

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

Anno 42

N. 1-2 – gennaio-giugno 2024

Editoriale

ROBERTO GUARASCI

*Quarant'anni di «AIDAinformazioni»
(1984-2024)*

Contributi su invito

PIERO INNOCENTI

*Que reste-t-il de nos amours? Arti
tradizionali di un possibile moderno Trivio:
Archivistica, Bibliografia, Documentazione*

Contributi

FRANCESCO AMATO, ANTONELLA
BENVENUTO, SILVIA CENITI, MARIA
TERESA CHIARAVALLI, CLAUDIA
LANZA, ERIKA PASCERI

*Indagine e analisi comparativa delle
procedure di codifica nella Terapia del
Dolore in Italia*

ANDREA CAPACCIONI

*Sull'affidabilità delle bibliografie generate
dai chatbot. Alcune considerazioni*

ALESSANDRO MAISTO

*La dimensione Testuale del Videogioco.
Classificazione dei transcript dei
videogiochi basata sul lessico*

ALEXANDER MURZAKU, PONTISH
YERAMYAN, CURT ANDERSON, STEVEN
BUXBAUM, RUBEN DIAZ, MARIELLE
LERNER, ARMENUI MINASYAN, HAZEL
MITCHLEY, JODIE-ANN PENNANT, MIA
SHANG, BRISA SPEIER BRITO

*Discovering and documenting brilliance.
A novel multimodal annotation method*

CAMILLA NAPPI

*Le document : miroir des enjeux
institutionnels et de l'évolution linguistique
des transitions écologique et énergétique en
France (2007-2022)*

Testimonianze

PIERO CAVALERI, FERRUCCIO DIOZZI

*La Documentazione in Italia e il ruolo di
Paolo Bisogno: una rapida evoluzione*

PAOLA GARGIULO, LUCIA MAFFEI

*Domenico (Ingo) Bogliolo. Profilo di un
navigatore nell'Information Science*

AIDAinformazioni Anno 42 – N. 1-2 – gennaio-giugno 2024

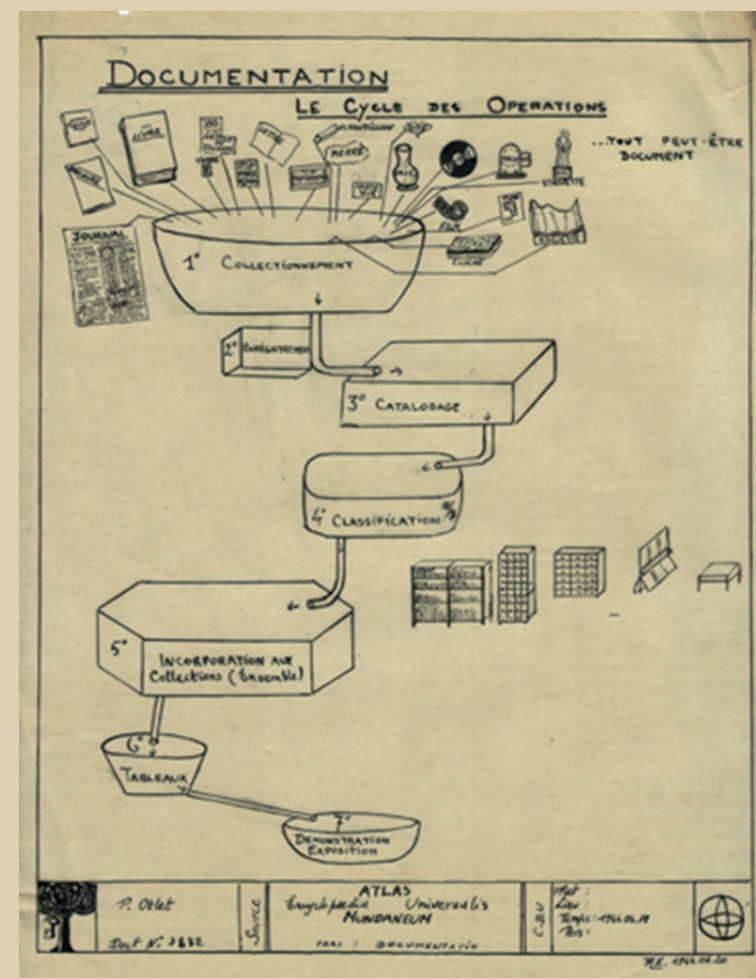
AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

NUMERO 1-2

ANNO 42

GENNAIO-GIUGNO 2024



mundaneum

In copertina
Disegno di Paul Olet, Collections Mundaneum, centre d'Archives, Mons (Belgique).

ISBN 979-12-5965-407-6

ISSN 1121-0095



CACUCCI  EDITORE
BARI

AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

Fondata nel 1983 da Paolo Bisogno

Proprietario della rivista:

Università della Calabria

Direttore Scientifico:

Roberto Guarasci, *Università della Calabria*

Direttore Responsabile:

Fabrizia Flavia Sernia

Comitato scientifico:

Anna Rovella, *Università della Calabria*;

Maria Guercio, *Sapienza Università di Roma*;

Giovanni Adamo, *Consiglio Nazionale delle Ricerche †*;

Claudio Gnoli, *Università degli Studi di Pavia*;

Ferruccio Diozzi, *Centro Italiano Ricerche Aerospaziali*;

Gino Roncaglia, *Università della Toscana*;

Laurence Favier, *Université Charles-de-Gaulle Lille 3*;

Madjid Ihadjadene, *Université Vincennes-Saint-Denis Paris 8*;

Maria Mirabelli, *Università della Calabria*;

Agustín Vivas Moreno, *Universidad de Extremadura*;

Douglas Tudhope, *University of South Wales*;

Christian Galinski, *International Information Centre for Terminology*;

Béatrice Daille, *Université de Nantes*;

Alexander Murzaku, *College of Saint Elizabeth, USA*;

Federico Valacchi, *Università di Macerata*.

Comitato di redazione:

Antonietta Folino, *Università della Calabria*;

Erika Pasceri, *Università della Calabria*;

Maria Taverniti, *Consiglio Nazionale delle Ricerche*;

Maria Teresa Chiaravalloti, *Consiglio Nazionale delle Ricerche*;

Assunta Caruso, *Università della Calabria*;

Claudia Lanza, *Università della Calabria*.

Segreteria di Redazione:

Valeria Rovella, *Università della Calabria*

Editrice: Cacucci Editore S.a.s.

Via D. Nicolai, 39 – 70122 Bari (BA)

www.cacuccieditore.it

e-mail: riviste@cacuccieditore.it

Telefono 080/5214220

AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

«AIDAinformazioni» è una rivista scientifica che pubblica articoli inerenti le Scienze dell'Informazione, la Documentazione, la Gestione Documentale e l'Organizzazione della Conoscenza. È stata fondata nel 1983 quale rivista ufficiale dell'Associazione Italiana di Documentazione Avanzata e nel febbraio 2014 è stata acquisita dal Laboratorio di Documentazione dell'Università della Calabria. La rivista si propone di promuovere studi interdisciplinari oltre che la cooperazione e il dialogo tra profili professionali aventi competenze diverse, ma interdipendenti. I contributi possono riguardare topics quali Documentazione, Scienze dell'informazione e della comunicazione, Scienze del testo e del documento, Organizzazione e Gestione della conoscenza, Terminologia, Statistica testuale e Linguistica computazionale e possono illustrare studi sperimentali in domini specialistici, casi di studio, aspetti e risultati metodologici conseguiti in attività di ricerca applicata, presentazioni dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Settore 11/A4 e censita per le Aree 10 – Scienze dell'antichità, filologico-letterarie e storico-artistiche; 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche; 12 – Scienze giuridiche; 14 – Scienze politiche e sociali, così come dall'ARES (Agence d'évaluation de la recherche et de l'enseignement supérieur) che la annovera tra le riviste scientifiche dell'ambito delle Scienze dell'Informazione e della Comunicazione. La rivista è, inoltre, indicizzata in: ACNP – Catalogo Italiano dei Periodici; BASE – Bielefeld Academic Search Engine; ERIH PLUS – European Reference Index for the Humanities and Social Sciences – EZB – Elektronische Zeitschriftenbibliothek – Universitätsbibliothek Regensburg; Gateway Bayern; KVK – Karlsruhe Virtual Catalog; Letteratura Professionale Italiana – Associazione Italiana Biblioteche; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Summon™ – by SerialsSolutions; Ulrich's; UniCat – Union Catalogue of Belgian Libraries; Union Catalog of Canada; LIBRIS – Union Catalogue of Swedish Libraries; Worldcat.

I contributi sono valutati seguendo il sistema del *double blind peer review*: gli articoli ricevuti dal comitato scientifico sono inviati in forma anonima a due referee, selezionati sulla base della loro comprovata esperienza nei topics specifici del contributo in valutazione.

AIDAinformazioni

Anno 42

N. 1-2 – gennaio-giugno 2024

CACUCCI  EDITORE
BARI

PROPRIETÀ LETTERARIA RISERVATA

© 2024 Cacucci Editore – Bari

Via Nicolai, 39 – 70122 Bari – Tel. 080/5214220

<http://www.cacuccieditore.it> e-mail: info@cacucci.it

Ai sensi della legge sui diritti d'Autore e del codice civile è vietata la riproduzione di questo libro o di parte di esso con qualsiasi mezzo, elettronico, meccanico, per mezzo di fotocopie, microfilms, registrazioni o altro, senza il consenso dell'autore e dell'editore.

Sommario

Editoriale

ROBERTO GUARASCI, Quarant'anni di «AIDAinformazioni» (1984-2024) 9

Contributo su invito

PIERO INNOCENTI, Que reste-t-il de nos amours? Arti tradizionali di un possibile moderno Trivio: Archivistica, Bibliografia, Documentazione 17

Contributi

FRANCESCO AMATO, ANTONELLA BENVENUTO, SILVIA CENITI, MARIA TERESA CHIARAVALLOTI, CLAUDIA LANZA, ERIKA PASCERI, Indagine e analisi comparativa delle procedure di codifica nella Terapia del Dolore in Italia 59

ANDREA CAPACCIONI, Sull'affidabilità delle bibliografie generate dai chatbot. Alcune considerazioni 81

ALESSANDRO MAISTO, La dimensione Testuale del Videogioco. Classificazione dei transcript dei videogiochi basata sul lessico 95

ALEXANDER MURZAKU, PONTISH YERAMYAN, CURT ANDERSON, STEVEN BUXBAUM, RUBEN DIAZ, MARIELLE LERNER, ARMENUI MINASYAN, HAZEL MITCHLEY, JODIE-ANN PENNANT, MIA SHANG, BRISA SPEIER BRITO, Discovering and documenting brilliance. A novel multimodal annotation method 117

CAMILLA NAPPI, Le document : miroir des enjeux institutionnels et de l'évolution linguistique des transitions écologique et énergétique en France (2007-2022) 145

Testimonianze

PIERO CAVALERI, FERRUCCIO DIOZZI, La Documentazione in Italia e il ruolo di Paolo Bisogno: una rapida evoluzione 163

PAOLA GARGIULO, LUCIA MAFFEI, Domenico (Ingo) Bogliolo. Profilo di un navigatore nell'Information Science 169

Editoriale

Quarant'anni di «AIDAinformazioni» (1984-2024)

Roberto Guarasci*

Le celebrazioni di anniversari, seppur di pubblicazioni scientifiche, oscillano quasi sempre tra la celebrazione del bel tempo andato e la prospezione del radioso sol dell'avvenir per usare uno storico paradigma ripreso in una recente pellicola cinematografica. In mezzo vivono una nutrita congerie di eventi, episodi, frammenti anche di rilevante valore emotivo dei personaggi e delle azioni che in quel lasso di tempo sono stati attori o spettatori di quanto accaduto. È come se cento, mille diari personali venissero fusi in un unico flusso ininterrotto di ricordi ai quali spesso si fatica a dare una struttura ed una organizzazione. Poi ci sono i contributi di quanti quel percorso non hanno condiviso o lo hanno condiviso in minima parte e lo reinterpretano secondo i loro schemi concettuali e le proprie credenze. La storia di una rivista scientifica è – però – la rivista stessa con tutti gli autori che nel tempo ne hanno permesso la vita e l'evoluzione tanto più quando la testata si pone come espressione di un dominio estremamente specialistico interessato, nei decenni, da così tante evoluzioni da cambiarne completamente la connotazione e le dimensioni mettendone finanche in dubbio la stessa necessità di sopravvivenza.

Il primo numero di «AIDAinformazioni» (1984) è poco più di un atto di nascita e la qualificazione come bollettino dell'omonima associazione dalla quale mutua il nome. Anche l'indicazione della periodicità risente dello spirito volontaristico dell'organismo che stava nascendo: «la periodicità del bollettino dipenderà in buona parte dalla quantità di notizie a disposizione: la redazione farà del suo meglio per acquisirne direttamente il maggior numero possibile»¹.

* Dipartimento di Culture, Educazione e Società (DiCES), Università della Calabria, Rende (CS), Italia. roberto.guarasci@unical.it.

¹ «A conferma del carattere inizialmente divulgativo di «AIDAinformazioni» è possibile citare l'esito della richiesta di un contributo che il 25 giugno 1994, con riferimento all'art. 25 della Legge 5 agosto 1981, n. 41674, l'AIDA [Associazione Italiana Documentazione Avanzata], per il tramite dell'allora Presidente Antonio Petrucci, rivolse al Ministero per i Beni Culturali e Ambientali, Divisione VIII Editoria – Ufficio Segreteria della Commissione per le pubblicazioni periodiche di elevato valore cultu-

È l'anno del lancio del Macintosh 128k di Apple che rivoluzionerà il mondo dei personal computer e della profezia orwelliana la cui tripartizione geografica – Oceania, Eurasia ed Estasia – sembra quasi echeggiare quella della visione “ternaria della rivista” ovvero oggetto, soggetto, servizi [della scienza della documentazione].

Qualche anno prima, nel 1980, Paolo Bisogno – direttore e fondatore della rivista – aveva pubblicato *Teoria della Documentazione* (Bisogno 1980) libro ostico e complesso ma che rappresenta il manifesto italiano della documentazione o almeno della sua elaborazione teorica. I decenni precedenti specie quelli contigui al secondo dopoguerra avevano conosciuto numerose realizzazioni pratiche – quasi tutte industriali – come testimoniano i centri di documentazione attivi sul territorio nazionale (*La Documentazione in Italia* 1952) con la significativa eccezione pubblica del Centro Nazionale di Documentazione del Consiglio Nazionale delle Ricerche (CNR) che costituirà in gran parte la radice dalla quale originerà l'Istituto di Studi e Ricerche sulla Documentazione Scientifica diretto dallo stesso Bisogno². Carente era stata però l'elaborazione concettuale vista anche la tiepida accoglienza se non l'aperta avversione che il nostro Paese aveva riservato alle teorie di Paul Otlet e dei suoi epigoni (Fumagalli 1895; Guarasci e Taverniti 2013).

Pur richiamandosi a più riprese ai grandi filoni europei delle scienze del documento – Otlet in primis – Bisogno estremizza le posizioni di Briet³ sulla sinergia profonda con le scienze dell'informazione e l'informatica aggiungendovi un'ampia parte di diffusione della cultura scientifica e politica della scienza. «AIDAinformazioni» è al tempo stesso un bollettino di divulgazione e una rivista scientifica che ospita contributi trasversali alle discipline che, nella visione di Bisogno, erano concorrenti alla gestione del documento, dalla linguistica “quantitativa” – oggi diremmo computazionale – alla retorica alla

rile. Nella richiesta Petrucci dichiarava tra le altre cose che il periodico, nel corso del 1993, non era stato né venduto né distribuito per abbonamento, ma era stato soltanto inviato ai soci e che, essendo l'organo di una Associazione professionale, era frutto della collaborazione volontaria dei soci e che pertanto non aveva personale specificatamente dedicato. A tale richiesta, il Ministero, in data 27 aprile 1995, rispose negativamente adducendo la seguente motivazione: “Rivista di aggiornamento e informazione, priva di impostazione scientifica” (Folino e Pasceri 2017, 62).

² Con provvedimento del Presidente del CNR n. 1556 del 17 febbraio 1966 il Centro Nazionale di Documentazione Scientifica e Tecnica e l'Ufficio Studi e Sviluppo vengono riuniti in un unico organismo denominato Servizio studi e rilevazioni articolato in due sezioni relative alle competenze degli organismi che lo avevano originato. La direzione del Servizio è affidata a Maurizio Giorgi, dell'Istituto idrografico della marina, mentre la sezione Rilevazioni è affidata a Maria Pia Carosella. Con successivo provvedimento n. 1993 del 2 dicembre 1968 il Servizio studi e rilevazioni viene soppresso. Nel frattempo il 15 luglio 1968 viene approvata la costituzione di un Laboratorio sulla Ricerca e sulla Documentazione del quale viene nominato reggente Paolo Bisogno. Il Laboratorio, nel febbraio 1976, diventerà Istituto di Studi sulla Ricerca e Documentazione Scientifica (ISRDS) (cfr. Guarasci 2011).

³ Renée Marie Hélène Suzanne Briet (Parigi, 1 febbraio 1894 – 13 febbraio 1989). Sull'opera della Briet si veda anche (Castellucci e Mori 2022).

logica. Il linguaggio è il collante tra le discipline perché «è il mezzo essenziale per ogni forma di comunicazione e, nell'ambito di questa, di documentazione e di informazione» (Bisogno 1980, 10). Gli indici dei primi anni di vita della rivista – fino al 1989 – riflettono questa connotazione di bollettino informativo dell'associazione con qualche contributo di più ampio respiro. Gli anni Novanta vedono un progressivo allargamento degli orizzonti con un focus crescente sulla definizione della figura professionale del documentalista e la concettualizzazione del termine documento come elemento fondante. Accanto ai tradizionali temi dell'automazione e dell'informazione dei servizi documentali e delle biblioteche cominciano a farsi strada contributi con una visione prospettica in linea con le tendenze del Paese. Sono gli anni di una crescente attenzione verso la digitalizzazione della Pubblica Amministrazione come risposta quasi taumaturgica alle consolidate inefficienze di alcuni settori dell'apparato dello Stato⁴ e sono anche gli anni dell'emanazione della legge 241/90 che, con la sua ridefinizione del concetto di documento amministrativo, apre la strada a tutta la normativa successiva sui documenti elettronici, ma sono anche gli anni di Tangentopoli⁵ e del cambiamento strutturale dell'assetto dello Stato e del rapporto tra i Poteri che non sarà ininfluente per la relazione dei cittadini con la memoria dei fatti. Nei titoli della rivista cominciano ad essere sempre più presenti – di fianco ai tradizionali temi dell'informazione e della documentazione – analisi e rassegne degli scenari e dei programmi europei con dei timidi tentativi di riprodurre in Italia figure professionali ed esperienze consolidate in paesi contermini. Forte è il legame con la Francia sia per i legami culturali che per la politica capillare di diffusione della documentazione e dell'utilizzo delle fonti documentali da loro messa in essere. Sul finire degli anni Novanta le leggi Bassanini aumentano l'autonomia funzionale delle università e la capacità di modellare i corsi di studio e i conseguenti insegnamenti e la rivista ospita un numero crescente di contributi sull'insegnamento della Documentazione e le conseguenti figure professionali in uscita ma è, ormai, una battaglia di retroguardia. Si è in gran parte consumata la dicotomia tra le scienze del testo, sempre più orientate al documento come patrimonio culturale e le scienze dell'informazione orientate al dato ed alla sua utilizzazione in una società che comincia a cogliere il valore crescente della conoscenza come vero valore competitivo e che è martellata quotidianamente dall'assioma carta uguale vecchio digitale uguale nuovo. Il numero 1 del 1999 si apre con il ricordo del fondatore e direttore Paolo Bisogno prematuramente scomparso il 6 febbraio dello stesso anno. Gli succede alla direzione della rivista Piero Innocenti⁶. I contributi degli anni successivi si spostano sempre più sulle scienze del libro e

⁴ Per una cronologia delle riforme della PA degli ultimi 25 anni si veda (ForumPA 2016).

⁵ Per una breve storia si veda (Diamanti 2008).

⁶ Direttore responsabile sarà fino alla cessione della rivista all'Università della Calabria Mario De Gregorio.

sulle applicazioni digitali a quel contesto che aveva – nonostante l'avversione iniziale – comunque culturalmente inglobato alcuni degli elementi costitutivi della documentazione olettiana come testimoniano i frequenti e continuativi momenti di collaborazione tra AIB (Associazione Italiana Biblioteche) e AIDA (Folino e Pasceri 2017, 76-77). Dall'*Information Literacy* alla letteratura grigia al *document delivery* e all'editoria elettronica la rivista segue l'evoluzione delle scienze del testo nel Paese ma, perdendo progressivamente il rapporto con la comunità di riferimento che diventa sempre più esile, anche i contributi palezano una crescente eterogeneità non sempre riconducibile ad un disegno unitario. Nel 2012 con il DM 76 l'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca (ANVUR) acquisisce la competenza della classificazione delle riviste⁷. «AIDAinformazioni» viene classificata come scientifica diminuendo ulteriormente la sua appetibilità come sede di pubblicazione. Nel frattempo l'Associazione si avvia ad una prevista chiusura e nel febbraio 2014 la rivista viene ceduta all'Università della Calabria che la detiene a tutt'oggi. Dal 2016 è rivista di classe A. Ma questo è il presente.

Riferimenti bibliografici

- ANVUR. n.d. “Riferimenti Normativi.” Consultato il 15 maggio 2024. <https://www.anvur.it/attivita/classificazione-delle-riviste/riferimenti-normativi/>.
- Bisogno, Paolo. 1980. *Teoria della Documentazione*. Milano: Franco Angeli.
- Castellucci, Paola, e Sara Mori. 2022. *Suzanne Briet nostra contemporanea*. Milano: Mimesis.
- Diamanti, Ilvo. 2008. *1992. Tangentopoli*. Bari: Laterza.
- Folino, Antonietta, e Erika Pasceri. 2017. *L'Associazione Italiana per la Documentazione Avanzata. Storia e Archivio*. Roma: Aracne.
- ForumPA. 2016. *25 anni di riforme della PA: troppe norme, pochi traguardi. La riforma Madia vista da quattro prospettive di analisi*. Digital360. https://www.astrid-online.it/static/upload/fpa_/fpa_25-anni-di-riforme_19_12_16.pdf.pdf.
- Fumagalli, Giuseppe. 1895. “La Conferenza Internazionale Bibliografica di Bruxelles e il Repertorio Bibliografico Universale.” *Rivista delle Biblioteche e degli Archivi* 6 (9-10): 130-44.

⁷ La competenza dell'ANVUR in materia è sancita dal comma 2 dell'allegato B del DM 76/2012 e riconfermata dal comma 4 dell'allegato D del DM 120/2016. Ai sensi di quest'ultimo, per i settori c.d. “non bibliometrici” (identificati dal comma 1 dello stesso allegato), l'Agenzia «anche avvalendosi di esperti e revisori anonimi, determina e aggiorna regolarmente, pubblicandoli sul proprio sito istituzionale, a) l'elenco di tutte le riviste di carattere scientifico dotate di ISSN; b) il sottoinsieme delle riviste di carattere scientifico definite “di classe A”» (ANVUR, n.d.).

Guarasci, Roberto, e Maria Taverniti. 2013. *Paul Otlet, il Rapport per l'Istituto Internazionale di Agricoltura*, Roma: Aracne.

Guarasci, Roberto. 2011. "La memoria della Scienza: L'Archivio Tecnico Italiano e il Centro Nazionale di Documentazione Scientifica." In *Archivi Privati*, a cura di Roberto Guarasci e Erika Pasceri, 195-218. Roma: CNR.

La Documentazione in Italia. Atti del Congresso Mondiale di Documentazione. 1952. Roma: CNR.

Contributo su invito

Que reste-t-il de nos amours?

Arti tradizionali di un possibile moderno Trivio:
Archivistica, Bibliografia, Documentazione

Piero Innocenti*

Abstract: The text is a journey through memory exploring the evolution of Book and Document Sciences over the last forty years, inspired by a 1942 song by Charles Trenet. The essay primarily addresses the question “what has become” of documentation science, only briefly touching upon “what will become.” It traces the history of Archival Science, Bibliography, Library Science, and Documentation, emphasizing the value and centrality of knowledge systematization. The text then delves into the complex issues related to the production, use, and preservation of digital documents from an interdisciplinary perspective, while not neglecting the publication of the Records in Contexts-Conceptual Model by the International Council on Archives, which unifies and replaces four pre-existing international standards. This overview concludes with a quote from Jules Verne, suggesting a connection with the past and the uncertainty of the future.

Keywords: Archival Science, Bibliography, Library Science, Documentation, Book and Document Sciences.

1. Nella buca del suggeritore

Il titolo è rubato a una coinvolgente canzone di Charles Trenet, rivolta a lenire con dosi massicce di sentimentalismo chi fosse, o si sentisse, giovane durante uno dei momenti peggiori della Francia del secolo scorso (1942), con la *république* a rischio di asservimento allo straniero. Ho tratto licenza di menzionare *amours* attingendo a una pubblicazione di settore la cui citazione esibisco come lasciapassare: «Archivio è una parola d’amore. Per la vita» (Valacchi 2017, 20). Ciò spieghi l’intrusione della musica, e di Cupido, nell’austera dimensione della scienza.

Vorrei sviluppare questo scritto il più possibile seguendo alla lettera quanto mi è stato ventilato nel cortese invito (non declinabile: me lo ha rivolto, sopravvalutandomi, uno stimato collega e amico da molti anni) a

* Già Professore ordinario. Dipartimento di Storia e culture del testo e del documento, Università degli Studi della Tuscia, Viterbo. piero.innocenti1945@gmail.com.

produrre un testo che offra “spunti di riflessione sulle scienze del testo e del documento nell’ultimo quarantennio”. Se si richiede riflessione, pare ovvio si sottintendano due domande: “che ne è”, e anche “che ne sarà”: in gran parte, saranno abbozzi, temo. Tratto solo la prima, delle due domande; alla seconda cercherò di accennare brevemente in chiusura. Sono grato dei termini dell’invito: sistematizzare un arco di tempo lungo quasi quanto un’intera vita di lavoro avrebbe sfidato la mia *Selbstbewusstsein* a una gara perduta in partenza: da ex liceale, al logoro, in Italia, “autocoscienza” preferisco Kant (come, *si parua licet*, Paolo Bisogno nella bella citazione kantiana posta in sua memoria da Basili in esergo di «AIDAinformazioni», no. 3 (luglio-settembre) del 2003; e penso anche al modo in cui lo ricorda (Castellucci 2003, 59-70)).

La precedente frazione di tempo, 1983-2003, era stata in gran parte dedicata al consolidamento della comunità associativa e della sua testata, «AIDAinformazioni». L’acronimo che ne apre il titolo vale Associazione Italiana per la Documentazione Avanzata, fondata nel 1983, la cui storia è inquadrata nel volume di bilancio *AIDA20*, curato da Basili e Bogliolo (2003): non sarà inutile ricordarne il contenuto¹, anche se oggi ad esso si può guardare da una prospettiva di scorcio maggiore nel lavoro di (Folino e Pasceri 2017), che ne presenta e inquadra l’archivio, e nel *corpus* delle annate della rivista, esse stesse “documento” e “monumento”, per usare un binomio della storiografia del secondo Novecento che vale qui da prolessi.

L’indicazione delle quattro decadi a ritroso conduce con lo scarto di un anno nelle vicinanze di (Lodolini 1985), *Organizzazione e legislazione archivistica italiana*, caposaldo storico-dottrinario nato da un insieme di diuturne ricognizioni territoriali anticipate a porzioni in rivista col titolo *Itinerari archivistici*; la pubblicazione, rimodellata e ripresa più volte, è rimasta vitale almeno fino al 2005, prefata da Spadolini, primo titolare del Ministero per i beni culturali di nuova istituzione (1974). Nel complesso, è periodo che si chiude al giro di boa della seconda di due frazioni, vent’anni ciascuna, di un tragitto verso un orizzonte improvvisamente meno definito: ne è spia il gran numero di titoli più spesso di intervento occasionale che progettati, recuperabili all’informazione solo da registrazione

¹ Nel fascicolo, aperto da un editoriale di (De Gregorio 2003), (Maffei 2003) celebra i venti anni di AIDA; (Baldazzi e Carosella 2003), ne ricordano le tappe importanti. Basili presenta il fascicolo, poi tratta, con Reale, dell’Istituto di studi sulla ricerca e documentazione scientifica (ISRDS), e in un terzo contributo del ricorrere della Documentazione nei convegni dei vent’anni precedenti. (Alberani e Poltronieri 2003) sviluppano la comparazione fra Documentazione e altre discipline dell’informazione, come anche (Carosella 2003) che parla di Documentazione comparata. Castellucci ricorda Paolo Bisogno e il suo contributo alla precoce fondazione della disciplina. Infine, si ha la bibliografia associativa sulla professione (Baldazzi 2003), il riassunto del fascicolo (Sannino) e alcune notizie su incontri ed iniziative dell’anno corrente.

randomizzata e particolarizzata, come si dirà brevemente più avanti (e più estesamente altrove, se ne capiterà l'occasione): è tematica orizzontale che fa riflettere sulla polverizzazione della ricerca, che parcellizzandosi talvolta sembra banalizzarsi.

La parola "spunto" mi assicura inoltre (altro fattore di tranquillità), il patrocinio di un avvocato d'ufficio affidabile, il Vocabolario Treccani *on line*, che così definisce uno dei significati possibili del termine: «Nel linguaggio teatrale, le prime parole di una battuta sussurrate dal suggeritore all'attore» (Treccani, n.d.). Mi comporterò dunque da suggeritore: né autore, né attore, semplice artiere, spigolando qualche idea e precisando infine che la testata ospitante mi sta a cuore non solo per il suo essere fonte di aggiornamento scientifico in un'area di ricerca che mi ha interessato da sempre fino ad oggi, ma anche perché mi trovai (2001) a collaborare per breve tempo alla organizzazione scientifica della testata (come altri si trovò in situazione analoga per rappresentarne la figura giuridica), allora trimestrale: tutto ciò è ricordato in «AIDAinformazioni», no. 1 (gennaio-marzo), ("Ai Lettori" 2000, 1-2), dove si menzionano le nomine di De Gregorio e di chi scrive negli incarichi rispettivamente di direttore responsabile e scientifico. Anche quello era un invito non declinabile, per gli stessi motivi di oggi: rispetto, stima, amicizia verso i proponenti. Era, per fortuna di ambedue le parti, un momento di transizione, trascorso il quale fu opportuno, di comune accordo, congedarsi da quella per me non infuocata esperienza.

2. Civiltà di parole: Archivistica, Bibliografia, Biblioteconomia, Documentazione

La sequenza delle parole nel titolo è alfabetica: ordinamento enunciativo cui si ricorre per la sua chiarezza, talvolta anche allo scopo di evitare rivalità e risse su questioni di gerarchia, frequentemente fomite di discordie all'interno della nostra area di studio e professionale. Anni or sono invece mi capitò (Innocenti 2008a, 47; Innocenti 2008b, 147) di esporre questioni di sostanza e di nominalismo terminologico seguendo «Una *climax* discendente: Bibliografia → Documentazione → Biblioteconomia → Archivistica [...] desunta dal *Vocabolario bibliografico* di Giuseppe Fumagalli (1940)» (l'autore era morto un anno prima) che trattando la Documentazione la definisce «ampliamento e perfezionamento della Bibliografia» (Fumagalli 1940, 150), osservando anche che a ciò portava la definizione datane da Paul Otlet nel 1934. Mai avrei pensato che quella remota citazione mi tornasse a servire, però oggi mi ritrovo a poter affermare che quella sintesi rappresenta in quel momento la più importante sistemazione in Italiano, a livello terminologico, della ospitalità della Documentazione fra le discipline del libro e del docu-

mento. Si potrebbe dire che concettualmente la posizione dell'Archivistica vi è sottostimata, ma è un effetto ottico dato dall'apparenza gerarchizzata (la lettura di (Ardolino 2022) è sufficiente qui a ricordare Fumagalli e le caratteristiche della sua informazione): l'enunciato complanare restituito dall'ordine alfabetico ripristina, mi sembra, i parallelismi. L'intuizione del 1940 è per così dire "sbocciata" solo più tardi, in modo apparentemente non-logico; è impossibile però non tenere conto del progressivo isolamento politico italiano per le sanzioni del 1935, culturale per l'autarchia conseguente, definitivo e internazionale con la guerra. Solo il libero confronto nella pace potrà riavviare il commercio intellettuale internazionale: nell'insieme degli scritti *Bibliographic Organisation*, Shera e Egan (1951) raccolgono gli atti di un convegno a Chicago dell'anno precedente; è un bilancio della ripartenza, a partire da considerazioni sull'inventario, per così dire, dei danni. Vi si può leggere, fra le altre questa lucida valutazione di uno dei partecipanti, ministro luterano e studioso di Archeologia, *patronus* dei lavori sui Rotoli del Mar Morto:

al relativo impoverimento della cultura umanistica si devono aggiungere gli effetti distruttivi delle guerre che abbiamo subito, guerre che hanno reso possibili enormi avanzamenti della scienza, conferendole un prestigio quasi oracolare e distruggendo nello stesso tempo gran parte dello sforzo umanistico, dislocando una parte rilevante del modello tradizionale delle relazioni colte. Questo, in ultima analisi, si rivelerà altamente salutare in futuro, ma per l'immediato gli effetti sui materiali e sugli strumenti dell'uso professionale sono devastanti, in particolare per quanto concerne prospettive e speranze bibliografiche (Kraeling 1951, 117).

La prima raccolta bibliografica delle pubblicazioni in Italiano relative all'archivistica (così suona il titolo di (Perrella 1963)), arriva anni dopo. L'autore si presenta in veste presso che penitenziale, definendo l'archivistica italiana "Cenerentola delle discipline" e con l'intento di sollevarne le sorti spiega la sua minuziosa intrapresa e il criterio seguito nell'offrirla in pubblico. All'interno dell'opera – a mia conoscenza per la prima volta in una bibliografia autonoma e monografica, e non più nella voce di cinque righe di un lessico tecnico – alla "documentazione" viene conferita personalità *uti singula* all'interno della tassonomia adottata per distribuire la materia. Perrella registra 1411 pubblicazioni, distribuite in poco meno di una trentina di partizioni e indicizzate da > 700 termini di soggetto e > 800 fra autori e titoli. La XIV partizione s'intitola *Documentazione e fotodocumentazione*, annovera 28 *item* di data successiva al 1949, pari all'1,9% del totale dei titoli registrati: era la tecnologia più avanzata dell'epoca per riprodurre e duplicare e giusto alla metà degli anni Sessanta si andava arricchendo con la diffusione (tardiva in Italia) della xerocopia. Poco, ma > 0, e rappresenta attenzione dedicata a una piccola quantità, bastevole però per segnare un passo avanti. Perrella era all'e-

poca Soprintendente degli Archivi di Lazio, Umbria e Marche, come risulta dal suo messaggio di adesione al IV Convegno degli Archivistici Ecclesiastici (Roma, Settembre 1962); delle sue pubblicazioni a me note la prima data al 1949, l'ultima al 1975. Il suo nome è legato con autorialità di secondo grado, da traduttore, almeno a un altro titolo della sua produzione a stampa. È che titolo: cinque anni dopo, pazientemente e meritoriamente, egli pubblica nel 1968 la traduzione italiana di (Brenneke 1953) sui fondamenti dell'archivistica in Europa, *opus postumum* salvato dalla ricostruzione su lascito di testi e appunti di lezioni da parte del curatore, Wolfgang Leesch (altro grande archivistica). Quanto alla bibliografia, nell'*Avvertenza* della traduzione di Perrella si legge: «L'elenco bibliografico, finito di compilare nel dicembre 1951, comprende la più importante bibliografia relativa alla archivistica teorica ed alla tecnologia archivistica, mentre, per quanto riguarda la bibliografia relativa alla storia archivistica, include solo quella sugli archivi statali; per gli archivi esteri è stata presa in considerazione solo la bibliografia introduttiva, quella generale e quella sugli archivi statali centrali» (Brenneke 1968, 484, 485-580). Quest'altro imponente elenco registra 1438 *item*, creando così non solo un manuale fondante ma anche la prima bibliografia speciale retrospettiva di Archivistica. Il testo di Brenneke reca in esergo la traduzione in Tedesco di una frase da (Casanova 1928, 505): «La conservazione degli atti corrisponde ad un bisogno innato dell'umanità, bisogno che l'ignoranza potrà pur calpestare, ma sopprimere non mai». Perrella la ripristina, ovviamente, nell'originale. Nel primo rigo della trattazione, sotto, Brenneke rinnova la professione di elogio e rispetto per il collega, di lui poco più anziano e che gli sarebbe sopravvissuto:

Il Casanova, che è l'autore della migliore opera generale sugli archivi che noi possediamo, divide l'archivistica in: archivistica pratica o tecnologica (archiviconomia), la quale ha come fine principale quello di elaborare le basi scientifiche del trattamento esterno del materiale archivistico (ricezione, custodia, conservazione e restauro); archivistica propriamente detta, ossia archivistica teorica (archivistica pura), che studia le regole per la sua utilizzazione e quindi i metodi per l'ordinamento interno e si occupa della natura e delle varie specie di archivi; diritto archivistico, al quale si aggiunge, in fine, anche la storia degli archivi e dell'archivistica (Brenneke 1968).

Archivistica Cenerentola? forse: ma sguattera o Principessa? *Sum leo quia nominor leo*, i riconoscimenti derivanti dalle due trattazioni sono molto alti, e la descrivono come ricca di sfaccettature e di pieghe riposte. Per trovare fra quelle pieghe la Documentazione bisogna addentrarsi molto nella lettura di Brenneke per trovare la formulazione, che consenta a Perrella di mettere una maiuscola anche al termine italiano che incidentalmente sembra translitterare *Dokumentation*, ma in realtà lo retroconverte:

Si raccoglie quindi ora non più solo il materiale scritto e stampato prodotto dalla gestione degli affari (redazioni commerciali, regolamenti aziendali, statuti sociali, scritti pubblicitari, listini di prezzi, lettere circolari), ma anche generalmente tutto il materiale documentario relativo all'economia | politica, che interessi ai fini statistici e di informazione nell'ambito delle imprese al cui servizio si trova l'archivio (statistiche, scritti celebrativi e commemorativi, giornali e periodici economici, ritagli di stampa, disegni, piante, illustrazioni). I grandi archivi regionali, come il renano-vestfalico o l'Archivio economico della Saar, servono da ufficio informazioni e da centro statistico per l'economia del loro distretto; gli archivi aziendali offrono alla loro azienda le basi per l'ulteriore sviluppo tecnico e la programmazione commerciale. Queste due attività sono state insieme raccolte sotto il moderno concetto di 'Documentazione' [*Dokumentation*], la cui funzione è la raccolta, l'ordinamento e la utilizzazione dei documenti di qualunque genere – quindi non solamente del materiale archivistico o manoscritto –, che possano avere importanza quale materiale documentario [*Tatsachenmaterial*] per determinati fini pratici e teorici. Con questa duplice attribuzione di funzioni il moderno archivio economico ritorna alla duplice funzionalità degli archivi antichi, che fu persa soltanto per l'unilaterale orientamento storico del secolo XIX, e diventa quindi un modello per gli archivi moderni in generale, il cui compito è ora, di riacquistare, slargando le sue finalità storiche e teoriche, quell'importanza attuale e pratica che un tempo fu loro (Brenneke 1968, 475-476).

Qui viene sottolineato come sia *anche* il riaffiorare di un concetto antico, in realtà, il passaggio da Archivistica dell'ottocentesco metodo storico a Documentazione "moderna", disciplina cui non ripugna, se possibile e necessario, usare algoritmi automatici di prossimità terminologica relativa ai contenuti, mentre in quella archivistica, come è stato ricordato commentando il luogo di Brenneke in un manuale universitario del 2018: «il documento già alla sua origine si vede assegnare uno spazio amministrativo e archivistico prefissato all'interno del tutto, e tale complesso può essere interrogato da un terzo seguendo la strada tracciata dalla formazione dell'archivio» (Santoro 2018, 172). Il termine "documentazione" (*Dokumentation*) ricorre sia nel testo italiano, sia nel testo tedesco esplicitamente ≥ 30 volte. Quanto a *Dokument(-en) / Urkunde(-en)*, documento/i, inutile contare, lo si trova si può dire ad ogni riga. Le parentesi quadre nel testo citato *supra* sono di Perrella, che usa l'accorgimento per traslare la sinonimia del Tedesco nel termine italiano che possa unificarla. In questo caso, egli riporta sia *Dokumentation* sia il *Tatsachenmaterial* dell'originale, sotto l'ombrello del sostantivo italiano "documentazione", nostro deverbale da "documentare".

Per comprendere ciò che ha dovuto affrontare Perrella per la resa interlinguistica della terminologia fra due lingue di ceppo diverso, si dovrà

ricordare di come in decenni più vicini a noi la più numerosa poliglossia abbia sollevato dubbi autorevoli, come (Lodolini 1987) su le incertezze terminologiche di *DAT* (*Dictionary of Archival Terminology 1984*), troppo disattento a quelli che l'interpretariato interlinguistico chiama oggi i "falsi amici" (ci torna a proposito di altro testo (Carucci 2022, 79)). Sono questioni di metodo che rendono necessario riprendere un momento de Saussure e Le Goff, cronologicamente distanti, ambedue vissuti al passaggio di due secoli scalati fra loro; campi diversi, intuizioni convergenti. Non si allude a singoli luoghi delle loro opere, ma alle metodologie: chi ha riflettuto su scienze del testo e del documento durante l'ultimo quarantennio ha dovuto fare i conti più o meno consapevolmente con quel grande precedente (molti a questo punto non esiterebbero a citare il monito di Verdi a Florimo, 1871, «Torniamo all'antico, sarà un progresso»); quindi risaliamo alla fonte, la parola.

Se "denotazione" va al *focus* e "connotazione" all'alone di contesto (sviluppati o subito) di un termine, ci si può chiedere a quale *focus* rimandi il significante, e se sia raggiungibile senza porlo all'interno di un sistema di plurimi assi ortogonali, ma disorientati: o a scansione pendolare secondo l'alternanza, oppure in concomitanza, di diacronia e sincronia. Il termine "documentazione" si presta peculiarmente a oscillare fra "denotando" e "connotando" se inserito nel campo di forze dia-/sincronico, cui soggiacciono i semantemi: trae origine infatti da "documento", il pro-nipote ultimo di una lunghissima discendenza verticale dal Latino *documentum*, persa la desinenza (o derivando dall'ablativo) in séguito alla perdita della flessione nel lento passaggio interlinguistico Latino → Italiano. Sono sì attestati anche *documentare* e *documentatio*, ma hanno la caratteristica di creazioni di laboratorio (nate morte) dei grammatici della tarda latinità, cui dedicarono il loro studio prima (Lindsay 1917), poi (Flammini 1990), e altri. Raramente usati, rappresentano uno scarto laterale dalla linea che dai latini *documen* → *documentum* porta ai nostri "documento" → "documentare" → "documentazione". Grande dunque il successo della radice *document-* per fecondare la ben più vasta area neo-latina e oltre: nella sincronia del momento in cui sto scrivendo, popola il panorama linguistico d'Europa, tranne pochissime isole non raggiunte dal radicale latino, nelle quali sopravvivono tradizioni "altre". Per mera curiosità probatoria (quindi "documentale"), sono queste le forme attestate: *document* (catalano, francese, inglese, olandese); *documento* (castigliano, galiziano, portoghese); *documentul* (rumeno); *doiciméad* (irlandese); *Dokument* (tedesco), *dokument* (cèco, bosniaco, croato, maltese, polacco, slovacco, sloveno), *документ* (*dokument*, russo, serbo, ucraino); *dokumentas* (lituano); *dokumentet* (danese, norvegese, svedese); *dokumenti* (estone); *dokumentua* (basco); *dokumentu* (lettone); *dokumentum* (ungherese); *dokumint* (frisone); *döküman* (turco); *документот* (*dokumentot*: macedone);

Документът (*Dokumentŭt*, bulgaro). Nascono invece da altre radici: *asia-kirja* (finlandese); *y ddogfen* (gallese); *το έγγραφο* (*to éngrafo*: neo-greco); *an sgrìobhainn* (gaelico di Scozia); *skjalið* (islandese).

Ida Kapp (1884-1979: v. *Thesaurismata* (Hiltbrunner, Hildegard, e Tietze 1954) e (Schrickx 2020)), filologa dal profilo biografico e intellettuale mitteleuropeo, è famosa non solo per essere la prima e, pare, unica donna che arrivò ad addottorarsi con Wilamowitz; non solo per la sua eccellente edizione della *Ecale* di Callimaco, ma anche per i suoi anni di redattrice, poi redattrice-capo, del *Thesaurus Linguae Latinae* (da sempre *ThLL*). In tale veste, firmò la stesura delle colonne che trattano *documentum*, forma a sua volta ampliata per filiazione di *docum-en/docim-en (-inis)*, di chiara derivazione da *doceo*, “insegnare”, tal che Kapp cita Lucrezio, unico autore che lo usa prima di Terenziano Mauro, nel senso di *exemplum monendo docens*, “esempio che spiega ammonendo” ((*ThLL*, n.d. 5,1: 1803-1809), *De rerum natura* 6, 391-392: *halent | pectore perfixo, documen mortalibus acre*, scolasticamente traducibile «esalino [le fiamme] dal petto trafitto, aspro monito ai mortali»). La tassonomia semantica profilata da Kapp è questa:

I. **id quo quis docetur**: A. praevalente notione demonstrandi. 1. fere i. q. exemplum quod docet. 2. fere i. q. indicium ex quo aliquid manifestum fit, [&c.]. 3. de re ipsa quae docet, i. q. terminus quo fines agrorum indicantur. B. praevalente notione admonendi. 1. fere i. q. exemplum ad admonendum vel maxime deterrendum propositum saepe accedit enuntiatum finale, [&c.]. 2. fere i. q. ad discendum res aut imitandas aut ad imitationem aliquo modo aptas propositum. C. praevalente notione probandi. 1. fere i. q. exemplum quod probat aliquid, probatio. 2. fere i. q. testimonium, argumentum. 3. sensu iuridiciali: quodcumque iudicio adfertur ad aliquid probandum. 4. de re ipsa quae probat. II. **id quod ad docendum pertinet, fere i. q. instructio**. III. **id quod docetur: fere i. q. praeceptum** (*ThLL*, n.d., 5.1803-1809).

La “connotazione” del documento ha massimo rilievo di interpretazione quando lo si tratti lontano dalla sfera della produzione, entrati nella sfera dell’uso sociale o economico.

Caso di scuola che la mia generazione sentì illustrare al terz’anno di Università: i documenti allegati ad un contratto di compravendita rogato danno valore *erga omnes* all’atto, ma singolarmente vanno ricondotti al contesto della loro produzione originale in caso di usi estranei alla validazione dell’atto cui sono allegati. Ma cito volentieri la meno afosamente giuridica connotazione documentale di una comunicazione elettronica con cui sono entrato in contatto nel corso delle ricerche per questi spunti: aprendo l’O-PAC (Online Public Access Catalogue) del Sistema bibliotecario di Ateneo dell’Università di Siena per un controllo, ne venni salutato così «Utilizziamo i cookie per essere sicuri che tu possa avere *la migliore esperienza sul nostro sito*.

Se continui ad utilizzare questo sito *noi assumiamo che tu ne sia felice* [miei i corsivi]», così in (Università di Siena, Sistema Bibliotecario di Ateneo, n.d.). Il più recente augurio ricevuto era del 13 marzo 2024: non ho resistito, e richiesto dal sistema di un commento di *feedback* ho disciplinatamente inviato pollice alzato e cuoricino. La non richiesta (e assai dubbia) inferenza, che ho interpretato come auspicio di frequente uso del servizio, e assai subdola quanto ai *cookie*, mi ha riportato non a quaranta ma a cinquanta anni or sono quando, lavorando in Bodleian Library e fruendo nei suoi accoglienti *toilet* maschili della carta igienica, la scoprii in qualche modo “ufficiale”, ciascun foglietto ammonendo, stampigliato ed esclamativo: “And now, wash your hands, please!” La prescrizione non avrebbe raggiunto probabilmente la stessa efficacia se appesa alla parete su un cartello. Quanto alla diacronia, il ricordare gli onnipresenti cartelli negli esercizi pubblici della mia infanzia che ammonivano “La persona educata non bestemmia e non sputa per terra” vale ancor oggi per dire che se anche ciò avrà dato scarsi risultati, rimane documento di quanto si sputasse in terra e si bestemmiasse nell’Italia del 1950-1955.

