



Università
di Catania

ASSOCIAZIONE per
l'INFORMATICA UMANISTICA
e la CULTURA DIGITALE

Consiglio Nazionale
delle Ricerche

ME.TE. DIGITALI

MEDITERRANEO IN RETE TRA TESTI E CONTESTI

ATTI DEL XIII CONVEGNO ANNUALE
AIUCD 2024



28 - 30 MAGGIO
MONASTERO DEI BENEDETTINI
PIZZA DANTE, 32 CATANIA

ISBN 978-88-942535-8-0



Copyright ©2024 AIUCD
Associazione per l'Informatica Umanistica e la Cultura Digitale



Il presente volume e tutti i contributi sono rilasciati sotto licenza Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). Ogni altro diritto rimane in capo ai singoli autori.
This volume and all contributions are released under the Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). All other rights retained by the legal owners.

A cura di: Di Silvestro Antonio; Spampinato Daria (2024). Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale AIUCD, Catania 28-30 maggio 2024, Università di Catania.

Editing: Denise Bruno; Christian D'Agata; Laura Mazzagufò; Francesca Prado; Eliana Vitale; Alessandro Zammataro.

Ultimo accesso agli URL in data 15 maggio 2024.

Si prega di notificare all'editore ogni omissione o errore si riscontri: [aiucd.segreteria \[at\] aiucd.org](mailto:aiucd.segreteria@aiucd.org)
Please notify the publisher of any omissions or errors found: [aiucd.segreteria \[at\] aiucd.org](mailto:aiucd.segreteria@aiucd.org)

Il programma della conferenza AIUCD 2024 è disponibile online <https://aiucd2024.unict.it/programma/>
The AIUCD 2024 Conference Program is available online <https://aiucd2024.unict.it/programma/>

I contributi pubblicati nel presente volume hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima mediante *double-blind peer review* sotto la responsabilità del Comitato di Programma di AIUCD 2024.

All the papers published in this volume have received favourable reviews by experts in the field of DH, through an anonymous double-blind peer review process under the responsibility of the AIUCD 2024 Program Committee.

Chair

Antonio Di Silvestro (Università di Catania)

Daria Spampinato (CNR Istituto di Scienze e Tecnologie della Cognizione)

Comitato di programma / Program committee

Emmanuela Carbé (Università di Siena)

Massimo Cultraro (CNR Istituto di Scienze del Patrimonio Culturale)

Christian D'Agata (Università di Catania)

Antonio Di Silvestro (Università di Catania)

Greta Franzini (Eurac Research)

Maurizio Lana (Università del Piemonte Orientale)

Cristina Marras (CNR Istituto del Lessico intellettuale europeo e Storia delle Idee)

Marco Mazzone (Università di Catania)

Ouafae Nahli (CNR Istituto di Linguistica Computazionale "Antonio Zampolli")

Marianna Nicolosi-Asmundo (Università di Catania)

Marina Paino (Università di Catania)

Giuseppe Palazzolo (Università di Catania)

Jonathan Prag (University of Oxford Merton College)

Daria Spampinato (CNR Istituto di Scienze e Tecnologie della Cognizione)

Rachele Sprugnoli (Università di Parma)

Francesco Stella (Università di Siena)

Segreteria scientifica / Scientific Secretariat

Liborio Barbarino (Università di Catania)

Denise Bruno (Università di Catania)

Giulia Cacciatore (Università di Catania)

Giuseppe Canzoneri (Università di Catania)

Elisa Conti (Università di Catania)

Milena Giuffrida (Università di Catania)

Miryam Grasso (Università di Catania)

Francesca Prado (Università di Catania)

Emilio M. Sanfilippo (CNR Istituto di Scienze e Tecnologie della Cognizione)

Eliana Vitale (Università di Catania)

Alessandro Zammataro (Università di Catania)

Comunicazione istituzionale: Claudia Cantale (Università di Catania) e Area Per la Comunicazione dell'Università di Catania (ACOM).

Institutional communication: Claudia Cantale (University of Catania) and the Area for Communication of the University of Catania (ACOM)

Supporto tecnico: Rosario Agrò, Area della Terza Missione dell'Università di Catania, per la consulenza e la progettazione grafica dei materiali informativi del convegno.

Technical support: Rosario Agrò, Third Mission Area of the University of Catania, for advice and graphic design of the conference information materials.

Enti organizzatori / Organisers

AIUCD; Università di Catania: Dipartimento di Scienze Umanistiche; CNR Istituto di Scienze e Tecnologie della Cognizione; CINUM: Centro di Informatica Umanistica dell'Università di Catania.

Supporter

CLARIN-IT; Neperia Group; Storage; programma Piaceri 2020-2022, Linea 1; Parmalat-Sole.

