

Un *General Data Protection Regulation* (GDPR) non molto *general*

A *General Data Protection Regulation* (GDPR) *not really general*

Elena Bougleux



Edizione digitale

URL: <https://journals.openedition.org/aam/4098>

DOI: 10.4000/aam.4098

ISSN: 2038-3215

Editore

Dipartimento Culture e Società - Università di Palermo

Notizia bibliografica digitale

Elena Bougleux, «Un *General Data Protection Regulation* (GDPR) non molto *general*», *Archivio antropologico mediterraneo* [Online], Anno XXIV, n. 23 (1) | 2021, online dal 30 juin 2021, consultato il 02 juillet 2021. URL: <http://journals.openedition.org/aam/4098> ; DOI: <https://doi.org/10.4000/aam.4098>

Questo documento è stato generato automaticamente il 2 juillet 2021.



Archivio antropologico mediterraneo è distribuita con Licenza Creative Commons Attribuzione - Non commerciale - Non opere derivate 4.0 Internazionale.

Un *General Data Protection Regulation* (GDPR) non molto *general*

A General Data Protection Regulation (GDPR) *not really general*

Elena Bougleux

Premessa

- 1 Nel maggio del 2018 è entrato in vigore il nuovo regolamento dell'Unione Europea in materia di trattamento dei dati personali e di privacy, noto come GDPR ovvero General Data Regulation Protection. Nei mesi successivi, e precisamente nel corso della sua assemblea generale di luglio, la Società Italiana di Antropologia Culturale (SIAC) ha promosso l'istituzione di una commissione di lavoro sui temi dell'etica nella ricerca e della protezione dei dati relativi alla ricerca antropologica, in particolare dedicata a raccogliere informazioni accurate sulle procedure e le prassi che gli antropologi in Italia hanno storicamente adottato per la gestione dei dati etnografici, nella prospettiva di organizzare la transizione verso il "nuovo mondo" regolato dal GDPR. I lavori di questa commissione, che sono ancora in corso e che hanno visto coinvolti rappresentanti di tutte le associazioni antropologiche italiane, saranno riportati in dettaglio in un articolo in preparazione per il prossimo numero di AAM. D'altra parte, durante i lavori è parso subito evidente quanto fosse necessario non solo approfondire la conoscenza dei termini giuridici del GDPR in relazione alle sue declinazioni nazionali e alle sue parti più specificamente attuative, quanto capire anche quale fosse il suo spirito etico, a quale tipo di rischi e a quale tipo di dati da proteggere si riferisse il legislatore europeo, e in che misura il nuovo GDPR avrebbe aiutato, o piuttosto avrebbe rischiato di penalizzare, la ricerca antropologica.
- 2 Nel febbraio 2021 la SIAC ha organizzato un webinar pubblico, in cui numerosi relatori hanno discusso gli elementi emersi dalla ricognizione sulle pratiche, hanno esposto un'analisi accurata del GDPR e dato una valutazione dei diversi dibattiti suscitati in Europa dalla sua applicazione. In questo contributo e nei saggi che seguono, l'intervista a Giulia Barrera e il testo di Roberta Raffaetà, si intende dare conto della discussione

emersa dal webinar dello scorso febbraio¹, senza ripercorrerne esattamente le tracce, ma raccogliendo gli elementi su cui è assolutamente necessario discutere e riflettere ancora, alla luce delle nuove stringenti normative che ormai sono in vigore, e che condizionano il nostro modo di fare ricerca sin dalle sue premesse: come fare domanda per un finanziamento, come relazionarsi con le fonti, come intraprendere il lavoro sul campo.

Cosa sono i dati

- 3 I dati con cui lavora la conoscenza non sono mai effettivamente dati, consegnati, *given*, viceversa sono sempre costruiti, pensati e ripensati, ‘coevoluti’. Potrebbe sembrare irriuale dare avvio a questa riflessione sul GDPR in modo così assertivo, ma si tratta di un punto di partenza fondamentale su cui è necessario fare molta chiarezza. Ad essere *costruiti* infatti non sono solo i dati delle scienze sociali, fragili esiti processuali spesso non riproducibili, di cui dirò estesamente più avanti e che sono il vero tema di questa riflessione. Costruiti sono anche i *data* delle scienze naturali, incredibilmente chiamate tuttora in qualche contesto “scienze esatte”, a partire dalla cui interpretazione epistemologica fuorviante si muove tutto il discorso sulla necessità di regolamentazione e protezione dei dati della ricerca, che si estende poi anche alla nostra disciplina.
- 4 Entro dunque nel merito delle diverse tipologie di dati, e mi arrischio in questa disamina comparativa perché nel corso della mia vita da ricercatrice ho avuto occasione di manipolare dati di assai diversa natura, provenienza ed entità. I *data* con cui lavorano le scienze naturali vengono considerati, specialmente al di fuori dei contesti che li producono, come oggetti cognitivi stabili e definitivi, come entità statiche su cui fondare affidabilmente il resto della conoscenza, il processo di costruzione della conoscenza. Nelle scienze naturali da oltre sessant’anni, e in tutte le altre discipline in tempi più recenti, si parla esplicitamente di *robust data set* per indicare quelle basi di dati che non solo sono ampie in senso quantitativo, e quindi forse più capaci di cogliere elementi di una data realtà, ma sono anche solide (*robust*) nel senso di attendibili e confermate. Un *robust data set* è costituito da dati raccolti coerentemente, tutti nello stesso modo, attraverso un metodo invariante, ovvero che non risente di perturbazioni da parte di fattori esterni, e non muta nel tempo. Un *robust data set* contiene inoltre dati puliti, ovvero dati già privati di elementi valutati superflui ai fini della sua specifica missione, e ordinati secondo la logica che poi servirà alla loro lettura. Tali dettagli mettono bene in luce quanto non tutti i database siano *robust*, anzi evidenziano che la maggior parte dei database non lo è affatto. I normali *data set* sono infatti insieme più o meno disordinati di informazioni ancora tutte da organizzare, raccolte in ordine sparso e senza una vera sequenzialità, prima di diventare *robust* e dunque forse utili ai fini della ricerca quantitativa. Il ricorso a un aggettivo qualitativo quale *robust* per fissare esattamente il carattere di un oggetto prettamente quantitativo come un dataset dovrebbe destare qualche sospetto: si rende necessario fare riferimento al processo di raffinamento e organizzazione dei dati al fine di poter consolidare il concetto unico e statico di dato nel senso di *given*, consegnato, finale e concluso, poiché se preso solo il dato non lo è, né consegnato, né finale, né concluso.
- 5 Fatta questa premessa di lessico, è necessario anche osservare la storia recente delle scienze sociali, dove negli ultimi decenni spicca la novità interessante portata dagli