Insomma: la connotazione, pur meno specifica, può vincerla, quanto a forza e durata, sulla denotazione: il contesto percepito nell’istantaneità di viene cogente rispetto al *prima-durante-dopo* di una frase: sulla rilevanza di questo tema ai fini dell’indicizzazione rimando alle trattazioni di (Buizza 2019, *Indicare quasi la stessa cosa: appunti di indicizzazione interlinguistica*) e (Fabbrizzi 2019, *I contesti dell’indicizzazione*) ambedue in *Viaggi a bordo di una parola* 2019 (Lucarelli, Petrucciani, e Viti 2019): pubblicazione ricca anche per gli altri contributi ospitati, tutti utili alla comunità professionale di riferimento: in particolare *Un catalogatore cartaceo di fronte alla Rete*, del grande Carlo Revelli (2019), ultimo, a mia notizia, dei suoi scritti pubblicati in vita. Parlando di parole, devo menzionare infine altri due contributi il primo è (Guerrini 2011), che formula una proposta in stato nascente di “tesoro delle scienze del documento”, coerente col quadro della riflessione in quel momento. Il secondo è quello di (Penzo Doria 2022), che lancia una sfida interessante quanto al peso specifico delle parole e al metodo con cui lo si determina; vi si passa in rassegna, allegando una sessantina di voci bibliografiche, una serie di definizioni di “archivio” sottoposte a scrutinio terminologico allo scopo di arrivare a una nuova definizione per via semantica ma anche epistemologica. La rivista *on line* su cui è stato pubblicato ne allega un riassunto chiaro e senza sforzo lo trascivo:

L’articolo propone una definizione della parola “archivio” a valle di un’indagine sugli orientamenti dottrinali in materia, analizzando – parola per parola – i lemmi in cui si compone la nuova proposta. Pur non esistendo una definizione valida in ogni contesto, il panorama dottrinale sembra soffrire in larga misura di una visione eterogenea, a valle di un cammino scientifico rigoroso, ma den-

so di imprevisti concettuali e lessicali. Ogni parola – con accezioni e contesti – non costituisce un significante semplice, ma rinvia a un bagaglio proprio di ogni tradizione disciplinare e culturale. Per queste ragioni, tenteremo di fornire una visione d'insieme, dal particolare al generale, delle definizioni di archivio (Penzo Doria 2022).

Nel filone delle bibliografie speciali autonome bisogna arrivare al 2000 per imbattersi in (Massabò Ricci e Carassi 2000), repertorio fortemente selettivo dal titolo italo-franco-inglese, con tre presentazioni, che si dichiara in linea di diretta successione a Perrella, pur precisando la differenza principale di metodologia. Presenta scritti di teoria archivistica italiana, scegliendo, registrando e indicizzando 63 pubblicazioni uscite «successivamente all'emanazione della legge archivistica del 1963, che ha segnato in termini normativi il coronamento del lungo dibattito iniziato nel 1861 con l'unificazione nazionale. [...] Le poche citazioni anteriori al 1963 inserite nella presente raccolta nascono pertanto o dal carattere di pietra miliare dell'opera prescelta o dalla sua stretta relazione con temi ancora oggi in discussione» (Massabò Ricci e Carassi 2000, 17). I due autori vengono dalla tradizione professionale e scientifica dell'archivistica torinese; la caratteristica peculiare del lavoro è la massima cura applicata nella lettura attenta di quanto poi riassume. Fra titoli e riassunti ricorre una quarantina di volte il termine “documentazione”, qui con la minuscola perché, se non ho visto male, solo in tre casi è connotato da sfumatura disciplinare e non solo a sinonimo di “patrimonio” (per risalire all'origine: (Casanova 1928) nelle sue 541 pagine usa praticamente ad ogni pagina la radice *Archiv-*; ≥ 2 volte la radice *Bibliograf-*; ≥ 7 volte la radice *Bibliotec-*; ≥ 9 volte il termine “Documentazione”; ≥ 4 volte “Indice/i”; e mai “Indicizzazione”²).

Non si può procedere ulteriormente, però, nella cronologia senza segnalare due intercalazioni che si frappongono fra (Perrella 1963) e (Massabò Ricci e Carassi 2000), datate l'una 1965, l'altra 1997 (ma 1999). Diversamente importanti quanto a modalità e stadio evolutivo, trattano di documentazione, di organizzazione e recupero dell'informazione, di indicizzazione, rendendosi conto (embrionalmente quella del 1965, più articolatamente la successiva) che occorre guardare orizzontalmente ai diversi contesti di fruizione, quindi archivî e biblioteche (come da sempre), ma anche luoghi di produzione e di lavoro e finalizzati a manufatti e artefatti, a servizi, a cultura (del passato e in atto), fino alla mera evasione di tempo libero; studiando i piani di interconnessione atti ad esprimere la migliore informazione possibile per il miglior servizio non schiacciato sull'ovvietà del consenso generale ma che persegua l'accrescimento di conoscenza, correlato al livello

² Il sondaggio linguistico, praticato su tutte le opere citate in queste pagine, dà risultati espliciti ma non consolidati e oltre tutto richiede lunghe precisazioni al contorno: rimane per ora un mio piacere solitario.

dell'obiettivo, individuale e sociale (si torna con questo al petulante Socrate del *Protagora?* in un certo senso, sì): potrebbe non essere peregrina l'idea che Fumagalli abbia concepito la Documentazione derivata dalla Bibliografia semplicemente considerando che dai libri o si estraggono le informazioni o sono muti.

La prima intercalazione è *Documentazione 1-3* (1965). La pubblica il (Servizio informazione e diffusione del Comitato nazionale per la produttività - CNP 1965) con sede in Roma³; consta di tre volumi di quasi 500 carte in totale, si presenta come dattiloscritto cartulato, non paginato, riprodotto a ciclostile. Il suo secondo volume, intitolato *La documentazione in azienda*, è chiuso da una criptobibliografia di circa 230 *item*, per una trentina di carte, curata da Bruno Balbis; che aggiornano e in certa misura integrano Perrella, e ciò a così breve distanza di tempo è di per sé rilevante; il finale del contributo informa: «È in corso di costituzione l'Associazione Italiana dei Documentalisti e Tecnici dell'Informazione, AIDI, che avrà la sua sede provvisoria presso il Comitato Nazionale per la Produttività - Roma». Molto più della nota bibliografica, per altro diligente, interessa la presenza dei saggi che esplorano a fondo, per l'epoca, le faccette della materia dal punto di vista archivistico, biblioteconomico, documentale. Il saggio più durevole è quello di (Casamassima 1965), che sarà preso a punto di partenza per la produzione del *Nuovo Soggettario* durante il primo decennio 2000 (in merito si diffonde (Cheti 2008)); il resto del *parterre* di autori, a me noti o meno noti, consiste prima di tutto ancora di Balbis, sul lavoro documentario⁴. Seguono F. Barberi (ben noto già allora nel mondo delle biblioteche), ancora su soggettazione e classificazione; A. M. Stein (funzionaria delle Dogane di Roma secondo Opac Sbn), su brevetti e su cataloghi commerciali e industriali; A. Mei, (ingegnere, esperto di tecnologia e di pubblica amministrazione), sull'archivio moderno; M. T. Ronchi (nota ai lettori di questa rivista e al mondo bibliotecario, se non altro per il vivace omaggio in (Bogliolo 2003)), sulla classificazione a

³ Comitato istituito nei primi anni di costruzione della Repubblica; intensamente attivo negli anni Cinquanta, le carte si trovano parte in Università cattolica del Sacro Cuore (sia Gemelli sia Franceschini ebbero ruolo nella iniziativa politica), parte in Archivio centrale dello Stato. Non è il caso di approfondire qui la possibile *trash-news* circolata il 13 novembre 2020 quando si pubblicava con enfasi: «Nel 2016 il Consiglio dell'Unione Europea raccomandava agli stati membri di dotarsi di un Comitato Nazionale per la Produttività. Con la legge di bilancio per il 2021 l'Italia ha deciso di istituire questo organismo, stanziando 2,5 milioni di euro all'anno» (Canna 2020). Idea franata nel giro di pochi giorni, per altri motivi, nessuno ha segnalato il probabile precedente.

⁴ Va annotato, su Balbis, che di due suoi lavori sulla situazione in Italia della disciplina e del suo insegnamento (Balbis 1950; 1951) tornava (Boustany 1991). (Guarasci 2022, 149), ne definisce la posizione "estremamente restrittiva", recensendo la trad. italiana di un libro di Briet e ricordando Otlet; nel 2011 egli aveva studiato le carte del Centro Nazionale di Documentazione Scientifica, che Balbis aveva diretto dal 1956, prima di passare al Segretariato della Camera dei Deputati.

faccette; infine Alessandro D'Alessandro, a me colpevolmente rimasto ignoto, su "classificazione" dei documenti d'archivio. (A mezzo secolo di distanza, e su tutt'altra casistica, (Greco 2011) riflette sulle criticità derivanti dal tentar di classificare un archivio: che stavolta però è digitale). La pubblicazione è rimasta presso che clandestina: rarefattasi verosimilmente a causa delle precarie caratteristiche materiali e forse anche perché diffusa in circuito ristretto. Al momento dell'ultimo controllo, 7 biblioteche italiane la dichiaravano posseduta in tutto o in parte; British Library ne possiede 1 copia, parziale; assente in Opac di BNCF (Biblioteca Nazionale Centrale di Firenze), KIT (Karlsruhe Institute of Technology), OCLC (Online Computer Library Center), LC Washington (The Library of Congress of Washington). Ne ho copia personale, dono augurale di un collega più anziano il giorno in cui ricevette il mio giuramento in BNCF.

La seconda intercalazione è di (Galluzzi e Valentino 1997), atti di un seminario (©1997), arrivati alla stampa nel dicembre 1998, contrassegnati da "finito di stampare" 1999. Questa, diversamente dalla precedente, è presente in (Massabò Ricci e Carassi 2000, 46) che la ascrive al 1997, per poi riassumerne la sezione archivistica col no. di scheda 53 (86-87). Qui la si segnala più estesamente per l'ampio rilievo anche di altre aree, trasversalmente attraversate da costante preoccupazione documentalistica, come si evince dalle presentazioni che si devono a due *grand commis* istituzionali e socio-economici dell'epoca (Maccanico 1997, V-XIII; Imperatori 1997, XV-XIX) e dalle due introduzioni di un accademico già allora di alto profilo internazionale (Galluzzi 1997, XXI-XXVIII, 3-39), la prima di politica della cultura, la seconda su *Nuove tecnologie e funzione culturale dei musei*. Il filo conduttore della trasversalità si prolunga nella sezione di saggi, con tematiche a tutte le quali sottostà interesse per la documentazione, spesso in quanto trattazione del materiale tutelato, ma anche in quanto deontologia disciplinare: E. Lo Sardo (archivista di Stato), "La soglia di Gutenberg: Machinae e archivi"; P. Innocenti (all'epoca Teoria e tecniche della classificazione e catalogazione, Viterbo), "Tecnologie informatiche e struttura tradizionale della biblioteca"; C. Pettenati (CERN), "La biblioteca virtuale: Problemi e opportunità"; P. Coural (Louvre), "L'impatto delle nuove tecnologie sullo scambio culturale"; T. De Mauro, "Computer, scuola e beni culturali"; W. Santagata (Scienze delle Finanze, Torino), "L'economia del museo nell'epoca delle reti-mercato telematiche"; F. Antinucci (CNR e Berkeley), "Beni artistici e nuove tecnologie"; V. Cappellini (Ingegneria, Firenze), "Telematica per i beni culturali"; P. Nancarrow (Imprimatur ALCS, Londra), "La protezione di voci digitalizzate del patrimonio culturale"; P. A. Valentino (Economia urbana, Roma), "Criteri e metodi per la scelta delle tecnologie informatiche applicabili ai beni culturali". La bibliografia, pur se relegata in note finali, è approfondita.

3. Ingresso nel nuovo secolo, fra manutenzione e innovazione

Il penultimo anno del secolo precedente aveva visto l'introduzione in ambito universitario del ciclo didattico 3 + 2; gli anni successivi nel ricordo rimangono caratterizzati da fluide incertezze. L'introduzione in prima applicazione della novità, col conseguente "smaltimento" dei laureati quadriennali (non è cinismo retrospettivo: andò così), portò a riflettere su una spesso intensa e confusa quotidianità didattica. Capitava di dividersi in Ateneo fra burocrazia deliberativa ed esecutiva, in aula fra insegnamenti di vario titolo e "peso" (dei nuovi CFU – Crediti Formativi Universitari), magari su diverse sedi che usavano bilance e pesi diversi: non ho ricordo che qualcuno si sia chiesto quale rapporto avesse ciò che stavamo facendo con l'articolo 33 della Costituzione: non penso alla prima frase, ripetuta spesso quanto ovviamente, ma alla seconda, «La Repubblica detta le norme generali sull'istruzione, &c.»; "generalì", non minuziose fino all'ossessione dei dettagli... Insegnamenti stratificati da una tradizione finalizzata alla preparazione dei quadri di medio ed alto livello nella organizzazione del corpo sociale entravano in crisi davanti ad una diversa distribuzione didattica da reinventare (e già su quella mi ero espresso criticamente in sede internazionale (Innocenti 1990)). Avrebbe potuto essere un problema a breve raggio; in quante attività si procede per invenzione → rodaggio → finissaggio → messa a regime. Ma nella realtà socio-politica i corpi intermedi, teorici recettori della manodopera qualificata dall'Università, si stavano contestualmente avviando a cambiare assetto o addirittura ad estinguersi in progressiva diminuzione di assorbimento; il che toglieva fascino e "presa" all'obiettivo di crescita culturale, non più molla anche di crescita personale e sociale. A mia conoscenza solo in un saggio del 2020 ho trovato un'analisi accettabile di ciò, che arriva a parlare di emergenza casuale dei ceti dirigenti, anche politici.

Non so se ad altri accadde lo stesso, per me fu complicato ripensare i fondamentali disciplinari alla luce del nuovo panorama, imponeva fretta piuttosto del rigore e calma che l'osservazione richiede. Per affidarsi a supporti letterari: pensavo che Mastronardi, 1962 col suo maestro di Vigevano e Vittorio Enzo Alfieri, 1969 con la sua "onagrocrazia" fossero cronisti, ma anche profeti. Quando la riforma è diventata "maggiorenne", 2017, un commento di Marzio Bartoloni, con l'austera e sana impassibilità dei numeri di "Il Sole 24 Ore", formulava senza *uis* polemica una considerazione, ospitata da un sito sindacale che indirettamente la faceva propria; la richiamo come punto di vista esterno, per esser sicuro di non essere allucinato dal mio personale *meminisse iuuat*:

Insomma il "3+2" è stato un flop, come diceva già nel 2010 l'ex ministro dell'Istruzione, Mariastella Gelmini? Inumeri sembrerebbero dire proprio di

sì. Alla riforma del 1999 – che con il Dm 509 ha introdotto per la prima volta in Italia la novità del “3+2” e dei crediti formativi – sono seguiti altri provvedimenti legislativi che, tra il 2004 e il 2008, hanno provato a ridisegnare la fisionomia degli atenei. Ma la sostanza non è cambiata, come certificano i dati delle iscrizioni all’università: nell’anno accademico 2000-2001 (l’ultimo con i vecchi diplomi) gli immatricolati erano 284mila. Da allora in poi, dopo un primo boom coincidente con l’avvio della riforma che ha fatto registrare un picco con 308mila matricole nel 2006-2007, c’è stata un’inesorabile discesa. Chiusa l’anno scorso con una mini-risalita a 275mila matricole, che a conti fatti significa comunque 10mila iscritti in meno rispetto a 15 anni prima (Bartoloni 2017).

Lungo il decennio cercai e trovai agio e sollievo spostando il baricentro del fare sull’attività di consulenza con altre amministrazioni dello Stato, nei limiti consentiti dalla mia posizione a tempo pieno, e soprattutto con la partecipazione ad alcuni consorzi di dottorato di ricerca: Firenze, animato da Cardini e Regoliosi, Siena, e soprattutto quello specifico di settore di Udine, fondato e diretto da Mauro Caproni (amico e collega di una carriera incominciata il medesimo giorno), e da molti altri colleghi, vecchi e nuovi, per lo più anch’essi amici; dottorati frequentati per fortuna da un buon numero di allievi di valore: i nomi di molti di loro oggi si leggono negli OPAC e nelle bibliografie di settore. Dal 2009 avevo assunto la direzione di «Culture del testo e del documento. Le discipline del libro nelle biblioteche e negli archivi», di Vecchiarelli, cui collaboravo da quando era «Culture del testo»⁵; potevo quindi continuare a cercare di fare e di far fare ricerca, condividendo in questo modo rapporti coi colleghi di quella stagione.

Mi rendo conto di essermi diffuso troppo, ma volevo spiegare perché credo di poter sopporre che ancora oggi sia tangibile l’imbarazzo con cui presentai, 2003, la pubblicazione del CNR dal titolo *Information Literacy in Europe*, ricco volume di panoramica della situazione europea curato da Carla Basili (2003a), che nello stesso anno si dedicava alle cure di «AIDAinformazioni» 3/2003 (Basili 2003b; 2003c) e AIDA20 2003 (la seconda con Bogliolo) (Basili e Bogliolo 2003). Licenziai il brevissimo testo apprezzando sì che la disciplina documentale fosse finalmente stata organicamente riconosciuta nella vita universitaria a livello di didattica e di ricerca, senza però spingermi molto oltre sulla situazione in Italia (per nulla su quella europea); accennai solo alla rivoluzione che si profilava per l’indicizzazione bibliografica:

⁵ L’arco 1995-2009 è esplorato da (Di Filippo e Fedolino 2021) con indici progressivo e semantico: di Documentazione si parla una decina di volte, di Archivistica cinque o sei, di Bibliografia una trentina. È in corso l’indicizzazione dell’intera consistenza: le frequenze del periodo successivo si profilano diverse. Riflessioni sull’analisi semantica di periodici di settore, fra cui questo, sono svolte in (Calvitti 2007a; 2007b; 2008), poi in (Calvitti 2011).

nuove istanze [provengono] dal contesto contemporaneo – da molti denominato “Società dell’informazione” che Basili coglie proponendo questo lavoro. Molto è stato fatto, sin qui, sul versante dell’offerta di servizi informativi, anche tenendo conto della dirompente tecnologia informatica. È tempo ora di volgere l’attenzione alla domanda di informazione, che prolifera sia come conseguenza di tecnologie pervasive, sia come effetto di trasformazioni nel mercato del lavoro; e di preoccuparsi perciò di come promuovere e diffondere una cultura dell’informazione, anche al di fuori dei confini professionali specialistici. Da un paio d’anni, la Documentazione fa parte organicamente della riforma dei settori scientifico-disciplinari che, in ambito accademico, governano il reclutamento e le attribuzioni didattiche, e fanno da riferimento alle aggregazioni del personale della ricerca (fondi di ricerca, progetti co-finanziati, istituzione di dottorati, etc.). Essa cioè è stata inserita, insieme con le più tradizionali Archivistica generale e speciale, Bibliografia e Biblioteconomia e via a seguire, in un’area unica: col risultato di far vedere più da vicino, sia analogicamente, sia contrastivamente, quale sia il suo profilo in relazione a discipline sorelle e/o cugine.

Contemporaneamente, pare si siano moltiplicati (e questo lavoro ne fa fede) i luoghi in cui sulla documentazione non solo si pratica, ma si riflette. Che vi sia integrazione fra luoghi della pratica e momenti della riflessione teorica non è una novità né in generale, né all’interno della comunità bibliotecaria e documentaria. Basterà ricordare come uno dei più importanti monumenti della biblioteconomia di questo secolo, il *Soggettario delle biblioteche italiane* (1956), abbia trovato un momento di discussione particolarmente alto, dopo il primo decennio di applicazione, proprio su una rivista di documentazione aziendale, nel 1966 (Innocenti 2003, XI-XII: si sta citando (Casamassima 1965), implicitamente e con errore di data).

Nonostante il suo scarso significato, il breve accenno veniva cortesemente menzionato poche settimane dopo da (Guarasci 2003). Il suo titolo ribadisce in sottotesto il legame fra il mondo bibliotecario e quello documentale “Libri, documenti e altre storie. L’insegnamento della Documentazione nelle università italiane”; egli ne riprese cinque anni dopo il secondo *colon* promuovendo una riflessione a più voci col manuale *Dal documento all’informazione* (Guarasci 2008), da lui curato coinvolgendo una quantità di amici e colleghi (fra cui chi scrive, oltre a G. Adamo, E. De Francesco, D. Gambarara-A. Givigliano, M. Lancia-A. Lapicciarella, S. Pigliapoco, R. Rolli, A. Romiti, A. Rovella, G. Scandurra, S. Pigliapoco); all’interno, oltre all’“Introduzione” (non firmata, ma sua) esplicitò ulteriormente il proprio pensiero contribuendo coi due saggi, l’introduttivo “Che cos’è la documentazione” e “Indicizzazione e classificazione: Concetti generali”; in quest’ultimo si manifesta l’impulso a considerare l’indicizzazione come metodologia e pratica ancipite, non specificamente archivistica né specificamente bibliografica: suggerimento reso ancora più evidente, quasi quindici anni dopo, dalla collaborazione (Guarasci e Guer-

rini 2022): il secondo è quello del curatore di *Biblioteconomia* (Guerrini e Crupi 2007), sottolineando, come ci si può continuare ad augurare, la necessaria convergenza fra due punti di vista talvolta in passato divergenti, se non competitivi.

Se (Guarasci, Rovella, e Taverniti 2011) parlano ormai come necessità di parlare di archivio digitale al livello politico-culturale-amministrativo dell'Unione Europea, nel 2010, sul versante archivistico, era uscita di Zanni Rosiello (leva 1931, grande carriera fino a direttrice d'archivio a Bologna) una testimonianza autobiografica di una archivista, ma anche di una "lettrice", che traeva dal passato slancio verso il futuro. Nel 1981, in piena attività, si era posta la potenzialmente imbarazzante domanda: *Gli Archivi di Stato: una forma di sapere «segreto» o pubblico?* (Eraclito di Efeso aveva risposto divinizzando l'archivio...); e ciò avveniva contestualmente al bilancio della discussione sulla materia, di Lodolini (1981) "L'ordinamento dell'archivio: nuove discussioni". La riflessione del 2010 (valida ancora oggi, appare sintomo di una sensibilità diffusa) riporta in attualità preoccupazioni di quasi quarant'anni prima, alla vigilia della spadoliniana costituzione del Ministero per i beni culturali, e da essa, si afferma, non vinte:

Ricordo che quando uscì – era il 1973 – il quaderno di Italia nostra dal titolo *Gli archivi: proposte di collaborazione per una migliore tutela*, lo lessi d'un fiato. [...] "Che gli archivi siano da considerare, di pieno diritto, beni culturali è verità tanto ovvia che può apparire del tutto superfluo ricordarla. Eppure cause di varia natura hanno finora impedito che questa verità fosse accettata e messa in pratica in tutte le sue conseguenze." È questo l'incipit del quaderno [...4 ss.] Stanno altresì emergendo archivi di donne. Recuperarli non è stato facile. Le donne sono state a lungo escluse da professioni e ruoli di un certo rilievo e sono state, fino a tempi recenti, debolmente consapevoli della loro identità di genere, dell'importanza di autorappresentarsi e di trasmettere alla posterità tracce di sé. Hanno lasciato perciò una esigua quantità di documentazione, spesso mescolata a quella dell'altro genere. Va inoltre ricordato [...] che la maggior parte della documentazione archivistica è da rapportare a chi aveva (e chi ha) una qualche forma di potere, perlomeno il potere della scrittura. Chi non sapeva (e non sa) scrivere ha lasciato (e lascia, anche se in tempi a noi vicini le cose sono in parte cambiate) esigue tracce di sé. Non pochi storici se ne sono resi conto e hanno cercato di colmare i vuoti presenti nelle sia pure enormi masse documentarie che sono state conservate. Quelli interessati a epoche più lontane hanno studiato [...] i segni di croce che, al posto delle firme, sono stati tracciati in determinati documenti [...]. Gli storici contemporaneisti hanno fatto ricorso alle fonti orali, in particolare alle interviste, per far parlare di sé [...] persone che non l'avevano mai fatto prima e la cui partecipazione a eventi o vicende storiche coevi non aveva lasciato alcuna traccia nella relativa documentazione. In tempi a noi vicini, inoltre, la riduzione dell'analfabetismo, la diffusione di nuove tecnologie che ha portato

fra altro a un crescente e più libero uso della scrittura [...] Oggi inoltre non ci si occupa soltanto di documentazione cartacea, ma anche di documentazione digitale ([...] sia quella prodotta direttamente su supporti elettronici, sia quella che dall'analogico viene trasposta in digitale). [...] l'avvento di Internet, ha rappresentato se non proprio una rivoluzione, come spesso è stata un po' enfaticamente definita, certamente un radicale cambiamento in ogni settore della vita sociale e delle relative possibilità comunicative. Il cambiamento ha investito anche il mondo degli archivi. Oramai, [...] vengono sempre più frequentemente posti problemi concreti e avanzate meditate riflessioni sulla natura e sulle particolari caratteristiche dei materiali di vario tipo presenti in Internet. [...] Occuparsi di conservazione-selezione-trasmissione di documentazione digitale è - come è noto - cosa ben diversa dall'occuparsi di documentazione cartacea. Infatti quando si ha a che fare con il mondo cartaceo, nel momento stesso in cui si salva la cosa in senso materiale, si salva anche il relativo contenuto. [...] In ambiente digitale invece in una situazione in cui tutto è fluido, evanescente, facilmente manipolabile e soggetto a una continua obsolescenza, conservare vuol dire trasformare. Vuol dire, [...] adottare nuovi formati di memorizzazione. [...] La mancanza di fisicità degli archivi fa sì che non si conservano oggetti, cose, ma piuttosto la capacità di riprodurli. Il che ha ovviamente aperto scenari del tutto nuovi. [...] Al momento i documenti informatici non hanno, almeno in Italia, soppiantato quelli cartacei con cui continuano a convivere più o meno pacificamente. Ma è pressoché certo, soprattutto se si guarda a ciò che sta avvenendo in altri paesi, che in futuro avremo sempre più a che fare con archivi digitali, con archivi cioè costituiti da documenti che nascono, vengono conservati, trasmessi e utilizzati tramite tecnologie informatiche. Già ora tante sono le risorse *on line* di cui possiamo servirci: cataloghi, inventari, strumenti di ricerca in genere, libri, riviste, documenti d'archivio, banche dati, eccetera. Sarebbe perciò inutile, anzi sbagliato, assumere atteggiamenti negativamente sussiegosi nei confronti di quanto si trova sul web e rimanere nostalgicamente attaccati al piacere tattile di posare le mani soltanto su carte polverose, ingiallite, odorose d'antico (anche se il fascino che da esse promana resta tuttora intatto, come Arlette Farge ci ha, e con maestria, descritto nel suo gradevole *Le goût de l'Archive* [del 1989: i dati completi in (Corbin 1991), che lo ha recensito sulle *Annales*, N.d.A.]). A contatto con le risorse di Internet bisogna piuttosto imparare a distinguere tra corretta e attendibile informazione, insidiose manipolazioni e frastornante rumore; occorre perciò impratichirsi sulle strategie di ricerca da seguire e sui modi di leggere e al meglio utilizzare quanto si trova *on line* (Zanni Rosiello 2010, 4-10).

Anche il modo in cui l'autrice conclude ha valore tuttora per quanto concerne il fruire della documentazione:

Ma vorrei ora riprendere il discorso sull'esigenza, espressa nel quaderno di Italia nostra, riguardante il patrimonio archivistico che dovrebbe essere oggetto della "massima fruizione da parte della generalità dei cittadini." In

anni recenti la consapevolezza che la documentazione archivistica possa essere traccia importante per conoscere il passato, sia esso vicino o lontano, non è più diffusa soltanto tra gli storici specialisti o tra gli studiosi in genere che fanno ricorso a questo o quel tipo di documenti d'archivio per approfondire le loro ricerche. Ha raggiunto altre fasce di pubblico, i cosiddetti *amateurs*, interessati a ricostruire la storia della propria famiglia, del paese in cui vivono, della casa in cui abitano, dei luoghi che frequentano, e così via. Un pubblico Interessato in altre parole a mantenere viva la propria personale memoria, a ricostruirsi una propria particolare identità in tempi, come quelli che stiamo vivendo, in cui siamo attaccati, con il rischio di essere inghiottiti, dai tanti tentacoli di una indifferenziata globalità appiattita sul presente. Sono atteggiamenti spesso apprezzabili. Anche se non sempre esenti da insidie e rischi, soprattutto se per identità si intende soltanto il proprio personale vissuto o la costruzione di una rigida armatura che serva a difendersi da altri comportamenti e altre culture. Per cercare di dare risposta alla domanda sempre più diversificata che proviene da parte di un pubblico non più esclusivamente elitario, si parla sempre più spesso di fruizione, di valorizzazione (concetto del resto già presente nei dibattiti politico-culturali degli anni Settanta come risulta anche da molti degli interventi contenuti nel quaderno di Italia nostra). E se ne parla talvolta a proposito, talaltra a sproposito. [...] manifestazioni pensate e costruite per coinvolgere ampi strati di pubblico restano spesso eventi effimeri, senza conseguenze di rilievo [...]. Interventi meno appariscenti, basati su studi severi e su minuziose e prolungate ricerche, sono troppo spesso tacciati di antiquata erudizione e bollati come cose d'altri tempi. Con effetti negativi per una corretta e meditata circolazione di cultura e di saperi. Sarebbe invece, a mio parere preferibile che, al fine di allargare la consapevolezza di tradizioni ancora valide e di estendere conoscenze in parte o del tutto nuove, ci fossero maggiore circolarità e scambi più frequenti tra cultura alta e cultura diffusa, tra ricerca specialistica e attività di valorizzazione [...]. Gli archivi sono una memoria fatta a sua volta di tante e diverse memorie pubbliche e private, segno e traccia della molteplicità dei poteri e della grande varietà dei numerosi soggetti pubblici e privati che li hanno posti in essere. Sono altresì il risultato di una complessa e stratificata sedimentazione connessa ai modi tramite i quali è stato trasmesso "il passaggio del ricordo attraverso le generazioni", per usare una felice osservazione fatta da Marc Bloch... Ma sono anche altro. Se, lasciando da parte il linguaggio specialistico, volessi rappresentare in modo immaginifico le emozioni e le suggestioni che può darci un archivio, ricalcherei le espressioni che Giorgio Manganelli ha usato a proposito della biblioteca. "Una biblioteca è molte, strane, inquietanti cose; è un circo, una balera, una chimera, un incantesimo, una magheria, un viaggio per la terra, un viaggio al centro della terra, un viaggio per i cieli; è silenzio, ed è una moltitudine di voci; è sussurro ed è urlo; è favola, è chiacchiera, è discorso delle cose ultime, è memoria, è riso, è profezia, soprattutto, è un infinito labirinto, ed un enigma che non

vogliamo sciogliere, perché la sua misteriosa grandezza dà un oscuro senso alla nostra vita” (Zanni Rosiello 2010 [10-13])⁶.

Qui cita Manganelli, nel 2017, con *Archivi di scrittori e scrittrici. Note a margine*, regalerà di nuovo al lettore non necessariamente perito di archivistica una sgargiante rassegna sugli archivi di Gadda, Manzini, Meneghelli, Romano: saggio di interdisciplinarietà, altro che *margini adscripta!* Lo scritto andava ad allinearsi a tentativi intrapresi qualche decennio prima (l'archivio fiorentino intitolato a Bonsanti, il centro di Pavia fondato da Maria Corti), e sul versante non umanistico ma tecnico-scientifico aveva trovato un suo cespite di riflessione nel lavoro di Guarasci 2011 su *La Memoria della Scienza*.

Sul versante bibliotecario nel 2010 si ha la prima tappa applicativa, concluso il lavoro teorico, del cosiddetto “Sistema nuovo soggettario”, relativo all'indicizzazione (il più prossimo dei punti di contatto fra le tre arti del moderno Trivio: Archivistica, Bibliografia e Documentazione) che nel febbraio 2007 così era stato salutato a Firenze da Carla Guiducci Bonanni, già direttrice della BNCF (1988-1995) poi sottosegretaria ai beni culturali e ambientali nel governo Dini (gennaio 1995-maggio 1996), all'epoca della presentazione in Palazzo Vecchio assessore alla cultura per il comune di Signa (provincia di Firenze):

Io credo che la storia del Soggettario sia emblematica e importante. Non è cessata mai in questi 50 anni... la prima edizione del Soggettario fu eclatante come questa... È una storia bella una storia che va al di là delle biblioteche perché è una storia che parla di passioni, di interesse, di vero piacere di lavorare, perché credo che il miglior premio che si possa avere sia quello del piacere di lavorare insieme, di ottenere dei risultati, di poter partecipare con gli altri ai risultati e di poter in qualche modo capire che quello che si fa è un passo avanti per il futuro.

Come si è lavorato al cantiere del *Nuovo Soggettario*, fino alla presentazione in pubblico, lo racconta Antonia Ida Fontana (a capo della BNCF dal 1996 al 2010) in un contributo comparso anch'esso in *Viaggi a bordo di una parola* (Lucarelli, Petrucciani, e Viti 2019), cui si può rimandare come a protagonista della vicenda. In (Biblioteca Nazionale Centrale di Firenze, n.d.) si dà il quadro dell'ultimo aggiornamento, dichiarato a marzo 2024, ivi si trova esposta anche la citazione da Guiducci Bonanni; in *Nuovo Soggettario. Manuale applicativo*, curato da M.C. Giunti e Anna Lucarelli (2024) (attuale reggente di BNCF) si legge invece il protocollo applicativo. L'ultimo aggiornamento cartaceo risale, al momento, alla coedizione AIB (Associazione Italiana Biblioteche) - BNCF di *Nuovo Soggettario* (Biblioteca Nazionale Centrale di Firenze 2021).

⁶ Sono fra [] le indicazioni di pagina riferite al .pdf tratto dalla fonte.

L'esperienza in quel cantiere ha propiziato in chi ne ha fatta pratica come collaboratore, tirocinante, o studente in visita alcuni contributi di vita propria: così è per (Calvitti 2007a; 2007b; 2008) (cui si è già accennato); così è per (Grillo 2008), tesi di dottorato di analisi semantica di materiale retrospettivo (leggi e bandi di antico regime di BNCRVE2), poi divenuta ampia e sistematica raccolta in (Grillo 2014), con ulteriori contributi esplicativi, anche a circolazione digitale, qui non segnalati. Così è per (Lucarelli 2008), che si occupa di trattamento conservazione e descrizione di materiale fotografico, sviluppando un suo intervento a un convegno internazionale di qualche anno prima (Bolzano, 2005); chi ricorda i tempi di ICONCLASS, tardi anni Ottanta, e conseguenti trattazioni delle immagini (su cui faceva il punto (Rossi 1996)), coglie il senso dell'occuparsi di tale materia. La quale, mai considerata in (Casanova 1928), ha in (Perrella 1963) ≥ 10 occorrenze; altrettante in Documentazione in azienda 1965; si dilata in (Brenneke 1968) che ne presenta sì ≥ 10 , ma una di loro contiene il rimando a 25 ulteriori item (pp. 576-577, no. 1368-1392); in (Innocenti 1997): ≥ 10 ; in (Massabò, Ricci, e Carassi 2000): ≥ 1 ; in (D'Angiolini 2002): ≥ 5 ; in (Guarasci 2008): ≥ 15 ; in (Cavallaro 2008) ≥ 60 (la gran parte concentrate nel saggio di Lucarelli menzionato appena sopra); (Gatto 2011) analizza la presenza nel Web e formula una ipotesi di metaopac (successivamente e da altri realizzati, in varie forme); in *Meminisse iuvabit* (Flöss e Franzoi 2022) (la miscellanea Chistè su cui più avanti) ≥ 10 . Si può parlare di una linea di tendenza all'emersione ormai e da tempo definitiva, se si pensa che già in *Memoria lunga* (Bertolucci e Pensato 1985) se ne possono constatare i prodromi, accanto all'esigenza non solo di archiviare ma anche di "archivizzare" (se posso di nuovo giocare con le parole) le fonti orali e l'etnomusicologia: a Carpitella 1985, colpevolmente sfuggito alla bibliografia degli scritti del musicologo uscita nella «Nuova rivista musicale italiana», del 1992 (26, no. 3-4, pp. 523-572): caso classico di trasversalità mancata, ce ne sono a centinaia, non è un delitto capitale, solo un segno di sciatteria.

4. Un tuffo (con cautela) nel digitale

Produzione, uso e conservazione del documento digitale sono temi oggi correntemente all'ordine del giorno, trattati per lo più all'interno di singole prospettive disciplinari che ne valorizzano singole faccette di maggior interesse specifico. A questo nodo si rivolge (Alfier 2020) (attività di archivista informatico dal 2009), proponendo un approccio interdisciplinare al documento nativo digitale: caratteri strutturali, sua trattazione nella teoria del diritto, nella documentazione, nell'archivistica, persino nelle riflessioni più recenti della diplomatica, che comincia a interessarsi al più ampio sistema di documentazione, condizionato dall'ordinamento giuridico, dal contesto sociale e culturale e

dallo scenario tecnologico. Si disegna insomma lo stato dell'arte del contemporaneo sistema della documentazione digitale, senza illudersi circa potenzialità e criticità. Tutto ciò, a parere di chi scrive, è la necessaria formalizzazione in atto a scopi di pubblica utilità a magari anche istituzionalizzati (come è intrinseco al ruolo di "documento" dal punto di vista giuridico, valido cioè *erga omnes*) di una tentazione risalente a molto prima di oggi, in aree lontane da quella istituzionale, attinenti piuttosto alla volontà di libera espressione organizzata da ipotesi collettive autogestite in forma più o meno riservata, per non dire segreta.

Wikipedia nasce nel 2001 ed è l'enciclopedia libera che liberamente e gioiosamente usiamo: queste pagine, sicuramente il wikinauta se ne sarà accorto, altro non sono che un gioco di Lego con materiali da essa derivati. Nella sua pagina di accoglienza, versione 6 Aprile 2024 h 10:24, si leggeva: «Chiunque può contribuire alle voci esistenti o crearne di nuove, affrontando sia gli argomenti tipici delle *enciclopedie tradizionali*, &c.» e fermiamoci qui perché abbiamo assistito non solo a una rivendicazione di genealogia diretta da Diderot e d'Alembert, fondatori di quella forma di pubblicazione che ben presto si consolidò come primo genere di opere generali trattato nelle bibliografie (Lucarelli (2011) dà al lettore una simpatica lista "semiseria" di enciclopedie e di thesauri); ma anche, implicitamente, ad esempio all'altissima operazione intellettuale del gruppo matematico Nicolas Bourbaki (pseudonimo scelto a burla), nato nel lontano 1933 e fiorente di pubblicazioni di alto livello almeno fino al 2016, dove il sapere di gruppo veniva privilegiato sul contributo individuale. Bella genealogia: ma se è vero che il lavoro collegiale è garanzia di convergenza di più competenze, viceversa l'anonimato del prodotto finale sempre in movimento e *in fieri* (che visibilmente ricalca l'*Enciclopedia galattica* nella *Trilogia della Fondazione* di Isaac Asimov) non è affatto un passo avanti verso la garanzia di autorevolezza. Se l'anonimato non è di per sé falso o di cattiva qualità, stanno però nell'inalterabile marmo cartaceo del libro sia il prodotto del marchio Diderot & d'Alembert, sia quello di Bourbaki, mentre certo inquina il percorso documentale attraverso cui si è arrivati alla conclusione il prodotto di Wikipedia. Lo so bene che è percorribile, ma non è immediatamente sotto gli occhi del lettore. Non è una critica, ma una riflessione: anche le norme giuridiche sono pubbliche (se no, non valgono), ma quando vanno avanti per autocitazione a catena di norme precedenti, alla terza il lettore abbandona; *bon gré malgré*, si rivela quale tecnica di occultamento.

Un bel caso di organizzazione documentale autogestita e mimetizzata nel campo delle culture del testo tentò OuLiPo, ovvero Ouvroir de Littérature Potentielle, anno di nascita, in ambito cartaceo, il 1960, fondatore Raymond Queneau con François Le Lionnais, divinizzato dai risultati qualitativi raggiunti e dall'ingresso in enciclopedie ufficiali, come la nostra

Treccani: partecipava anche Perec “gnomo e cabalista” (così (Calvino 1982) su un quotidiano). Nel 1985 auspice una pubblicazione dell’Università di Bologna, nasce la traduzione italiana del circolo francese: “Opificio di Letteratura Potenziale”, siglata in OpLePo, nel 1990 il gruppo ha consacrazione notarile. Subito attratto dai preavvisi della seduzione digitale, vi transita, ma il suo sito sembra ora aggiornarsi a ritmo lento (Ouvroir de littérature potentielle, n.d.)⁷. La per ora abbastanza ricca documentazione viaggia appunto sul sito, che si fregia a significante di una remota frase di Calvino, nel frattempo scomparso: «Ogni esempio di testo costruito secondo regole precise apre la molteplicità ‘potenziale’ di tutti i testi virtualmente scrivibili secondo quelle regole, e di tutte le letture virtuali di quei testi» (*La filosofia di Raymond Queneau*, Calvino 1981, ora in Calvino 1995, 1429). L’oplepiano (così si autodefiniscono gli adepti) che ha ripescato la citazione dal saggio ha così rivendicato ulteriormente l’intenzione di OpLePo di prolungare la strada OuLiPo (già del resto lo dice il nome tradotto...).

Scendendo dall’Empireo delle idee alte, dove stanno Illuminismo, speculazione matematica, letteratura di rango, l’idea dell’autorialità eteronima anche in tempi recentissimi fu fatta propria da altri pseudonimi collettivi quali Luther Blissett, oppure Wu Ming, per stare ai più noti e radicati dentro il digitale: il primo, sbucato non *post* 1994, anno della canonizzazione iconica, ha fatto parlare di sé lasciando innumerevoli tracce, anche giudiziarie, almeno fino al 2015; il secondo tuttora lo si cita (e si autocita) perché nel 1999 NON andò in cinquina per lo Strega (Ming 2019): ci hanno riflettuto almeno (De Pascale 2009; Maniscalco 2019), e grande folla di occorrenze emerge da una ricerca onomastica nel Web⁸.

Indagare su fenomeni che in talune circostanze possono persino riguardare procedimenti di attività giudiziaria è materia che si fa fluida, anche dal punto di vista ideologico, e porta nell’area trattata da *Le forme del falso* (Caporale et al. 2022), dove è peculiarmente orientato il saggio di (Orsini 2022), “Falsitates et fraudes omnes evitare. Il ruolo dell’archivio nella produzione del falso storico”. E qui non posso non ricordare che dedicai la mia prolusione universitaria, 1988 (Innocenti 1989), al *double lie* della spia, falsificatrice della propria identità per scoprire la Verità, a fini superiori. In campo istituzionale (ammesso e non concesso che vi possa essere qualcosa

⁷ Controllato il 14 Marzo 2024, l’ultimo documento registrato è del Giugno 2023.

⁸ Un grande nome della filologia classica ottocentesca, che di tradizione frammentaria se ne intendeva, affermò senza battere ciglio che se di Eschilo sono sopravvissute solo sette tragedie, su 73 titoli attribuiti, ciò significa solo che evidentemente erano le più belle. Protetto da questa non discutibile *auctoritas*, me la sento di postillare (anche se solo in nota) che tormentarsi per conservare tutta la documentazione nata digitale potrebbe anche costituire, alla lunga, un falso problema: da Eschilo ci separano due millenni e mezzo circa, ma continuiamo a goderne...

di più da sempre istituzionale dello spionaggio), vorrei annotare che l'incontro quasi casuale con una pagina Web della Soprintendenza archivistica e bibliografica dell'Emilia Romagna (18 marzo 2024) mi ha regalato una carrellata retrospettiva di grande interesse, che si apre elencando trattazioni all'incirca dell'ultimo quindicennio spiegandone così la presenza nella pagina contigua:

Vengono proposti in questa sezione strumenti di ricerca elaborati su iniziativa della Soprintendenza archivistica e bibliografica dell'Emilia-Romagna, strumenti di lavoro, atti di convegni organizzati dalla Soprintendenza anche in collaborazione con altri enti e istituzioni, interventi e materiali vari relativi all'attività di tutela, di valorizzazione, di studio e di elaborazione scientifica svolta dalla Soprintendenza, dai suoi funzionari e dai suoi collaboratori. Alcuni dei materiali, nella versione presentata qui o in successive rielaborazioni, sono stati anche pubblicati a stampa. Altri invece sono inediti. Di ciò si dà debitamente conto nella presentazione o nelle note dei contributi stessi (Soprintendenza archivistica e bibliografica dell'Emilia-Romagna, n.d.).

La cronologia della pagina è presentata dal produttore in ordine cronologico inverso, allo scopo di offrire subito al consultante il dato più recente; qui invece diamo sunto in cronologia diretta dal vecchio al nuovo degli argomenti trattati e loro autori: l'archivio Vitali Verga, fondo plurisecolare dal Duecento al primo Novecento nell'Archivio di Stato di Parma (Michelangelo Ferraro), 2010; recensione del saggio di Orietta Filippini su Carlo Cartari (1614-1697) e l'Archivio di Castel Sant'Angelo (Linda Giuva), 2010; il documento digitale come fonte di prova forense, trattato in .ppt con inusitata chiarezza (Luciana Duranti), 2010; storia globale e riviste locali (Mauro Maggiorani), 2012; i poli archivistici (Stefano Vitali), 2013; Salvatore Alongi e Leonardo Mineo presentano ambedue il manuale di Giuva-Guercio, 2014; archivî come bene culturale (M. Maggiorani), 2014; didattica in archivio e storia locale (M. Maggiorani), 2014; archivî emiliani dopo il terremoto (S. Vitali), 2014; la Resistenza negli archivî femminili dell'Emilia Romagna (Maria Lucia Xerri), 2014; documenti e archivî nella sanità elettronica (curatore Stefano Pigliapoco), 2016; musei di macchine (Giampiero Romanzi), 2017; le fonti per la storia del '68 a Bologna (M. Maggiorani), 2018; profili di tutela e strategie archivistiche (G. Romanzi), 2018; pratiche digitalizzate negli archivî comunali (G. Romanzi), 2019; un archivio d'impresa floro-vivaistica (M. Maggiorani), 2022. La cronologia della pagina è presentata dal produttore in ordine cronologico inverso, allo scopo di offrire subito al consultante il dato più recente; qui invece diamo sunto in cronologia diretta dal vecchio al nuovo degli argomenti trattati e loro autori: l'archivio Vitali Verga, fondo plurisecolare dal Duecento al primo Novecento nell'Archivio di Stato di Parma (Michelangelo Ferraro), 2010; recensione del saggio di Orietta Filippini su Carlo

Cartari (1614-1697) e l'Archivio di Castel Sant'Angelo (Linda Giuva), 2010; il documento digitale come fonte di prova forense, trattato in .ppt con inusitata chiarezza (Luciana Duranti), 2010; storia globale e riviste locali (Mauro Maggiorani), 2012; i poli archivistici (Stefano Vitali), 2013; Salvatore Alongi e Leonardo Mineo presentano ambedue il manuale di Giuva-Guercio, 2014; archivi come bene culturale (M. Maggiorani), 2014; didattica in archivio e storia locale (M. Maggiorani), 2014; archivi emiliani dopo il terremoto Stefano Vitali), 2014; la Resistenza negli archivi femminili dell'Emilia Romagna (Maria Lucia Xerri), 2014; documenti e archivi nella sanità elettronica (curatore Stefano Pigliapoco), 2016; musei di macchine (Giampiero Romanzi), 2017; le fonti per la storia del '68 a Bologna (M. Maggiorani), 2018; profili di tutela e strategie archivistiche (G. Romanzi), 2018; pratiche digitalizzate negli archivi comunali (G. Romanzi), 2019; un archivio d'impresa floro-vivaistica (M. Maggiorani), 2022. È evidente il florilegio di temi, che necessitano di semantica per essere collegati fra sé eventualmente, o ad altro, antecedente, contestuale o successivo che sia; e dove. La cronaca è una pesca a strascico nel calendario che mette in fila il pescato; altro è individuare, i fili d'Arianna che colleghino ad altre piste, a partire da altri luoghi di produzione: però è comunque un punto di partenza, che evidenzia come i segnali di novità (anche sottoforma di criticità sottovalutata) si vadano a nascondere in sedi appartate, seppelliti da materiali interessanti sotto altri punti di osservazione, allotrì quanto ad essenza.

Tornando al 2022: è l'anno di *Non solo carta*; esce in Trentino, che tratta di *Gestione documentale e conservazione degli archivi digitali* (Bortoli, Boschi, e Lazzeri 2022); così la prefatrice, M. Guercio:

Il volume presenta un'analisi attenta del lavoro svolto negli ultimi dieci anni per la formazione e gestione degli archivi digitali sia a favore delle strutture provinciali sia a supporto degli enti locali. Gli autori si misurano con tutti i problemi aperti che la lunga esperienza condotta a Trento in questo campo ha reso evidenti nel corso del tempo: Carlo Bortoli responsabile del capitolo introduttivo sulle difficoltà e sulle incertezze della conservazione digitale e della successiva analisi dell'evoluzione e dei requisiti dei sistemi P.I.Tre. per la gestione documentale, Judith Boschi autrice del capitolo sul piano unico di conservazione degli atti della Provincia autonoma e Annamaria Lazzeri che approfondisce il difficile nodo dello scarto in ambiente digitale. [...] Emergono con nettezza i principi su cui poggia l'intera progettazione del sistema documentario provinciale, gli stessi peraltro che la disciplina archivistica ha sempre affermato e che trovano robuste conferme nella presentazione in questa sede delle soluzioni realizzate in questi anni: – la necessità di disporre di una visione dinamica dell'archivio anche quando si opera per garantire la necessaria persistenza ai suoi oggetti e ai loro reciproci legami; – l'esigenza di pianificare gli strumenti tenendo sempre conto del quadro complessivo che collega le buone pratiche, la normativa, i bisogni operativi

delle persone e delle istituzioni; – il riconoscimento che l’implementazione applicativa ha costante bisogno di confronti, di verifiche, di monitoraggio. Alla conservazione digitale nella forma attuale stabilita dal legislatore gli autori riconoscono, giustamente, un ruolo non autonomo rispetto alla formazione dei documenti e delle aggregazioni: un intervento di salvaguardia e protezione precoce che – pur non garantendo ancora la conservazione permanente degli archivi – assicura un presidio sia pure iniziale e ancora da perfezionare in collaborazione con altri soggetti pubblici (Guercio 2022, IX-XIII: IX-X).