Chair di area/ Track chair

Le culture digitali nel Mediterraneo

Cristina Marras (CNR Istituto del Lessico intellettuale europeo e Storia delle Idee)

Paola Moscati (CNR Istituto di Scienze del Patrimonio Culturale)

Archivi ed edizioni digitali

Christian D'Agata (Università di Catania)

Greta Franzini (Eurac Research)

Analisi computazionale dei testi

Angelo Mario Del Grosso (CNR Istituto di Linguistica Computazionale "Antonio Zampolli")

Simone Reborra (Università di Verona)

Ontologie e Semantic Web

Marianna Nicolosi Asmundo (Università di Catania)

Francesca Tomasi (Università di Bologna)

Preservazione della memoria e del patrimonio digitale

Fabio Ciraci (Università del Salento)

Anna Maria Marras (Università di Torino)

Lista dei revisori /List of reviewers

Maristella Agosti (Università di Padova), **Stefano Allegrezza** (Università di Bologna), **Chiara Alzetta** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Liborio Barbarino** (Università di Catania), **Nicola Barbuti** (Università di Bari Aldo Moro), **Stefano Bazzaco** (Università di Verona), **Benedetta Bessi** (Università Ca' Foscari di Venezia), **Andrea Bolioli** (ricercatore indipendente), **Paolo Bonora** (Università di Bologna), **Federico Boschetti** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Dominique Brunato** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Flavia Bruni** (Università Gabriele D'Annunzio di Chieti-Pescara), **Marina Buzzoni** (Università Ca' Foscari di Venezia), **Alberto Campagnolo** (Université Catholique de Louvain/KULeuven), **Anna Cappellotto** (Università di Verona), **Emmanuela Carbé** (Università di Siena), **Vittore Casarosa** (CNR Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo" – Università di Pisa), **Fabio Ciotti** (Università di Roma "Tor Vergata"), **Fabio Ciraci** (Università del Salento), **Elisa Conti** (Università di Catania), **Salvatore Cristofaro** (CNR Istituto per il Lessico Intellettuale Europeo e Storia delle Idee), **Christian D'Agata** (Università di Catania), **Elisa D'Argenio** (HUN-REN Hungarian Research Centre for Linguistics), **Mauro De Bari** (Università di Bari Aldo Moro), **Riccardo Del Gratta** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Angelo Mario Del Grosso** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Saulo Delle Donne** (Università del Salento), **Giorgio Maria Di Nunzio** (Università di Padova), **Antonio Di Silvestro** (Università di Catania), **Filippo Diara** (Università di Torino), **Giulia Fabbris** (Università Ca' Foscari di Venezia), **Riccardo Fedriga** (Università di Bologna), **Franz Fischer** (Università Ca' Foscari di Venezia), **Greta Franzini** (Eurac Research), **Francesca Frontini** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Daniele Fusi** (Stuttgart University & VeDPH – Università Ca' Foscari di Venezia), **Carola Gatto** (Università del Salento), **Lucia Giagnolini** (Università di Bologna), **Emiliano Giovannetti** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Milena Giuffrida** (Università di Catania), **Edmondo Grassi** (Università San Raffaele di Roma), **Miryam Grasso** (Università di Catania), **Alessandro Iannella** (Università di Cagliari - Università di Pisa – Università di Torino), **Paola Italia** (Università di Bologna), **Maurizio Lana** (Università del Piemonte Orientale), **Pietro Maria Liuzzo** (Bibliotheca Hertziana), **Dominique Longrée** (Université de Liège), **Francesco Mambrini** (Università Cattolica del Sacro Cuore di Milano), **Tiziana Mancinelli** (Istituto Italiano di Studi Germanici), **Anna Maria Marras** (Università di Torino), **Cristina Marras** (CNR Istituto del Lessico intellettuale europeo e Storia delle Idee), **Federico Meschini** (Università della Toscana), **Alessio Miaschi** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Paolo Monella** (Università Sapienza di Roma), **Ouafae Nahli** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Marianna Nicolosi-Asmundo** (Università di Catania), **Giuseppe Palazzolo** (Università di Catania), **Valentina Pasqual** (Università di Bologna), **Gianluca Pavan** (Università di Roma "Tor Vergata"), **Giulia Pedonese** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Jonathan Prag** (University of Oxford Merton College), **Simone Reborra** (Università di Verona), **Giulia Renda** (Università di Bologna), **Roberto Rosselli Del Turco** (Università di Torino), **Enrica Salvatori** (Università di Pisa), **Emilio M. Sanfilippo** (CNR Istituto di Scienze e Tecnologie della Cognizione), **Eva Sassolini** (CNR Istituto di Linguistica Computazionale "Antonio Zampolli"), **Pietro Sichera** (CNR Istituto di Scienze e Tecnologie della Cognizione), **Daniele Silvi** (Università di Roma "Tor Vergata"), **Elena Spadini** (University of Basel), **Daria Spampinato** (CNR Istituto di Scienze e Tecnologie della Cognizione), **Linda Spinazzè** (Università Ca' Foscari di Venezia), **Rachele Sprugnoli** (Università di Parma), **Francesco Stella** (Università di Siena), **Cecilia Tamagnini** (Università di Bologna), **Timothy Tambassi** (Università Ca' Foscari di Venezia), **Francesca Tomasi** (Università di Bologna), **Marco Venuti** (Università di Catania), **Fabio Vitali** (Università di Bologna).