Science and Technology Studies, STS², inaugurati – forse a sua insaputa – da Thomas Kuhn nel 1962 e poi sviluppati in multiple direzioni, aprendo una lunga stagione di analisi e critica epistemologica ai processi di costruzione dei dati nelle scienze naturali. Grazie alle discussioni accese dagli STS, chi oggi assume ancora una posizione determinista nei confronti dell'esito della propria ricerca (evenienza rara) collocata all'interno delle scienze naturali fa una professione di ignoranza del dibattito sollevato dalle STS, oppure deve esplicitare e motivare – direi quasi giustificare – la propria posizione epistemologica in modo molto netto e convincente (evenienza ancora più rara).

- 6 Il contributo principale delle STS è stato quello di obbligare tipi diversi di ricercatori a confrontare sul campo i rispettivi metodi di lavoro, riconoscendo nel metodo dell'altro una capacità inaspettata di arrivare a dare risposte per domande che, nel proprio campo e con il proprio metodo, non si sarebbero mai neanche poste. Gli “uni” e gli “altri” sono stati intercambiabilmente i ricercatori delle scienze sociali e di quelle naturali, vicendevolmente sorpresi dalle potenzialità del confronto, con la sostanziale differenza che quelli delle scienze naturali hanno visto “indebolirsi” la propria pregressa nozione di *dato*, divenuta via via più contingente e processuale, mentre per quelli delle scienze sociali si è trattato quasi sempre di confermare la validità del proprio metodo, certamente “debole”, instabile e precario, ma incredibilmente fecondo, quanto affascinante e utile per tutti, specialmente nelle situazioni di crisi³. Nello stesso tempo, in certi ambienti delle scienze sociali si è anche iniziato ad avvertire un certo qual fascino per il dato quantitativo, per la sua capacità di descrivere visivamente e permettere confronti, che pure era stato considerato per tutto il postmodernismo come il male assoluto⁴.
- 7 Non sempre, va detto, gli studi nelle STS e le loro contaminazioni hanno prodotto interessanti e feconde sintesi, ma se non altro sono state aperte delle porte, e abbattuti alcuni steccati di reciproca ignoranza e mis-conoscenza⁵. Un esempio recente e molto adatto a mostrare la crescita di valore del dialogo interdisciplinare acceso dalle STS è fornito dalla ricerca sui temi ambientali: a partire dalla consacrazione ricevuta dal quinto Rapporto Generale dell'IPCC (*Assessment Report 5*, WG II, 2014), il coinvolgimento delle scienze sociali nella ricerca sulle crisi dell'ambiente e sugli impatti del cambiamento climatico è stato sempre crescente, mostrando sempre più spesso quanto siano necessari gli sconfinamenti dei ricercatori di specifiche discipline verso le discipline degli altri. Vorrei osservare che l'AR5 ha stabilito la necessità – e anche l'urgenza – di considerare l'apporto delle scienze sociali per descrivere, gestire e mitigare la crisi climatica già nel 2014, a partire da una letteratura scientifica internazionale che a sua volta era già ampiamente disponibile⁶. Invece il Piano Nazionale della Ricerca italiano per il settennio 2021-27 delega ancora alle scienze naturali tutta la gestione del “pacchetto ambiente”, riproducendo una visione vecchia, parziale e pure economicamente dannosa, in quanto si priva di contributi ricchi, documentati e creativi dei ricercatori che anche in Italia hanno lavorato sulle “scienze sociali del clima”⁷.
- 8 Dunque, riportando il focus sulla natura dei dati: proprio mentre con grande fatica le scienze naturali (già esatte) escono dalla torre di cristallo, e nelle more di crisi ambientali sempre diverse e più frequenti (inclusa la pandemia da Covid, che si configura come una crisi ambientale sindemica⁸), il mondo si riempie di tecnologie: la comunicazione su scala globale è appaltata a pochissimi attori privati monopolisti, l'università e tutto il settore educativo si consegnano a mani legate a pochi provider

della rete, la gestione dei *big data* (ovvero le nostre stesse biografie, con tutti i dettagli inclusi) è affidata a oscurissimi *data analytics*, che ci comprano e ci vendono a caro prezzo, come se fossimo noi stessi dei *dati* (la vicenda della compravendita dei profili FB da parte di *Cambridge Analytica* per influenzare la campagna elettorale di Trump è un caso eloquente). Osservo solo per completezza che il valore miliardario di WhatsApp che Facebook è stata disposta a pagare per comprarla, e inglobarla nel suo database, ha pochissimo a che fare con il valore reale e materiale delle infrastrutture e del personale che lavora a WhatsApp, e molto a che fare con il numero dei suoi utenti – il valore di WhatsApp è infatti decuplicato in dieci anni, grazie alla “sola” crescita degli utenti, e non certo quella dei suoi dipendenti professionisti delle IT, né dei loro stipendi⁹.