L’*Introduzione* di Carlo Bortoli a proposito del mestiere dell’archivista conclude fra realismo ed ottimismo: «La conservazione digitale è oggi tra i temi principali [...] La trasformazione digitale [...] non può essere ignorata ma può essere opportunamente sfruttata. In questo senso l’accentuazione delle istanze legate alla dimensione tecnologica non ha per fortuna provocato l’estinzione della specie degli archivisti» (Bortoli 2022, 10). Impressiona anche in calce al volume (pp. 127-129) la tavola di ben 28 precedenti titoli della serie «Archivi del Trentino: fonti, strumenti di ricerca e studi», nata nel 1999, cui appartiene questo efficace 29esimo numero. Al no. 27 appena menzionato era uscito, datato anch’esso 2022 ma finito di stampare nell’ultimo mese del 2021, *Meminisse iuvabit* (Flöss e Franzoi 2022), volume miscelaneo per onorare l’ottantesimo compleanno di Pasquale Chisté (l’ho già menzionato come fonte), una vita al servizio dell’Amministrazione provinciale di Trento vissuta da funzionario di raffinata cultura, solido buon senso e grande competenza del settore archivistico-bibliotecario, dove è arrivato fino al grado massimo della carriera, congedandosene nel 2004. Dell’opera, accanto a interventi su archivi del passato vicino e lontano, interessanti ma remoti ai fini del discorso che stiamo andando a concludere, vorrei proporre due *margini adscripta* a luoghi di maggiore attualità per esprimere loro il mio consenso.

Il primo è (Carucci 2022, 79): l’argomentazione verte su una tecnicità giuridica, quale si riflette nel titolo del saggio: “La tutela dei dati personali negli Archivi storici dopo l’approvazione del Regolamento europeo”, e si sviluppa per una ventina di pagine, molto dense; è preceduta da un atto di sfiducia, quasi una catilinaria, nei confronti degli sforzi di standardizzazione a oltranza del linguaggio, attraverso l’incolta onnipresenza di una lingua inglese ridotta ai minimi termini di come si parla all’accettazione in aeroporto:

L’approvazione del *Regolamento europeo* in materia di tutela dei dati personali, approvato nel 2016, ha introdotto alcune modifiche alla normativa italiana che ne hanno reso necessario un adeguamento alle disposizioni europee. Queste, come è noto, prevalgono sulla normativa nazionale. | Il testo europeo è molto complesso e la pessima traduzione in italiano, che sembra ignorare il linguaggio tecnico dell’archivistica, ne rende più difficile la com-

preensione. Dobbiamo pertanto familiarizzare con espressioni inconsuete che è necessario cercare di equiparare a concetti che siamo abituati a esprimere in maniera diversa.

Si ripete qui la critica già formulata da (Lodolini 1987) su le incertezze terminologiche di (*DAT* 1984), troppo disattento a quelli che l'interpretariato interlinguistico chiama oggi i "falsi amici". Si spera, non senza pessimismo, che questo richiamo, del tutto sottoscrivibile, abbia migliore sorte. Nella stessa pubblicazione (Giorgi 2022, 178-179n), dal canto suo riprende considerazioni da lui già presentate nel 2020 al seminario, *La crisi degli archivi, un'emergenza democratica*, organizzato da ANAI e da AIDUSA, e valorizza due efficaci citazioni, che nella grande distanza di tempo intercorrente fra loro, formulano lo stesso richiamo:

Penso alle prime pagine di Croce, *Teoria e storia della storiografia*, ove tra le altre riflessioni si segnala: "Una storia senza relazione col documento sarebbe una storia inverificabile; e poiché la realtà della storia è in questa verificabilità, e la narrazione nella quale si viene concretando è narrazione storica solo in quanto è esposizione critica del documento (intuizione e riflessione, coscienza e autocoscienza, ecc.), una storia di quella sorta, priva di significato e di verità, sarebbe inesistente in quanto storia" [sta citando (Croce 1989, 1) N.d.A.]; più in generale, sul tema rimanda a (D'Angiolini e Pavone 1973), 1672 e ss., e continua: Sul "rischio concreto che, dopo secoli di attenta conservazione della documentazione archivistica, si possa spegnere letteralmente la luce in luoghi simbolo dell'Italia migliore, lasciando spazio a chi ritiene che si possa fare storia senza documenti, reinventare il passato o vivere in un eterno presente, senza memoria" ha posto recentemente l'attenzione [e cita *Chiudiamo gli archivi* (Russo 2019) N.d.A.].

Ancora prima, nel no. 23 della collana, *Oblío, tempo, cultura ed etica* (Guastalla e Lazzeri 2019), era stata pubblicata una *satira lanx* di saggi e riflessioni presentati in convegno da ANAI nell'arco di tempo 2015-2018. Fra le ≥ 25 occorrenze di 'documentazione' che si trovano al suo interno, normalmente indicando il materiale, un paio di volte la disciplina (la riflessione sulla quale però sottende a molte delle pagine), è possibile risalire, p. 28n., ad un riferimento alla giurisprudenza sul diritto digitale quale in (Pascuzzi 2010) (autore presente in rassegna a proposito di altri argomenti). Ma anche all'esposizione così persuasiva di Pettenella, circa la diffrazione fra due discipline:

L'arte attinge senza sosta al materiale d'archivio, proprio come accadeva a Boccioni; ma testimonia il rapido moltiplicarsi dei *media*, della materia e dei supporti che hanno dato forma alla documentazione, dal secolo scorso alla contemporaneità. Guardando al gigantesco patrimonio documentario degli archivi novecenteschi, gli artisti contemporanei sembrano esprimere qualche parola significativa nel delicato passaggio dalla memoria individuale all'espe-

rienza universale (Pettenella 2019, 137).

Concludo ricordando che il 21 dicembre 2023 è stata data la notizia (*to whom it may concern*) che il 10 novembre è venuto alla luce RiC 2023 (Records in Contexts–Conceptual Model (International Council on Archives, Expert Group on Archival Description 2023)). Riporto soltanto le ultime righe del testo che definiscono in sunto la non banale ambizione:

RiC-CM reconciles, integrates, builds on, and replaces the four existing standards: General International Standard Archival Description (ISAD(G)); International Standard Archival Authority Records–Corporate Bodies, Persons, and Families (ISAAR(CPF)); International Standard Description of Functions (ISDF); and International Standard Description of Institutions with Archival Holdings (ISDIAH). RiC-CM 1.0 can be downloaded from <https://www.ica.org/app/uploads/2023/12/RiC-CM-1.0.pdf> or <https://github.com/ICA-EGAD/RiC-CM/blob/master/current-version/RiC-CM-1.0.pdf>. Anyone interested in learning more about or discussing RiC with others is encouraged to join the RiC user group (International Council on Archives, Expert Group on Archival Description 2023).

Dai molti, l'uno, dunque: non male, se funziona. Tutto ho letto, non tutto ho capito, però mi sono sentito in famiglia; se non fratello, cugino, magari di secondo grado o acquisito, ma ben disposto grazie a un tirocinio ormai lungo. Quando ho cominciato a bazzicare biblioteche, le tavole della Legge erano le “regole del 1956” nonché il *Manuale del catalogatore*; le RICA (Regole Italiane di Catalogazione per Autori) erano in gestazione, e REICAT (Regole Italiane di CATalogazione) in mente Dei; la revisione della Dewey era in corso. Il *Nuovo Soggettario*, dai primi Duemila in poi, ha fornito nuove piste da seguire, per non parlare della lunga serie di ISBD (International Standard Bibliographic Description), di uno dei quali nel lontano 1979 fui, per la mia piccolissima parte di componente di vasta commissione internazionale, coinvolto. Chi a sua volta ha bazzicato archivi, ha frequentato le sigle di cui alla citazione appena offerta, in procinto di essere amalgamate. E così anche il prossimo paio di generazioni di professionisti, magari di nuovo accomunati, come i predecessori, sotto l'unica dicitura “Archivi e Biblioteche”, potrebbe essere stato fornito di materia su cui formarsi, o anche discutere: talvolta, ho paventato, anche di lana caprina.

Frequentiamo per l'ultima volta una saggezza antica: Ovidio enuncia il concetto di “autolesionismo” giustapponendo a un emistichio finale l'emistichio iniziale successivo: *Video meliora proboque | deteriora sequor*. È ben noto che lo misero in endecasillabi sia Petrarca sia Foscolo. Tornò a voltarlo in endecasillabo, giocando di nuovo come il predecessore augusteo, su due emistichî diversi (“Vedo il bene, l'approvo, | e seguò il male”), anche l'autore di una traduzione: siccome è rimasto a me ignoto, a scanso di sanzioni, mi affretto

a citare il copyright del sito di Rete che me ne ha informato: “© 2000-2005 Miti3000.it”. Amen.

5. «Quien sabe? répondit Thalcave»

Le pagine che vado a chiudere son state aperte da una canzone di Trenet, intimistica, che l'anno dopo (1943) fu seguita da una più “pubblica”, *Douce France*. Alle orecchie di un ragazzo degli anni Cinquanta/Sessanta ambedue furono spazzate via, roba da vecchie zie a favore per le tappe che da *Rock Around the Clock* (1954), *Oh, Carol!* (1959), *The Lion Sleeps Tonight* (1961) portano a Beatles e Rolling Stones, e infine alla maggiore età. Quell'ormai post-adolescente forse recuperò con atto di contrizione *Que reste-t-il* ascoltandola nel *Baisers volés* di Truffaut, per richiamare la privatissima mancanza di futuro della post-modernità. Mi sono sentito ancora come quell'ipotetico adolescente (non lo ero, ma talvolta *de te fabula narratur*) quando mi imbattei in *Ce que nous devons défendre*, di Hazard, scritto e pubblicato di getto nell'anno della mobilitazione generale: leggera *brochure* di cartoncino, 12 facciate in tutto. Lo feci mio casualmente nel novembre 1980, per pochissimi fiorini ungheresi, da un venditore di libri usati, sul Bastione dei pescatori, collina di Buda, vicino al Dipartimento universitario di Storia, dove eravamo in visita Sergio Bertelli, Franco Gaeta, Carlo Pincin e io. Tibor Klaniczay, che aveva organizzato l'incontro per il quale eravamo in missione, me lo vide in mano, lodò l'acquisto, declinò la immediata mia disponibilità a offrirglielo in dono, e ci raccontò a tavola che aveva letto da ragazzo *La Crise* in prima edizione, appassionandosi durevolmente, e parlò diffusamente dell'importanza di quel testo patriottico, senza farmi pesare l'ignoranza, per altro da me non negata.

Da allora in poi mi fu chiaro quanto dovesse essere stata significativa la sequenza di quelle due canzoni (la seconda bandiera esplicita) del 1942-43, seguite a giro da *La mer* che nel 1946 inondò il dopoguerra della Francia tornata libera. E mi resi conto anche che a quella domanda sul futuro posta dal titolo della prima aveva già risposto, forse in modo ricorsivo, un altro documento importantissimo della cultura popolare francese, e da lì europea: «– Ne nous reverrons-nous jamais? – *Quien sabe? répondit Thalcave*, en levant son bras vers le ciel. Ce furent les dernières paroles [...], qui se perdirent dans le souffle du vent » (Verne 1868, 214).

Magari il vento le ha portate, le parole, a Bob Dylan, per la sua fantastica lirica del 1963: chi sa...

Riferimenti bibliografici

- “Ai lettori.” 2000. *Aidainformazioni*, anno 18, no. 1 (gennaio-marzo): 1-2.
- «AIDAinformazioni». 2003. anno 21, no. 3 (luglio-settembre).
- Alberani, Vilma, e Elisabetta Poltronieri. 2003. “La Documentazione rispetto alle altre discipline dell’informazione.” *AIDAinformazioni*, anno 21, no. 3 (luglio-settembre): 19-45.
- Alfieri, Alessandro. 2020. *Il sistema di documentazione digitale*. Milano: Editrice Bibliografica.
- Alongi, Salvatore. 2014. “Caratteri della manualistica archivistica in Italia nel XX secolo.” <https://sab-ero.cultura.gov.it/risorse/materiali/interventi>.
- Ardolino, Enrico Pio. 2022. “Fumagalli, Giuseppe.” In *Dizionario bio-bibliografico dei bibliotecari italiani del XX secolo*. AIB. <https://www.aib.it/aib/editoria/dbbi20/fumagalli.htm>.
- Balbis, Bruno. 1950. “La documentazione.” In *Il libro e le biblioteche: Atti del Primo Congresso bibliologico francescano internazionale: 20-27 Febbraio 1949*, vol. 1 di *Conferenze di carattere generale*, 261-319. Romae: Pontificium Athenaeum Antonianum.
- Balbis Bruno. 1951. *L'insegnamento professionale della documentazione in Italia*, a cura del Consiglio nazionale delle ricerche, Centro nazionale di documentazione scientifico tecnica. Roma: CNR.
- Baldazzi, Anna, e Maria Pia Carosella. 2003. “AIDAvant’anni: Alcune date da ricordare.” *AIDAinformazioni*, anno 21, no. 3 (luglio-settembre): 13-16.
- Baldazzi, Anna. 2003. “Bibliografia AIDA sulla professione.” *AIDAinformazioni*, anno 21, no. 3 (luglio-settembre): 117-23.
- Bartoloni, Marzio. 2017. “«3+2», una riforma fallita.” *Il Sole 24 Ore*, 8 giugno 2017. <https://m.flcgil.it/rassegna-stampa/nazionale/3-2-una-riforma-fallita.flc>.
- Basili, Carla, e Domenico Bogliolo, a cura di. 2003. *AIDAvanti: Vent’anni di AIDA: La documentazione fra teoria e applicazioni: Atti del 7 convegno nazionale AIDA*. Roma: AIDA.
- Basili, Carla, e Emanuela Reale. 2003. “Documentazione scientifica e documentazione sulla scienza nella missione dell’ISRDS.” *AIDAinformazioni*, anno 21, no. 3: 71-91.
- Basili, Carla, a cura di. 2003a. *Information Literacy in Europe 2003. A First Insight into the State of the Art of Information Literacy in the European Union*. Roma: CNR.

- Basili, Carla. 2003b. "Introduzione." *AIDAinformazioni*, anno 21, no. 3: 11.
- Basili, Carla. 2003c. "Tappe salienti della Documentazione nel periodo 1983-2003 nella rassegna dei convegni AIDA, ASIS e FID." *AIDAinformazioni*, anno 21, no. 3: 105-15.
- Bertolucci, Paola, e Rino Pensato, a cura di. 1985. *La memoria lunga: Le raccolte di storia locale dall'erudizione alla documentazione: Atti del Convegno realizzato in collaborazione con l'Istituto superiore regionale etnografico di Nuoro e l'Associazione italiana biblioteche, sezione sarda. Cagliari, 28-30 aprile 1984*. Milano: Editrice Bibliografica.
- Biblioteca nazionale centrale di Firenze (BNCF). 2021. *Nuovo soggettario: Guida al sistema italiano di indicizzazione per soggetto*, 2. ed. interamente rivista e aggiornata, a cura di Anna Lucarelli. Roma - Firenze: AIB - BNCF.
- Biblioteca Nazionale Centrale di Firenze (BNCF). n.d. "Nuovo soggettario." Consultato il 14 marzo 2024. <https://thes.bncf.firenze.sbn.it/>.
- Bogliolo, Domenico. 2003. "La Martinelli. Un incontro." *AIDAinformazioni*, anno 21, no. 1 (gennaio-marzo): 95-97.
- Bortoli, Carlo, Judith Boschi, e Annamaria Lazzeri. 2022. *Non solo carta: Gestione documentale e conservazione degli archivi digitali della Provincia autonoma di Trento*. Prefazione di Mariella Guercio. Archivi del Trentino: fonti, strumenti di ricerca e studi 29. Trento: Provincia autonoma di Trento. Ufficio beni archivistici, librari e Archivio provinciale.
- Bortoli, Carlo. 2022. "Introduzione." In *Non solo carta: Gestione documentale e conservazione degli archivi digitali della Provincia autonoma di Trento*. Prefazione di Mariella Guercio. Archivi del Trentino: fonti, strumenti di ricerca e studi 29, 1-10. Trento: Provincia autonoma di Trento. Ufficio beni archivistici, librari e Archivio provinciale.
- Brenneke [*aliter*: Brennecke], Adolf. 1953. *Archivkunde: Ein Beitrag zur Theorie und Geschichte des Europäischen Archivwesens*. Bearbeitet nach Vorlesungsnachschriften und Nachlaßpapieren und ergänzt von Wolfgang Leesch. Leipzig: Kohler & Amelang (poi: 1970; Saur Verlag 1988, 1993).
- Brenneke, Adolf. 1968. *Archivistica. Contributo alla teoria ed alla storia archivistica europea*. Tradotto da Renato Perrella. Archivio della Fondazione italiana per la storia amministrativa 1, Monografie, ricerche ausiliarie, opere strumentali 6. Milano: Giuffrè.
- Boustany, Joumana. 1991. "La bibliologie en 1950 en Italie." *Bulletin d'Informations Internationales de Bibliologie* 14, 4. trim. <https://web.archive.org/web/20080905172043/http://www.docinfos.com/Articles/7-Articles/558-la-bibliologie-en-1950-en-italie>.

- Buizza, Pino. 2019. "Indicare quasi la stessa cosa: Appunti di indicizzazione interlinguistica." In *Viaggi a bordo di una parola: scritti sull'indicizzazione semantica in onore di Alberto Cheti*, a cura di Anna Lucarelli, Alberto Petrucciani e Elisabetta Vinci, 33-49. Roma: AIB.
- Calvino, Italo. 1981. *La filosofia di Raymond Queneau*. Torino: Einaudi.
- Calvino, Italo. 1982. "Perec, gnomo e cabalista." *La Repubblica*, 6 marzo 1982.
- Calvino, Italo. 1995. "La filosofia di Raymond Queneau." In *Saggi 1945-1985*, a cura di Mario Barenghi, 1410-30. I Meridiani, 2 v. Milano: Mondadori.
- Calvitti, Tiziana. 2007a. "Digitalizzazione ed indicizzazione di alcuni periodici di discipline del libro: Un progetto di thesauro." Tesi di dottorato di ricerca, Università di Udine.
- Calvitti, Tiziana. 2007b. "Un progetto di tesoro di discipline del libro." In *Una Mente colorata: Studi in onore di Attilio Mauro Caproni per i suoi 65 anni*, a cura di Cristina Cavallaro, studi promossi, raccolti, ordinati da Piero Innocenti, 465-84. Manziana: Vecchiarelli. Il libro e le letterature.
- Calvitti, Tiziana. 2008. "Le fonti nei vocabolari controllati: Il caso del Nuovo Soggettario." In *Piccoli scritti di biblioteconomia per Luigi Crocetti (10 marzo 2007-10 marzo 2008)*, a cura di Cristina Cavallaro, promossi, raccolti, ordinati da Piero Innocenti. 3-33. Manziana: Vecchiarelli.
- Calvitti, Tiziana. 2011. "Da «Culture del testo» a «Culture del testo e del documento»: Un quadro complessivo della rivista." In *Books seem to me to be pestilent things: Studi in onore di Piero Innocenti per i suoi 65 anni*, a cura di Varo A. Vecchiarelli e Cristina Cavallaro, 851-61. Manziana: Vecchiarelli.
- Canna, Franco. 2020. "L'Italia si dota di un Comitato Nazionale per la Produttività." *Innovation Post*, 13 novembre 2020. <https://www.innovationpost.it/attualita/politica/litalia-si-dota-di-un-comitato-nazionale-per-la-produttivita/>.
- Caporale, Marina, Cristina Demaria, Daniela Donato, Anna Maria Lorusso, e Francesco Mazzucchelli. a cura di. 2022. *Le forme del falso*. Bologna: University Press. https://icar.cultura.gov.it/fileadmin/risorse/News/54_2022_Forme_falso.pdf.
- Carosella, Maria Pia. 2003. "AIDA e ADBS: Elementi per una 'documentazione comparata'". *AIDAinformazioni*, anno 21, no. 3 (luglio-settembre): 93-101.

- Carucci, Paola. 2022. "La tutela dei dati personali negli Archivi storici dopo l'approvazione del Regolamento europeo." In *Meminisse iuvabit studi in onore di Pasquale Chistè*, a cura di Lydia Flöss e Stefania Franzoi, 79-102. Archivi del Trentino: fonti, strumenti di ricerca e studi 27. Trento: Provincia autonoma di Trento Soprintendenza per i Beni culturali, Ufficio Beni archivistici, librari e Archivio provinciale.
- Casamassima, Emanuele. 1965. "La soggettazione." In *Documentazione in azienda 2*, a cura del Servizio Informazione e Diffusione del CNP, 235-56. Roma: Servizio Informazione e Diffusione del CNP.
- Casanova, Eugenio. 1928. *Archivistica*, Roma (Siena, Arti Graf. Lazzeri), XVI.
- Castellucci, Paola. 2003. "Paolo Bisogno: La precoce fondazione della disciplina della Documentazione." *AIDAinformazioni*, anno 21, no. 3 (luglio-settembre): 59-70.
- Cavallaro, Cristina, a cura di. 2008. *Piccoli scritti di Biblioteconomia per Luigi Crocetti (10 Marzo 2007-10 Marzo 2008)*, promossi, raccolti, ordinati da Piero Innocenti. Manziana: Vecchiarelli.
- Cheti, Alberto. 2008. "La lezione di Emanuele Casamassima sulla soggettazione." In *Piccoli scritti di biblioteconomia per Luigi Crocetti (10 marzo 2007-10 marzo 2008)*, a cura di Cristina Cavallaro, promossi, raccolti, ordinati da Piero Innocenti, 53-86. Manziana: Vecchiarelli.
- Corbin, Alain. 1991. Recensione di *Le goût de l'archive*, Arlette Farge, Paris: Editions du Seuil, 1989. *Annales. Economies, sociétés, civilisations*, 46^e année, no. 3: 595-97.
- Croce, Benedetto. 1989. *Teoria e storia della storiografia*, a cura di Giuseppe Galasso. Milano: Adelphi (ed. orig. 1916).
- D'Angiolini, Piero, e Claudio Pavone. 1973. "Gli archivi." In *Storia d'Italia, V, t. 2, I documenti*, 1659-91. Torino: Einaudi.
- D'Angiolini, Piero. 2002. *Scritti archivistici e storici*, a cura di Ezelinda Altieri Magliozzi; introduzione di Claudio Pavone. Pubblicazioni degli archivi di Stato. Saggi 75. Roma: Ministero per i beni e le attività culturali, Direzione generale per gli archivi.
- De Gregorio, Mario. 2003. "Editoriale." *AIDAinformazioni*, anno 21, no. 3 (luglio-settembre): 5.
- De Pascale, Gaia. 2009. *Wu Ming: Non soltanto una band di scrittori. Con un'intervista a Wu Ming*. Altrescritture 3. Genova: Il melangolo.
- Dictionary of Archival Terminology (DAT)*. 1984. München; New York; Paris: K. G. Saur.

- Di Filippo, Maria Chiara, e Beatrice Fedolino. 2021. “«Culture del testo», poi «Culture del testo e del documento»: Indici. Parte I. 1995-2003.” *Culture del testo e del documento. le discipline del libro nelle biblioteche e negli archivi* n. s. 22, no. 64 (gennaio-aprile): 93-144.
- Duranti, Luciana. 2010. “Il documento digitale come fonte di prova: produzione, gestione e conservazione.” <https://sab-ero.cultura.gov.it/risorse/materiali/interventi>.
- Fabbrizzi, Andrea. 2019. “I contesti dell’indicizzazione.” In *Viaggi a bordo di una parola: Scritti sull’indicizzazione semantica in onore di Alberto Cheti*, a cura di Anna Lucarelli, Alberto Petrucciani e Elisabetta Vinci, 77-89. Roma: AIB.
- Flammini Giuseppe. 1990. “Prolegomeni alla ‘recensio plenior’ degli ‘Hermeumata Pseudodositheana’.” *Giornale italiano di filologia* 42: 3-43.
- Flöss, Lydia, e Stefania Franzoi, a cura di. 2022. *Meminisse iuvabit. Studi in onore di Pasquale Chistè*. Archivi del Trentino: fonti, strumenti di ricerca e studi 27. Trento: Provincia autonoma di Trento, Soprintendenza per i Beni culturali, Ufficio Beni archivistici, librari e Archivio provinciale.
- Folino, Antonietta, e Erika Pasceri. 2017. *L’Associazione Italiana per la Documentazione Avanzata: Storia e archivio*, Enumera 10. Ariccia: Aracne editrice internazionale.
- Fumagalli, Giuseppe. 1940. *Vocabolario bibliografico*, a cura di Giuseppe Boffito e Giovanni De Bernard. Biblioteca di bibliografia italiana 16. Firenze: Olschki.
- Galluzzi, Paolo, e Pietro Valentino, a cura di. [© 1997, ma 1999]. *I formati della memoria: Beni culturali e nuove tecnologie alle soglie del terzo millennio*. Firenze: Giunti.
- Galluzzi, Paolo. 1997. “Nuove tecnologie e funzione culturale dei musei.” In *I Formati della memoria. Beni Culturali e nuove tecnologie alle soglie del terzo millennio*, a cura di Paolo Galluzzi e Pietro Valentino. 3-39. Firenze: Giunti.
- Gatto, Giada. 2011. “La valigia di Mary Poppins: La fotografia su Web: Descrizione di alcuni cataloghi e banche dati per la fotografia su Web. Con una proposta di creazione di un Opac specifico.” In *Books seem to me to be pestilent things: Studi in onore di Piero Innocenti per i suoi 65 anni*, a cura di Varo Augusto Vecchiarelli e Cristina Cavallaro, 941-54. Manziana: Vecchiarelli.

- Giorgi, Andrea. 2022. "La crisi degli Archivi di Stato." In *Meminisse iuvabit studi in onore di Pasquale Chistè*, a cura di Lydia Flöss e Stefania Franzoi, 177-83. Archivi del Trentino: fonti, strumenti di ricerca e studi 27. Trento: Provincia autonoma di Trento Soprintendenza per i Beni culturali Ufficio Beni archivistici, librari e Archivio provinciale.
- Giunti, Maria Chiara, e Anna Lucarelli, a cura di. 2024. "Nuovo Soggettario: Manuale applicativo." https://thes.bncf.firenze.sbn.it/Manuale_applicativo.pdf.
- Giuva, Linda, e Maria Guercio. 2014. *Archivistica. Teorie, metodi, pratiche*. Roma: Carocci.
- Greco, Francesco. 2011. "La classificazione nelle banche dati del Poligrafico e Zecca dello Stato: Scenari attuali e prospettive future." In *Books seem to me to be pestilent things: Studi in onore di Piero Innocenti per i suoi 65 anni*, a cura di Varo Augusto Vecchiarelli e Cristina Cavallaro, 955-61. Manziana: Vecchiarelli.
- Grillo, Manuela. 2008. *Indicizzazione semantica di bandi, manifesti e fogli volanti: Una proposta*. In *Piccoli scritti di biblioteconomia per Luigi Crocetti (10 marzo 2007-10 marzo 2008)*, a cura di Cristina Cavallaro, promossi, raccolti, ordinati da Piero Innocenti. 121-45. Manziana: Vecchiarelli.
- Grillo, Manuela. 2014. *Leggi e bandi di antico regime*. Bibliographica 8. Carpeghe: Documenta.
- Guarasci, Roberto. 2003. "Libri, documenti e altre storie. L'insegnamento della Documentazione nelle università italiane." *AIDAinformazioni*, no. 3 (luglio-settembre): 47-55.
- Guarasci, Roberto, a cura di. 2008. *Dal documento all'informazione*. Milano: Iter.
- Guarasci Roberto. 2011. "La Memoria della Scienza, L'Archivio Tecnico Italiano e il Centro Nazionale di Documentazione Scientifica." In *Archivi Privati*, a cura di Roberto Guarasci e Erika Pasceri, 195-218. Roma, CNR.
- Guarasci, Roberto. 2022. "Paul Otlet e Suzanne Briet: Note in margine al libro di Paola Castellucci e Sara Mori." *AIDAinformazioni*, no. 1-2 (gennaio-giugno): 147-50.
- Guarasci, Roberto, Anna Rovella, e Maria Taverniti. 2011. "Archivi digitali e profili formativi nell'UE." In *Books seem to me to be pestilent things: Studi in onore di Piero Innocenti per i suoi 65 anni*, a cura di Varo A. Vecchiarelli e Cristina Cavallaro, 1295-1302. Manziana: Vecchiarelli.
- Guarasci, Roberto, e Mauro Guerrini, 2022. *Cos'è l'indicizzazione*. Milano: Editrice Bibliografica.

- Guastalla, Anna, e Annamaria Lazzeri, a cura di. 2019. *Oblio, tempo, cultura ed etica: Saggi e riflessioni dai convegni ANAI 2015-2018*. Archivi del Trentino: fonti, strumenti di ricerca e studi 23. Trento: Provincia autonoma di Trento. Soprintendenza per i Beni culturali. Ufficio Beni archivistici, librari e Archivio provinciale.
- Guercio, Maria. 2022. "Prefazione." In *Non solo carta: Gestione documentale e conservazione degli archivi digitali della provincia autonoma di Trento*, a cura di Carlo Bortoli, Judith Boschi e Annamaria Lazzeri, IX-XIII. Archivi del Trentino: fonti, strumenti di ricerca e studi 29. Trento: Provincia autonoma di Trento. Ufficio beni archivistici, librari e Archivio provinciale.
- Guerrini, Mauro 2011. "Per un tesoro delle scienze del documento." In *Books seem to me to be pestilent things: Studi in onore di Piero Innocenti per i suoi 65 anni*, a cura di Varo Augusto Vecchiarelli e Cristina Cavallaro, 1121-26. Manziana: Vecchiarelli.
- Guerrini, Mauro, e Gianfranco Crupi, diretta da. 2007. *Biblioteconomia. Guida classificata*, a cura di Stefano Gambari. Presentazione di Luigi Crocetti. Milano: Editrice Bibliografica.
- Hiltbrunner, Otto, Hildegard Kornhardt, und Franz Tietze, herausgegeben von. 1954. *Thesaurismata: Festschrift für Ida Kapp zum 70. Geburtstag*. München: C.H. Beck.
- Imperatori, Gianfranco. 1997. "Premessa." In *I Formati della memoria. Beni Culturali e nuove tecnologie alle soglie del terzo millennio*, a cura di Paolo Galluzzi e Pietro Valentino. XV-XIX. Firenze: Giunti.
- Innocenti, Piero. 1989 [ma: 1990]. "L'imperatore ha le orecchie d'asino: Esercizi di lettura sul paradigma della spia." In *Annuario della Università degli studi della Basilicata. A. A. 1987-1988*, vol. 2, 68-106. Potenza: UniBas. Lo scritto è del 1988.
- Innocenti, Piero. 1990. "Bibliography: Teaching and Profession." *Libraries & Culture. A Journal of Library History* 25, no. 3 (Summer): 471-73.
- Innocenti, Piero. 1997. "I luoghi della memoria scritta." *Culture del testo* 3, no. 7 (gennaio-aprile): 39-47.
- Innocenti, Piero. 2003. "Presentazione." In *Information Literacy in Europe. A First Insight into the State of the Art of Information Literacy in the European Union*, a cura di Carla Basili, XI-XII. Roma: CNR.
- Innocenti, Piero. 2008a. "Bibliografia, biblioteconomia, documentazione." In *Dal documento all'informazione*, a cura di Roberto Guarasci, 43-76. Milano: ITER.

- Innocenti, Piero. 2008b. "Bibliografia, biblioteconomia, documentazione." In *Piccoli scritti di biblioteconomia per Luigi Crocetti (10 marzo 2007-10 marzo 2008)*, a cura di Cristina Cavallaro, promossi, raccolti, ordinati da Piero Innocenti. 147-81. Manziana: Vecchiarelli
- International Council on Archives, Expert Group on Archival Description (ICA-EGAD). 2023. "Records in Contexts-Conceptual Model." ICA, November 10, 2023. <https://www.ica.org/resource/records-in-contexts-conceptual-model/>.
- Kraeling, Carl H. 1951. "The Humanities." In *Bibliographic Organisation. Fifteenth Annual Conference of the Graduate Library School. July 24-29, 1950*, edited by Jesse Hauk Shera and Egan Margaret Elizabeth, 109-26. Chicago: The University of Chicago Press.
- Lindsay, Wallace Martin. 1917. "The Philoxenus Glossary." *The Classical Review* 31, no. 7 (November): 158-63.
- Lodolini, Elio. 1981. "L'ordinamento dell'archivio: nuove discussioni." *Rassegna degli Archivi di Stato* 41: 38-56, poi in *Archivistica*. Milano: Franco-Angeli, 7. ed. 1995, 285-301.
- Lodolini, Elio. 1985. *Organizzazione e legislazione archivistica italiana: Dall'Unità d'Italia alla costituzione del Ministero per i beni culturali e ambientali*, prefazione di Giovanni Spadolini. 3. ed., Bologna: Patron.
- Lodolini, Elio. 1987. "Note di terminologia archivistica: Archivio, archivistica, archivista, archivistico, -a." *Archivio Storico Italiano* 145, no. 1, 531 (gennaio-marzo): 99-115.
- Lucarelli, Anna. 2008. *Quando una collezione speciale chiede ospitalità ad una grande Biblioteca*. In *Piccoli scritti di biblioteconomia per Luigi Crocetti (10 marzo 2007-10 marzo 2008)*, a cura di Cristina Cavallaro, promossi, raccolti, ordinati da Piero Innocenti. 183-201. Manziana: Vecchiarelli.
- Lucarelli, Anna. 2011. "Enciclopedie e Thesauri, "essenze" e "proprietà". Una lista semiseria." In *Books seem to me to be pestilent things: Studi in onore di Piero Innocenti per i suoi 65 anni*, a cura di Varo A. Vecchiarelli e Cristina Cavallaro, 975-78. Manziana: Vecchiarelli.
- Lucarelli, Anna, Petrucciani Alberto, e Elisabetta Viti, a cura di. 2019. *Viaggi a bordo di una parola: Scritti sull'indicizzazione semantica in onore di Alberto Cheti*. Roma: AIB.
- Maccanico, Antonio. 1997. "Presentazione." In *I Formati della memoria. Beni Culturali e nuove tecnologie alle soglie del terzo millennio*, a cura di Paolo Galluzzi e Pietro Valentino. V-XIII. Firenze: Giunti.

- Maffei, Lucia. 2003. "Per i venti anni di AIDA." *AIDAinformazioni*, no. 3 (luglio-settembre): 9-10.
- Maniscalco, Zuela. 2019. *Wu Ming: Dalla fake news al romanzo 'neostorico' 54*, prefazione di Massimiliano Tortora. Universitas. Palermo: La Zisa.
- Massabò Ricci, Isabella, e Marco Carassi, a cura di. 2000. *Scritti di teoria archivistica italiana. Rassegna bibliografica: Ecrits de théorie archivistique italienne: Writings on Italian Archival Theory*. Roma: Ministero per i Beni e le Attività Culturali, Ufficio Centrale per i Beni Archivistici.
- Mineo, Leonardo. 2014. "Un nuovo «manuale» di archivistica: alcune riflessioni." <https://sab-ero.cultura.gov.it/risorse/materiali/interventi>.
- Ming, Wu. 2019. "Vent'anni fa, la non-apparizione di Luther Blissett al #PremioStrega: #20annidiQ." Giap. 9 luglio 2019. <https://www.wumingfoundation.com/giap/2019/07/20annidiq-premio-strega/>.
- Non solo carta. Gestione documentale e conservazione degli archivi digitali della Provincia autonoma di Trento*, a cura di Carlo Bortoli, Judith Boschi e Annamaria Lazzeri. Archivi del Trentino: fonti, strumenti di ricerca e studi 29. Trento: Provincia autonoma di Trento. Ufficio beni archivistici, librari e Archivio provinciale.
- Orsini, Pasquale. 2022. "Falsitates et fraudes omnes evitare. Il ruolo dell'archivio nella produzione del falso storico." In *Le forme del falso*, a cura di Marina Caporale Caporale, Cristina Demaria, Daniela Donato, Anna Maria Lorusso, e Francesco Mazzucchelli, 15-26. Bologna: University Press. https://icar.cultura.gov.it/fileadmin/risorse/News/54_2022_Forme_falso.pdf.
- Ouvroir de littérature potentielle. n.d. Consultato il 14 marzo 2024. <https://www.oulipo.net/>.
- Pascuzzi, Giovanni. 2010. *Il diritto dell'era digitale*. Bologna: Il Mulino.
- Penzo Doria, Gianni. 2022. "Una nuova definizione di archivio." *Italian Journal of Library, Archives and Information Science* 13, 2:156-73. <https://doi.org/10.36253/jlis.it-465>.
- Perrella, Renato. 1963. *Bibliografia delle pubblicazioni italiane relative all'archivistica. Rassegna descrittiva e guida*, Quaderni della Rassegna degli archivi di Stato 24. Roma: [s. n. (Siena, La Galluzza)].

- Pettenella, Paola. 2019. "Lungo il Novecento, Archivi di artisti tra futuro e passato." In *Oblio, tempo, cultura ed etica: Saggi e riflessioni dai convegni ANAI 2015-2018*, a cura di Anna Guastalla e Annamaria Lazzeri, 131-38. Archivi del Trentino: fonti, strumenti di ricerca e studi 23. Trento: Provincia autonoma di Trento. Soprintendenza per i Beni culturali. Ufficio Beni archivistici, librari e Archivio provinciale.
- Revelli, Carlo. 2019. "Un catalogatore cartaceo di fronte alla rete." In *Viaggi a bordo di una parola: scritti sull'indicizzazione semantica in onore di Alberto Cheti*, a cura di Anna Lucarelli, Alberto Petrucciani e Elisabetta Vinci, 175-82. Roma: AIB.
- Rossi, Marielisa. 1996. "La fotografia in biblioteca. Una rassegna." *Culture del testo. Rivista italiana di discipline del libro* 2, no. 4 (gennaio-aprile): 7-22.
- Russo, Saverio. 2019. "Chiudiamo gli archivi? Meno soldi e meno addetti per i Beni Culturali." *Historia Ludens*, 13 gennaio 2019. <https://www.historialudens.it/diario-di-bordo/324-chiudiamo-gli-archivi.html>.
- Sannino, Rosa. 2003. "Abstract." *AIDAinformazioni*, no. 3 (luglio-settembre): 125-28.
- Santoro, Raffaele. 2018. "La registrazione di protocollo e il titolario." In *Gli ordinamenti originari degli archivi*, 168-74. Trieste: EUT.
- Schricks, Josine. 2020. "Frauen am Thesaurus linguae Latinae 1: Ida Kapp (1884-1979)." In *Parerga*, 6 maggio 2020. <https://doi.org/10.58079/so73>.
- Servizio informazione e diffusione del Comitato nazionale per la produttività. a cura di. 1965. *Documentazione 1-3 [1965]*. I. *La documentazione. 1. Il mondo documentario*, a cura del Servizio informazione e diffusione del C.N.P., Roma: Comitato nazionale per la produttività, VII, 93 cc. II. *La documentazione in azienda. 2. Il lavoro documentario. Bibliografia*, 94-333 cc. III. *La documentazione in azienda. 3. La strumentazione documentaria*, 334-495 cc.
- Shera, Jesse Hauk, and Margaret Elizabeth Egan, edited by. 1951. *Bibliographic Organisation. XVth Annual Conference of the Graduate Library School. July 24-29, 1950*, Chicago: The University of Chicago Press.
- Soprintendenza archivistica e bibliografica dell'Emilia-Romagna. n.d. "Materiali." Consultato il 14 marzo 2024. <https://sab-ero.cultura.gov.it/risorse/materiali>.
- Thesaurus Linguae Latinae (ThLL)*. Open Access. n.d. Consultato il 14 marzo 2024. <https://publikationen.badw.de/de/thesaurus/lemmata>.

- Treccani. n.d. “Spunto.” Consultato il 7 febbraio 2024. <https://www.treccani.it/vocabolario/spunto2/>.
- Università di Siena, Sistema Bibliotecario di Ateneo. n.d. “Archivi: Glossario dei termini Archivistici Usati.” Consultato il 14 marzo 2024. <http://www.sba.unisi.it/baec/archivi-glossario>.
- Valacchi, Federico. 2017. *Archivio. Concetti e parole*. Milano: Editrice bibliografica.
- Verne, Jules. 1868. *Les enfants du capitaine Grant: Voyage autour du monde*. Paris: Hetzel.
- Zanni Rosiello, Isabella. 1981. “Gli Archivi di Stato: una forma di sapere «segreto» o pubblico?” *Quaderni storici* 16, n° 47 (2), *Culture del lavoro* (Agosto), 624-38.
- Zanni Rosiello, Isabella. 2010. “Il quaderno di Italia nostra sugli archivi: un testo da rileggere.” Soprintendenza archivistica e bibliografica dell’Emilia-Romagna: Interventi. <https://sab-ero.cultura.gov.it/risorse/materiali/interventi>.
- Zanni Rosiello, Isabella. 2017. “Archivi di scrittori e scrittrici. Note a margine.” *Critica letteraria*, anno 17, no. 3: 605-30.

Contributi

Indagine e analisi comparativa delle procedure di codifica nella Terapia del Dolore in Italia

Francesco Amato*, Antonella Benvenuto**, Silvia Ceniti*,
Maria Teresa Chiaravalloti***, Claudia Lanza, Erika Pasceri****

Abstract: This paper aims to analyse the coding practices used by Pain Therapy Centers in Italy, focusing on some of the main diagnostic procedures, diagnoses and types of care setting. Pain Therapy Centers have been invited to fill out a structured survey and the results have been analysed both quantitatively and qualitatively, in order to detect variability in the use of clinical coding systems. The findings have revealed the lack of uniformity in a relatively new discipline which has undergone a significant regulatory reorganization and has been greatly improved by technological-instrumental advances. This work will lay the groundwork towards an assessment of Pain Therapy care quality in Italy and the consequent efforts that are needed in order to define a common care model to ensure prompt and appropriate patient care.

Keywords: Pain Therapy, Coding systems, Diagnosis codes, Procedures codes, Comparative analysis.

1. Introduzione

Il dolore cronico in Italia è caratterizzato da un'epidemiologia estesa e presenta elevati costi socio-sanitari (Stubhaug et al. 2024). Tuttavia, in assenza di un modello assistenziale uniforme, risulta difficile disporre di una stima precisa della malattia e dei costi reali da essa derivanti. Nei Centri di Terapia

* Unità operativa complessa (UOC) Terapia dolore e Cure Palliative, Azienda Ospedaliera di Cosenza (AOCS), Cosenza, Italia. f.amato@aocs.it; s.ceniti@aocs.it.

** Unità operativa semplice dipartimentale (UOSD) Monitoraggio e controllo delle attività di ricovero, AOCS, Cosenza, Italia. antonella.benvenuto37@gmail.com.

*** Istituto di Informatica e Telematica, Consiglio Nazionale delle Ricerche (IIT-CNR), Rende (CS), Italia. maria.chiaravalloti@iit.cnr.it.

**** Dipartimento di Culture, Educazione e Società, Università della Calabria, Rende (CS), Italia. claudia.lanza@unical.it; erika.pasceri@unical.it.

Gli autori hanno lavorato collegialmente alla stesura dell'articolo, tuttavia Claudia Lanza ha redatto il paragrafo "Introduzione"; Silvia Ceniti e Antonella Benvenuto hanno redatto i paragrafi "Inquadramento normativo" e "Metodologia"; Erika Pasceri e Maria Teresa Chiaravalloti hanno redatto i paragrafi "Analisi dei dati" e "Conclusioni"; Francesco Amato ha supervisionato il lavoro.

del Dolore, accanto alle terapie farmacologiche prescritte in base al meccanismo patogenetico del dolore e ai fattori di stimolazione, vengono praticate procedure mininvasive che trovano la loro indicazione a fronte dell'insuccesso dei farmaci o della comparsa di effetti collaterali dei farmaci, così come il prolungarsi del quadro clinico di dolore. In questo contesto, non esistendo un sistema di codifica specialistico per le prestazioni di Terapia del Dolore, per molte procedure vengono utilizzati codici generici, comuni a trattamenti impiegati anche per motivi diversi da quello antalgico. Tutto ciò ha delle ripercussioni sia in ambito clinico, in quanto si perdono informazioni granulari circa i trattamenti e le procedure effettuate, ma anche economico, poiché i codici di rimborso conseguentemente generati sono da considerarsi potenzialmente inappropriati.

Se da un lato è evidente la necessità di pensare e lavorare in maniera innovativa e sinergica per poter addivenire ad erogare servizi sanitari di alta qualità, è altrettanto vero che si devono sviluppare modelli di investimento sostenibili e caratterizzati dall'appropriatezza della cura e della spesa. In questo modo è possibile garantire certezza attuativa sia di efficacia clinica, per assicurare il miglior stato di salute, sia di economicità procedurale, convertendo e riducendo la spesa storica attribuita a percorsi sanitari obsoleti e non più caratteristici dei bisogni epidemiologici. È auspicabile l'uso di criteri efficienti che rispecchino quanto più possibile i principi di eticità comparando il costo al beneficio che deriva dalla nuova tecnologia o terapia. Questa è anche la direzione verso cui si muove l'Unione Europea. Infatti, le più recenti direttive comunitarie prevedono procedure più strutturate per la valutazione e promozione dei dispositivi medici al fine di rispettare il rapporto costo-efficacia. Diversi studi di economia sanitaria hanno affrontato le problematiche relative alla valutazione del costo del dolore cronico, con l'obiettivo di stimare sia i costi diretti del trattamento sanitario, sia i costi indiretti, come ad esempio quelli derivanti dalle perdite di produttività determinate dalle assenze dal lavoro dei pazienti affetti da dolore cronico. Ad oggi, si stima che in Italia questa patologia colpisca circa 10 milioni di pazienti (Breivik et al. 2006; Melotti et al. 2009; Apolone et al. 2009; Toccaceli et al. 2019) con una conseguente incidenza sulla spesa sanitaria pubblica di 61,9 miliardi di euro all'anno (Sanità24 2023).

Tuttavia, questo dato potrebbe risultare sottostimato dal momento che non tiene conto né dei costi sanitari sostenuti nel privato, né di quelli derivanti da un utilizzo inappropriato dei farmaci analgesici (Allegrì et al. 2015).

Le procedure praticate nei Centri di Terapia del Dolore risultano sempre più efficaci grazie alla progressiva continuità d'uso e al miglioramento tecnologico delle soluzioni praticate. Per tale motivo, durante tutte le fasi del processo assistenziale è necessario che vi sia una collaborazione attiva con i Centri di Terapia del Dolore, oltre che una loro supervisione sul percorso di cura del paziente. Dal punto di vista organizzativo, la necessità di offrire servizi che

mettano il paziente al centro del percorso di cure deve esplicarsi attraverso momenti sanitari sinergici e non dispersi territorialmente. In quest'ottica, quindi, il processo erogativo sanitario dovrebbe configurarsi come una sorta di "staffetta" tra professionisti, strutture ed ambiti assistenziali per gestire in modo coerente e complementare il complesso iter che parte dalla diagnosi, passa per il trattamento ed esita nel follow up del paziente. È necessaria, quindi, la costruzione di una *practice* clinica omogenea, in cui il medico, sottraendosi alla logica dell'autoreferenzialità, possa trattare tempestivamente e in maniera globale il dolore per erogare prestazioni sanitarie ad alta complessità in un ambito organizzativo multidisciplinare, multiprofessionale e che si avvale dell'utilizzo di alte tecnologie. In questa concatenazione di "prese in carico" chi gestisce la fase di diagnosi realizza un traguardo intermedio di salute e predispone le condizioni per la fase successiva di trattamento, che potrebbe essere svolto all'interno di altre strutture ed attraverso altre competenze professionali. L'output complessivo di questo processo sanitario fornisce indicazioni circa la qualità della risposta assistenziale rispetto al problema di salute.

