



Dati, modelli e interpretazioni: verso una fase filosofica nelle *digital humanities*

Teresa Numerico
Università di Roma 3
teresa.numerico@uniroma3.it

Abstract

This talk aims at unmasking the idea of data as conceived in most of the literature on Big Data, as well as in some of distant reading practices. This idea is based on the view that data, particularly if it is digitally represented, keeps its purity and is not contaminated by any kind of subjective interpretation, while its unique interpretation is that provided by the algorithmic interrogation of the data. I will discuss the issue in relation to the original genealogy of digital data, preserved in computer memory. According to Joseph Licklider – originator of the concept of the ‘fund of knowledge’ as a feasible digital project – the flow of data contained in an automatically accessible repository does not exhibit any special characteristic and does not substantially differ from the object of its representation. I will argue, instead, that data exists only according to a chosen interpretation and that it is accessible only through an implicit acceptance of its built-in meaning - that is, a meaning that depends on the fetishization of quantification or of another general model in which we want to capture it. The final part of the paper

suggests some new issues that need to be addressed in the new agenda of the critical digital humanities.

Keywords: distant reading, datification, critical digital humanities

1. Introduzione

La tesi generale di questo articolo riguarda lo statuto epistemologico da attribuire ai dati raccolti nel contesto disciplinare delle *digital humanities*. Sarebbe, infatti, ingenuo ritenere che i dati raccolti nell'ambito della datificazione delle ricerche testuali delle discipline umanistiche esistano in una forma pura. La loro costituzione è invece il frutto di un'interpretazione, forse implicita e talvolta inscritta nei programmi usati per trattare questi oggetti di ricerca, ma non per questo dotata di una caratterizzazione meno ermeneutica e soggettiva di altre strategie e di altri metodi di organizzazione dell'oggetto testuale. La rappresentazione consiste nella capacità di modellare il fenomeno oggetto di ricerca facendo tacere alcune caratteristiche e amplificandone altre che sono al centro dell'interesse epistemologico e ontologico del ricercatore. Il fenomeno della feticizzazione della quantificazione non riguarda solo le scienze umanistiche, ma entro tale ambito disciplinare pesa particolarmente l'assenza di una riflessione epistemologica matura e attenta, che invece già avviene in altri contesti delle scienze sociali nei quali i metodi della quantificazione e della datificazione stanno già ottenendo risultati eclatanti¹.

Il processo della datificazione riguarda ormai anche le aree umanistiche della conoscenza, dopo aver investito le scienze fisiche, biologiche e le neuroscienze e, infine, le scienze sociali. Ci troviamo quindi in un paradigma di generale trasformazione che costruisce meccanismi impliciti per la costruzione dei dati e li interpreta attraverso la costruzione di algoritmi opachi e oscuri², senza permettere alcun riscontro epistemologico sulle dinamiche di funzionamento di questi dispositivi per l'attribuzione di significato ai dati. Già in molti campi si lavora a una critica di questi strumenti e dei loro meccanismi di funzionamento, pure implicitamente discriminatori³. È opportuno aprire nell'ambito delle *digital humanities* uno spazio di discussione critica e politica che investa metodi tecniche infrastrutture e programmi di lavoro adottati nei laboratori di ricerca.

¹ Drucker (2011) e Rouvroy & Berns (2013).

² Pasquale (2015).

³ Cfr. O'Neil (2016), Pasquale (2015), Noble (2018).

Il presente articolo assume il compito di discutere alcuni metodi della datificazione nell'ambito delle *digital humanities*, per aprire una riflessione critica sull'uso epistemologico di questi strumenti. L'oggetto principale di analisi di questo lavoro sono le tecniche del *distant reading*, che uno dei metodi al momento di maggior successo e sui quali si appuntano le aspettative di rinnovamento e di innovazione delle discipline umanistiche a contatto con la digitalizzazione. Il *distant reading* è una pratica degli studi letterari resa famosa dal lavoro di ricerca e divulgazione di Franco Moretti (2013), fondatore dello *Stanford Literary Lab*. L'idea principale di questo nuovo ambito di ricerca è che sia impossibile sapere qualcosa sulla letteratura attraverso il *close reading*, cioè la lettura diretta di alcuni testi presi in esame. La letteratura è molto più dell'insieme dei testi che un critico può leggere nell'ambito della sua breve vita. Per analizzare il fenomeno letterario nella sua complessità e soprattutto nella sua ampiezza è necessario allontanarsi dalla lettura diretta dei testi e appoggiarsi alla capacità di memoria e gestione dei dati offerta dai computer. Il *distant reading* quindi include una serie di pratiche di ricerca e critica letteraria che si appoggiano agli strumenti computazionali per fare ipotesi e offrire soluzioni ai problemi di critica letteraria. I *Pamphlet* dello *Stanford Literary Lab*⁴ sono esempi interessanti e significativi di questa nuova pratica di ricerca e fanno emergere temi e tecniche per l'approfondimento letterario che ci interrogano sulla professione del futuro del critico e soprattutto sulla sua formazione e sulle competenze che gli sono richieste.

L'obiettivo dell'articolo è concentrare l'attenzione su alcuni aspetti critici delle pratiche di quantificazione del fenomeno letterario, mostrando i rischi epistemologici, sociali e politici di scelte disciplinari talvolta prive di un adeguato lavoro di analisi teorica relativa alle scelte effettuate. L'articolo propone, inoltre, alcune riflessioni, che muovono del lavoro di Berry e Fagerjord (2017), per una revisione del campo di ricerca delle *digital humanities* nel senso di includere una valutazione epistemologica sui metodi adottati e sulle loro conseguenze epistemologiche. Tale analisi critica, le cui caratteristiche non possono che essere meta-disciplinari, dovrebbe fare parte integrante del campo degli studi umanistici mediati dal computer, coinvolgendo gli studiosi in una definizione di metodi e pratiche capaci di produrre risultati originali, oltre che epistemologicamente solidi.

La progressiva quantificazione riguarda tutti i settori delle cosiddette *humanities* e propone un cambio di paradigma di studi per queste discipline

⁴ Vedi la sezione §4 per una descrizione dettagliata dei processi e delle tecniche del *distant reading*.

che ne avvicini metodi e risultati alle discipline scientifiche più tradizionali che hanno attraversato il passaggio alla quantificazione in tempi passati.

Ai fini del presente lavoro la più generale tendenza nel senso della matematizzazione e automazione dei metodi di analisi viene analizzata esclusivamente in relazione alle pratiche di analisi letteraria e la critica si concentra in particolare sulla datificazione della testualità letteraria, che viene presentata come una ricostruzione più precisa e ipoteticamente più adeguata ai fenomeni oggetto di ricerca. Nell'articolo affronterò con uno sguardo critico la scelta dei modelli e delle infrastrutture effettuata nella critica computazionale e discuterò come tali scelte rischino di essere strumenti addomesticati dai meccanismi di produzione e riproduzione di categorie neoliberali. L'articolo si propone, tra l'altro, di esercitare una critica alla ragione computazionale sulle orme di Longo⁵, con particolare riguardo ai metodi della seconda ondata di ricerche delle *digital humanities*, in particolare quello del *distant reading* e della critica letteraria computazionale.

L'articolo si propone di contribuire alla costituzione delle *critical digital humanities*⁶ con l'intento di analizzare alcuni dei metodi usati per mostrarne gli aspetti pregiudiziali che presiedono alla modellizzazione dei diversi fenomeni della testualità, intesa nella sua complessità materiale e culturale.

Il presente articolo è organizzato quindi con una prima premessa metodologica che sviluppa temi che non saranno trattati in dettaglio nell'articolo, ma la cui articolazione costituisce l'ipotesi portante su cui si fonda l'impianto critico del lavoro su rappresentazione letteraria digitale e *distant reading*.

Nel paragrafo 3 affronterò il tema della costruzione dei modelli nell'ambito delle *digital humanities* e delle difficoltà di scegliere e separare epistemologicamente tra *modelli di* una certa tendenza alla ricostruzione testuale e alle sue pratiche storiche e filologiche e *modelli per* una rappresentazione che possa intervenire sul testo e agirvi produttivamente dentro.

Dopo aver chiarito il complesso statuto epistemologico dei modelli digitali per la testualità (con speciale riguardo a quella letteraria, anche se non esclusivamente concentrati su quella), nel paragrafo 4 fornirò un'analisi

⁵ Longo (2009).

⁶ Le *critical digital humanities* sono un ambito critico delle *digital humanities* la cui crescita è sollecitata in Berry & Fagerjord (2017).

critica delle scelte di metodi e modelli proposte nell'ambito del *distant reading* dallo *Stanford Literary Lab*. Il paragrafo si addentra anche nella discussione di alcune scelte epistemologiche compiute in alcuni dei pamphlet pubblicati nel sito. La scelta dei *pamphlet* da analizzare è dettata dall'esigenza di discutere quelli che pongono più esplicitamente questioni epistemologiche, che si addentrano a giustificare il metodo utilizzato in funzione dei risultati ottenuti. Laddove, insomma, la discussione si spinge ad argomentare intorno alle strategie per raggiungere un'interpretazione creativa e originale dei testi presi in esame.

Nel successivo paragrafo 5 analizzo invece i problemi di politica accademica sulla costruzione di un *expertise* necessario per la formazione degli umanisti digitali. È evidente che se si sceglie di formare gli studiosi futuri soprattutto nell'ambito disciplinare tecnico, si decide di abbandonare di costruire le competenze critiche ed epistemologiche che sono molto importanti nella costruzione di un esperto di *humanities*. Maggiori sono le richieste di *training* tecnico minori sono le possibilità di un *training* critico e ermeneutico per queste nuove figure professionali. Non c'è una netta contrapposizione in teoria sulla formazione di figure con competenze procedurali, ma è necessario tenere a mente gli obiettivi, gli scopi specifici delle discipline dell'ambito delle *humanities*. Altrimenti si rischia una lenta eutanasia di questi campi di ricerca in favore della costruzione di tecnici che sono fungibili a tutti gli ambiti di ricerche, sia quelle più specificamente umanistiche, sia quelle più specificamente ingegneristiche e di programmazione.