- 9 La repentina tecnologizzazione e quindi *datizzazione* dell'esistenza privata e pubblica di tutti, avvenuta attraverso la creazione e lo scambio di informazioni quantitative, ha fatto scivolare il discorso sulla natura e sulla fragilità del *dato/given* verso una direzione ben diversa da quella che veniva, e viene tuttora, intesa nella ricerca e che invece è determinata dalle politiche di mercato. Sono queste infatti a stabilire come avvengono la raccolta e la gestione globale di informazioni, quelle sì fisse e statiche, quali sono i dati biografici, medici, bancari, telefonici, valutativi, ecc. La legge europea sulla protezione dei dati nota come GDPR nasce dunque in questo contesto caratterizzato da un duplice problema: da una parte, la necessità di gestire, per la prima volta nella storia, valanghe impensabili di dati personali, a totale disposizione di pochi privati, con tutte le conseguenze relative alle potenzialità manipolatorie e politiche insite nel possesso di tale massa di dati; dall'altra, la nuova e ubiqua sensibilità verso il concetto stesso di *dato* e verso la sua controversa natura, suscitata globalmente da un'esperienza di ricerca quali le STS, non *mainstream* ma certamente nota anche a chi non vi ha direttamente contribuito, che ha messo in dialogo permanente, e direi definitivo, i diversi tipi di dati.
- 10 L'esperienza di confronto sulla natura dei diversi tipi di dati promossa dalle STS ha avviato inoltre l'esplorazione di un terreno ibrido, promettente e creativo, ma proprio per questo denso di rischi e di possibili fraintendimenti, a cominciare da quello della denominazione dell'oggetto del comune interesse. Tutti hanno cominciato a parlare di dati, riferendosi però a entità della conoscenza ben diverse. Ad oggi, la domanda sull'autorialità, sulla possibilità dialettica e sulla proprietà degli esiti della ricerca anche all'interno dei progetti in scienze sociali, sin dalle premesse di contenuto, sia quando i progetti vengono sottoposti a valutazione, fino alla consegna dei risultati, viene posta quasi esclusivamente in termini di *data management*: anche in antropologia si parla paradossalmente di “generazione dei dati”, “scambio di dati”, “interpretazione dei dati”. Il *dato* è diventato il fulcro della discussione sul contenuto, la sua archiviazione è diventata il nodo problematico principale del progetto.
- 11 Di fronte a questo scivolamento, al contrario, vorrei osservare quanto sia limitante e in certi casi proprio sbagliato parlare in termini di *dato* riferendosi alla materia prima plastica e dinamica con cui lavora la conoscenza, antropologica ma non solo: sia dal punto di vista del rigore del ragionamento (chi dà cosa?) sia dal punto di vista della premessa epistemologica implicita che questo presuppone (esiste in sé qualcosa da dare); concordo con coloro che rigettano la nozione di *dato, given*, consegnato e statico, sempre e a maggior ragione nella ricerca sociale. La premessa epistemologica implicita è infatti quella secondo cui, da qualche parte, esiste un'informazione solida e attendibile (al pari di un dato telefonico o bancario), che può essere raggiunta e rivelata

grazie a un opportuno processo di ricerca, attraverso il quale l'informazione medesima viene presa, catturata, acquisita, o più gentilmente e semplicemente data (*given*) al ricercatore. Il ricercatore in questa visione è quindi una sorta di "ricevente del dato", e lo diventa a maggior ragione quando il processo di interrogazione della sua materia di studio è andato a buon fine – cioè il dato si è dimostrato tale, fisso e stabile. La questione, come si osserva facilmente, diventa rapidamente ontologica e varia sostanzialmente in base alla natura, alla qualità e all'oggetto (e non, lo voglio sottolineare, "in base al processo") della ricerca attivata dal ricercatore; a seconda dei casi sarà più o meno legittimo, opportuno o addirittura inutile porsi la domanda dell'esistenza primaria dell'informazione in quanto entità dotata di un suo statuto ontologico, pregresso alla ricerca. Sto quindi affermando che, per molta parte della ricerca, non esiste una verità pregressa da svelare e dunque nessun dato da dare.

- 12 Quanto appena detto apparirà molto anomalo per alcuni ricercatori, mentre per altri la stessa affermazione sarà scontata e forse ovvia: questa varietà di reazioni è la prima prova tangibile che sicuramente i *dati* non sono gli stessi per tutti, e non possono essere trattati in maniera indifferenziata riferendosi agli ambiti della vita sociale, meno che mai alle discipline della ricerca. Quando discutiamo delle entità della conoscenza manipolate in antropologia e in biologia, in antichistica e in scienza dei materiali, osserviamo come la situazione autoriale, dialettica, proprietaria – e dunque tutta l'architettura che investe la giurisdizione del *dato* – viene posta in modo sostanzialmente diverso. Assumere come ipotesi queste diversità ontologiche tra i dati di diversa provenienza e contenuto significa saper entrare nel merito della dinamica e della specificità delle forme della conoscenza, ed è esattamente il contrario di ciò che assume il GDPR.
- 13 Il dibattito sulla legittimità e sulla in/opportunità di trattare i dati della ricerca in modo indifferenziato rispetto alle discipline, con particolare attenzione all'antropologia, è stato affrontato in dettaglio da Peter Pels *et al.* in un forum ospitato sulla rivista dell'EASA¹⁰. I colleghi hanno messo a fuoco in modo esemplare la specificità del concetto di *dato* per la ricerca etnografica, e il dibattito merita di essere letto per esteso.