Gli obiettivi principali del presente studio riguardano l'identificazione e il confronto di più tipologie di procedure antalgiche caratterizzanti la Terapia del Dolore, ponendo particolare attenzione ai casi in cui a tali procedure risultano associati dei DRG¹ non pertinenti per la spesa complessiva e per la tipologia di ricovero. Inoltre, lo studio vuole porre le basi per un successivo lavoro di uniformazione delle procedure, dei DRG associati e dei relativi setting assistenziali, costruendo una matrice di corrispondenza a partire dai relativi sistemi di codifica al fine di rendere comparabile, e quindi valutabile, l'attività svolta dai Centri delle Unità operative di Terapia del Dolore (Hardy 2022). Lo studio è stato condotto grazie ad una collaborazione sinergica tra professionisti nell'ambito della Terapia del Dolore ed esperti in sistemi di gestione e organizzazione della conoscenza in ambito clinico.

2. Inquadramento normativo

Un modello assistenziale funzionale dovrebbe prevedere una necessaria e stretta collaborazione tra l'ospedale, ovvero il momento in cui vengono messe in atto le cure specialistiche, e le strutture sanitarie dislocate sul territorio, responsabili delle iniziali fasi di individuazione della patologia e di indirizzamento verso uno specifico percorso di cura, oltre che della fase di follow up. Questo modello trova la sua realizzazione nella costituzione di una Rete, per come declinato dalla Legge n. 38 del 15 marzo 2010 "Disposizioni per garan-

¹ Raggruppamento omogeneo di diagnosi (Diagnosis Related Group, DRG). Il DRG è un codice specifico, che permette di classificare tutti i malati dimessi da un ospedale in gruppi omogenei in base alle risorse impegnate per la loro cura (ISS 2020).

tire l'accesso alle cure palliative e alla terapia del dolore” sul territorio nazionale e recepita dalle singole Regioni. Il modello organizzativo delle Reti di Terapia del Dolore dovrebbe consentire da un lato l'adeguato trattamento del dolore cronico, nonché la continuità assistenziale del paziente, dall'altro di ottenere una stima reale della prevalenza del dolore cronico e dei relativi costi diretti e indiretti per il Servizio Sanitario Nazionale (SSN), al fine di realizzare una pianificazione attenta alla spesa e alle esigenze di salute del paziente in accordo con quanto previsto dal Piano Nazionale delle Cronicità (Ministero della Salute 2019). La Terapia del Dolore rappresenta anche uno degli obiettivi primari del Piano Sanitario Nazionale (Ministero della Salute 2019-2021), che prevede la definizione di un modello di rete regionale di centri specialistici, integrati con la medicina del territorio, in grado di garantire il diritto di accesso alla Terapia del Dolore su tutto il territorio regionale a tutti i cittadini, di ogni età ed in ogni luogo di cura, compreso il domicilio del paziente. Tuttavia, lo stato di attuazione della Legge 38/2010 è caratterizzato da una forte disomogeneità a livello regionale e locale, creando difficoltà di accesso ai servizi da parte dei cittadini. È, pertanto, auspicabile sviluppare un modello assistenziale uniforme, che abbia come riferimento normativo quanto è esplicitato dall'Intesa sottoscritta in sede di Conferenza Stato-Regioni (2012) in termini di requisiti minimi e di criteri di accreditamento, e che preveda l'istituzione, attraverso appositi provvedimenti regionali e aziendali, di una struttura specificatamente dedicata al coordinamento della Rete. Gli obiettivi attesi dalla Rete di Terapia del Dolore in ogni Regione sono:

- istituzione di una struttura organizzativa con funzioni di coordinamento e monitoraggio delle attività della Rete. Essa dovrà interfacciarsi con le Aziende, con i professionisti della Rete di Terapia del Dolore e con la Rete delle cure palliative al fine di garantire la piena applicabilità della normativa vigente e la fruibilità dei servizi previsti. Pertanto, la struttura organizzativa dovrà essere supportata da un sistema informativo che si integri con quello regionale. Inoltre, la struttura potrà individuare *best practices* da divulgare in altri contesti;
- attivazione di tutti i nodi della Rete sia a livello territoriale che ospedaliero, garantendo la realizzazione di percorsi diagnostici, terapeutici e riabilitativi condivisi e predisponendo procedure clinico-assistenziali e organizzativo-gestionali della Rete di Terapia del Dolore uniformi a livello territoriale e, possibilmente, conformi ad un modello assistenziale nazionale. Risulta fondamentale la realizzazione del sistema informativo regionale, che permetta la comunicazione e la tracciabilità dei dati in tutti i nodi della Rete di Terapia del Dolore, che sia capace di interoperare con i sistemi informativi delle altre Reti regionali;
- tutela del cittadino nell'accesso alla Rete di Terapia del Dolore;

- mantenimento della continuità delle cure mediante l'integrazione dei diversi setting assistenziali. Tale finalità potrà essere raggiunta mediante il supporto dei sistemi informativi che consentono la condivisione dei flussi informativi tra professionisti, soprattutto tra quelli ospedalieri e quelli che operano sul territorio.

Ancora all'interno di una prospettiva normativa, il Decreto ministeriale 21 novembre 2018 "Aggiornamento dei codici delle specialità cliniche e delle discipline ospedaliere" ha inserito il codice 96, identificativo della disciplina di Terapia del Dolore, per la rilevazione delle attività ospedaliere sulla Terapia del Dolore. Tale passaggio risulta di grande rilevanza per documentare il controllo del dolore in ambito ospedaliero o ambulatoriale, in quanto vi è l'obbligo per le Aziende sanitarie e ospedaliere di censire nel sistema informativo sanitario le Unità di degenza ospedaliera dedicate alla disciplina Terapia del Dolore, i relativi posti letto e l'annotazione nelle Schede di Dimissione Ospedaliera (SDO)² dei ricoveri effettuati nelle unità operative dedicate. Più di recente è stato stipulato l'Accordo del 27 luglio 2020 in sede di Conferenza Stato-Regioni che, ai sensi della Legge 38/2010, approva il documento "Accreditamento delle reti di terapia del dolore" (Rep. Atti n. 119/CSR). In tale Accordo le Regioni e le Province autonome si impegnano a recepirlo entro dodici mesi dalla data di approvazione, inserendo contenuti relativi all'accreditamento della Rete volti alla qualificazione dei percorsi di Terapia del Dolore e ad attivare un sistema di monitoraggio.

Pertanto, le aziende sanitarie, ospedaliere, ospedaliere universitarie, gli istituti di ricovero e cura a carattere scientifico, nella rilevazione dei modelli "HSP.12 - Posti letto per disciplina delle strutture di ricovero pubbliche ed equiparate" e "HSP.13 - Posti letto per disciplina delle case di cura private", dovranno fare riferimento alle tabelle aggiornate, nelle quali è stato inserito il codice 96 relativo alla Terapia del Dolore. L'identificativo di disciplina per la Terapia del Dolore ha dato la possibilità di effettuare un'analisi puntuale delle prestazioni erogate dai Centri di Terapia del Dolore presenti sul territorio nazionale sia in termini quantitativi che qualitativi, pertanto essa rappresenta un passo importante, anche se ancora non sufficiente, per il riconoscimento dell'autonomia e della dignità di questa disciplina (Conferenza Stato-Regioni 2020).

² Tale scheda è stata istituita con Decreto del Ministro della Sanità del 28 dicembre 1991, come strumento ordinario per la raccolta delle informazioni relative ad ogni paziente dimesso dagli istituti di ricovero pubblici e privati in tutto il territorio nazionale per attivare il flusso informativo e al fine di garantire l'omogeneità dei dati rilevati sull'intero territorio nazionale, così da consentirne elaborazioni, aggregazioni e confronti a vari livelli.

3. Metodologia

La prima fase dello studio ha riguardato una ricognizione delle strutture HUB presenti sul territorio nazionale che dichiarano di erogare prestazioni connesse alla Terapia del Dolore. In particolare, sono stati individuati 17 centri HUB di Terapia del Dolore, a cui è stata richiesta la compilazione di un Case Report Form (CRF), al fine di rilevare i codici delle procedure erogate e delle relative diagnosi, i setting assistenziali, i DRG, e le relative tariffe applicate. Sul totale delle strutture contattate, 11 hanno correttamente compilato il CRF, 3 solo parzialmente, mentre 3 non lo hanno compilato. La Figura 1 mostra i centri HUB contattati che hanno compilato totalmente o parzialmente il CRF.



Figura 1: Centri HUB di Terapia del Dolore che hanno compilato totalmente o parzialmente il CRF.

In assenza di un archivio unico nazionale dei Centri specialistici di Terapia del Dolore, la ricognizione è stata effettuata a partire da informazioni reperite direttamente in rete dai portali delle diverse Aziende sanitarie e, laddove necessario, le informazioni reperite sono state integrate mediante contatto telefoni-

co. La rilevazione dei dati è stata effettuata in un arco temporale compreso tra il 14 dicembre 2022, data di avvio dell'indagine, e il 28 aprile 2023, giorno di chiusura dell'indagine e avvio dell'analisi dei risultati.

Preliminarmente alla fase di analisi dei dati, è stato necessario procedere ad una fase di normalizzazione dei dati inseriti da ciascun centro HUB al fine di ripulire e rendere omogenea la base dati di partenza. In particolare, sono stati rimossi refusi e sono stati revisionati i dati immessi in modo differente o non coerente (ad es., la denominazione del setting abbreviata, o codici ICD9-CM³ non trascritti correttamente). Questa fase ha consentito di analizzare le metodologie e le pratiche di codifica adottate dai diversi centri HUB coinvolti nell'indagine sia per la parte dei codici relativi alle procedure che alle diagnosi. Le tabelle di seguito riportate elencano il sottogruppo di procedure antalgiche oggetto del presente studio (Tab. 1). Esse sono 6 delle 13 rilevate mediante l'indagine presso i Centri HUB partecipanti e sono state selezionate per il presente lavoro di analisi in quanto più specifiche per la disciplina Terapia del Dolore e perché associate a DRG considerati "critici" (Tab. 2), cioè quelli più frequentemente utilizzati, ma con maggiore difformità sia nei setting che nelle codifiche correlate a diagnosi e procedure.

Codice ICD9-CM Procedura antalgica	Denominazione ICD9-CM Procedura antalgica
03.6	Separazione di aderenze del midollo spinale e delle radici dei nervi
03.90	Inserzione di catetere nel canale vertebrale per infusione di sostanze terapeutiche o palliative
03.93	Impianto o sostituzione di elettrodo/i di neurostimolatore spinale
03.94	Rimozione di elettrodo/i del neurostimolatore spinale
80.59	Altra distruzione di disco intervertebrale
86.06	Inserzione di pompa di infusione totalmente impiantabile

Tabella 1: Elenco codici e denominazioni ICD9-CM delle procedure antalgiche oggetto del presente studio.

³ L'International Classification of Diseases - Clinical Modification 9th revision (ICD9-CM) è il sistema utilizzato per la codifica delle procedure e le diagnosi correlate. La classificazione prevede dei codici a 2, 3, 4 o 5 cifre con la punteggiatura che sta ad indicare la specificità della procedura o diagnosi da codificare.

Codice DRG	Descrizione DRG
007	Interventi su nervi periferici e cranici e altri interventi su sistema nervoso con CC
008	Interventi su nervi periferici e cranici e altri interventi su sistema nervoso senza CC
009	Malattie e traumatismi del midollo spinale
035	Altre malattie del sistema nervoso senza CC
461	Intervento con diagnosi di altro contatto con i servizi sanitari
463	Segni e sintomi con CC
464	Segni e sintomi senza CC
468	Intervento chirurgico esteso non correlato alla diagnosi di dimissione principale
500	Interventi su dorso e collo eccetto per artrodesi vertebrale senza CC
531	Interventi sul midollo spinale con CC
532	Interventi sul midollo spinale senza CC

Tabella 2: Elenco codici e denominazioni DRG associati alle procedure analizzate.

4. Analisi dei dati

L'indagine è stata strutturata in modo tale da permettere la rilevazione dei dati necessari ad effettuare le seguenti analisi:

- centri HUB di Terapia del Dolore che effettuano una determinata procedura antalgica;
- centri HUB di Terapia del Dolore che effettuano una determinata procedura antalgica in un determinato setting (degenza/DH, ambulatoriale);
- codici di procedura utilizzati dai centri HUB di Terapia del Dolore rispetto ad uno stesso codice di diagnosi;
- codici DRG utilizzati dai centri HUB di Terapia del Dolore rispetto ad uno stesso codice di procedura e valutazione della corrispondente tariffazione.

Per quanto riguarda il primo punto di analisi, la Tabella 3 riporta la sintesi numerica degli HUB che effettuano le procedure antalgiche oggetto di analisi. Questo dato ha permesso di verificare la diffusione di ciascuna delle procedure analizzate nei centri HUB dislocati sul territorio nazionale.

Codice procedura antalgica	N° di HUB che effettuano la procedura
03.6	10
03.90	9
03.93	13
03.94	7
80.59	6
86.06	10

Tabella 3: Sintesi numerica dei centri HUB che effettuano le procedure analizzate.

Per quanto riguarda, invece, l'analisi dei codici di diagnosi associati all'espletamento delle procedure antalgiche oggetto del presente studio, nelle Tabelle 4, 5 e 6 sono riportate a titolo esemplificativo le situazioni relative a tre delle sei procedure antalgiche oggetto del presente studio. In esse risulta evidente che sono molteplici e diversificati i codici di diagnosi che portano alla realizzazione di una stessa procedura antalgica. Grazie al fatto che ICD9-CM è una classificazione con struttura gerarchica, è possibile desumere l'appartenenza dei codici di diagnosi alle branche specialistiche di patologia.

Denominazione HUB	Codice Diagnosi in ICD9-CM
Cosenza	353.8; 724.02
Foggia	722.83
Garbagnate	338.28
Careggi	349.2
Pavia	724.02
Potenza	722.83; 724.2-3; 722.52; 337.1
Napoli	353.4
Roma - Tor Vergata	338.4; 724.2; 349.2
Verona	722.0-9; 724.0-9
Torino	338.29

Tabella 4: Codici di diagnosi utilizzati nei centri HUB per la procedura 03.6 *Separazione di aderenze del midollo spinale e delle radici dei nervi*.

Denominazione HUB	Codice Diagnosi in ICD9-CM
Cosenza	353.8; V53.02; V53.2; 996.2
Catania	V53.02; 338.4
Foggia	353.1
Garbagnate	338.28-29; 996.2
Careggi	440.20; 729.2
Monaldi	353.2-4
Pavia	353.4; 337.21-22; 531.3; 353.0; 353.1; 355.0; 338.4; 784.0
Potenza	*manca indicazione di codice ICD-9-CM
Ravenna	338.4; 053.13; 355.1; 722.83; 724.02; 596.54; 440.22
Napoli	353.4; 996.2
Roma - Tor Vergata	353.8; 722.83; 338.24; 338.4
Torino	338.29; V53.02
Verona	355.0-9; 356.0-9; 357.0-9; V53.02; 724.1- 2; 723.3; 784.0; 346.01; 346.11; 346.21; 346.81; 346.91

Tabella 5: Codici di diagnosi utilizzati nei centri HUB per la procedura 03.93 *Impianto o sostituzione di elettrodi di neurostimolatore spinale.*

Denominazione HUB	Codice Diagnosi in ICD9-CM
Cosenza	338; 338.3; V53.09
Garbagnate	344.1; V53.02
Careggi	729.2; 781.0
Pavia	334.1
AOU Pisana	338.4
Potenza	340; 335.2
Ravenna	V58.82
Napoli	353.4
Torino	338.29
Verona	338.29; 338.3; 729.2

Tabella 6: Codici di diagnosi utilizzati nei centri HUB per la procedura 86.06 *Inserzione di pompa di infusione totalmente impiantabile.*

I dati raccolti hanno permesso, inoltre, di analizzare per una stessa procedura antalgica il setting assistenziale di erogazione, il relativo DRG e la corrispondente tariffazione. Nelle Tabelle seguenti (Tabb. 7-9) è riportato questo tipo di analisi per alcune delle procedure antalgiche oggetto del presente studio. Nelle Figure seguenti (Figg. 2-4) è, invece, illustrato graficamente l'andamento della tariffazione per una stessa procedura antalgica.

Denominazione HUB	DRG	Setting assistenziale
Cosenza	532	RO
Basilicata		
Roma - Tor Vergata		
Foggia	500	RO
Garbagnate		
Pavia		
Careggi	532	DH/RO
Pavia	500	DH
Verona		
Basilicata	532	DH
Napoli		
Basilicata	531*	RO con cc
Roma - Tor Vergata	461	RO
Torino	461	DAY SURGERY

Tabella 7: Codici DRG e setting assistenziale utilizzati nei centri HUB per la procedura 03.6 *Separazione di aderenze del midollo spinale e delle radici dei nervi*. L'asterisco (*) indica "con complicanze".

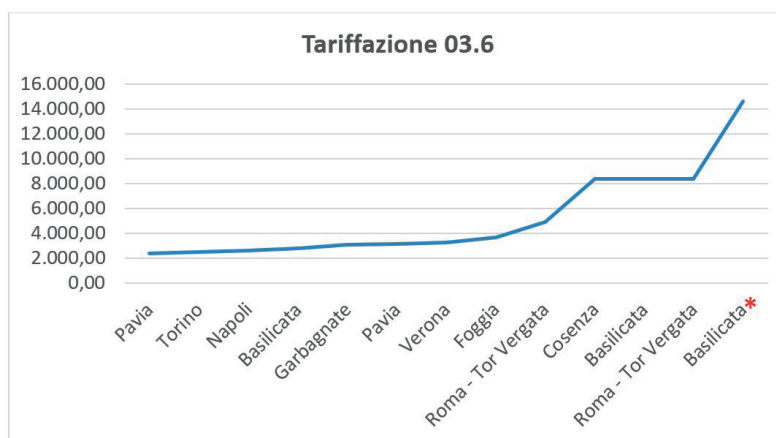


Figura 2: Tariffazione nei centri HUB per la procedura 03.6 *Separazione di aderenze del midollo spinale e delle radici dei nervi*. L'asterisco (*) indica "con complicanze".

Denominazione HUB	DRG	Setting assistenziale
Cosenza	532	RO
Catania		
Foggia		
Garbagnate		
Garbagnate		
Monaldi		
Careggi		
Pavia		
Pavia		
Napoli		
Roma - Tor Vergata		
Careggi	468	RO
Basilicata	*DRG manca	DH/RO
Ravenna	461	RO
Pavia		
Napoli	531*	DH
Roma - Tor Vergata	531*	RO
Torino	461	DAY SURGERY CON P./RO
Torino	532	DAY SURGERY
Verona	532	DH

Tabella 8: Codici DRG e setting assistenziale utilizzati nei centri HUB per la procedura 03.93 *Impianto o sostituzione di elettrodi di neurostimolatore spinale*. L'asterisco (*) indica "con complicanze".

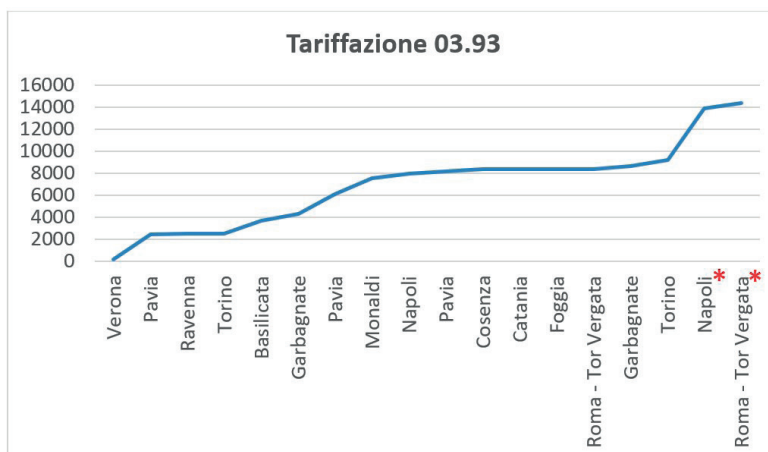


Figura 3: Tariffazione nei centri HUB per la procedura 03.93 *Impianto o sostituzione di elettrodi di neurostimolatore spinale*. L'asterisco (*) indica "con complicanze".

Denominazione HUB	DRG	Setting assistenziale
Cosenza	461	RO
Ravenna		
Cosenza	008	RO
Garbagnate		
Careggi		
Basilicata		
Pavia		
Basilicata	007	RO con cc
Napoli	007	RO
Torino	461	DAY SURGERY CON P./RO
Verona	461	DH
Verona	*manca DRG	*manca setting

Tabella 9: Codici DRG e setting assistenziale utilizzati nei centri HUB per la procedura 86.06 *Inserzione di pompa di infusione totalmente impiantabile*. L'asterisco (*) indica "con complicanze".

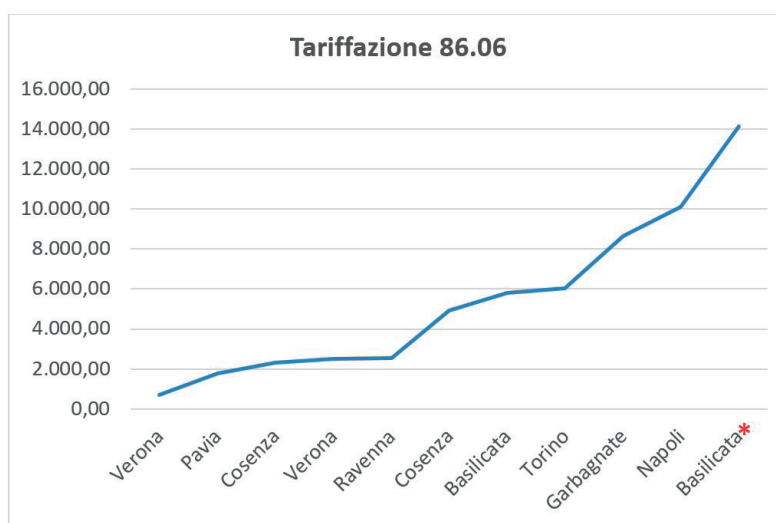


Figura 4: Tariffazione nei centri HUB per la procedura 86.06 *Inserzione di pompa di infusione totalmente impiantabile*. L'asterisco (*) indica "con complicanze".

Da questa prima analisi emerge un certo discostamento tra i vari centri. È possibile, infatti, evidenziare che la stessa procedura viene effettuata in setting assistenziali diversi e con un diverso utilizzo delle risorse, in termini di spesa complessiva e razionale di distribuzione, e, di conseguenza, una discrepanza tra tariffazione e DRG-specifico. Risulta dunque necessario porre le basi per un confronto tra pari al fine di rivalutare il tipo di utilizzo del setting in base alla procedura e alla tecnologia utilizzata.

Come evidenziato in precedenza, alcuni dati inclusi nella base di dati risultano essere mancanti o non inseriti correttamente. Trattandosi di schede compilate manualmente sono stati individuati refusi ed errori di battitura che hanno necessariamente comportato – per alcune analisi – una visione parziale dell’insieme. In questi casi è stato deciso di mantenere comunque l’indicazione dell’HUB e la procedura effettuata per tenere traccia dell’informazione, anche se questa – essendo parziale – non consente un confronto diretto con gli altri centri circa l’associazione fra il codice della procedura e quello della diagnosi.

Dal punto di vista clinico è altresì possibile notare che per la stessa procedura vengono generati DRG diversi (medici, chirurgici), ma non sempre tale discrepanza risulta giustificata dal numero di diagnosi correlabili ed è solo in parte associabile all’utilizzo di codifiche multiple di procedure⁴.

L’analisi effettuata circa l’associazione dei codici di procedure antalgiche ai codici di diagnosi evidenzia informazioni utili per i professionisti sanitari da utilizzare come base di discussione per definire una comune metodologia di codifica. Da quanto emerge, infatti, si fa sempre più evidente la necessità di una condivisione di linee guida per l’utilizzo dei codici in modo omogeneo e coerente, che dovrebbe considerare la globalità del ricovero, tenendo conto sia della tipologia di procedura antalgica eseguita, che della diagnosi, del setting utilizzato e dei costi realmente sostenuti quali costo del personale, costo dei materiali consumati e delle apparecchiature utilizzate.

Un aspetto di rilievo che è emerso durante l’analisi è relativo alle diverse denominazioni che una stessa procedura assume nei centri HUB interpellati, e che non sempre coincide con la denominazione ufficiale del sistema ICD9-CM. Questa discrepanza è ascrivibile al fatto che alcuni concetti clinici non trovano un equivalente semantico nelle denominazioni di procedure del sistema di codifica ICD9-CM, perlopiù perché rappresentativi di pratiche cliniche altamente specialistiche e/o implicanti tecnologie nuove, non ancora descritte dai codici ICD9-CM. Questo fa sì che i codici di procedure di ICD9-CM diventino dei “contenitori” di prestazioni con un diverso grado di dettaglio clinico, e quindi conseguenzialmente anche con un costo differente⁵.

Di seguito si riportano le evidenze rilevate per alcune delle procedure oggetto del presente studio e i relativi codici ICD9-CM di diagnosi associati.

Codice e denominazione procedura ICD9-CM: 03.6 Separazione di aderenze del midollo spinale e delle radici dei nervi

⁴ La codifica multipla è una procedura che prevede la possibilità di inserire più codici ICD9-CM per la formulazione di una diagnosi laddove un codice da solo non sia esaustivo.

⁵ È altresì necessario precisare che, coerentemente con gli obiettivi di regionalizzazione del SSN, alle Regioni viene lasciata comunque l’autonomia di definire i propri tariffari, da utilizzare come importante strumento di policy per incentivare l’appropriatezza ospedaliera.

Denominazione HUB	Denominazione locale procedura	Codice diagnosi ICD9-CM
HUB Ospedali Riuniti di Foggia	Separazione di aderenze del midollo spinale e delle radici dei nervi	722.83
HUB Calabria: PO Mariano Santo Cosenza	Separazione di aderenze del midollo spinale e delle radici dei nervi (epidurolisi ed epiduroscopia con lisi aderenze)	353.8
HUB Garbagnate: ASST Rhodense	Peridurolisi in periduroscopia	338.28
HUB AOU Careggi Firenze	Epiduroscopia	349.2
HUB Pavia	Separazione di aderenze del midollo spinale e delle radici (con periduroscopio)	724.02
HUB Pavia	Separazione di aderenze del midollo spinale e delle radici (con catetere armato secondo rackz)	724.02
HUB Potenza	Separazione di aderenze del midollo spinale e delle radici dei nervi	722.83
HUB San Giuliano Napoli	Adesiolisi caudale	353.4
HUB Torino: Ospedale Molinette	Peridurale selettiva con catetere orientabile st. Reed	338.29
HUB Verona: AOU Integrata	Radiofrequenza pulsata sul ganglio o sulle radici nervose del midollo	722
HUB Verona: AOU Integrata	Peridurolisi /periduroscopia	722

Codice e denominazione procedura ICD9-CM: 03.93 Impianto o sostituzione di elettrodo/i di neurostimolatore spinale

Denominazione HUB	Denominazione locale procedura	Codice diagnosi ICD9-CM
HUB Calabria: PO Mariano Santo Cosenza	Impianto o sostituzione di elettrodo di neurostimolatore spinale (provvisorio)	353.8
HUB Catania: PO San Marco Policlinico	Impianto o sostituzione di elettrodo/i di neurostimolatore spinale	V53.02

Denominazione HUB	Denominazione locale procedura	Codice diagnosi ICD9-CM
HUB Ospedali Riuniti di Foggia	Impianto / Sostituzione di Elettrocatteter Midollare	353.1
HUB Ospedali Riuniti di Foggia	Impianto o sostituzione di elettrodo/i di neurostimolatore spinale	353.1
HUB Ospedali Riuniti di Foggia	Collocazione / Sostituzione di Neurostimolatore Spinale + generatore di impulsi a canale doppio, non specificato se ricaricabile	353.1
HUB Ospedali Riuniti di Foggia	Collocazione / Sostituzione di Neurostimolatore Spinale + altro generatore di impulsi	353.1
HUB Ospedali Riuniti di Foggia	Collocazione / Sostituzione di Neurostimolatore Spinale + altro generatore di impulsi a canale singolo ricaricabile	353.1
HUB Ospedali Riuniti di Foggia	Collocazione / Sostituzione di Neurostimolatore Spinale + altro generatore di impulsi a canale doppio ricaricabile	353.1
HUB Garbagnate: ASST Rhodense	Impianto elettrodo/i SCS fase test cervicale	338.28, 338.29
HUB Garbagnate: ASST Rhodense	Impianto elettrodo/i DRG fase test lombare	338.28, 338.29
HUB Garbagnate: ASST Rhodense	Sostituzione di elettrodo/i per SCS	996.2
HUB Garbagnate: ASST Rhodense	Sostituzione di elettrodo/i per DRG	996.2
HUB AOU Careggi Firenze	Impianto di elettrodo spinale fase trial	729.2
HUB Ospedale Monaldi	Impianto o sostituzione di elettrodo/i di neurostimolatore spinale	353.3, 353.4
HUB Pavia	Impianto o sostituzione di elettrodo di neurostimolatore	353.4
HUB Potenza	SCS Impianto-sostituzione elettrodo	*mancante
HUB Ravenna: Ospedale Santa Maria delle Croci	Impianto di stimolatore trial	338.4

Denominazione HUB	Denominazione locale procedura	Codice diagnosi ICD9-CM
HUB San Giuliano Napoli	Impianto di neurostimolatore midollare	353.4
HUB San Giuliano Napoli	Sostituzione elettrodo midollare	996.2
HUB Università Tor Vergata	Altri disturbi delle radici e dei plessi nervosi	353.8
HUB Università Tor Vergata	Sindrome postlaminectomia, regione lombare	722.83
HUB Torino: Ospedale Molinette	SCS provvisorio	338.29
HUB Verona: AOU Integrata	Test di elettro-stimolazione percutanea nel canale vertebrale (compreso sacrale)	355
HUB Verona: AOU Integrata	Collocazione di impianto definitivo di elettrodo/i di neurostimolatore spinale con generatore di impulsi	V53.02
HUB Verona: AOU Integrata	Sostituzione di elettrodo/i di neurostimolatore spinale	V53.02
HUB Verona: AOU Integrata	Sostituzione di neurostimolatore	V53.02

Codice e denominazione procedura ICD9-CM: 86.06 Inserzione di pompa di infusione totalmente impiantabile

Denominazione HUB	Denominazione locale procedura	Codice diagnosi ICD9-CM
HUB Calabria: PO Mariano Santo Cosenza	Inserzione di pompa totalmente impiantabile	338
HUB Calabria: PO Mariano Santo Cosenza	Inserzione di pompa totalmente impiantabile	338.3
HUB Calabria: PO Mariano Santo Cosenza	Sostituzione di pompa totalmente impiantabile	V53.09
HUB Catania: PO San Marco Policlinico	Impianto Pompa Intratecale	338.4

Denominazione HUB	Denominazione locale procedura	Codice diagnosi ICD9-CM
HUB Ospedali Riuniti di Foggia	Inserzione di pompa di infusione totalmente impiantabile + Altri interventi sul midollo spinale e sulle strutture del canale vertebrale	353.1
HUB Garbagnate: ASST Rhodense	Impianto di Pompa per Terapia Intratecale	344.1
HUB Garbagnate: ASST Rhodense	Sostituzione di Pompa per Terapia Intratecale	V53.02
HUB AOU Careggi Firenze	Impianto pompa infusione subaracnoidea	729.2
HUB AOU Careggi Firenze	Impianto pompa spasticità	781.0
HUB Pavia	Inserzione di pompa totalmente impiantabile	334.1
HUB AOU Pisana	Impianto pompa	338.4
HUB Ravenna: Ospedale Santa Maria delle Croci	Inserzione di pompa totalmente impiantabile	340, 335.2
HUB Ravenna: Ospedale Santa Maria delle Croci	Inserzione di pompa totalmente impiantabile	V58.82
HUB Napoli: Ospedale Pascale	Impianto di pompa intratecale	353.4
HUB Torino: Ospedale Molinette	Pompa itt	338.29
HUB Verona: AOU Integrata	Inserzione pompa di infusione	338.3
HUB Verona: AOU Integrata	Inserzione pompa di infusione	338.29
HUB Verona: AOU Integrata	Test per infusione intratecale di farmaci: con inserzione del catetere o test a bolo continuo	729.2

5. Conclusioni

In linea con le direttive comunitarie che prevedono procedure più strutturate per la valutazione dei dispositivi medici e per una corretta valutazione di costo-efficacia (Cabina di Regia 2017), lo studio ha inteso effettuare un con-

fronto sulle procedure antalgiche, caratterizzanti la disciplina della Terapia del Dolore, per poter approfondire diversi aspetti che le caratterizzano e fornire ai diversi livelli di programmazione, indicazioni utili che mirino a garantire e migliorare la qualità assistenziale del paziente con dolore. Lo studio condotto ha consentito di effettuare una ricognizione sull'utilizzo a livello nazionale dei sistemi di codifica di ambito clinico relativamente al dominio specialistico della Terapia del Dolore. Esso è da considerarsi come una base di partenza per definire delle linee guida di indirizzo, che traggano beneficio dalla contestualizzazione desumibile dalle analisi effettuate, circa un utilizzo uniforme ed appropriato delle codifiche relativamente alla procedura, al DRG associato, al setting assistenziale e alla relativa tariffa applicata. Un utilizzo appropriato delle codifiche consente infatti di effettuare delle analisi periodiche e puntuali che possano essere di supporto alle istituzioni per una tempestiva sorveglianza sanitaria nonché per un'efficiente programmazione, oltre che per un monitoraggio costante delle procedure disponibili sul territorio nazionale nell'ambito della Terapia del Dolore (O'Malley et al. 2005).

Una gestione adeguata del paziente affetto da dolore richiede, difatti, che si possa disporre di un sistema di assistenza continuativa, che permetta la realizzazione di piani di cura personalizzati a lungo termine, la razionalizzazione dell'uso delle risorse e il miglioramento della qualità della vita, prevenendo le disabilità e la non autosufficienza. Considerando le dimensioni epidemiologiche e l'impatto altamente invalidante del dolore ed al fine di perseguire la finalità di efficientamento e razionalizzazione delle risorse disponibili e al contempo promuovere qualità ed esiti positivi nei trattamenti, è necessario erogare in maniera appropriata ed uniforme su tutto il territorio nazionale prestazioni ad alta complessità altamente innovative e minimamente invasive per dare una risposta alla sempre crescente domanda di salute. Le procedure antalgiche oggi in uso hanno visto negli ultimi decenni una fioritura di attività chirurgiche mininvasive e di attività per il trattamento del dolore acuto e cronico/persistente con utilizzo di metodiche innovative, che in alcuni casi sostituiscono interventi più invasivi in auge negli anni precedenti.

Tuttavia, il sistema di remunerazione degli ospedali del SSN si basa ancora sulle tariffe nazionali definite dal DM del 18 ottobre 2012, relative alla classificazione DRG, versione 24, introdotta nel 2009. Si riscontra, pertanto, una certa differenza tra le Regioni relativamente alle procedure di Terapia del Dolore in riferimento ai setting assistenziali, ai codici di diagnosi e ai codici di intervento utilizzati e conseguentemente alla relativa valorizzazione economica. Queste differenze potrebbero essersi ingenerate anche perché la classificazione DRG attualmente in uso non ha seguito l'evoluzione clinica e tecnologica della Terapia del Dolore. Inoltre, in Italia, a differenza di altri Paesi europei, manca un organo competente per l'utilizzo dei sistemi di codifica di ambito clinico che possa aggiornare, formare, monitorare l'utilizzo di sistemi di codi-

fica e classificazione per la compilazione dei documenti sanitari come previsto dalla legislazione corrente.

Lo studio ha messo in evidenza le criticità nell'utilizzo dei sistemi di codifica dedicati in maniera appropriata ed una conseguente difformità nel sistema tariffario attuale che necessita di un confronto fra gli stakeholder di settore al fine di arrivare all'elaborazione di linee guida per le varie patologie trattate e le relative procedure. Tali linee guida dovrebbero consentire di: i) uniformare le modalità di utilizzo dei sistemi di codifica delle procedure e delle diagnosi; ii) definire processi e percorsi condivisi che possano rappresentare la base per le successive valutazioni di qualità effettuabili utilizzando le informazioni contenute nelle SDO; iii) disciplinare il costo delle prestazioni eseguite in considerazione delle tecniche operatorie e dei dispositivi utilizzati.

Alla luce dei risultati emersi, il presente studio evidenzia la necessità di una riconsiderazione profonda della strutturazione di codici di diagnosi e procedure nel settore della Terapia del Dolore, al fine di rivedere e aggiornare periodicamente la classificazione DRG, in collaborazione con i clinici, per includere nella stima dei costi le evoluzioni circa la tipologia delle cure e delle tecnologie utilizzate. Questo consentirebbe da un lato di garantire una maggiore equità e coerenza nella programmazione sanitaria e nell'allocazione delle risorse e, dall'altro, di orientare il sistema verso un maggiore efficientamento, introducendo un meccanismo in grado di controllare la capacità produttiva degli erogatori (Cerbo and Langiano 2004).

Riferimenti bibliografici

- Allegri, Massimo, Carlo Lucioni, Silvio Mazzi, and Giulio Serra. 2015. "Social cost of chronic pain in Italy." *GRHTA Global & regional Health Technology Assesment* 2(1): 33-42. <https://doi.org/10.5301/GRHTA.5000187>.
- Apolone, Giovanni, Oscar Corli, Augusto T. G. Caraceni, Emanuele Negri, Silvia Deandrea, Mauro Montanari, Maria Teresa Greco, and Cancer Pain Outcome Research Study Group (CPOR SG) Investigators. 2009. "Pattern and quality of care of cancer pain management. Results from the Cancer Pain Outcome Research Study Group." *British journal of cancer* 100(10): 1566-74. <https://doi.org/10.1038/sj.bjc.6605053>.
- Breivik, Harald, Beverly Collett, Vittorio Ventafridda, Rob Cohen, and Derek Gallacher. 2006. "Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment." *European journal of pain (London, England)* 10(44): 287-333. <https://doi.org/10.1016/j.ejpain.2005.06.009>.
- Cabina di Regia. 2017. *Programma Nazionale HTA Dispositivi Medici - Documento Strategico*.

- Cerbo, Michele, e Teresa Langiano. 2004. "L'impatto a livello nazionale del sistema di remunerazione." In *I Drg in Italia: un successo o un'occasione perduta?*, a cura di Nicola Falcitelli e Teresa Langiano, 85-115. Bologna: il Mulino.
- Conferenza Stato-Regioni. 2012. "Intesa sulla proposta del Ministro della salute, di cui all'articolo 5 della legge 15 marzo 2010, n. 38, di definizione dei requisiti minimi e delle modalità organizzative necessari per l'accreditamento delle strutture di assistenza ai malati in fase terminale e delle unità di cure palliative e della terapia del dolore. (SALUTE). Intesa ai sensi dell'articolo 8, comma 6, della legge 5 giugno 2003, n. 131." Governo italiano. <http://archivio.statoregioni.it/DettaglioDoc9753.html?ID-Doc=37447&IdProv=10992&tipodoc=2&CONF=CSR>.
- Conferenza Stato-Regioni. 2020. "Accordo 27 luglio 2020 ai sensi della legge 15 marzo 2010, n. 38, sul documento 'Accreditamento delle reti di terapia del dolore'." Roma: Ministero della Salute. https://www.salute.gov.it/imgs/C_17_pubblicazioni_2584_allegato.pdf.
- Hardy, Lynda R. 2022. *Health Informatics. An Interprofessional Approach*. Elsevier.
- ISS (Istituto Superiore di Sanità). 2020. "Sistema di classificazione dei pazienti." 27 gennaio 2020. <https://www.iss.it/sistema-di-classificazione-dei-pazienti>.
- Legge 15 marzo 2010, n. 38. "Disposizioni per garantire l'accesso alle cure palliative e alla terapia del dolore." *Gazzetta Ufficiale Serie Generale*, 19 marzo 2010, n. 65.
- Melotti, Rita Maria, Dekel Samolsky, Gedaliahu Boaz, Francesca Carosi, Ennio Ricchi, Paolo Chiari, Rocco D'Andrea, and Gian Franco Di Nino. 2009. "Categories of congruence between inpatient self-reported pain and nurses evaluation." *European journal of pain (London, England)* 13(9): 992-1000. <https://doi.org/10.1016/j.ejpain.2008.11.019>.
- Ministero della Salute. 2012. "Decreto del Ministero della Salute 18 ottobre 2012. Remunerazione prestazioni di assistenza ospedaliera per acuti, assistenza ospedaliera di riabilitazione e di lungodegenza post acuzie e di assistenza specialistica ambulatoriale." *Gazzetta Ufficiale Serie Generale*, 28 gennaio 2013, n. 23, Suppl. Ordinario n. 8.
- Ministero della Salute. 2018. "Decreto 21 novembre 2018 Aggiornamento dei codici delle specialità cliniche e delle discipline ospedaliere." *Gazzetta Ufficiale Serie Generale*, 14 febbraio 2019, n. 38.
- Ministero della Salute. 2019. *Linee guida sul trattamento del dolore cronico*. Roma: Ministero della Salute.

- Ministero della Salute. 2019-2021. *Piano della Performance 2019-2021*. https://www.salute.gov.it/portale/ministro/documenti/pianoPerformance/Piano_performance_2019-2021.pdf.
- O'Malley, Kimberly, Karon F. Cook, Matt D. Price, Kimberly Raiford Wildes, John F. Hurdle, and Carol M. Ashton. 2005. "Measuring diagnoses: ICD code accuracy. Health services research." *Health services research* 40(5 pt 2): 1620-39. <https://doi.org/10.1111/j.1475-6773.2005.00444.x>.
- Sanità24. 2023. "Dolore cronico, ne soffrono 9,8 milioni di italiani e pesa su privati e Ssn per 62 miliardi." *Il Sole 24 Ore*. 21 dicembre 2023. https://www.sanita24.ilsole24ore.com/art/aziende-e-regioni/2023-12-21/dolore-cronico-ne-soffrono-98-milioni-italiani-e-pesa-privati-e-ssn-62-miliardi-102912.php?uuid=AF0Ppp8B&refresh_ce=1.
- Stubhaug, Audun, Johan Liseth Hansen, Sara Hallberg, Anders Gustavsson, Anne Elise Eggen, and Christopher Sivert Nielsen. 2024. "The costs of chronic pain—Long-term estimates." *European Journal of Pain* 28(6): 960-77. <https://doi.org/10.1002/ejp.2234>.
- Toccaceli, Virgilia, Nadia Francia, Isabella Cascavilla, e Michael Tenti, a cura di. 2019. *Dolore cronico in Italia e suoi correlati psicosociali dalla "Indagine europea sulla salute"*. Rapporti ISTISAN, 23-28.

Sull'affidabilità delle bibliografie generate dai chatbot. Alcune considerazioni

Andrea Capaccioni*

Abstract: The aim of this paper is to examine the use of chatbots in academic settings, focusing on certain aspects of the reliability of bibliographies generated by artificial intelligence (AI). This study uses data obtained from querying ChatGPT 3.5 and Bard to produce a bibliography on a specific topic (Open Access Books). The analysis of the query results, along with comparisons to similar studies, has highlighted the low level of accuracy in the bibliographic citations generated by chatbots and the prevalence of phenomena such as hallucinations and confabulations in AI, which can affect the accuracy of the information. The chatbots use real author and editor names and existing titles that are consistent with the prompt terms, but they frequently use them incorrectly.

Keywords: Artificial Intelligence, Chatbot, Fake References, Hallucinations, Confabulations.

1. Introduzione

Con il lancio del primo chatbot di nuova generazione, nel novembre del 2022, nel mondo accademico è ripresa con vigore la discussione sulle potenzialità dell'intelligenza artificiale (d'ora in poi IA) in ambito scientifico. I chatbot sono programmi di IA basati su modelli linguistici, ovvero software addestrati a ricavare informazioni da grandi quantità di dati in gran parte presi da Internet, che simulano una conversazione con un essere umano e spesso sono utilizzati nei siti web e nelle applicazioni mobili per fornire assistenza ai clienti o per rispondere alle domande degli utenti. Quando i chatbot vengono interrogati, non avviano un processo di ricerca ed estrazione delle informazioni, ma elaborano le loro risposte utilizzando algoritmi che considerano i termini presenti nei prompt e si basano su stime probabilistiche e previsioni statistiche generate da grandi modelli linguistici e da un lungo addestramento¹.

* Dipartimento di Lettere-Lingue, Letterature e Civiltà antiche e moderne, Università degli Studi di Perugia, Italia. andrea.capaccioni@unipg.it.

¹ Per una introduzione generale ai chatbot si rinvia a (Roncaglia 2023, in particolare ai capitoli 9. "Il contesto: IA e reti neurali" e 10. "Le IA generative").

I chatbot hanno il potenziale per migliorare e offrire nuove opportunità nell'ambito dell'istruzione superiore e della ricerca. Tuttavia, sono emersi alcuni aspetti critici. Sono state evidenziate, per esempio, implicazioni di natura etica legate all'impiego nella ricerca e nella didattica (Stokel-Walker 2022). L'UNESCO ha promosso una guida che contiene un utile riepilogo delle principali problematiche nell'ambito dell'*higher education* (Sabzalieva e Valentini 2023). I chatbot sono in grado di generare idee per progetti o fornire indicazioni sulle fonti, assistono nella ricerca in repository e banche dati, traducono in molte lingue, contribuiscono a migliorare la qualità dei testi dal punto di vista linguistico e dello stile, sono di supporto alle attività di revisione paritaria (Srivastava 2023), contribuiscono perfino a generare articoli scientifici (Buruk 2023; Macdonald et al. 2023). Ci sono degli ambiti però in cui queste tecnologie hanno evidenziato dei limiti, per esempio nella capacità di elaborare bibliografie conformi agli standard accademici.

Il presente contributo intende esaminare alcuni aspetti dell'affidabilità delle bibliografie generate dall'IA anche con l'ausilio dei dati ottenuti attraverso l'interrogazione di ChatGPT 3.5 e Bard (attualmente non più disponibile)². Va specificato che si tratta di una semplice query che ha generato un campione limitato che crediamo tuttavia possa fornire una esemplificazione attendibile per comprendere il funzionamento di queste tecnologie e in più costituire la testimonianza di una fase della loro rapida evoluzione, come dimostra la decisione di Google di sostituire Bard con Gemini.

ChatGPT 3.5 e Bard, sviluppati rispettivamente da OpenAI e Google, sono progettati per simulare conversazioni con esseri umani e destinati a un pubblico ampio e per questo più diffusi e facilmente raggiungibili, anche perché accessibili in modo gratuito. Il primo si distingue nei compiti che richiedono la comprensione e la risposta a domande basate su fatti, il secondo si caratterizza per la capacità di generare testi creativi. Ciò si deve anche al fatto che i due chatbot lavorano su differenti set di dati la cui consistenza non siamo in grado di conoscere nel dettaglio in quanto OpenAI e Google forniscono solo informazioni generiche. ChatGPT utilizza un *large supervised training dataset* basato anche su informazioni prese da Internet, utilizzando per esempio i dati raccolti nel Web da soggetti terzi come (Common Crawl, n.d.) (Brown et al. 2020, 3), mentre Google ha addestrato LaMDA (Language Model for Dialogue Applications), il linguaggio utilizzato da Bard, «on a dataset created from public dialog data and other public web documents» (Thoppilan et al., 2022, 4).

Può essere utile aggiungere qualche altra informazione. ChatGPT 3.5 (*chat generative pre-trained transformer*) è stato reso disponibile al pubblico da OpenAI, un'azienda di ricerca e distribuzione di IA fondata da Elon Musk (dimessosi nel 2018), Sam Altman e altri, a fine novembre del 2022 (Mahyoob,

² In Internet si può trovare anche la denominazione Bard 2.0.

Algaraady, e Alblwi 2023). Il chatbot, che può contare su un dataset che ha raggiunto l'ordine dei terabyte, è in grado di apprendere le relazioni tra le parole e di generare testi che si avvicinano a quelli che potrebbero essere elaborati dall'uomo. ChatGPT è basato sull'architettura *transformer*, una rete neurale progettata per compiti di elaborazione del linguaggio naturale (NLP) particolarmente adatta per la generazione di testi. Questo chatbot nella versione da noi presa in esame, non ha accesso, per scelta del produttore, a dati internet aggiornati (il limite è fissato nel mese di settembre del 2021). Per completare l'informazione, diremo che OpenAI ha sviluppato altri due modelli (che non prenderemo in esame), consultabili prevalentemente previa sottoscrizione di un abbonamento, denominati GPT-4 e GPT-4o che presentano diverse funzionalità avanzate.

Bard, il chatbot di Google, è stato lanciato agli inizi del 2023 (Knight 2023) ed è basato su un modello linguistico di intelligenza artificiale definito *factual*, sviluppato da Google AI, in grado di comprendere il significato di parole e frasi, analizzare il contesto di una conversazione, fornire informazioni. Bard, come abbiamo detto, è stato addestrato su una grande quantità di dati e ciò gli ha consentito di generare testi, tradurre, scrivere diversi tipi di contenuti, rispondere alle domande. Dal febbraio del 2024 Bard non è più consultabile nelle modalità originarie, Google ha infatti ufficializzato l'avvio di Gemini (Google, n.d.), un progetto presentato come parte della riorganizzazione dei propri modelli di IA (Pichai e Hassabis 2023).

