaduta diretta sulla progettazione e implementazione del database condiviso con gli altri partner del progetto: *“C'è stato un grosso lavoro di progettazione anteriore che ha portato alla definizione di quei database [contraddistinto da un] approccio multi-livello e integrato, nel senso di coniugare più livelli*

*di analisi e di integrare procedure che fino ad allora erano state utilizzate separatamente [...] coniugare un livello contestuale con un livello individuale, quindi modelli e le pratiche organizzative nei call center per vedere come potessero avere delle ricadute sulle percezioni individuali, sullo stress lavoro correlato, sulla soddisfazione, sull'alienazione, sulle dimensioni della qualità della vita lavorativa, sulla percezione dell'ambiente e qualsivoglia. [Quindi] sono stati progettati due strumenti di rilevazione: una scheda di analisi del contesto che prevedeva di essere compilata sia dai responsabili aziendali sia da quelli sindacali per reperire una serie di informazioni sulle pratiche organizzative che in ciascuno dei call center coinvolti nel campione venivano impiegate; e poi un questionario molto articolato che prevedeva di rilevare le dimensioni del benessere o malessere individuale, e poi tutta una serie di percezioni sulla qualità dell'ambiente di lavoro, sulla qualità delle condizioni organizzative, e quindi cercava di mettere in relazione sostanzialmente le varie dimensioni della qualità della vita lavorativa. La scheda di analisi del contesto doveva servire a mettere in relazione le condizioni organizzative di ciascun call center con le relative percezioni individuali. Quindi la struttura della matrice prevedeva che per ciascun operatore di call center venissero incollati i report relativi alle condizioni organizzative che provenivano dalla scheda di contesto da ciascuno dei 24 call center coinvolti nella ricerca. Quindi per ciascun operatore si avevano proprietà globali, di contesto, e anche analitiche relative alla composizione della forza lavoro del call center e anche proprietà individuali, e questo consentiva di adottare una logica multilivello” (R2U).*

La base empirica che fornisce l’infrastruttura di ricerca costituisce una materia grezza che deve essere ulteriormente trattata, elaborata e analizzata, secondo procedure adeguate alla natura delle informazioni raccolte. Una delle fasi di trattamento del dato, riportata dagli intervistati, è costituita dall’estrazione delle variabili significative, posteriore alla ricognizione dell’intero insieme di dati: *“abbiamo fatto un unico database, [...] e lì abbiamo costruito noi, togliendo tutto quello che non ci interessava, la parte che serviva a noi per la nostra ricerca” (R3D)*. Termini quali “interconnessione” e “interoperabilità” ricorrono in alcune esperienze di ricerca in cui il dato dell’infrastruttura principale veniva considerato insufficiente per l’indagine da compiere e quindi usato a complemento di dati tratti da infrastrutture secondarie: *“sono state unificate... informazioni presenti in più database; [abbiamo,] per quanto è stato possibile, cercato di integrare le informazioni presenti su uno dei sistemi informatizzati, che era quello degli studenti, con altri sistemi [...], quindi le*

*informazioni di due dataset separati... sono state unificate... informazioni presenti in più database” (R1U). Infine, la rilavorazione del dato è emersa come pratica usuale preliminare all’attuazione dell’analisi: “noi abbiamo preso quei dati e li abbiamo rilavorati. In modo particolare, i dati di Sapienza non erano organizzati in modo longitudinale, quindi abbiamo costruito delle matrici in modo da ricostruire attraverso la matricola tutta la carriera di ogni singolo studente. Si tratta di diverse centinaia di migliaia di carriere di studenti” (R3U).*

Interessanti considerazioni sono, infine, emerse dalle parole dei testimoni privilegiati, che hanno sottolineato come *“l’infrastruttura un po’ guida anche la tua agenda [di ricerca]”* e dunque permette di ridefinire gli obiettivi cognitivi, *“cambiare strada”* e valutare altre possibilità. In particolare, un testimone ha riferito come in una sua esperienza di sfruttamento di materiale empirico, l’infrastruttura abbia totalmente guidato la formazione della domanda di ricerca e di come questo approccio induttivista abbia condotto a uno studio originale, seppur adattato alle caratteristiche della risorsa utilizzata: *“il ragionamento è stato: cosa abbiamo nei dati che abbiamo che può “fittare” una domanda di ricerca che nessuno mai ha osservato che è ragionevole, che si fa in un tempo breve, e che si può attaccare con i dati? E la risposta è stata: “studiamo la crescita dell’università!”. Ovviamente uno può studiare la crescita dell’università in tanti modi. Avendo quella struttura, abbiamo fatto un disegno di tipo quantitativo. Abbiamo fatto un modello semplice perché non è un panel, quindi non si potevano fare troppe cose e ci abbiamo fatto il paper. In qualche modo è una cosa nuova perché nessuno ha mai studiato la crescita dell’università nel tempo su un sample sufficientemente significativa da fare un’inferenza statistica”*. Lo stesso intervistato fa infine notare che, al contrario, gli è anche accaduto di ritenere insufficienti i dati di un’infrastruttura partendo da una domanda pregressa e dunque di aver deciso di imputare in essa dati originali, incrementandone il potenziale informativo: *“In certi casi, certe domande possono suggerire come ampliare un’infrastruttura. Mi manca un dato, mi metto a lavorare per produrlo ed aggiungerlo”*.

#### *4.3.3 Collaborazioni a livello locale, nazionale o internazionale*

Tra le declinazioni della domanda di ricerca formulata nell’ambito del disegno del presente studio, un focus particolare è stato dedicato agli effetti dell’utilizzo di infrastrutture di ricerca sulle dinamiche aggregative degli scienziati sociali a livello locale (di dipartimento o di istituto), nazionale e internazionale. Si è inteso dunque esplorare la possibilità che le

risorse informative condivise tra ricercatori abbiano potuto supportare effettive collaborazioni per la produzione di nuovi prodotti di ricerca, ponendosi tra le altre cose come mezzi per annullare i confini geografici dell'organizzazione del lavoro scientifico.

Il quadro emerso dalle interviste riguardo l'effettività delle collaborazioni derivate dalla disponibilità delle infrastrutture di ricerca si è rivelato piuttosto composito: poco più della metà degli intervistati ha caratterizzato le infrastrutture di ricerca quali strumenti effettivamente abilitanti allo sviluppo di *output* di ricerca condivisi (citandone esplicitamente alcuni nel corso delle narrative), e dunque forieri di collaborazioni effettive, la cui durata può manifestare caratteri di contingenza o (più raramente) protrarsi nel tempo; un piccolo numero di ricercatori ha lamentato di essere riuscito a mettere a frutto solo parzialmente le potenzialità collaborative legate all'utilizzo di piattaforme comuni (sia a livello di pubblicazioni che di sviluppo di costituzione di *network*); infine due quinti degli intervistati hanno evidenziato invece come lo sfruttamento congiunto delle infrastrutture di ricerca non abbia inciso significativamente sulla produzione di *output* co-autorati. All'interno di quest'ultimo gruppo un numero limitato di intervistati non ha affatto riscontrato, in base alla propria esperienza, un legame tra nascita di nuove collaborazioni effettive e sfruttamento di infrastrutture di ricerca, riconoscendo ad esse un potenziale relazionale solo teorico.

Una prima notazione rispetto a un cambiamento generale del modello collaborativo è emersa dalle parole di una professoressa ordinaria, per la quale, rispetto al passato, la nascita delle collaborazioni per mezzo dello sfruttamento congiunto di infrastrutture di ricerca conduce all'abolizione dei fattori di distanza fisica, condensandosi invece attorno a oggetti virtuali condivisi: *“Era un modello un po' più cartaceo e legato ai treni e agli aerei [...] noi eravamo intorno a una rivista e c'incontravamo periodicamente”* (R3D). Tuttavia, nei casi in cui l'utilizzo comune di infrastrutture abbia generato la produzione di *output* condivisi, la maggior parte degli intervistati ha dichiarato di aver esperito collaborazioni *a partire dal coinvolgimento in progetti specifici* per cui lo sforzo collaborativo per la costruzione *ex-novo* di un'infrastruttura si è riverberato nella produzione di articoli e report a firme congiunte. L'infrastruttura ha rappresentato il veicolo per sostanziare un'istanza collaborativa già indotta *ex ante* dalla partecipazione comune al progetto: *“Il progetto ha fatto nascere le collaborazioni, che è di solito il modo standard finora di utilizzare infrastrutture come quelle”* (R1D). Le collaborazioni si sono dunque concretizzate con colleghi, perlopiù italiani e/o colleghi di dipartimento, che avevano concorso al progetto e alla creazione comune

dell'infrastruttura; tale circostanza fa desumere che essi fossero già presenti nel *network* scientifico degli intervistati *“perché senza quella collaborazione, quella fiducia reciproca, quella conoscenza che poi porta alla costruzione di un gruppo non si arriva a nessun database. Quindi intanto a monte c'è la conoscenza, la collaborazione, degli interessi in comune, e quindi la generazione di un database”* (R2D). Il contenuto dell'infrastruttura, in grado di soddisfare le esigenze legate a una domanda di ricerca comune, ha condotto alla trasformazione dell'affinità tra ricercatori in collaborazione effettiva nella produzione scientifica, a patto che i co-autori condividessero visioni e istanze comuni sull'ontologia del dato (*“voglio lavorare con gente che tratta i dati come oggetti sociali da interrogare [e non come elementi dal carattere autoevidente] – R2D*) e sulle modalità del suo sfruttamento. L'impossibilità di collaborare con ricercatori esterni ai progetti è peraltro legata alle regole degli stessi (*cf.* par. 2.3), per cui l'infrastruttura creata viene condivisa a terzi in rari casi.

Anche per quanto riguarda la circostanza, rappresentata in poche occasioni, della formazione di collaborazioni internazionali, si riscontra in genere una possibilità di produzione di *output* condivisi solo nei casi di concorso nella creazione di infrastrutture da progetto specifico. Nei casi di infrastrutture generate da progetto europeo che ospitino dati *cross-sectional* provenienti da diversi Paesi prodotti con la medesima metodologia, l'istanza collaborativa emersa a livello di implementazione dell'infrastruttura può tradursi in termini di collaborazioni effettive in studi comparativi congiunti: *“La fase di collaborazione c'è stata in fase di costruzione del dataset \*\*\* [...]. [Mi ha consentito] di produrre un risultato che potesse essere di interesse per partner internazionali. Ciò significa che nel momento in cui ho lavorato su due prodotti: un report, lavorando sui database dell'università, e uno sull'apertura dei programmi open, il fatto di aver potuto includere casi diversi e realtà diverse nell'analisi, non solo nazionali ma di altri Paesi, mi ha permesso di coinvolgere nella stesura del lavoro anche altri ricercatori stranieri che ovviamente potevano confrontarsi sul dato e sulle mie argomentazioni. Se il lavoro si fosse basato prettamente su dati nazionali molto probabilmente non ci sarebbe stata la collaborazione di altri ricercatori”* (R2D). La produzione di articoli scientifici co-autorati con colleghi stranieri può essere inoltre facilitata dal fatto che i soggetti coinvolti nella *partnership* che dà vita all'infrastruttura supportino la medesima visione dell'impresa scientifica come processo comune: *“abbiamo costruito questo database che ha consentito di creare una piccola infrastruttura di ricerca tra noi e un gruppo di ricercatori della Charles University di Praga [...]. Noi abbiamo bisogno di massimizzare*

*le risorse che abbiamo, di mettere in comune queste risorse che abbiamo, perché l'unica possibilità è di pensarci dentro una comunità, cioè dentro un insieme di persone che condividono dei valori scientifici” (R2U).* Infine, le potenzialità di collaborazione internazionale, quando non sfruttate, vengono comunque percepite quali possibilità concrete a cui poter attingere in un tempo futuro: *“nello specifico non mi è mai però capitato di lavorare a stretto contatto con qualcuno che non sia italiano, però il fatto di aver visto e aver lavorato su dataset diversi secondo me è un punto a favore, nel senso che se dovesse capitarmi di riuscire ad avere una collaborazione con una persona che sta dall'altra parte del mondo, riusciamo a capirci sostanzialmente” (R1U).*

La disponibilità di infrastrutture *indipendenti da progetti*, disponibili a livello italiano o europeo, viene quasi mai citata quale opportunità per la produzione di studi originali da sviluppare attraverso collaborazioni. In queste poche occasioni, secondo un professore associato, tali collaborazioni si sviluppano in ragione della complementarità di competenze tecniche e della diversificazione nelle possibilità di accesso all'infrastruttura: *“sul piano delle competenze, nel senso che a volte c'è bisogno di collaborare con qualcuno perché ti mancano le competenze di trattamento d'informazione, banalmente competenze informatiche di collegamento di database relazionali... anche solo per quello... qualcuno se la cava da solo magari per fare un minimo, se vuoi fare di più c'è bisogno di qualcuno che sa muoversi e sa costruire database relazionali e un piano. L'altro piano che può sviluppare competenze è la proprietà. Io collaboro con qualcuno per mettere in comune le nostre proprietà, cioè i nostri accessi” (R2U).* L'opzione “strumentale” nella formazione di collaborazioni legate allo sfruttamento delle infrastrutture viene posta in risalto anche da un'altra ricercatrice sociale, che ha riscontrato *“un impatto dirompente nelle reti di ricerca nel numero di collaborazioni, anche [nell']estensione della collaborazione, [nel]l'uso di dati molto differenti rispetto alle indagini fatte in analisi primaria fatte in prima persona su piccoli campioni [...]*”, facilitato dal fatto che nei gruppi di ricerca che hanno utilizzato le infrastrutture lei sia riuscita a inserirsi quale portatrice di forti competenze statistiche incarnando *“un ruolo di collaborazione con tanti soggetti di questo dipartimento, proprio perché nell'ambito di ogni progetto c'è la necessità di uno statistico che magari consentisse una strategia più rigorosa o l'applicazione di un modello matematico più rigoroso” (R2D).* Un raro caso in controtendenza per quanto riguarda la nascita di nuove collaborazioni derivato dall'utilizzo di infrastrutture di caratura internazionale è riscontrabile nella testimonianza di un giovane

ricercatore che, alla domanda specifica se lo sfruttamento di infrastrutture di ricerca abbia influito sulle sue possibilità di realizzare ricerche in collaborazione con altri ricercatori italiani o stranieri nel corso della sua carriera, ha risposto citando una serie di esempi in successione: *“Sì, in termini di collaborazioni effettive il mio network si è proprio allargato. Si è proprio decisamente allargato. [...] essendo database più o meno standard, ci possiamo scambiare le sintassi e quindi aggiungere elementi e così via. E quindi, per fare un esempio, sull'OCSE-PISA mi è capitato di scrivere con la collega \*\*\*, con la quale abbiamo fatto un paio di articoli in italiano e in inglese, sul PISA con \*\*\*, con \*\*\* e \*\*\* sul PIAAC; attualmente sto lavorando invece su \*\*\* che prende l'indagine sulla matematica e le scienze, mettendo i dati in serie storica. Quindi sì, ti permette di fare tutti questi lavori... [...] ragioniamo insieme sulle analisi da fare, poi magari le analisi vengono fatte pure singolarmente, però ci scambiamo i diversi modelli di analisi e poi arriviamo a dei risultati che in qualche modo appunto sintetizzano il lavoro svolto” (RIU)*. Lo stesso intervistato ha manifestato il suo entusiasmo rispetto allo sfruttamento delle potenzialità collaborative, avvenuto sin dall'inizio della sua carriera, anche in relazione all'utilizzo di risorse informative della World Bank nell'ambito di un progetto sulla comparazione dei distretti industriali nell'Europa Orientale ed Occidentale che ha coinvolto anche un economista e un ingegnere: *“è stato un caso di felice cooperazione su un progetto di ricerca molto specifico che ci faceva vedere appunto una serie di interscambi, e quindi ha funzionato sostanzialmente bene. Ed è stata proprio l'infrastruttura che ci ha fatto lavorare insieme... perché se non mancava l'occasione...” (RIU)*. In questo caso, dunque, è stata la disponibilità di un'infrastruttura di tipo preconstituito a generare l'opportunità collaborativa, peraltro con colleghi esterni ai *network* dell'intervistato.