Per chi è il GDPR

- 14 Come rilevato da Barrera nell'intervista pubblicata in questo numero, l'urgenza di legiferare in materia di protezione dei dati viene dettata alla Commissione Europea non certo dalla necessità di proteggere chi si trova coinvolto in una ricerca, ma piuttosto di proteggere l'utente (non più identificato come cittadino, e neanche come informatore) dall'uso non trasparente dei propri dati sensibili da parte delle grandi corporation della comunicazione, da quelle telefoniche a quelle del web, a cui è *obbligatorio* cederli in cambio dell'accesso ai rispettivi servizi, in un processo di accondiscendenza non negoziabile e spesso non reversibile. Che i dati personali abbiano un valore commerciale è ormai certo, ma altrettanto certo è che la maggior parte dei *dati* della ricerca non ne ha alcuno. E qui emerge la prima contraddizione tra la (giusta) tutela dell'utente di fronte allo strapotere economico delle corporation della comunicazione e la (dubitabile) necessità di proteggere gli informatori della ricerca con cui abbiamo lavorato, come antropologi, per mesi o anni. Se parliamo della protezione delle identità di chi è coinvolto in ricerche su delinquenza, prostituzione, traffico di stupefacenti,

terrorismo, stiamo ponendo una questione ovvia nella ricerca antropologica; esiste infatti una nutrita letteratura su cui in questa sede non è il caso di insistere, in quanto è sempre stata cura dei ricercatori delle scienze sociali di anonimizzare le fonti, e in ogni caso di proteggerle. Certamente è possibile migliorare le strategie di anonimizzazione, ma questo non richiede un regolamento europeo da recepire in tutti gli stati membri. Come emerso dal dibattito in sede EASA, il presupposto del rapporto tra ricercatore e informatore è quello di fiducia – e dunque protezione – reciproca, mentre quello tra multinazionale delle comunicazioni e cliente è di soggiacenza – e dunque sfiducia e diffidenza preventiva.

- 15 Che cosa succede poi se da parte del ricercatore vengono violate le consegne previste dalla protezione? Il GDPR entra nel merito del tipo di protezione e delle conseguenze della sua violazione, che consistono in multe salatissime pensate per scoraggiare chiunque, gli enti di ricerca per primi, ma che certamente non scoraggiano una grande corporation che notoriamente mette a budget una cospicua voce *casualties*, incidenti, per coprire le cause legali a cui inevitabilmente va incontro la sua attività – se non altro perché opera in contesti giuridici diversi, e spesso non vale proprio la pena di adattare le politiche aziendali ad un contesto che promette di rendere commercialmente di più delle multe che infligge.
- 16 Pels e colleghi entrano ancor più nel vivo delle due ragioni per cui il GDPR non si adatta e non parla della ricerca antropologica, ovvero non descrive utilmente la specificità dei suoi dati. La prima ragione è il già citato problema epistemologico: in antropologia il dato è inevitabilmente non solo costruito, ma co-costruito con il s/oggetto della ricerca, con l'informatore, con il contesto, con il gruppo. Un antropologo da solo è una contraddizione in termini, senza contesto di relazioni la ricerca antropologica semplicemente non è *data*. Per non dire di una delle istanze fondamentali delle ricerche antropologiche, secondo cui la nostra disciplina “dà voce a ...”, capovolgendo la direzione del senso di *dare*. Ammesso che si possa usare il termine “dato”, e Pels non si sottrae al suo utilizzo, senza dubbio quelli della ricerca etnografica sono dei co-dati, delle co-informazioni, co-prodotte insieme al contesto, nel corso della relazione e grazie alla relazione che sta alla base della ricerca (Pels *et al.* 2018: 391). Quindi, se il dato appartiene a entrambi gli interlocutori di un'intervista, oppure a tutti i partecipanti di un forum, chi è ad avere bisogno di protezione? e chi necessita di essere protetto da chi? Entrambi, tutti?
- 17 La questione così posta è paradossale perché è la risposta a una domanda mal posta. Nessuno infatti dà qualcosa a qualcun altro nella relazione etnografica, o piuttosto, ogni relazione etnografica è una relazione di scambio, e come sanno bene gli antropologi le relazioni di scambio sono biunivoche, reciproche, ricorsive e complesse. Quindi il già molto discusso concetto di dato/*given* non si può applicare al dato etnografico – e così anche, ne deducono Pels e colleghi, a quest'ultimo non si può applicare il GDPR nella sua forma generale.
- 18 La seconda specificità del dato etnografico rilevata nel Forum è la differenza strutturale tra dato grezzo e dato processato. Il discorso si fa assai interessante, e ci riporta direttamente alla discussione iniziale introdotta dalle STS sulla modalità di costruzione del dato, questa volta *dato processato*, nelle scienze naturali. In nessun paper di ricerca infatti verranno mai messi a disposizione i dati grezzi (*row data*) delle scienze naturali. Tutte le scienze naturali dall'avvento dei computer nelle loro procedure, cioè almeno dagli anni Sessanta, si sono modellate intorno ai processi di *data analysis*. Interi istituti

di ricerca e schiere di ricercatori globali non si occupano più di trovare dati, cioè generare informazioni nuove, bensì processano (*alias* ripuliscono, organizzano, rendono *robust*) i dati prodotti da altri. La ricerca è diventata, nel senso che essa coincide esattamente con, un grande processo di analisi dei dati. I dati in sé sono solo file di numeri, matrici e sequenze, che da soli non dicono proprio nulla, sono muti e inerti, e vanno filtrati, selezionati, normalizzati, e interpretati (ma si dice “letti” per evitare rischi di soggettivismo) per arrivare a mostrare – forse – un *dato processato* di qualche significato. Se si cambia il processo di *data analysis*, pur partendo dalla stessa base di dati ne risulta modificato completamente il dato processato, e dunque il risultato della ricerca.