2. “Create a bibliography”: gli errori dell'IA

Il 17 dicembre del 2023 abbiamo chiesto a ChatGPT 3.5 e Bard di elaborare una bibliografia su di un determinato argomento. I due chatbot sono stati messi alla prova utilizzando il seguente *prompt*, ovvero l'istruzione impartita all'IA in modo che possa compiere l'azione richiesta: “Create a bibliography about Open Access Books”. L'interrogazione è stata formulata in inglese perché online sono disponibili più informazioni in questa lingua (Fig. 1) e per questo motivo abbiamo pensato che avremmo potuto ottenere un numero maggiore di risposte. L'espressione “to create a bibliography” è attestata in lingua inglese, in particolare in ambito universitario (es. “The purpose of this course is to learn how to create a bibliography”; “Students are required to create a bibliography at the end of their research paper, listing all the sources they consulted”). Si è preferito far ricorso a *bibliography* rispetto a *references* (o *list of references*) in quanto, pur essendo termini considerati simili, la bibliografia fa riferimento a un insieme più ampio di informazioni relative alle fonti consultate per la stesura di un articolo scientifico o di una monografia.

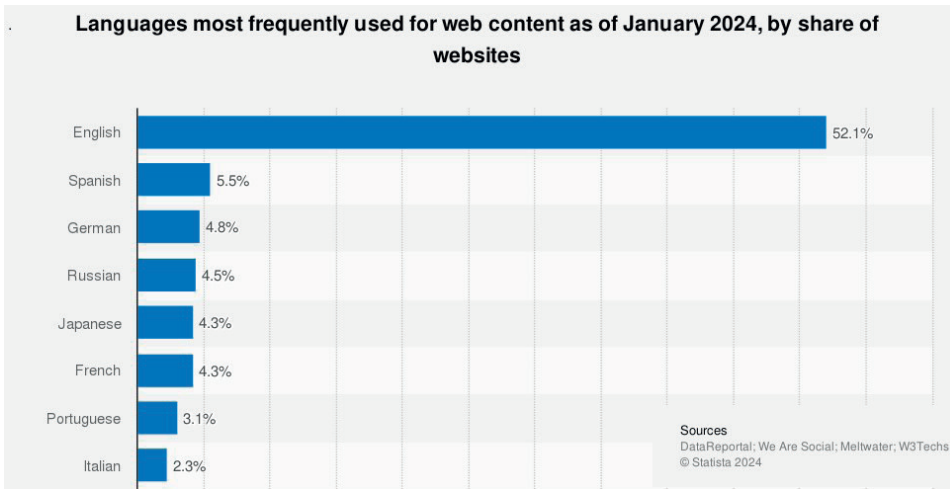


Figura 1: Lingue maggiormente usate per i contenuti nel web (DataReportal; W3Techs 2024).

La scelta dell'oggetto della bibliografia è ricaduta sulla “monografia scientifica ad accesso aperto”, conosciuta in lingua inglese con l'espressione *Open Access Book* (OAB), in quanto è un argomento da noi approfondito in altre sedi e questo ha reso più facile la verifica del grado di correttezza delle citazioni bibliografiche³. Nonostante sia stato scelto un argomento molto specifico e, come in precedenza ricordato, sia stata effettuata una singola interrogazione, riteniamo che i risultati ottenuti possano fornire informazioni utili⁴. In primo luogo, essi documentano il modo in cui i chatbot elaborano una bibliografia in un preciso momento della loro esistenza. Si tratta di un aspetto non trascurabile, in quanto l'IA è in grado di migliorarsi costantemente grazie al feedback degli utenti e all'intervento di azioni correttive (anche umane) introdotte dagli stessi gestori dei chatbot. Inoltre, i dati ottenuti possono contribuire a fornire, anche grazie al confronto con i risultati di altre ricerche (Agrawal, Mackey, e Kalai 2023; Athaluri et al. 2023; Day 2023; McGowan et al. 2023), un quadro più dettagliato sul modo in cui l'IA elabora le bibliografie. Torneremo su questi aspetti nel paragrafo successivo.

Per facilitare la lettura dei risultati abbiamo suddiviso le citazioni bibliografiche elaborate dai chatbot in tre categorie: citazione corretta, citazione falsa, altro argomento (Tab. 1). I dati ottenuti a seguito della formulazione del prompt sono stati trascritti mantenendo lo stile citazionale utilizzato dai chatbot (Tab. 2). Le valutazioni, a seguito della nostra analisi, sono inserite subito

³ Sul tema dell'OAB si veda in particolare (Capaccioni e Ranfa 2023) e (Capaccioni 2021). Altri studiosi hanno scelto criteri diversi per il riscontro dei risultati. Per esempio, Terence Day ha verificato l'affidabilità dei riferimenti bibliografici forniti da ChatGPT prendendo come riferimento cinque articoli pubblicati da una rivista scientifica (Day 2023).

⁴ I risultati di una ricerca impostata in modo simile sono consultabili in (McGowan et al. 2023).

dopo le singole citazioni e sono contrassegnate da uno a tre asterischi: * citazione corretta; ** citazione falsa; *** altro argomento. Tra parentesi rotonde sono indicati gli elementi della citazione che non risultano coerenti. Va inoltre fatta una importante precisazione: in entrambi i chatbot le citazioni risultate false e riferite ad articoli e libri non fanno riferimento allo specifico argomento degli OAB da noi richiesto, bensì a quello più generale dell'Open Access.

<i>citazione corretta</i>	tutti gli elementi (autore, titolo, editore, anno di pubblicazione, ecc.) che compongono la citazione sono corretti.
<i>citazione falsa</i>	uno o più elementi che compongono la citazione non risultano coerenti.
<i>altro argomento</i>	l'opera citata non corrisponde all'argomento richiesto nel prompt.

Tabella 1: Suddivisione delle citazioni bibliografiche.

Nella risposta di ChatGPT sono presenti 10 contributi (tutti accompagnati da un breve abstract, ma privo di indicazione del Digital Object Identifier-DOI): 4 libri, 4 rapporti⁵, 1 articolo; 1 contributo è risultato non attinente all'argomento prescelto. Delle 9 citazioni prese in considerazione ne risultano: 2 corrette e 7 false (Tab. 2).

ChatGPT	Bard
<p>Suber, P. (2012). "Open Access." The MIT Press. *</p> <p>A foundational work by Peter Suber providing an in-depth exploration of the open access movement, its principles, and its impact.</p> <p>Eve, M. P. (2014). "Open Access and the Humanities: Contexts, Controversies, and the Future." Cambridge University Press. *</p> <p>Examines the challenges and opportunities of open access in the humanities, offering insights into the specific context of this field.</p> <p>Schöpfel, J., & Prost, H. (Eds.). (2013). "Open Access to Scientific and Technical Information: State of the Art and Future Trends." De Gruyter Saur. ** (autori, editore, anno)</p>	<p><i>Books</i></p> <p>Suber, P. (2022). Open Access: Why It Matters. MIT Press. ** (titolo, anno)</p> <p>This book provides a comprehensive overview of open access, discussing its history, benefits, and challenges. It also covers the different models of open access publishing.</p> <p>Bawden, D. (2020). Open Access in Theory and Practice: The Theory-Practice Relationship and Openness. Routledge. ** (autori)</p> <p>This book explores the relationship between theory and practice in open access, examining philosophical, legal, and economic aspects of open access alongside its practical implementation.</p>

⁵ Nella categoria "libri" abbiamo compreso le monografie, gli atti di convegno, ecc.; in quella dei "rapporti" abbiamo inserito i contributi che non si configurano né come articoli, né come libri.

<p>A collection of essays providing a comprehensive overview of open access in scientific and technical disciplines, addressing current trends.</p> <p>Chan, L., & Costa, S. (Eds.). (2013). "Open Access and the Humanities: Contexts, Controversies, and the Future." Cambridge University Press. ** (autori, anno)</p> <p>An anthology that delves into the impact and challenges of open access in humanities research, offering diverse perspectives.</p> <p>Morrison, H. (2015). "Sustaining the Knowledge Commons." An OCLC Publication. ** (anno, editore)</p> <p>Explores economic and strategic issues related to sustaining open access initiatives, emphasizing the importance of a knowledge commons.</p> <p>Swan, A. (2010). "The Open Access Citation Advantage: Studies and Results to Date." JISC. ** (editore)</p> <p>A report summarizing studies on the citation advantage of open access articles, contributing to the ongoing discussion on the impact of open access.</p> <p>Jubb, M. (2016). "Monitoring the Transition to Open Access: A Report for Universities UK." Research Information Network. ** (anno)</p> <p>Examines the progress and challenges in the transition to open access in the UK, offering insights into institutional perspectives.</p> <p>OAPEN. (2019). "Open Access Books in the Humanities and Social Sciences: An Overview of the History and Current Situation." OAPEN. ** (autore, titolo, anno, editore)</p> <p>Provides a comprehensive overview of the historical development and current state of open access books in the humanities and social sciences.</p> <p>Bosman, J., Kramer, B., & Otte, W. (2006). "The Economics of Open Access Publishing." Information Economics and Policy, 18(4), 369-381. ** (autore, titolo, editore, anno)</p>	<p>Landsman, K. (2020). Foundations of Quantum Theory: From Classical Concepts to Operator Algebras. Springer Nature. *** (altro argomento)</p> <p>This open access book offers an in-depth exploration of the foundations of quantum theory, starting from classical concepts and progressing to operator algebras.</p> <p>Nordhoff, S. (2019). Cookbook for Open Access Books. Open Book Publishers. ** (editore, anno)</p> <p>This hands-on guide provides practical advice for authors and publishers on how to create and publish open access books.</p> <p><i>Journal Articles</i></p> <p>Solomon, D. J., & Björk, B.-C. (2022). Developing Open Access Journals: A Practical Guide. Routledge. ** (autore/i, editore, anno+ non è un articolo)</p> <p>This practical guide provides step-by-step instructions on how to establish and manage an open access journal.</p> <p>Levy, N. (2022). Bad Beliefs: Why They Happen to Good People. Oxford University Press. *** (altro argomento, non è un articolo)</p> <p>This engaging book explores the psychology of belief formation and the reasons why people sometimes hold false beliefs.</p> <p>Duschinsky, R. (2022). Mentalizing and Epistemic Trust: The Work of Peter Fonagy and Colleagues at the Anna Freud Centre. Oxford University Press. *** (altro argomento, non è un articolo)</p> <p>This book examines the role of mentalizing in human relationships and its impact on epistemic trust.</p> <p><i>Web Resources</i></p> <p>Directory of Open Access Books (DOAB): https://www.doabooks.org/ *</p> <p>DOAB is a community-driven discovery service that indexes and provides access to scholarly, peer-reviewed open access books.</p>
---	--

<p>A scholarly article that conducts an economic analysis of open access publishing, shedding light on the economic aspects of this publishing model.</p> <p>Fuchs, C., & Sandoval, M. (Eds.). (2013). "Critique, Social Media and the Information Society." Routledge. *** (altro argomento)</p> <p>Explores the intersection of open access with social media and the information society, providing a critical perspective on contemporary issues.</p> <p>Remember to check for the latest publications, as the field of open access is dynamic, and new research and insights are continually emerging.</p>	<p>OAPEN: https://www.oapen.org/ *</p> <p>OAPEN is an online library and publication platform that works with publishers to build a quality-controlled collection of open access books.</p> <p>MIT Press Open Access Books: https://mitpress.mit.edu/open-access-at-mit-press/books/ *</p> <p>MIT Press offers a variety of open access books across a range of disciplines.</p> <p>IntechOpen Open Access Books: https://www.intechopen.com/books *</p> <p>IntechOpen is a leading publisher of open access books in science, technology, and medicine.</p> <p>Fonti</p> <p>www.waikato.ac.nz/library/guidance/guides/open-access-information</p>
---	--

Tabella 2: I risultati del prompt: "Create a bibliography about open access books". Legenda: * citazione corretta; ** citazione falsa; *** altro argomento.

Tra parentesi tonde sono indicati gli elementi non corretti.

Osserviamo ora il tipo di manipolazione messo in atto dal chatbot prendendo in esame più da vicino una delle citazioni false che trascriviamo nello stile citazionale originale:

Bosman, J., Kramer, B., & Otte, W. (2006). "The Economics of Open Access Publishing." *Information Economics and Policy*, 18(4), 369-381.

Va fatto notare che la rivista «*Information Economics and Policy*» esiste e presenta il numero 18, fascicolo 4 pubblicato nel 2006, al suo interno però non c'è traccia del contributo segnalato⁶. Con lo stesso titolo si può invece rintracciare un altro articolo, non presente nell'elenco generato dal chatbot, di Christian Zimmermann, *The economics of open access publishing*, pubblicato però nella rivista «*Economic Analysis and Policy*» (39, 1, 2009, pp. 49-53, [https://doi.org/10.1016/S0313-5926\(09\)50054-9](https://doi.org/10.1016/S0313-5926(09)50054-9)).

Da questo accertamento è possibile cogliere alcune caratteristiche del modo in cui ChatGPT genera citazioni false che però apparentemente sembrano vere. Il chatbot impiega nomi di autori ed editori reali e titoli esistenti, spesso coerenti con i termini del *prompt*, ma li utilizza in modo non corretto.

⁶ Cfr. (Science Direct, n.d.). Anche la paginazione dell'articolo risulta errata.

Bard, nella sua risposta, propone 11 contributi (con breve abstract, senza DOI) suddivisi, ed è questo un aspetto che lo differenzia dal modo in cui ChatGPT presenta i dati bibliografici, in categorie: *books*, *journal articles* e *web resources*; con l'aggiunta della segnalazione di una "fonte"⁷. Va notato che il chatbot risulta impreciso nella classificazione delle pubblicazioni in quanto inserisce tre monografie tra le riviste accademiche. Al termine del riscontro risultano utili ai fini della nostra analisi 4 libri e 4 risorse web, mentre 3 contributi non si riferiscono all'argomento prescelto. Sono dunque 8 i contributi presi in considerazione: 4 riferiti a monografie e 4 a risorse reperibili in Internet. Nessuno dei volumi risulta descritto correttamente, le descrizioni bibliografiche presentano le caratteristiche che abbiamo riscontrato in ChatGPT. Bard genera infatti citazioni false ricorrendo all'utilizzo di autori ed editori reali e titoli esistenti. Anche in questo caso ricorriamo a un esempio. Nella bibliografia è presente l'opera, considerata un punto di riferimento per lo studio dell'accesso aperto, *Open access* di Peter Suber:

Suber, P. (2022). *Open Access: Why It Matters*. MIT Press.

La citazione, in apparenza verosimile, contiene ben 2 indicazioni errate: il sottotitolo (*Why it matters*) e l'anno di prima pubblicazione in quanto è indicato il 2022 anziché il 2012. Le 4 risorse web presenti nell'elenco di Bard corrispondono invece a risorse esistenti.

3. Allucinazioni, confabulazioni e accuracy

I risultati ottenuti dall'interrogazione mostrano che ChatGPT ha generato un numero più alto di citazioni false rispetto a Bard. In un'altra ricerca in cui è stato effettuato un confronto simile, nell'ambito però della letteratura in psichiatria, si sono registrati esiti diversi. Secondo (McGowan et al. 2023, 5) ChatGPT ha fornito un elenco di 35 citazioni, di cui solo il 6% corrisponde effettivamente a opere reali, mentre Bard ha generato otto citazioni false. Questi risultati ci dicono che entrambi i chatbot a prescindere dalle loro "competenze" nei diversi ambiti disciplinari, che crediamo possano variare in quanto essi utilizzano differenti basi di dati per il loro addestramento, hanno dimostrato di essere inaffidabili nella generazione di citazioni bibliografiche. Il modo in cui ChatGPT e Bard elaborano le false citazioni bibliografiche è, come abbiamo visto, simile e conferma l'esistenza di un fenomeno noto agli studiosi con il nome di *allucinazione*⁸. Un termine preso in prestito dalla

⁷ Si tratta di un sito universitario che contiene informazioni generali sull'OA (The University of Waikato, n.d.).

⁸ Il fenomeno generale delle allucinazioni dell'IA è molto studiato; per un'ampia introduzione

psicologia e riferito a quelle persone che percepiscono come reale ciò che è immaginario. Nell'ambito dell'IA l'allucinazione (*hallucination*) si può definire «a phenomenon where AI generates a convincing but completely made-up answer» (Athaluri et al. 2023).

Questo tipo di allucinazione può essere attribuita a un problema specifico che presentano i modelli di NLP e indica appunto la generazione di contenuti falsi basati su dati reali. OpenAI ha riconosciuto fin da subito questo tipo di problemi e le difficoltà incontrate nel risolverli:

ChatGPT sometimes writes plausible-sounding but incorrect or nonsensical answers. Fixing this issue is challenging, as: (1) during RL training, there's currently no source of truth; (2) training the model to be more cautious causes it to decline questions that it can answer correctly; and (3) supervised training misleads the model because the ideal answer depends on what the model knows (opens in a new window), rather than what the human demonstrator knows (OpenAI 2022).

L'allucinazione può dunque essere considerata una sorta di "limite" (OpenAI 2022) dell'IA, oppure un "effetto collaterale", o meglio il risultato «di previsioni statistico-probabilistiche basate su grandi modelli linguistici e su un lungo addestramento, in parte autonomo e in parte supervisionato o per rinforzo: non vi è dunque nessuna "copiatura" meccanica delle informazioni incamerate attraverso il corpus di testi di partenza, e non vi è neanche un'operazione di estrazione dal corpus delle informazioni considerate più rilevanti rispetto al prompt dell'utente» (Roncaglia 2023, cap. 11).

Sulla base di queste considerazioni, possiamo comprendere come l'analisi delle allucinazioni presupponga la conoscenza della natura e del funzionamento dei modelli linguistici su cui sono basati i chatbot. Tuttavia, questo fenomeno, in particolare se studiato in ambito bibliografico, non può prescindere dall'osservazione delle risposte che forniscono i chatbot. Seguendo questo metodo, un gruppo di studiosi newyorkesi ha proposto alcune osservazioni che ci aiutano a comprendere meglio il fenomeno (McGowan et al. 2023, 5). Grazie all'esame dei risultati ottenuti a seguito di interrogazioni di tipo "bibliografico" («What has been published regarding suicidal behavior in individuals at high risk for psychosis?») effettuate su ChatGPT e Bard, i ricercatori sono giunti alla conclusione che il riferimento alle allucinazioni in questi casi deve essere considerato improprio. Le allucinazioni sarebbero percettive per natura e verrebbero prodotte dunque senza uno stimolo esterno, mentre le *fake references* dovrebbero essere assimilate a un disturbo nella produzione del linguaggio conosciuto con il nome di confabulazione (*confabulation*). Per i linguisti, la confabulazione consiste nel ricordare in modo errato o inventare del tutto le

dedicata alle allucinazioni generate da modelli di linguaggio naturale (NLG) si veda (Ji et al. 2023) e (Smith 2023).

informazioni per colmare le lacune della memoria⁹. L'esame della correttezza (*accuracy*¹⁰) delle citazioni generate dai chatbot, ovvero la possibilità di determinare la presenza di informazioni accurate e/o di *fabricated citations* (citazioni false o inventate), fornisce un prezioso strumento per testare l'affidabilità di alcune azioni dei modelli di IA e verificarne i progressi.

Perché dunque i chatbot “confabulano”? Gli studiosi americani suppongono che il motivo sia rintracciabile nel fatto che gli *output* (i risultati di un processo) dei modelli linguistici di grandi dimensioni (appartenenti al gruppo dei Large Language Models-LLM) utilizzati dai chatbot sono per natura probabilistici e fanno ricorso a stime basate sulla *semantic similarity* tra due elementi di testo¹¹. Nello specifico, questi chatbot utilizzano modelli di co-occorrenza basati sulla probabilità che due o più parole appaiano vicine in un testo l'una all'altra. Nel momento in cui sono interrogati, i chatbot cercano parole (token) «that co-occur often with query terms in its training data» e poi ne trovano altre che co-occorrono con quelle parole in un processo di iterazione che porta a generare una risposta coerente: «the end result is a very convincing pastiche from up to millions of probabilistically linked bodies of text» (McGowan et al. 2023, 5).

L'ipotesi della confabulazione è convincente, anche se sono auspicabili ulteriori ricerche, in quanto sembra favorire un'interpretazione più accurata delle caratteristiche delle false bibliografie e può così contribuire a contrastare meglio il fenomeno. Da tempo, informatici e linguisti in particolare, stanno lavorando per cercare di correggere questo limite dei chatbot. Una delle tecniche più note è il Reinforcement Learning from Human Feedback (RLHF) che prevede l'intervento dell'uomo sia nel controllo del set di dati di addestramento, sia nella verifica del gradimento (umano) delle risposte generate (OpenAI 2022). Ricordiamo anche la Retrieval-Augmented Generation (RAG) una modalità di addestramento dei LLM che favorisce la fruizione di informazioni esterne in tempo reale permettendo di generare risposte più accurate.

Gli sviluppatori di OpenAI non hanno dubbi sul fatto che nel tempo le allucinazioni e le confabulazioni svaniranno, ma c'è ancora del lavoro da fare. Il gruppo di studiosi americani il cui lavoro abbiamo altre volte menzionato fa notare che la versione di marzo 2023 di ChatGPT 3.5 e quella di giugno 2023 di Bard, entrambe sottoposte alla tecnica del RLHF, mostrano ancora bassi

⁹ Altri studi hanno esaminato il fenomeno della confabulazione proposto dagli studiosi americani, si veda per esempio (Emsley 2023).

¹⁰ Nell'ambito dell'IA con il termine *accuracy* si intende la misura di quanto un sistema informatico progettato per elaborare automaticamente testo in linguaggio naturale produce *output* corretti e fedeli rispetto ai dati di *input* o alle informazioni di riferimento. Si veda anche (Estevez et al. 2022, 9).

¹¹ Come abbiamo visto in precedenza, G. Roncaglia propone un'analisi per molti versi simile (Roncaglia 2023, cap. 11).

tassi di *accuracy* nella generazione delle citazioni bibliografiche (McGowan et al. 2023, 5). Gli stessi risultati da noi raccolti nel dicembre del 2023 ed esposti nel presente contributo confermano questi limiti (Tab. 2).

Gli sforzi compiuti dai produttori di IA per migliorare le prestazioni dei chatbot sono innegabili. Osserviamo anche lo sviluppo di strumenti più specifici che ricorrono all'uso dell'IA (es. Elicit, Paperpile, ecc.) per supportare la comunicazione scientifica. Tuttavia in questa fase, all'interno della quale includiamo anche il periodo qui preso in esame, i chatbot di OpenAI e Google sono caratterizzati da una accentuata sperimentazione che propone diverse soluzioni, ancora non definitive e soprattutto poco efficaci. Durante la seconda metà del 2023, per esempio, è stata segnalata la propensione di ChatGPT 3.5, quando gli è stato chiesto di elaborare una bibliografia su un argomento, a fornire fonti di riferimento generiche (siti web, directory, ecc.), anziché elenchi di citazioni bibliografiche¹². Nei mesi successivi, questa tendenza è stata corretta, e il chatbot è tornato a generare liste di risorse bibliografiche più specifiche¹³.

Gli utenti, in particolare quelli che operano in un contesto accademico, devono essere consapevoli dei limiti dei chatbot e adottare un approccio critico nell'utilizzo delle bibliografie generate da questi strumenti.

Riferimenti bibliografici

Agrawal, Ayush, Lester Mackey, and Adam Tauman Kalai. 2023. "Do Language Models Know When They're Hallucinating References?" *arXiv*, January 2023. <https://doi.org/10.48550/arxiv.2305.18248>.

Athaluri, Sai Anirudh, Sandeep Varma Manthena, V.S.R. Krishna Manoj Kesapragada, Vineel Yarlagadda, Tirth Dave, and Rama Tulasi Siri Duddumpudi. 2023. "Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing Through ChatGPT References." *Curēus* 15 (4). <https://doi.org/10.7759/cureus.37432>.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. "Language Models are Few-Shot Learners." *arXiv*, May 2020. <https://arxiv.org/abs/2005.14165>.

¹² «For example, whereas ChatGPT in March 2023 provided a list of citations, most of which were fabricated, in June 2023, it no longer provided any citations. While this change did not increase the number of accurate citations, it reduced the number of fabricated citations to zero, thereby eliminating the need for time intensive verification» (McGowan et al. 2023, 5).

¹³ Abbiamo sottoposto a ChatGPT 3.5 in data 14 giugno 2023 il *prompt* già usato in precedenza: "Create a bibliography about Open Access Books", il risultato è stato un elenco di risorse bibliografiche (di cui non abbiamo controllato il livello di accuratezza). Nello stesso giorno la medesima richiesta è stata rivolta a Google Gemini che ha invece generato un breve elenco di fonti.

- Buruk, Oğuz Oz. 2023. "Academic Writing with GPT-3.5 (ChatGPT): Reflections on Practices, Efficacy and Transparency". In *26th International Academic Mindtrek Conference*, 144-53. Tampere, Finland: ACM. <https://doi.org/10.1145/3616961.3616992>.
- Capaccioni, Andrea. 2021. "Origins and Developments of the Open Access Books." In *Digital Libraries. Advancing Open Science*. <https://doi.org/10.5772/intechopen.95357>.
- Capaccioni, Andrea, e Elena Ranfa. 2023. "Gli editori e le monografie scientifiche ad accesso aperto: un rapporto in evoluzione." *Sistema Editoria. Rivista internazionale di studi sulla contemporaneità* 1 (1): 79-94. <https://doi.org/10.14672/se.v1i1.2124>.
- Common Crawl. n.d. Consultato il 19 maggio 2024. <https://commoncrawl.org/>.
- DataReportal; W3Techs. 2024. "Languages most frequently used for web content as of January 2024, by share of websites." Statista, January 2024. <https://www.statista.com/statistics/262946/most-common-languages-on-the-internet/>.
- Day, Terence. 2023. "A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT." *The Professional Geographer* 75 (6): 1024-27. <https://doi.org/10.1080/00330124.2023.2190373>.
- Emsley, Robin 2023. "ChatGPT: these are not hallucinations – they're fabrications and falsifications." *Schizophrenia* 9 (1). <https://doi.org/10.1038/s41537-023-00379-4>.
- Estevez, Almenzar Marina, David Fernandez Llorca, Emilia Gomez Gutierrez, and Fernando Martinez Plumed. 2022. "Glossary of Human-centric Artificial Intelligence." *JRC Publications Repository*. <https://doi.org/10.2760/86066>.
- Google. n.d. "Gemini." Consultato il 19 maggio 2024. <https://gemini.google.com>.
- Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Yejin Bang, Andrea Madotto, and Pascale Fung. 2023. "Survey of Hallucination in Natural Language Generation." *ACM Computing Surveys* 55 (12): 1-38. <https://doi.org/10.1145/3571730>.
- Knight, Will. 2023. "Meet Bard, Google's Answer to ChatGPT." *WIRED*, February 6, 2023. <https://www.wired.com/story/meet-bard-googles-answer-to-chatgpt/>.
- Macdonald, Calum, Davies Adeloye, Aziz Sheikh, and Igor Rudan. 2023. "Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis." *Journal of Global Health* 13 (February). <https://doi.org/10.7189/jogh.13.01003>.

- Mahyoob, Mohammad, Jeehaan Algaraady, and Abdulaziz Alblwi. 2023. "Proposed Framework for Human-like Language Processing of ChatGPT in Academic Writing." *International Journal of Emerging Technologies in Learning* 18 (14): 282-93. <https://doi.org/10.3991/ijet.v18i14.41725>.
- McGowan, Alessia, Yunlai Gui, Matthew Dobbs, Sophia Shuster, Matthew Cotter, Alexandria Selloni, Marianne Goodman, Agrima Srivastava, Guillermo A. Cecchi, and Cheryl M. Corcoran. 2023. "ChatGPT and Bard Exhibit Spontaneous Citation Fabrication during Psychiatry Literature Search." *Psychiatry Research* 326 (August): 115334. <https://doi.org/10.1016/j.psychres.2023.115334>.
- Open AI. 2022. "Introducing ChatGPT." November 30. <https://openai.com/blog/chatgpt/>.
- Pichai, Sundar, and Demis Hassabis. 2023. "Introducing Gemini: Our largest and most capable AI model." December 06. <https://blog.google/technology/ai/google-gemini-ai/#sundar-note>.
- Roncaglia, Gino. 2023. *L'architetto e l'oracolo: Forme digitali del sapere da Wikipedia a ChatGPT*. Bari; Roma: GLF editori Laterza. EPUB.
- Sabzalieva, Emma, and Arianna Valentini. 2023. *ChatGPT and artificial intelligence in higher education: quick start guide*. UNESCO International Institute for Higher Education in Latin America and the Caribbean. <https://unesdoc.unesco.org/ark:/48223/pf0000385146>.
- Science Direct. n.d. "Information Economics and Policy." Consultato il 19 maggio 2024. <https://www.sciencedirect.com/journal/information-economics-and-policy>.
- Smith, Craig S. 2023. "Hallucinations could blunt ChatGPT's success." *IEEE Spectrum*, March 29. <https://spectrum.ieee.org/ai-hallucination>.
- Srivastava, Mashrin. 2023. "A day in the life of ChatGPT as an academic reviewer: Investigating the potential of large language model for scientific literature review." *Preprint. Open Science Framework*, February 16. <https://doi.org/10.31219/osf.io/wydct>.
- Stokel-Walker, Chris. 2022. "AI Bot ChatGPT Writes Smart Essays-Should Professors Worry?" *Nature*, December 9. <https://doi.org/10.1038/d41586-022-04397-7>.
- The University of Waikato. n.d. "Open Research." Consultato il 19 maggio 2024. www.waikato.ac.nz/library/guidance/guides/open-access-information.
- Thoppilan, Romal, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, et al. 2022. "LAMDA: Language Models for Dialog Applications." *arXiv*, January 2022. <https://doi.org/10.48550/arxiv.2201.08239>.

La dimensione Testuale del Videogioco

Classificazione dei *transcript* dei videogiochi basata sul lessico

Alessandro Maisto*

Abstract: In this work, we explore the textual dimension of video games. Despite their pronounced visual and interactive characteristics, video games can be regarded as documents due to their narrative and communicative elements. Our research delves into this textual dimension to automatically generate rating tags associated with offensive language, violence, and the presence of drugs. We utilized a dictionary of English slang, compiled from various online sources and manually annotated with four categories: Slang, Violence, Drugs, and Discrimination. The resulting electronic dictionary facilitated the automatic assignment of the three rating tags with high precision. It has also been employed to classify video games based on their lexical content. The two classification tasks – by rating tags and by lexical dimension – could pave the way for an automatic warning system capable of analyzing the full textual dimension of a video game.

Keywords: Videogame transcript corpus, Automatic videogame rating, Bad language, Slang dictionary, Text analysis.

1. Introduzione

In questo lavoro proponiamo uno studio computazionale della dimensione testuale dei videogiochi, esplorando, in particolare, il potere classificatorio del linguaggio volgare, violento e discriminatorio associato allo slang presente nei dialoghi dei videogiochi in inglese. In particolare, si offrirà una analisi approfondita della dimensione testuale dei videogiochi, analizzata in ottica di classificazione automatica del rating di età tramite un approccio di estrazione terminologica, basato sullo slang.

Per definizione, il videogioco è un prodotto multimediale interattivo. Dal punto di vista puramente scientifico, le indagini sui videogiochi si sono concentrate sulla loro dimensione sociologica, sul design e sulla loro costruzione ed in parte anche sugli aspetti psicologici della loro fruizione. Tuttavia, in

* Dipartimento di Scienze Politiche e della Comunicazione, Università degli Studi di Salerno, Fisciano (SA), Italia. amaisto@unisa.it

quanto prodotto multimediale, il videogame può essere considerato anche un documento (Aarseth 2004; Juul 2005), provvisto di una forte correlazione con immagini, suoni e altri elementi extra-linguistici, ma con una chiara e spiccata dimensione testuale.

Il testo, all'interno di un videogame, è diventato sempre più centrale man mano che la complessità di questi prodotti di intrattenimento (ma anche culturali) è andata crescendo. Volendo identificare i momenti di evoluzione del fenomeno videoludico negli ultimi 50 anni, cioè dalla loro nascita fino ad oggi, possiamo identificare diverse fasi durante le quali il testo ha assunto una certa importanza: benché i primi videogiochi fossero sostanzialmente costruiti grafici, dove l'intrattenimento era basato principalmente nello sviluppare una coordinazione occhio/mano tale da sconfiggere la macchina, già nei primi anni si svilupparono una quantità di giochi di natura puramente testuale. Questi giochi, definiti *avventure testuali* iniziarono a diffondersi negli anni Ottanta anche sulla scia del successo dei cosiddetti *giochi di ruolo* (GDR, o RPG dall'inglese *Role-Playing Game*) (Heritage 2021). Consistevano, sostanzialmente, nella descrizione testuale di una situazione e nella richiesta di un comando testuale specifico per risolvere la situazione.

Successivamente, questo tipo di gioco si è evoluto, trasformandosi in quelle che furono definite *avventure grafiche*, dove il testo era ancora un elemento centrale del gioco, ma veniva accompagnato da un meccanismo *punta e clicca* localizzato in un riquadro grafico (Marchiori et al. 2011).

Con lo sviluppo delle *console* e con l'avanzamento dei sistemi informatici ed il conseguente arricchimento dei comparti grafici, le componenti extra-linguistiche presero il sopravvento, ma senza poter mai escludere completamente la dimensione linguistica. Trattandosi di esperienze simulate, narrative e comunicative, infatti, il linguaggio non poteva che conservare la propria importanza, soprattutto in generi come il *gioco di ruolo*.

Ridotto alla propria dimensione testuale (i cosiddetti *transcript*) il videogame deve essere considerato un documento e può essere analizzato da un punto di vista linguistico attraverso una analisi computazionale.

In quanto documento, il videogame ha delle caratteristiche che lo distinguono da tutte le altre tipologie di documenti, inclusi quelli più simili legati, ad esempio, al linguaggio dei prodotti audiovisivi. Uno di questi elementi è, indubbiamente, l'interattività, intesa come la possibilità, da parte del giocatore, di scegliere una di varie possibili frasi in risposta ad un dialogo.

L'importanza della dimensione linguistica dei videogiochi è attestata anche dalla presenza, all'interno degli indicatori utilizzati per attribuire un *rating* di età minima del giocatore, di elementi di chiara matrice testuale. Uno di questi, ad esempio, riguarda il linguaggio volgare (*Bad Language*) come proposto dal sistema Europeo *Pan European Game Information* (PEGI), ma anche la discriminazione, che viene espressa soprattutto a livello linguistico o l'uso di Dro-

ghe. Tra questi indicatori, quello del linguaggio volgare è, inoltre, tra quelli più comunemente riscontrati assieme all'indicatore di violenza.

In questo contesto, dunque, possiamo sintetizzare gli obiettivi dell'articolo come segue:

1. Presentare una risorsa in lingua inglese per lo studio della dimensione testuale dei videogame comprendente un dizionario elettronico dello slang ed un corpus composto da 51 *transcript* di videogiochi per un totale di oltre 700.000 token;
2. Proporre un approccio classificatorio automatico introduttivo che possa dimostrare l'efficacia dell'analisi linguistica nell'attribuzione di etichette di *warning* ai videogiochi.

Nel secondo paragrafo illustreremo brevemente il funzionamento del sistema PEGI ed i suoi principali difetti. Successivamente, nel paragrafo 3, parleremo degli studi effettuati sulla dimensione testuale del videogioco. Il paragrafo 4 descriverà nel dettaglio la risorsa sviluppata, partendo dal dizionario elettronico per poi arrivare al corpus. Il paragrafo 5 illustrerà la metodologia di classificazione adottata e l'esperimento effettuato sul corpus. In fine, nel paragrafo 6 concluderemo mostrando possibili sviluppi e implicazioni del nostro lavoro.

2. Il Sistema PEGI

I sistemi di valutazione dei videogiochi (*Video Game Rating Systems*) rappresentano un indicatore fondamentale per ragioni sociali, psicologiche, mediche e per aspetti legali. La multidisciplinarietà delle implicazioni dell'attribuzione di questo *rating* è attestata anche dalla varietà di settori di provenienza dei professionisti che si occupano di determinare le metodologie di valutazione. Le valutazioni, poi, sono di competenza dei produttori dei videogiochi che affidano a *tester* e ad esperti di giurisprudenza una eventuale revisione del rating.

Il Sistema di valutazione PEGI prevede, infatti, che i produttori propongano una lista di descrittori per il proprio prodotto sotto forma di risposte ad un questionario contenente 37 domande. Un sistema automatico individua nelle risposte la presenza di contenuti inappropriati proponendo l'aggiunta o la rimozione di alcuni indicatori. Alla fine di questo procedimento automatico avviene una fase di verifica manuale che può ulteriormente modificare le etichette attribuite al gioco (Felini 2015). Solo in poche occasioni, il software e i suoi *transcript* vengono effettivamente analizzati da un'audience specializzata e diversificata.

Il sistema PEGI utilizza una lista ristretta di etichette che include Linguaggio volgare, Paura, Sesso/nudità, Droghe, Discriminazione, Gioco d'azzardo, Violenza e Acquisti in gioco (dal 2018).

Le etichette di età sono numeriche e indicano l'età minima consigliata per un videogioco. Queste etichette sono cinque e includono 3, 7, 12, 16 e 18 anni.

Sottoporre i propri prodotti ad un *rating systems* come PEGI favorisce i produttori, che mostrano il proprio interesse per la salvaguardia dei clienti, evitando di rilasciare prodotti di dubbia moralità. Inoltre, trattandosi di un sistema europeo, si riduce il rischio di incorrere in sanzioni dai singoli stati dell'Unione. Infatti, nel mondo, esistono diversi organismi nazionali o sovranazionali che valutano l'impatto dei videogiochi e ne regolano l'utilizzo, e ciascuno di essi tende ad utilizzare indicatori diversi e a proporre punteggi diversi. Un videogioco come *The Sims 4* ad esempio, viene classificato da PEGI come 12+ a causa di "violenza non realistica tra umani", mentre è considerato 18+ da RARS, che è l'organismo che opera in Russia, *Matures (violence and sexual references)* da ACB (Australia) e *Teens (obscene humor, sexual themes, violence)* da ESRB (Nord America).

Tra le critiche principali a questi sistemi, tuttavia, possiamo annoverare il fatto che i *rating* sono attribuiti sulla base di una conoscenza parziale del prodotto. Diversi studiosi del settore hanno suggerito una revisione dei processi e dei criteri di valutazione, favorendo un approccio più dettagliato e preciso (Dogruel e Joeckel 2013; Felini 2015; Humphreys e Wang 2017).

Un approccio testuale, ad esempio, favorirebbe un'analisi completa e rapida dei *transcripts* portando alla luce possibili elementi importanti ai fini della valutazione del gioco sia dal punto di vista dall'uso di linguaggio volgare che da altri punti di vista (identificazione di indicatori testuali di violenza e nomi di droghe). L'uso di strumenti di questo tipo non dovrebbe sostituire un'analisi manuale del gioco, ma suggerire la presenza di tali indicatori al valutatore e favorire la loro ricerca nell'intero documento, e non solo in parziali campioni di testo.

3. Contesto e studi correlati

I principali studi sul linguaggio e sui videogiochi riguardano il loro uso nell'apprendimento di una lingua straniera (Klimova e Jaroslav 2017). In questo senso, però, numerosi studi si sono concentrati su giochi sviluppati appositamente per l'apprendimento e dunque per il cosiddetto *Game-Based Learning* (Wright, Betteridge, and Buckb 2006).

Esiste, tuttavia, un filone afferente alla Linguistica dei Corpora che considera il linguaggio del videogioco come base per studi linguistici e sociolinguistici di specifici fenomeni. Ne è un esempio il lavoro di Ray (2019) che ha analizza-

to le traduzioni della terminologia fantascientifica all'interno dei videogiochi. Fekete e Porkolab (2019) hanno utilizzato una metodologia computazionale per l'analisi dei toponimi nella serie di videogiochi *The Elder Scrolls*. In particolare, i due autori hanno esaminato oltre 1.000 nomi di luoghi inventati da un punto di vista morfologico ed onomastico.

Heritage (2021) ha, invece, studiato il problema del genere all'interno di diversi videogiochi, dimostrando l'importanza delle rappresentazioni di genere nel dominio e la varietà di sfumature che vengono attribuite a personaggi maschili e femminili in un corpus di oltre 300 mila token estratti da 10 diversi videogiochi.

Ensslin (2012), come anche Gee (2014), ha analizzato i videogiochi dal punto di vista dell'analisi del discorso. Ensslin, in particolare, ha dimostrato come questi prodotti multimediali funzionino come mezzo e obiettivo di comunicazione, favorendo la nascita di nuovi vocabolari, nuovi significati, generi testuali e pratiche discorsive.

Goorimoorthee (2019) ha analizzato gli accenti dei personaggi non giocatori dei videogiochi allo scopo di verificare l'assenza di diversità culturale nei giochi della *BioWare*, uno studio di produzione che ha fatto dell'inclusività sociale dei propri titoli un elemento distintivo.

A mettere in relazione slang e videogiochi c'è il lavoro di Balteiro (2019), che però analizza le conversazioni tra giocatori e non la dimensione testuale del videogioco. Ensslin e Finnegan (2019), invece si sono concentrati sull'uso del linguaggio volgare nelle cosiddette comunità "co-sit" (co-situate) di giocatori.

L'approccio proposto nel presente articolo si ispira fortemente al lavoro di Maisto et al. (2021), proponendo, tuttavia, una serie di modifiche che hanno condotto al miglioramento dei risultati ottenuti. In particolare, in Maisto et al. le etichette di *Bad Language*, *Violence* e *Drugs* erano estratte sulla base della frequenza relativa, mentre in questo lavoro proporremo un approccio più efficace basato esclusivamente sul conteggio dei termini. Infatti, l'idea alla base di queste modifiche è che, indipendentemente dalla dimensione del testo, la sola presenza di poche espressioni volgari, violente o discriminatorie possa rappresentare un indicatore sufficiente all'attribuzione dell'etichetta corrispondente. Sono stati aggiunti, inoltre, una serie di termini referenti alla sfera sessuale.

4. Risorse linguistiche

Lo slang rappresenta una sfida nella lessicografia ed in particolare nello sviluppo di dizionari specifici. James B. McMillan (1978) considera il problema del lessico slang come un problema riguardante la sua stessa definizione. La varietà di etichette e l'impossibilità di suddividere lo slang in sottocategorie rende impossibile una sua analisi di tipo linguistico, sociolinguistico o psico-

linguistico. Ciononostante, negli anni ne sono state date varie definizioni e sono stati prodotti diversi studi sul tema.

Benché sia stato spesso considerato un fenomeno marginale o bollato come volgare (De Klerk 1990), lo slang ha una connotazione sociopsicologica che emerge nelle varie definizioni che ne sono state date nel corso degli anni. Citando Amari (2010), Walt Whitman nel 1885 definì questo registro linguistico come «un tentativo dell'umanità di fuggire dal puro letteralismo e di esprimersi senza limiti». Più recentemente, il *Concise Oxford Dictionary* (2006) ha definito lo slang come «un linguaggio informale più comunemente parlato che scritto che è tipicamente ristretto ad un particolare contesto o gruppo». *L'Historical Dictionary of American Slang* (Lighter 1994) sottolinea come il suo carattere irriverente faccia dello slang uno strumento di opposizione all'autorità e alla conformità linguistica, non per forza di natura sovversiva, ma utile ad indicare un'attitudine comune che permette il riconoscimento di un individuo all'interno di un particolare gruppo sociale (Amari 2010).

Partridge (1933) elencava diverse ragioni dell'utilizzo dello slang, tra cui il gioco o il divertimento, la sdrammatizzazione, un uso per differenziarsi, per essere pittoreschi, per brevità, per arricchire il linguaggio, per rendere più concreto l'astratto, per intimare con qualcuno, mostrando la propria appartenenza ad un gruppo e per comunicare nel codice segreto di quel gruppo. Per queste sue caratteristiche lo slang tende naturalmente ad esprimere concetti legati a criminalità, stupefacenti o sesso, utilizzando spesso un tono "indecoroso" (Dumas e Lighter 1978). Secondo Green (2011) oltre 5.000 parole appartenenti al lessico dello slang riguardano la criminalità, mentre quelle che identificano le droghe sono quasi 4.000. Numerosissimi sono anche i termini razzisti (circa 1.500, di cui la maggior parte riferita alle persone di colore o agli ebrei) e termini che riguardano la morte (oltre 1.000).

Lo slang, inoltre, è un tipo di linguaggio che fa dell'ambiguità, della metafora e del cambiamento di significato il proprio elemento distintivo (Mattiello 2005). Queste caratteristiche lo avvicinano al linguaggio dell'arte e della poesia (Amari 2010). Lo slang annovera termini di uso più o meno comune utilizzati in senso figurato e spesso dispregiativo. Questi usi nascono da una "esuberanza linguistica" che porta all'introduzione di nuove parole o di nuovi significati a parole già esistenti (Mencken 1967). Ne sono un esempio emblematico numerosi nomi di droghe (*coke* per *cocaina*; *baby* per *marijuana*, ecc.), relativi al sesso (*banana* per *pene*, *backdoor* per *ano*, *fruit* per *omosessuale*), alle bevande alcoliche (*jar*, *piss* o *belt*). Questa ambiguità, che rende oscuro, indeterminato o poco trasparente il significato di queste parole, deriva da meccanismi associativi complessi ed eterogenei, che spaziano dalla metonimia (*jar* per bevanda alcolica), alla metafora o da processi più oscuri che coinvolgono altri slang (*belt* è associato ad esempio al termine "a hit" che indica metaforicamente un cattivo post-sbornia) (Mattiello 2005).

Le caratteristiche qui citate, fanno dello slang un ottimo indicatore di elementi caratterizzanti alcune delle etichette utilizzate da PEGI e da altri sistemi di *rating* per classificare i videogiochi. Tuttavia, per poter rendere il lessico slang fruttuoso a questo scopo è necessario suddividerlo in diverse parti, ciascuna riguardante uno degli ambiti di valutazione del rating di un prodotto e ibridarlo con termini scientifici o di gergo estratti da domini specifici come quello delle droghe o quello militare.

SsVdd (*Slang-sex-Violence-drugs and discrimination*) è una risorsa lessicale terminologica costruita per la valutazione e la classificazione dei *transcript* dei videogiochi. La risorsa è stata sviluppata in lingua inglese e contiene oltre 6.000 termini monorematici e polirematici a cui sono stati associati quattro tag riguardanti le diverse categorie di appartenenza.

Come mostrato dalla tabella 1, la risorsa contiene le seguenti categorie di parole:

- *Slang (generico)*: consiste in una raccolta di termini volgari o generalmente associati al cosiddetto *bad language*. Al suo interno troviamo termini riguardanti insulti di ogni tipo o relativi a discriminazione, sesso e sessismo. Tra questi termini troviamo nomi come *asshole* (stronzo), *bollocks* (bugie), ma anche *dinkey* (pene), verbi come *chup* (disapprovare) o *uneat* (vomitare);
- *Sesso*: il tag Sex contiene parole con espliciti riferimenti ad atti sessuali come, ad esempio, *carnal knowledge* (conoscenza carnale), *orgy* (orgia) o *striptease*;
- *Violenza*: con questo tag sono identificati termini standard e gergali relazionati alla violenza, come, tra gli altri, i nomi di armi o di oggetti associati alla guerra. I termini identificati dal tag *violenza* includono parole come *kill* (uccidere), *ammunition* (munizioni), *assassin* (assassino), *hatred* (odio), *terrorism* (terrorismo);
- *Droghie*: contiene termini generici, commerciali, scientifici e gergali che indicano le droghe o il loro consumo. In questa sezione troviamo *methamphetamine*, *cocaine*, *pentobarbital* ma anche *sniff* (sniffare), *nose candy* (droghe inalate), *speedball*;
- *Discriminazione*: si tratta di termini che riguardano discriminazioni di genere o di natura etnica, regionale o religiosa. Troviamo termini come *burrhead* (persona di colore), *chinky* (di origine asiatica), *guido* (italo-americano), *shiksa* (donna attratta da persone di origine ebrea), *schizo* (schizofrenico), *wacko* (pazzo o drogato).

Categoria	Tag	Numero
Slang	Slang	3.795
Sesso	Sex	208
Indicatori lessicali di Violenza	Violence	635
Nomi di droghe	Drugs	533
Termini Discriminatori	Discrimination	1.094
TOTALE		6.122

Tabella 1: Suddivisione in categorie del dizionario sVdd.

La fonte dalla quale sono stati estratti i termini del dizionario *SsVdd* è, principalmente, il web, ed in particolare alcune pagine che riportano termini associati e neologismi dello slang. Le pagine esaminate sono *urbandictionary.com*, *theonlineSlangdictionary.com*, *peevish.co.uk/Slang*. Per i nomi di droghe, è stato necessario affidarsi alle definizioni di *wikipedia*. Molte definizioni sono state poi verificate nella terza edizione del *NTC's dictionary of American Slang and colloquial expressions* (Spears 2000).

Il dizionario necessita di un costante aggiornamento per poter conservare la propria efficacia pratica, dato che lo slang è un linguaggio in costante mutazione (Mattiello 2005; Benedict e Munro 1997; Eble 1996; Andersson e Trufgill 1992), che genera numerosi nuovi termini, li dimentica o ne modifica il significato in un breve lasso di tempo. Si tratta di un linguaggio estremamente sensibile ai mutamenti sociali, culturali, di costume e perfino tecnologici. Una variabilità simile è possibile trovarla anche nei nomi di droghe, dato che è possibile che, con il tempo, vengano prodotte nuove sostanze stupefacenti.