I simili si attraggono. La valutazione letteraria sulle piattaforme di digital social reading

Gabriele Vezzani¹, Simone Rebora², Massimo Salgaro³

¹Università di Verona, Italia - gabriele.vezzani@univr.it

²Università di Verona, Italia - simone.rebora@univr.it

³Università di Verona, Italia - massimo.salgaro@univr.it

ABSTRACT

La rivoluzione digitale che interessa una parte sempre crescente del campo letterario ha avuto fra i suoi effetti principali quello di attribuire al lettore comune un'inedita centralità. L'impossibilità di concepire i lettori come soggetti passivi, diretti da cause esterne, impone di superare i modelli della valutazione letteraria in cui il valore attribuibile a un'opera viene visto come il risultato della decisione di un ristretto gruppo di lettori esperti, suggerendo l'applicazione di criteri più inclusivi. In questo paper, proponiamo di considerare la valutazione letteraria come parte di un complesso sistema di relazioni tra singoli lettori e mostriamo, usando la teoria dei grafi, come le dinamiche interne a tale sistema possano portare all'emergere di criteri valutativi condivisi. Al fine di ottenere dati reali su cui basare la nostra analisi, ci si è focalizzati sulla piattaforma di digital social reading Goodreads.

PAROLE CHIAVE

Valutazione letteraria; digital social reading; teoria dei grafi; recensioni online.

1. INTRODUZIONE

Come è stato ormai notato da diversi studiosi [15, 18], nel corso degli ultimi anni il panorama letterario ha subito profonde trasformazioni dovute all'influenza delle tecnologie digitali. Per quanto riguarda la ricezione, e più in particolare la valutazione delle opere letterarie, il principale mutamento a cui si è assistito consiste nell'inedita centralità di cui gode ora il lettore comune.

Prima che essa sconfinasse nell'ecosistema digitale, la valutazione della letteratura era appannaggio quasi esclusivo di una ristretta cerchia di 'professionisti', come autori, critici letterari, ricercatori, e insegnanti. A questi soltanto era riservato il privilegio di diffondere le proprie opinioni riguardo a cosa dovesse considerarsi letteratura di qualità e cosa no, mentre i lettori 'inesperti' erano visti come una massa acritica ed omogenea, capace di esprimersi soltanto mediante l'acquisto delle opere prescritte dai critici stessi [20] o dagli ingranaggi dell'industria culturale [1]. Oggi, grazie a piattaforme come Goodreads o Wattpad, nonché alla crescente circolazione di materiale letterario su Instagram o Twitter (ora X), a questo sistema ne subentra uno estremamente più fluido, in cui ogni lettore ha la possibilità di esprimere il proprio giudizio e influenzare in maniera diretta il panorama letterario. Su queste piattaforme, l'attenzione e l'attribuzione di valore, non più dirette da organismi istituzionali, diventano le monete di uno scambio democratico ed orizzontale.

La digitalizzazione del campo letterario porta così al crollo di ogni ideale monolitico di qualità. La nuova libertà d'espressione di cui gode il lettore comune fa sì che il panorama della valutazione letteraria si popoli dei criteri più disparati, dimostrando quanto poco difendibile sia ormai l'idea di un unico canone, quale manifestazione di un ideale unico e trasversale di qualità. Appaiono ora evidenti tanto la pluralità dei canoni quanto il fatto che la validità di questi è sempre relativa ad una (o più) comunità di lettori ed è pertanto situata in un preciso contesto storico e, soprattutto, sociale. La rivoluzione digitale impone insomma di ripensare le teorie secondo cui veniva tradizionalmente compresa la valutazione delle opere letterarie, costruendo un modello capace di rendere conto del contributo di qualunque lettore indipendentemente dal suo inquadramento istituzionale, un modello che ponga al centro l'individuo, la sua interazione col libro e il contesto sociale all'interno del quale questa ha luogo.

Un simile modello dovrebbe, prima di tutto, essere in grado di dare ragione del fatto che, pur fondata sulle azioni e opinioni di singoli individui, la valutazione delle opere letterarie tenda ad organizzarsi secondo criteri condivisi in maniera più o meno omogenea da determinati gruppi di lettori. Già Bourdieu [5] notava come a diverse classi sociali corrispondessero diversi sistemi di preferenze. Seppure l'almeno apparente determinismo che nel pensiero del sociologo francese fa dipendere il gusto dalle dinamiche di classe sia stato oggetto di critica [12, 15], la presenza di criteri valutativi socialmente determinati rimane un dato difficilmente contestabile [16].

Il nostro modello dovrà quindi essere in grado di spiegare la presenza di criteri valutativi condivisi come fenomeno emergente, risultato sistemico e globale delle interazioni locali fra i singoli agenti che compongono il sistema all'interno

del quale vengono recepite le opere letterarie. Si può compiere un primo passo nell'affrontare tale problematica considerando che il nostro gusto, e pertanto anche il sistema delle nostre preferenze in ambito letterario, è fortemente legato alla nostra identità [9]. Ricerche provenienti dall'ambito della psicologia sociale [8,19] hanno mostrato come ciascuno di noi sia costantemente alla ricerca di feedback esterni in grado di confermare e legittimare i valori che, da noi interiorizzati, costituiscono la base della nostra identità. Anche il gusto può essere considerato come il risultato di un atto di interiorizzazione di valori, di natura non morale ma estetica [10]. Ciò ci aiuta a capire quanto importante sia per noi che altri condividano le nostre preferenze e idiosincrasie, così legittimando i valori che si manifestano tramite esse. Possiamo ipotizzare che la ricerca di tale legittimazione proceda, per i lettori, secondo due canali: 1) rafforzando i legami che li uniscono a soggetti che condividono i loro stessi gusti e allentando quelli con soggetti che non li condividono, contribuendo così alla formazione di comunità fondate su gusti condivisi; 2) lasciandosi influenzare, tanto nella scelta di nuovi libri da leggere che nella valutazione di libri già letti, dalle opinioni maggioritarie all'interno della comunità di appartenenza [2], contribuendo così a rafforzare l'omogeneità di quest'ultima.