Tornando al tema della corrispondenza tra collaborazione nell'implementazione dell'infrastruttura e collaborazione nella produzione di *output* di ricerca da essa derivati, tale circostanza non appare affatto scontata e spesso deriva dalle caratteristiche del progetto entro cui l'infrastruttura nasce. Come anticipato, infatti, la possibilità di produrre *output* condivisi a partire dall'accesso comune ad un'infrastruttura può rimanere solo teorica. Come sottolineato dall'esperienza di mancata collaborazione di un professore associato nell'ambito di un progetto che ha coinvolto 4 unità operative dislocate geograficamente in altrettante città italiane: *“C'è stata collaborazione, sì, c'è stata sicuramente per arrivare alla matrice finale [...]. Il risultato collaborativo è stato parziale...Al momento di pubblicare abbiamo deciso*

*di dividerci, per cui le due università del sud hanno prodotto un volume e noi [Università Sapienza] un altro. Il loro più descrittivo, il nostro andava più a vedere le variabili in relazione tra loro. [...]. L'Università di Milano si è defilata, e le altre due hanno lavorato per proprio conto. Abbiamo collaborato solo per fare prima della pubblicazione un convegno finale di presentazione dei risultati in cui ci siamo divisi le parti.* Lo stesso soggetto individua la ragione del distacco nella diversità degli obiettivi cognitivi che le diverse unità avevano intenzione di raggiungere sfruttando la medesima infrastruttura: *“il mio tentativo [di studio sperimentale era] osteggiato anche da una delle unità di ricerca che non riusciva a coglierne il senso”*. Scetticismo sul legame tra infrastrutture di ricerca e nascita di nuove collaborazioni è espresso infine da una professoressa associata: *“non sono del tutto convinta che siamo ad un livello di maturità tale per cui l'infrastruttura di ricerca, almeno nella mia esperienza ribadisco, può aver prodotto quella collaborazione. [...] che l'output sia arrivato proprio in forza di, non ne sono del tutto sicura. È più opera della buona volontà dei ricercatori coinvolti...più che dell'infrastruttura” (R2D)*.

Allontanandoci dal campo delle collaborazioni effettive per la produzione di prodotti co-autorati, quasi tutti gli intervistati hanno sottolineato che lo sfruttamento di risorse comuni ha il potenziale, a volte sfruttato a volte non sfruttato, rafforzare i *network* relazionali, tuttavia non sostituendo *pattern* di rete già costituiti. Solo un intervistato propone una visione peculiare del legame tra infrastrutture e reti – quella delle *reti lasche* – legate all'entrata e uscita in molteplici reti di ricercatori che sviluppano infrastrutture di volta in volta diverse: *“Secondo me, il senso dell'infrastruttura è anche questo...io prendo un'autostrada per andare al mio paese... non è che voglio passare la mia vita in quell'autostrada, la prendo, la uso quando mi serve e altre volte userò altre strade. Questo fa parte della ricchezza che la rete consente di sviluppare, perché è una moltiplicazione di risorse, e anche la possibilità di allargare queste infrastrutture, saltando da un'infrastruttura all'altra...perché saltando da un'infrastruttura all'altra si allargano queste infrastrutture, si creano dei... anche dei punti di collegamento tra le infrastrutture” (R2U)*.

In chiusura, è importante capire come i testimoni privilegiati, creatori e primi utilizzatori di infrastrutture di tipo europeo, descrivano il legame tra le infrastrutture di ricerca e le proprie opportunità collaborative. Citando le loro esperienze, essi percepiscono le infrastrutture di ricerca quali fattori che possano assicurare continuità a collaborazioni esistenti e costituirsi come base per il mantenimento di nuove collaborazioni costruite a

partire da esse. L'aver concorso alla creazione di tali risorse ha posizionato tali soggetti in una condizione di vantaggio competitivo nella creazione di nuove collaborazioni: *“Allora più che altro il fattore essenziale è che io alcune infrastrutture le ho create e quindi questo mi ha dato una posizione e una possibilità di collaborazione che altri non hanno avuto, semplicemente perché io ho dei dati unici. Non potevo avere le medesime collaborazioni senza l'infrastruttura [...] avere una conoscenza tacita sull'infrastruttura che nessuno ha, ha creato possibilità di collaborazione con altri che hanno dei grossi dataset”*. La visibilità dell'infrastruttura appare come elemento determinante nella stipula di collaborazioni soprattutto internazionali, prima di tutto perché dà modo di mettere in contatto tra loro creatori di infrastrutture potenzialmente interoperabili: *“alcune delle nostre infrastrutture di ricerca hanno una visibilità tale che consente una facile rintracciabilità nel panorama nazionale ed europeo... [...] ed io collaboro con \*\*\* [cita un famoso professore olandese] non perché sia simpatico ma perché anche lui ha una riconoscibile, specifica infrastruttura che in qualche modo comunica con la nostra”*. Sia dal lato della creazione che dello sfruttamento le infrastrutture hanno *“incentivato collaborazioni nuove che non c'erano prima...”* Significativa, a questo proposito, l'ammissione di un testimone su alcuni lavori collaborativi che *“non sarebbero nati se non ci fosse stato lo sforzo di creare questa infrastruttura, sì. [Ad esempio ho] fatto un articolo di presentazione che è stato presentato ieri a Washington con un greco che io non avrei mai incontrato in vita mia se non avessi partecipato a quest'infrastruttura europea sulle tecnologie e il patrimonio culturale”*.

#### 4.3.4 Propensione all'interdisciplinarietà

Connesso in parte all'approfondimento sulle collaborazioni, il tema degli effetti di interdisciplinarietà interna ed esterna<sup>72</sup> relativo allo sfruttamento di infrastrutture di ricerca è stato affrontato in tutte le interviste che costituiscono la base empirica di questo studio. Come già emerso in una delle testimonianze riportate nel sub-paragrafo precedente, relativamente alla collaborazione di un sociologo con un economista e un ingegnere nell'utilizzo di risorse

---

<sup>72</sup> Seguendo le definizioni di interdisciplinarietà proposte dalla Roadmap ESFRI 2016, per “interdisciplinarietà interna” si intende l'integrazione di diversi sottocampi all'interno della stessa disciplina per affrontare le domande di ricerca proposte, con la locuzione “interdisciplinarietà esterna” si fa riferimento all'integrazione di diverse aree disciplinari per affrontare questioni di ricerca che mirano generalmente a risolvere un complesso problema scientifico (cfr. ESFRI, 2016).

informative della World Bank<sup>73</sup>, la disponibilità di insiemi di dati e indicatori integrati e armonizzati offre potenzialmente ai ricercatori la possibilità di ampliare il panorama informativo relativo connesso ai propri interessi di ricerca, consentendo di approfondire tematiche di interesse da prospettive differenti e di pubblicare lavori interdisciplinari. Tale circostanza è riconosciuta, in potenza, dagli intervistati, tanto che in alcuni casi emerge una consapevolezza di cospicue potenzialità interdisciplinari di alcune infrastrutture istituzionalizzate, come OCSE-PISA: *“lì ci lavorano intanto i sociologi, gli informatici, i pedagogisti, gli psicologi, gli economisti dell'istruzione che da un certo punto di vista contribuiscono allo strumento questionario PISA in continua modifica”* (R2D). L'aumento dell'interazione tra scienze sociali e altre scienze è stato citato, in un paio di casi, quale una delle caratteristiche più salienti del recente periodo, e questo nuovo approccio ha posto in evidenza la questione della creazione di infrastrutture ibride nell'ambito di progetti di ricerca, che aggregano dati derivanti da diversi domini, come nell'esempio di un dirigente di ricerca che ha evidenziato, per la formazione di un'infrastruttura, l'instaurazione di un *“collegamento con gli altri settori coinvolti come architetti, gli storici dell'arte, gli urbanisti”* (R3D).

L'analisi delle interviste ha tuttavia posto in rilievo che, sebbene quasi sempre gli intervistati abbiano evidenziato l'efficacia della condivisione di risorse informative nella produzione di collaborazioni interdisciplinari nell'ambito di progetti di ricerca di ateneo o nazionali, è talvolta mancato un concreto riferimento ad articoli o libri co-firmati in cui sia riscontrabile tale orientamento, quale ricaduta del lavoro comune. Solo in qualche occasione il riferimento ad *output* specifici dal contenuto interdisciplinare è emerso più chiaramente, come in questo caso riportato da un professore associato: *“nell'analisi di alcune politiche sociali per il governo locale, ho lavorato con ingegneri, con medici e con urbanisti. Avevamo una piccola infrastruttura, un database dove conferivamo alcuni dati raccolti... a parte i dati di seconda mano raccolti da varie fonti...e poi abbiamo fatto un'indagine sociale con strumenti metodologici della sociologia, la quale però nel disegno di questi strumenti e all'utilizzo di questi strumenti hanno collaborato anche questi esperti coi loro punti di vista. Quindi, rispetto alla stessa infrastruttura, abbiamo fatto un output più ricco, reciprocamente”* (R2D). Questa considerazione porta a desumere che la ricchezza informativa e le potenzialità delle infrastrutture sfruttate nell'ambito di una *partnership*, anche in presenza di colleghi rappresentanti di settori o sub-settori scientifici differenti, si siano declinate, nella

---

<sup>73</sup> Cfr. par. 4.3.3.

maggior parte dei casi, in lavori orientati alla materia di competenza. A supporto di questa considerazione, è da citare l'esperienza di un ricercatore *junior*: *“nel caso della ricerca sul database di Twitter, un dataset unico ha permesso di fare varie analisi sotto diversi punti di vista. È stato il punto di vista del network che si è creato e quindi un certo tipo di analisi, un altro tipo di analisi era più geografico e riferita al posto da cui erano partiti questi tweet e poteva essere un altro focus con altri obiettivi. Quindi sicuramente dataset più ricchi di informazioni permettono di approcciare al dato e alla ricerca da vari punti di vista [...] il gruppo era composto da geologi, vulcanologi, comunicatori, sociologi, ognuno ha analizzato dal suo punto di vista i dati”* (R1U). La produzione di prodotti di ricerca interdisciplinari non è addirittura emersa in alcun caso per quanto riguarda lo sfruttamento di infrastrutture europee precostituite, mentre è stata riscontrata in un solo caso riguardo all'utilizzo di dati messi a disposizione da World Bank.

In generale, comunque, come si accennava, una sempre maggiore presa di coscienza del valore della pratica della ricerca interdisciplinare nelle scienze sociali è stata rilevata da più di un ricercatore sociale, e la relazione diretta tra questo fenomeno e la presenza o la costituzione di infrastrutture di ricerca è stata ribadita in alcune occasioni: *“Certo adesso anche a livello di progettualità, di progetti di ricerca... anche a livello di Ateneo adesso si spinge più verso l'interdisciplinarietà per i progetti in cui sono coinvolti aree scientifiche differenti e facoltà differenti, ma anche nell'ottica di proporre progetti a livello europeo c'è una maggiore tendenza a costituire delle reti di ricerca anche con persone differenti e di avere come output una risorsa, un'infrastruttura”* (R2D). L'esistenza delle infrastrutture condivise, anche a livello locale di istituzione universitaria, è considerata come opportunità per superare la frammentazione della ricerca: *“io spesso con la mia esperienza noto che non essendoci degli archivi condivisi, io e i miei colleghi di altri dipartimenti o addirittura di altre facoltà marciamo parallelamente su direttrici di ricerca, magari ottenendo anche risultati che possono essere non solo messi in condivisione ma anche fruttuosi per i follow-up di ricerca e non lo sappiamo”* (R2D). Ancora, secondo l'esperienza di una ricercatrice *junior*, le infrastrutture di ricerca conferiscono spazio a metodi di ricerca integrativi e trasversali riguardanti l'intero settore delle scienze sociali: *“personalmente mi ha portato a scoprire la bellezza e l'importanza della “mixed method research” [...]. Le grandi masse di dati non sono autoevidenti, hanno bisogno di essere scandagliate in moltissimi modi e allora essere gruppo e condividere significa proprio che tutte quelle competenze interdisciplinari*

*serviranno ad arrivare a risultati di qualità, e probabilmente apriranno la strada a nuovi studi più piccoli” (R1D). Sulla stessa scia, l’utilizzo di infrastrutture appare, nella testimonianza di un professore associato, “auspicabile in [un’] ottica di approccio multilivello integrato e l’integrazione può essere intesa come integrazione tra diversi ambiti disciplinari”<sup>74</sup> (R2U). Nel caso appena riportato, l’opzione di interdisciplinarietà interna è stata perseguita già in fase di progetto per la creazione *ex-novo* di un’infrastruttura. Certamente, se in sede di progettazione la domanda di interdisciplinarietà viene prevista, l’infrastruttura viene poi sfruttata in modalità poliprospectica, come riportato da un’intervistata in riferimento ad una banca dati costituita da dati secondari Istat: “era un tema di ricerca condiviso nell’ambito di questo dipartimento da diversi ricercatori, anche se da diversi punti di vista rispetto agli stessi dati. Quindi abbiamo pensato di condividere queste risorse in questa struttura: partendo dallo stesso file di microdati, di ottenere diverse visioni dei dati a seconda degli obiettivi e delle curiosità scientifiche dei diversi autori. [L’output] è un volume collettaneo in cui hanno scritto circa 10 ricercatori diversi con ottiche differenti, leggendo aspetti differenti degli stessi dati. Quindi secondo me il valore aggiunto di questa condivisione è questa, che a partire dagli stessi dati, che sono stati rilevati da una fonte differente che è appunto Istat, per motivi diversi dagli obiettivi cognitivi di ciascun ricercatore...” (R2D).*

In riferimento a infrastrutture già presenti a livello nazionale ed europeo, seppure sia opinione comune che esse “contengono al loro interno dati ed informazioni che riguardano più ambiti e argomenti perché vanno ad indagare le opinioni degli individui o come si relaziona la società con varie tematiche che possono andare dall’ambiente al lavoro [fa riferimento alle banche dati Istat]... da questo punto di vista, un’infrastruttura interdisciplinare supporta la mia attività di ricerca che secondo me per definizione è interdisciplinare perché mi permette di avvicinarmi a diversi temi e anche toccare diverse discipline” (R1D), un’intervistata evidenzia come nella sua esperienza concreta di sfruttamento di un’infrastruttura con possibili implicazioni interdisciplinari, non ne abbia sfruttato appieno le potenzialità, poiché guidata da una domanda di ricerca molto mirata e incardinata su un solo sub-settore scientifico: “L’interdisciplinarietà, questo è il mio punto di

---

<sup>74</sup> Citando la propria esperienza relativa alla creazione e condivisione di un’infrastruttura di ricerca, lo stesso professore ha sottolineato come abbia “cercato di comporre le unità di ricerca anche guardando alla diversa specializzazione. Erano tutte interne alla Sociologia, ma alcune erano specializzate nella sociologia del lavoro e dell’organizzazione in maniera tale da poter recuperare quelle competenze che sia io che il Professor \*\*\* come sociologi generali potevamo non avere o avere meno. Però sicuramente avrebbe giovato in tutt’e due i progetti la presenza anche di figure esterne alla disciplina sociologica...”.

vista, nasce dalla domanda di ricerca. La domanda di ricerca nel mio caso non era interdisciplinare, di conseguenza le potenzialità magari di un approccio interdisciplinare da me non sono state sfruttate appieno” (R2D). In un altro paio di casi è stata addirittura riferita l’assenza di ricadute interdisciplinari dallo sfruttamento dell’infrastruttura: in un caso perché il contenuto non permetteva collegamenti con altre discipline, nell’altro perché le potenzialità non sono state sfruttate a causa di mancanza di competenze nel gruppo di ricerca. E a proposito del mancato sfruttamento delle infrastrutture per finalità di lavoro interdisciplinare, è da tenere presente la questione posta da un professore ordinario riguardo le ricadute valutative di lavori collaborativi e interdisciplinari derivanti dallo sfruttamento comune di infrastrutture: *“Più è complesso il tema dell’infrastruttura, più è variegato il contesto socioculturale delle persone che ci lavorano e più aumentano le possibilità di lavoro interdisciplinare, che è un bene dal punto di vista sociale. Va detto però al ricercatore individuale che poiché il suo lavoro verrà valutato da matrici disciplinari specializzate, il lavoro interdisciplinare può essere assolutamente improduttivo sul piano accademico. [...] Database ed infrastrutture che sono rivolti ad interventi su sistemi complessi sono naturalmente di accesso ed utilizzazione interdisciplinare, ma hanno questo problema [se gli output derivanti entrano in valutazione]”* (R3U).