Come abbiamo detto, il dizionario include sia parole semplici che parole composte polirematiche. In particolare, le parole composte, o *MultiWord Expressions* (MWE) risultano essere più efficaci nella ricerca semantica delle espressioni di slang, perché sono generalmente meno ambigue (Elia 2002). Nei testi è possibile riscontrare determinati termini sia con uso ordinario e dunque privo di volgarità, sia nella loro accezione più volgare. Il termine *beaver* indica comunemente il castoro, ma in determinati contesti viene utilizzato come sinonimo di vagina. Allo stesso termine è possibile riferirsi anche tramite delle MWE che di norma sono ambigue, ma che poco probabilmente troveremo con alta frequenza nel loro significato ordinario. Un esempio è il termine *bearded oyster*. Tuttavia, nel linguaggio slang, molte MWE risultano comunque di elevata ambiguità e dunque di difficile interpretazione. Sempre relativa ai genitali femminili, infatti possiamo trovare il termine *promised land*, o riferito ad andare di corpo il termine *make a deposit*.

Questo fenomeno risulta chiaro anche ad una rapida analisi delle concordanze all'interno dei corpora dell'inglese. Il termine *bearded oyster* appare nel corpus *enTenTen21* (Jakubiček et al. 2013) 11 volte, di cui almeno 6 con il

significato di *vagina* («hint: a friend was on #21 and we hadn't even seen the Pussyfooters, the Camel Toe High Steppers, the Bearded Oysters», *thechicory.com*), in un caso con il significato di uomo barbuto («The father, a venerable, bearded oyster, of august appearance and solemn deportment, was much mortified that one of his family should happen to be so sickly», *gutenberg.ca*) e 4 in senso letterale («Add a dozen bearded oysters, that have been well rinsed in their own liquor», *chestofbooks.com*). Per *promised land* (66.492 occorrenze) o per *make a deposit* (circa 1.000 occorrenze), la stragrande maggioranza dei risultati riguarda il significato ordinario delle due sequenze di parole.

I quattro tag proposti all'interno del nostro dizionario non devono essere univoci, nel senso che ad ogni parola non deve essere associato un solo tag. I tag multipli possono derivare sia dalla ricchezza insita nel lessico dello slang e dunque nella polisemia che contraddistingue la maggior parte dei termini, ma anche dal fatto che una serie di termini definiti slang possono identificare crimini e criminali, droghe, forze dell'ordine, morte o assassinio, abuso e discriminazione razziale. Un esempio di entrata del nostro dizionario è il seguente:

ass crack;Slang
badonkadonk;Slang+Sex
batty;Slang+Discrimination

Nell'esempio sopra vediamo che il lemma è separato dalle informazioni ad esso relative da un punto e virgola. Ad ogni entrata del dizionario sono stati associati i cinque tag relativi alle classi individuate (*Slang*, *Sex*, *Violence*, *Drugs* e *Discrimination*).

Un termine, quindi, può ricevere un doppio tag sia per via della propria polisemia, sia, semplicemente, perché appartiene contemporaneamente alla categoria Slang e a quella Sex. *badonkadonk* ad esempio appartiene allo slang, ma indica contemporaneamente il sedere di una donna afroamericana, mentre *batty* può essere utilizzato sia per indicare un sedere che un omosessuale.

I tag multipli rappresentano il 2,25% del dizionario (140 parole) e riguardano principalmente i termini *slang* (128) che ricevono un secondo tag di tipo *sex* (27), *drugs* (17), *discrimination* (56) o *violence* (28). Ci sono poi 9 termini etichettati come *violence* che ricevono il tag *drugs* 5 volte, quello *discrimination* 3 volte ed una volta quello *sex*. Anche il tag *drugs* si trova in associazione con il tag *discrimination* in 3 occasioni.

Il dizionario elettronico è liberamente consultabile e scaricabile alla pagina (Github, n.d.a.).

I *transcript* selezionati per l'esperimento sono stati rintracciati sul web in varie piattaforme in cui utenti comuni e appassionati possono condividere i file contenenti i dialoghi o i testi dei videogiochi raccolti manualmente durante l'esperienza di gioco. Per le caratteristiche interattive dei videogiochi, infatti, non è facile estrarre ogni possibile dialogo, dato che spesso il giocatore

ha diverse possibilità che, come detto, generano risposte diverse e guidano la conversazione verso sponde diverse. Secondo Heritage (2021), esistono quattro metodi per estrarre i dialoghi dei videogiochi: utilizzare un software che estrae tutte le battute, giocare ogni possibile permutazione di testo ed annotarla, usare i *transcripts* creati dai fan o siti web specializzati che hanno utilizzato una delle prime tre opzioni.

I portali utilizzati per la costruzione del nostro corpus sono principalmente due: *transcripts.fandom.com* and *game-scripts.fandom.com*. Da questi due portali sono stati estratti 50 *transcripts* relativi a giochi pubblicati nel lasso temporale che va dal 1998 al 2020. Le dimensioni totali del corpus sono di 701.384 token.

La scelta dei titoli selezionati è stata effettuata sulla base di una serie di criteri non rigidi, tra cui figura, innanzitutto, la disponibilità del *transcript* al momento della raccolta. Verificata la disponibilità sulle piattaforme indicate sopra, i successivi criteri di scelta adottati sono stati i seguenti:

1. *Popolarità*: i titoli inseriti nel corpus hanno goduto o godono tutt'ora di buona popolarità o sono riconosciuti universalmente come capisaldi di determinati generi di videogiochi;
2. *Temi e Generi*: nella selezione effettuata si è cercato di includere titoli appartenenti a quanti più generi differenti, facendo però attenzione anche alle tematiche o alle diverse ambientazioni. La maggioranza dei giochi, tuttavia, appartiene ai generi in cui la dimensione testuale ha una maggiore importanza come gli RPG o i giochi di *Azione/Avventura*;
3. *Ratings*: in ultimo sono stati raccolti titoli con valutazioni differenti, dai giochi adatti a tutti (3+) a quelli esclusivamente per adulti (18). Vista la natura dell'esperimento che illustreremo nelle prossime pagine, tuttavia, abbiamo prediletto titoli che permettessero di far risaltare la presenza/assenza degli indicatori che cercheremo di evidenziare automaticamente.

Le date di pubblicazione dei titoli selezionati tendono a concentrarsi negli anni a partire dal 2010. Un buon numero di titoli (16) sono del primo decennio del 2000, mentre solo 2 titoli sono precedenti. Questa disparità è dovuta principalmente alla difficoltà di rintracciare i *transcript* dei videogiochi troppo datati, dovuta al fatto che le *fandom* utilizzate sono sorte solo in un momento successivo.

Per quanto riguarda la classificazione PEGI, abbiamo una predominanza di titoli per adulti (PEGI 18) seguita da titoli per adolescenti (PEGI 16) e bambini in età scolare (PEGI 7 e 7+). Solo pochi titoli sono riservati a bambini (PEGI 3) e preadolescenti (PEGI 12). In questo caso, si è scelto di privilegiare titoli più corposi dal punto di vista testuale, i cui *transcript* contenessero un numero di token abbastanza alto. La media dei token per *transcript* più alta,

infatti, riguarda i videogiochi per adulti (circa 20.000 token) seguita dai PEGI 16 (circa 17.000). Le altre categorie hanno medie estremamente più basse (intorno ai 4.000 token ciascuna), ma sono state inserite proprio come importante controprova dell'efficacia del modello su testi di dimensioni inferiori.

I generi scelti sono quelli più popolari e includono principalmente giochi di azione (19 titoli, 15.000 token in media), i giochi di ruolo (8 titoli, 27.000 token in media), i cosiddetti *picchiaduro* (3 titoli, 5.000 token in media), gli *sparatutto* (3, 9.000 token), *Hack and slash* (3, 5.000 token), *party e platform* (6, 3.000 token), gli *sportivi* (3, 3.000 token) e gli *horror* (6, 18.000 token).

Di seguito osserviamo in che modo si strutturano i *transcript*:

[Dark lane. The main character wounded a goat-like demon. It lies on the ground.]

Starring GARCIA HOTSPUR

Garcia: The bullet train is here, hellmonkey.

Demon: As if you mortals can be saved with one squeeze. Kill me, and I shall be replaced by another and another... and another still. You cannot point that pet gun of yours at all demonkind. And wherever you are not looking is where the greatest threat shall be. One at a time, we shall seize the treasures of your his as spoils, and leave only emptiness and despair my last gifts to you.

Garcia: Just don't forget to wrap them, Puta Claus.

Demon: By the way, Hotspur... How is your dear sweet Paula? Is she "hanging in there"?

Garcia: Fuck you!

[He blows the demon head out with his gun and run away. He runs home and finds his girlfriend hanged.]

L'esempio sopra è tratto dalla prima scena del videogioco *Shadow of the Damned*. Possiamo subito notare la presenza di alcune descrizioni dei principali elementi visivi della scena. Abbiamo poi un discorso diretto tra due personaggi. All'interno del discorso diretto, i termini di slang sono molti (*the bullet train, hellmonkey, Puta Claus, hanging in there, Fuck you*). Sia nelle descrizioni che nel discorso diretto, invece, è possibile trovare riferimenti alla violenza (*wounded, kill, gun, hanged*).

Trattandosi di dialoghi trascritti dagli utenti, nella maggior parte dei casi non sono presenti gli elementi interattivi, ma solo le linee di dialogo scelte dal giocatore durante la sua partita. Dunque, si tratta comunque di *transcript* parziali. Questa mancanza, tuttavia, è in parte mitigata dal fatto che in genere, gli sviluppatori tendono a rendere obbligatorie molte delle linee di dialogo riproponendole più volte durante la scena finché non vengono scelte. Visto questo tipo di approccio, i dialoghi effettivamente mancanti rappresentano una piccola parte della dimensione testuale del gioco. Il corpus è stato reso disponibile per il download alla pagina (Github, n.d.b). I diversi *transcript* sono stati inseriti in un unico file testuale (.txt) e separati dal seguente indicatore di fine documento: "--END--".

5. Estrazione dei termini

Come abbiamo detto, vista l'importanza della componente visiva, non è possibile stabilire automaticamente il rating di un videogioco basandosi esclusivamente sulla sua dimensione testuale. Tuttavia, attraverso l'estrazione di termini appartenenti al linguaggio dello slang e ad un lessico caratterizzato da specifiche emozioni, è possibile far emergere la presenza di indicatori legati alla presenza di linguaggio volgare, violenza ed uso di droghe.

L'esperimento che proponiamo in questo paragrafo vuole dimostrare come tali indicatori possano fungere da supporto ai sistemi di rating nell'attribuzione delle specifiche etichette di *warning*, andando ad approfondire più dettagliatamente la dimensione testuale di un videogioco spesso relegata ad attore secondario nelle fasi di analisi dei titoli.

La metodologia proposta è basilare ma si dimostra effettiva e meriterebbe un ulteriore approfondimento. Gli *step* che hanno portato al risultato finale sono così riassunti:

1. Estrazione dei termini del dizionario *SsVdd* e del dizionario *NRC*, relativamente ai tag *Anger*, *Fear* e *Disgust*, dai *transcript* dei videogiochi selezionati;
2. Formulazione di tre macro-etichette tramite l'accorpamento di diversi tag;
3. Calcolo delle frequenze relative per ciascuna macro-etichetta;
4. Calcolo dei margini di assegnazione dei tag;
5. Valutazione dei risultati.

Per quanto riguarda l'indicatore di violenza, benché il numero di termini associati a questa feature nel nostro dizionario siano un buon numero, si è deciso di integrare la risorsa con informazioni di tipo emozionale estratte da un dizionario delle emozioni spesso utilizzato in task di *emotion detection*. Il dizionario usato è l'*NRC Affect Intensity Lexicon* (Mohammad e Turney 2013), disponibile gratuitamente per il download sul web. Questo dizionario contiene oltre 10.000 entrate lessicali a cui sono stati attribuiti 8 punteggi riguardanti otto emozioni basilari. I punteggi di associazione tra le parole e le emozioni di base sono stati assegnati tramite un questionario sottoposto ad un elevato numero di soggetti umani tramite una campagna di *crowdsourcing*. Da questo ampio dizionario abbiamo deciso di estrarre solo i termini che presentavano una associazione con tre emozioni in particolare: rabbia (1.515 parole), paura (1.798 parole) e disgusto (1.108 parole).

Dal corpus sono stati estratti un totale di circa 37.000 termini (circa il 5,2% del corpus) di cui, tuttavia, un buon numero appartenente all'*Affect Intensity Lexicon*. Questi ultimi sono stati aggiunti al conteggio delle parole relative alla violenza. Se raggruppiamo le occorrenze dei termini per i quattro tag, Slang,

Violenza, Droghe, possiamo notare che oltre 32.000 parole sono etichettate come Violenza, circa 5.000 appartengono allo Slang, e 800 a Droghe. Se prendiamo in esame i type estratti, cioè le singole occorrenze di ogni parola, abbiamo circa 650 parole estratte per Slang e Violenza e circa 110 per Droghe.

Osservando i termini estratti, ci accorgiamo, però, della presenza di innumerevoli parole ambigue presenti in ciascuna categoria. Tra le parole afferenti al dizionario delle Droghe troviamo parole come *blow* (105 occorrenze), *idiot* (46), *green* (46) tra le più presenti. Troviamo, con un minor numero di occorrenze anche parole non ambigue come *medicine* (23), *acid* (16), *cigarette* (9), *pills* (6), *morphine* (2). I termini che abbiamo definito Slang presentano, a loro volta, una simile incidenza di parole ambigue, benché, al primo posto, troviamo la parola *fucking* con 452 occorrenze. Per quanto riguarda i termini che indicano Violenza, non sembra che l'aggiunta delle circa 4.500 parole del *lessico delle emozioni* abbia portato ad un aumento del numero di termini estratti. Il numero di type estratti per violenza (661), infatti, è molto simile al totale di type afferenti al tag Slang (640). Tuttavia, il numero dei singoli token estratti è estremamente maggiore per violenza. Questo effetto è stato in parte influenzato dal fatto che i termini del *lessico delle emozioni* sono, a volte, molto frequenti, ma è dovuto, principalmente, all'alta frequenza di termini violenti nei giochi selezionati. Le parole *dead*, *kill*, *fight* e *soldier* contano rispettivamente 658, 955, 593 e 382 occorrenze e sono tra le parole con maggior frequenza del corpus. Se uniamo le parole *dead*, *death* e *die* otteniamo un totale di 1.340 occorrenze.

Calcolando la *Term Frequency/Inverse Document Frequency* (TF/IDF) delle parole nei testi, abbiamo un'ulteriore conferma dell'incidenza della terminologia violenta e spesso militare che caratterizza i videogiochi selezionati. Il TF/IDF rappresenta un valore di peso di una parola che rapporta la sua frequenza in un determinato testo alla frequenza della stessa parola in tutti i testi. Più essa appare nelle diverse sezioni del corpus, meno efficace sarà nel determinare le caratteristiche distintive di una specifica sezione. Abbiamo estratto le 100 parole con il valore di TF/IDF più alto per ogni testo e le abbiamo contate per verificare in quanti testi questi termini avessero un peso determinante. Una parola come *blood*, appartenente alla categoria violenza, è tra le 100 parole più importanti in più della metà dei giochi. Lo stesso vale per *guards*. Poco dietro troviamo termini come *shooting* e *destroyed*.

Al termine dell'estrazione, i punteggi dei vari dizionari sono stati combinati in modo da creare un indice che riguardasse tre indicatori nello specifico:

- *Indice di linguaggio volgare*: calcolato combinando i valori di occorrenza delle parole di Slang, Discriminazione e Sesso;
- *Indice di violenza*: associando i termini violenti a quelli estratti dal *lessico delle emozioni* e standardizzando per la dimensione del testo;
- *Indice di droghe*: contenenti i termini del dizionario delle droghe.

Nonostante la presenza di numerosi termini ambigui, i risultati ottenuti sembrano confermare l'efficacia di questa prima estrazione nell'identificare i tre indicatori all'interno dei giochi selezionati. Nella tabella 3 possiamo osservare alcuni dei valori limite ottenuti sulla base della frequenza relativa dei termini dei tre indicatori.

Indicatore	Valore Minimo	Valore Massimo	Media
Linguaggio volgare	<10 (Ninja Gaiden, Medieval)	>300 (Cyberpunk 2077)	85.27
Violenza	0.05 (Disney Infinity 3.0)	0.15 (Mortal Kombat vs DC Universe)	0.09
Droghe	1 (Toy Story 3, Cars 3: Driven to win)	55 (Cyberpunk 2077)	12.8

Tabella 3: Punteggi minimi, massimi e media dei tre indici.

Tra i giochi privi di linguaggio volgare troviamo *Ninja Gaiden*. Si tratta di un gioco di genere *hack and slash* dei primi anni 2000, etichettato come PEGI 16+ ed ESRB M a causa della violenza dei contenuti. Il tono dei dialoghi, tuttavia, non sembra essere caratterizzato da volgarità o slang.

???: "I am Maith. You will never get pass me."

Ryu: "Then I must defeat you!"

Maith: "You are as bold as your father. But he is a much better swordsman."

Ryu: "You know my father?"

Maith: "Come and fight, young Hayabusa!"

(GAMEPLAY)

Ryu: "You killed my father."

Maith: «Killed? It is true that we fought. But your father is alive.»

Ryu: «Lair!»

Maith: «No, it is not a lie. If you proceed further, you will see him. But it will be the last thing you see.»

Il tono della conversazione è teso, ma non si incorre mai in espressioni volgari. Ne è un esempio l'uso dell'espressione *liar*, preferita a numerose possibili alternative slang (*bullshitter*, *fibber*, *storyteller*).

Il secondo caso di videogioco privo di espressioni volgari è *Medieval*, un gioco del 1998 contraddistinto da un linguaggio di stile medievale e dunque caratterizzato proprio dall'assenza di uno slang moderno.

First Mate: Captain, I thought you ought to know we have a stowaway on board, I've told the men to scour the decks for him.

Captain: Good, I want that scurvy dog dangling from yon yardarm by his bowels!

First Mate: Is that really necessary? Couldn't we just give him a good tongue lashing

and drop him off at the next port?

Possiamo notare un elemento quantomeno gergale come *scurvy dog*, che appartiene però ad un linguaggio “piratesco”, soprattutto se accompagnato da sequenze come *dangling from yon yardarm by his bowels*. Si tratta dunque di un linguaggio fortemente stereotipato su elementi tipici della narrativa di avventura del ‘900.

Il gioco che presenta un numero maggiore di termini volgari è, invece, *Cyberpunk 2077*. Anche in questo caso, la valutazione sembra abbastanza corretta, trattandosi di un videogioco per maggiorenni ambientato tra la malavita di una città immaginaria del prossimo futuro.

Stout: Now listen close. This piece of shit, Antony Gilchrist is he your contact? Is he the one who leaked intel on the convoy?

V: That guy? Never seen him before.

Militech Guard: Checks out.

V: Listen, I know where the transport is. I can help you. Just want a favor in return.

Stout: Hmm...

Gilchrist: I told you! I fucking told you! I'm not the mole! Jesus Christ!

Stout: Shut him up.

Gilchrist: Unhand me now before I-- ungh!

L'intero *transcript*, così come l'esempio sopra, risulta caratterizzato da una fortissima presenza di slang e parolacce come *piece of shit* o *fucking*.

Per quanto riguarda l'indice di violenza, troviamo un punteggio molto basso per un gioco di corse di macchine come *Cars 3*. Benché ridotti al minimo, i termini che indicano combattimenti o scontri sono presenti anche in giochi come *Super Mario Party* o *Toy Story 3*. In ambito automobilistico, invece, la presenza di parole che indicano violenza è molto bassa.

Sono, al contrario, molto frequenti i termini violenti in uno dei videogiochi di lotta famosi proprio per la propria crudezza e per le scene sanguinose, *Mortal Kombat vs DC Universe*. Benché in questa versione del 2008 il franchise abbia introdotto alcuni eroi della *DC Comics*, l'universo fumettistico di Batman e Superman, bisogna ricordare come il primo *Mortal Kombat* fece così tanto scalpore al proprio rilascio che il caso fu portato al Senato degli Stati Uniti.

DARK KAHN: You seek to destroy me. Good...

[Everyone begins to hold their heads in pain as the rage surges through them.]

DARK KAHN: Give in to your rage. Unleash your aggression. FIGHT!!!

[The two forces, with the exception of Raiden and Superman, charge at each other. Just before everyone collides, the scene cuts to darkness. It is replaced by pictures of the ensuing brawl. Liu Kang throws a flying kick at Green Lantern. Jax and Lex Luthor pound away at each other. Wonder Woman blocks a high kick from Kitana. Scorpion throws his spear as Flash dives in with his fist. Baraka and Deathstroke draw their blades...]

Anche in questo breve dialogo, accompagnato da una descrizione delle immagini, è facile notare la costante presenza di termini legati al combattimento (*destroy, fight, kick, spear, blades*), al dolore (*pain*), e alla rabbia (*rage*).

In fine, abbiamo l'indice di Droghe, dove il videogioco *Cyberpunk 2077* ottiene il punteggio più alto. Non si tratta di un punteggio altissimo, considerato che la frequenza dei nomi di droghe è di 55 su circa 73 mila token; tuttavia, è il punteggio più alto di un videogioco, ed è coerente con quello che sappiamo della trama. Il videogioco è ambientato in un futuro distopico di tipo *cyberpunk* dove le persone si fanno impiantare arti e organi cibernetici per potenziare le proprie capacità, rischiando di perdere il contatto con la realtà. I nomi di droga individuati riguardano più farmaci e altre sostanze che droghe vere e proprie (*analgesic, painkiller*). I videogiochi per bambini e quelli *fantasy* o ambientati nel medioevo, invece, sono caratterizzati dalla completa assenza di queste parole. Oltre a *Toy Story 3* e *Cars 3*, tra giochi in cui le droghe sono assenti troviamo anche *Dark Souls II: Scholar of the First Sin* e *Assassin's Creed: Bloodlines*.

Una volta estratte le occorrenze dei termini afferenti al dizionario *SsVdd*, abbiamo valutato automaticamente la presenza dei tre indicatori nei videogiochi, verificando poi l'etichetta attribuita ad ogni gioco dai due principali sistemi di valutazione, PEGI e ESRB. In particolare, abbiamo stabilito tre livelli di rischio per ogni videogame:

- Rischio Zero: videogiochi privi di *warnings* di tipo *violence, bad language* o *Drugs*;
- Rischio Uno: giochi a cui è stata attribuita una delle tre etichette prese in esame;
- Rischio Due: videogiochi in cui l'etichetta è preceduta dai termini *insane/intense/strong*.

La scelta di attribuire un'etichetta ad un videogioco è stata effettuata stabilendo un margine di assegnazione. Questo margine, benché differente per i tre indicatori, è uniforme per tutti i videogiochi selezionati. La tabella 4 presenta i valori di assegnazione ottimizzati per ogni categoria. I punteggi sono stati stabiliti tramite una sperimentazione che ha preso il via dal calcolo dei quantili su percentuali del 33 e del 67%. Una fase di ottimizzazione del modello, effettuata testando un gran numero di valori, ha permesso di modificare i quantili per ottenere risultati migliori. In particolare, è stato necessario aumentare di 10 punti il quantile del livello 1 per *Bad Language*, e duplicare i valori per *Drugs*. I quantili di *Violence*, al contrario sono stati modificati in maniera contrastante, aumentando il livello uno e diminuendo il due.

Indice	Livello Due	Livello Uno
Bad Language	>80	>60
Violence	>0.105	>0.05
Drugs	>30	>22

Tabella 4: Valori soglia per l'attribuzione delle etichette.

Come abbiamo detto, per l'indicatore di violenza, vista la presenza di un enorme numero di termini generici estratti grazie al lessico emozionale, sono stati impiegati i valori di frequenza relativa. L'*Affect Intensity Lexicon*, infatti, contiene un numero molto alto di parole generiche etichettata con emozioni specifiche come rabbia, paura o tristezza. Con l'etichetta *Anger*, ad esempio, possiamo trovare parole come *brutality* (con un punteggio di 0.964), ma anche *flagrant*, *impermeable*, *clamor* o *intense* (0.455). Visto il numero di occorrenze estratte tramite questi termini, si è preferito utilizzare il punteggio di frequenza per diminuire il loro impatto in giochi caratterizzati da *transcript* più grandi.

La tabella 5 presenta i valori di precisione ottenuta nell'identificazione delle etichette per i videogiochi selezionati.

	Bad Language	Violence	Drugs	Media
Valutazione del livello di rischio	0.62	0.54	0.66	0.61
Valutazione della presenza dell'indicatore	0.78	0.86	0.86	0.83

Tabella 5: Valutazione dei risultati.

Prendendo in considerazione il livello di rischio (zero, uno e due) delle etichette, il sistema è in grado di riconoscere la presenza di un *warning* con una precisione media del 61%. Benché questo risultato non sia trascurabile, l'obiettivo, come abbiamo detto, non è attribuire un'etichetta, ma suggerire la presenza di eventuali indicatori e favorire un controllo manuale del software prima del suo rilascio. Prendendo in considerazione la capacità del sistema di predire la sola presenza/assenza di un *warning*, il punteggio di precisione è di gran lunga superiore (83%).

Nella tabella 7 possiamo osservare le differenze nella valutazione dei videogiochi in cui gli indicatori sono presenti (Positivi) e quelli in cui non lo sono (Negativi).

	SsVdd		
	Precisione Positivi	Precisione Negativi	F-score
Bad Language	0.83	0.73	0.78
Violence	0.98	0.14	0.92
Drugs	0.64	0.94	0.72

Tabella 6: Precisione del modello per Positivi e Negativi.

Il dato più significativo è legato alla precisione sui negativi dell'indicatore Violenza. Il punteggio estremamente basso ottenuto può essere dovuto all'e-

strazione dei termini generici presenti nell'*Affect Intensity Lexicon*, che induce il modello a considerare come positivi la maggior parte dei testi analizzati. Seppur in maniera meno marcata, la dimensione ridotta del tag *Drugs* produce un risultato inverso nella rispettiva etichetta. Il punteggio dell'etichetta *Bad Language*, invece, risulta più equilibrato, ma genera un F-score inferiore a quello di *Violence*.

Alla luce dell'avvento dei *Large Language Models* (LLM) è necessario proporre un confronto con le capacità di questi modelli. I LLM, o IA generative, sono dei costrutti software, basati sulla tecnologia delle reti neurali, addestrati su enormi quantità di dati linguistici ed in grado di comprendere e produrre il linguaggio in maniera quasi del tutto indistinguibile da un essere umano. Questi modelli possono essere considerati un'evoluzione dei modelli di Semantica Distribuzionale di tipo predittivo (Lenci 2023). A differenza di questi, però, i LLM sono addestrati su una quantità di dati estremamente maggiore (GPT-3, ad esempio, era addestrato su un corpus di 500 miliardi di parole), e con una complessità della rete neurale molto superiore.

Per il confronto, abbiamo utilizzato ChatGPT, basato sul modello *GPT-4* (aprile 2024) sottoponendo al *prompt* la seguente query: «Given the 51 texts in the file, each separated by the --END-- sequence, estimate the presence of drugs in each sequence and present the results in a table with the following columns: A: number of the sequence; B: Description of content; C presence of drugs (high, moderate, low)».

ChatGPT ha prodotto le tabelle richieste affermando, nel caso della violenza e del linguaggio volgare, che «this table summarizes the thematic content of [violence, ranging from discussions and threats to physical confrontations and strategic actions involving violence, across the sequences in the document | explicit language based on the dialogue and themes present in the text]». Per quanto riguarda l'uso di droghe, pur modificando la *query*, abbiamo ricevuto in risposta una semplice estrazione di elementi lessicali o una assegnazione random dell'etichetta. I risultati ottenuti da ChatGPT con gli altri due indicatori, tuttavia, sono di gran lunga inferiori alle aspettative, riuscendo a raggiungere un valore massimo di precisione per la violenza di 0.31, considerando due distinti livelli, e di 0.68 considerando solo la presenza, e per il *Bad Language* rispettivamente di 0.27 e 0.45. L'impressione, tuttavia è che, sottoponendo all'analisi di ChatGPT ogni *transcript* singolarmente, i risultati possano essere migliori.

Al contrario, chiedendo al software di indicare un possibile rating di età per ogni videogioco abbiamo ottenuto una risposta estremamente positiva, con una correlazione tra i rating automatici e quelli di PEGI dello 0.89 (correlazione di Pearson). In questo caso, ChatGPT ha prodotto la risposta basandosi su valutazioni interne che fanno affidamento alle sue specifiche conoscenze linguistiche di natura semantica, applicate al testo. Dato che nel corpus non

sono presenti i titoli dei singoli videogiochi, probabilmente ChatGPT non ha fatto uso di conoscenze pregresse su tali titoli, e possiamo considerare valido questo secondo esperimento.

6. Conclusioni

In questo articolo abbiamo presentato una risorsa elettronica formata da un dizionario elettronico del linguaggio slang ed un corpora composto da *transcript* di videogiochi in inglese, ed un sistema per classificare automaticamente i *transcript* dei videogiochi. Pur trattandosi di prodotti multimediali interattivi, i videogiochi possono essere considerati dei testi per la loro natura comunicativa e narrativa. L'uso di questi prodotti culturali, tuttavia, è regolato da specifiche etichette che allertano sui rischi di esposizione dei minori a contenuti non idonei a determinate fasce d'età. Le agenzie di *rating*, tuttavia, non sempre concordano nel classificare i videogiochi, perché le loro analisi prendono in considerazione elementi parziali del prodotto.

Il sistema di classificazione proposto è in grado di allertare le agenzie su eventuali rischi legati alle etichette relative a linguaggio volgare, violenza e uso di droghe, sulla base del dizionario elettronico dello slang. Il modello proposto è stato testato sul corpus dei *transcript* che include 50 videogiochi pubblicati tra gli anni 90 ed il 2020. I risultati sono incoraggianti, in quanto con questa semplice metodologia è possibile identificare correttamente circa l'80% dei videogiochi in cui uno dei tre indicatori è presente.

Benché l'attribuzione di etichette di *warning* rappresenti una possibile applicazione pratica per le risorse elettroniche presentate, l'analisi della dimensione testuale dei videogiochi può proporsi come *step* fondamentale per lo studio di fenomeni sociali o psicologici. Vista la natura sociale dello slang, l'estrazione di questi termini potrebbe fungere da base per uno studio più dettagliato della natura sociologica e sociolinguistica dei videogiochi più di tendenza, portando alla luce, ad esempio, il desiderio di evasione sociale dei giocatori. Allo stesso modo, da un'analisi approfondita dei dati, anche su scala diacronica, si potrebbe portare alla luce l'evoluzione di fenomeni come la discriminazione razziale o sessuale, analizzando l'influenza che questo media ha sulle giovani generazioni di giocatori. La risorsa che abbiamo presentato, pertanto, sarà a disposizione di ricercatori di diverse discipline, interessati a svariate tematiche legate al linguaggio utilizzato nei videogiochi.

In un possibile prosieguo di questo progetto ci proponiamo di lavorare ad un sistema per la disambiguazione dei termini slang. Ciò potrebbe essere possibile sia lavorando sulla sintassi della frase, sia analizzando i termini tramite un modello di Semantica Distribuzionale di tipo contestuale, cioè in grado di carpire il significato di un termine polisemico dal contesto in cui appare.

Riferimenti bibliografici

- Aarseth, Espen J. 2004. "Quest Games as Post-Narrative Discourse Espen Aarseth." In *Narrative across media: The languages of storytelling*, edited by Marie-Laure Ryan, 361-76. Lincoln: University of Nebraska Press.
- Amari, Jorgen. 2010. "Slang lexicography and the problem of defining Slang." In *The fifth international conference on historical lexicography and lexicology*, 1-11. University of Oxford.
- Andersson, Lars-Gunnar, and Peter Trudgill. 1992. *Bad Language*. London: Penguin Books Downes.
- Balteiro, Isabel. 2019. "Lexical and morphological devices in gamer language in Fora." In *Approaches to videogame discourse: Lexis, interaction, textuality*, edited by Astrid Ensslin and Isabel Balteiro, 39-57. Bloomsbury Academic. <https://doi.org/10.5040/9781501338489.0008>.
- Benedict, Jennifer, and Pamela Munro. 1997. *U.C.L.A. Slang 3: [a dictionary by 25 U.C.L.A. students]*, 3-28. Westwood: Department of Linguistics, University of California.
- De Klerk, Vivian. 1990. "Slang: A male domain?" *Sex Roles* 22: 589-606.
- Dogruel, Leyla, and Sven Joeckel. 2013. "Video Game Rating Systems in the US and Europe: Comparing Their Outcomes." *International Communication Gazette* 75(7): 672-92. <https://doi.org/10.1177/1748048513482539>.
- Dumas, Bethany K., and Jonathan Lighter. 1978. "Is Slang a word for linguists?" *American Speech* 53(1): 5-17.
- Eble, Connie C. 1996. *Slang & sociability: In-group language among college students*. The University of North Carolina Press.
- Elia, Annibale. 2002. "Discorso scientifico e linguaggio settoriale. Un esempio di analisi lessico-grammaticale di un testo neuro-biologico." *Simboli, linguaggi e contesti* 2: 71-85. Carocci. <https://www.iris.unisa.it/handle/11386/1061667>.
- Ensslin, Astrid, and John Finnegan. 2019. "Bad language and bro-up cooperation in co-sit gaming." *Approaches to videogame Discourse: Lexis, interaction, textuality*, 139-56. Bloomsbury Publishing.
- Ensslin, Astrid. 2012. *The language of gaming*. Basingstoke: Palgrave Macmillan.
- Fekete, Tamás, and Ádám Porkoláb. 2019. "From Arkngthand to Wretched Squalor: Fictional place-names in universe." *ICAME Journal* 43(1): 23-58.
- Felini, Damiano. 2015. "Beyond Today's Video Game Rating Systems: A Critical Approach to PEGI and ESRB, and Proposed Improvements." *Games and Culture* 10(1): 106-22. <https://doi.org/10.1177/1555412014560192>.

- Gee, James Paul. 2014. *Unified discourse analysis: Language, reality, virtual worlds and video games*. Routledge.
- Github. n.d.a. "Slang and Videogames." Consultato il 15 maggio 2014. <https://github.com/amatusNLP/Slang-and-Videogames/blob/main/Ss-Vdd.txt>.
- Github. n.d.b. "Slang-and-Videogames." Consultato il 15 maggio 2014. <https://github.com/amatusNLP/Slang-and-Videogames>.
- Goorimoorthee, Tejasvi, Adrianna Csipo, Shelby Carleton, and Astrid Ensslin. 2019. "Language Ideologies in Videogame Discourse: Forms of Sociophonetic Othering in Accented Character Speech." In *Approaches to Videogame Discourse: Lexis, Interaction, Textuality*, edited by Astrid Ensslin and Isabel Balteiro, 269-87. London: Bloomsbury Academic. <http://dx.doi.org/10.5040/9781501338489.0020>.
- Green, Jonathon, 2011. "Some thoughts on Slang." *Revista Alicantina de Estudios Ingleses* 24: 153-71. <https://doi.org/10.14198/raei.2011.24.06>
- Heritage, Frazer. 2021. "Language, Gender, and Videogiochi." In *Language, Gender and Videogiochi*, edited by Frazer Heritage, 27-61. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-74398-7_2.
- Humphreys, Ashlee, and Rebecca Jen-Hui Wang. 2018. "Automated text analysis for consumer research." *Journal of Consumer Research* 44(6): 1274-1306.
- Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý, and Vít Suchomel. 2013. "The TenTen corpus family." In *7th international corpus linguistics conference CL*, 125-27.
- Juul, Jesper. 2011. *Half-real: Video games between real rules and fictional worlds*. MIT press.
- Klimova, Blanka, and Kacet Jaroslav. 2017. "Efficacy of computer games on language learning." *Turkish Online Journal of Educational Technology-TOJET* 16(4): 19-26.
- Lenci, Alessandro, 2023. *Understanding natural language understanding systems. A critical analysis*. <https://arxiv.org/pdf/2303.04229>.
- Lighter, Jonathan E. 1994. *Historical Dictionary of American Slang*. Random House. Oxford University Press.
- Maisto, Alessandro, Giandomenico Martorelli, Antonietta Paone, and Serena Pelosi. 2021. "Extracting video games rating labels from transcript files." *Internet of Things* 16: 100439.

- Marchiori, Eugenio J., Ángel Del Blanco, Javier Torrente, Iván Martínez-Ortiz, and Baltasar Fernández-Manjón. 2011. "A visual language for the creation of narrative educational games." *Journal of Visual Languages & Computing* 22(6): 443-52.
- Mattiello, Elisa. 2005. "The pervasiveness of Slang in standard and non-standard English." *Mots Palabras Words* 6: 7-41.
- McMillan, James B. 1978. "American Lexicology 1942-1973." *American Speech* 53(2): 141-63.
- Mencken Henry Louis. 1967. "American Slang, Its origin and history." In *The American Language*, edited by Raven I. McDavid and David W. Maurer, 4th edition with supplements, 702-61. New York, Alfred A Knopf.
- Mohammad, Saif M., and Peter D. Turney. 2013. "Crowdsourcing A Word-Emotion Association Lexicon." *Computational Intelligence* 29(3): 436-65. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>.
- Partridge, Eric. 1933. *Slang Today and Yesterday*. London: Routledge and Kegan Paul.
- Ray, Alice. 2019. "Playing with the language of the future: The localization of science-fiction terms in videogames." In *Approaches to videogame discourse: Lexis, interaction, textuality*, edited by Astrid Ensslin and Isabel Balteiro, 87-115. Bloomsbury Academic.
- Spears, Richard A. 2000. *NTC's dictionary of American Slang and colloquial expressions*. NTC Publishing Group. <http://thuvienso.bvu.edu.vn/handle/TVDHBRVT/14818>.
- Wright, Andrew, David Betteridge, and Michael Buckby. 2006. *Games for language learning*. Cambridge University Press.

Discovering and documenting brilliance

A novel multimodal annotation method

Alexander Murzaku*, Pontish Yeramyan, Curt Anderson, Steven Buxbaum, Ruben Diaz, Marielle Lerner, Armenui Minasyan, Hazel Mitchley, Jodie-Ann Pennant, Mia Shang, Brisa Speier Brito

Abstract: The pursuit of understanding human brilliance has long fascinated scholars, practitioners, and observers across various domains of performance. Capturing and understanding human brilliance is an important endeavor, as it allows us to gain insight into the motivations behind extraordinary achievements. However, brilliance remains a complex phenomenon that eludes simple characterization. In this paper we propose a methodology for investigating brilliance. We use structured interviews with high-achieving individuals to document the language surrounding their accomplishments, and code this language using seven distinct themes. Each interview is then annotated using a multimodal annotation schema that captures a variety of linguistic and paralinguistic features, including phonetic information, hand gesture, and eye gaze. This system allows us to discover and document the tacit knowledge that underlies human brilliance and make inroads in understanding the full communicative expression of brilliance.

Keywords: Multimodal Annotation, Tacit Knowledge, Thematic Analysis, Corpus Analysis, Human Brilliance.

1. Introduction

The pursuit of understanding human brilliance has long fascinated scholars, practitioners, and observers across various domains of performance, be it in the arts, sciences, sports, or leadership. Efforts to define brilliance, as well as the exploration into what drives exceptional performance, follow various approaches. Simonton uses his historiometric approach to reverse engineer genius, by investigating «the most eminent figures of history, whether creators or leaders, using various at-a-distance quantitative methods, including computer content analysis of primary documents such as creative products

* Saint Elizabeth University, Morristown, NJ, USA and Leveraging Genius Institute at Gap International, Springfield, PA, USA. amurzaku@steu.edu and alex.murzaku@gapinternational.com. All co-authors are members of the Leveraging Genius Institute at Gap International and can be contacted at first.last@gapinternational.com.

or private correspondence» (Simonton 2016, 4). However, brilliance remains a complex phenomenon that eludes simple characterization. This complexity necessitates innovative methodologies to uncover the underlying elements that compose it.

The methodology described in this paper explores what contributes to brilliance through interviews and offers the possibility of a nuanced analysis facilitated by multimodal annotations. When participants articulate their journeys, philosophies, and reflections, they provide clues to the mental frameworks and attitudes that underpin their performance. However, the richness of these verbal accounts often contains subtleties and nuanced expressions of thought and emotion that traditional analysis methods may overlook. Multimodal annotation is a method that leverages multiple data types — textual, auditory, visual, and kinesthetic cues — to enrich the interpretation of conversation. This composite approach allows researchers to not just listen to what is said but to also tune into how it is expressed. Multimodal analysis offers a glimpse into not only what people do during their expression of brilliance and how they do it, but also what drives them, what triggers their moments of genius. This paper proposes a method to capture and document what creates brilliance.

2. Brilliance as tacit knowledge

Human history is filled with many documented examples of brilliant minds and talents that have influenced the path of human progress. From Socrates to Shakespeare, from Da Vinci to Einstein, from Beethoven to John Lennon, each has left an indelible mark on their respective fields, demonstrating how individual brilliance can drive forward entire domains of human endeavor.

The role of brilliance in human evolution can be seen as a catalyst for growth, for adaptation, and for transformation. It has enabled humanity to transition from basic survival to advancing human knowledge and technology, pushing the boundaries of what it means to be human and how we understand and interact with the world around us.

Examples of brilliance are manifested in the corporate world as well through innovative business strategies, groundbreaking products that we all use, effective leadership, and transformative organizational cultures. Brilliant minds like Steve Jobs at Apple, Jeff Bezos at Amazon, Larry Page and Sergey Brin at Google, or Elon Musk at Tesla show that brilliance does not rest solely on individual genius but also on the ability to foster innovation, adapt to change, and execute visionary strategies that drive businesses — and sometimes entire industries — forward.

While brilliance often involves a unique combination of talent, creativity, and insight that seems inherently personal, organizations can implement practices to capture and disseminate these brilliant sparks. Capturing and

documenting the insights, methodologies, and creative processes of brilliant individuals in knowledge repositories allows organizations to preserve this valuable information. However, it remains in the domain of information; what is missing is capturing brilliance in action, brilliance in performance.

Brilliance in performance, or genius, refers to the outstanding display of skill, talent, and ability routinely transcending standard expectations. It's the combination of innate aptitude, rigorously honed skills, deep understanding, and often an innovative or creative approach that separates merely good performances from truly exceptional ones.

Documenting and preserving brilliance in performance, whether it pertains to artistic achievements, stellar scientific contributions, or exceptional leadership, involves capturing not only the tangible outcomes of such efforts but also the processes, insights, and nuances that characterize them. This is a multifaceted endeavor that demands careful attention to both the tangible and intangible aspects of achievements. What is collected in this endeavor are two forms of knowledge labeled as explicit knowledge and tacit knowledge (Polanyi 1958). Brilliance lives in tacit knowledge.

Explicit knowledge is verbalized, documented and easily transferred. In contrast, tacit knowing is a way of knowing without telling, highlighting the nature of certain types of knowledge - knowledge that one can act upon but finds difficult to fully articulate or transfer solely through written or spoken word (Polanyi 1966, 18). «[T]acit knowledge is knowledge hidden from the consciousness of the knower» point out Olomolaiye and Egbu (2005, 770), and Casonato and Harris (1999, quoted in Dampney, Busch, and Richards 2002, 3) observe that «[t]acit knowledge includes skills, experiences, insight, intuition and judgment. It is typically shared through discussion, stories, analogies and person-to-person interaction; therefore, it is difficult to capture or represent in explicit form». There is tacit knowledge in the understandings, skills and philosophies developed by societies with long histories of interaction with their natural surroundings. For rural and indigenous peoples, local knowledge informs decision-making about fundamental aspects of day-to-day life. This knowledge is integral to a cultural complex that also encompasses language, systems of classification, resource use practices, social interactions, ritual and spirituality (UNESCO 2014). Discussing tacit teaching, Miton and De-Deo (2022) point out that explicit knowledge can become tacit once the code required to express it is lost, and tacit knowledge can become explicit in cases where it can be codified. They found that tacit knowledge is acquired through socialization and sharing direct experience and interactions with others.

In a series of studies, Guarasci et al. (2008, 2010) explore the tacit knowledge of master craftsmen whose products are arguably the output of decades, sometimes centuries, of accumulated experiences, driven by what we consider brilliance or genius. They propose the creation of formalized workflows and

methods for the development of knowledge repositories necessary for guaranteeing continuity and succession related to specialized practices. In the field of goldsmith handcraftsmanship, for example, they identified a strong need to represent the regularly used knowledge and know-how of master craftsmen that allows for the production of exclusive and highly valuable products, without the need to describe and structure the whole production cycle.

According to Polanyi (1966, 17) true knowledge lies in our ability to use it. How can we use the brilliance of a master? Chess players, stresses Polanyi, enter into a master's spirit by rehearsing the games the master played to discover what the master had in mind, in an effort to acquire that tacit knowledge (Polanyi 1966, 30). We propose that a more direct path is needed to reliably access brilliance.

The multimodal analysis proposed in this paper allows for a deep exploration of the "why": why the chess master spends untold hours to study and improve, why the master goldsmith in Calabria keeps visualizing and creating works of beauty, what contributes to their brilliance, to their genius. The "why" is the most opaque part of tacit knowledge and its transmission is made possible by a thorough analysis of not only knowledge as linguistic content but also the manner in which it is expressed by the expert.

Language is the standard medium of knowledge documentation and transfer, but representing an individual's brilliance cannot rely on traditional means to study and preserve this tacit knowledge. Furthermore, tacit knowledge, when transferred, is often shared through face-to-face interactions. To capture the multidimensionality of genius and its expression requires a holistic approach to language analysis beyond a strictly unimodal view.

3. Language is multimodal

Traditional linguistic analysis is often unimodal, focusing on information expressed in a single modality and overlooking other linguistic and paralinguistic information that might be simultaneously present in other modalities. For instance, a mainstream semantic analysis might center around linguistic units such as morphemes, words, phrases, or constructions while idealizing away from non-syntactic information that might nevertheless contribute to the meaning of an utterance. In face-to-face interaction, though, speakers and hearers simultaneously use and integrate visuospatial and speech modalities during communication. If the analysis of language privileges a single channel, this ignores the wealth of additional information available when language is used face-to-face, including modulation of pitch and intonation, hand gestures, eye gaze, facial expression, and body posture. These are not just add-ons to the speech signal but are highly integrated into producing and understanding utterances. Based on this observation, there has been a growing movement to

analyze language as a multimodal phenomenon that is expressed simultaneously in visual and auditory channels (Holler and Levinson 2019; Perniss 2018).

Supporting this view that language is a multimodal phenomenon is abundant evidence that hearers use information from multiple modalities when understanding language. The classic example of this is the McGurk effect (McGurk and MacDonald 1976; MacDonald and McGurk 1978), where a voice saying [ba] is paired with a face articulating [ga], resulting in the perception of [da] in order to reconcile contradictory auditory and visual information. Visual information is used in other ways as well; for instance, speech is processed, interpreted, and even produced faster when facial articulatory movements can be seen (van Wassenhove, Grant, and Poeppel 2005; Drijvers and Holler 2023).

In addition to the visual modality corresponding to the face and mouth, there is evidence that co-speech gestures are also exploited by listeners and speakers. Adult speakers often use gesture when speaking (McNeill 1992; 2005; Kendon 2004; Goldin-Meadow 2003; Clark 1996), while children produce gestures even before speech, and the presence of gesture predicts certain linguistic milestones in language acquisition (see Goldin-Meadow 2014 for an overview). Gestures that mismatch a prior spoken word or are semantically incongruent with the spoken word show an N400 effect, an event-related potential (ERP) component in electroencephalogram (EEG) studies that is argued to index semantic anomaly (Kutas and Federmeier 2011). Listeners are sensitive to the temporal overlap between co-speech gestures and speech (McNeill 1992; Church, Kelly, and Holcomb 2013; Loehr 2007), and the integration of gesture with speech is partially determined by the emotional valence of the utterance (Levy and Kelly 2020; Kelly and Tran 2023).

Information from the visual modality can be linked to the utterance syntactically, semantically, and pragmatically. For instance, speakers are better at verbally communicating a two-dimensional shape to another speaker when allowed to use hand gestures (Graham and Argyle 1975), with the effect more pronounced when the shape was designed to be more difficult to communicate. This shows gesture may carry semantic information that can be understood by listeners. Iconic gestures contribute to the meaning of an utterance as well. A person might say “I didn’t see it coming” while pantomiming the image of a car blind-siding them. Kelly et al. (1999) show that participants, when asked to summarize what was said, include in their summaries information that was contributed solely through iconic gestures. Co-speech gestures also contribute to how utterances cohere in a discourse (Lascarides and Stone 2009).