Le piattaforme di *digital social reading* [17] offrono un perfetto laboratorio per testare simili teorie. In particolare, Goodreads, incorporando le funzionalità di un tradizionale *social network* e permettendo quindi ai propri utenti di stringere amicizie e seguirsi a vicenda, si presenta come il setting ideale per studiare la componente sociale della valutazione letteraria. Essendo l'intera piattaforma dedicata all'espressione dei propri gusti in fatto di letteratura, è possibile interpretare l'amicizia fra due utenti come una casistica del fenomeno generale secondo cui due individui che condividono gusti simili tendono a consolidare i legami che li uniscono. Sulla base di questo assunto, abbiamo formulato le seguenti ipotesi:

- Un network costruito seguendo i rapporti di amicizia fra gli utenti di Goodreads tenderà ad organizzarsi secondo una forte struttura modulare, rispecchiando quanto si è detto sopra riguardo l'emergere di comunità basate sulla condivisione di gusti simili (**H1**).
- All'interno di tali comunità, il gusto degli utenti sarà più omogeneo di quanto non lo sarebbe sull'intero network. In particolare, ipotizziamo che la probabilità che un utente A abbia letto gli stessi libri di un utente B sarà maggiore se A e B appartengono alla stessa comunità di quanto non lo sarebbe se fossero estratti casualmente dall'intero network (**H2a**). Inoltre, ci aspettiamo che, qualora appartengano alla stessa comunità, A e B tenderanno a valutare in maniera simile opere simili, ovvero che l'appartenenza a una data comunità abbia un effetto statisticamente significativo sulla valutazione delle opere (**H2b**).

2. METODOLOGIA

2.1. Dati

Al fine di ottenere un campione casuale di utenti, si è partiti dalla *sitemap* di Goodreads¹. Qui, in una serie di files xml costantemente aggiornata, si trovano elencati gli utenti della piattaforma, ordinati in base all'ultimo periodo che li ha visti attivi sulla stessa. Si sono così scaricati i link alle pagine di 263,822 utenti, dai quali si è estratto in maniera casuale un campione di 1000 unità. Con uno script Python², servendosi del pacchetto Selenium in congiunzione con la piattaforma Dockers, sono stati salvati i link alle pagine degli utenti che figuravano come amici di ciascuno dei soggetti di partenza. La stessa operazione è stata svolta altre due volte, ogni volta partendo dal gruppo di amici creato all'iterazione precedente.

I dati ricavati mediante la procedura appena descritta sono stati organizzati nella forma di un network, nel quale ogni nodo rappresenta un utente, connesso tramite archi non-direzionali ai nodi corrispondenti alle sue amicizie. Complessivamente, il network è formato da 20,492 nodi e 23,379 archi. Per avere la possibilità di replicare eventuali risultati, abbiamo ripetuto la medesima procedura partendo da un diverso set di mille utenti (sempre estratti casualmente dagli originari 263,822, escludendo quelli selezionati nel primo round) e attivando lo script Python quattro volte invece di tre. In questo caso, si è ottenuto un network composto da 59,880 nodi e 69,750 archi.

In entrambi i casi, la distribuzione dei gradi dei nodi (il numero di archi da essi posseduti) si è rivelata fortemente asimmetrica, con una grande porzione del network (~90%) costituita da nodi aventi una sola connessione³. Dal momento che i legami fra lettori (qui operazionalizzati come amicizie online) rappresentano un elemento centrale della nostra

¹ <https://www.goodreads.com/siteindex.user.xml>

² Tutti gli script utilizzati per questo articolo sono disponibili al link <https://github.com/GVezzani/AIUCD24.git>

³ Con ogni probabilità, ciò è dovuto prevalentemente al modo in cui si sono costruiti i grafi, scaricando, ad ogni round di scraping, le amicizie degli utenti le cui informazioni erano state scaricate nel corso del round precedente. Non potendo continuare il processo fino ad esaurire tutti gli utenti della piattaforma, necessariamente ci si è dovuti ad un certo punto arrestare. Così, non si sono scaricati i dati relativi alle amicizie degli utenti dell'ultimo round, i quali quindi figurano, nella maggior parte dei casi, come aventi un solo amico (l'utente del round precedente partendo dal quale si è arrivati a loro).

domanda di ricerca, si è deciso di concentrare ulteriori indagini soltanto su una porzione dei due grafi, escludendo da essi i nodi con troppe poche connessioni. Dopo aver tentato diverse configurazioni (vd. il grafico nella Fig. 1), è risultato che