Esempio di dataset accademico: la VQR

- 19 Per rendere chiaro questo abisso semantico tra dato grezzo e dato processato, e come prova trasversale della dinamicità della natura dei dati processati tanto nelle scienze naturali che in quelle sociali, ricorro a un esempio noto a chiunque abbia lavorato nell'università in Italia negli ultimi dieci-quindici anni, che riguarda la procedura di VQR e le distorte (fragili, precarie) rappresentazioni della realtà accademica che questa ci ha fornito. Quando sono stati resi disponibili i “dati grezzi” delle prime due VQR del 2004-2010 e 2011-2014, ovvero i dati sulla produttività scientifica direttamente forniti dagli atenei, alcuni ricercatori hanno fatto un esperimento interessante, ovvero hanno applicato i criteri di valutazione utilizzati nella seconda VQR (2011-2014) ai dati forniti per la prima (2004-2010). I criteri di valutazione sono qui i cosiddetti algoritmi che, a partire dai giudizi sui singoli *prodotti* di ricerca (qui mi scuso per il termine) hanno configurato la valutazione dei dipartimenti e degli atenei. Nei termini del discorso precedente, i dati grezzi della produttività scientifica sono stati ripuliti, organizzati, resi *robust* nel corso della VQR per arrivare a dei dati processati che hanno permesso il confronto a livello nazionale, hanno cioè reso possibile la compilazione di una graduatoria da cui sono dipesi finanziamenti, incentivi ecc. La VQR, sottolineo, ha reso confrontabili le performance di discipline bibliometriche e non, ha prodotto valutazioni uniche per dipartimenti con eterogenei campi di ricerca, sottoposti ad eterogenei algoritmi di valutazione.
- 20 Il risultato dell'interessante esperimento è consistito in una nuova classifica delle università (cioè una nuova valutazione presupposta oggettiva delle università) basata di nuovo sui dati grezzi della prima VQR, ma completamente diversa dalla precedente. Questa nuova classifica – vecchi dati e nuovi criteri, cioè nuovi processi di analisi dei dati, nuovi algoritmi – avrebbe potuto essere confrontata più sensatamente con la seconda VQR, essendo generata attraverso le medesime procedure (naturalmente questo poteva accadere solo in un discorso ipotetico, che non si è mai svolto in sede istituzionale). Interessante, eppure anche questo confronto non sarebbe stato del tutto “giusto”, non sarebbe stato cioè un confronto tra pari: perché nel frattempo, tra la prima e la seconda VQR sono cambiati i criteri per il conferimento dei prodotti di ricerca (qui ci vorrebbe un altro articolo per discutere la semantica di “prodotto” e quanto sia impreciso e fuorviante usare questo termine per descrivere ciò che la ricerca fa).
- 21 Tra la prima e la seconda VQR infatti è cambiata la fisionomia del database all'origine della valutazione, cioè è cambiato l'assetto d'insieme delle informazioni predisposte per

essere processate da algoritmi (che a loro volta erano cambiati). Dunque al fine di ottenere una valutazione comparativa – la VQR viene ripetuta periodicamente affinché si possano comparare le sue valutazioni nel tempo – sono state processate basi di dati con strutture diverse, e attraverso criteri diversi¹¹. È diventato quindi palese e chiaro agli occhi di tutti che con gli stessi “dati” si possono costruire molti e discordanti “prodotti” (l’ANVUR converrà che si possano chiamare prodotti di ricerca anche gli esiti delle sue valutazioni), e che le classifiche prodotte dalla VQR negli anni non sono confrontabili in senso stretto; forse lo sono in senso lato, cioè storicamente, in modo contingente, ovvero considerata la società (!), la situazione economica (!) e il collettivo umano che le ha prodotte (!!). Quanto meno è possibile dire che le classifiche della VQR sono *sia* socialmente *che* scientificamente *costruite*.