Some types of phonetic information can be integrated into an utterance in ways that are troublesome for mainstream approaches. For instance, formal semantics often identifies morphemes as the unit of semantic composition, but lengthening a vowel can sometimes signal semantic information that’s not

strictly expressed by the morpheme itself (e.g., *it was a loong movie*), complicating traditional approaches to compositionality. Phonetic information is also used to convey pragmatic information such as irony and sarcasm, which are not necessarily found in the morphemic content of the utterance. Bryant and Fox Tree (2002) extracted sarcastic and non-sarcastic utterances from radio talk shows and presented them to experiment participants; participants were unable to distinguish which utterances were sarcastic or non-sarcastic based on the text of the utterance but could do so when the sound recordings were played. Filled pauses, like *uh* and *um*, control the flow of the discourse (Clark and Fox Tree 2002; Cossavella and Cevasco 2021), again operating distinctly from the morphemes in the utterance. And while the precise mapping between emotions and prosody is not perfectly understood (Larrouy-Maestri, Poeppel, and Pell 2024), experiments have shown that emotions and attitudes can be recognized based on the quality of voice alone (Pell 2006; Dara, Monetta and Pell 2008; Monetta, Cheang and Pell 2008). These examples show that we might gain better insight into language by setting aside assumptions about what can contribute to semantics and pragmatics, namely assumptions that understanding is based purely on words or morphemes alone.

This is a small sampling of the literature that shows that paralinguistic features are closely linked to speech, but it makes a case that language is expressed through multiple modalities simultaneously. Importantly, as Massaro (1987) points out, not only are these paralinguistic features used to encode part of the message, but they are also used in the interpretation of the message. There has therefore been a growing movement for corpora of spoken communication to include more than just transcripts of the spoken word, as transcripts cannot adequately capture the contribution of paralinguistic cues in verbal communication.

Examples of multimodal corpora that go beyond words are, among others, the SmartKom Corpus (Schiel, Steininger and Türk 2002), the Augmented Multi-party Interaction Corpus (Carletta et al. 2005), and the HuComTech Corpus (Abuczki and Ghazaleh 2013). These corpora include annotations at a variety of levels, including speech transcription, dialogue acts, group activity, topics and subtopics, abstractive and extractive summaries, where participants focused their attention, emotions expressed by facial expressions, emotion and physical cues, prosody, speech, intonation, head and hand gestures, hand shapes, movements, posture, and touches. Annotations such as these, in conjunction with annotations for pitch and body movement, have also been used to contribute to the development of multimodal sentiment analysis, which is used to automatically detect the positive, negative or neutral expression of emotion in a variety of modalities, as opposed to only texts (Gandhi et al. 2023).

As part of the movement for multimodal analyses of communication, we take the approach that if we want to understand the motivations behind brilliant people or what generates brilliance more generally, focusing solely on their words is not sufficient, as we may miss aspects of their language that signal precisely those motivations we are looking for. In this article we propose a method we believe can enable the discovery of what underpins brilliance. In the following section, we describe our approach to creating an annotated multimodal corpus that can be used to study the mindsets of exceptional individuals.

4. Methodology for discovering brilliance

4.1. In-depth structured interviews

In-depth structured interviews were conducted with 76 high-achieving individuals via video conference, obtaining verbal permission to record both video and audio. Participants were asked to prepare a list of accomplishments and were prompted to choose one accomplishment on which to focus during the interview. Participants were led through a question-answer protocol over a period of roughly 60-90 minutes.

Following the structured interviews, interviewers provided participants with a visual schematic of key excerpts of their own speech that represent the mindset components that lead to brilliance (see Yeramyán and Jackson 2014 for a description of the structured interview method and training). These excerpts are henceforth referred to as “genius excerpts.” As described in white papers by Yeramyán and Jackson (2014; 2015), approximately ten thousand genius excerpts extracted from interview transcripts and notes were classified by a team of linguists based on a lexical semantic analysis of the text. However, the identified classes did not account for everything the participants expressed during their interviews. The themes displayed in Table 1 are derived from subsequent efforts to complement the text-based analysis (explicit knowledge) with a holistic analysis of the interview experience (tacit knowledge), which resulted in a more refined thematic grouping. These seven major themes characterize the mindsets that underlie the brilliant accomplishments of high-achieving individuals.

POSSIBILITY / CREATION / CHOICE / FREEDOM
PURPOSE / CONTRIBUTION / WORLD IMPACT / LEGACY
EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND
CURIOSITY / DISCOVERY / LEARNING / GROWING
AT-STAKENESS ¹ / CONVICTION / RESOLVE / PERSEVERANCE
JOY / EXCITEMENT / EXHILARATION / SPIRIT
TOGETHERNESS / RELATIONSHIP / CONNECTION / LOVE

Table 1. Seven themes of brilliance (Gap International 2021).

As Braun, Goudbeek, and Kraemer (2020) point out, there are disparities between self- and other-annotation. They conclude that discrepancies between authors' intentions and readers' perceptions might indicate a mismatch between the authors' intended affect and the affect experience of others. To minimize this discrepancy, participants were prompted to classify their own genius excerpts according to these seven themes; they were able to select up to three of the seven themes for each genius excerpt.

4.2. Interview Annotation

4.2.1. *Full interview annotation*

The first phase of analysis focused on the theme of EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND. Four of the structured interviews were selected for an in-depth examination. One annotator watched the interviews in their entirety and selected sections where they noted the theme. A team of annotators then watched these sections and collectively identified the linguistic and paralinguistic features indicating the presence of the theme. The team aligned on which features were most salient and categorized them into linguistic categories: lexical, phonetic, rhetorical, syntactic; and paralinguistic categories: eye gaze, facial expressions, body movement, and hand gestures. The EUDICO Linguistic Annotator (ELAN) video and audio annotation software (The Language Archive 2023) was selected for annotation, and a tier was created for each category. Throughout this annotation phase, there were between four and six annotators (as some annotators joined the project after the initial phase).

Once the feature tags were finalized, the annotators individually watched each interview in full and tagged the features in sections they identified as relevant to the theme of EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND. All tags were applied at the word or phrase level. In addition to an-

¹ Here "AT-STAKENESS" is a neologism used to mean the condition of having something at stake.

notating features of this theme, annotators also used a dedicated ELAN tier to flag sections where they perceived any of the other six themes of brilliance to be present in the interviewee's speech, with the intent to use this information in the following phase of research. All practical instructions on implementing feature tags were documented for reference in a set of annotation guidelines.

For two lexical features of the theme, absolute and intensifier, automatic tagging was used to aid annotators as they navigated through the entire interview. The team collectively composed a word list with relevant lexemes. For example, the list of absolutes included lexemes like *any*, *all*, *every*, *each*, *always*, *never* and the list of intensifiers included lexemes such as *very*, *really*, *highly*, *so*, *super*. The algorithm tagged any instance of the lexemes, which produced considerable noise. For this reason, the automatically identified tags were not incorporated directly into the data set, but instead used as a tool to direct the annotator's attention to relevant tokens. The automatic tagging flagged all instances of the lexemes on a dedicated tier in ELAN, and the annotator tagged only relevant tokens on the lexical tier. For instance, an annotator would ignore an instance of *so* as a conjunction (e.g. *so I left*) but tag an instance of *so* as an intensifier (e.g. *so determined*).

In the first phase of annotation, the paralinguistic features were tagged both manually and automatically. A video analysis software tool based on open-source libraries was created to automatically detect and tag visible facial expressions, emotions, eye gaze, and body movements (e.g. *furrowed brow*, *surprise*, *anger*, *happiness*, *widen eyes*, *lean in*, *lean out*, etc.). As visual analysis software uses frame-by-frame analysis (24 frames per second), it was decided to record only the dominant tag during the pronunciation of a word segment (e.g., if the enunciation of a word took 0.5 seconds, among the 12 tags generated for the 12 frames, only the most frequent tag would be reported for that time interval). Still, there was a significant overproduction of tags which caused more noise than signal for the purpose of this research. A persistent issue with automatic annotation was a high error rate (e.g., the software would interpret the V-shape of certain glasses as a furrowed brow and therefore tag it as *angry*). For this reason, although automatic annotations of visual cues were produced for all videos, they were not included in the final analysis.

4.2.2. *Full interview annotation review*

After the initial annotation, each structured interview was reviewed by another annotator. The reviewer watched the interview in full, checking all annotations as well as unannotated sections. The reviewer made notes in a separate review tier about any proposed changes and flagged newly added annotations.

The reviewer and annotator then met to discuss disagreements or new annotations made by the reviewer and align on the final annotations. Any cases for which the reviewer and annotator could not reach an agreement were brought to the whole annotation team for a final decision. In addition to the changes agreed upon by the reviewer and annotator, there was a list of sanctioned changes and additions which the reviewer could make without consulting the annotator. Sanctioned changes and additions included those made to reflect updates to annotation guidelines or new annotations on the phonetic tier.

4.2.3. *Partial interview annotation*

A second approach was taken for analysis of the remaining six themes, using portions of each of the 76 structured interviews rather than analyzing the entire interviews. A research cycle was designed to study one theme of brilliance at a time, as follows.

The team began with preliminary analysis to determine the linguistic features of the theme at hand using the genius excerpts, as classified by the participants. For this analysis, annotators used a sample of ten genius excerpts from each participant's interview to identify the linguistic features for the given theme. The team aligned on which features were most salient and categorized them into linguistic categories: lexical, phonetic, rhetorical, semantic and syntactic. The paralinguistic categories from the EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND theme were carried over to this analysis: eye gaze, facial expressions, body movement, and hand gestures. For each theme of brilliance, all feature tags were defined and documented in annotation guidelines for reference during annotation.

Once the feature tag set was defined, a preliminary annotation phase was carried out with a subset of the structured interviews to evaluate the relevance of the tags and to note considerations in implementation. Each annotator completed annotation of one randomly chosen interview, annotating 3-4 sections using each of three annotation passes (as described below). In an alignment meeting, annotators proposed changes to the feature tags, which were agreed upon by the whole team and documented in the annotation guidelines. After the feature tag sets were finalized, the team annotated all participant interviews using ELAN, meeting daily to resolve issues or disagreements on their use. Between six and ten annotators completed annotations for a given theme of brilliance.

Portions of each interview were selected by conducting three annotation passes for each interview. The first pass used each participant's genius excerpts for each theme, according to their thematic classification. The genius excerpts were located in the interviews using a matching algorithm; the team annota-

ted features in these sections and in the utterances immediately preceding and following. The second pass consisted of iterating through sections that had been previously flagged as corresponding to one of the six themes of brilliance, identified during the annotation of EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND described above.

The third pass focused on lexemes that were tagged automatically using a matching algorithm. The annotation team created lexical sets composed of lexemes that are semantically similar to the themes of brilliance. For example, the POSSIBILITY / CREATION / FREEDOM / CHOICE theme's lexical set included words like *opportunity, innovation, potential, invent*. These lexical sets were selected by the annotators based on review of the genius excerpts and word embedding techniques based on language generators. The first iteration of this matching algorithm used lemmas to detect multiple forms of a lexeme (e.g., *believe, believed, believing*). It was noted that the algorithm was missing target lexemes that were subsequently found by the annotator. A stemming approach eventually replaced the lemmatization approach to capture more lexemes, which yielded a more effective set of automatically tagged lexemes. Annotators then checked tagged lexemes and the surrounding utterances for the presence of the brilliance theme at hand.

Some features were identified across themes, and therefore their corresponding tags were carried over from annotation of one theme to the next. Two such features were the use of absolutes and intensifiers. The automatic tagging of absolutes and intensifiers, as described above for analysis of the EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND theme, was also available as an ELAN tier to annotators to reference in all three annotation passes.

The three passes were effective in directing annotators' attention to sections of the interview relevant to the theme being studied. Taken together, the passes ensured that the selected portions were indeed relevant. Each annotation pass did at times direct annotators to sections that were already viewed in another pass; however, each pass also identified additional sections to annotate.

4.2.4. *Partial interview annotation review*

Each annotated interview was reviewed by another member of the annotation team. The reviewer checked all annotations and labeled their agreement for each annotation on a separate agreement tier using binary documentation: 1 for agreement and 0 for disagreement. The reviewer could also input additional annotations to the annotated sections.

The reviewer and annotator then met to discuss disagreements or new annotations made by the reviewer and align on the final annotations. Any cases for which the reviewer and annotator could not reach an agreement were brought to the whole annotation team for a final decision. In addition to the

changes agreed upon by the reviewer and annotator, there was a list of sanctioned changes and additions which the reviewer could make without consulting the annotator. Sanctioned changes included: those made to reflect updates to annotation guidelines or new annotations of tags that were well-defined and less subject to annotator judgement (e.g. phonetic features). In rare cases, newly defined feature tags that were straightforward in their use (e.g. questions) were also included as a sanctioned change. Any changes made, both new annotations and revisions to existing annotations, were counted and marked on a separate change tier.

Figure 1 lays out the major steps of the annotation process as conducted for the partial interview approach.

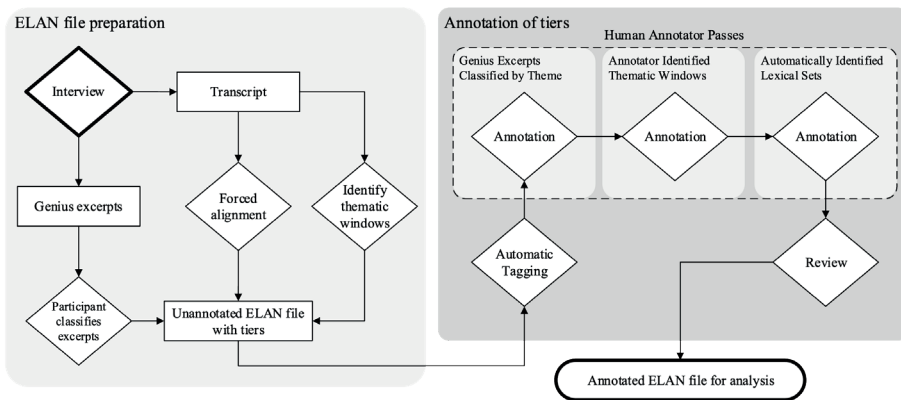


Figure 1: Visualization of the file preparation and annotation process, starting with the structured interview and ending with an annotated ELAN file.

4.2.5. *Partial interview annotation inter-rater reliability*

To measure the consistency and reliability of the subjective annotation system, the team used an inter-rater reliability metric that tests whether the annotator and reviewer's judgments were the same at a rate significantly higher than one would expect due to chance. It is commonly used in place of percent agreement as a way of evaluating the agreement rate while considering the occurrence of chance agreements.

The metric, Cohen's κ (kappa), uses the relative observed agreement among raters, and the hypothetical probability of chance agreement to evaluate the agreement rate in relation to the possibility of chance agreements (Cohen 1960). Table 2 shows the κ scores and the agreement rate between the annotators and the reviewers for each theme.

Brilliance theme	Cohen's κ	Agreement
POSSIBILITY / CREATION / CHOICE / FREEDOM	0.478	95.8%
PURPOSE / CONTRIBUTION / WORLD IMPACT / LEGACY	0.456	91.9%
CURIOSITY / DISCOVERY / LEARNING / GROWING	0.488	97.8%
AT-STAKENESS / CONVICTION / RESOLVE / PERSEVERANCE	0.439	89.2%
JOY / EXCITEMENT / EXHILARATION / SPIRIT	0.463	93.2%
TOGETHERNESS / RELATIONSHIP / CONNECTION / LOVE	0.450	91%

Table 2: Kappa statistic and agreement rate by brilliance theme.

According to the guidelines set forth by Landis and Koch (1977), all of the κ statistics are within the Moderate Agreement range (0.4 to 0.6). This indicates that while there were disagreements about the subjective annotations, the reviewer and the annotator had identical judgments at a rate significantly higher than one would expect from chance. This suggests a relatively high degree of alignment among annotators regarding feature tags and their use.

4.2.6. *Interview annotation summary*

In summary, two distinct annotation processes were used; one in which the entire interview was viewed, and one in which only selected portions of the interview were viewed. In the first approach, all sections of the interview identified as relevant to the theme were annotated according to the annotator's discretion. In the second approach, the annotators used a more systematic approach to identify sections relevant to the theme, incorporating participant input to do so. The second approach allowed the annotation team to annotate the remaining themes in one sixth of the time that was required for the first approach, while the first provided annotators with a comprehensive view of the interview data. Future investigations may adopt the full interview annotation method for other themes in order to draw a direct comparison between annotation approaches and more fully evaluate the merits of each.

4.2.5 *Acoustic dimension of annotations*

Speech offers a natural and complex way to convey information about our emotional and mental state. In conversational or spontaneous speech, such information is often expressed unintentionally, yet small acoustic changes can create large perceptual effects heard by the listener. Many studies have been conducted to determine which acoustic correlates are significant in the study of emotional speech (Weninger et al. 2013; Kamiloğlu, Fischer, and Sauter 2020; Kacur et al. 2021). Similarly, acoustic correlates have been identified for

diagnosing various health conditions (Roark et al. 2011; Meilán et al. 2014; Rektorova et al. 2016; Sara et al. 2020). While there have been attempts to examine the acoustic correlates of complex expressions like charisma (Hirschberg and Rosenberg 2005), leadership (Weninger et al. 2012), and cooperation (Tognetti et al. 2020), little has been done to study the acoustic correlates of brilliance. In the present multimodal study, we aim to identify acoustic correlates of each brilliance theme by extracting acoustic features used in the previously mentioned studies. We hypothesize that acoustic patterns map onto clusters which can classify a particular brilliance theme.

Transcripts for each interview were obtained using the default Automatic Speech Recognition output in Video Text Tracks (VTT) (World Wide Web Consortium 2019) format, and later were improved using WhisperX (Bain et al., 2023) and underwent automatic text-to-speech alignment via Montreal Forced Aligner (MFA) (McAuliffe et al. 2017) using an American English pronunciation dictionary. Automatic alignment offers sufficiently accurate parsing of speech and pause regions for use in speech-based measures (Yuan et al. 2023). Output alignments resulted in 52 TextGrids with word-level and phone-level alignments for 76 speakers². TextGrids with word-level alignments were imported into ELAN to reference during manual annotation. Speech disfluencies in conversational speech (i.e. hesitations, interruptions, filled pauses) and low recording quality created some misalignment between text and audio. However, these were manually corrected during annotation. Similarly, mistranscribed words were manually changed to the correct word in accordance with the audio; these corrected transcripts were imported from the annotation of one theme to the next. A separate ‘Pause Tier’ was manually generated in ELAN from empty intervals on the word-tier to ensure perceived pauses were annotated with time (ms) accuracy. Pause regions were defined as intervals where no acoustic signal was detected by MFA.

Our proposed method of annotation is based on existing research on emotions and voice, in which both acoustic features and listeners’ perception is measured (Kamiloglu, Fischer, and Sauter 2020; Dai, Fell, and MacAuslan 2009; Erickson et al. 2016; Kacur et al. 2021). Prior to extracting acoustic features, the data was manually annotated in ELAN for perceptual effects of acoustics in a “Phonetic Tier”. These perceptual tags³ were simply defined as: *pitch change*, *emphatic stress*, *rhythm*, and *pause*. Perceptual annotation of the data offered an indication if the brilliance theme in question was acoustically perceived by the annotator or not. Thus, tag labels were not meant to have a

² The difference in number of TextGrid files and number of speakers is due to multiple speakers per interview. Each TextGrid file corresponds to one interview, some containing multiple speakers.

³ Perceptual tags for acoustics were the same for all themes except for EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND, as this was the first theme annotated. Therefore tags for EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND were simplified to the current tags.

one-to-one relationship with the acoustic feature extracted for that region, but rather an indication that the given brilliance theme was phonetically marked compared to the surrounding speech.

Our approach to acoustic feature extraction includes features reported in the literature as relevant in emotional speech detection and speech used to detect health conditions. Acoustic features across four domains will be extracted in Praat (Boersma and Weenink 2024), namely (i) pitch, (ii) voice quality, (iii) intensity, and (iv) rate (temporal measures) (Tab. 3). Multiple acoustic measures across these domains have proven best for classification results (Kacur et al. 2021; Kamiloğlu, Fischer, and Sauter 2020). The table in Appendix III shows all acoustic features to be extracted (*fundamental frequency, formant frequencies, harmonics to noise ratio, intensity and amplitude, shimmer, speech rate, articulation rate, and pauses*) and their corresponding tags, perceptual correlates, and definitions. Praat scripts for extracting acoustic features were adapted for the data to iterate on a by-speaker basis and parameters were adjusted to account for variables shown affect acoustics (i.e. gender). These scripts were run on small batches of the Genius Excerpts to see preliminary acoustic trends and adjust script parameters. These scripts iterate over voiced regions of speech unless the acoustic feature specifically calls for unvoiced regions.

5. From qualitative to quantitative annotation analysis: a preliminary sample

After the annotation process, certain linguistic and paralinguistic features were observed to appear more prominently in some themes over others. The annotation process itself offered ongoing qualitative insights into the motivations and emotional triggers that generate brilliance, such as what features should we look for and how should they be tagged. That very process created the basis for a hypothesis. Based on those hypotheses, we proceeded with a quantitative analysis of the generated annotations. Below is a preliminary analysis of two quantitative features, one from the lexical domain and one from the acoustic domain.

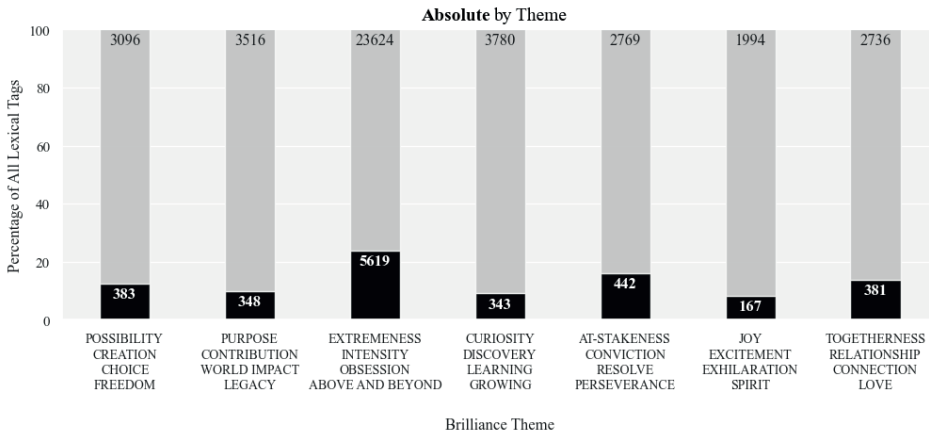


Figure 2: Absolute lexical tag proportion by theme.

Brilliance theme	Proportion of <i>absolute</i> tags to Lexical tags	Total <i>absolute</i> tags
POSSIBILITY / CREATION / CHOICE / FREEDOM	0.124	383
PURPOSE / CONTRIBUTION / WORLD IMPACT / LEGACY	0.099	348
EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND	0.238	5619
CURIOSITY / DISCOVERY / LEARNING / GROWING	0.091	343
AT-STAKENESS / CONVICTION / RESOLVE / PERSEVERANCE	0.160	442
JOY / EXCITEMENT / EXHILARATION / SPIRIT	0.084	167
TOGETHERNESS / RELATIONSHIP / CONNECTION / LOVE	0.129	381
Weighted mean proportion for all themes	0.206	7683

Table 3: *absolute* lexical tags as a proportion of all lexical tags, organized by brilliance theme.

Table 3 shows the number of *absolute* as a proportion of the total lexical tags for each theme. The *absolute* tag was used to indicate the presence of lexemes that indicated an all or nothing meaning (e.g. *always*, *every time*, *never*, *total*). The hypothesis that absolutes were a more salient feature of the EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND theme than the other themes was tested using a one-sample proportion test.

The test measured whether there was a statistically significant difference between the proportion of *absolute* to all lexical tags in the EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND theme (0.238) than for the

overall weighted mean proportion of *absolute* to all lexical tags (0.206). The one-sample proportion test revealed that the higher-than-average proportion of *absolute* for the EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND theme was significant $\chi^2 = 141.1$, $p < 0.0001$. This indicates that in sections where the EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND theme was noted by annotators, the interviewees were more likely to use absolutes in their speech, suggesting that the use of absolutes is an indicative feature of the expression of EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND.

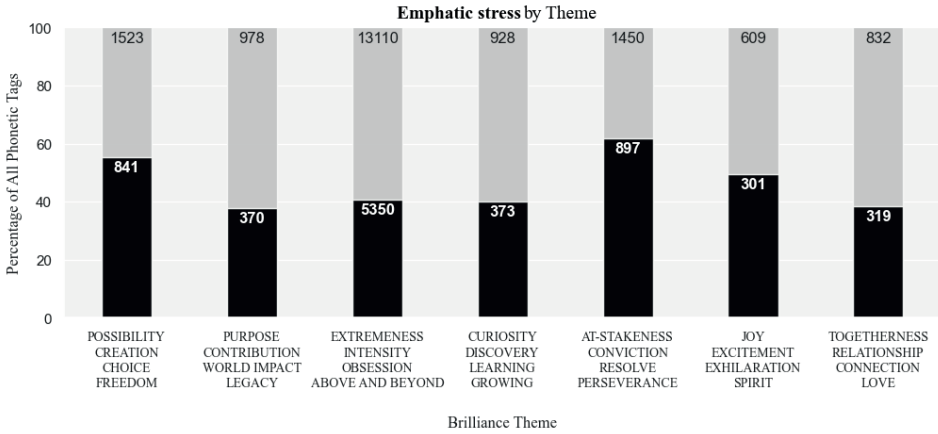


Figure 3: *emphatic* stress tag proportion by theme.

Brilliance theme	Proportion of <i>emphatic stress</i> tags to Phonetic tags	Total <i>emphatic stress</i> tags
POSSIBILITY / CREATION / CHOICE / FREEDOM	0.552	841
PURPOSE / CONTRIBUTION / WORLD IMPACT / LEGACY	0.378	370
EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND	0.408	5350
CURIOSITY / DISCOVERY / LEARNING / GROWING	0.402	373
AT-STAKENESS / CONVICTION / RESOLVE / PERSEVERANCE	0.619	897
JOY / EXCITEMENT / EXHILARATION / SPIRIT	0.494	301
TOGETHERNESS / RELATIONSHIP / CONNECTION / LOVE	0.383	319
Weighted mean proportion for all themes	0.445	8451

Table 4: *emphatic stress* tags as a proportion of all Phonetic tags, organized by brilliance theme.

The *emphatic stress* tag was used when the annotator perceived word-level stress used to express the theme, beyond the normal lexical stress. Table 4

shows the *emphatic stress* tags as a proportion of the total Phonetic tags for each theme.

The *emphatic stress* tag was prevalent throughout the seven themes. Analysis showed it to be most prominent in sections noted for the AT-STAKENESS / CONVICTION / RESOLVE / PERSEVERANCE theme. A one-sample proportion test was performed to determine if there was a statistically significant difference between the proportion of *emphatic stress* in the AT-STAKENESS / CONVICTION / RESOLVE / PERSEVERANCE theme (0.619) and the overall weighted mean proportion of *emphatic stress* tags (0.445). The one-sample proportion test indicated that there is a statistically significant difference, $\chi^2 = 176.27$, $p < 0.0001$. The results of this test show that the difference from the average proportion of *emphatic stress* across all themes is statistically significant. This suggests that *emphatic stress* was more likely to appear in sections where the AT-STAKENESS / CONVICTION / RESOLVE / PERSEVERANCE theme was noted, as compared to the other themes.

6. Conclusions and future work

The examination of what causes human brilliance is an area of inquiry which needs methodologies that are not only useful for explicit forms of knowledge, but can also adequately capture tacit knowledge. The implementation of a multimodal annotation approach for the analysis of structured interviews with individuals sharing their brilliance has revealed the complex interaction of cognitive factors that underpin exceptional accomplishments. The diversity of these factors requires the expansion of traditional analytical frameworks to embrace the different data types generated in conversation, such as textual transcripts and acoustic and visual data. In defining the methodology proposed here, every effort was made to minimize annotators' perception biases and to maximize the participants' intention through linguistic and behavioral analysis of the interviewee.

The data generated by this annotation process enables a systematic uncovering of patterns related to mindset and attitudes, while maximizing the codification of tacit knowledge within narrative accounts of brilliance. By employing quantitative measures of the cooccurrence of themes of brilliance with factors such as acoustic features, contextually relevant gestures, and semantic clusters, a more structured examination of the attributes associated with brilliance is made possible.

Future explorations into the domain of brilliance may benefit significantly from the incorporation of artificial intelligence (AI) and the automation of analytical processes. AI's capabilities in managing extensive datasets and identifying complex patterns could greatly augment the depth and efficiency of analysis. Automating elements of the annotation process, including emotional

tone assessment, thematic tagging through natural language processing, and pattern identification, could streamline research methodologies, allowing for a focus on more interpretive aspects of analysis.

The increased attention to the role of multimodal interaction with Large Language Models enables the development of annotation tools specifically designed for multimodal data, aiming to improve both the precision and user experience for researchers. We are also considering the integration of machine learning algorithms to refine pattern detection capabilities and to enhance the predictability of traits associated with brilliance based on annotated interviews. Automation combined with careful human fine-tuning will enable longitudinal research studies to capture the developmental trajectory of brilliance, providing insight into its dynamism and adaptability. Future research directions also include expanding and balancing the participant pool across various disciplines, cultural backgrounds, and personal histories to ensure a more inclusive representation of brilliance.

In summary, the scholarly pursuit of understanding human brilliance necessitates an ongoing commitment to methodological innovation and interdisciplinary exploration. The multimodal annotation method proposed within this study offers a foundational approach for dissecting the intricate web of brilliance. With the prospective integration of AI and automation, research into human potential is poised for advancements that promise richer insights and more nuanced understandings of how brilliance manifests and could be cultivated across human activity.

References

- Abuczki, Ágnes, and Esfandiari Baiat Ghazaleh. 2013. "An overview of multimodal corpora, annotation tools and schemes." *Argumentum* 9(1): 86-98. <https://core.ac.uk/download/pdf/161005866.pdf>.
- Bain, Max, Jaesung Huh, Tengda Han, and Andrew Zisserman. 2023. "WhisperX: Time-accurate speech transcription of long-form audio." In *Proceedings of Interspeech 2023*, 4489-93. <https://doi.org/10.21437/Interspeech.2023-78>.
- Boersma, Paul, and Weenink, David. 2024. *Praat: doing phonetics by computer* [Computer program]. V6.1.09. <http://www.praat.org/>.
- Braun, Nadine, Martijn Goudbeek, and Emiel Krahmer. 2020. "Emotional Words-The Relationship of Self-and Other-Annotation of Affect in Written Text." *Cognitive Science*: 430-36.
- Bryant, Gregory A., and Jean E. Fox Tree. 2002. "Recognizing verbal irony in spontaneous speech." *Metaphor and Symbol* 17(2): 99-117. https://doi.org/10.1207/S15327868MS1702_2.

- Carletta, Jean, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, et al. 2005. "The AMI meeting corpus: A pre-announcement." In *International workshop on machine learning for multimodal interaction*: 28-39. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/11677482_3.
- Casonato, R., and K. Harris. 1999. "Can an Enterprise Really Capture 'Tacit Knowledge': We answer two top questions on knowledge management from the Electronic Workplace 1999 Conference." *Gartner Group Research Note Select Q&A 16*.
- Church, R. Breckinridge, Spencer Kelly, and David Holcombe. 2013. "Temporal synchrony between speech, action and gesture during language production." *Language, Cognition and Neuroscience* 29(3): 345-54. <https://doi.org/10.1080/01690965.2013.857783>.
- Clark, Herbert H. 1996. *Using language*. Cambridge University Press. <https://doi.org/10.2277/0521561582>.
- Clark, Herbert H., and Jean E. Fox Tree. 2002. "Using *uh* and *um* in spontaneous speaking." *Cognition* 84(1): 73-111. [https://doi.org/10.1016/S0010-0277\(02\)00017-3](https://doi.org/10.1016/S0010-0277(02)00017-3).
- Cohen, Jacob. 1960. "A coefficient of agreement for nominal scales." *Educational and psychological measurement* 20(1): 37-46. <http://dx.doi.org/10.1177/001316446002000104>.
- Cossavella, Francisco, and Jazmín Cevalco. 2021. "The importance of studying the role of filled pauses in the construction of a coherent representation of spontaneous spoken discourse." *Journal of Cognitive Psychology* 33(2): 172-86. <https://doi.org/10.1080/20445911.2021.1893325>.
- Dai, Keshi, Harriet Fell, and Joel MacAuslan. 2009. "Comparing Emotions Using Acoustics and Human Perceptual Dimensions." In *CHI '09 Extended Abstracts on Human Factors in Computing Systems CHI EA '09*: 3341-46. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/1520340.1520483>.
- Dampney, Kit, Peter Busch, and Debbie Richards. 2002. "The meaning of tacit knowledge." *Australasian Journal of Information Systems* 10(1): 3-13. <https://doi.org/10.3127/ajis.v10i1.438>.
- Dara, Chinar, Laura Monetta, and Marc D. Pell. 2008. "Vocal emotion processing in Parkinson's disease: Reduced sensitivity to negative emotions." *Brain Research* 1188: 100-11. <https://doi.org/10.1016/j.brainres.2007.10.034>.
- Drijvers, Linda, and Judith Holler. 2023. "The multimodal facilitation effect in human communication." *Psychonomic Bulletin & Review* 30: 792-801. <https://doi.org/10.3758/s13423-022-02178-x>.

- Erickson, Donna, Chunyue Zhu, Shigeto Kawahara, and Atsuo Suemitsu. 2016. "Articulation, Acoustics and Perception of Mandarin Chinese Emotional Speech." *Open Linguistics* 2(1): 620-35. <https://doi.org/10.1515/opli-2016-0034>.
- Gandhi, Ankita, Kinjal Adhvaryu, Soujanya Poria, Erik Cambria, and Amir Hussain. 2023. "Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions." *Information Fusion* 91: 424-44. <https://doi.org/10.1016/j.inffus.2022.09.025>.
- Gap International. 2021. "The Leveraging Genius Conference 2021." Online, February 22-24, 2021, <https://m.eventsinamerica.com/events/the-leveraging-genius-conference-2021/retail-wholesale/ecommerce/5obagzt-duc3696mz>.
- Goldin-Meadow, Susan. 2003. *Hearing gesture: How our hands help us think*. Harvard University Press.
- Goldin-Meadow, Susan. 2014. "Widening the lens: What the manual modality reveals about language, learning and cognition." *Philosophical Transaction of the Royal Society B* 369: 20130295. <https://doi.org/10.1098/rstb.2013.0295>.
- Graham, Jean Ann, and Michael Argyle. 1975. "A cross-cultural study of the communication of extra-verbal meaning by gestures." *International Journal of Psychology* 10(1): 57-67. <https://doi.org/10.1080/00207597508247319>.
- Guarasci, Roberto, Antonietta Folino, Alessia Cosentino, Elena Cardillo, and Maria Taverniti. 2008. "Gestion et formalisation de la connaissance tacite." In *Sistemes d'Information & Intelligence economique* 1: 564-76. IHE éditions.
- Guarasci, Roberto, Elena Cardillo, Antonietta Folino, and Maria Taverniti. 2010. "Multilingual taxonomic and terminological structures of a domain." In *Globalization and the Management of Information Resources - Papers from the International Conference, Sofia, Bulgaria, 12-14 November 2008*, 508-18. St. Kliment Ohridski University Press.
- Hirschberg, Julia Bell, and Andrew Rosenberg. 2005. "Acoustic/Prosodic and Lexical Correlates of Charismatic Speech." In *Proceedings of Eurospeech 2005*. <https://doi.org/10.7916/D8Q52Z3Z>.
- Holler, Judith, and Stephen C. Levinson. 2019. "Multimodal language processing in human communication." *Trends in Cognitive Sciences* 23(8): 639-52. <https://doi.org/10.1016/j.tics.2019.05.006>.

- Kacur, Juraj, Boris Puterka, Jarmila Pavlovicova, and Milos Oravec. 2021. "On the Speech Properties and Feature Extraction Methods in Speech Emotion Recognition." *Sensors* 21(5): 1888. <https://doi.org/10.3390/s21051888>.
- Kamiloğlu, Roza G., Agneta H. Fischer, and Disa A. Sauter. 2020. "Good Vibrations: A Review of Vocal Expressions of Positive Emotions." *Psychonomic Bulletin & Review* 27(2): 237–65. <https://doi.org/10.3758/s13423-019-01701-x>.
- Kelly, Spencer D., and Quang-Anh Ngo Tran. 2023. "Exploring the emotional functions of co-speech hand gesture in language and communication." *Topics in Cognitive Science* 00: 1-23. <https://doi.org/10.1111/tops.12657>.
- Kelly, Spencer D., Dale J. Barr, Ruth Breckinridge Church, and Katheryn Lynch. 1999. "Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory." *Journal of Memory and Language* 40: 577-92. <https://doi.org/10.1006/jmla.1999.2634>.
- Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge University Press.
- Kutas, Marta, and Kara D. Federmeier. 2011. "Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP)." *Annual Review of Psychology* 62: 621-47. <https://doi.org/10.1146/annurev.psych.093008.131123>.
- Landis, J. Richard, and Gary G. Koch. 1977. "The measurement of observer agreement for categorical data." *Biometrics*: 159-74. <http://dx.doi.org/10.2307/2529310>.
- Larrouy-Maestri, Pauline, David Poeppel, and Marc D. Pell. 2024. "The sound of emotional prosody: Nearly 3 decades of research and future directions." *Perspectives on Psychological Science*. <https://doi.org/10.1177/17456916231217722>.
- Lascarides, Alex, and Matthew Stone. 2009. "Discourse coherence and gesture interpretation." *Gesture* 9(2): 147-80. <https://doi.org/10.1075/gest.9.2.01las>.
- Levy, Rachel S., and Spencer D. Kelly. 2020. "Emotion matters: The effect of hand gesture on emotionally valenced sentences." *Gesture* 19(1): 41-71. <https://doi.org/10.1075/gest.19029.lev>.
- Loehr, Daniel. 2007. "Aspects of rhythm in gesture and speech." *Gesture* 7(2): 179-214. <https://doi.org/10.1075/gest.7.2.04loe>.
- MacDonald, John, and Harry McGurk. 1978. "Visual influences on speech perception processes." *Perception and Psychophysics* 24(3): 253-57. <https://doi.org/10.3758/BF03206096>.

- Massaro, Dominic W. 1987. *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, New Jersey: Lawrence Erlbaum.
- McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, and Morgan Sonderegger. 2017. "Montreal Forced Aligner: trainable text-speech alignment using Kaldi." In *Proceedings of the 18th Conference of the International Speech Communication Association*: 498-502. <https://doi.org/10.21437/Interspeech.2017-1386>.
- McGurk, Harry, and John MacDonald. 1976. "Hearing lips and seeing voices." *Nature* 264: 746-48. <https://doi.org/10.1038/264746a0>.
- McNeill, David. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- McNeill, David. 2005. *Gesture and thought*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226514642.001.0001>.
- Meilán, Juan José G., Francisco Martínez-Sánchez, Juan Carro, Dolores E. López, Lymarie Millian-Morell, and José M. Arana. 2014. "Speech in Alzheimer's Disease: Can Temporal and Acoustic Parameters Discriminate Dementia?" *Dementia and Geriatric Cognitive Disorders; Basel* 37(5-6): 327-34. <https://doi.org/10.1159/000356726>.
- Monetta, Laura, Henry S. Cheang, and Marc D. Pell. 2008. "Understanding speaker attitudes from prosody by adults with Parkinson's disease." *Journal of Neuropsychology* 2: 415-30. <https://doi.org/10.1348/174866407X216675>.
- Miton, Helena, and Simon DeDeo. 2022. "The cultural transmission of tacit knowledge." *Journal of the Royal Society Interface* 19(195): 20220238. <https://doi.org/10.1098/rsif.2022.0238>.
- Olomolaiye, Anthony, and Charles Egbu. 2005. "Tacit vs. explicit knowledge - the current approaches to knowledge management." In *Second Scottish Conference for Postgraduate Researchers of the Built and Natural Environment (Probe)*: 503-11. <https://www.irbnet.de/daten/iconda/CIB10682.pdf>.
- Pell, Marc D. 2008. "Cerebral mechanisms for understanding emotional prosody in speech." *Brain and Language* 96(2): 221-34. <https://doi.org/10.1016/j.bandl.2005.04.007>.
- Perniss, Pamela. 2018. "Why we should study multimodal language." *Frontiers in Psychology* 9. <https://doi.org/10.3389/fpsyg.2018.01109>.
- Polanyi, Michael. 1958. *Personal Knowledge: Towards a Post - Critical Philosophy*. London: Routledge and Kegan Paul.
- Polanyi, Michael. 1966. *The Tacit Dimension*. Garden City, NY: Doubleday & Company Inc.

- Rektorova, Irena, Jiri Mekyska, Eva Janousova, Milena Kostalova, Ilona Eliassova, Martina Mrackova, Dagmar Berankova, Tereza Necasova, Zdenek Smekal, and Radek Marecek. 2016. "Speech Prosody Impairment Predicts Cognitive Decline in Parkinson's Disease." *Parkinsonism & Related Disorders* no. 29 (August): 90-95. <https://doi.org/10.1016/j.parkrel-dis.2016.05.018>.
- Roark, Brian, Margaret Mitchell, John-Paul Hosom, Kristy Hollingshead, and Jeffrey Kaye. 2011. "Spoken Language Derived Measures for Detecting Mild Cognitive Impairment." In *IEEE Transactions on Audio, Speech, and Language Processing* 19(7): 2081–90. <https://doi.org/10.1109/TASL.2011.2112351>.
- Sara, Jaskanwal Deep Singh, Elad Maor, Barry Borlaug, Bradley R. Lewis, Diana Orbelo, Lliach O. Lerman, and Amir Lerman. 2020. "Non-Invasive Vocal Biomarker Is Associated with Pulmonary Hypertension." *PLOS ONE* 15(4): e0231441. <https://doi.org/10.1371/journal.pone.0231441>.
- Schiel, Florian, Silke Steininger, and Ulrich Türk. 2002. "The SmartKom Multimodal Corpus at BAS." In *Proceedings of the International Language Resources and Evaluation Conference (LREC): 200-06*. <http://www.lrec-conf.org/proceedings/lrec2002/pdf/49.pdf>.
- Simonton, Dean Keith. 2016. "Reverse engineering genius: historiometric studies of superlative talent." *Annals of the New York Academy of Sciences* 1377 (1): 3-9. <https://doi.org/10.1111/nyas.13054>.
- The Language Archive. 2023. *ELAN* [Computer program] V.6.7. Max Planck Institute for Psycholinguistics. <https://archive.mpi.nl/tla/elan>.
- Tognetti, Arnaud, Valerie Durand, Melissa Barkat-Defradas, and Astrid Hopfensitz. 2020. "Does He Sound Cooperative? Acoustic Correlates of Cooperativeness." *British Journal of Psychology* 111(4): 823–39. <https://doi.org/10.1111/bjop.12437>.
- UNESCO. 2014. "Inventory-making: a cumulative in-depth study of periodic reports." UNESCO Intangible Cultural Heritage. Accessed April 8, 2024. <https://ich.unesco.org/en/focus-on-inventory-making-2014-00876>.
- van Wassenhove, Virginie, Ken W. Grant, and David Poeppel. 2005. "Visual speech speeds up the neural processing of auditory speech." In *Proceedings of the National Academy of Sciences* 102 (4): 1181-86. <https://doi.org/10.1073/pnas.0408949102>.
- Weninger, Felix, Florian Eyben, Björn Schuller, Marcello Mortillaro, and Klaus Scherer. 2013. "On the Acoustics of Emotion in Audio: What Speech, Music, and Sound Have in Common." *Frontiers in Psychology* 4: 292. <https://doi.org/10.3389/fpsyg.2013.00292>.

Weninger, Felix, Jarek Krajewski, Anton Batliner, and Björn Schuller. 2012. “The Voice of Leadership: Models and Performances of Automatic Analysis in Online Speeches.” *IEEE Transactions on Affective Computing* no. 3 (October): 496-508. <https://doi.org/10.1109/T-AFFC.2012.15>.

World Wide Web Consortium (W3C). 2019. “WebVTT: The Web Video Text Tracks Format.” <https://www.w3.org/TR/webvtt1/>.

Yeramyian, Pontish, and Eric Jackson. 2014. “Tapping into Genius.” <https://www.gapinternational.com/insights/details/index/tapping-into-genius>.

Yeramyian, Pontish, and Eric Jackson. 2015. “How Purposeful Mindsets Drive Performance.” <https://www.gapinternational.com/insights/details/index/how-purposeful-mindsets-drive-performance>.

Yuan, Jiahong, Wei Lai, Christopher Cieri, and Mark Liberman. 2023. “Using Forced Alignment for Phonetics Research.” In *Chinese Language Resources*, edited by Chu-Ren Huang, Shu-Kai Hsieh, and Peng Jin, 289–301. Text, Speech and Language Technology 49. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-38913-9_17.

Appendix I

Tags and tiers for EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND brilliance theme

EXTREMENESS / INTENSITY / OBSESSION / ABOVE AND BEYOND	
Tier	Tags
Pragmatic	semantic syntactic
Rhetoric	lexical repetition syntactic repetition contrast personification escalation negation minimizing figurative expanding enumeration semantic doubling opposition certainty hyperbole pronoun shift

Word Choice	emotional charge absolute intensifier pragmatic modal verb number interjection
Hands	one hand spatial two hands spatial one hand word reinforcer two hands word reinforcer
Eye Gaze	up right up center up left down right down center down left focused gaze closed
Body	lean in lean out shoulder movement fix posture head nod affirmative head nod negative
Face	tense surprise anger happiness sad fear disgust smile narrow eyes widen eyes
Phonetic	stress breathy lower pitch raise pitch lower volume raise volume lengthening

7. Appendix II

The following table lists tags and tiers for the following brilliance themes:

- joy / excitement / exhilaration / spirit
- possibility / creation / choice / freedom
- curiosity / discovery / learning / growing
- at-stakeness / conviction / resolve / perseverance
- togetherness / relationship / connection / love
- purpose / contribution / world impact / legacy

Tags specific to certain themes are listed below the dashed line.

	JOY	POSSIBILITY	CURIOSITY	AT-STAKENESS	TOGETHERNESS	PURPOSE
	EXCITEMENT	CREATION	DISCOVERY	CONVICTION	RELATIONSHIP	CONTRIBUTION
	EXHILARATION	CHOICE	LEARNING	RESOLVE	CONNECTION	WORLD IMPACT
	SPIRIT	FREEDOM	GROWING	PERSEVERANCE	LOVE	LEGACY
Tier	Tags					
Semantic and Thematic	pragmatic universal embodied					
	magnitude	magnitude	magnitude	modal verb	empathy	for others
	joy from others	openness	joy from others	agency	interdependence	greater good
	for others	agency	for others	negative	environment	commodification
	success	modal verb	success	motivation	mutual benefit	knowing thyself
	gratitude		gratitude			values
	experiential		experiential			metaphysical
Rhetoric	figurative					
	negation lexical repetition syntactic repetition					
	semantic repetition	contrast uncertainty	contrast semantic repetition	certainty	contrast semantic repetition	contrast semantic repetition
Eye Gaze	up right					
	up left					
	up center					
	down right					
	down left					
	down center					
	focused gaze					
	closed					
Hands	one hand spatial					
	two hands spatial					
	one hand word reinforcer					
	two hands word reinforcer					
Body	lean in					
	lean out					
	shoulder movement					
	fix posture					
	head affirmative					
	head negative					
Face	tense					
	smile					
	narrow eyes					
	widen eyes					
Phonetic	Pause					
	emphatic stress					
	pitch change					
	rhythm					
	laughter			Laughter		
Syntactic	Conditional					
	reflexive	question	continuation question reflexive	continuation causative		causative comparative

	Absolute Intensifier emotional charge brilliance theme lexeme					
Lexical	spatio- temporal perception metaphysical agentive verb modal verb	spatio- temporal perception	spatio- temporal perception metaphysical agentive verb modal verb	spatio- temporal motion verb	agentive verb modal verb	spatio- temporal agentive verb modal verb magnitude scope

8. Appendix III

Acoustic features and their corresponding perceptual tag (used in annotation), perceptual correlate, and definition.

Acoustic Feature	Tag(s)	Perceptual Correlate	Definition
F0 (fundamental frequency)	pitch change, emphatic stress	Pitch	Lowest periodic cycle of the acoustic signal
Jitter	pitch change	Pitch irregularity	Frequency instability of F0
F1, F2, F3 (formant frequencies)	pitch change	Voice quality	Concentration of acoustic energy around first, second, and third formants
HNR (harmonics to noise ratio)	pitch change	Voice quality	Mean ratio of quasi-periodic to non-periodic signals across time segments
Intensity and amplitude	emphatic stress	Loudness	Measures of energy in the acoustic signal
Shimmer	pitch change	Loudness irregularity	Amplitude instability of F0
Speech rate	rhythm	Speed of speech	Number of complete elements (i.e. words / phonemes / syllables) per time unit
Articulation rate	rhythm	Speed of speech	Number of complete elements (i.e. words / phonemes / syllables), per time unit, excluding pauses
Pauses	pause	Speed of speech	Interval of silence between linguistic units

Le document : miroir des enjeux institutionnels et de l'évolution linguistique des transitions écologique et énergétique en France (2007-2022)

Camilla Nappi*

Abstract: The main aim of this contribution is to highlight the crucial importance of documentation in a study on the discursive evolution of the terms *ecological transition* and *energy transition*. The study is based on an analysis of discourses issued by the French Ministry of the Environment over the period from 2007 to 2022. After examining the role of documentation in the advancement of linguistic research, its importance in the specific context of this study will be analysed, introducing the employed approach, the investigated corpus, and the methodology that has been used. Concrete examples drawn from the corpus will be illustrated, highlighting the dual functionality of documentation in the chosen approach. From this perspective, this study aims to underline that documentation can serve both as a material and as a contextual source, providing access to valuable linguistic information while helping to interpret socio-political realities and contemporary issues.