Il paragrafo 6 affronta specificamente ma sinteticamente i rischi a cui vanno incontro gli studiosi che si affidano nei campi umanistici esclusivamente a metodi *computer assisted* o *computer based*. In questo paragrafo si prendono in esame tutti i problemi che sono stati illustrati in dettaglio nel resto dell'articolo e si avanzano delle soluzioni e delle proposte per contrastare e limitare i rischi, adottando delle pratiche che possano massimizzare i vantaggi e tenere sotto controllo le derive più pericolose dei metodi computazionali applicati all'analisi della testualità. Il paragrafo finale (7) segnala la necessità di una svolta filosofica ed epistemologica nell'ambito delle *digital humanities* capace di tenere dentro la sfera della disciplina la riflessione sui metodi, sui modelli e sulle pratiche adottate, evitando di definire delle procedure standard per realizzare le analisi testuali. In particolare si suggerisce di mantenere vivo lo sguardo critico sui metodi computazionali utilizzati, ai fini di una molteplicità di possibili soluzioni metodologiche ed ermeneutiche. Si segnala la necessità di applicare alle *digital humanities* le stesse prospettive che si adottano

nell'ambito degli *science and technology studies* che concentrano il proprio dibattito intorno alla critica dei metodi, che restano sempre carichi di una implicita teoria, la cui natura e i cui obiettivi è necessario esplicitare. Un'altra area che richiede il continuo ripensamento è lo studio delle soluzioni archivistiche. È evidente che la memoria del computer e il processo di vicariare stabilmente la funzione di ricordo umana offrono grandi opportunità e presentano molte domande irrisolte agli studiosi. Un archivio è costruito scegliendo i reperti da conservare in una regolazione della conservazione, che è il frutto di un giudizio sulla sua rilevanza. È necessario quindi articolare quello che sta dentro e quello che sta fuori dall'archivio⁷ e come vengono conservati i documenti che devono essere analizzati. L'ambito delle ricerche critiche può inoltre svolgersi anche intorno alla scelta di aree dominanti e dominate della disciplina. La centralità dell'inglese per esempio nelle strutture dei programmi adottati implicitamente adotta una centralità della lingua che non può che danneggiare la specificità degli studi umanistici anche quelli strettamente testuali. La questione della centralità dell'inglese è un caso più specifico di una più generale necessità dello sviluppo di un'area di *critical digital humanities* che si concentri su uno studio genealogico della riorganizzazione culturale rappresentata dal computer e dalla digitalizzazione. L'obiettivo è una costante riflessione sul ruolo del dispositivo digitale con la sua storia e il suo sviluppo nell'ambito di una riorganizzazione generale della conoscenza, dei suoi metodi e dei suoi obiettivi.

2. Premessa epistemologica sulla datificazione e sulle possibili conseguenze per le discipline umanistiche

In questa premessa metodologica si trattano questioni che non verranno discusse in dettaglio nell'articolo, ma che ne costituiscono il presupposto imprescindibile. La logica e la cibernetica tra gli anni Trenta e Quaranta del secolo scorso hanno orientato la costruzione di un dispositivo elettronico a programma memorizzato, comunemente noto come computer che ha trasformato profondamente nel tempo l'organizzazione della conoscenza. Anche le discipline umanistiche sono profondamente interessate da quello che si potrebbe considerare un mutamento di paradigma, tale da rinnovare alcune cruciali scelte metodologiche disciplinari. In particolare è nel contesto di una nuova strategia per la creazione di modelli della testualità che abbiamo visto nascere e svilupparsi la digitalizzazione, e le sue pratiche

⁷ Cfr. Derrida (1995)

discorsive e metodologiche che si trovano al centro della riflessione delle *digital humanities*⁸.

Nell'articolo *La memoria e la rete*⁹ ho affrontato il tema dello status ontologico ed epistemologico della quantificazione dei contenuti (non solo letterari) in relazione a una genealogia del concetto di memoria digitale del computer. Secondo Joseph Licklider (1965) – iniziatore dell'idea di “fondo della conoscenza” – il flusso dei dati contenuto in un magazzino digitale accessibile automaticamente non produce nessuna trasformazione sui dati e non fa nessuna differenza tra il contenuto della memoria e l'oggetto della rappresentazione conservato nelle sue partizioni¹⁰.

Come è stato suggerito da Lewis Mumford¹¹ se decidiamo di concentrare la nostra attenzione sulla quantità, stiamo trasformando la nostra capacità ermeneutica. La teoria di un *surplus* di significato dovuto all'automazione computazionale dei dati, senza implicare una qualsiasi analisi soggettiva, consolida l'autonomia del sistema, sia tecnicamente che simbolicamente, con delle conseguenze non sempre affidabili per i risultati delle discipline.

Ci sono molti analisti che ipotizzano la fine della teoria, il più famoso dei quali è Chris Anderson, uno dei cofondatori di Wired, che in un famoso e citatissimo articolo del 2008¹² discute la trasformazione epistemologica alla quale ci sta conducendo l'enorme quantità di dati disponibili in tutti i campi della ricerca scientifica, con particolare riguardo alle scienze sociali e umane. La tesi di Anderson è che ormai la quantità di dati da interpretare è talmente cresciuta da rendere inutile il ricorso a teorie e modelli che sarebbero comunque sempre parziali e in ultima analisi scorretti, e che sempre più spesso sono completamente inadeguati a dare senso alla grande quantità di dati. La spiegazione causale che i modelli tentano di ricostruire, sia pure approssimativamente, sarebbe soppiantata dalla possibilità di costruire correlazioni talmente stringenti tra serie complesse di ingenti quantità di dati, da permettere di costruire interpretazioni e previsioni, principale obiettivo dell'analisi dei dati.

L'argomento di Anderson si basa su un assunto che non viene discusso e che ne sostiene l'impalcatura, secondo il quale ci sarebbe una differenza profonda tra la natura dei dati che, come i fatti, sarebbero automaticamente a disposizione di chi fa ricerca, e quella dei modelli interpretativi che

⁸ Cfr. Fiorimonte *et al.* (2015, capp. 1-2).

⁹ Numerico (2017).

¹⁰ Licklider (1965).

¹¹ Mumford (1970).

¹² Cfr. Anderson (2008).

sarebbero invece frutto di scelte epistemologiche teoriche e per questo considerati non oggettivi e inaffidabili. Ma come suggerisce anche il titolo di un testo molto interessante a cura di Lisa Gitelman, *Raw data is an oxymoron*¹³ è difficile credere che i dati si possano dare senza essere stati organizzati e senza che una configurazione sia intervenuta a mettere ordine al loro interno. I dati sono costruiti come le teorie e i modelli.

Sembra, infatti, credibile che i dati bruti (*raw* vuol dire crudo in inglese) non esistano¹⁴. Esiste inevitabilmente una struttura che trasforma il contenuto perché ne costruisce i confini (il *corpus* di un certo periodo storico nell'ambito di uno strumento per l'analisi dei dati come Ngram¹⁵, messo a disposizione da Google, attraverso il suo magazzino di libri digitalizzati, Google Books) o ne definisce la descrizione (il meccanismo con cui si attribuiscono tag sintattici o semantici alle unità di significato). È chiaro, perciò, che è necessario nei progetti di digitalizzazione e *distant reading* definire le tecniche e i modelli che utilizziamo per "cucinare" i dati. Dobbiamo, infatti, preparare i dati perché gli algoritmi possano maneggiarli e creare le correlazioni e le interpretazioni di interesse per il progetto di ricerca. Se anche i dati sono costruiti subiscono gli stessi problemi. Possono essere oggetto di pregiudizi, o di scelte soggettive e parziali esattamente come la modellizzazione epistemologica.

Dobbiamo accordarci che la correlazione tra serie di dati e la ripetizione di certi schemi identificati possa costituire un'interpretazione o una spiegazione del testo, come è implicitamente suggerito da strumenti come Google Ngram; abbiamo bisogno di scegliere quali correlazioni abbiano senso nella nostra rappresentazione. Come indicato da Gillespie¹⁶ è cruciale esercitare un approfondimento critico sulla capacità interpretativa degli algoritmi e delle azioni preparatorie che eseguiamo sui dati perché siano gestiti e spiegati dagli algoritmi.

Andrew Piper¹⁷ nell'editoriale di lancio del primo numero del *Journal of cultural analytics*, una delle riviste di punta di una disciplina dallo stesso

¹³ Gitelman (2013).

¹⁴ Leonelli (2018).

¹⁵ Google Ngram viewer è uno strumento che consente di calcolare la frequenza di una serie di termini all'interno di un *corpus* definito a partire dalla collezione di Google Books. I *corpora* vanno dal 1500 al 2008 e possono essere specializzati soprattutto nell'ambito di testi a stampa scritti in inglese. Lo strumento ha molti limiti che vanno dal riconoscimento dell'OCR e quindi dalla possibilità dell'errore, alla mancanza di un *corpus* completo soprattutto nelle lingue meno rappresentate in Google Books, alle scelte di traslitterazione per le lingue che non usano l'alfabeto latino ecc.

¹⁶ Gillespie (2014).

¹⁷ Piper (2016).

nome, inventata da Manovich il cui obiettivo è descrivere le nuove prospettive di una disciplina che affronti l'analisi culturale a partire dalla digitalizzazione, afferma che: «we can think of computation as cultural studies' *pharmakon*, as Plato intended it – a remedy that can also be a poison». Possiamo interpretare veleno nei termini di una trasformazione disciplinare che non consente alle altre metodologie di sopravvivere nel nuovo scenario, senza nemmeno comprendere gli effetti epistemici opachi delle nuove prospettive e interfacce adottate (cfr. Rockwell & Sinclair (2016), sul rapporto tra ermeneutica e dispositivi infrastrutturali). È interessante notare che il termine *pharmakon* sia centrale anche nella filosofia della tecnologia di Bernard Stiegler¹⁸, un notevole studioso francese. L'ambivalenza del *pharmakon* getta una luce problematica anche sui rapporti di potere di natura disciplinare tra esperti di *humanities* e di tecnologia.