GDPR e antropologi/he sul campo

- 22 Dal momento che ciò che conta è quindi il dato processato, cioè il risultato, è lecito chiedersi perché all’antropologia e alla ricerca in generale il GDPR chieda di rendere trasparenti i dati grezzi. I dati etnografici grezzi sono per noi le voci registrate, le note di campo, gli appunti in brutta copia, le riprese video non montate, tutti materiali delicati costruiti sulla, e grazie alla, fiducia, che niente hanno a che vedere con i dati da proteggere dalla metamorfosi loro imposta da un “processo” di normalizzazione e standardizzazione. Sull’inutilità della condivisione dei dati antropologici grezzi Pels ritorna descrivendo un esempio lampante di frode accademica (manipolazione e falsificazione della base dei dati), venuto alla luce recentemente nei Paesi Bassi non già attraverso la condivisione della base di dati – che era stata fatta con i pari dal ricercatore fraudolento – ma piuttosto grazie all’attenta *peer-review* di giovani studiosi che hanno considerato “del tutto inverosimili” i risultati presentati, con tanta determinazione da costringere l’autore a rivelare lui stesso la manipolazione del proprio database (Pels *et al.* 2018: 400). Ancora una volta, la posizione giuridica per garantire la sicurezza assunta dal GDPR si mostra inservibile nella protezione della (buona) ricerca.
- 23 Proseguendo sulla falsariga della specificità del dato etnografico Pels *et al.* (2018: 408) tornano sul concetto di co-costruzione introdotto, e si soffermano sulla indispensabile relazione di fiducia che deve intercorrere tra l’etnografo e i suoi co-autori (citiamo le espressioni di Pels *et al.*) affinché la ricerca etnografica possa cominciare a svolgersi. Senza relazione di fiducia non avviene alcuno scambio di informazioni, e senza informazioni e condivisioni, come già detto, non c’è ricerca. La nozione di fiducia è essenziale da evidenziare come prerequisito della nostra ricerca perché è esattamente ciò che manca nella relazione tra le corporation della comunicazione e i suoi utenti. In questa seconda relazione infatti i dati scambiati non sono né co-costruiti né co-evoluti, ma sono pretesi a partire da una posizione di forza ad un utente che non può sottrarsi, e che dunque instaura con il database una relazione sospettosa come automatismo di difesa (Pels *et al.* 2018: 412).
- 24 La relazione di fiducia nella nostra ricerca è molto più antica del GDPR. Negli statuti della ricerca antropologica americana sin dagli anni ’40 del Novecento si trovano evidenziati i criteri del *no harm* e di *reciprocal trust*, visti come assunti indispensabili senza i quali la ricerca antropologica proprio non sussiste¹². Talvolta alcuni dati sono “rubati” anche dal più specchiato antropologo: questi li potremmo disconoscere e

valutare meno nobili degli altri, ma in ogni caso sono stati rubati in una circostanza costruita sulla fiducia, che a sua volta ha reso possibile il furto. Posso fare qui un esempio legato ad una mia ricerca, svolta grazie a delle persone che avevano la mia piena fiducia, e viceversa, le quali mi hanno permesso l'accesso ad archivi che altrimenti non avrei raggiunto, le cui informazioni non erano pensate per essere condivise, né con me né attraverso una ricerca del mio tipo. Questi sono rientrati come *dati grigi* nella mia documentazione (report interni dell'azienda, linee guida per lo sviluppo strategico, contenuti dei corsi di formazione, ecc.) e non sono stati dischiusi per l'accesso ad altri ricercatori su alcuna piattaforma. Tuttavia nelle fasi di costruzione e poi valutazione della mia ricerca io li ho condivisi con i colleghi vicini, con chi ha corretto e rivisto le mie pubblicazioni, ottenendo l'assenso delle mie fonti alla condivisione caso per caso. Così alla base della mia ricerca è rimasta la relazione di fiducia, cosa che sarebbe stata impossibile se avessi perso il controllo sulla condivisione, come richiede ora il GDPR. Anzi direi di più, alla relazione di fiducia che stava alla base del mio accesso ai dati si è aggiunta in qualche caso la richiesta di dischiudere miratamente i suddetti dati grigi, ricevuta da più di un informatore – direttamente e per interposta persona – in un contesto di mantenuta fiducia, e forse anche necessità reciproca. Nel caso della ricerca di cui sto parlando infatti, avevamo la certezza assoluta che il soggetto disvelato dall'apertura degli archivi non potesse essere in alcun modo danneggiato da tale decisione, al contrario, i soggetti da proteggere sarebbero stati semmai i ricercatori, e le loro reti di relazioni.

- 25 Il motivo per cui indugio su questo caso è perché esemplifica come ogni ricerca etnografica sia un caso a sé, ogni relazione di fiducia e di interdipendenza sia unica e non replicabile, ogni infrazione di codice ha un contesto, una motivazione, un destinatario – tutti elementi che devono essere letti insieme per valutare la circostanza dell'infrazione, caso per caso. Come può la compilazione di un questionario di *disclaimer* sostituire questa variabilità di contesti e situazioni?
- 26 Infine, ma non certo in ordine di importanza, si pone per la nostra ricerca un problema di relazione istituzionale: i nostri interlocutori durante le delicate fasi decisionali in cui si negoziano le condizioni di privacy da inserire nei progetti sono gli Uffici Ricerca e/o Trasferimento Tecnologico e/o Relazioni Internazionali delle università e degli enti che ospitano i progetti. Tali uffici, salvo rarissime eccezioni, sin dall'inizio della loro esistenza hanno gestito i dati provenienti dalla ricerca quantitativa, e sono forgiati sulle relative esigenze: il dato è quello che arriva dalle scienze esatte, e in quanto tale va protetto come si fa ad esempio con un brevetto. Il brevetto è una forma di protezione della ricerca, ma è possibile ottenerlo una volta stabilito e corrisposto il suo prezzo, mentre è inutile sottolineare che, al contrario, alcuni dati della ricerca sociale non devono essere condivisi mai, a nessun prezzo. Oggi, con la ricerca che si frammenta in una molteplicità di sfumature, e le procedure del GDPR che vengono applicate alla lettera essenzialmente per proteggere al massimo grado le università dal rischio di penali fuori scala per i loro budget, tali uffici prevedono ancora, e di nuovo, un solo tipo di trattamento per tutti i tipi di ricerca, di fatto azzerando tutte le differenze tra i dati (quali/quantitativo, grezzo/processato, privato/condiviso) di cui abbiamo parlato finora (cfr. Raffaetà in questo numero).
- 27 La standardizzazione della procedura per tutti i tipi di ricerca – e del pensiero – già penalizzante per come è prevista nella normativa sul GDPR, diventa ingestibile a livello locale, quando gli interlocutori non sono titolati, a causa del ruolo amministrativo che

ricoprono, a ricorrere al pensiero critico e non possono certo dedicarsi, data la missione che hanno, all'interpretazione giuridica sottile. Due elementi, la lettura critica e l'interpretazione giuridicamente corretta della normativa ma antropologicamente flessibile, che invece sono stati elaborati dalla commissione interassociativa che ha lavorato al GDPR dal 2018, e devono ora essere esplicitati e messi a disposizione di tutti gli antropologi che fanno ricerca in Italia attraverso i canali di comunicazione ufficiale della SIAC. Tali elementi e concrete indicazioni saranno presentati in un articolo di Francesca di Declich in preparazione per il prossimo numero di AAM.