Dalle parole di un testimone privilegiato, emerge come l’interdisciplinarità sia un elemento ricorrente in fase di creazione dell’infrastruttura, che può riverberarsi non per forza nella produzione di un singolo prodotto interdisciplinare, ma nella costituzione di contenitori, quali riviste (*“[che contengano] approfondimenti quasi sempre di tipo interdisciplinare... e comunque c’erano riferimenti relativi a tutte le discipline coinvolte [nella costituzione dell’infrastruttura], in cui vari autori, a volte individualmente a volte in concorso con colleghi provenienti da altre discipline, soddisfano obiettivi cognitivi più o meno complessi e integrabili sotto diverse prospettive. Un altro testimone, ricordando la propria produzione, ha al contrario sottolineato, rispetto alla relazione tra infrastrutture e interdisciplinarità, la persistenza di rigidità epistemiche che non vengono superate: “ho piuttosto l’impressione che un’infrastruttura, siccome tende a riflettere una tradizione epistemologica ben precisa, tende piuttosto a ridurre la collaborazione disciplinare più che aumentarla o favorire certe collaborazioni. È chiaro che magari lavori con gli informatici. Per esempio tra economia ed informatica c’è una buona compatibilità... ma tra un sociologo ed un economista... non li avvicini mica”*).

## 4.4 Valutazioni sulle ricadute dell'utilizzo di infrastrutture sui prodotti di ricerca: aspetti positivi e criticità

### 4.4.1 Aspetti qualitativi relativi agli output sviluppati da infrastruttura

Una rilevante dimensione che si è inteso approfondire nell'ambito di questo studio ha riguardato le ripercussioni percepite dai ricercatori sugli *output* di ricerca per i quali sono state utilizzate risorse informative aperte e/o condivise, quali le infrastrutture di ricerca disponibili tramite *web* o create nell'ambito di progetti specifici. Aderenza agli obiettivi cognitivi, rafforzamento della capacità di rappresentazione del fenomeno, istanze di affidabilità e accuratezza del dato, robustezza dello stesso, opportunità di comparazione rappresentano alcuni degli aspetti emersi dalle parole degli intervistati in riferimento alla qualità dei propri prodotti di ricerca derivati da infrastrutture.

Un primo aspetto più volte emerso dalle interviste ha riguardato il parallelismo tra qualità del dato contenuto nell'infrastruttura e qualità del prodotto di ricerca scaturito dal suo sfruttamento. Dalle parole di un'intervistata la percezione della ricaduta appare evidente: *“la qualità è stata influenzata in termini positivi perché entrambi i prodotti [derivati dall'utilizzo di infrastrutture di ricerca] ed entrambe le infrastrutture godono di quel grado di affidabilità e di robustezza [...]. Quindi il prodotto che ne è scaturito è un prodotto affidabile e robusto rispetto al dato utilizzato”* (R2D). Nel caso presentato, l'intervistata riferiva di due esperienze tratte dall'utilizzo di infrastrutture di tipo europeo, ma anche quando si tratta di infrastrutture derivanti da progetto, aspetti di qualità scientifica relativi all'affidabilità del dato vengono esaltati, in particolare da un professore associato: *“Quello che enfatizzerei è la possibilità di utilizzare dati affidabili e in qualche misura certificabili. [...] Il fatto di poter usufruire di dati affidabili, altamente affidabili, non voglio dire certificati o certificabili, ma altamente affidabili, costruiti in modo tale da poter essere utilizzati anche per differenti scopi di ricerca”* (R2U). Dato che tutte le tipologie di infrastrutture fanno riferimento alla supervisione di *team* di ricerca (dei quali gli utenti possono talvolta far parte), gli intervistati manifestano generalmente piena fiducia nella qualità del contenuto della risorsa, specie se i creatori o le fonti hanno un riconoscimento internazionale (*“allora io che non sono molto giovane guardo molto alla credibilità della fonte e quindi utilizzo solo quelle che secondo me sono credibili”* – R3D) o se il corredo di metadati è sufficientemente ricco da permettere la ricostruzione delle fasi di produzione del dato: *“È chiaro che se uso i dati della Banca Mondiale quantomeno li*

*riconosco come dati ufficiali, mainstream, standardizzati a cui non attribuisco una verità assoluta, però li prendo come dei dati solidi. La solidità non è perenne non è che la solidità equivale a verità, ma significa che quei dati hanno una storia” (R2D).*

Il fatto di poter contare su un punto di accesso unico (*single entry-point*) a set di dati pre-organizzati è percepito come un fattore di vantaggio sull’*output* di ricerca soprattutto in riferimento alla velocità di produzione dei lavori. In particolare, le infrastrutture di caratura internazionale, oltre a dare percezione e/o consapevolezza di affidabilità, attendibilità e robustezza, rappresentano risorse che, grazie alla loro pre-organizzazione, consentono di individuare rapidamente le variabili rilevanti su cui indagare (*“Il plus di qualità è il poter basare la creazione del prodotto su qualcosa che è già stato prestrutturato, preorganizzato e predefinito, poter costruire su qualcosa di solido e sul quale si può rielaborare e successivamente crescere di più; invece se si deve ricominciare da zero il percorso è più lungo e magari più difficile e rischioso perché può prendere strade sbagliate. Invece se hai un’infrastruttura che ti consente di avere una base da cui partire, quello sicuramente...”* – R3D) e anticipano il lavoro di gestione degli errori non campionari<sup>75</sup>, per esempio in riferimento al contatto con le fonti primarie per l’ottenimento delle informazioni (*“un altro punto forte poteva essere la mancanza di bias [...] ma allo stesso tempo mancano i problemi che possono derivare da un’intervista anche strutturata come quella di un questionario soprattutto quando chiedi una cosa avvenuta nel passato quindi non molto recente”* – R1U).

Il risparmio di tempo, se i dati provengono da fonti differenti, è più volte evidenziato (*“[ho avuto accesso a] informazioni organizzate e sistematizzate che io altrimenti, anche in altri lavori precedenti, ho dovuto reperire da più fonti, con un grado di precisione ed accuratezza inferiore”* – R2D) e talvolta accompagnato da valutazioni molto interessanti sul rapporto tra produzione di un lavoro scientifico e istanze di valutazione dello stesso, nei termini della possibilità di ricevere *feedback* più rapidi rispetto alla produzione individuale, come emerge dalle parole di questo giovane ricercatore: *“Velocità non significa qualità; però se io riesco a produrre un paper in breve tempo [da sottoporre] poi a peer review, in un tempo ridotto è ovvio che la mia capacità di produzione e di affinamento dello strumento migliora.*

---

<sup>75</sup> Gli errori non campionari concorrono, insieme ai campionari, nella determinazione dell’entità dell’errore complessivo di uno studio statistico e rappresentano la misura dell’accuratezza dei valori contenuti (Brasini *et al.*, 2002). Possono essere suddivisi in tre tipologie, secondo Lessler e Kasbeek (1992): errori di lista (assenza di biunivocità nella corrispondenza tra lista e popolazione oggetto d’indagine); errori di mancata risposta (dovuti al rifiuto, all’impossibilità o all’indisponibilità a rispondere da parte dei soggetti contattati per l’indagine); errori di misurazione (generati tutte le volte in cui, durante un’indagine, viene rilevato un valore diverso da quello reale).

*Quindi, da questo punto di vista, la ricaduta cognitiva è indiretta ma c'è. Quindi io riesco ad avere un feedback immediato, più veloce su quello che sto facendo e so se sto prendendo una cantonata o sto lavorando bene sempre sulla base del double blind che mi viene fatto sugli stadi d'avanzamento che presento. In questo senso, assolutamente sì, può avere un effetto positivo anche da un punto di vista cognitivo perché comprime dei tempi che sarebbero dei tempi morti, dandomi la possibilità di usare al meglio il mio tempo” (R1U).*

Derivando spesso da progetti a scala medio-grande, le infrastrutture di ricerca offrono generalmente evidenti vantaggi sulla capacità di rappresentazione dei fenomeni in virtù di una copertura campionaria difficilmente eguagliabile tramite studi individuali (“*Sicuramente abbiamo raggiunto un campione notevole di più di 1700 \*\*\*, forse non avremmo potuto averlo*” (R2U). L'alto numero di record e variabili a disposizione è particolarmente apprezzato poiché offre una visione globale dei fenomeni oggetto di studio, sfruttabile in molteplici lavori, come testimoniato da due ricercatori: “*ho un'ampiezza tale che mi consente di elaborare una ricerca sicuramente molto più interessante rispetto ai risultati, non meramente esplorativa, con un numero di casi estremamente limitato, che va a compensare la carenza di dettaglio del dato*” (R2D); (“[Ho particolarmente apprezzato] *la completezza di alcune informazioni presenti. Ovviamente io avrei potuto scendere nel dettaglio di chiedere il singolo esame dato o non dato all'università. Era possibile tecnicamente e l'ho fatto in altri lavori. La completezza, come dicevo, delle informazioni presenti nel dataset*” (R1U). La ricchezza del dato produce inoltre consapevolezza su aspetti dei fenomeni studiati non presi inizialmente in considerazione: “*Quello che per me è stato interessante è che mi ha fatto vedere delle cose che non avevo mai visto. Sicuramente mi ha portato la possibilità di vedere delle sfaccettature che non avevo considerato*” (R3D); “*Queste indagini poi hanno questa caratteristica di avere dei moduli, loro li chiamano “nested”, “annidati”, dei moduli dentro altri che ti permettono di andare ad approfondire delle tematiche specifiche*” (R1U).

Un altro importante aspetto connesso alle infrastrutture e alla ricchezza informativa dei dati in esse contenuti riguarda l'opportunità di effettuare comparazioni a livello regionale o internazionale sfruttando l'omogeneità dei metodi di collezione. La potenzialità di comparazione, specialmente nei progetti europei, è avvertita a stadi successivi rispetto all'implementazione dell'infrastruttura: “*io ho avuto la raccolta dati per l'Italia: fin tanto che io ho lavorato sulla mia singola realtà non ho percepito le potenzialità come infrastruttura di ricerca. Quando le informazioni nazionali sono state integrate dalle altre informazioni dei*

*Paesi partner, allora effettivamente la ricchezza dei dati e la novità del dato che era contenuto mi ha dato, col tempo e con l'uso, anche la percezione delle potenzialità che avevano questi dati che erano stati raccolti...*” (R2D). A questo proposito, un ricercatore fa notare che oggi, grazie alla disponibilità delle infrastrutture, vi è una maggiore possibilità di comparare i dati tra Paesi piuttosto che confrontare i risultati che su quei dati si sono avuti: *“È una dimensione assolutamente fondamentale il fatto di avere dati così organizzati. Prima no [...], si comparavano i risultati. Mentre invece in questo caso io posso comparare gli odd-ratio, i modelli per nazione, addirittura io posso fare un modello dove la nazione è la variabile di controllo. Quindi vado a vedere una struttura latente e vedo poi come le nazioni influiscono... oppure posso “dummizzare” per gruppo di nazioni, e così via”* (R1U). Tutto ciò apre anche per l’ambito sociologico potenzialità maggiori per lo svolgimento di analisi comparative: *“Io mi ricordo che i grandi tentativi di analisi comparativa... non è che non ci sono stati prima, ma si contavano su una mano, e poi soprattutto la loro diffusione era abbastanza limitata, ad esempio grazie ad Eurobarometro. Invece adesso sta diventando, su ambiti anche tra loro diversi, un modo di lavorare più comune”* (R2U). Quando la copertura geografica della risorsa si accompagna anche ad una copertura dei dati nel tempo, ecco che emergono ulteriori possibilità per analisi longitudinali che possono soddisfare ulteriori bisogni cognitivi: *“Quello che abbiamo rilevato [utilizzando l’infrastruttura] è che ci sono, almeno negli ultimi dieci anni, alcune tendenze che emergono e che si modificano, anche se il periodo non è lunghissimo e quindi stiamo confrontando con quello che è successo nei precedenti dieci anni”* (R2D). Anche nell’ottica di un utilizzo dell’infrastruttura come puro supporto di confronto per le risultanze di studi individuali, è particolarmente apprezzata la presenza di dati che consentono l’apertura sia ad una prospettiva *cross-sectional* che ad una longitudinale: *“una ricerca di maggiore qualità perché è legata ad un discorso di comparabilità dei propri risultati con altri risultati che porta ad un rafforzamento della qualità del risultato”* (R1D).

Talvolta, l’approccio alle infrastrutture nella conduzione di uno studio ha dato la possibilità di concentrarsi non sulle grandi tendenze, bensì sui dati anomali o *outlier*, che hanno costituito per questa professoressa associata uno spunto di riflessione per arricchire il suo paper: *“[...] tirar fuori i dati piccoli... perché a me non interessano le grandi tendenze, perché sulle grandi tendenze se tu pensi alle ovvietà, quelle ovvietà sono sicuramente sostenute dai dati. Però a me interessa vedere i dati delle due punte perché lì si vedono i veri*

*cambiamenti. Quindi con \*\*\* [cita collaboratore] guardiamo i dati “piccoli”, non i dati “grandi” (R2D).* La possibilità di indagare a fondo su strutture latenti e dati imprevisi si incrocia con la concatenazione di diversi database, se l’infrastruttura utilizzata è di tipo distribuito *“Il mio interesse sul PISA è un interesse prettamente tematico, nel senso che soprattutto le ultime due wave sono estremamente ricche e poi concatenano diversi database, quindi mi permettono di testare delle ipotesi o dei segmenti di ipotesi che con altri tipi di database non posso analizzare... ad esempio, fare delle analisi multilivello, delle scomposizioni della varianza che prima non potevo fare, oppure incrociare le risposte degli studenti, dei loro genitori e degli insegnanti con dati ecologici” (R1U).* Infine, l’ampia disponibilità di dati può dar luogo a test di metodologie che possono incidere sul cambiamento del modo di fare ricerca del singolo ricercatore: *“Sicuramente [l’utilizzo dell’infrastruttura] è stato determinante. Senza i risultati di ricerca e senza testare la fecondità di questo modello metodologico non credo che l’avrei promosso, né avrei variato il mio modo di fare ricerca” (R2U).*

Quando il lavoro collaborativo su un’infrastruttura genera un volume collettaneo con l’approfondimento di diverse subtematiche, ecco che il plus qualitativo del singolo studio acquisisce rilevanza nel contesto dell’insieme delle indagini prodotte da un *team* di ricercatori. La prospettiva di integrazione e triangolazione della ricerca è stata esaltata in particolare da una professoressa associata *“Quello che mi ha dato in più [l’utilizzo dell’infrastruttura] non è quello che ho scritto io, ma è il fatto che abbiamo scritto in tanti sugli stessi dati. E quest’ottica, non di triangolazione ma di più, cioè di avere diversi punti di vista a partire dalla stessa risorsa di dati...non è tanto il mio capitolo o i miei capitoli, ma è il fatto che gli stessi dati sono risorsa per tante persone e possono essere letti in tanti modi, ovviamente avendo un obiettivo comune, un tema di fondo comune che era lo stile di vita nel tempo libero e la partecipazione culturale” (R2D).*

Quanto ai testimoni privilegiati, essi hanno inteso prevalentemente soffermarsi su benefici relativi all’*output* di ricerca riguardo il recepimento del prodotto da parte della comunità scientifica, come la credibilità connessa all’infrastruttura stessa che ricade sugli elaborati generati a partire da essa. Due ricercatori hanno dichiarato che eventuali obiezioni di carenza sulla qualità dei dati da parte di un *reviewer* potrebbero essere accantonate rispetto alla risonanza “mediatica” dell’infrastruttura e alla sua strutturazione progressiva come solida risorsa di ricerca: *“In molti casi la qualità dei dati non era abbastanza buona, ma abbastanza*

*buona per dimostrare che in linea di principio i metodi funzionano [...] La qualità dei dati nel frattempo è diventata così buona che non si possono semplicemente rifiutare i risultati sulla base di una tale semplice critica” e dal fatto che altri ricercatori hanno sfruttato la stessa risorsa per studi simili: “Se io dico che ho raccolto dati in 30 Paesi diversi da fonti diverse, il referee mi dice magari che ci sono dei problemi di comparabilità. Se io invece gli dico i dati vengono da \*\*\* [cita il nome di un’infrastruttura europea], ha un manuale, ha i sistematic data checks, i referee sono molto più calmi. [...] Quando l’infrastruttura è stata usata per 20-30 paper e questi 20-30 paper hanno “sputato” risultati ragionevoli, uno comincia a capire che tutto sommato il dato è utilizzabile [...] Il paper ha guadagnato in credibilità”. Altro aspetto particolarmente apprezzato riguarda la fruizione del complesso apparato esperienziale connesso all’implementazione dell’infrastruttura: a questo proposito, un ricercatore ha dichiarato che: “Trovare davanti a delle informazioni già metabolizzate e strutturate permette di costruire il discorso in maniera più sicura. Il prodotto di ricerca che vai a costruire, il prodotto tuo di ricerca è evidentemente supportato dall’esperienza già raccolta all’interno dell’infrastruttura”. Anche la dimensione della gestione professionale dell’infrastruttura comporta un vantaggio qualitativo sui prodotti di ricerca, ad esempio relativamente ai controlli di qualità sui dati condotti in maniera intensiva e sistematica, specie per le risorse di caratura europea: “È uno strumento gestito da informatici che quindi supporta una serie di servizi, tipo routine di controlli interni di qualità e così via, che sono state largamente automatizzate ed è standardizzata nel senso che c’è un manuale metodologico. [...] Il microdato è sufficientemente armonizzato”. Infine, un ultimo aspetto rilevato riguarda la crescente tendenza alla geolocalizzazione del dato presente nelle grandi infrastrutture, circostanza che consente di rispondere a certe domande di ricerca a cui non si pensava in partenza, per esempio quelle legate alla localizzazione delle unità statistiche.*

#### *4.4.2 Criticità nell’utilizzo delle infrastrutture precostituite*

A controbilanciare le evidenze positive dichiarate dagli intervistati rispetto alle loro esperienze, l’ultima parte dell’intervista ha inteso stimolare nei ricercatori una riflessione su eventuali criticità, problemi o svantaggi connessi all’uso di infrastrutture di ricerca. In particolare, rilevato l’utilizzo non prevalente di infrastrutture di interesse europeo, o comunque precostituite a prescindere da progetti specifici, rispetto allo sfruttamento di risorse a cui gli stessi intervistati hanno concorso alla creazione, l’interesse si è rivolto

principalmente a raccogliere l'opinione sui *caveat* relativi all'accostamento a questo tipo di infrastrutture.