Evelyn Fox Keller¹⁹, una epistemologa esperta soprattutto di biologia, suggerisce che una volta adottata una metodologia o una modellizzazione trasformiamo definitivamente l'oggetto di osservazione scientifica. Applicare il nuovo modello materiale al testo come una dimensione epistemologica del tutto nuova richiede una notevole attenzione alla scelta delle interfacce, cioè delle infrastrutture computazionali, le cui conseguenze sull'oggetto di ricerca possiamo solo vagamente prevedere e sulle quali non abbiamo concordato preventivamente. Sarebbe auspicabile esercitare una consapevolezza preventiva che la selezione dei metodi, con i connessi strumenti tecnici necessari per ottenere i risultati desiderati, possa definitivamente distruggere l'oggetto precedente dell'analisi letteraria imponendo un nuovo progetto di interfaccia, un modello per filtrarlo, tradurlo e comprenderlo. L'effetto ontologico sulla trasformazione della realtà dell'oggetto dell'investigazione scientifica è una questione di politica della ricerca e richiede agli studiosi e alle studiose una scelta di posizionamento critico ed etico. Possiamo sperimentare liberamente, ma dobbiamo comunque prendere in considerazione il significato e gli effetti degli strumenti tecnici che scegliamo di adottare per ottenere una nuova rappresentazione dell'oggetto di ricerca.

È molto difficile ostacolare il processo di datificazione in sé. Come suggerisce Stiegler²⁰, le tecniche di esternalizzazione della memoria non sono nate con la costruzione degli algoritmi, la storia dell'umanità è segnata

¹⁸ Cfr. Stiegler (2015) e Stiegler (2015a).

¹⁹ Keller Fox (1991).

²⁰ Cfr. Stiegler (2015) e Stiegler (2015a).

dall'evoluzione tecnologica. Tuttavia – e qui sta il punto da discutere se non vogliamo ricadere in un paradigma deterministico – la scelta di quali adottare per esternalizzare la memoria e del significato da attribuire loro è e resta una scelta politica. Una scelta di politica della conoscenza.

Le conseguenze della visione di Stiegler sui meccanismi di controllo e orientamento della memoria esternalizzata vengono proposte da Rouvroy e Berns²¹ che introducono il concetto di governamentalità algoritmica a proposito dell'uso e della costruzione dei dati e della loro interpretazione a partire dalle tracce lasciate dagli utenti dei servizi digitali, che percepiscono la perdita di controllo sulla loro vita come qualcosa di inevitabile, incontrollabile e prestabilito.

In questa premessa metodologica ho sollevato delle questioni che non riguardano solo l'epistemologia delle *digital humanities*, ma ho proposto una riflessione generale sui processi della datificazione delle discipline, e sulla tendenza a favorire la valutazione degli aspetti esclusivamente quantitativi dei fenomeni in oggetto.

3. Quantificazione e modellizzazione delle *digital humanities*

Il processo di costruzione della scienza moderna passa per una fase di matematizzazione delle conoscenze, con l'obiettivo della misurabilità dei fenomeni sotto osservazione²². L'informatica e la sua antesignana cibernetica sono strumenti che contribuiscono, attraverso il meccanismo della digitalizzazione a una facilitazione dei processi di datificazione, processi che innescano una dimensione quantitativa e misurabile che include quindi anche l'oggetto delle discipline umanistiche. La misurabilità o la presunta quantificabilità dei dati usata come metodo per la costruzione di nuovi modelli e soprattutto nuove evidenze sperimentali è al centro del dibattito nelle *digital humanities*, dal momento che pone problemi di significazione e automatizza potenzialmente le pratiche interpretative, almeno parzialmente.

Secondo Willard McCarty l'attività di questa disciplina è la creazione di modelli attraverso l'uso del computer. Il progetto del suo articolo del 2004 è lavorare a una modellizzazione della creazione dei modelli nell'ambito delle scienze umane a partire da tre questioni fondamentali:

²¹ Rouvroy & Berns (2013).

²² Koyré (2000).

[...] to identify humanities computing with an intellectual ground shared by the older disciplines, so that we may say how and to what extent our field is of as well as in the humanities, how it draws from and adds to them; at the same time to reflect experience with computers "in the wild"; and to aim at the most challenging problems, and so the most intellectually rewarding future now imaginable.²³

La tesi di McCarty si configura seguendo la distinzione di Geertz tra *modelli di* e *modelli per*, individuando nel concetto di modello almeno un doppio scenario «By ‘modeling’ I mean the heuristic process of constructing and manipulating models; a ‘model’ I take to be either a representation of something for purposes of study, or a design for realizing something new»²⁴.

Il modello per le discipline umanistiche digitali serve sia come comprensione dei fenomeni, quindi è un modello di qualcosa che cerca di spiegare. Allo stesso tempo però ha anche un carattere manipolativo. Si tratta di rappresentare in modo nuovo le domande conoscitive poste dal testo, un modo che sia più potente e che offra una spiegazione migliore e più comprensiva dei diversi fenomeni testuali. Approfondendo ancora il tema del ruolo epistemologico dei *modelli di*, McCarty introduce il rapporto tra modello e rappresentazione. Un modello è sempre una rappresentazione di un fenomeno che vuole spiegare e comprendere. Il contesto della rappresentazione viene immediatamente associato a quello dell'intelligenza artificiale e in particolare dell'area della *Knowledge Representation* (KR), ma è evidente che questo argomento sollevi delle perplessità perché, sebbene la rappresentazione computazionale sia basata sulla logica, non tutti i fenomeni si prestano a essere rappresentati attraverso la logica²⁵. La conclusione alla quale perviene McCarty è che, l'uso di questi modelli computazionali come strumenti di spiegazione dei fenomeni delle *humanities*, richiede una certa cautela epistemologica:

I have argued, ours is an experimental practice, using equipment and instantiating definite methods, for the skilled application of which we need to know what we are doing as well as it can be known. I have labeled the core of this practice “modeling,” and suggested how, properly understood, modeling points the way to a computing that is of as well as in the humanities: a continual process of coming to know by manipulating representations.²⁶

²³ McCarty (2004, 254-255).

²⁴ McCarty (2004, 255).

²⁵ McCarty (2004, 261).

²⁶ McCarty (2004, 265).

Il processo pratico di manipolazione dell'oggetto della ricerca umanistica attraverso l'uso di strumenti computazionali non può essere criticato come elemento di innovazione e crescita delle *humanities*; tuttavia resta aperto il problema della spiegazione metodologica, della capacità di costruire previsioni e dell'eventualità di trovare cose nuove e inaspettate attraverso l'uso di questi strumenti.

Partendo dalla riflessione epistemologica, è necessario rendere conto delle pratiche sperimentali, dei modelli computazionali adottati e della conseguente capacità di previsione di risultati efficaci come esito di quegli esperimenti, frutto delle teorie computazionali assunte come punto di vista.

Methods are explicit, actions definite, results forthcoming, yet we have been unable fully and persuasively to articulate the intellectual case for the means by which these results are produced. Hence the just-a-tool status of computing, the not-a-discipline slur, the tradesman's entrance or other back door into the academy.²⁷

La domanda che resta senza risposta ma che sta al cuore dell'intervento di McCarty riguarda la questione antica e di difficile soluzione «to the very hard, old problem of discovery: how does this happen?» come avviene la scoperta, e ha come corollario il rischio che l'uso di modelli e pratiche che non siano adeguatamente sostenute da una riflessione epistemologica possa comportare l'entrata della dimensione commerciale nell'ambito delle ricerche accademiche. Questo problema è presente nell'ambito delle *digital humanities*, ed è esplicitamente affrontato in un articolo sull'etica dei *big data* dal titolo: *Is it Research or is it Spying? Thinking-Through Ethics in Big Data AI and Other Knowledge Sciences*²⁸. L'articolo propone una riflessione sulla purezza etica delle analisi testuali. I metodi di *text analysis*, usati nelle *digital humanities*, infatti, sono simili alle tecniche usate per spiare la corrispondenza e le interazioni delle persone connesse attraverso la rete.

Il ruolo dei modelli nella scienza è stato oggetto di interesse anche da parte di Rosenblueth e Wiener²⁹ agli albori delle ricerche sulla cibernetica, che poco dopo prese l'avvio ufficiale³⁰. L'intento dell'articolo sui modelli del 1945 era lanciare un nuovo orizzonte transdisciplinare che avrebbe

²⁷ McCarty (2004, 265).

²⁸ Berendt *et al.* (2015).

²⁹ Rosenblueth & Wiener (1945).

³⁰ Wiener (1948/1961). Per un approfondimento di questo tema della costruzione dei modelli e del loro carattere artificiale e costruttivo del fenomeno di studio vedi Cordeschi (2002).

dovuto caratterizzare la nuova area di ricerca, il cui campo d'azione si trovava nella «no-man's land between the various established fields»³¹. I due autori dichiaravano che il modello migliore di un gatto è sempre un gatto e che se noi vogliamo usare un modello (o per meglio dire, “una rappresentazione”) inevitabilmente siamo costretti a oscurare parte del fenomeno in oggetto in una “*closed box*”. La loro idea era che per risolvere i problemi della ricerca si sceglievano sempre le variabili convenienti per la soluzione migliore del problema in questione:

It is obvious, therefore, that the difference between open-box and closed-box problems, although significant, is one of degree rather than of kind. All scientific problems begin as closed-box problems, i.e., only a few of the significant variables are recognized. Scientific progress consists in a progressive opening of those boxes.

Many of these small compartments may be deliberately left closed, because they are considered only functionally, but not structurally important.³²

Ciò che viene chiuso nella scatola, o potremmo dire più precisamente rimosso, è un elemento fondamentale di cosa vogliamo far emergere come caratteristica saliente di un testo o di un altro artefatto oggetto delle *digital humanities*. La scelta delle variabili da chiudere nella scatola definisce lo spazio di dominio del metodo e l'area epistemologica delle possibilità di spiegazione. È rilevante la scelta dell'interfaccia o più in generale dell'infrastruttura³³ che viene usata per ottenere la soluzione del problema. La posizione del metodo di ricerca e delle infrastrutture disponibili, algoritmi, software, *corpora* di dati definisce il campo delle possibilità fenomeniche da ricostruire e da ricercare, ma nella maggior parte dei casi non viene determinato da considerazioni di tipo epistemologico, sebbene acquisti, una volta assunto, un ruolo decisivo sia per la ricostruzione dell'oggetto della ricerca, sia per la proposta di eventuali metodi di spiegazione (*modelli di* e *modelli per*, come suggerito da McCarty³⁴ *infra*). La scelta del modello e della conseguente infrastruttura per la ricerca esercita un potere non solo sugli altri ricercatori che si sentono impegnati in questo stesso ambito di ricerca, ma soprattutto sull'oggetto della ricerca, sulla sua rappresentazione e sulle possibilità dell'interpretazione rese agibili o ostacolate, fino a cancellarle³⁵.