BIBLIOGRAFIA

AAA, *American Anthropological Association*: www.americananthro.org.

Alliegro E. V., 2020 *Out of Place out of Control. Antropologia dell'ambiente in crisi*, CISU Roma.

Barnes B., Bloor D., Henry J., 1996 *Scientific Knowledge: A Sociological Analysis*, Athlone and Chicago University Press, London.

Bloomberg: www.bloomberg.com/europe.

Bloor D., 1976 *Knowledge and social imagery*, Routledge, London.

Bonifacio V., Vianello R. (a cura di), 2020 *Il ritmo dell'esperienza. Dieci casi etnografici per pensare i conflitti ambientali*, CLEUP, Padova.

Breda N., 2020 «La Mongolia tra natura "incontaminata" e cambiamenti climatici. Gli esempi dagli zud e dalle zone umide», in Bonifacio V., Vianello R. (a cura di), *Il ritmo dell'esperienza. Dieci casi etnografici per pensare i conflitti ambientali*, CLEUP, Padova: 91-126.

Cartwright N., 1983 *How the laws of physics lie*, Oxford University Press, Oxford.

CIS, Critical Infrastructure Studies: <https://cistudies.org>.

Clifford J., 1988 *The Predicament of Culture: Twentieth Century Ethnography, Literature, and Art*, Harvard University Press, Boston

De Koning M. et al, 2019 «Guidelines for anthropological research: Data management, ethics, and integrity», in *Ethnography*, a. XX, vol. II: 170-174.

Forbes: www.forbes.com.

Geertz C., 1973 *The Interpretation of Cultures*, Basic Books, New York.

Goodman A., Heath D., Lindee S. (eds.), 2003 *Genetic Nature/Culture: Anthropology and Science Beyond the Two Culture Divide*, University of California Press, Berkeley.

Greenpeace International: www.greenpeace.org/international.

Hacking I., 1983 *Representing and Intervening*, Cambridge University Press, Cambridge.

Haraway D. J., 1997 *Modest_Witness@Second_Millennium. Female-Man@_ Meets_OncoMouse™*, Routledge, New York.

Harding S., 1991 *Whose Science? Whose Knowledge?*, Cornell University Press, Ithaca, New York.

- Knorr Cetina K., 1981 *The manufacture of knowledge: an essay on the constructivist and contextual nature of science*, Pergamon Press, Oxford.
- Kuhn T., 1962 *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.
- IPCC, 2007 *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, IPCC, Geneva.
- IPCC, 2014 *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, New York.
- Latour B., 1987 *Science in action: how to follow scientists and engineers through society*, Harvard University Press, Cambridge, Massachusetts.
- Lynch M., 1993 *Truth in Context, An Essay on Pluralism and Objectivity*, MIT Press, Cambridge, Massachusetts.
- Marcus G., Clifford J., 1986 *Writing Culture. The Poetics and Politics of Ethnography*, University of California Press, Berkeley.
- Pels P., et al., 2018 «Forum: Data management in anthropology: the next phase in ethics governance?» in *Social Anthropology/Anthropologie Social*, a. XXVI, vol. III: 391-413 <https://onlinelibrary.wiley.com/doi/full/10.1111/1469-8676.12526>.
- Pickering A., 1984 *Constructing Quarks: A sociological history of particle physics*, University of Chicago Press, Chicago.
- Rabinow P., 1996 *Essays on the Anthropology of Reason*, Princeton University Press, Princeton.
- Roars – Returns on Academic Research: www.roars.it/online.
- Rorty R., 1991 *Objecivity, Relativism, and Truth. Philosophical Papers - Vol. I*, Cambridge University Press, Cambridge.
- Van Aken M., 2020 *Campati per aria*, Elèuthera, Milano.

NOTE

1. Il webinar “General Data Protection Regulation in Antropologia”, organizzato da chi scrive e dal Direttivo della SIAC, si è svolto il 4/02/2021 ed è visibile qui: www.siacantropologia.it/appuntamenti/general-data-protection-regulation-in-antropologia. Hanno partecipato Francesca Declich, Giulia Barrera, Roberta Raffaetà; ha coordinato Alessandro Simonicca, Presidente del Collegio dei Proviri della SIAC.
2. Qui si veda una ricostruzione della principale genealogia delle STS secondo la Harvard Kennedy School di Harvard: <https://sts.hks.harvard.edu/about/whatissts.html>. Qui la pagina principale di EASST (*European Association of Science and Technology Studies*), la principale associazione professionale degli STS in Europa: https://easst.net/sts_journal. Qui una bibliografia a cura di studiosi indipendenti, non esaustiva ma sempre in progress: <https://cistudies.org/critical-infrastructures-bibliography/ci-bibliography-sts> (ultime consultazioni delle pagine citate: 06/2021).
3. Interessante notare la proliferazione dei programmi accademici in STS ai livelli (BA, MA, PhD), tutti sbilanciati verso le analisi della contemporaneità, recensiti da Cornell University già nel 2014: <https://sts.cornell.edu/sites/sts/files/STS%20Programs%20Ver%20V.pdf>.