Un primo ordine di possibili criticità ha fatto riferimento a una deriva per così dire “conformista” della produzione scientifica ed è stato evidenziato in particolare da quattro ricercatori. Il primo paventa il rischio per cui “*ci si appiattisce su quello che già preesiste e si rinuncia a condurre magari percorsi di ricerca più stimolanti. Infatti credo che sia cresciuto il numero di contributi... parlo di articoli su riviste, prevalentemente incentrati su analisi secondarie dei dati... e siano diminuite quelli incentrati sui dati raccolti*” (R2U). Secondo due altri giovani ricercatori, il disegno della ricerca di chi ha progettato e poi svolto le *survey* contenute nei grandi database internazionali produce una standardizzazione definitoria delle unità di analisi e degli indicatori che può, per certi versi, rivelarsi un vantaggio in termini di comparabilità della ricerca, ma che provoca anche un condizionamento sulla fase creativa connessa al disegno delle investigazioni. Se infatti “*l'accessibilità a fonti di dati secondari, anche di grande densità e di grande ampiezza, ha sicuramente il vantaggio di rendere la ricerca sociale avvantaggiata rispetto alla produzione di conoscenza... cioè non si parte più da zero, oltre ad offrire possibilità di elaborazione ad hoc che potrebbero produrre dei contributi di ricerca di rilievo*” (R1U), d'altra parte “*poter accedere direttamente ad un database significa perdere nei termini di progettazione autentica, originale*” (R1D). Molto significativa, infine, è la testimonianza di un frequente utilizzatore dell'infrastruttura internazionale OCSE-PISA, il quale rileva che l'appiattimento degli obiettivi cognitivi derivato dalla disponibilità delle infrastrutture può andare di pari passo con quello delle metodologie di interrogazione dei dati e degli apparati teorici che fanno riferimento alle ricerche: “[...] *buona parte di questi database sono pensati per certe finalità; tu puoi fare anche altre cose, però molto spesso nella letteratura internazionale si è anche sviluppata una certa forma di conformismo. Quindi chi fa analisi sul PISA usa certa letteratura, con determinate metodologie, e se vai fuori da quello sei out*” (R1U). D'altronde però, continua lo stesso ricercatore, il fatto di piegare le esigenze cognitive alla disponibilità di infrastrutture di ricerca disponibili è, soprattutto per giovani ricercatori, un'esigenza che scaturisce dalla difficoltà di finanziamenti per ricerche basate su dati primari: “*questo ha poi a che fare con i finanziamenti della ricerca. In Italia è difficilissimo creare survey ex novo... cioè... perché o riesci a prendere un FIRB o un SIR, altrimenti con un progetto d'Ateneo o qualcos'altro tu non riesci a farlo*” (R1U).

Un secondo ordine di perplessità ha a che fare con la fiducia riposta dal ricercatore nel dato già organizzato, a maggior ragione se vi è una scarsità di metadati a supporto dell'infrastruttura. Soprattutto i ricercatori più esperti segnalano che *“il problema più grosso è quello sulla qualità degli indicatori...nelle banche dati c'è un problema grossissimo che è la qualità degli indicatori e come sono costruiti”* e ancora *“non tutti i dati, materiali e documenti sono di qualità. Quindi, dovrebbe esserci un controllo a monte prima di inserirle in queste infrastrutture, in modo che il ricercatore si senta invogliato ad utilizzarle perché sa che al loro interno può trovare informazioni che siano credibili”*. Ci si chiede *“come si fa a rendere ispezionabile una ricerca”* dato che *“lo svantaggio più grosso è che ti devi fidare perché non hai il controllo della situazione. E questo è un problema”* (R3U). Insomma, per alcuni ricercatori l'affidabilità e la robustezza del dato sono elementi che non vanno dati per scontati ma andrebbero verificati di volta in volta; e in questo senso l'esperienza viene considerata quale elemento dirimente nel valutare la credibilità della fonte (*“anche in questo campo l'esperienza, la sensibilità di riconoscere problemi da segnali indiretti è decisiva”*) (R2U). Il pensiero dunque va ai giovani che dovrebbero abbandonare, secondo le parole di un altro ordinario, la mentalità da *“supermercato del microdato”* dal momento che *“il microdato va spiegato perché lo chiedi, all'interno di quale contesto argomentativo e di ricerca, e va quindi costruito. Anche il dato dell'infrastruttura cosiddetta tale è un dato costruito”* (R3U). Ancora, nelle parole di un professore associato emerge la necessità di un approfondimento del rapporto tra l'infrastruttura e il suo contesto di produzione: *“Da un punto di vista degli svantaggi, quello a cui bisogna stare attenti è il modo in cui questi dati vengono costruiti quindi [...] stare attenti ai metadati, quindi come sono stati selezionati, rilevati, quali sono stati gli strumenti di rilevazione, qual è stata la popolazione di riferimento, il periodo d'indagine, anche andare a reperire i questionari con cui i dati sono creati perché il dato, come dico sempre ai miei studenti, non ha autonomia semantica, ma il significato che possiamo dare al dato dipende dal modo in cui è stato costruito”*. Se è vero che il dato non possiede autonomia semantica, i pericoli si amplificano quando si propende per un cieco approccio induttivista tale per cui si perde il senso stesso della ricerca, come nota questo giovane ricercatore: *“Il rischio è che uno s'innamora tanto del dato e dell'analisi del dato che si dimentica che, come diceva il buon Franco Ferrarotti, “La sociologia è una scienza empirica teoricamente orientata [...] cioè noi creiamo delle teorie e delle ipotesi e le andiamo a testare empiricamente, se vogliamo fare seriamente il nostro lavoro. Il lavoro teorico non*

*può essere eliso e messo in secondo piano. L'uso di [...] queste grandi fonti di dati e così via, corre il rischio di mettere al primo posto l'aspetto tecnico procedurale – lo vedo coi miei tesisti della magistrale – per cui gli studenti vogliono subito fare il modello di regressione, stimare, invece dico “no, bisogna prima fare la domanda!” (R1U).*

Solo in pochi casi si è fatto riferimento a difficoltà contingenti relative all'utilizzo *pratico* delle infrastrutture di tipo europeo o scaturite da progetti esterni. Un professore ordinario ha segnalato che le competenze informatiche richieste talvolta non appaiono in linea con quelle possedute da parte dei ricercatori, anche se nativi digitali: *“c'è il rischio che entrare all'interno dei processi delle infrastrutture di ricerca innovativi richieda una conoscenza che non è solo quella naturale per un digital native o per un giovane ricercatore che è più esperto ad utilizzare gli strumenti dei suoi maestri” (R3U)*. Tale concetto è stato ribadito da un professore associato che ha sottolineato inoltre che il tempo di adattamento per la comprensione dell'utilizzo di un'infrastruttura può essere abbastanza lungo: *“[queste infrastrutture sono] numerose e richiedono spesso una specializzazione e una conoscenza tale che prima, ad esempio, non era necessaria per le infrastrutture di ricerca tradizionali. Per cui in termini di tempo, di conoscenza, anche di applicazione richiedono ovviamente un maggior impegno del ricercatore” (R2U)*. Infine, un primo ricercatore ha riferito come alcune di esse si configurino quali “compartimenti stagni” incapaci di comunicare con altre risorse, sollevando un problema di interoperabilità: *“certo poi ci sono questioni di interoperabilità tra le varie infrastrutture, l'assenza di interoperabilità tra queste è un problema” (R3D)*.

Il fenomeno dell'autolimitazione dello sviluppo di idee originali e dello sfruttamento acritico delle informazioni è ribadito da uno dei testimoni privilegiati, che ammonisce di *“non prendere quell'infrastruttura come oro colato o sopravvalutare l'infrastruttura rispetto alla visione personale”* ed emerge nelle riflessioni di un altro testimone, quale fattore foriero di virtuosità nella rapidità di accesso a informazioni su grandi campioni, ma anche di appiattimento sulle variabili disponibili. Quest'ultimo aspetto è ribadito anche dal terzo testimone che fa notare come rispetto a obiettivi cognitivi di grande portata è possibile che *“ogni tanto [nelle infrastrutture disponibili] possono mancare delle variabili...”*, tuttavia rivelando infine un atteggiamento molto pragmatico: *“ad un certo punto quando uno deve pubblicare deve essere pragmatico...uno tira fuori un oggetto che è pubblicabile”*. Il pragmatismo legato all'esigenza generale di pubblicare si riverbera anche nell'atteggiamento rispetto alla robustezza del dato in essa contenuto: *“Un esempio è la \*\*\*: è un caso classico*

*dove c'è un migliaio di paper che usano gli stessi dati senza porsi il problema. Quindi c'è il rischio che certi problemi metodologici dei dati vengano dimenticati [...] in certi casi uno si arrende, e quindi non li affronta e basta, tanto alla fine è la pubblicabilità che conta”.*

#### **4.5 Caratteristiche desiderabili delle infrastrutture**

Nel corso delle interviste, i ricercatori si sono espressi su quali dovessero essere, secondo la loro visione, i requisiti e le caratteristiche che un'infrastruttura di ricerca dovrebbe possedere per essere considerata una risorsa desiderabile per condurre la propria attività scientifica. Sullo sfondo di una percezione di comunicazione insufficiente rispetto alle infrastrutture già disponibili a livello nazionale e internazionale (si paventa la creazione di “*un portale che racchiuda per argomenti i dataset disponibili a livello mondiale o europeo*” – R1U), gli intervistati hanno posto in modo marcato l'attenzione su temi quali l'accesso aperto, la qualità del dato e la possibilità di interconnessioni interne ed esterne all'infrastruttura.

Quanto al tema dell'accesso, i ricercatori sociali presentano opinioni diversificate rispetto al grado desiderabile di apertura delle infrastrutture. C'è chi, come una professoressa associata, si è schierato per un accesso liberalizzato: “*Secondo me l'accesso dovrebbe essere...a me verrebbe da dire libero, cioè quest'accesso dovrebbe essere il più possibile aperto*” (R2D), e chi, come un professore ordinario, prospetta esplicitamente l'inclusione nel target delle infrastrutture di non addetti ai lavori, in nome di un approccio alla *citizen science* (cfr. par. 1.3) che comunque dovrebbe pur far leva sull'intermediazione del sistema scientifico nel rapporto tra dati e pubblico, perché: “*la disponibilità dovrebbe essere sia pur mediata ma resa di pubblico dominio, affinché le persone sappiano [...] se avessimo avuto la disponibilità costante di qual è, per esempio, la “rischiosità” sul territorio dei vaccini, ci saremmo resi conto che alcune patologie sono endemiche, che in realtà quello che tutti temono è in forte calo*”. Un approccio più cauto è emerso dalle parole di un'assegnista di ricerca, che si è manifestata più prudente rispetto all'accesso generalizzato, prospettando invece margini di controllo sull'utilizzo del dato: “*riguardo all'accesso lo vorrei aperto ma non so se tutti agli utenti o solo alla comunità scientifica per intero. Questa è una domanda che mi pongo e a cui non so dare ancora una risposta. Una modalità di accesso aperta ma*

*controllata, nel senso che se io accedo devo comunque dare dei miei riferimenti, nel senso che se uso i dati impropriamente devo essere individuato” (R1D).* Due intervistati hanno adottato la prospettiva di un approccio libero all’interno della comunità scientifica, seppur legato a vincoli di “contribuzione scientifica”. I modelli proposti concordano sul principio di lealtà, basato sul fatto che: *“non si può semplicemente utilizzare un’infrastruttura senza contribuire allo sviluppo di questa infrastruttura [...] anche semplicemente diffondendo quest’infrastruttura, per esempio allertando o comunicando a colleghi l’esistenza di quell’infrastruttura, le possibilità che offre, provando nei limiti delle proprie possibilità a svilupparne gli aspetti” (R2U);* in particolare, per lo sfruttamento aperto a terzi di infrastrutture nate nell’ambito di progetti, una ricercatrice propone di offrire uno “scambio di competenze” *“per produrre un’analisi particolare che io non so fare e che dia a te ovviamente visibilità ma che porti il gruppo a produrre un risultato su un versante che non saprebbe raggiungere” (R1D).* Un giovane ricercatore si è soffermato sulla necessità di rendere gli accessi alle infrastrutture sempre gratuiti, a fronte della scarsità di finanziamenti imposta dal sistema della ricerca italiano, mentre un ricercatore più esperto ha espresso una preferenza per infrastrutture che *“non richiedano una fortissima specializzazione e che non pongano delle barriere di tipo sia strutturale, funzionale che di replicabilità” (R2D).* Su questa scia, in riferimento all’abilità tecnica nell’utilizzo delle infrastrutture, un primo ricercatore ha ribadito che l’infrastruttura debba essere *user-friendly* e consentire all’utente una consultazione quanto più possibile autonoma. Riguardo ancora l’usabilità delle infrastrutture, un ricercatore *junior* ha sostenuto il problema del legame commerciale tra alcune infrastrutture e specifici *software* informatici che consentono di fruire dei loro contenuti: *“Uno dei rischi grandi di queste infrastrutture è che ci sia il prevalere degli interessi commerciali di chi sviluppa grandi software” (R1U).*

Sul lato della qualità dei contenuti, si distinguono le dimensioni dell’affidabilità del dato, legato a doppio filo con quello della trasparenza dei processi di produzione dello stesso, e della tempestività degli aggiornamenti. Prevedibilmente, più volte viene sottolineato come il dato debba essere attendibile, perché *“gira tanta di quella immondizia sul web...” (R3D)* o comunque provenire da fonte autorevole. Tuttavia, quel che emerge con più forza è la richiesta della disponibilità di metadati a corredo dell’infrastruttura che garantiscano l’ispezionabilità rispetto ai metodi di raccolta del dato: *“parlando di ricerca sociologica, non so come si è arrivati a un dato campione, non so come sono state condotte determinate*