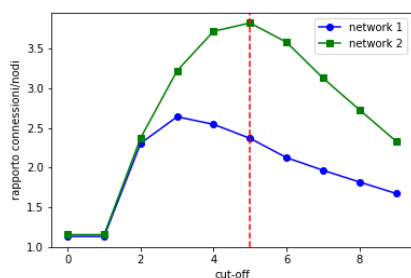


Figura 1. Numero di archi per nodo a diversi valori del cut-off.

il valore ottimale per tale soglia fosse 3 per il primo network e 5 per il secondo. Per mantenere coerenza fra le due analisi e massimizzare il numero di connessioni all'interno dei network, si è scelto di escludere da entrambi i network i nodi di grado inferiore a 5. I grafi così ottenuti sono composti, rispettivamente, da 625 nodi e 1622 archi e 708 nodi e 2706 archi.

Infine, con un secondo script Python, si sono scaricati i dati relativi allo storico delle letture di ciascun utente in entrambi i network. In particolare, sono stati salvati i titoli dei libri letti e, quando disponibile, la valutazione ad essi assegnata (da 1 a 5). In totale, gli utenti nel primo grafo hanno letto 38,538 libri diversi, quelli nel secondo 61,321.

Purtroppo, pochi di questi libri sono stati valutati da un numero di utenti tale da rendere possibile (o comunque significativa) un'analisi statistica. Per fare un esempio, un modello di regressione a effetti misti (si veda sotto per maggiori dettagli) richiede che si abbiano almeno cinque osservazioni per ogni livello di ciascuno degli effetti random inclusi. I libri che soddisfano tale condizione, e che quindi avrebbero potuto essere analizzati con tale metodo, non sono che il 3% del totale. Per ovviare a questa impasse, si è deciso di aggregare le opere per genere letterario.

Le informazioni relative al genere di un'opera sono presenti su Goodreads sotto forma di tags generati dagli utenti stessi. Nel corso dello scraping di questi dati si è incorsi in un problema affine a quello appena esposto: per la maggior parte delle opere nel nostro corpus nessun utente aveva generato alcun tag. Ciò ha reso per noi impossibile determinare il genere di 84% dei libri nel nostro corpus. Come è evidente dalla statistica appena citata, nonostante le sue limitazioni, questo approccio ci avrebbe permesso di lavorare con più opere di quante ne avremmo avute a disposizione basando le nostre analisi sui singoli libri, e si è scelto pertanto di aggregarle per genere in fase di analisi, escludendo quelle prive di genere.

2.2. Analisi

Al fine di testare l'ipotesi **H1**, entrambi i grafi sono stati analizzati mediante l'algoritmo di Louvain [4]. In breve, questo metodo è usato per trovare le comunità in cui si suddivide un network, ovvero gruppi di nodi all'interno dei quali si ha una densità di connessioni maggiore rispetto a quella che si otterrebbe selezionando casualmente un insieme di nodi della stessa numerosità. Il coefficiente di modularità, che per un grafo non pesato e non direzionale come i nostri può variare da un minimo di -0.5 a un massimo di 1 [7], misura la forza della struttura modulare del grafo.

Per testare l'ipotesi **H2a**, per ciascun utente di ciascuno dei due network si è poi calcolato:

- La probabilità di aver letto almeno un libro in comune con un altro utente all'interno della stessa comunità (gruppo 1) e la probabilità di aver letto almeno un libro in comune con un altro utente casuale (gruppo 2).
- Il numero medio di letture in comune con utenti appartenenti alla stessa comunità (gruppo 1) e il numero medio di letture in comune con utenti dell'intero network (gruppo 2).

L'influenza che l'appartenenza a una data comunità esercita sulle valutazioni delle opere letterarie (**H2b**) è stata verificata sia globalmente, sull'interrezza di entrambi i grafi, che concentrandosi su esempi rappresentati da specifici generi letterari. In primo luogo, si è creato un dataset delle valutazioni, elencando per ciascuna l'identificativo dell'utente e della comunità d'appartenenza, il genere letterario del libro valutato e la valutazione stessa (da 1 a 5). Servendosi del pacchetto lme4 di R [3], si è costruito un modello di regressione lineare a effetti misti, al quale si è richiesto di predire le valutazioni degli utenti basandosi su un'intercetta e sugli effetti random rappresentati da: i singoli lettori, il genere dei libri valutati e, per gli stessi generi, le comunità di appartenenza degli utenti (effetto random innestato nel precedente). In parole povere, l'intuizione alla base del modello è che: a) diversi lettori valutano secondo criteri diversi, b) generi diversi vengono valutati diversamente e c) utenti delle stesse comunità tendono a valutare in modo simile libri appartenenti a un medesimo genere letterario. Naturalmente, il terzo punto è dove risiede il nostro interesse. Per valutarne la significatività statistica si è usata la funzione 'anova' di R, confrontando le performance del nostro modello con quelle di un secondo, identico al precedente se non per l'esclusione dell'effetto di nostro interesse. L'intento, qui, consisteva nel verificare se l'aggiunta della variabile relativa alle comunità di appartenenza degli utenti comportasse un miglioramento significativo rispetto alla performance un modello che includesse come effetti random soltanto quelli relativi agli utenti e ai generi letterari dei libri valutati.