Anche interessante il programma degli *Anthropocene Campus* che si sono sviluppati con la *mission* di far confluire nella didattica *social data* e *science data*: www.anthropocene-curriculum.org (ultima consultazione 06/2021).

4. Qui mi riferisco in particolare al linguistic turn (Russell, Foucault, Quine), che in antropologia prende corpo con *The Interpretation of Cultures* (Geertz 1973), *Writing Culture: The Poetics and Politics of Ethnography* (Clifford, Marcus 1986), *The Predicament of Culture: Twentieth Century Ethnography, Literature, and Art* (Clifford 1988).

5. Solo per fornire una cornice bibliografica, riporto qui alcune delle voci più autorevoli del dibattito dentro e intorno alle scienze naturali, che hanno portato a far vacillare in modo definitivo e assai produttivo la nozione di dato come *given*: Bloor 1976, Latour 1979, Cetina-Knorr 1981, Cartwright 1983, Hacking 1983, Pickering 1984, Harding 1991, Rorty 1991. Ovviamente la lista è largamente incompleta.

6. Fino al Rapporto precedente dell'IPCC, l'Assessment Report IV, 2007, la gestione delle tematiche di mitigazione e gestione della crisi climatica era considerata competenza delle sole discipline di fisica, geologia, climatologia e ingegneria del clima, con le relative estensioni ai biomi botanici e zoologici, ma non a quelli umani.

7. Cfr. tra gli altri, tra i contributi più recenti in Italia: Alliegro 2020, Bonifacio e Vianello 2020, Breda 2020, M. Van Aken 2020.

8. Si possono vedere i dettagli di questa correlazione secondo il direttore della *School of Public Health di Harvard* www.hsph.harvard.edu/c-change/subtopics/coronavirus-and-climate-change/. Questa la correlazione secondo il rapporto di Greenpeace: www.greenpeace.org/international/press-release/47042/who-covid-19-report-clear-link-between-biodiversity-loss-zoonotic-disease/ (ultima consultazione 06/2021).

9. Versione di Forbes: www.forbes.com/sites/parmyolson/2014/10/06/facebook-closes-19-billion-whatsapp-deal/. Versione di Bloomberg: www.bloomberg.com/news/features/2020-12-09/facebook-fb-plans-to-turn-messaging-app-whatsapp-into-a-moneymaking-business (ultima consultazione 06/2021).

10. Pels *et al.*, Forum: Data management in anthropology: the next phase in ethics governance? In *Social Anthropology, Anthropologie Sociale* 2018, <https://onlinelibrary.wiley.com/doi/full/10.1111/1469-8676.12526>.

11. Qui è possibile ricostruire come la valutazione d'insieme sia prodotta da quella dei singoli, D. Pinchera pubblicato su ROARS 2013, www.roars.it/online/vqr-classifiche-ed-errori-di-misura/. Qui si può consultare l'articolo di G. De Nicolao sulla propagazione e amplificazione degli errori, sempre su ROARS, 2016: www.roars.it/online/una-lezioncina-di-aritmetica-per-il-consiglio-direttivo-dellanvur (ultime consultazioni 06/2021).

12. Qui e nelle pagine seguenti è possibile consultare gli articoli inclusi nel codice etico della AAA (*American Anthropological Association*) <http://ethics.americananthro.org/ethics-statement-1-do-no-harm/> in particolare focalizzati sui principi di responsabilità professionale e di esclusione del danno (ultima consultazione 06/2021).

RIASSUNTI

La ricerca antropologica e in generale quella sociale si trovano ad affrontare le restrizioni e le nuove procedure di gestione dei dati stabilite dalla nuova regolamentazione europea, entrata in

vigore nel 2018. Prima di entrare nel merito tecnico di cosa sia ancora possibile fare e di come poterlo fare nel rispetto della regolamentazione vigente, il saggio si chiede in che misura il concetto di dato per come inteso dal GDPR si adatti al caso della nostra disciplina. Il saggio analizza la genesi e la specificità dei dati antropologici, principalmente di quelli etnografici, affronta le questioni epistemologiche e procedurali poste dalla diversa natura dei dati in antropologia e in altre discipline del sapere, e richiama varie tradizioni di studi – in filosofia, storia, critica culturale e linguistica – che hanno approfondito il tema della genesi della conoscenza, contribuendo a mettere in luce la limitatezza e i rischi di un approccio che propone una gestione unica per una molteplicità di circostanze della ricerca assai variabile e non riducibile.

Anthropological research and social research in general are facing the restrictions and the new data management procedures established by the new European regulation, established in 2018. Before analyzing the technical merits of what is still possible to do and how to do it in compliance with current regulations, the essay discusses to what extent the concept of data as understood by the GDPR fits the case of our discipline. The essay analyzes the genesis and specificity of anthropological data, mainly ethnographic data, and raises epistemological and procedural issues related to the different nature of data in anthropology and in other disciplines. Recalling various traditions of studies – in philosophy, history, cultural criticism and linguistic – which have deeply discussed the genesis of knowledge, the essay highlights the limitations and risks of an approach that proposes a single management for a multiplicity of variable and non-reducible research circumstances.

INDICE

Keywords : data construction, database, algorithm, protection, ethnography

Parole chiave : costruzione del dato, database, algoritmo, protezione, etnografia

AUTORE

ELENA BOUGLEUX

Dipartimento di Lingue, Letterature e Culture Straniere, Università di Bergamo
elena.bogleux@unibg.it