*interviste. Io devo avere tutte le conoscenze utili a che possa valutare nei termini del livello di qualità quel database. Se quel database mi convince, io ho risparmiato molto a monte e posso effettivamente progredire più velocemente per quello che sono i miei interessi di studio” (R1D).* L’idea è quella che la forza dell’infrastruttura coincida con la sua credibilità, e tale aspetto è legato al controllo delle fasi della ricerca che hanno portato alla sua implementazione: *“un’infrastruttura di ricerca, se deve essere di supporto all’attività di ricerca, [...] ovviamente il ricercatore ha necessità di comprendere anche come si è evoluto un processo di ricerca e quindi di vederne tutti gli step non solo l’output finale [...]” (R3U).* Nell’esplicitazione dei processi, è vista anche la possibilità di annidamento di un *know-how* replicabile per successive esperienze di ricerca: *“insomma il know-how presente in quella struttura specifica può essere importante per uno studioso che intende replicare o a fare una cosa diversa ma partendo da quello che è già stato fatto” (R1U).* L’aggiornabilità, infine, è paventata quale requisito auspicabile ed essenziale per sottrarre l’infrastruttura all’obsolescenza: *“[una caratteristica desiderabile delle infrastrutture è] la tempestività dell’aggiornamento del dato...altrimenti sono inutili. L’impossibilità di rendere il dato aggiornato ha reso quasi inutile il \*\*\* se non per indagini retrospettive” (R2D).*

Tra le caratteristiche desiderabili delle infrastrutture, alcuni ricercatori indicano, inoltre, la possibilità di interconnessioni interne ed esterne. Col primo termine si intende, riprendendo le parole di un professore associato, la possibilità di creare delle piattaforme informative che siano in grado di coniugare dati *“con una struttura diversa, nel senso che provengano da strumenti di rilevazione diversi”*, perseguendo una *“logica di integrazione fra approcci che si sono sviluppati in maniera antagonista” (R2U).* La logica di un approccio integrato nella costruzione di contenuti interni è atta a stabilire collegamenti tra livelli di analisi differenti nell’ambito dello studio del medesimo fenomeno e tra diverse tipologie di dati, andando a superare certi steccati legati ai paradigmi classici delle scienze sociali: *“credo invece molto nella possibilità che da questo paradigma scissorio, che vede anche l’antagonismo tra approcci quantitativi e qualitativi, si possa passare ad un paradigma invece integrato e quindi i requisiti che vedo nelle infrastrutture è che si muovano su più livelli di analisi combinando approcci che si sono sviluppati in maniera separata” (R2U).* Sul fronte delle interconnessioni esterne, è più volte ribadita la necessità che le infrastrutture siano potenzialmente interoperabili con altre, così anche da favorire potenzialità interdisciplinari negli interessi di ricerca: *“dal mio punto di vista, riprendendo il tema a me molto caro*

*dell'interdisciplinarietà, se si costruissero delle infrastrutture che dialogano e fanno dialogare le discipline, cosa che fanno forse meglio di quelle tradizionali, se aumentassero questa loro potenzialità, secondo me avrebbero vinto la loro sfida” (R2D).*

I testimoni privilegiati di questo studio hanno fornito un punto di vista peculiare sulle caratteristiche desiderabili delle infrastrutture, derivato sia dalla loro esperienza personale di utilizzo che dal rapporto con altri utenti dell'infrastruttura. Il primo testimone ha fatto cenno al bisogno di aprire l'infrastruttura a quanti più scienziati interessati, garantendo “operatività pratica” nel caso di piattaforme virtuali, “cioè deve essere un sito veloce dove le informazioni sono recuperabili in maniera rapida ed efficace”. La capacità di interoperatività con altre infrastrutture, anche afferenti a differenti discipline, è posta in primo piano insieme alla garanzia di validità del dato. Le caratteristiche di accessibilità e usabilità sono state ribadite anche dal secondo testimone, che però fa emergere principalmente un'istanza di professionalizzazione nella gestione e nella politica di mantenimento. Vengono auspicati “un ragionevole livello di standardizzazione [e] degli standard di consistenza interna, di qualità metodologica, di documentazione che devono essere necessariamente più alti”, insieme ad una capacità di aggiornamento che deve costituire la cifra distintiva di una risorsa infrastrutturale di tipo europeo. Il terzo testimone ha fatto leva sul bisogno di legittimare gli investimenti nella creazione di un'infrastruttura di ricerca attraverso la misurazione del ritorno per la società nel suo complesso in termini di acquisizione di conoscenza spendibile, ma si è rivelato prudente riguardo al grado di apertura desiderabile, specie riguardo all'ideale della *citizen science*, citato più sopra, considerato ancora difficilmente attuabile: “Penso che dobbiamo stare attenti all'apertura [...], è in genere qualcosa che può essere molto utile, molto produttivo anche per assicurarsi che i risultati della ricerca siano dei “doni” per la società [...]; dall'altra parte, è un modo per cercare di garantire la qualità scientifica e la validità di ciò che stiamo facendo. [...] Ciò tuttavia implica requisiti abbastanza elevati da parte dei professionisti, dei cittadini, delle aziende, qualunque attore coinvolto nel processo. Penso che questo sia ancora veramente poco chiaro, un problema che deve essere risolto ... penso che sia più un'idea che la realtà”.

## CONCLUSIONI

Il tema del cambiamento nei modi di produzione della conoscenza per effetto dello sfruttamento delle infrastrutture di ricerca è stato ancora poco affrontato, e la letteratura è caratterizzata da un numero limitato di approfondimenti in relazione alle ripercussioni nell'ambito delle scienze non dure. Il presente lavoro, muovendo da una caratterizzazione della pratica di ricerca contemporanea fondata su caratteri di condivisione e apertura del dato, potenziamento delle strutture interattive tra ricercatori e innestata su una progressiva ricezione delle innovazioni tecniche in campo ICT, ha inteso implementare uno studio esplorativo sugli effetti generati dallo sfruttamento delle infrastrutture di ricerca nel lavoro degli scienziati sociali. Nello specifico, la *domanda di ricerca* da cui lo studio è originato ha inteso indagare, attraverso l'implementazione di uno *studio di caso* applicato alla comunità dei sociologi e riferito al maggior polo universitario e al maggior ente di ricerca pubblico italiani, su quale fosse il cambiamento indotto dall'utilizzo di tali risorse nelle attività che presiedono alla produzione di nuova conoscenza. In particolare, sono stati approfonditi gli effetti sulle strategie di ricerca, sull'instaurazione di collaborazioni, sulla tendenza all'interdisciplinarietà. Un'attenzione particolare è stata inoltre riservata a quale fosse la valutazione, da parte dei ricercatori sociali, delle ricadute sui prodotti della ricerca derivati dallo sfruttamento di tali risorse in termini di vantaggi o problematiche. Sono state realizzate 15 interviste semi-strutturate a ricercatori sociali di diversa età, genere e posizione accademica e a 3 testimoni privilegiati che potessero costituire un *benchmark* di riferimento su pratiche avanzate. L'analisi sistematica dei contenuti testuali è stata svolta secondo i metodi dell'*analisi tematica* (Braun e Clarke, 2006). I risultati della ricerca sono stati riassunti nelle sezioni che seguono.

### *Cosa sono le infrastrutture di ricerca nella visione degli scienziati sociali?*

Le definizioni di infrastruttura di ricerca proposte a livello nazionale (Programma Triennale della Ricerca 2010-2012) e internazionale (European Commission, 2008; European Science Foundation, 2013) includono una moltitudine di referenti, di cui solo un sottoinsieme è ascrivibile al campo delle scienze sociali (German Council of Science and Humanities, 2011). Seguendo alcuni osservatori, le infrastrutture di ricerca sono percepite dai ricercatori in modo vario e nebuloso: per alcuni rappresentano un *framework* organizzativo che abilita

alla collaborazione e alla condivisione di dati e risultati di ricerca; per altri si sostanziano nel contenuto informativo a cui è offerto un accesso (*cf.* Renschler *et al.*, 2013; Farago, 2014). Nell'ambito della somministrazione delle interviste non sono state utilizzate definizioni operative preliminari del concetto di infrastruttura di ricerca per mantenere aperto il carattere esplorativo dell'indagine. L'analisi ha reso evidente come tra gli scienziati sociali non esista un'unica opzione definitoria generale, precisa e condivisa, ma tanti significati quanti quelli che i ricercatori derivano da visioni personali sull'assetto del lavoro scientifico e quanto puntualmente sperimentato nell'esperienza di ricerca. La consapevolezza di aver utilizzato una risorsa definibile come infrastruttura di ricerca può tra l'altro pervenire anche solo a posteriori, a margine di una riflessione sul lavoro collaborativo intrapreso. A fronte delle varie declinazioni (*cf.* 4.1), è tuttavia emersa una generale tendenza verso l'inquadramento delle infrastrutture di ricerca in termini di risorse informative che supportano il percorso della ricerca e veicolano le relazioni tra gli scienziati: *database* derivati da progetto o resi disponibili a livello internazionale, banche dati, archivi di dati sono stati qualificati quali aggregatori di competenze e informazioni, la cui sempre più ampia diffusione sta consentendo di svolgere il lavoro scientifico in modo diverso rispetto al passato, abilitando alla produzione di nuova conoscenza attraverso un'impresa collaborativa (laddove l'infrastruttura sia co-creata) o tramite l'accesso a dati di seconda mano (laddove si fruisca di un'infrastruttura terza). Tali referenti sono stati talvolta designati quali "oggetti frontiera", che assumono significato in quanto elementi del sistema sociotecnico della ricerca che collegano ambienti, strumenti e personale di ricerca al fine di perseguire l'avanzamento della conoscenza.

#### *Che tipo di infrastrutture utilizzano e per quali fini?*

Nel panorama delle scienze sociali si riscontra sempre più frequentemente l'utilizzo di infrastrutture di ricerca. Sebbene a livello internazionale le istituzioni scientifiche premano verso l'utilizzo di infrastrutture per le scienze sociali armonizzate disponibili da remoto, come quelle che stanno sorgendo in ambito europeo (ESFRI, 2016), spesso i ricercatori lavorano su infrastrutture isolate, derivanti dall'implementazione di progetti specifici, mirate alla domanda locale e non coordinate con altre basate su tematiche analoghe (Dutton e Meyer, 2009; *cf.* par. 2.3). Una volta ravvisata quale sia la rilevanza dello sfruttamento delle infrastrutture nella comunità degli scienziati sociali del caso studio ci si chiede se, dato il *trend* al rialzo delle pubblicazioni derivanti dallo sfruttamento delle risorse di interesse

europeo (*cf.* Fitzgerald, 2016), i ricercatori si siano rivolti a questo tipo di risorse o abbiano continuato a privilegiare quelle derivanti da progetto. Dalle interviste è emerso che l'utilizzo delle infrastrutture ha prodotto un impatto crescente nel tempo nella generazione di nuovi prodotti della ricerca, sebbene sia stata riscontrata un'incidenza più limitata nell'esperienza degli intervistati più avanti nella carriera. Tra coloro che si trovano a un livello di carriera d'entrata o mediano, l'impatto sulla produzione scientifica è stato percepito come rilevante, anche in ragione del fatto che la scarsità di risorse finanziarie per il settore della ricerca si è tradotta nella necessità di sfruttare risorse di dati precostituite e accessibili da remoto. Nonostante un discreto numero di intervistati (soprattutto i più giovani) abbia indicato l'utilizzo in carriera di infrastrutture di ricerca di interesse europeo, come l'Eurobarometro, di caratura internazionale come i database derivanti da indagini OCSE, e l'utilizzo delle banche dati nazionali o internazionali sia ricorso un discreto numero di volte, sono state le infrastrutture *project-driven*, sviluppate nell'ambito di progetti nei quali i ricercatori fossero coinvolti, a dimostrarsi le più utilizzate. Mentre le infrastrutture derivate da progetto hanno quasi sempre generato studi di tipo analitico, le infrastrutture di interesse europeo/caratura internazionale e le banche dati, oltre che a fini di elaborazione dei dati in esse contenuti, sono state utilizzate come supporto di sfondo per motivi di semplice consultazione o per avvalorare metodologie da applicare ad altri insiemi di dati. Il riscontro della prevalenza di utilizzo di infrastrutture derivate da progetto è spiegabile in prima istanza con la poca informazione nell'ambito delle scienze sociali soprattutto sulle risorse di interesse europeo attualmente disponibili. Ma vi è un altro aspetto da tenere in considerazione: l'istanza del controllo processo di formazione del dato può ancora rivestire ancora un certo peso tra gli scienziati sociali, soprattutto su quelli più esperti.

*Le infrastrutture esercitano un'influenza sul processo di ricerca in termini di progettazione?*

Nel quadro della nuova scienza collaborativa, le infrastrutture precostituite o co-create offrono nuove opportunità intellettuali, consentendo di approfondire domande di ricerca non perseguibili dai ricercatori attraverso studi implementati in modo individuale (Nielsen, 2011, tr. it. 2012). Il "diluvio di dati" consentito dal proliferare degli archivi digitali implica inoltre una produzione di risultati scientifici fondata sempre più spesso su procedimenti induttivi basati sul dato e gestiti grazie a nuovi strumenti tecnologici e risorse digitali per la ricerca (Hey *et al.*, 2009). Questa tendenza, riconoscibile sul fronte delle scienze dure (si vedano i

numerosi studi derivati dallo sfruttamento della *GenBank*), è presente anche nel campo delle scienze sociali? In che modo la disponibilità *ex-ante* o la creazione *ex-novo* di infrastrutture di ricerca produce effetti sul modo di impostare il disegno d'indagine? Nonostante la prevalenza di utilizzo delle infrastrutture derivate da progetto, nei ricercatori sociali fa breccia un generale atteggiamento verso la preliminare verifica della disponibilità di basi di dati preorganizzate, da prendere eventualmente in considerazione rispetto all'implementazione di una collezione di dati originali, soprattutto in un periodo di carenza di finanziamenti adeguati (decisivi nella scelta di "disegnare" studi attraverso infrastrutture di ricerca già precostituite si rivelano fattori quali la convenienza economica e il risparmio di tempo). Tale verifica può precedere allo sfruttamento dei dati dell'infrastruttura per la produzione di studi analitici (nei casi dei ricercatori più giovani), all'alimentazione parziale delle nuove infrastrutture *project-driven*, ma può anche essere utile solo a livello di supporto di sfondo per l'ideazione di studi futuri. A dispetto di questa tendenza, nelle testimonianze dei ricercatori più esperti si è ravvisato un attaccamento ad uno stile di ricerca che rifugge dal riutilizzo di dati secondari. Coloro che hanno utilizzato infrastrutture scaturite da progetto, hanno indicato la piena coincidenza tra disegno della ricerca e progettazione dell'infrastruttura, evidenziando il *condizionamento* della sua disponibilità sulla possibilità stessa di effettuare gli studi riportati nelle narrative. Il dato più interessante che è emerso dall'analisi (anche nelle parole dei testimoni privilegiati) consiste nel fatto che l'infrastruttura utilizzata, sia essa precostituita o generata da progetto, ha favorito una modalità di approccio maggiormente induttivista rispetto al passato nell'implementazione degli studi, tale per cui domande di ricerca non previste inizialmente nei disegni sono state create *ex-post*, per via della conduzione di ragionamenti e riflessioni "dentro le matrici". Ciò ha tendenzialmente consentito, secondo gli intervistati, una maggiore libertà creativa nel processo di ricerca stimolando molteplici ridefinizioni degli obiettivi cognitivi e talvolta la sperimentazione di nuove metodologie o tecniche di analisi.

*Quali sono gli effetti delle infrastrutture sulle dinamiche di collaborazione?*

In un contesto di crescente ricezione delle evoluzioni tecniche in campo ICT sia per fini comunicativi che computazionali (*cfr.* par. 1.2.2), la creazione e la disponibilità di risorse condivise quali le infrastrutture di ricerca stanno incentivando la produzione di lavori collaborativi e influenzando l'organizzazione del lavoro scientifico (Nielsen, 2011, tr. it.