³¹ Wiener (1948/1961, 2).

³² Rosenblueth & Wiener (1945, 319).

³³ Cfr. Bowker *et al.* (2009).

³⁴ McCarty (2004).

³⁵ Cfr. Keller Fox (1991).

La rimozione di quello che non può essere preso in esame in quella rappresentazione del fenomeno oggetto di studio, nel caso delle *digital humanities* cancella spesso irrimediabilmente altre possibili strade interpretative. Talvolta persino l'oggetto testuale dell'analisi nella sua versione predigitale viene perduto, adottando, nel tessuto stesso dell'oggetto di ricerca, l'infrastruttura promossa e configurata dalla digitalizzazione. Si pensi per esempio alle poesie nelle quali la struttura spaziale dei versi è parte integrante del loro significato. Nel momento in cui si rappresenta un testo di Palazzeschi per esempio attraverso una trascrizione digitale per quanto accurata, è impossibile rendere conto della dimensione spaziale del verso nel foglio. Un simile aspetto, irriproducibile allo stato delle tecniche di digitalizzazione è cancellato per sempre per chi si limita a un accesso al testo in formato digitale, per quanto si possa lavorare a una codifica dettagliata. Tale processo di rimozione di alcune variabili rilevanti non avviene sempre in maniera consapevole o profonda, favorendo l'efficienza della traduzione digitale anche a costo di minimizzare possibili errori. Un altro esempio è il caso di Ngram che fa riferimento ai *corpora* messi a disposizione da Google Books e che non solo non prestano attenzione al problema dell'eshaustività del *corpus* preso in esame, ma non ritengono nemmeno rilevante correggere la grande quantità di errori di trascrizione dovuti al funzionamento piuttosto carente dell'OCR (*Optical Character Recognition*)³⁶, né si interroga sulle conseguenze delle scelte di translitterazione adottate quando il *corpus* non è scritto in caratteri latini.

Se in un'analisi di *distant reading* di un *corpus* di scritti si tralasciano alcuni scritti o se si decide di non prestare attenzione alla polisemicità di alcune parole presenti in un contesto fortemente diacronico, mentre si attiva per quei termini una precisa ipotesi interpretativa, quelle parti mancanti o quelle scelte interpretative sono definitive e ineludibili, anche per chi affronterà i testi successivamente. Spesso, infatti, questi problemi sono considerati marginali e quindi implicitamente taciuti, rendendo più difficile la ricerca di chi volesse mettere in discussione il metodo prescelto.

La difficoltà di stabilire la rilevanza epistemologica dei modelli e delle rappresentazioni adottate, suggerito da McCarty, è persistente e anzi rischia di trasformarsi in un intervento (nel senso di Hacking³⁷) senza un'adeguata epistemologia corrispondente. Il metodo viene usato come obiettivo, dal momento che gli strumenti digitali sono disponibili e costituiscono una tentazione per lo studioso di discipline umanistiche di rinnovare il proprio

³⁶ Cfr. Roth *et al.* (2017), per un uso del sistema che mostra anche i suoi limiti.

³⁷ Hacking (1983).

armamentario teorico. Una simile attitudine spinge per una definizione delle *digital humanities* come:

Shared set of methods, tools, archives, across the humanities, without a disciplinary focus of its own, beyond that of building better tools and archives (creating tools for ‘intervening’) . [...] [the term DH] become used [...] to describe the broad kinds of research work they undertake or even a tactical or discursive construction.³⁸

La concentrazione sull’uso dei metodi disponibili senza un’adeguata critica all’aspetto teorico e alla loro portata interpretativa induce Berry e Fagerjord a invocare la necessità di introdurre nell’ambito delle *digital humanities* una riflessione epistemologica sui progetti in corso, una sorta di osservatorio culturale sulla scelta dei metodi da adottare. L’importanza di una fase di ricerche critiche che non siano esclusivamente concentrate sulla pratica sperimentale dell’uso dei metodi digitali appare ai loro occhi ineludibile.

Nella logica di costruire l’ambito della riflessione sulle *digital critical humanities* nel prossimo paragrafo discuterò come la costruzione di modelli e teorie giochi implicitamente un ruolo molto importante, se non cruciale anche nell’ambito delle *digital humanities* che più pretende di allontanarsi dalla necessità epistemologica di costruire spiegazioni teoriche. Nella prossima sezione si discuteranno i metodi e gli obiettivi del *distant reading* per mostrare come le scelte epistemologiche, talvolta implicite, possano condurre a dei risultati non del tutto motivati che dovrebbero essere oggetto di un attento esame critico.

4. Moretti e lo *Stanford Literary Lab*

Nella prima fase delle *digital humanities* il tema che aveva catturato l’attenzione degli studiosi era quello dei rapporti tra software, rappresentazione dei testi e la natura dei testi stessi. Tutte le discussioni nell’ambito della *Text Encoding Initiative* (TEI) vertevano sulla rappresentabilità del testo nei processi di codifica. Tale attenzione implicava concentrarsi sulle pratiche di rappresentazione digitale della testualità, sia nel senso delle scelte che dovevano caratterizzare la codifica dei testi, sia nel senso di esaminare i criteri da adottare nel processo di traduzione della testualità nel contesto digitale, attività che viene chiamata *textual scholarship*. Questa disciplina comprende la filologia, la storia del testo e il

³⁸ Berry & Fagerjord (2017, 33).

suo interesse critico nel senso del processo di interpretazione, anche nell'ambito della cosiddetta attività di *close reading* che queste pratiche di digitalizzazione avrebbero potuto rendere più efficace, a certe condizioni³⁹.

La seconda fase delle *digital humanities* è invece più concentrata sui meccanismi di quantificazione di grandi quantità di dati di carattere testuale e sulle possibilità epistemologiche in termini di misurabilità dei fenomeni indipendentemente dall'esercizio di un controllo sui processi di codifica. In questa fase più recente l'interesse degli studiosi è stato maggiormente impegnato a sviluppare algoritmi per la gestione di grandi quantità di dati e per la costruzione di meccanismi interpretativi basati sulla costruzione di grafi, mappe e categorizzazioni circa la presenza di certe espressioni in contesti testuali molto ampi. Il principale attore di questa trasformazione può essere considerato Franco Moretti che con i suoi testi rivoluzionari ha aperto la strada alle ricerche sul cosiddetto *distant reading*⁴⁰. La modellizzazione di grandi strutture sociali interpretative e le conseguenti capacità di spiegazione suggerite da Moretti⁴¹ richiedono una grande consapevolezza critica ed epistemologica, che non sempre è in azione negli articoli entusiasti dei neofiti delle nuove tecniche di quantificazione. Se si tratta di un semplice cambio di prospettiva e di una ricerca strutturalista allora non ci sono difficoltà o novità epistemologiche, si tratta della riproposizione di metodi già sperimentati con alterne fortune. Se invece si intende affermare che gli strumenti computazionali e l'infrastruttura tecnologica rappresentino una vera novità negli studi e nei risultati di critica letteraria, allora dobbiamo vagliare con attenzione tale posizione. È necessario mettere in campo un chiarimento sull'effetto di spiegazione del modello o della metodologia se vogliamo evitare il rischio di una *petitio principii*.

Per dare il senso della trasformazione che questo nuovo campo di ricerche propone e promuove ci riferiamo al *Pamphlet 16* che Moretti ha scritto nel 2016 (il testo è poi confluito nella pubblicazione *Canon/Archive* del 2017). La sua tesi è che il *computational criticism*, come lui preferisce chiamare le *digital humanities*, viene definito come rapporto tra misure quantitative e la definizione dei concetti usati per costruire i modelli di spiegazione della critica.

It's fascinating, how a series of quantitative measurements enters into a dialogue with concepts, and slowly transforms them. Slowly. Forget the hype

³⁹ Per una storia delle DH, cfr. Earhart (2015).

⁴⁰ Per una genealogia del *distant reading*, cfr. Underwood (2017).

⁴¹ Cfr. Moretti (2007) e Moretti (2013).

about computation making everything faster. Yes, data are gathered and analyzed with amazing speed; but the explanation of those results – unless you're happy with the first commonplace that crosses your mind – is a different story; here, only patience will do.⁴²

La tesi di Moretti – che resta il *maître à penser* del gruppo di studiosi raccolti intorno allo *Stanford Literary Lab*, centro di ricerche che ha promosso il *quantitative formalism* o meglio del *computational criticism*, nuovo nome per il *distant reading* – è che si tratti di una trasformazione nel senso della quantificazione delle ricerche letterarie e che questa rivoluzione porterà lentamente delle novità nell'ambito della critica letteraria. L'idea di Moretti apre un nuovo campo di ricerche nell'ambito delle *digital humanities*, un campo che si differenzia da quello originario della rappresentazione dei testi.

In questa seconda fase di ricerche non conta tanto in che modo i testi che costituiscono i *corpora* siano rappresentati, se siano presenti completamente integri o se si tratti solo di una parte del *corpus* da prendere in considerazione, e nemmeno se i testi siano trascritti in modo molto accurato o usando un preciso standard di marcatura. L'elemento che costituisce il maggior interesse è la disponibilità di una grande quantità di dati da analizzare. L'obiettivo delle ricerche è l'elaborazione di tecniche di analisi che permettano di mettere a valore la grande quantità di dati disponibili e soprattutto il processo stesso di quantificazione dei dati letterari. Il progetto di quantificare e identificare degli elementi distintivi per la loro quantità costituisce una sorta di punto di arrivo "*beyond concepts*" come suggerisce Moretti adottando questo titolo per una delle sezioni del suo saggio⁴³ e affermando che la quantificazione testuale produca:

[...] a form of thinking that fuses together the formulation and the operationalization of concepts, leaving them often half-implicit as concepts, while liberating their full force as algorithms.⁴⁴

L'indicazione di Moretti è che la quantità possa darci delle indicazioni di qualità e possa essere più costitutiva di altre forme di pensiero rispetto alla definizione stessa dei concetti a cui far approdare la critica computazionale: «the main contribution of computational criticism to literary study may well come from these Centaur-like creatures, half script,

⁴² Moretti (2016, 5).