Infine, per ciascuno dei generi letterari nel dataset è stato eseguito un test ANOVA per verificare se ci fossero delle differenze statisticamente significative fra le valutazioni di utenti appartenenti a diverse comunità.

3. RISULTATI

Per quanto riguarda l'ipotesi **H1**, l'impiego dell'algorithmo di Louvain ha permesso di dividere con successo entrambi i grafi in comunità di utenti. Nel primo grafo, sono state individuate 70 comunità, per un coefficiente di modularità pari a 0.8. Nel secondo, si sono trovate 28 comunità, con un coefficiente di 0.61.



Figura 2. Porzione del secondo grafo con comunità. La visualizzazione è stata realizzata col software Gephi, usando l'algorithmo ForceAtlas 2

Simili risultati sono stati ottenuti anche per l'ipotesi **H2a**. Il primo test di Mann-Whitney sul primo grafo ha dato risultati statisticamente significativi ($u = 273,279$, $p < 0.001$, $r_{pbi} = 0.3$), dimostrando che la probabilità di aver letto almeno un libro in comune è maggiore di 0.11 per utenti appartenenti alla stessa comunità di quanto non lo sia per utenti casuali. Il secondo test di Mann-Whitney ha dimostrato ($u = 262,937$, $p < 0.001$, $r_{pbi} = 0.14$) che due utenti appartenenti allo stesso gruppo hanno letto, in media, 0.5 libri in comune in più rispetto a due utenti casuali. Come si può notare nella Tabella 1, tali risultati sono stati confermati anche per il secondo grafo.

In riferimento all'ipotesi **H2b**, il modello lineare a effetti misti è stato adattato nel primo grafo utilizzando *Restricted Maximum Likelihood estimation* (REML) e il criterio di convergenza è stato raggiunto con un valore di 17,306. Per quanto riguarda gli effetti di nostro interesse, la varianza nelle valutazioni dovuta al genere del libro valutato equivale a 0.009 (std. dev. = 0.09), mentre

quella relativa alla comunità ('genere:comunità', in quanto siamo interessati ai voti di utenti appartenenti a comunità diverse su libri dello stesso genere) a 0.02 (st. dev. = 0.14). Il confronto con un modello dal quale era stato escluso il secondo degli effetti menzionati, ha dimostrato che la sua aggiunta comporta un miglioramento di performance statisticamente significativo ($X^2 = 24.03$, $p < 0.01$).

Per alcuni generi, il test ANOVA ha dato risultati statisticamente significativi. È il caso, per fare soltanto due esempi, dei generi 'Historical Fiction' ($F(45, 288) = 1.7$, $p < 0.01$) e 'Romance' ($F(41, 1092) = 4.0$, $p < 0.01$). Tuttavia, per ragioni che verranno discusse in seguito, nel 74% dei casi è stato impossibile raggiungere la soglia di significatività.

Anche in questo caso (cfr. Tabella 1) la replicazione sul secondo grafo ha confermato i risultati ottenuti.

Test	Risultati
Mann-Whitney 1	$u = 333,747$, $p < 0.001$, $r_{pbi} = 0.38$, differenza fra medie = 0.13
Mann-Whitney 2	$u = 293,454$, $p < 0.001$, $r_{pbi} = 0.07$, differenza fra medie = 0.35
Modello a effetti misti	'genere: varianza = 0.02(st. dev. = 0.15), 'genere:comunità': varianza = 0.005 (st. dev. = 0.07)
Confronto fra modelli	$X^2 = 4.72$, $p = 0.02$
ANOVA	Risultati statisticamente significativi nel 24% dei casi

Tabella 1. Risultati dei test eseguiti sul secondo network.

4. DISCUSSIONE

I risultati dell'applicazione dell'algorithmo di Louvain a entrambi i network esaminati confermano l'ipotesi **H1**, mostrando come la rete di interazioni fra gli utenti di Goodreads si organizzi per strutture modulari, o per comunità. È bene notare che, dal momento che Goodreads non fornisce dettagliate informazioni demografiche relative ai propri utenti, non è possibile escludere categoricamente l'ipotesi che i gruppi trovati siano il riflesso di fattori esterni, come ad esempio il livello di istruzione o la classe sociale dei soggetti. Proprio perché tali informazioni non sono accessibili sulla piattaforma, tuttavia, possiamo presumere che esse non svolgano un ruolo preponderante nel determinare i comportamenti degli utenti (fra cui anche la scelta di stringere amicizie), e che questi siano basati prevalentemente sui dati che effettivamente circolano sul sito, ovvero dati relativi alla valutazione di opere letterarie. Tale ipotesi è in linea col secondo dei risultati della nostra immagine, ovvero l'aver mostrato come la struttura del network di relazioni sociali su Goodreads rispecchi differenze di gusti e criteri valutativi.