2012; ESFRI, 2016). Alcuni studi riferiti a settori di *hard science*, hanno rivelato che la condivisione di *database* di ricerca può stimolare la formazione di nuovi legami senza tuttavia sostituire i *pattern* di collaborazione già formati e le interazioni a livello locale (Hine, 2006; Binz-Schaft *et al.*, 2014). Nell'ambito di un'analisi di impatto sulle infrastrutture di ricerca nel panorama scientifico effettuata da Hüsing *et al.* (2010, *cf.* par. 2.3.1), si sono riscontrate differenze nell'effetto delle infrastrutture sui processi collaborativi tra i diversi campi scientifici: in particolare è stata rilevata una minore influenza sulla collaborazione evidenziata dagli studiosi afferenti alle scienze umane e sociali<sup>76</sup>. Sulla base di queste considerazioni, si è supposto che le infrastrutture possano sostenere il *networking* dei ricercatori sociali, supportando le reti di collaborazione già formate. Tuttavia, dato il crescente utilizzo delle infrastrutture, si è inoltre ipotizzato che predispongano alla formazione di nuovi legami, possibilmente travalicando i confini geografici delle organizzazioni di appartenenza.

Dall'analisi delle interviste è emerso un quadro quantomeno interlocutorio sul piano della produzione di *output* collaborativi per effetto dell'utilizzo delle infrastrutture. Se i testimoni privilegiati hanno riconosciuto le infrastrutture di ricerca quali piattaforme che possono assicurare continuità a legami esistenti e contribuire alla nascita di nuovi, i ricercatori inclusi nel campione di questo studio hanno riportato esperienze di vario tipo. Nell'ambito delle infrastrutture nate da progetto, gli intervistati, pur riconoscendo che il lavoro comune di implementazione possa *potenzialmente abilitare* alla nascita di nuove collaborazioni, hanno esperito la creazione di legami, perlopiù a livello di dipartimento o a livello nazionale, e raramente a livello internazionale (a seconda del livello territoriale dei progetti), non in tutte le esperienze, e - in tutti i casi positivi - con colleghi che avevano concorso alla creazione comune dell'infrastruttura e quindi già presenti nel loro *network*. L'esclusione di colleghi esterni è spiegabile con la circoscrizione all'accesso alle infrastrutture alla sola *partnership* che ha dato vita al progetto, circostanza connaturata ai progetti nazionali ed europei. Al netto delle esperienze positive, la corrispondenza tra collaborazione nell'implementazione dell'infrastruttura e collaborazione nella produzione di *output* di ricerca non è apparsa affatto scontata o automatica, e spesso riconosciuta solo sulla carta, derivando dalle caratteristiche del progetto entro cui l'infrastruttura nasce o, come avviene tipicamente, dalla convergenza degli obiettivi cognitivi e da un'omogeneità di visione circa l'ontologia del dato. In definitiva,

---

<sup>76</sup> Si ricorda tuttavia che gli stessi autori hanno qualificato questi risultati come poco significativi a causa del basso numero di risposte ricevute nell'ambito della *survey* per alcuni ambiti disciplinari (*cf.* 2.3.1).

nella maggior parte dei casi l'infrastruttura ha rappresentato il veicolo per sostanziare un'istanza collaborativa già indotta *ex ante* dalla partecipazione comune ai progetti, dimostrando di aver rafforzato *network* già esistenti attraverso la creazione di *output* comuni. Quanto all'instaurazione di *nuove* collaborazioni nella creazione di *output* a partire da infrastrutture precostituite di rilevanza internazionale, quest'evenienza è stata citata esplicitamente solo in un caso, da un giovane ricercatore. In generale, la disponibilità "esterna" di infrastrutture di ricerca nell'ambito delle scienze sociali, specie di quelle di interesse europeo, non sembra aver impresso un significativo segnale di discontinuità nell'instaurazione di *nuove* dinamiche aggregative, se non in un numero limitato di casi.

### *E sull'inclinazione alla ricerca interdisciplinare?*

Molti dei temi di ricerca che si porranno all'attenzione degli scienziati sociali, come quelli legati alle *grand challenge* del mondo contemporaneo, richiederanno una stretta collaborazione con ricercatori afferenti ad altre discipline (OCSE, 2013). Molte delle infrastrutture di interesse europeo sono state già concepite come programmi a lungo termine che incoraggiano l'interdisciplinarietà, portando a convergenze nel processo di produzione della conoscenza nel tentativo di superare le frontiere tra discipline individuali (ESFRI, 2016). Si è ipotizzato in questo studio che la disponibilità di infrastrutture consenta ai ricercatori di ampliare il panorama informativo incidendo sulla formazione di domande di ricerca orientate alla produzione di studi interdisciplinari. Sebbene dalle interviste sia emersa una presa di coscienza del valore della pratica della ricerca interdisciplinare nelle scienze sociali e si siano sottolineate le potenzialità della condivisione di risorse informative nella produzione di studi interdisciplinari, la grande maggioranza degli *output* derivati da infrastrutture si è declinata in lavori orientati alla materia di competenza. La produzione di prodotti interdisciplinari conseguenti a sfruttamento di infrastrutture di interesse europeo non è addirittura emersa in alcun caso; mentre è stata ravvisata in un solo caso riguardo all'utilizzo di basi di dati precostituite di livello internazionale (dati World Bank). Nei casi delle infrastrutture derivate da progetto, se l'opzione di interdisciplinarietà è stata perseguita già in fase di creazione di un'infrastruttura e nella fase della raccolta dati, la risorsa è stata solo in pochi casi sfruttata in modalità poliprospectica. Ne consegue dunque che, anche quando presente, la possibilità di generare un lavoro interdisciplinare non appare essere stata colta dagli intervistati in tutte le sue potenzialità a causa o della formazione di domande di ricerca

molto mirate (l'opzione di interdisciplinarietà è spesso tralasciata anche dai testimoni privilegiati) o della carenza di competenze complementari nel gruppo di ricerca. L'istanza di interdisciplinarietà non appare dunque essere stata particolarmente stimolata per effetto delle infrastrutture di ricerca nel caso di studio esaminato.

*Quali vantaggi nella produzione di nuovi output di ricerca i ricercatori sociali riconoscono come effetto delle infrastrutture?*

La modalità di produzione della conoscenza fondata sulla creazione e condivisione di infrastrutture di ricerca introduce l'opportunità di sfruttare nuovi metodi e strumenti informatici che consentono a singoli ricercatori e *team* di ricerca di condurre analisi approfondite e di carattere vario su volumi di osservazioni non paragonabili col passato. L'implicazione più diretta che è possibile derivare da tale quadro riguarda la possibilità di produrre *output* connotati da una maggiore qualità. Nonostante il report del 2010 di Hüsing *et al.* abbia rilevato tra gli scienziati sociali un entusiasmo minore, comparato agli altri campi disciplinari, rispetto all'importanza attribuita alle infrastrutture sulle ricadute qualitative in riferimento agli *output* prodotti (Hüsing *et al.* 2010, *cfr.* par. 2.3.1), il crescente utilizzo delle stesse negli ultimi anni ha indotto a supporre che la valutazione dei ricercatori sociali non si rivelasse più così tiepida. Si è ipotizzato dunque i sociologi intervistati potessero caratterizzare come effetti dello sfruttamento delle infrastrutture di ricerca alcuni aspetti relativi agli *output* generati a partire da esse, quali il migliore adeguamento agli obiettivi cognitivi e una qualche efficacia rispetto alle istanze di comparazione, copertura e accuratezza del dato. Dall'analisi delle interviste, l'ipotesi della concretizzazione di una serie di vantaggi che i ricercatori riconoscono riversarsi sugli *output* prodotti è stata confermata.

Nelle esperienze raccolte, l'utilizzo delle infrastrutture ha implicato consistenti agevolazioni nella capacità di rappresentazione dei fenomeni in virtù di una copertura campionaria difficilmente eguagliabile tramite studi individuali e consentito comparazioni a livello regionale o internazionale grazie all'omogeneità dei metodi di raccolta. Dal punto di vista prettamente cognitivo, sono state segnalate le opportunità che le infrastrutture hanno fornito rispetto (i) all'emersione di strutture latenti innestate nei fenomeni (più facilmente riscontrabili rispetto al passato grazie all'evoluzione tecnica dei *software* di elaborazione); (ii) alla possibilità di scegliere se concentrarsi sulle grandi tendenze o sui dati imprevisti; (iii) alla triangolazione dei risultati di ricerca con altri studi generati dalla medesima infrastruttura.

Come rilevato dai testimoni privilegiati, anche per i ricercatori selezionati per lo studio di caso, la credibilità connessa all'infrastruttura ricade in qualche modo sugli elaborati generati a partire da essa. La scelta di utilizzo dell'infrastruttura avviene in ragione dell'affidabilità, dell'attendibilità e della robustezza dei dati in essa presenti, proprietà che incontrano generalmente la piena fiducia da parte degli intervistati specie se i creatori o le fonti godono di un riconoscimento internazionale, e comunque in ragione del fatto che l'implementazione delle basi di dati, quando sviluppate da progetto, oltre ad essere corredate da un *set* consultabile di metadati, fanno riferimento alla supervisione di un *team* del quale essi possono far parte. Tali caratteristiche di qualità del dato contenuto nell'infrastruttura generano una valutazione positiva sui prodotti derivati anche in termini di solidità dei risultati e validità della ricerca. Per quanto riguarda le infrastrutture cui i partecipanti non concorrono alla creazione, i ricercatori rilevano in particolare che, grazie alla loro pre-organizzazione e ai controlli di qualità, il loro utilizzo preserva le elaborazioni da eventuali errori non campionari presenti nelle matrici. Infine, il risparmio di tempo nelle analisi, connesso all'utilizzo delle infrastrutture, viene valutato quale potenziale abilitatore di una maggiore qualità della ricerca per via della possibilità di rispondere ai *feedback* e alle critiche sui lavori scientifici con maggiore tempismo e flessibilità, potendo contare sulla sistematizzazione dei dati e sulla duttilità dei *software* di interrogazione.

In definitiva, l'accesso a insiemi di dati organizzati attraverso infrastrutture di ricerca ha prodotto un effetto riconoscibile per i ricercatori sociali: lo sviluppo di prodotti di ricerca migliori. I ricercatori ritengono infatti che la presenza delle infrastrutture abbia migliorato la loro capacità di intraprendere e produrre ricerche di alta qualità e ciò può essere correlato sia a una maggiore disponibilità di dati che alla presenza di dati robusti. Prescindendo dalle ricadute sulla formazione delle domande di ricerca riportate più sopra, gli studi citati nelle narrative sono stati valutati positivamente perché approfonditi, rigorosi, e caratterizzati dal riferimento a popolazioni più ampie. Un problema che tuttavia rimane aperto riguardo l'associazione tra utilizzo delle infrastrutture e qualità della ricerca inerisce al fatto che, se è vero i ricercatori riconoscono alle infrastrutture una serie di ricadute positive sugli *output*, alcuni di essi manifestano qualche ritrosia rispetto all'utilizzo di infrastrutture precostituite di interesse europeo o caratura internazionale. Lo studio ha inteso a questo proposito dunque raccogliere le motivazioni connesse a questa circostanza, riscontrandole su due livelli principali: il primo riguarda il condizionamento del lavoro di ricerca che può esprimersi in

una sorta di conformismo espresso dall'appiattimento a obiettivi cognitivi precostituiti in qualche modo "suggeriti" dall'infrastruttura (sorprendentemente evidenziato anche dai ricercatori più giovani, ma da essi accettato e superato); il secondo ordine riguarda il tema dell'ispezionabilità della ricerca, laddove non vi siano garanzie specifiche sulla qualità del dato. Da ciò deriva che il livello globale di accordo con le dichiarazioni sulla qualità della ricerca associata all'utilizzo di infrastrutture si scontra talvolta con la tendenza conservatrice dei ricercatori sociali a sviluppare sempre nuove risorse per soddisfare le loro particolari esigenze. Un'ulteriore tema di riflessione inerisce invece all'aspettativa che la maggiore qualità derivata dall'utilizzo delle infrastrutture abbia reso più probabile la circostanza che i prodotti citati abbiano prodotto un certo impatto sia dal punto di vista accademico che sotto l'aspetto sociale ed economico, a seconda delle tematiche sviluppate. Nel presente studio non è stata rilevata evidenza di tali benefici, tuttavia tali dimensioni possono essere esplorate attraverso future indagini caratterizzate dal tentativo di catturare impatti più ampi.

#### *Prospettive di evoluzione dell'utilizzo delle infrastrutture nelle scienze sociali*

Le infrastrutture di ricerca fisiche e digitali stanno acquisendo un ruolo centrale nella scienza contemporanea aperta, collaborativa e basata sul dato, consolidandosi sempre più spesso quali prerequisiti per l'avanzamento della conoscenza in tutti i campi scientifici. Se tale circostanza si presenta in tutta la sua evidenza nel caso delle scienze naturali (si pensi al CERN, il più grande laboratorio al mondo di fisica delle particelle, o alla *GenBank*, la banca dati *open access* sulle sequenze dei nucleotidi alimentata da migliaia di laboratori in tutto il mondo; *cf.* Nielsen, 2011, tr. it. 2012), quando si parla di *scienze sociali* la circostanza appare meno scontata (*cf.* Farago, 2014), sebbene ormai da una decina d'anni si riscontri un progressivo riconoscimento dell'importanza dello sviluppo e dell'implementazione di tali risorse anche per le c.d. scienze non dure, testimoniato dall'impegno della comunità scientifica europea verso la creazione e il mantenimento di infrastrutture di ricerca a livello europeo, e la creazione di nuove, affinché promuovano il progresso scientifico, la collaborazione internazionale e interdisciplinare (*cf.* European Commission, 2015; ESFRI, 2016). In virtù della potenziata capacità di catturare e lavorare dati e indicatori in modi inediti rispetto al passato attraverso infrastrutture di ricerca generate da attività collaborative – come le collezioni di dati derivate da progetti di ricerca specifici o da periodiche *survey* internazionali (come nel caso delle infrastrutture di interesse europeo afferenti al quadro

CESSDA, *cf.* par. 2.4.1) – anche il dominio della ricerca sociale si trova a navigare su un “oceano” di risorse informative che ha il potenziale di ispirare approcci innovativi ai processi di ricerca e di strutturare in modo peculiare le azioni e le relazioni dei ricercatori. Lo studio di caso qui presentato ha raffigurato il quadro di una comunità scientifica ancora alle prese con le *prime fasi* del processo di transizione verso un modo di produzione della conoscenza basato sulle infrastrutture di ricerca, soprattutto con riferimento a quelle condivise su vasta scala. Se da una parte è vero che la produzione scientifica nel campo degli studi sociologici fa sempre più affidamento a tali risorse e i disegni di ricerca tengono in maggiore considerazione la loro disponibilità, dall’altra parte i ricercatori hanno dimostrato di sfruttare solo parzialmente – per poca informazione o ritrosia – le potenzialità connesse all’accesso alle infrastrutture di interesse europeo o precostituite a livello internazionale (sono soprattutto le giovani leve a coglierne le potenzialità), preferendo un maggiore controllo sul processo di ricerca attraverso l’implementazione di infrastrutture *ad hoc*. Il lavoro sulle infrastrutture legate a progetto, dimostratosi la modalità prevalente di utilizzo, in genere esclude dalle procedure d’accesso ricercatori esterni ai progetti, limitando le potenzialità collaborative al raggio dei colleghi inseriti in *network* già esistenti. Ulteriore conseguenza di tale circostanza riguarda la limitata aggregazione di competenze attorno alle infrastrutture per la produzione di studi interdisciplinari, se non già presenti nei gruppi di ricerca e laddove i dati disponibili lo consentano. Sono gli stessi ricercatori ad auspicare che le infrastrutture di ricerca per le scienze sociali, soprattutto quelle di tipo *project-driven*, allentino le restrizioni all’accesso esterno e la resistenza dei gruppi di ricerca e si spingano sempre più all’interoperabilità (*cf.* 4.5). Ciò consentirebbe una maggiore condivisione dell’impresa scientifica che implicherebbe da una parte l’annullamento dei confini geografici tra ricercatori e dall’altra lo stimolo all’aggregazione delle competenze. I vantaggi rilevati dal punto di vista qualitativo e cognitivo rispetto agli *output* segnano comunque la presa di coscienza delle opportunità connesse all’utilizzo delle infrastrutture nella produzione di nuova conoscenza.

I risultati emersi da questa ricerca costituiscono un primo passo verso la comprensione delle dinamiche connesse allo sfruttamento delle infrastrutture di ricerca nell’ambito delle scienze sociali e si pongono come base per un ricontrollo da operare attraverso studi successivi. L’obiettivo esplorativo del presente lavoro porta con sé una serie di evidenze, rilanciando l’apertura di interrogativi e proponendo uno stimolo al dibattito e alla riflessione sul tema.