⁴³ Moretti (2016).

⁴⁴ Moretti (2016, 6).

and half theory»⁴⁵. Insomma programmare è come pensare nella nuova critica, anzi forse si troverebbe in una posizione epistemologica privilegiata rispetto al pensiero perché offrirebbe una chiarezza categoriale e concettuale che la critica tradizionale non potrebbe mai garantire. Tutto l'opposto della cautela critica di McCarty⁴⁶ sui modelli e sul loro carattere di spiegazione computazionale di fenomeni che, come quelli letterari, potrebbero non avere i requisiti per essere trattati secondo questi metodi. L'obiettivo stesso dei progetti del *Literary Lab* è la creazione di strutture algoritmiche che possano costituire un corrispettivo categoriale del testo, fornire una comprensione della letteratura o della testualità imprevista e imprevedibile, completamente nuova rispetto ai modelli finora usati nella critica letteraria.

La conclusione del saggio di Moretti è che questo metodo di ricerca computazionale richieda «a double work of translation - from natural to literary objects, and from concepts to algorithms – became itself part of the point: it defined our view of the digital humanities as the form taken by a scientific-explanatory approach in the digital age»⁴⁷.

La necessità di un approccio scientifico-epistemologico dei metodi che Moretti invoca si scontra con la presenza di algoritmi non supervisionati che agiscono su una base dati e esercitano le loro implicite categorizzazioni, basate su un pregiudizio della programmazione che potrebbe anche essere e rimanere implicito per tutta la durata della ricerca, anche in presenza delle conseguenze interpretative della spiegazione adottata.

Approfondendo un altro *Pamphlet*, il numero 4 del *Literary Lab*, dal titolo *A quantitative literary history of 2958 nineteenth-century British Novels: the semantic cohort method*⁴⁸ scritto da Ryan Heuser e Long Le-Khac si possono notare i problemi metodologici che si pongono nel gruppo di ricerca. Secondo il *Pamphlet 4* la questione metodologica principale è: «How do we get from numbers to meaning?»⁴⁹ ma il passaggio epistemologico non sembra chiaro nemmeno agli autori che proseguono: «How to move from this kind of evidence and object to qualitative arguments and insights about humanistic subjects – culture, literature, art, etc. – is not clear»⁵⁰.

Commentando questo articolo Alan Liu scrive:

⁴⁵ *Ibidem*.

⁴⁶ McCarty (2004, cfr. § 2).

⁴⁷ Moretti (2016, 6-7).

⁴⁸ Heuser & Le-Khac (May 2012).

⁴⁹ Heuser & Le-Khac (2012, 46).

⁵⁰ *Ibidem*.

They help advance an important, general digital humanities goal that might be called *tabula rasa* interpretation—the initiation of interpretation through the hypothesis-free discovery of phenomena. Also shared by such digital humanities methods as topic modeling (a mathematics-based way to discern differentiated clusters of words that Heuser and Le-Khac use to corroborate their findings), the ideal in its purest form is what Heuser and Le-Khac call ‘an unsupervised method that generates topics without subjective input from users’ (28). That is, a computer should be able to read texts algorithmically and discover word cohorts or clusters leading to themes without acting on an initial concept from an interpreter looking to confirm a particular theme⁵¹.

La posizione di Liu è molto critica sull’interpretazione *tabula rasa* perché è consapevole che non esista un’interpretazione simile: il programma troverà le *cohorti* e i *cluster* che è stato addestrato a riconoscere, ma la mancanza di una avvertita posizione epistemologica sui metodi computazionali sembra invece governare le ricerche, avvalorando la discutibile tesi secondo cui i dati quantitativi sarebbero l’unico elemento capace di condurre verso l’identificazione di concetti interpretativi.

Gli autori del *Pamphlet 4* affermano che:

In moving from signal to concept, from numbers to meaning, what may be needed in fact is more numbers. More numbers and different kinds of numbers. As humanists, we may fear that gathering more quantitative data only moves us farther from the qualitative meaning we seek, but we’re suggesting that having more kinds of data actually moves us closer to finding meaning.⁵²

L’obiettivo epistemologico, ma anche interpretativo di questo tipo di lavoro sembra condurre alla ricerca di più dati, non importa come costituiti da dare in pasto a un algoritmo non supervisionato, cioè in grado di agire senza l’intervento umano. O meglio l’intervento umano consiste solo nella definizione dell’algoritmo della programmazione che dovrà attribuire senso ai dati allo stato bruto (*tabula rasa interpretation*).

La posizione sulla quale mi sento di dissentire profondamente riguarda l’impianto teorico di cui i diversi *Pamphlet* del *Literary Lab* sono espressione più o meno esplicita e più o meno convinta. Secondo Moretti e gli altri collaboratori per essere un buon critico computazionale sarebbe necessario insegnare agli studenti la capacità di produrre un codice tecnico senza formulare concetti interpretativi. Il programma costruito allo scuro di una categorizzazione funzionale dovrebbe implicitamente implementare un

⁵¹ Liu (2013, 214).

⁵² Heuser & Le-Khac (2012, 48).

metodo per riconoscere concetti interpretativi senza una valutazione critica o una scelta esplicita da parte dello studioso.

Nella prossima sezione cercherò di spiegare il conflitto di politica dei saperi. L'attenzione alle competenze tecniche del *digital scholar* rischiano di offuscare l'addestramento e la formazione delle sue capacità critiche finendo per spingere verso una figura che potrebbe occuparsi di testualità digitale dal punto di vista letterario, ma anche dal punto di vista di una *sentiment analysis* propria di certe ricerche di marketing. L'umanista digitale potrebbe essere una figura di frontiera che si applica al testo letterario, come a qualsiasi altro tipo di analisi di dati non ancora strutturati. Il rischio sarebbe che le capacità critiche non farebbero più parte integrante del bagaglio di uno studioso umanista, ma sarebbero una sorta di competenza laterale non più essenziale. Tale esito finirebbe per dissolvere un certo tipo di preparazione, a favore di una serie di competenze procedurali e tecniche, competenze di tipo meramente gestionale.

5. Conflitto tra discipline e politica della ricerca: come devono essere educati gli umanisti digitali?

La critica computazionale è un caso specifico di una tendenza generale già in atto in diverse discipline anche in ambito umanistico: l'idea che con l'aiuto del computer sia sempre più frequente analizzare i problemi attraverso la manipolazione di grandi quantità di dati, invece che con l'aiuto di modelli e metodi di spiegazione. Tale tesi è diventata di dominio pubblico con la pubblicazione su *Wired* dell'articolo sulla fine della teoria dove Anderson⁵³ discuteva una tendenza già in atto da tempo in molti campi di ricerca scientifici, di marketing e relativi ai grandi progetti statali di sorveglianza dei cittadini per motivi di sicurezza.

La tendenza alla quantificazione nelle scienze umanistiche viene discussa già negli anni Settanta. Secondo Mumford⁵⁴ le scienze spingevano per una trasformazione quantitativa e questo valeva anche per le discipline umanistiche:

Though this program for the automatic mass production of knowledge originated in science, it has been imitated in the humanities, particularly in American universities, as a sort of status symbol, to underwrite budget requests in competition with the physical and social sciences, and to provide a quantitative measure for professional promotions. Whatever the original breach

⁵³ Anderson (2008) di questo articolo abbiamo già ampiamente discusso *infra*.

⁵⁴ Mumford (1970).

between the sciences and the humanities, in method they have now – *pace* Charles Snow! – become one. Though they run different assembly lines, they belong to the same factory. The mark of their common deficiency is that neither has given any serious consideration to the results of their uncontrolled automation.⁵⁵

La spinta alla quantificazione e datificazione che include anche le discipline umanistiche, oltre a essere una trasformazione nel senso di una loro proceduralizzazione che contrasta con il carattere critico ed ermeneutico che le caratterizza, contestualmente promuove una redistribuzione delle competenze.

In un'intervista al *Los Angeles Review of Books* a marzo del 2016 Moretti dichiarò che era necessario insegnare ai potenziali studenti di questo nuovo settore della critica computazionale la capacità per pensare in termini di struttura di codifica e infatti secondo lui in futuro i risultati più promettenti «will come from scripts that are half scripts/half cultural, literary, historical concept»⁵⁶. Il tema di quali siano le competenze da attribuire agli studiosi di nuovo tipo apre una discussione sul conflitto di potere interdisciplinare.

Se finora gli studiosi di discipline umanistiche dovevano avere delle specifiche competenze critiche che potevano essere acquisite senza dover studiare corsi tecnici, in questo nuovo scenario non sarà più così: per trovare le risposte ai problemi di critica umanistica ci sarà bisogno anche di competenze tecniche, o addirittura principalmente di quelle, che sanciranno il nuovo assetto disciplinare delle *humanities* che verranno.

Anche Geoffrey Rockwell e Stéfan Sinclair nel libro *Hermeneutica*, in cui discutono delle potenzialità delle nuove tecniche applicate alle pratiche interpretative, sostengono che in questo nuovo orizzonte dobbiamo scontare alcuni rischi, tra i quali:

That entanglement might lead to the corruption of humanistic scholarship. The practices of software engineering could be the first step in the disappearance of the humanities.⁵⁷

La risposta degli autori consiste in una riflessione sulla necessità di preservare non le pratiche classiche che hanno caratterizzato le *humanities*, ma piuttosto una conversazione etica e vitale su quello che significa essere umani. L'obiettivo che si propone *Hermeneutica* è proprio permettere di

⁵⁵ Mumford (1971, 181).

⁵⁶ Moretti (2016a).

⁵⁷ Rockwell & Sinclair (2016, 20).

pensare attraverso la tecnologia a come tutelare il senso delle *humanities* e non i metodi che hanno finora usato⁵⁸.