Ciò è evidente, in primo luogo, nello storico delle letture degli utenti. Come giustamente notano Heydebrand e Winko [13], la scelta di un libro da leggere può essere considerata come un atto valutativo a tutti gli effetti. A questo proposito, i risultati di entrambi i test di Mann-Whitney dimostrano, confermando l'ipotesi **H2a**, che le scelte di due utenti appartenenti alla stessa comunità sono più coerenti di quanto non lo siano quelle di due utenti casuali. In secondo luogo, i risultati del modello di regressione lineare a effetti misti dimostrano, confermando l'ipotesi **H2b**, che gli individui appartenenti alla stessa comunità tendono a condividere simili criteri valutativi, e quindi a valutare opere appartenenti allo stesso genere in maniera più coerente di quanto non lo farebbero due utenti casuali.

Tuttavia, i risultati dei test ANOVA impongono cautela nell'interpretazione. Il fatto che sia stato possibile trovare risultati statisticamente significativi soltanto per una porzione dei generi presi in considerazione riflette, infatti, la complessità dell'effetto in esame. È innanzitutto possibile che, nel caso di molti generi, semplicemente non si sia raggiunto un numero di lettori per comunità abbastanza elevato da garantire la significatività statistica dei risultati. Due potenziali fonti di variabilità devono essere considerate. Per primi, i generi stessi. Per fare soltanto un esempio, una comunità di utenti che condividono la passione per la science fiction sarà molto più omogenea nelle sue valutazioni di generi affini (distopia, fantasy...) e lascerà ai suoi membri molta più 'libertà' nel valutare opere di generi più distanti (poliziesco, romanzo storico...). La stessa variabilità si può ipotizzare, poi, in relazione agli utenti. Bisogna considerare infatti che, nella realtà, gli individui possiedono gusti poliedrici e non possono essere pensati come appartenenti a un singolo gruppo di lettori. Questo fa sì che, nelle comunità da noi individuate, siano presenti membri più o meno 'eretici', più o meno in linea con gli standard del gruppo.

A seguito di simili considerazioni, potremmo aspettarci che ogni genere, raggiunta una numerosità campionaria tale da controllare per i fattori di variabilità appena esposti, presenti differenze significative nelle valutazioni attribuitegli da lettori appartenenti a diverse comunità. Ciò, tuttavia, non avviene. Nella Figura 3, vediamo a confronto i diagrammi a scatola delle valutazioni relative ai generi 'LGBT' e 'Spirituality'.

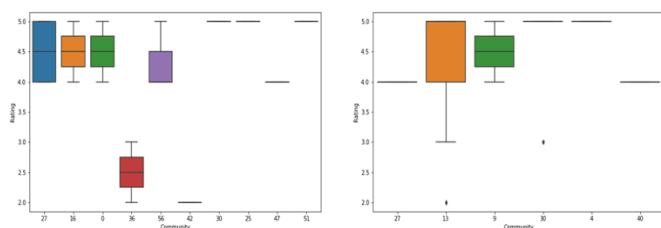


Figura 3. Confronto fra il diagramma a scatola relativo alle valutazioni per i generi 'LGBT' (a sinistra) e 'Spirituality' (a destra)

numero maggiore (19) di quelli che hanno fatto lo stesso per libri afferenti al primo (16). Tuttavia, il genere 'LGBT' permette di rilevare una differenza statisticamente significativa fra le valutazioni provenienti da diverse comunità di utenti, riferita in particolare alla comunità numero 36, che sembra non apprezzare particolarmente questi libri. Al contrario, per quanto riguarda il genere 'Spirituality', tutte le valutazioni si assestano fra il 4 e il 5, dimostrando una maggiore concordia. È chiaro, insomma, che l'ambiguità dei risultati ottenuti non dipende solamente da questioni

legate alla numerosità campionaria. Piuttosto, sembra che essa sia dovuta alle comunità stesse e alle caratteristiche dei generi valutati. Alcuni di questi, infatti, si dimostrano particolarmente adatti a catturare le differenze di gusto fra determinate comunità, mentre altri non svolgono alcun ruolo significativo in tal senso.

Come interpretare un simile risultato? Vicini, in questo, alla lezione di Bourdieu [6], possiamo considerare il panorama artistico come un campo. In ogni sua fase, tale campo può essere immaginato come uno spazio le cui dimensioni corrispondono a determinate opere, determinati generi o determinati movimenti letterari (in base al livello di risoluzione che si sceglie per la propria analisi). Un certo gusto altro non sarebbe che una *posizione* occupata in tale spazio, o, meglio, una presa di posizione rispetto alle dimensioni (opere, generi, movimenti) sulle quali esso è definito. La multidimensionalità del campo letterario è ciò che permette che due gusti siano in parte sovrapposti, pur rimanendo distinti: nulla, infatti, impedisce a due persone (o a due gruppi) di avere simili pareri relativamente a un certo genere, ma differire fortemente nel valore che attribuiscono a un altro.

I nostri risultati dimostrano come la presenza di criteri valutativi condivisi non debba necessariamente (nonostante, talvolta, possa) essere ricondotta all'azione di agenti istituzionali o determinanti socioeconomiche. Essa può anche essere interpretata come il risultato della libera azione di singoli individui che, esplorando il campo letterario e confrontandosi con altri individui impegnati nella medesima attività, contribuiscono all'emergere di una omologia fra le posizioni da loro occupate nel campo stesso (gusti) e nella rete di rapporti sociali che li unisce gli uni agli altri.