## Riferimenti bibliografici

Anderson, C. (2008), *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, in “Wired”, vol. 16, no. 7, June 23, 2008, disponibile all'indirizzo web: [www.wired.com/science/discoveries/magazine/16-07/pb\\_theory](http://www.wired.com/science/discoveries/magazine/16-07/pb_theory) (ultimo accesso: agosto 2017).

Anderson, P. (2007), *What is Web 2.0? Ideas, technologies and implications for education*, JISC, Bristol.

Antonijević, S., Dormans, S., Wyatt, S. (2013), “Working in Virtual Knowledge: Affective Labor in Scholarly Communication”, in Wouters, P., Beaulieu, A., Scharnhorst, A. & Wyatt, S., *Virtual Knowledge: Experimenting in the Humanities and the Social Sciences*, MIT Press, Cambridge, MA.

Arcila-Calderón, C., Calderín M., Aguaded, I. (2015), *Adoption of ICTs by communication researchers for scientific diffusion and data analysis*, in “El profesional de la información”, v. 24, n. 5, pp. 526-536.

Atkins, D., Borgman, C.L., Bindoff, N., Ellisman, M., Fellman, S., et. al. (2010), *Building a UK foundation for the transformative enhancement of research and innovation. Report of the international panel for the 2009 review of the UK research councils e-science programme*, Research Council, London.

Baccini, A. (2010), *Valutare la ricerca scientifica – Uso e abuso degli indicatori bibliometrici*, Il Mulino, Bologna.

Barkhuus, L., Brown, B. (2012), *The Sociality of Fieldwork: Designing for Social Science Research Practice and Collaboration*, in “Proceedings of GROUP 2012”, ACM Press, pp. 35-44.

Barnes, B. (1974), *Scientific Knowledge and Sociological Theory*, Routledge and Kegan Paul, London and Boston.

Bietz, M.J., Lee, C.P. (2009), *Collaboration in Metagenomics: Sequence Databases and the Organization of Scientific Work*, in Wagner, I, Telloğlu, H., Balka, E., Simone, C., Ciolfi, L. (eds.), “ECSCW 2009: Proceedings of the 11th European Conference on Computer Supported Cooperative Work”, 7-11 September 2009, Vienna, Austria, Springer-Verlag, London, pp. 243-262.

Binz-Scharf, M.C., Kalish, Y., Paik, L. (2014), *Making Science: New Generations of Collaborative Knowledge Production*, in “American Behavioral Scientist”, November 2014, pp. 1-11.

Bloor, D. (1976), *Knowledge and Social Imagery*, Routledge & Kegan Paul, London.

Bloor, D. (1981), *The strengths of the strong programme*, in “Philosophy of Social Science”, 11, pp. 199-213.

Borgman, C.L. (2007), *Scholarship in the Digital Age. Information, Infrastructure, and the Internet*, MIT Press, Cambridge.

Borgman, C.L. (2012), *The Conundrum of Sharing Research Data*, in “Journal Of The American Society For Information Science And Technology”, 63(6), pp. 1059–1078.

Boulton, G., Campbell, P., Collins, B., Elias, P., Hall, W., Laurie, G., O’Neill, O., Rawlins, M., Thornton, J., Vallance, P. & Walport, W. (2012), *Science as an open enterprise*, The Royal Society, London.

Bowker, G.C., Baker, K., Millerand, F., Ribes, D. (2010), “Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment”, in Hunsinger, J., Klastrup, L., Allen, M.M. (eds.), *International Handbook of Internet Research*, Springer, Dordrecht, pp 97-117.

Brasini, S., Freo, M., Tassinari, F., Tassinari, G. (2002), *Statistica aziendale e analisi di mercato*, Il Mulino, Bologna.

Braun, V., Clarke, V. (2006), *Using thematic analysis in psychology*, in “Qualitative research in psychology”, 3, pp. 77-101.

Briggs, C.L. (1986), *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research*, Cambridge University Press, Cambridge.

Bucchi, M. (2004), *Sociologia della scienza*, in “Nuova informazione bibliografica”, Anno I, N. 3/ luglio-settembre.

Callon, M. (1991), “Techno-economic Networks and Irreversibility”, in Law, J. (Ed.) *A Sociology of Monsters? Essays on Power, Technology and Domination. Sociological Review Monograph*. Routledge, London, 38: 132-161.

Camussone, P.F., Cuel, R., Ponte, D. (2011), “Internet and innovative knowledge evaluation processes: new directions for scientific creativity?”, in D’Atri, A., Ferrara, M., George, J.F., Spagnoletti, P. (Eds.), *Information Technology and Innovation Trends in Organizations*, Springer-Verlag, Berlin Heidelberg, pp. 435-442.

Ciula, A. Nyhan, J., Moulin, C. (2011), *Research Infrastructures in the Digital Humanities*, ESF Science Policy Briefing 42, European Science Foundation (ESF), disponibile all’indirizzo web:  
[http://archives.esf.org/fileadmin/Public\\_documents/Publications/spb42\\_RI\\_DigitalHumanities.pdf](http://archives.esf.org/fileadmin/Public_documents/Publications/spb42_RI_DigitalHumanities.pdf) (ultimo accesso: agosto 2017).

Clarke, A. E., Fujimura, J.H. (edited by) (1992), *The Right Tools for the Job: At Work in Twentieth Century Life Sciences*, Princeton University Press, Princeton, NJ.

Clarke, V., Braun, V. (2013), *Successful qualitative research: A practical guide for beginners*, Sage, London.

Collins, A. (1992), "Toward a design science of education", in Scanlon, E., O'Shea, T. (Eds.), *New directions in educational technology*. Springer-Verlag, Berlin Heidelberg, pp. 15-22.

Conference on Open Access to Knowledge in the Sciences and Humanities (2003), *Dichiarazione di Berlino*, disponibile all'indirizzo web: [http://www.biblio.polito.it/open\\_access/info\\_generali/dichiarazioni\\_a\\_favore\\_dell\\_open\\_access](http://www.biblio.polito.it/open_access/info_generali/dichiarazioni_a_favore_dell_open_access) (ultimo accesso: agosto 2017).

Corbetta, P. (1999), *Metodologia e tecniche della ricerca sociale*, Il Mulino, Bologna.

Corbetta, P. (2003), *La ricerca sociale: metodologia e tecniche*, Il Mulino, Bologna.

CRUI (2009), Linee guida per gli archivi istituzionali, disponibile all'indirizzo web: [https://www.crui.it/images/allegati/pubblicazioni/2009/linee\\_guida\\_archivi\\_istituzionali.pdf](https://www.crui.it/images/allegati/pubblicazioni/2009/linee_guida_archivi_istituzionali.pdf) (ultimo accesso: agosto 2017).

Dagiral, E., Peerbaye, A. (2016), *Making Knowledge in Boundary Infrastructures: Inside and Beyond a Database for Rare Diseases*, in "Science & Technology Studies", 29(2), pp. 44-61.

Daston, L. (2009), *Science Studies and the History of Science*, in "Critical Inquiry" 35(4), pp. 798-813.

David, P.A. (2004), *Understanding the emergence of 'open science' institutions: Functionalist economics in historical context*, in "Industrial and Corporate Change", 13(4), pp. 571-589.

Davison, M.J. (2003), *The Legal Protection of Databases*, Cambridge University Press, Cambridge, UK.

Dutton, W.H. (2010). "Reconfiguring access in research: information, expertise, and experience", in Dutton, W., Jeffreys, P. (eds.), *World wide research. Reshaping the sciences and humanities*, The MIT Press, Cambridge, Massachusetts, pp. 1-19.

Dutton, W.H., Meyer, E.T. (2008), *e-Social Science as an Experience Technology: Distance From, and Attitudes Toward, e-Research*, presentation for the 4th International Conference on e-Social Science, Manchester, UK, 18-19 June 2008.

Dutton, W.H., Meyer, E.T. (2009), *Experience with New Tools and Infrastructures of Research: An exploratory study of distance from, and attitudes toward, e-Research*, in "Prometheus", vol. 27, no. 3, pp. 223-238.

Edwards, P.N. (2010), *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*, MIT Press, Cambridge, MA.

Elias, P. (2010), "Providing Data on the European Level", in German Data Forum, *Building on progress. Expanding the research infrastructure for the social, economic, and behavioral sciences*, Vol. 1, Budrich UniPress, Opladen, pp. 139-154.

ESFRI (2016), *Strategy Report On Research Infrastructures – Roadmap 2016*, Science and Technology Facilities Council, disponibile all'indirizzo web: [http://www.esfri.eu/sites/default/files/20160308\\_ROADMAP\\_single\\_page\\_LIGHT.pdf](http://www.esfri.eu/sites/default/files/20160308_ROADMAP_single_page_LIGHT.pdf) (ultimo accesso: agosto 2017).

European Commission (2008), *Proposal for a Council Regulation on the Community legal framework for a European Research Infrastructure (ERIC)*, disponibile su <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52008PC0467&from=EN> (ultimo accesso: agosto 2017).

European Commission (2015), *Social science and humanities - Networks of Research infrastructures supported by the European Union*, ISBN: 978-92-79-38685-5.

European Science Foundation (2013), *Research Infrastructures in the European Research Area. A report by the ESF Member Organisation Forum on Research Infrastructures*, Ireg Strasbourg, ISBN: 978-2-918428-94-7, disponibile su [https://portal.meril.eu/meril/downloads/mof\\_research\\_infrastructures.pdf](https://portal.meril.eu/meril/downloads/mof_research_infrastructures.pdf) (ultimo accesso: agosto 2017).

Falk, I., Guenther, J. (2006), *Generalising from qualitative research: case studies from VET in context*, 15th NCVET conference, 2006.

Farago, P. (2014), "Understanding How Research Infrastructures Shape the Social Sciences: Impact, challenges, and outlook", in Duşa, A., Nelle, D., Stock, G. and Wagner, G.G., *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, SCIVERO Verlag, Berlin, pp. 21-33.

Fitzgerald, R. (2016), "The European Social Survey Infrastructure: Past, Present And Future", keynote session presentation at 3rd International ESS Conference, 13-15th July 2016, Amphimax, University of Lausanne, Switzerland, disponibile su: <http://www.europeansocialsurvey.org/docs/about/conference/ESS-Conference-July-2016-Fitzgerald.pdf> (ultimo accesso: agosto 2017).

Fleck, L. (1935), *Entstehung und Entwicklung einer wissciologia senschaftliche Tatsache*, tr. it. *Genesi e sviluppo di un fatto scientifico: per una teoria dello stile e del collettivo di pensiero*, Il Mulino, Bologna, 1983.

Franzoni, C., Sauermann, H. (2014). *Crowd science: The organization of scientific research in open collaborative projects*, in "Research Policy", 43(1), pp. 1–20.

Friedmann, G. (1978), *The Anatomy of Work: Labor, Leisure, and the Implications of Automation*. Greenwood Press, Westport, Connecticut.

Gagliardi, D., Cox, D., Li, Y. (2015), "Institutional Inertia and Barriers to the Adoption of Open Science", in Reale, E., Primeri, E. (Eds.), *The Transformation of University Institutional and Organizational Boundaries*, Sense Publishers, Rotterdam/Boston/Taipei, pp. 107-133.

Gallino, L. (1995), "Sociologia della Scienza", in *Enciclopedia Italiana Treccani - V Appendice*, disponibile all'indirizzo web: [http://www.treccani.it/enciclopedia/sociologia-della-scienza\\_%28Enciclopedia-Italiana%29/](http://www.treccani.it/enciclopedia/sociologia-della-scienza_%28Enciclopedia-Italiana%29/) (ultimo accesso: agosto 2017).

Gallino, L. (2003), *Tecnologie della cultura, società in rete. Una sfida per la formazione universitaria*, in *Quaderni di sociologia*, 47(31), pp. 66-82

German Council of Science and Humanities (2011), *Recommendations on Research Infrastructures in Humanities and Social Sciences*, German Council of Science and Humanities, disponibile all'indirizzo web: [https://www.wissenschaftsrat.de/download/archiv/10465-11\\_engl.pdf](https://www.wissenschaftsrat.de/download/archiv/10465-11_engl.pdf) (ultimo accesso: agosto 2017).

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., Trow, M. (1994), *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*, SAGE, London.

Gigerenzer, G. (2001), "Digital Computer: Impact on the Social Sciences", in Smelser, N. J., Baltes, P. B. (eds.), *International Encyclopedia of the Social and Behavioral Sciences*, Vol. 6, Elsevier, Amsterdam, pp. 3684–3688.

Giles, J. (2005), *Internet Encyclopaedias Go Head to Head*, in "Nature", 438 (2005), pp. 900-901.

Glaser, J., Laudel, G. (2009a), *Identifying individual research trails* in Larsen, B. & Leta, J. (eds.), *Proceedings of the 12th International Society for Scientometrics and Infometrics (ISSI '09)*, Rio de Janeiro, Brazil, pp. 841-845.

Glaser, J., Laudel, G. (2009b), "On interviewing "good" and "bad" experts", in Bogner, A., Littig, B., Menz, W. (Eds.), *Interviewing experts*, Palgram MacMillan, London, pp. 117-138.

Glesne, C., Peshkin, A. (1992), *Becoming qualitative researchers: An introduction*, Longman, White Plains, NY.

Gorden, R.L. (1975), *Interviewing: Strategy, techniques and tactics*, Dorsey, Homewood, IL.

Gorgolewski, K. J., Margulies, D. S., Milham, M. P. (2013), *Making data sharing count: A publication based solution*, in "Frontiers in Neuroscience", 7(9), pp. 1-7.

- Heimeriks, G., Leydesdorff, L. (2012), *Emerging search regimes: measuring co-evolutions among research, science, and society*, "Technology Analysis & Strategic Management", 24:1, pp. 51-67.
- Hey, T., Tansley, S., Tolle, K. (2009), "Jim Gray on e-science: A transformed scientific method", in Hey, T., Tansley, S., Tolle, K. (Eds.). *The fourth paradigm. Data-intensive scientific discovery*, Microsoft Research, Redmond, Washington, pp. xvii-xxxi.
- Hine, C. (2006), *Databases as scientific instruments and their role in the ordering of scientific work*, in "Social Studies of Science", vol. 36, no. 2, pp. 269-298.
- Hüsing, T., Robinson, S., Barjak, F., Bendel, O., Wiegand, G., Eccles, K., Meyer, E., Schroeder, R., Kertcher, Z., Coslor, E. (2010), *eResearch2020: The Role of e-Infrastructures in the Creation of Global Virtual Research Communities. Final Report. Report to the European Commission*, European Communities, Bonn and Brussels.
- Ishak, N.M., Bakar, A.Y.A. (2014), *Developing sampling frame for case study: challenges and conditions*, in "World Journal of Education", Vol. 4, No. 3, pp. 29-35.
- JISC, British Library (2012), *Researchers of tomorrow. The research behavior of generation Y doctoral students*, The British Library and HEFCE, London.  
Report disponibile su:  
<http://www.jisc.ac.uk/media/documents/publications/reports/2012/Researchers-of-Tomorrow.pdf> (ultimo accesso: agosto 2017).
- Kennan, M.A. (2008), *Reassembling scholarly publishing: open access, institutional repositories and the process of change* (unpublished doctoral thesis), The University of New South Wales, Sydney, Australia, disponibile all'indirizzo web:  
<https://ils.unc.edu/~wildem/ASIST2009/Kennan.pdf> (ultimo accesso: agosto 2017).
- Knorr-Cetina, K. (1981), *The manufacture of knowledge: An Essay on Constructivism and Contextual Nature of Science*, Pergamon, Oxford.
- Knorr-Cetina, K. (1982), *Scientific Communities or Transepistemic Arenas of Research" A Critique of Quasi-Economic Models of Science*, in "Social Studies of Science", 12, pp. 101-130.
- Kuhn, T.S. (1962), *The structure of scientific revolution*, University of Chicago, Chicago, tr.it. *La struttura delle rivoluzioni scientifiche*, Einaudi, Torino, 1969.
- Latour, B., Woolgar, S. (1979), *Laboratory Life. The Social Construction of Scientific Facts*, Princeton University Press, Princeton.
- Latour, B. (1987), *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press, Cambridge, Massachussets.
- Latour, B. (1999), *On recalling ANT*, in "The Sociological Review", Volume 47, Issue S1 May 1999, pp. 15-25.