Non tutti sono di questo stesso avviso. In un articolo di Allington *et al.*⁵⁹ sempre su *Los Angeles Review of Books* si sferra un pesante atto d'accusa sulle scelte neoliberali delle *digital humanities* con la conseguente sottomissione della disciplina al volere di una politica tardo-capitalista che si propone di cancellare ogni barlume di capacità critica, interpretativa, soggettiva e autonoma dalle discipline umanistiche. La ricostruzione storica molto violenta sull'affermazione delle DH nelle università americane è stata ampiamente discussa, avversata e sostenuta aprendo un dibattito che non è ancora concluso⁶⁰.

Tornando alla posizione di Moretti sembra che la formazione dei nuovi studiosi di questa disciplina dovrebbe concentrarsi principalmente sulle capacità tecniche, mentre l'*expertise* di tipo critico risulterebbe meno rilevante, proprio per chi deve analizzare un testo e interpretarlo. Secondo questa visione gli studiosi di *digital humanities* non dovranno ricevere uno specifico *training* professionalizzante nelle materie umanistiche, oppure questo potrà essere piuttosto superficiale. Non sembrerebbe nemmeno necessario che l'addestramento dei nuovi studiosi sia saldamente in mano agli umanisti. Forse potrebbe essere sufficiente ingaggiare dei semplici informatici per la formazione dei nuovi studiosi del settore.

Come se appunto l'attività di programmazione fosse trasversale alle discipline e non avesse grande importanza il tipo di contenuti a cui viene applicato il programma. L'esperto di *Big Data* potrebbe costruire programmi per il *marketing* e la *sentiment analysis* nell'ambito dei *social network*, così come per grandi *corpora* letterari senza grandi differenze.

Il dibattito sulla formazione è teatro di un conflitto di potere tra le discipline, in cui sembrerebbe avere la peggio la formazione critica di tipo umanistico. Come è suggerito da Wendy Hui Kyong Chun sarebbe possibile invertire la coppia *bright/dark* a proposito delle DH:

*I want to propose that the dark side of the digital humanities is its bright side, its alleged promise—its alleged promise to save the humanities by making them and their graduates relevant, by giving their graduates technical skills that will allow them to thrive in a difficult and precarious job market.*⁶¹

⁵⁸ Rockwell & Sinclair (2016, 21).

⁵⁹ Allington *et al.* (2016).

⁶⁰ Cfr. Spahr *et al.* (2016), Brennan (2017) e Bond *et al.* (2017), solo per citare alcuni degli articoli di risposta.

⁶¹ Chun (2014, 2).

La sua critica non riguarda tutti i progetti di DH, molti dei quali lei ammira e ritiene che possano dare un contributo di qualità ai metodi usati per analizzare i big data perché a suo avviso:

[...] the humanities should play a big role in big data, not simply because we're good at pattern recognition (because we can read narratives embedded in data) but also, and more importantly, because we can see what big data ignores.⁶²

La posizione di Chun, che condivido pienamente, riguarda il generale entusiasmo intorno alla tecnologia e alla sua centralità nell'educazione:

I also want to stress that my sympathetic critique is not aimed at the humanities, but at the general euphoria surrounding technology and education. That is, it takes aim at the larger project of rewriting political and pedagogical problems into technological ones, into problems that technology can fix.⁶³

La studiosa è contraria alla tesi che la tecnologia possa riformulare e risolvere tutti i problemi politici e pedagogici, come se una volta trasformati in problemi tecnici la soluzione fosse lì a portata di mano rappresentata dalla semplice riformulazione come questione neutrale di natura tecnica. La sua critica può essere accostata a quella al tecno-soluzionismo portata al successo mediatico dal libro di Morozov, dal titolo: *To save everything click here*⁶⁴.

Anche nelle *humanities* potrebbe darsi che applicare le tecniche non porti a risolvere i problemi o che debba essere precisato quale sia l'apporto tecnico e quale quello soggettivo interpretativo dello studioso, che propone il quesito e definisce il metodo con il quale si può rispondere alla domanda critica. La domanda inoltre deve sempre essere nuova rispetto a quelle del passato e non essere un metodo nuovo per accreditare risposte già fornite. Sul tema della novità anche Moretti avanza qualche dubbio.

[...] because the digital humanities have presented themselves as a radical break with the past, and must therefore produce evidence of such a break. And the evidence, let's be frank, is not strong. What is there, moreover, comes in a variety of forms, beginning with the slightly paradoxical fact that, in a new approach, not everything has to be new.⁶⁵

⁶² *Ibidem.*

⁶³ Chun (2014, 3).

⁶⁴ Morozov (2013).

⁶⁵ Moretti (2016, 4-5).

La sua idea finale è che il metodo non porti necessariamente a delle novità, che si potrebbero definire come scoperte realizzate grazie alle tecniche usate. Eppure, secondo Moretti, resta la convinzione che scoprire la stessa cosa della critica tradizionale, usando un metodo diverso arricchisca comunque la capacità di analisi, che per lui resta empirica e insieme sociologica. E tuttavia, anche lui si rammarica che per il momento la critica computazionale non abbia prodotto innovazione nell'ambito dell'approfondimento critico sebbene ci sarebbero grandi opportunità di usare l'archivio digitale per approfondimenti del materialismo storico. Per ora la sua delusione a questo proposito è palpabile: «it would be nice if, one day, big data could lead us back to big questions»⁶⁶.

L'uso di questi metodi non porta a novità sul piano delle grandi questioni, dei grandi problemi che hanno articolato la critica letteraria del secolo scorso. Ma non c'è da stupirsi. Senza la formulazione di domande critiche importanti è difficile che possano supplire i metodi, per quanto tecnologici e innovativi questi possano essere.

6. Alcune riflessioni su rischi e antidoti dei metodi *computer-assisted* per l'interpretazione letteraria

In un recente articolo, Manovich⁶⁷, l'inventore della *cultural analytics* – la disciplina che studia i fenomeni culturali sul piano della loro rappresentabilità in forma digitale e della loro misurabilità secondo determinate caratteristiche – ha notato come ci siano alcune difficoltà a valutare i fenomeni culturali attraverso i metodi analitici inventati da questa disciplina.

Innanzitutto la realtà delle esperienze culturali costituisce spesso dei gruppi che non possono essere riconosciuti attraverso il semplice uso della statistica che tenderà a stabilire una media che non corrisponde a nessun *cluster* reale. Inoltre la disponibilità dei dati è molto eterogenea online. Manovich fa l'esempio delle banche dati che raccolgono immagini artistiche internazionali che sono spesso concentrate sulla produzione dell'occidente e non rendono conto di quella del resto del mondo. Anche nel caso si utilizzino le immagini di Instagram sulle città, si mostra come solo alcune aree delle città siano rappresentate, mentre altre siano completamente assenti. La popolazione online negli US è molto diversa da quella dei paesi africani, sia in termini di quantità di utenti sia in termini di interessi e stili di

⁶⁶ Moretti, (2016, 7).

⁶⁷ Manovich (2017).

produzione del contenuto. Si tratta solo di alcuni dei tanti esempi che mostrano che la datificazione mantiene al proprio interno alcune caratteristiche pregiudiziali e non corrisponde alla totalità degli oggetti, in questo caso delle immagini disponibili in assoluto.

La tesi di Manovich è che questo non infici i metodi della *cultural analytics*, ma imponga di tenere conto che le nostre banche dati non sono costruite in modo completo o democratico. Inoltre suggerisce che lo scopo della *cultural analytics* non sia ricostruire dei modelli generali, ma singole caratteristiche notevoli dei fenomeni culturali:

Rather than only treating culture as ‘data points’ that together create patterns that we want to discover, disregarding the individual points afterwards, Cultural Analytics should pay equal attention to both patterns and individual artifacts, experiences and interactions. As creators and audience members, we engage and enjoy concrete artifacts and experiences, and not patterns. [...] One of the goals of Cultural Analytics is to help us find truly unique artifacts in the infinite universes of media now being created.⁶⁸

Il passaggio è interessante perché segnala un cambio di direzione degli obiettivi delle tecniche della *cultural analytics*, affini a quelle della critica computazionale. Segnala, cioè, la necessità di riconoscere delle caratteristiche uniche negli artefatti culturali, che li rendono veramente speciali e che li fanno considerare, tra gli altri esempi di fenomeni affini, come quelli rilevanti.

Come sostiene Gillespie⁶⁹, quando usiamo degli algoritmi per interpretare i dati abbiamo bisogno che i dati siano già preparati per essere trattati attraverso il software. I dati devono essere normalizzati e ripuliti di eventuali errori o incongruenze. La manipolazione dei dati per renderli adatti al processo tuttavia non viene, però, considerata come elemento epistemologico che richieda una discussione preventiva. A questo proposito Berry e Fagerjord⁷⁰ reclamano l’istituzione di una vigilanza critica su alcuni rischi che potrebbero spingere il discorso disciplinare nelle braccia del potere e di una visione neoliberale della conoscenza umanistica.

Il primo rischio è pensare alla cultura in forma di dati, ipotesi considerata un vero e proprio anatema per le *humanities*. Un altro rischio sarebbe considerare l’informazione come una fonte di profitto per il capitalismo, purché che venga adeguatamente catturata. Un ultimo rischio potrebbe essere quello, suggerito da Rockwell e Sinclair di affidarsi

⁶⁸ Manovich (2017, 275).

⁶⁹ Gillespie (2014).

⁷⁰ Berry & Fagerjord (2017).

completamente al progresso attraverso le tecniche senza metterne in discussione gli aspetti critici⁷¹. Questo rischio è secondo loro inevitabile e attiene a qualsiasi attività di impegno verso il mondo che non abbia carattere cinico.

Per contrastare questi rischi possono essere messi in campo alcuni argomenti che costituiscono un potenziale antidoto e dei suggerimenti per vigilare sulle derive pericolose della quantificazione miope nelle discipline di natura critica come le *humanities*.