I processi alla base di una tale omologia possono essere di due tipologie principali:

- 1) processi di selezione, in base ai quali gli individui tendono a stringere legami sociali con coloro che possiedono gusti simili ai loro (pensiamo al caso di un utente che manda una richiesta d'amicizia all'autore di una recensione appena letta e con la quale si trova particolarmente d'accordo);
- 2) processi di influenza, in base ai quali gli individui tendono ad omologare i propri gusti a quelli dei membri dei gruppi sociali a cui appartengono (banalmente, ciascun utente sarà maggiormente disposto ad accettare i suggerimenti o le suggestioni provenienti dalla sua cerchia di amici, piuttosto che da sconosciuti).

Diversi studiosi, nell'ambito della sociologia [11, 14], hanno cominciato ad esplorare come queste due tipologie di processi intervengano nella determinazione del gusto estetico. Tuttavia, tali ricerche ignorano il caso della letteratura, concentrandosi piuttosto su musica e televisione. Inoltre, manca ancora un interesse rivolto specificamente a forme di socialità digitale, come quelle considerate nel presente studio. Un possibile sviluppo futuro per questo lavoro potrebbe

consistere proprio nel determinare in che misura i fenomeni qui individuati debbano essere attribuiti a dinamiche di selezione e quanto, invece, a dinamiche di influenza.

BIBLIOGRAFIA

- [1] Adorno, Theodore, e Max Horkeimer. *Dialectic of enlightenment*. London: Blackwell Verso, 1997.
- [2] Asch, Solomon. «Effects of group pressure upon the modification and distortion of judgments." Groups, Leadership and Men». In *Human Relations*, 177-190. Lancaster: Carnegie Press, 1971.
- [3] Bates, Douglas, Martin Mächler, Bolker Ben, e Steve Walker. «Fitting Linear Mixed-Effects Models Using lme4». *J. Stat. Soft.* 67 (2015): 1-48.
- [4] Blondel, Vincent, Guillaume Jean-Loup, Renaud Lambiotte, e Etienne Lefebvre. «Fast unfolding of communities in large networks». *J. Stat. Mech.* 2008, 1-12.
- [5] Bourdieu, Pierre. *La Distinction: Critique sociale du jugement*. Paris: Minuit, 1979.
- [6] Bourdieu, Pierre. *Les règles de l'art: genèse et structure du champ littéraire, Nouv. éd., revue et corrigée*. Paris: Éd. du Seuil, 2010.
- [7] Brandes, Ulrik, Daniel Dellinger, Marco Gaertler, Robert Gorke, Martin Hoefer, Zoran Nikoloski, e Dorothe Wagner. «On Modularity Clustering». *IEEE Trans. Knowl* 20 (2008): 172-188.
- [8] Burke, Peter, e Donald Reitzes. «An Identity Theory Approach to Commitment». *Social Psychology Quarterly* 54 (1991): 239-251.
- [9] Dolby, Nadine. «The Shifting Ground of Race: The role of taste in youth's production of identities». *Race Ethnicity and Education* 3 (2000): 7-23.
- [10] Fingerhut, Joerg, Javier Gomez-Lavin, Claudia Winklmayr, e Jesse Prinz. «The Aesthetic Self. The Importance of Aesthetic Taste in Music and Art for Our Perceived Identity». *Front. Psychol* 11 (2021).
- [11] Friemel, Thomas. «Network dynamics of television use in school classes». *Social Networks* 34 (2012): 346-358.
- [12] Hennion, Antoine. «Those Things That Hold Us Together: Taste and Sociology». *Cultural Sociology* 1 (2007): 97-114.
- [13] Von Heydebrand, Renate, e Simone Winko. *Einführung in die Wertung von Literatur: Systematik - Geschichte - Legitimation*. München: Schöningh, 1996.
- [14] Lizardo, Oscar. «How Cultural Tastes Shape Personal Networks». *Am Sociol Rev* 71 (2006): 778-807.
- [15] Murray, Simone. *The digital literary sphere: reading, writing, and selling books in the Internet era*. Baltimore: Johns Hopkins University Press, 2018.
- [16] Prior, Nick. «Critique and Renewal in the Sociology of Music: Bourdieu and Beyond». *Cultural Sociology* 5 (2011): 121-138.
- [17] Rebora, Simone, Peter Boot, Federico Pianzola, Brigitte Gasser, J.Berenike Herrmann, Maria Kraxenberger, Moniek M. Kuijpers, et al. «Digital humanities and digital social reading». *Digital Scholarship in the Humanities* 36 (2021): ii230-ii250.
- [18] Salgaro, Massimo. «Literary value in the era of big data. Operationalizing critical distance in professional and non-professional reviews». *Journal of Cultural Analytics* 7 (2022).
- [19] Stets, Jan, Scott Savage, Peter Burke, e Phoenicia Fares. «Cognitive and Behavioral Responses to the Identity Verification Process». In *Identity and Symbolic Interaction*, (a cura di) Richard Serpe, Robin Stryker, e Brian Powell, 65-88. Springer International Publishing, 2020.
- [20] Van Rees, Kees. «How a literary work becomes a masterpiece: On the threefold selection practised by literary criticism». *Poetics* 12 (1983): 397-417.