- Laudel, G. (2001), *Collaboration, creativity and rewards: why and how scientists collaborate*, in "International Journal of Technology Management", 22, pp. 762-781.
- Laudel, G., Glaser, J. (2007), *Interviewing scientists*, in "Science, Technology & Innovation studies", Vol. 3, No. 2, December 2007.
- Lauer, G. (2014), "Challenges for the Humanities: Digital Infrastructures", in Duşa, A., Nelle, D., Stock, G., Wagner, G.G., *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, SCIVERO Verlag, Berlin, pp. 35-38.
- Lee, C. P. (2007), *Boundary negotiating artifacts: Unbinding the routine of boundary objects and embracing chaos in collaborative work*, in "Computer Supported Cooperative Work: The Journal of Collaborative Computing", vol. 16, no. 3, pp. 307-339.
- Leibniz Association (2014), "Leibniz Research Alliance 'Science 2.0'", <http://www.leibniz-gemeinschaft.de/en/research/leibniz-research-alliances/science-20/> (ultimo accesso: agosto 2017).
- Lessler, J.T., Kasbeek, W.D. (1992), *Nonsampling Error in Surveys*, John Wiley & Sons Inc., New York.
- Leonelli, S. (2012), *When Humans Are the Exception: Cross-Species Databases at the Interface of Biological and Clinical Research*, in "Social Studies of Science", Vol. 42 issue 2, pp. 214-236.
- Lepori, B., Barré, R., Fillietreau, G. (2008), *New perspectives and challenges for the design and production of S&T indicators*, in "Research Evaluation", 17(1), pp. 33-44.
- Lossau, N. (2012), *An Overview of Research Infrastructures in Europe – and Recommendations to LIBER*, in "Liber Quarterly: The Journal of European Research Libraries", Volume 21 Issue 3/4 2012, pp. 313-329.
- Lynch, M. (1993), *Scientific practise and ordinary action: Ethnometodology and social studies of science*, Cambridge University Press, Cambridge.
- MacLuhan, M. (1964), *Understanding Media: The Extensions of Man*, McGraw Hill, New York, London.
- Mahapatra, A. (2010), *Researchers in a Brave New Web 2.0 World*, in "ACS Chem. Biol.", 2010, 5 (9), pp. 799–800.
- Martin, B. (2003), "The changing social contract for science and the evolution of the university", in Geuna, A., Salter, A.J., Steinmuller, W.E. (eds.), *Science and innovation. Rethinking the rationales for funding and governance*, Edward Elgar, Cheltenham, pp. 7-29.

Martinez-Uribe, L., Patrick, M. (2011), *Researchers' Database Behaviour and Use: A Literature Review*, Virtual Infrastructure with Database as a Service (VIDaaS) Project, JISC, disponibile all'indirizzo *web*:

<http://vidaas.oucs.ox.ac.uk/docs/VIDaaS%20Literature%20Review%20v1.0.pdf>  
(ultimo accesso: agosto 2017).

MERIL project (2011), *MERIL, Research Infrastructures of European relevance. A comprehensive inventory*, European Science Foundation, ISBN: 978-2-918428-45-9, disponibile su [http://archives.esf.org/fileadmin/Public\\_documents/Publications/meril.pdf](http://archives.esf.org/fileadmin/Public_documents/Publications/meril.pdf)  
(ultimo accesso: agosto 2017).

Merton, R.K. (1938), *Science, technology and society in XVII century England*, in "Osiris", Vol. 4, pp. 360-632, tr. it. *Scienza, tecnologia e società nell'Inghilterra del XVII secolo*, Franco Angeli, Milano, 1975.

Merton, R.K. (1968), *The Matthew Effect in Science*, in "Science", Vol. 159, No. 3810, pp. 56-63.

Merton, R.K. (1981), *La Sociologia della Scienza. Indagini Teoriche ed Empiriche*, Franco Angeli, Milano.

Monteiro, E., Hanseth, O. (1996), "Social shaping of information infrastructure: on being specific about the technology" in Orlikowski, W.J., Walsham, G., Jones, M.R., and DeGross, J.I. (edited by), *Information technology and changes in organizational work*, Chapman & Hall, London, pp. 325-343.

Montesperelli, P. (2008), "Esperienze e spunti di riflessione da una prospettiva "ermeneutica" di analisi delle interviste", in Cipriani, R. (a cura di) *L'analisi qualitativa*, Armando, Roma.

Mulkay, M.J., Edge, D.O. (1976), *Astronomy Transformed: The Emergence of Radio Astronomy in Britain*, John Wiley and Sons, Inc., New York.

National Academy of Sciences (1989), *Information Technology and the conduct of Research*, National Academic Press, Washington D.C.

Nielsen, M. (2011), *Reinventing discovery: The new era of networked science*, Princeton University Press, New Jersey, tr. it. *Le nuove vie della scoperta scientifica. Come l'intelligenza collettiva sta cambiando la scienza*, Einaudi, Torino, 2012.

NORFACE (2009), *Data and Research Infrastructure: Recommendations to the NORFACE partners and proceedings of the NORFACE Conference on Data & Research Infrastructure*, Oy: Helsinki.

Nowotny, H., Scott, P., Gibbons, M. (2001), *Re-Thinking Science. Knowledge and the Public in an Age of Uncertainty*, Polity Press, Cambridge.

O'Reilly, T. (2005), *What is Web 2.0? Design patterns and business models for the next generation of software*, articolo da blog, disponibile all'indirizzo *web*: <http://www.oreilly.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html> (ultimo accesso: agosto 2017).

OCSE (2007), *OECD Principles and Guidelines for Access to Research Data from Public Funding*, Organisation for Economic Co-Operation and Development, Paris, disponibile all'indirizzo *web*: <http://www.oecd.org/dataoecd/9/61/38500813.pdf> (ultimo accesso: agosto 2017).

OCSE (2013), *New Data for Understanding the Human Condition. International Perspectives*, OECD Global Science Forum Report on Data and Research Infrastructure for the Social Sciences, OECD.

Open Knowledge Foundation (2014), "Open Knowledge: What is Open?", disponibile all'indirizzo *web*: <https://okfn.org/opendata/> (ultimo accesso: agosto 2017).

Paccagnella, L. (2012), *Web 2.0: dalla divulgazione alla co-produzione della conoscenza scientifica? Opportunità e limiti*, in Proceedings "Comunicare Fisica 2012", Frascati Physics Series Italian Collection, pp. 220-223.

Pasquetto, I.V., Sands, A.E., Borgman, C.L. (2015), *Exploring Openness in Data and Science: What is "Open," to Whom, When, and Why?*, in Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community, St. Louis, Missouri, Article No. 141.

Patton, M.Q. (2002), *Qualitative research and evaluation methods (3rd ed.)*, Sage Publications, Thousand Oaks, CA.

Pearce, N. (2010), *A study of technology adoption by researchers. Web and e-science infrastructures to enhance research*, in "Information, communication & society", v. 13, n. 8, pp. 1191-1206.

Ponte, D., Simon, J. (2011), *Scholarly Communication 2.0: Exploring Researchers' Opinions on Web 2.0 for Scientific Knowledge Creation*, "Evaluation and Dissemination, Serials Review", 37:3, pp. 149-156.

Procter, R., Williams, R., Stewart, J., Poschen, M., Snee, H., Voss, A., Asgari-targhi, M. (2010), *Adoption and use of Web 2.0 in scholarly communications*, in "Philosophical Transactions A: Mathematical, Physical and Engineering Sciences", vol. 368, issue 1926, pp. 4039-4056.

Programma Triennale della Ricerca 2010-2012: glossario, disponibile all'indirizzo *web*: <http://www.miur.it/UserFiles/3248.pdf> (ultimo accesso: agosto 2017).

Reale, E., De Filippo, D.A., Gómez, I., Lepori, B., Potì, B., Primeri, E., Probst, C., Sanz Casado, E. (2011), *New uses of the institutional databases of universities: indicators of research activity*, in "Research Evaluation", 20(1), March 2011, pp. 47-60.

Renschler, B., Kleiner, I., Wernli, B. (2013), "Concepts and key features for understanding social science research infrastructures", in Renschler, B., Kleiner, I., Wernli, B., Farago, P., Joye, D. (Eds.), *Understanding Research Infrastructures in the Social Sciences*, Seismo Press, Zurich, pp. 11-18.

Rinaldi, A. (2014), *Spinning the web of open science*, in "EMBO Reports", 15(4), pp. 342-346.

RISIS project (2013), *DOW RISIS (313082) 2013-11-20-1*, disponibile su <http://risis.eu/official-documents/> (ultimo accesso: agosto 2017).

Rizzuto, C. (2010), *Inspiring excellence. Research infrastructures and the Europe 2020 Strategy*, Richard L. Hudson Science Business, disponibile all'indirizzo [web: https://ec.europa.eu/research/infrastructures/pdf/esfri/publications/esfri\\_inspiring\\_excellence.pdf](http://ec.europa.eu/research/infrastructures/pdf/esfri/publications/esfri_inspiring_excellence.pdf) (ultimo accesso: agosto 2017).

Rudolph, J.L. (2014), *Science Studies, Enciclopedia of Scienze Education*, Springer Netherlands.

Scheliga, K. (2015), *Communication Forms and Digital Technologies in the Process of Collaborative Writing* in Proceedings of the Second International Conference, INSCI 2015 Brussels, Belgium, May 27-29, 2015 Proceedings, LNCS 9089, pp. 113-122.

Seidman, I.E. (1991), *Interviewing As Qualitative Research: A Guide for Researchers in Education and the Social Sciences*, Teachers College Press, Columbia University.

Skogerbø, E. (2011), "Case studies and comparative analysis", presentazione nell'ambito di "MEVIT 4800 Perspectives and methods in media research", disponibile sul sito: [http://www.uio.no/studier/emner/hf/imk/MEVIT4800/h11/undervisningsmateriale/MEVIT4800%202011\\_2\\_250811.ppt.pdf](http://www.uio.no/studier/emner/hf/imk/MEVIT4800/h11/undervisningsmateriale/MEVIT4800%202011_2_250811.ppt.pdf) (ultimo accesso: agosto 2017).

Stake, R. (1995), *The Art of case study*, Sage, London.

Star, S.L. (1991), "Power, technologies and the phenomenology of conventions: on being allergic to onions" in Law, J. (ed.) *A Sociology of Monsters? Essays on Power, Technology and Domination*, Routledge, London, pp. 26-56.

Star, S.L. (edited by) (1995), *The Cultures of Computing*, Basil Blackwell, Oxford.

Star, S.L., Bowker G.C. (2007), "L'infrastruttura dei new media", in Lievrouw L.A., Livingston S. (a cura di), *Capire i new media. Culture, comunicazione, innovazione tecnologica e istituzioni sociali*, Hoepli, Milano.

Star, S.L., Ruhleder K. (1996), *Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces*, in "Information Systems Research", Vol. 7, No. 1, March 1996, pp. 111-134.

Swarnborn, P. (2010), *Case Study Research: What, Why and How?*, Sage, London.

Tashakkori, A., Teddlie, C. (2003), *Handbook of Mixed Methods in Social & Behavioral Research*, Sage, Thousand Oaks.

Trinczek, R. (2009), "How to Interview Managers? Methodical and Methodological Aspects of Expert Interviews as a Qualitative Method in Empirical Social Research", in Bogner, A., Littig, B., Menz, W. (eds.), *Interviewing Experts*, Palgrave Macmillan, Houndsmills, pp. 203-216.

Venkatesh, V., Morris, M., Davis, G., Davis, F. (2003), *User acceptance of information technology: Toward a unified view*, in "MIS quarterly", v. 27, n. 3, pp. 425-478.

Walker, D.W., Atkinson, M. P., Brooke, J. M., Watson, P. (eds.) (2011), *Special theme: E-science novel research, new science and enduring impact*, Philosophical Transactions of the Royal Society A, 369, issue 1949.

Ware, M., Monkman, M. (2008), *Peer review in scholarly journals: perspective of the scholarly community - an international study*. Publishing Research Consortium.

Weber, M. (1922), *Wirtschaft und Gesellschaft*, Mohr, Tubingen, trad. it. *Economia e Società*, Edizioni di Comunità, Milano, 1962.

Wilkins, J. S. (2008), *The roles, reasons and restrictions of science blogs*, in "Trends in Ecology & Evolution", 23(8), pp. 411-413.

Williams, R., Stewart, J., Slack, R. (2005), *Social learning in technological innovation: experimenting with information and communication technologies*, Edward Elgar, Aldershot, UK.

Yin, R.K. (2009), *Case Study Research: Design and Methods (fourth edition)*, Sage, Thousand Oaks, CA.

Ziman, J. (1998), *Essay on science and society: Why must scientists become more ethically sensitive than they used to be?*, in "Science", Vol. 282, Issue 5395, pp. 1813-1814.

Ziman, J. (2000), *Real Science: What It Is and What It Means*. Cambridge University Press, trad. it. *La vera scienza. Natura e modelli operativi della prassi scientifica*, Dedalo, Bari, 2002.

## **Allegato: Protocollo di intervista**

*L'intervista è volta a rilevare l'esperienza e i giudizi del ricercatore riguardo gli effetti dell'utilizzo delle infrastrutture di ricerca sul suo lavoro. Le domande saranno tese a chiarire il posizionamento dell'intervistato riguardo lo sfruttamento di queste risorse e verteranno su una narrativa che avrà ad oggetto uno o più output selezionati dalla produzione scientifica dell'intervistato. La chiusura sarà affidata alla richiesta di un giudizio complessivo sull'utilizzo di infrastrutture di ricerca nella produzione di conoscenza nell'ambito delle scienze sociali.*

*(introduzione)* Il lavoro dei ricercatori e la loro produzione scientifica si basano sempre di più sull'utilizzo di nuovi strumenti – legati alle innovazioni nel settore ICT – che consentono il passaggio da modalità tradizionali di accesso alle (e sistematizzazione delle) informazioni (biblioteca, *workshop*, informazioni non sistematizzate presenti sul *web*) a nuove modalità basate sull'utilizzo di infrastrutture di ricerca.

1. Se le fosse chiesto di dare una definizione di infrastruttura di ricerca, in che termini ne parlerebbe? Cosa sono per lei?
2. Lei usa o ha usato infrastrutture di ricerca nella sua attività di ricerca?
3. In termini generali, crede la disponibilità (o la creazione) di queste risorse possa aver modificato il suo modo di fare ricerca?
4. In che modo lo sfruttamento di queste risorse ha avuto un riflesso rispetto al complesso della sua produzione scientifica? Quali *output* hanno generato? Può farmi qualche esempio? *(4-alternativa)*. Guardando le pubblicazioni del suo curriculum mi sono soffermato sul prodotto *x*. Possiamo considerare questo prodotto come esempio di output di ricerca creato grazie all'utilizzo di una infrastruttura di ricerca?
5. Poter contare sulla disponibilità di un'infrastruttura di ricerca come incide in termini di disegno della ricerca e di implementazione? Può parlarmene in concreto facendo riferimento agli *output* che abbiamo selezionato/di cui mi ha parlato?
6. Quali sono per lei gli aspetti più rilevanti per cui la presenza e l'utilizzo dell'infrastruttura l'hanno aiutata ad ottenere una migliore qualità scientifica del prodotto realizzato?

7. L'utilizzo di questa infrastruttura è concluso o pensa di sfruttare nuovamente la/le infrastruttura/e che hanno generato il prodotto? In che modo?
8. Lei svolge o ha svolto ricerca interdisciplinare? (se sì) Rispetto alla sua esperienza, lo sfruttamento di infrastrutture di ricerca può agevolare lo sviluppo e l'approfondimento di temi di ricerca trasversali?
9. Lo sfruttamento di infrastrutture ha influito sulle sue possibilità di realizzare ricerca in collaborazione con altri ricercatori, italiani e stranieri? Se sì, come?
10. Quali sono i vantaggi (o gli svantaggi) che comporta l'utilizzo di infrastrutture di ricerca nel lavoro di uno scienziato sociale?
11. Rispetto alla sua esperienza, quali sono i requisiti che un'infrastruttura di ricerca dovrebbe avere?