Il primo antidoto è una critica della ragione informatica⁷². Se riconosciamo il carattere vincolato e storicamente limitato della nascita e dello sviluppo del computer e delle tecniche di programmazione, non possiamo non mantenere uno sguardo critico sulle tecniche che senza attenzione epistemologica si propongono di applicare il software a ogni problema⁷³. Calude e Longo⁷⁴ hanno anche ottenuto un risultato tecnico a proposito delle correlazioni spurie. Più aumenta la base dati, maggiori saranno le evidenze di correlazioni che sono del tutto casuali e non hanno ragione nella reciprocità di comportamento delle serie di dati delle variabili correlate. Nell'*abstract* dell'articolo si legge:

Using classical results from ergodic theory, Ramsey theory and algorithmic information theory, we show that this 'philosophy'[Big Data philosophy against science] is wrong. For example, we prove that very large databases have to contain arbitrary correlations. These correlations appear only due to the size, not the nature, of data. They can be found in 'randomly' generated, large enough databases, which—as we will prove—implies that most correlations are spurious. Too much information tends to behave like very little information.⁷⁵

Non ha senso perseguire l'uso di questi metodi che cercano delle correlazioni, le quali hanno solo un carattere arbitrario e non sono sintomo di nessuna regolarità significativa, a meno di spiegare esplicitamente le ragioni per cui quelle correlazioni trovate sono dotate di senso.

L'altro antidoto, proposto da Rockwell e Sinclair⁷⁶, riguarda i criteri che si usano per teorizzare adottando gli strumenti software nell'ambito delle discipline umanistiche e per creare prototipi destinati a testare quelle teorie, come Voyant – una *suite* di programmi concepiti per eseguire attività

⁷¹ Rockwell & Sinclair (2016, 21).

⁷² Longo (2009).

⁷³ Cfr. Numerico (2017) e Numerico (2018).

⁷⁴ Calude & Longo (2017).

⁷⁵ Calude & Longo (2017, 595).

⁷⁶ Rockwell & Sinclair (2016).

ermeneutiche *computer-assisted*. Vorrei segnalare alcune delle caratteristiche che gli autori individuano per tali strumenti e che dovrebbero facilitare la messa in discussione dei metodi prescelti anche attraverso una comunicazione trasparente circa il loro funzionamento e la loro progettazione.

They are not impenetrable black boxes; rather they are open to inspection. Rather than hiding their workings, they display them in a way that allows the user to understand their principles of operation.

They are open to manipulation in service of exploration and understanding [...].

They are supplemented by other materials [...] that encourage reflection on their use [...].

They force users to configure them in ways that provoke reflection.

They fail in interesting ways, and they show their limits so that when they fail you know that they have failed [...].⁷⁷

Adottare antidoti teorici e pratici rispetto ai metodi quantitativi usati per definire le interpretazioni dei dati, potrebbe metterci al riparo dai rischi peggiori posti dall'uso inavvertito di questi metodi. Ai fini di mantenere alta la guardia epistemologica sull'efficacia e la legittimità dell'adozione superficiale di questi metodi, è utile seguire dei suggerimenti offerti dall'ultimo libro di Berry e Fagerjord⁷⁸ volti a contrastare le possibili derive della nuova fase delle *digital humanites*. La loro proposta consiste in un allargamento di campo volto a includere una discussione sull'epistemologia e i metodi delle discipline umanistiche, dei loro valori e anche sulle competenze necessarie per affrontarne i problemi. Si tratterebbe a loro avviso di fornire un contributo alla "critica della cultura" e una nuova opportunità per affrontare questioni politiche ed etiche.

Tutte queste riflessioni ci suggeriscono che le *digital humanities* stanno attraversando una nuova proficua fase di assestamento e riorganizzazione sia del proprio statuto epistemologico sia dell'agenda delle prossime sfide da affrontare, sia dei potenziali nuovi oggetti della ricerca. Nel prossimo e ultimo paragrafo che funge anche da conclusione del discorso, pur mantenendo il carattere aperto di una riflessione sulle opportunità e i rischi che questa disciplina dovrà affrontare per essere finalmente matura, mi occuperò di fornire dei suggerimenti circa quali potrebbero essere le più proficue aree di intervento sulle quali esercitare le proprie competenze. La sfida è impellente e cruciale: studiare il ruolo della tecnologia nell'ambito

⁷⁷ Rockwell & Sinclair (2016, 166).

⁷⁸ Berry e Fagerjord (2017, p. 143).

delle scienze umane discutendo gli aspetti filosofici epistemologici e critici della nuova area di ricerca.

7. Un suggerimento per le *digital humanities*: una nuova fase filosofica

La proposta di una nuova fase delle *digital humanities* che sia più attenta alle questioni critiche, epistemologiche e politiche delle scelte disciplinari trae origine dalla ricerca di una definizione di cosa siano le discipline umanistiche in generale e come le *digital humanites* possano aumentarne la portata e ampliarne i metodi, senza mettere a rischio la natura storica di queste discipline, come conoscenze critiche e non strettamente procedurali. Feenberg⁷⁹ afferma che: «we reject the idea that there are appropriate techniques for everything, from forming friendship to enjoying Thanksgiving dinner». Ma se non ci sono procedure buone per ogni obiettivo, allora è necessario procedere alla valutazione critica dei metodi e soprattutto alle conseguenze politiche e all'analisi delle relazioni interdisciplinari che si instaurano a partire dalle scelte degli strumenti da adottare. Dobbiamo cioè applicare un approccio critico come quello suggerito da Berry a proposito di ciò che lui chiama il *computational turn* degli studi umanistici:

Computational techniques are not merely an instrument wielded by traditional methods; rather they have profound effects on all aspects of the disciplines. Not only do they introduce new methods, which tend to focus on the identification of novel patterns in the data as against the principle of narrative and understanding, they also allow the modularisation and recombination of disciplines within the university itself.⁸⁰

In questo paragrafo mi propongo di incrementare la consapevolezza nell'ambito delle discipline umanistiche, cercando di motivare, attraverso una riflessione teorica più ampia, la richiesta di un'investigazione sullo statuto epistemologico e quindi anche politico delle *digital humanities*. Al fine di spostare l'accento dall'entusiasmo per le nuove tecniche, ai loro aspetti critici ed epistemologici suggerisco di individuare alcune specifiche aree di intervento, e domande che ancora non hanno trovato una risposta definitiva.

Il primo tema sul quale sarebbe auspicabile soffermarsi è quello della lettura da lontano. Cosa significa decidere di "leggere" un testo da lontano e

⁷⁹ Feenberg (2009 p. 150).

⁸⁰ Berry (2011, p.13).

con l'aiuto di un dispositivo computazionale? Quali conseguenze produce sull'interpretazione di quel testo e sul senso dell'atto di interpretazione come il prodotto di una soggettività individuale su un testo unico, distinto da ogni altro. Farsi aiutare da un modello *data oriented* per categorizzare i testi e per le loro possibili interpretazioni ha degli effetti sull'attività ermeneutica che devono essere indagati e compresi a fondo, sebbene non si debbano demonizzare per principio nuove prospettive critiche e metodologiche sui testi. È necessario lavorare a una ricognizione critica dei nuovi processi in corso di realizzazione.

Un altro ambito di ricerche da promuovere consiste nell'applicare alle *digital humanities* la prospettiva degli *Science and Technologies Studies* (STS), come suggerisce anche Alan Liu. Se vogliamo spingere le *humanities* verso i metodi delle discipline più strettamente scientifiche, non possiamo ignorare il dibattito "contro il metodo" (per esempio le riflessioni di Feyerabend; Latour; Pickering) che esiste all'interno dell'epistemologia scientifica:

Any quest for stable method in understanding how knowledge is generated by human beings using machines founders on the initial fallacy that there are immaculately separate human and machinic orders, each with an ontological, epistemological, and pragmatic purity that allows it to be brought into a knowable methodological relation with the other— whether a relation of master and slave, cause and effect, agent and instrument, or another.⁸¹

Dobbiamo essere consapevoli come umanisti che non esiste nessuna "purezza dei metodi"⁸², perciò è necessario sceglierli con cautela: nella decisione, infatti, è insito anche il tipo di risultati che potremo/vorremo/saremo in condizione di ottenere. I metodi sono strumenti che contribuiscono a creare gli oggetti di ricerca e ne modificano per sempre lo *status* ontologico ed epistemologico⁸³.

Un altro ambito che ha bisogno di una riflessione profonda è quello relativo ai processi di archiviazione che sono quelli che sono stati più direttamente trasformati dall'uso del digitale. Innanzi tutto è importante riconoscere che nessun archivio potrebbe contenere tutto il materiale potenzialmente disponibile. Un archivio si definisce in base alle scelte di conservazione. Inoltre un sistema di archiviazione è caratterizzato dalla stabilità e continuità dei suoi documenti nel tempo. Gli archivi digitali si presentano come omnicomprensivi e non fanno scelte di conservazione,

⁸¹ Liu (2013, 416).

⁸² Cellucci (1998).

⁸³ Hacking (1983/1987).

eppure sono destinati a non garantire la continuità. I documenti conservati possono continuamente andare dispersi. La dinamica della conservazione dipende dalla durevolezza delle aziende che se ne fanno carico e dalla loro capacità di non disperdere i propri dati o per atti ostili o per malfunzionamento delle macchine e inoltre sono sempre in trasformazione. Secondo Lovink i *database* conservati online e contenenti materiale digitalizzato non costituiscono degli archivi, perché non sono protetti da un accesso diretto e perché hanno delle strategie di conservazione di lunga durata che non sono garantite dai server delle aziende Internet.

Una raccolta di archivio deve essere curata, mantenuta e preservata, ma questo non avviene nel caso degli archivi dei *Big Data*. Inoltre un archivio lavora contro il malfunzionamento e l'eventuale conservazione di informazioni false o non autentiche al proprio interno. Attività impossibile nei *repository* della rete⁸⁴. Possiamo ancora dire che l'archivio contenente tutto il "*fund of knowledge*" preconizzato da Licklider e poi realizzato dall'ampiezza della datificazione digitale online, sia ancora un archivio vero e proprio? Tutto questo dovrebbe essere oggetto della riflessione epistemologica e politica della nuova fase delle *digital humanities*.

È inoltre necessario rilevare il ruolo egemonico svolto dall'inglese nelle *digital humanities*. Tutti gli standard della disciplina provengono dalle ricerche in quella lingua e molto spesso gli studiosi sono costretti a usare la *koinè* inglese per poter partecipare al dibattito. Un esempio della evidente centralità dell'inglese nello sviluppo della disciplina si può riscontrare nell'uso dei tag dell'inglese nell'ambito della *Text Encoding Initiative* (TEI), il più importante progetto internazionale di rappresentazione della testualità in formato digitale⁸⁵. Tale stato di cose ha delle conseguenze culturali notevoli sulla progettazione degli strumenti e anche sullo stile epistemologico che vige nella disciplina, le cui conseguenze sarebbe importante indagare, magari usando le tecniche adottate nei *postcolonial studies*⁸⁶.

L'ultimo campo di ricerche sul quale impegnare le *critical digital humanities* è una riflessione genealogica sul ruolo e il significato della macchina digitale nella riorganizzazione della conoscenza, sia dal punto di vista epistemologico che politico. Qual è il ruolo svolto dal dispositivo nei processi di costruzione della conoscenza umanistica?⁸⁷

⁸⁴ Lovink (2016, cap. 11).

⁸⁵ Per maggiori informazioni sul funzionamento del consorzio internazionale della TEI consultare il sito web ufficiale <http://www.tei-c.org/>.

⁸⁶ Per uno studio approfondito del problema cfr. Fiormonte (2017).

⁸⁷ Cfr. Numerico (2018).

Si tratta quindi di un vasto programma di riorganizzazione disciplinare che richiede energie, risorse economiche, personali e una buona dose di tenacia e ostinazione per essere realizzato. Sarebbe importante investire in questa nuova ondata di ricerche critiche sulle *digital humanities* le competenze di studiosi europei, con speciale riguardo a quelli che abbiano una formazione filosofica ed epistemologica.

Bibliografia

- Allington, D., Brouillette, S., Golumbia, D. (2016), «Neoliberal Tools (and Archives): A Political History of Digital Humanities», *LARB*, May 2016. On-line:
<https://lareviewofbooks.org/article/neoliberal-tools-archives-political-history-digital-humanities/>
- Anderson, C. (2008), «The End of Theory: The Data Deluge That Makes the Scientific Method Obsolete», *Wired*, 23 July 2008. On-line:
http://www.wired.com/science/discoveries/magazine/16-07/pb_theory
- Berendt, B., Büchler, M., Rockwell, G. (2015), «Is it Research or is it Spying? Thinking-Through Ethics in Big Data AI and Other Knowledge Sciences», *Künstl Intell* (2015) 29: 223.
On-line:
<https://doi.org/10.1007/s13218-015-0355-2>
- Berry, D. M., Fagerjord, A. (2017), *Digital Humanities*, London, Polity Press.
- Berry, D. M. (2011), «The computational turn: thinking about the digital humanities», *Culture Machine*, vol. 12.
On-line:
<http://www.culturemachine.net/index.php/cm/article/view/440/470>
- Bond, S. E. *et al.* (2017), «‘Digital’ Is Not the Opposite of ‘Humanities’», in *The cronicle of higher education*, Nov. 2017.
- Borgman, C. (2007), *Scholarship in the Digital Age*, Cambridge (Mass.), MIT Press.

- Bowker, G. C., Baker, K., Millerand, F., Ribes, D. (2009), «Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment». In Hunsinger, J., Klastrup, L., Allen, M. (eds), *International Handbook of Internet Research*, Springer, Dordrecht.
- Brennan, T. (2017), «The Digital-Humanities Bust», in *The cronicle of higher education*, Oct. 2017.
- Calude, C. S., Longo, G. (2017), «The Deluge of Spurious Correlations in Big Data», in *Foundations of Science*, vol. 22 (3), pp.595-612.
- Cellucci, C. (1998), *Le ragioni della Logica*, Roma-Bari, Laterza.
- Cordeschi, R. (2002), *The discovery of the artificial*, Dordrecht, Springer Netherland.
- Drucker, J. (2011), «Humanities Approaches to Graphical Display», *Digital Humanities Quartely*, Vol. 5.1.
On-line:
<http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>
- Earhart, A. E. (2015), *Traces of the Old, Uses of the New: The Emergence of Digital Literary Studies*, University of Michigan Press, Ann Arbor, MI. On-line:
<https://quod.lib.umich.edu/cgi/t/text/textidx?cc=etlc;c=etlc;idno=13455322.0001.001;rgn=full%20text;view=toc;xc=1;g=dculture>
- Feenberg, A. (2009), «Critical theory of technology», in Kyrre, J. *et al.* (eds), *A companion to philosophy of technology*, Chichester, Wiley-Blackwell, pp. 146-153.
- Fiormonte, D. (2017), «Lingue, codici, rappresentanza. Margini delle Digital Humanities», in *Filologia digitale: problemi e prospettive. Accademia Nazionale dei Lincei*, Anno CDXIV - 2017. Contributi del Centro Linceo Interdisciplinare “Beniamino Segre”, Vol. 135, Roma, Bardi Edizioni, pp. 114-140.
- Fiormonte, D., Numerico, T., Tomasi, F. (2015), *The digital humanist. A critical inquiry*, New York, Punctum Books.

- Gillespie, T. (2014), «The relevance of algorithms», in Gillespie, T., Boczkowski, P., Foot, K. (eds), *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, MA., The MIT Press.
- Gitelman, L. (2013), *'Raw Data' is an oxymoron*, Cambridge, MA, MIT Press.
- Hacking, I. (1983), *Representing and intervening*, Cambridge Univ. Press, New York, (*Conoscere e sperimentare*, trad. it. di Prodi, R., Roma-Bari, Laterza, 1987).
- Heuser, R., Le-Khac, L. (2012), «A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method», *Stanford Literary Lab*.
On-line:
<http://litlab.stanford.edu/LiteraryLabPamphlet4>
- Keller Fox, E. (1991), *Conversazioni con Evelyn Fox Keller*, Milano, Elèuthera.
- Koyre, A. (2000), *Dal mondo del pressappoco all'universo della precisione*, Torino, Einaudi.
- Leonelli, S. (2018), *La ricerca scientifica nell'era dei big data*, Milano, Meltemi.
- Licklider, J. C. R. (1965), *Libraries of the future*, Cambridge, MA., The MIT Press.
- Liu, A., (2013) «The meaning of digital humanities», in *PMLA*, vol. 128, N. 2, March, pp. 409-423.
- Longo, G. (2009), «Critique of Computational Reason in the Natural Sciences», In Gelenbe, E., Kahane, J.-P. (eds), *Fundamental Concepts in Computer Science*, Imperial College Press, London pp. 43-70.
- Lovink, G. (2016), *Social Media Abyss*, London, Polity (*L'abisso dei social media*, trad. it. di Parrella B., Milano, Egea, 2017).

- Manovich, L. (2017), «Cultural Data: Possibilities and Limitations of the Digital Data Universe» in Grau, O., Coones, W., Rühse, V., *Museum and Archive on the Move. Changing Cultural Institutions in the Digital Era*, De Gruyter, Berlin, Boston, 259-276.
- Moretti, F. (2007), *Graphs, Maps, Trees: Abstract Models for Literary History*, London, Verso.
- Moretti, F. (2013), *Distant Reading*, London, Verso.
- Moretti, F. (2016), «Literature, Measured, Pamphlet 12», *Stanford Literary Lab*.
On-line:
<https://litlab.stanford.edu/LiteraryLabPamphlet12.pdf>.
- Moretti, F. (2016), «The Digital in the Humanities: An Interview with Franco Moretti», *Los Angeles Review of Books*, March 2016.
On-line:
<https://lareviewofbooks.org/article/the-digital-in-the-humanities-an-interview-with-franco-moretti/>
- Moretti, F., *et al.* (2017), *Canon/Archive: Studies in Quantitative Formalism*, New York, n+1.
- Morozov, E. (2013), *To save everything click here*, Perseus Publishing, New York, (*Internet non salverà il mondo*, trad. it. di Pannofino, G., Milano, Mondadori, 2014).
- Mumford, L. (1971), *The myth of the machine, the pentagon of power*, Secker and Warburg, London.
- Noble S.U. (2018), *The algorithms of oppression*, New York University Press, New York.
- Numerico, T. (2017) «La memoria e la rete» in Bertollini, A., Finelli, R., (a cura di) *Soglie del linguaggio. Corpo, mondi, società*, Roma, Roma Tre Press, pp. 81-102.
- Numerico, T. (2018), «La rivoluzione digitale», in Thermes, D., (a cura di), *Rivoluzione & Rivoluzioni*, Lanciano, Carabba “Agorà”, in corso di pubblicazione.

- O'Neil, C., 2016, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, St. Ives plc, Allen Lane, (*Armi di distruzione matematica. Come i big data aumentano la disuguaglianza e minacciano la democrazia*, trad. it. di Cavallini D., Milano, Bompiani, 2017).
- Pasquale, F. (2015), *Black box Society*, Cambridge MA. Harvard Univ. Press.
- Rosenblueth, A., Wiener, N. (1945), «The roles of models in science», in *Phil. Sci.*, vol. 12, pp. 316-322 (trad. it. «Il ruolo dei modelli nella scienza», in Somenzi, V., Cordeschi, R., (a cura di) *La filosofia degli automi*, Bollati Boringhieri, Torino, 1994, pp. 86-92).
- Roth, S., *et al.* (2017), «Futures of a distributed memory. A global brain wave measurement (1800–2000)», in *Technological Forecasting and Social Change*, Vol. 11, May 2017, pp. 307-323.
- Rouvroy, A., Berns, T. (2013), «Gouvernementalité algorithmique et perspectives d'émancipation: le disparate comme condition d'individuation par la relation», *Réseaux*, 2013/1 n° 177, p. 163-196.
On-line:
<http://www.cairn.info/revue-reseaux-2013-1-page-163.htm>.
- Spahr, J., *et al.* (2016), «Beyond Resistance: Towards a Future History of Digital Humanities», *Los Angeles Review of Books*, May 2016.
On-line:
<https://lareviewofbooks.org/article/beyond-resistance-towards-future-history-digital-humanities/>.
- Stiegler, B. (2015), *La société automatique 1. L'avenir du travail*, Paris, Fayard.
- Stiegler, B. (2015a), *Platone digitale*, Milano, Mimesis.
- Underwood, T. (2017), «A Genealogy of Distant Reading», *Digital Humanities Quarterly*, 2017, vol 11, n. 2.
On-line:
<http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html#eahrt2015>

Wiener, N. (1948/1964), *Cybernetics: or control and communication in the animal and the machine*, Cambridge, MA., The Mit Press (*La Cibernetica*, trad. it. di Barosso, G., Milano, Mondadori, 1968).