Keywords: Discourse Analysis, Terminology, Corpus Linguistics, Phrase-discursive Approach, Environmental Discourse Analysis.

1. Introduction

Depuis l'Antiquité jusqu'à nos jours, les documents ont joué un rôle fondamental dans la transmission des savoirs, en permettant de saisir, d'organiser et de partager des connaissances essentielles pour l'avancement de nos sociétés. Qu'ils soient matérialisés sous forme de manuscrits, livres imprimés, fichiers numériques ou bases de données, ces artefacts représentent des ressources essentielles d'information qui alimentent les réflexions et les recherches des linguistes, ainsi que les interactions humaines avec le monde. En tant que véhicules de savoir, le document et l'information forment donc un duo dynamique (Molinaro 2004) qui nourrit la recherche, l'apprentissage et la prise de décision dans un large éventail de domaines (Nossereau 2003, 1).¹

* Dipartimento di Studi Letterari, Linguistici e Comparati, Università di Napoli L'Orientale, Napoli, Italia. camilla.nappi@unior.it.

En linguistique le document revêt une importance particulière en tant que source de données textuelles (Vetulani 2000, 317). À cet effet, les linguistes exploitent les documents pour examiner la structure, la variation et l'évolution des langues. Ils se penchent également sur les différentes typologies de supports documentaires afin d'observer les tendances linguistiques et de développer de nouvelles théories (Vetulani 2000, 321). Au cours des dernières décennies, les recherches linguistiques ont vu l'émergence des premières études empiriques basées sur l'utilisation de corpus linguistiques (Teubert 2009, 185). Ces vastes ensembles de données textuelles, organisées selon des principes spécifiques, offrent aux chercheuses et chercheurs la possibilité d'extraire des informations pertinentes, de mener des recherches ciblées et de comparer les données linguistiques recueillies. Grâce à l'utilisation de logiciels de traitement automatique, les linguistes peuvent, entre autres, obtenir des informations précieuses sur le fonctionnement des langues, des discours, ainsi que des unités linguistiques qu'ils contiennent. Ainsi, la documentation demeure un élément central dans les recherches linguistiques, compte tenu du fait qu'elle offre aux linguistes une base solide pour leurs analyses et le développement des théories.

Cette contribution a pour objectif fondamental de souligner l'importance de la documentation dans une étude visant à explorer l'évolution discursive des termes *transition écologique* et *transition énergétique* à travers l'analyse des discours du ministère de l'Environnement en France pendant la période 2007-2022. Après avoir abordé le rôle de la documentation dans la transmission du savoir et dans le progrès des recherches linguistiques, nous examinerons de manière concrète son apport à notre étude, en présentant en détail notre démarche, le corpus exploré, ainsi que la méthodologie adoptée. Par la suite, nous analyserons des exemples concrets tirés de notre corpus, de manière à ce que nous puissions prouver la double fonctionnalité de la documentation au sein de notre approche. En ce sens, nous visons à mettre en exergue dans quelle mesure la documentation peut agir en tant que source matérielle et contextuelle permettant à la fois d'accéder à des informations linguistiques précieuses, et d'interpréter les réalités sociopolitiques et les enjeux contemporains.

2. Le document : fondement de la transmission des savoirs et pilier de la recherche linguistique

L'idée que la documentation est essentielle pour transmettre le savoir n'est pas nouvelle, mais elle s'est établie dès les Lumières (Juanals 2003). Orientée vers la compilation et la dissémination des connaissances, l'*Encyclopédie* dirigée par Diderot et d'Alembert a joué, en ce sens, un rôle majeur dans l'expansion du champ intellectuel de la société. À une époque où la connaissance était sou-

vent réservée à une élite restreinte, cet ouvrage lexicographique a jeté les bases de l'affirmation d'une culture de l'information accessible à tous (Juanals 2002, 54), un principe de démocratisation qui continue d'influencer notre façon de penser et d'interagir avec le savoir à l'heure actuelle.

Cependant, la conception du *document* en tant que « ressource matérielle d'information » a été formalisée au début du XX^e siècle, avec des figures pionnières dans le domaine des sciences de l'information telles que les documentalistes Otlet, Briet et Meyriat. Ces penseuses et penseurs ont contribué à mettre en lumière les aspects informatifs et fonctionnels du document, le décrivant comme un moyen qui permet d'enregistrer la pensée humaine et la réalité extérieure (Otlet 1934, 10), ainsi que comme un « signe tangible ou symbolique, conservé [...] dans le but de représenter, reconstituer ou prouver un phénomène ou physique ou intellectuel » (Briet 1951, 7). Meyriat a, ensuite, mis en avant le lien intrinsèque entre *document* et *information*, en le définissant comme « un objet qui supporte l'information, qui sert à la communiquer » (Meyriat 1981 ; Couzinet 2001, 143). Il transforme ainsi le document en un moyen d'acquérir de nouvelles informations ou s'instruire.

En linguistique la reconnaissance du document en tant que base empirique pour approfondir et étudier les phénomènes de la langue remonte aux années 1970 et va de pair avec l'essor des premières techniques d'exploitation des corpus (de Angelis 2017, 81).

Au tournant des années 1990 et 2000, à une époque marquée par l'avènement massif de la technologie, se développe la linguistique de corpus « qui s'intéresse à la langue en contexte sous la forme de grands ensembles de textes » (Williams 2005, 13). Cette discipline établit, entre autres, des pratiques de numérisation qui remettent en question certains concepts fondamentaux placés à la convergence de diverses approches linguistiques. Ainsi, les notions de *texte* et *document* s'entremêlent de manière de plus en plus complexe du fait de leur interrelation et de leur complémentarité. En effet, du point de vue linguistique, le concept de *document* peut être assimilé à celui de *texte-archive*, en mettant l'accent sur le support – physique ou virtuel – qui permet la conservation d'un texte (Maingueneau 2014, 35-36). Cette évolution conceptuelle s'accompagne également d'une interrogation croissante sur la relation entre *texte* et *discours*, notamment dans le cadre de la linguistique textuelle et de l'analyse du discours, où ces notions peuvent se chevaucher ou se différencier selon la perspective d'analyse adoptée, sans qu'aucune définition définitive n'émerge encore. Cette dynamique démontre donc que les notions de *document* et *documentation* convergent de plus en plus avec celles de *texte(s)*, *discours* et *corpus*. Par exemple, un document ou une documentation peuvent être vus comme un texte ou un ensemble de textes ou de discours disponibles sur un sujet donné et qui peuvent être analysés en tant que corpus.

D'une part, l'interrelation croissante entre ces concepts traduit l'impact de l'informatisation de la langue à travers les bases de données textuelles, domaine qui existait déjà avant l'ère des ordinateurs. En effet, la révolution technologique a considérablement amélioré les aspects qualitatifs et quantitatifs de l'exploration des données linguistiques, en enrichissant les corpus et leur taille, et en rendant des outils de traitement automatique plus accessibles (Ji 2017). À cet effet, les corpus sont devenus des instruments indispensables dans divers domaines linguistiques, notamment pour la description des langues, la rédaction de dictionnaires et grammaires, ainsi que pour l'enseignement des langues (Seifen et Bunkham 2022). Parallèlement, les chercheuses et chercheurs ont commencé à exploiter les corpus pour répondre à des questions sur le fonctionnement des langues, tester des hypothèses et observer plus finement les phénomènes linguistiques. L'analyse de corpus textuels se configure, pourtant, comme un passage fondamental dans la recherche linguistique actuelle, ce qui met en valeur le rôle de la documentation en tant qu'outil privilégié d'accès à l'information et à la connaissance dans ce contexte particulier.

3. Documenter la réalité sociopolitique et analyser l'évolution discursive des transitions écologique et énergétique en France (2007-2022)

La section précédente de notre contribution a souligné l'importance vitale du document comme moyen principal d'accès à la connaissance dans le domaine des recherches linguistiques contemporaines. Dans les paragraphes suivants, nous nous concentrerons particulièrement sur son utilité concrète au sein d'une étude linguistique. Pour ce faire, nous nous appuyerons sur notre recherche doctorale en cours axée sur l'évolution conceptuelle et discursive des notions de *transition écologique* et *transition énergétique* dans les discours du ministère de l'Environnement en France pendant la période 2007-2022.

Notre objectif principal consiste à prouver la dualité fonctionnelle des documents contenus dans notre corpus, et que nous détaillerons par la suite, en mettant en lumière leur capacité à capturer à la fois des aspects concrets de la réalité et des dynamiques linguistiques en évolution. Ainsi, notre analyse vise à souligner deux aspects interconnectés. D'une part, nous souhaitons montrer comment les documents contribuent à « enregistrer la pensée humaine et la réalité extérieure » et à « repenser ou prouver un phénomène ou physique ou intellectuel », conformément à la conception des documentalistes Otlet et Briet (cfr. par. 2). Dans notre contexte spécifique, les documents permettent, de fait, de relever les actions politiques et économiques de soutien aux transitions écologique et énergétique en France. D'autre part, nous cherchons à mettre en lumière la contribution des documents à la diffusion d'informations

sur le sujet et à l'avancement des réflexions portant sur des phénomènes linguistiques donnés. Cela implique spécifiquement d'observer les aspects phraséologiques et l'évolution discursive des termes étudiés au fil du temps, sous l'influence des événements extralinguistiques et sociopolitiques. Ensuite, nous exposerons brièvement le travail qui a nourri cette contribution, et mettrons l'accent sur la composition du corpus et les principes méthodologiques qui ont guidé notre analyse.

3.1. *Contexte d'étude et méthodologie : fondements du corpus et principes analytiques*

Notre recherche doctorale s'inscrit dans le champ émergent des études linguistiques sur l'environnement, telles que préconisées par plusieurs chercheuses et chercheurs (Chetouani et Tournier 1994 ; Dury 1999 ; L'Homme 2016 ; Bonnet et Geslin 2019 ; Rakotoelina et Reboul-Touré 2020 ; Balnat et Gérard 2022 ; Hamon et Paissa 2023 ; Altmanova et Cartier 2024) et essaie d'analyser les discours provenant du ministère de l'Environnement en France au cours des quinze dernières années, dans le but de comprendre la construction du défi sociétal des transitions écologique et énergétique. Cette étude adopte une approche théorique et méthodologique qui s'inspire à la fois des démarches de terminologie textuelle et diachronique (Dury et Picton 2009 ; Condamines 2018 ; Picton 2018), et de l'approche de la « formule » dans le cadre de l'analyse du discours institutionnel (Krieg Planque 2009 ; Krieg-Planque et Oger 2010). Notre perspective hybride permet ainsi d'examiner les influences extralinguistiques et contextuelles sur le comportement discursif des termes *transition écologique* et *transition énergétique*, tout en mettant en lumière leurs évolutions conceptuelles¹ au fil des années et les enjeux sociopolitiques et institutionnels qu'ils cristallisent à différents moments de l'histoire des politiques environnementales en France.

Pour notre recherche, nous avons constitué un corpus *ad hoc* nommé « Co-TEc-Ten », regroupant les prises de parole publiques de treize ministres de l'Environnement sous les présidences de Sarkozy et Macron². D'un point de

¹ Il est important de noter que l'approfondissement du parcours conceptuel des termes étudiés, ainsi que leur comportement en tant que « formule » selon la définition de Krieg-Planque, ne seront pas traités dans la présente étude. En effet, ces aspects sont abordés plus en détail dans notre travail de thèse, ainsi que dans une étude préliminaire sur ce sujet et qui est en cours de publication (Nappi 2024).

² Les ministres pris en compte dans notre analyse sont les suivants : Jean-Louis Borloo, Nathalie Kosciusko-Morizet, Nicole Bricq, Delphine Batho, Philippe Martin, Ségolène Royal, Nicolas Hulot, François de Rugy, Élisabeth Borne, Barbara Pompili, Amélie de Montchalin, Agnès Pannier-Runacher et Christophe Béchu. Ces interventions publiques représentent un échantillon représentatif de la diversité au sein de la majorité ministérielle, ce qui nous permet de ne pas prendre en considération les spécificités individuelles de chaque intervenant.

vue quantitatif, ce corpus présente une taille moyenne (Chiari 2007, 44-45) et comprend 327 textes, sélectionnés à partir de la « Collection des discours publics » du portail d'information Vie-publique.fr. En ce qui concerne le plan chronologique, lors de la création du corpus, nous avons adopté une approche en diachronie courte, en délimitant une fenêtre temporelle de quinze ans afin de saisir l'évolution conceptuelle et discursive des transitions écologique et énergétique. À cet égard, nous avons pris en considération deux événements socio-historiques et extralinguistiques majeurs qui ont marqué l'institutionnalisation des défis environnementaux : le Grenelle de l'environnement de 2007 et la création du ministère de la Transition énergétique lors de la formation du gouvernement de Borne en 2022. Par ailleurs, nous avons divisé notre corpus en trois sous-corpus – S1 (2007-2014), S2 (2015-2017) et S3 (2018-2022) – dans le but de cerner les moments clés qui ont façonné l'histoire des politiques publiques environnementales en France, depuis les premières mesures politiques en faveur des transitions jusqu'à leur transformation en tant que piliers des initiatives ministérielles axées sur l'environnement. Ces critères méthodologiques – chronologiques et quantitatifs – ont pourtant joué un rôle essentiel dans la mise en valeur de la documentation au sein de notre étude termino-discursive. En effet, ils ont permis de planifier la collecte de manière rigoureuse, en garantissant la fiabilité et la représentativité des données en vue de leur traitement informatique.

Après avoir organisé et préparé les données textuelles, nous avons mené une analyse outillée à l'aide du logiciel Sketch Engine, et nous avons adopté une approche lexicométrique et phraséologique (Fiala 1994 ; Mayaffre 2008). Cette méthodologie repose principalement sur l'analyse de la fréquence d'occurrence des termes étudiés, en explorant les concordances, les cooccurrences, les mots-clés (*KWIC*) et les segments répétés (*n-grams*).

Notre analyse quantitative du corpus a donc révélé des aspects phraséo-discursifs intéressants concernant les notions de *transition écologique* et *transition énergétique*, ainsi que les termes récurrents dans les discours ministériels de la période étudiée. Plus précisément, nous avons examiné le comportement lexico-syntaxique de ces termes, en observant leurs cooccurents fréquents et les structures lexicales et syntaxiques qui les accompagnent, ce qui a permis de déduire des informations significatives sur la manière à travers laquelle ils opèrent dans les discours analysés. Par ailleurs, nous avons établi une liste des mots et termes complexes les plus fréquents dans le corpus, et nous avons observé leur répartition dans les sous-corpus pour repérer les évolutions et variations au fil du temps.

En se basant sur ces résultats, les sections suivantes proposeront des exemples tirés de notre corpus afin de mettre en lumière le rôle central de la documentation ministérielle dans la démarche adoptée.

3.2. *Comprendre les transitions écologique et énergétique à travers la documentation ministérielle*

Notre analyse repose sur deux axes fondamentaux qui cherchent à mettre en relief la double fonctionnalité de la documentation ministérielle dans la compréhension des transitions écologique et énergétique. D'une part, nous aborderons les discours ministériels en tant que source matérielle fournissant un aperçu des actions et des engagements institutionnels en faveur de ces transitions. En parallèle, nous mettrons en évidence le rôle essentiel de ces discours dans l'évaluation de l'influence d'événements extralinguistiques sur l'évolution discursive des termes étudiés, en analysant leur fréquence et leur distribution temporelle. D'autre part, nous soulignerons également l'importance de la documentation en tant qu'observatoire de phénomènes linguistiques. À travers une analyse phraséo-discursive, nous examinerons les évolutions des intentions et identifierons les visions et les valeurs sous-jacentes à ces engagements. En étudiant les cooccurrences de *transition écologique* et *transition énergétique* au sein de notre documentation, nous cherchons ainsi à éclairer les préoccupations et les orientations du ministère au fil du temps.

3.2.1. *La documentation ministérielle : une fenêtre sur les dynamiques sociopolitiques et extralinguistiques*

À l'ère contemporaine où les préoccupations environnementales sont devenues cruciales à l'échelle mondiale, les transitions écologique et énergétique occupent une place centrale dans les agendas politiques. Ces changements, qui visent à réduire les impacts néfastes sur l'environnement, nécessitent d'actions concertées et d'engagements institutionnels solides. Dans ce contexte, la documentation ministérielle émerge, d'abord, comme une source de données importante pour évaluer ces actions et engagements. En tant que reflet tangible des politiques gouvernementales, les discours ministériels offrent un aperçu précieux des mesures concrètes mises en œuvre pour promouvoir ces transitions. À ce titre, l'identification des n-grammes les plus fréquents dans notre corpus permet de mettre en évidence les segments répétés qui révèlent les axes principaux du ministère de l'Environnement en matière de transitions écologique et énergétique. Parmi ceux-ci, figurent notamment des dénominations de mesures législatives telles que *Loi de programmation pour la transition énergétique* et *Loi climat (et) Résilience*, ainsi que des initiatives ministérielles comme *stratégie bas carbone* et *programmation pluriannuelle de l'énergie*. De plus, des événements clés de l'histoire des politiques climatiques en France, tels que *Grenelle de l'environnement* et *Débat national pour la transition énergétique*, sont aussi mentionnés dans notre corpus. En examinant de près les discours ministériels, nous pouvons également observer comment ces références

sont utilisées pour légitimer les actions entreprises et pouvoir communiquer les engagements institutionnels au public. Ces références sont susceptibles de constituer des indicateurs précieux de l'orientation politique et des actions concrètes menées pour favoriser les transitions écologique et énergétique.

En tant que support matériel, les discours ministériels permettent, entre autres, d'observer comment les événements extralinguistiques reflétés dans la documentation ministérielle ont pu influencer les actions gouvernementales et contribuer à l'institutionnalisation de certaines questions sociétales. L'analyse fréquentielle des occurrences des termes *transition écologique* et *transition énergétique*, ainsi que leur périodisation dans notre corpus offrent des exemples concrets de la manière dont les événements sociohistoriques ont façonné le destin discursif de ces notions. Ces tendances révèlent des périodes de stabilité, de croissance et de fluctuation dans l'utilisation de ces termes, souvent en réponse à des initiatives et des événements politiques majeurs. Initialement peu fréquents dans les discours ministériels entre 2008 et 2011, période marquée par l'adoption des lois issues du Grenelle de l'environnement (*Loi Grenelle I* de 2008 et *Loi Grenelle II* de 2009), ces termes connaissent une croissance significative de leurs occurrences depuis 2012, avec une prédominance marquée de *transition écologique* en 2013, suivie d'une forte augmentation de l'utilisation de *transition énergétique* en 2015. Un examen plus détaillé de leurs occurrences annuelles dans notre corpus révèle que les deux termes atteignent leur pic de fréquence entre 2012 et 2015, avant de décliner à partir de 2017. Entre 2018 et 2020, une période de fluctuations est observée, suivie d'une remontée depuis 2021. La périodisation des occurrences de *transition écologique* et *transition énergétique* dans notre corpus révèle donc des tendances intéressantes, susceptibles d'ouvrir la voie à plusieurs hypothèses significatives. Tout d'abord, les années 2012-2015 apparaissent comme une période de stabilisation conceptuelle et linguistique de ces notions dans les discours ministériels. Cette hausse d'utilisation coïncide, entre autres, avec le lancement d'initiatives ministérielles clés, telles que le Débat National pour la Transition Énergétique de 2012 et la création du Conseil National de la Transition Écologique en 2013, qui ont contribué à l'institutionnalisation de l'urgence écologique. En effet, elles ont abouti à l'approbation de la *Loi relative à la transition énergétique pour la croissance verte de 2015*. Deuxièmement, l'augmentation progressive des occurrences de *transition énergétique* suggère une montée en importance de ce défi au sein des préoccupations institutionnelles, peut-être en réponse aux priorités politiques internationales telles que la COP21 et l'Accord de Paris. En revanche, la stagnation des occurrences de *transition écologique* témoigne d'un certain décalage entre cet enjeu et les priorités politiques. D'ailleurs, la diminution des occurrences des deux termes à partir de 2017 peut être liée à l'accomplissement de la politisation des défis écologique et énergétique, symbolisée par le changement dénomiatif du ministère, qui intègre de manière

définitive le mot *transition* dans ses attributions pour les années à venir. Par la suite, l'augmentation des occurrences des deux termes entre 2021 et 2022 est à attribuer à l'avènement de la crise énergétique et de la guerre russo-ukrainienne, qui ont remis en avant les transitions écologique et énergétique parmi les préoccupations ministérielles. L'observation des occurrences annuelles de *transition écologique* et *transition énergétique* révèle donc l'influence significative d'événements clés de l'histoire des politiques environnementales nationales et internationales sur la trajectoire évolutive et discursive de ces termes au cours de la période étudiée. Dans ce contexte, la documentation joue donc un rôle essentiel, compte tenu du fait qu'elle permet d'identifier et mesurer ces influences.

3.2.2. *La documentation ministérielle comme observatoire de l'évolution discursive des termes au fil du temps*

Comme révélé par notre analyse, les discours ministériels témoignent des efforts gouvernementaux et institutionnels visant à encourager les transitions écologique et énergétique. De plus, nous estimons qu'ils offrent une perspective linguistique sur l'évolution des discours et des tendances qui influent sur les politiques et les actions du gouvernement dans ce domaine. En explorant la documentation ministérielle, nous sommes, en effet, en mesure de détecter les changements discursifs des termes, en illustrant ainsi leur adaptation en réponse aux évolutions politiques, sociétales et environnementales. Ainsi, nous cherchons à prouver comment ces transformations sont susceptibles de refléter des changements dans les priorités politiques, les avancées technologiques ou les pressions sociétales.

En adoptant une approche phraséo-discursive, nous avons exploré le profil combinatoire des termes complexes *transition écologique* et *transition énergétique*. Pour ce faire, nous avons utilisé la fonction *Word Sketch* du logiciel Sketch Engine, qui a notamment permis d'approfondir le comportement lexico-syntaxique de ces termes dans le corpus, en mettant en lumière leurs schémas grammaticaux ainsi que leurs collocations nominales, verbales, pronominales et prépositionnelles. Par ailleurs, nous avons employé les fonctions *Wordlist*, *N-grams* et *Keywords* pour analyser les mots-témoins et les termes clés présents dans les discours ministériels portant sur les transitions écologique et énergétique. Le but a été ainsi d'identifier les mots-thèmes fréquemment associés à ces transitions ainsi que les *topois* discursifs caractérisant les allocutions ministérielles sur ces défis. Cela a permis, entre autres, de comparer les sous-corpus afin de détecter les évolutions discursives spécifiques à des périodes données.

Le décryptage phraséo-discursif de *transition écologique*, qui inclut notamment l'analyse de ses cooccurrences les plus fréquentes dans notre corpus, a mis en relief des aspects intéressants concernant son comportement discursif. À ce titre, les substantifs fréquemment associés à ce terme, tels que *cohésion* et *développement*, semblent évoquer une intention de présenter la transition écologique de manière positive, en l'intégrant dans des initiatives plus larges de cohésion territoriale et de promotion de la durabilité :

- « Le simple fait d'avoir lié transition écologique et cohésion des territoires atteste le rôle essentiel des collectivités » (Discours du ministre Béchu, 13/09/2022) ;
- « L'initiative [...] en matière de déchets ou d'énergie, de transition écologique et de développement durable » (Discours de la ministre Royal, 03/04/2017).

Parallèlement, l'association de *transition écologique* à l'adjectif *solidaire*, ainsi que le recours à la marque d'énonciation de proximité *notre*, mettent en évidence sa dimension inclusive et équilibrée. Ce faisant, la transition écologique dévient un processus collectif où chacun est appelé à contribuer à sa réalisation :

- « La transition écologique et solidaire est une question, aussi, de mobilité » (Discours du ministre Hulot, 21/11/2017) ;
- « Pour réussir notre transition écologique » (Discours de la ministre, 07/10/2021).

Enfin, la collocation de ce terme avec des verbes axiologiques liés au noyau sémantique « action/mobilisation/concrétisation », tels que *réussir*, *faire*, *engager* et *accélérer*, renforce la perspective positive, tout en soulignant l'urgence des actions à entreprendre :

- « L'année qui va s'ouvrir est déterminante pour réussir la transition écologique » (Discours de la ministre Borne, 10/12/2019) ;
- « Encourager l'émergence de nouveaux financements pour engager fermement la transition écologique » (Discours de la ministre Bricq, 23/05/2012) ;
- « Enfin, le chamboulement géopolitique [...] vont accélérer la transition écologique » (Discours de la ministre Pompili, 01/04/2022).

En ce qui concerne l'analyse phraséo-discursive de *transition énergétique*, des nuances se dégagent, en plaçant ce terme dans une perspective davantage économique et en soulignant son impact sur les enjeux environnementaux. En effet, nous observons une récurrence collocationnelle avec des mots tels que *croissance*, *économie*, *biodiversité* et *dérèglement* :

- « En évoquant la transition énergétique et la croissance verte, nous inventons le futur » (Discours de la ministre Royal, 19/02/2015) ;

- « C'est maintenant chose faite [...] les chantiers de l'économie circulaire et de la transition énergétique » (Discours de la ministre Royal, 12/11/2014) ;
- « Le Président de la République avait fixé [...] la transition énergétique et la biodiversité » (Discours de la ministre Batho, 05/04/2013).

L'analyse des cooccurrences verbales de *transition énergétique* révèle, entre autres, une perspective proactive et opérationnelle associée à ce terme. En effet, la présence de verbes axiologiques liés au noyau sémantique « intensification/anticipation/urgence » (e.g. *sécuriser*, *accentuer* et *anticiper*) suggère une volonté d'intensifier les efforts et de planifier stratégiquement des mesures pour relever les défis de la transition énergétique :

- « Ces atouts [...] sécurisent notre transition énergétique » (Discours de la ministre Royal, 01/10/2014) ;
- « C'est donc [...] accentuer la transition énergétique » (Discours de la ministre Royal, 16/01/2015) ;
- « On a anticipé la transition énergétique » (Discours de la ministre Royal, 09/08/2014).

Par la suite, l'analyse fréquentielle des mots-témoins et des termes clés présents dans les discours ministériels sur les transitions écologique et énergétique a permis de mettre en lumière les implications sémantiques et les tendances émergentes dans ces discours. Il est intéressant de noter que le mot *énergie* surpasse en fréquence le mot *transition* dans le corpus, avec 1861 occurrences. Cette prédominance souligne un intérêt marqué pour les enjeux énergétiques et confirme l'existence de *topoïs* discursifs privilégiés dans les allocutions ministérielles étudiées. Cette focalisation sur l'énergie se manifeste linguistiquement par la présence d'un vocabulaire spécifique à ce domaine. L'analyse temporelle des occurrences révèle parallèlement une augmentation significative du mot *énergie* entre 2018 et 2022 (sous-corpus S3), ce qui suggère une évolution de l'importance accordée à ce sujet au fil du temps, potentiellement influencée par des événements socio-historiques majeurs, tels que la crise énergétique de 2021.

D'autre part, l'analyse des n-grammes et des termes clés dans les sous-corpus a révélé les évolutions discursives associées aux transitions écologique et énergétique au fil des années. Les discours ministériels de la période 2007-2014 se démarquent comme une phase initiale de réflexion, de débat public et d'élaboration de cadres législatifs pour orienter les politiques environnementales. Cela se manifeste par l'utilisation fréquente de mots liés aux ressources énergétiques (e.g. *énergies renouvelables* et *gaz*), ainsi que par la mention des problématiques environnementales (e.g. *effet de serre* et *réchauffement climatique*). De plus, des événements marquant de l'histoire des politiques publiques environnementales (e.g. *Grenelle de l'environnement* et *Débat national sur la transition énergétique*) sont également abordés dans ces discours. Cependant, les allocutions

ministérielles de la période 2015-2017 témoignent d'une phase intermédiaire caractérisée par des politiques plus concrètes, une réglementation renforcée et une intégration plus étroite des aspects financiers liés à la transition. Parmi les termes clés et les n-grammes récurrents figurent pourtant : *prix carbone*, *tarification du carbone* et *obligation verte*. Enfin, les discours ministériels de la période 2018-2022 mettent en lumière une attention particulière aux enjeux géopolitiques liés à l'énergie et symbolisent le passage de la sensibilisation à la mise en œuvre concrète des politiques énergétiques. En ce sens, la diversification des termes relatifs à la sphère énergétique tels que *nucléaire*, *efficacité énergétique* et *sécurité d'approvisionnement*, ainsi que l'augmentation des termes liés à la production et à la consommation d'énergie, montrent la transition vers une gestion opérationnelle plus prédominante. Par ailleurs, la mention fréquente du *gaz russe* souligne une préoccupation croissante pour les questions géopolitiques relatives à l'approvisionnement en gaz, ce qui met en exergue l'importance de la sécurité énergétique et des relations internationales dans les discours ministériels contemporains.

L'analyse de la documentation ministérielle offre donc un aperçu précieux de l'évolution discursive des transitions écologique et énergétique au fil du temps. En étudiant les cooccurrences, les n-grammes et les termes clés, nous avons pu dégager les trajectoires discursives spécifiques à chaque période, en révélant les priorités, les orientations politiques et les préoccupations gouvernementales en matière de transitions écologique et énergétique. Ainsi, la documentation ministérielle apparaît comme un outil essentiel pour éclairer également la complexité des défis permanents liés à ces transitions.

4. Conclusion

L'objectif de cette contribution a été de mettre en lumière le rôle de la documentation dans l'étude de l'évolution discursive des termes *transition écologique* et *transition énergétique* à travers un corpus de discours provenant du ministère de l'Environnement en France pendant la période 2007-2022.

En analysant ces discours, nous avons ainsi dévoilé la double nature de la documentation en tant que fenêtre sur les réalités socio-politiques et extralinguistiques et observatoire des phénomènes linguistiques. D'une part, les discours ministériels fournissent un contexte textuel précieux pour comprendre les initiatives institutionnelles et les mesures législatives pour promouvoir ces transitions, ainsi que pour évaluer l'impact des événements socio-historiques majeurs sur l'utilisation de ces termes. En ce sens, les politiques nationales et internationales, ainsi que des événements clés tels que l'adoption d'une loi ou l'organisation d'un débat, sont autant de facteurs qui peuvent influencer sur l'utilisation des termes. Cette observation montre de manière évidente le pouvoir des acteurs institutionnels dans la création et la consolidation de concepts

clés dans l'arène publique et politique. D'autre part, les discours ministériels servent également d'outil d'analyse pour étudier les phénomènes linguistiques dans des domaines émergents, à l'instar des transitions écologique et énergétique. Notre analyse phraséo-discursive a donc révélé les nuances et les variations dans l'usage des termes *transition écologique* et *transition énergétique* dans le contexte ministériel ainsi que les visions et les valeurs qui sous-tendent ces concepts. Parallèlement, nous avons également repéré les mots-thèmes et les *topoïs* discursifs qui les accompagnent, tout en montrant comment ils reflètent des changements dans les priorités politiques et les pressions sociétales au fil du temps. En saisissant les discours et les actions des institutions, la documentation dépasse sa matérialité pour devenir le reflet authentique des réalités sociopolitiques d'une époque donnée, en offrant ainsi une perspective éclairante sur les événements, les transformations et les défis majeurs rencontrés au cours d'une période historiquement significative.

Références bibliographiques

- Altmanova, Jana, Emmanuel Cartier, et Silvia Domenica Zollo. 2024 (à paraître). "Variations terminologique dans le domaine de la biodiversité et du changement climatique." *Repère Dorif*.
- Balnat, Vincent, et Gérard Christophe, sous la direction de. 2022. *Néologie et environnement. Neologica* 16.
- Bonnet, Valérie, et Albane Geslin. 2019. "Les mots de l'écologie, 25 ans après. Circulation des discours et des notions." *Mots. Les langages du politique* 119 : 9-14. <https://doi.org/10.4000/mots.24186>.
- Briet, Suzanne. 1951. *Qu'est-ce que la documentation ?* Paris : ÉDIT.
- Chetouani, Lamria, et Michel Tournier, sous la direction de. 1994. *Environnement, Écologie, Verts. Mots* 39. www.persee.fr/doc/mots_0243-6450_1994_num_39_1_1891.
- Chiari, Isabella. 2007. *Introduzione alla linguistica computazionale*. Roma: Editori Laterza.
- Condamines, Anne. 2018. "Nouvelles perspectives pour la terminologie textuelle." In *Terminology & Discourse/Terminologie et discours*, sous la direction de Jana Altmanova, Maria Centrella et Katherine E. Russo, 1-13. Bern : Peter Lang.
- Couzinet, Viviane, sous la direction de. 2001. *Jean Meyriat, théoricien et praticien de l'information-documentation*. Paris : ADBS édition.

- de Angelis, Rossana. 2017. "La linguistique de corpus à l'épreuve du numérique : textes, textures, documents." *Dossiers d'HEL, Analyse et exploitation des données de corpus linguistiques* 11 : 81-95. https://hal.science/hal-01511280/file/Dossier_HEL_11_9_De_Angelis.pdf.
- Dury, Pascaline. 1999. "Les variations sémantiques en terminologie : étude diachronique et comparative appliquée à l'écologie." In *Sémantique des termes spécialisés*, sous la direction de Valérie Delavigne et Myriam Bouveret, 17-33. Rouen : Publications de l'Université de Rouen.
- Dury, Pascaline, et Aurélie Picton. 2009. "Terminologie et diachronie : vers une réconciliation théorique et méthodologique ?" *Revue française de linguistique appliquée* 14 (2) : 31-41. <https://doi.org/10.3917/rfla.142.0031>.
- Fiala, Pierre. 1994. "L'interprétation en lexicométrie. Une approche quantitative des données lexicales." In *Le lexique : construire l'interprétation*, sous la direction de Simone Lecointre et Danielle Leeman, *Langue française* 103 : 113-122. www.persee.fr/doc/lfr_0023-8368_1994_num_103_1_5731.
- Hamon, Yannick, et Paola Paissa. sous la direction de. 2023. *Discours environnementaux. Convergences et divergences. Lingue d'Europa e del Mediterraneo*. Roma : Aracne.
- Ji, Yujing. 2017. "Linguistique de corpus : une introduction." *Arts, langage, apprentissages*. 14 juin 2017. <https://arlap.hypotheses.org/9517>.
- Juanals, Brigitte. 2002. "L'encyclopédie, des Lumières au numérique : migration d'une utopie." *Communication et langages* 131 : 53-65. https://www.persee.fr/doc/colan_0336-1500_2002_num_131_1_3125.
- Juanals, Brigitte. 2003. *La culture de l'information : du livre au numérique*. Paris : Hermès science publications.
- Krieg-Planque, Alice. 2009. *La Notion de « Formule » en analyse du discours. Cadre théorique et méthodologique*. Besançon : Presses Universitaires de Franche-Comté.
- Krieg-Planque, Alice, et Claire Oger. 2010. "Discours institutionnels. Perspectives pour les sciences de la communication." *Mots. Les langages du politique* 94 : 91-96. <https://doi.org/10.4000/mots.19870>.
- L'Homme, Marie-Claude. 2016. "Terminologie de l'environnement et sémantique des cadres." In *SHS Web of Conferences* 27, *Congrès Mondial de Linguistique Française-CMLF, 04 novembre 2016*. https://www.shsconferences.org/articles/shsconf/pdf/2016/05/shsconf_cmlf2016_05010.pdf.
- Maingueneau, Dominique. 2021. *Discours et analyse du discours*. Paris : Armand Colin.

- Mayaffre, Damon. 2008. "L'entrelacement lexical des textes. Cooccurrences et lexicométrie." *Texte et Corpus* 3 : 91-102. <https://hal.science/hal00553808/document>.
- Meyriat, Jean. 2001. "Document, documentation, documentologie." In *Jean Meyriat, théoricien et praticien de l'Information-Documentation*, sous la direction de Viviane Couzinet, 143-59. Paris : ADBS Éditions.
- Molinaro, Fabrice. 2004. "Document et information." *Les Infostratèges*. 16 janvier 2004. <https://www.les-infostrateges.com/article/document-et-information>.
- Nappi, Camilla. 2024. "Les notions de 'transition écologique' et 'transition énergétique' au prisme du discours institutionnel : entre variation et enjeux environnementaux." *Repère DoRiF* 30, Variations terminologiques et innovations lexicales dans le domaine de la biodiversité et du changement climatique. Roma : DoRiF Università. <https://www.dorif.it/reperes/camilla-nappi-les-notions-de-transition-ecologique-et-transition-energetique-au-prisme-du-discours-institutionnel-entre-variation-et-enjeux-environnementaux/>.
- Nossereau, Alain. 2004. "Le document comme contenant, contenu et médium. Les reformulations du numérique." https://archivesic.ccsd.cnrs.fr/sic_00001115.
- Otlet, Paul. 1934. *Traité de documentation : le livre sur le livre*. Brussels : Éditions Mundaneum.
- Picton, Aurélie. 2018. "Terminologie outillée et diachronie : éléments de réflexion autour d'une réconciliation." *ASp* 74 : 27-52. <https://doi.org/10.4000/asp.5255>.
- Rakotoanelina, Florimond, et Sandrine Reboul-Touré. 2020. "La biodiversité en discours : communication, transmission, traduction." *Les Carnets du Cediscor* 15. <https://doi.org/10.4000/cediscor.1591>.
- Seifen, Karl, et Nichuta Bunkham. 2022. "Apports des corpus linguistiques : étude de « aller » et « venir » en français et en thaï." *Moussons* 40 : 65-90. <https://doi.org/10.4000/moussons.9804>.
- Teubert, Wolfgang. 2009. "La linguistique de corpus : une alternative [version abrégée]." *Semen* 27 (1) : 185-211. <https://doi.org/10.4000/semen.8914>.
- Vetulani, Grażyna. 2007. "Quelques exemples d'analyse des corpus en vue de la traduction." *Studia Romanica Posnaniensia* 25 : 317-25. <https://doi.org/10.14746/strop.2000.2526.029>.
- Williams, Geoffrey. 2005. *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes.

Testimonianze

La Documentazione in Italia e il ruolo di Paolo Bisogno: una rapida evoluzione

Piero Cavaleri*, Ferruccio Diozzi**

«...la documentazione è prima di tutto analisi e categorizzazione e solo successivamente diviene attività che si esplica attraverso procedure e con linguaggi adeguati»

Paolo Bisogno

1. Una considerazione preliminare

Per tutti coloro i quali hanno avuto parte nello sviluppo della Documentazione in Italia esistono probabilmente due punti fermi: da un lato il debito di riconoscenza intellettuale e umana nei confronti di Paolo Bisogno; dall'altro la crisi della Documentazione stessa, sia come disciplina che come ambito applicativo, sviluppatasi nel corso dei primi anni del nuovo secolo. Nella qualità di ex presidenti di AIDA, Associazione Italiana Documentazione Avanzata, oltre che come persone tuttora pienamente impegnate nella società dell'informazione, abbiamo accettato con grande piacere l'invito della redazione di «AIDAinformazioni», a preparare uno dei contributi previsti per il numero speciale che celebra i quaranta anni di vita della rivista. Lo facciamo ragionando, in maniera assolutamente preliminare, su questi due elementi sperando di concorrere alla loro migliore conoscenza.

* Già presidente AIDA (Associazione Italiana Documentazione Avanzata). cavaleri.piero@gmail.com.

** Già presidente AIDA ferruccio.diozzi@gmail.com.

2. La Documentazione in Italia negli ultimi decenni del ventesimo secolo

I documentalisti in Italia nella seconda parte del ventesimo secolo hanno svolto un ruolo cruciale nella gestione, nell'organizzazione e nella diffusione delle informazioni in vari settori, tra cui biblioteche, archivi e centri di informazione.

È all'interno della loro comunità che si sono sviluppate le idee più interessanti per innovare e rendere adeguato alle esigenze di un mondo in costante e rapido cambiamento – cambiamento consistente in larga parte nella sempre maggiore necessità di utilizzare l'informazione per produrre, agire e imparare – il sistema per la produzione, raccolta e distribuzione delle informazioni e dei documenti che le veicolano.

La consapevolezza dei documentalisti dell'impatto delle tecnologie elettroniche sul sistema di distribuzione delle informazioni, in primo luogo delle informazioni tecnico-scientifiche, è stata fondamentale per introdurre anche in Italia il tema della necessità di creare strutture atte a favorire la circolazione delle informazioni, aiutare la crescita di una cultura dell'informazione come bene produttivo e promuovere la formazione di personale adeguato per far fronte alle nuove esigenze.

I documentalisti sono stati responsabili della proposta di innovazione nella catalogazione, nell'indicizzazione e nella conservazione di un'ampia gamma di documenti – in particolari nei campi della ricerca e delle applicazioni industriali - garantendo l'accessibilità e il reperimento delle informazioni per i ricercatori, gli studiosi ma anche per il pubblico in generale. Inoltre, hanno contribuito allo sviluppo e all'implementazione di sistemi e standard di gestione delle informazioni, facendo progredire il campo della Documentazione e della Scienza dell'informazione. Il loro lavoro è stato fondamentale per sostenere la ricerca, l'istruzione e la conservazione del patrimonio culturale del Paese (Carosella e Valenti 1983).

3. L'approccio di Bisogno

Nel contesto qui su illustrato la rilevanza della figura di Paolo Bisogno è indubbia sia per il contributo scientifico fornito alla disciplina sia per le politiche di sviluppo della professione che egli seppe suscitare e per un suo sintetico profilo biografico rimandiamo a Petrucciani (2006).

Sul primo versante occorre partire dalla capacità di Bisogno di aggiornare il paradigma disciplinare della Documentazione innestando, sulla classica elaborazione di Paul Otlet e della sua scuola, approcci e stimoli nuovi che Bisogno traeva assieme dall'osservazione della realtà in evoluzione e dalla presa in carico di punti di vista e sollecitazioni nuove che interessavano certo la Documentazione ma anche le discipline vicine come la Biblioteconomia e l'Archivistica

e altri campi di ricerca, dalla Epistemologia alla Sociologia della Conoscenza. Da questo punto di vista resta fondamentale il testo di Bisogno del 1979 (Bisogno 1979) e sono molto importanti i ragionamenti condotti sulla sua eredità intellettuale (Castellucci 2003; 2007a; 2007b).

Sul versante applicativo essere stato assieme artefice della nascita e dello sviluppo di un istituto di ricerca come il CNR/ISRDS, Istituto di studi sulla ricerca e documentazione scientifica e di un'associazione scientifica e professionale come l'AIDA, Associazione Italiana Documentazione Avanzata, è la migliore testimonianza di un impegno lontano da ogni separatezza tra elaborazione disciplinare e definizioni di approcci operativi. In altre parole Bisogno voleva consolidare quanto in ambito scientifico si andava sviluppando in una dimensione istituzionale ed associativa in modo tale che sviluppi conoscitivi, con l'aggiornamento continuo dei paradigmi di riferimento, e operatività viaggiassero sempre paralleli. Tutto ciò in una robusta visione d'insieme che dimostrava la chiarissima percezione delle trasformazioni epocali, già evidenti negli anni ottanta del secolo scorso, come si può facilmente dedurre da quanto Bisogno (1985, 19-20) scriveva:

Siamo tutti d'accordo che la maggior parte dei sintomi di cattivo funzionamento della società odierna deriva dal fatto che un mondo ristretto, composto di piccoli gruppi, si è dilatato in un tempo relativamente breve, sino a divenire una grande società. Questa trasformazione non guidata, in taluni ambiti tumultuosa, in altri accelerata ma sempre ineguale, ha causato disordini, incongruenze, in ogni parte della vita sociale. Infatti la tendenza al superamento di molteplici confini (geografici, politici, ideologici, razziali, religiosi, di costume, ecc.) provocata dall'azione raziocinante della scienza, suscita contrasti e squilibri perché muta il quadro di riferimento logico-psicologico dell'individuo e della società. E ormai sappiamo che se è vero che l'ambiente sociale stimola con le sue richieste le risposte della scienza è anche vero che la società è talmente influenzata dai risultati scientifici da modificare le sue esigenze e le sue sollecitazioni.

Rileggendo queste righe si apprezzano pienamente le capacità di analisi e di *forecast*, di previsione sociale di una personalità come quella di Bisogno, soprattutto se le si confrontano con le molte visioni unilineari dello sviluppo del consorzio umano che proliferarono a livello *mainstream* in quegli anni (Fukuyama 1992).

La fine prematura non ha permesso a Bisogno di sviluppare il suo impegno anche nel nuovo millennio e chi lo ha seguito e studiato sa quanto sarebbe stata importante una sua presenza, soprattutto per affrontare le difficoltà molto forti che hanno poi determinato una vera e propria crisi della Documentazione.

4. La crisi della documentazione nel ventunesimo secolo

La diminuzione dell'importanza dei documentalisti negli ultimi venti anni può essere attribuita a diversi fattori il primo e il più rilevante dei quali è stata

sicuramente l'innovazione tecnologica. La crisi della documentazione come disciplina e professione è stata determinata in primo luogo da come si sono evolute le tecnologie del trattamento automatico delle informazioni che per un lungo periodo erano state individuate come il fattore distintivo della documentazione stessa rispetto alla biblioteconomia "classica".

L'avvento di Internet ha trasformato le modalità di accesso alle informazioni, in primo luogo della documentazione scientifica. Questo cambiamento ha portato a una minore dipendenza dai metodi tradizionali di gestione dei documenti e a una maggiore automazione dei compiti precedentemente svolti dai documentalisti.

La distribuzione degli strumenti di calcolo negli ambienti di lavoro e la loro connessione in rete ha spinto le organizzazioni – pubbliche e private – a ricercare una sempre maggiore efficienza e una diminuzione dei costi grazie all'adozione di processi e algoritmi automatizzati per l'organizzazione delle informazioni.

L'indirizzo preso dall'innovazione tecnologica ha comportato inoltre profondi cambiamenti in tutte le professioni legate al trattamento delle informazioni, facendo sì che il ruolo dei documentalisti si sia diversificato o fuso con altri ruoli, come gli scienziati dell'informazione, gli analisti di dati e i manager della conoscenza. Ciò ha portato a un'attenuazione dei confini tradizionali e a uno spostamento dell'attenzione verso pratiche più ampie di gestione delle informazioni.

Il susseguirsi di crisi economiche seguite da periodi di sviluppo intenso hanno accresciuto le pressioni economiche e i vincoli di bilancio nel settore pubblico e in quello privato portando a una riduzione dei finanziamenti per i servizi bibliotecari e archivistici tradizionali, con conseguente ridimensionamento o ristrutturazione delle posizioni di documentalista trasferendo sui ricercatori e sul personale in genere l'onere di compiere direttamente le ricerche informative.

Complessivamente questi processi hanno fatto sì che le competenze richieste nelle professioni dell'informazione si siano fortemente modificate, con una crescente richiesta di conoscenze informatiche e per l'analisi e la gestione dei dati. I documentalisti in gran parte erano privi di queste competenze, mentre la disciplina Documentazione rispetto a queste esigenze si è dovuta confrontare con altre discipline meglio attrezzate per definirne i termini di impiego.

Nonostante questi fattori abbiano determinato la pratica scomparsa, almeno in Italia, della Documentazione sia come disciplina che come professione, è necessario ricordare che i principi fondamentali elaborati nel suo ambito ed in particolare della Documentazione italiana - organizzare, conservare e facilitare l'accesso alle informazioni - rimangono essenziali, anche se in forme evolutive nel contesto più ampio della gestione delle informazioni e delle biblioteche digitali come è stato notato in Diozzi (2017) e in Folino e Pasceri (2017).

5. Conservare la memoria

Viviamo, non è una novità, sotto la maledizione dell'eterno presente per cui, troppo spesso, l'accelerazione che la società contemporanea ha acquisito come sua caratteristica principale tende a stravolgere gli equilibri del consorzio umano. Non solo nell'attualità, ma anche nella politica e nell'economia, ci si trova di fronte ad una tendenza semplificatrice che banalizza l'importanza dei fenomeni, pretende spesso di offrire soluzioni a problemi complessi con ricette che oscillano tra l'ovvio e l'irrealizzabile.

I diversi campi del sapere e le buone pratiche professionali sono anche essi minacciati da questo spirito del tempo.

I primi sono obbligati a rispondere alle "esigenze" presentate e dimostrare così la propria "produttività" ed un esempio per tutti è certamente rintracciabile nel rapporto di "amore-odio" instauratosi, durante la pandemia da Covid 19, tra una parte dell'opinione pubblica e la comunità scientifica di riferimento. Malgrado questa si sia mossa complessivamente bene non ha avuto un adeguato riconoscimento e sconta anzi l'insensibilità ai messaggi operativi più volte trasmessi. Messaggi che "hanno il torto" di mettere in discussione le abitudini correnti e obbligano a pensare al medio periodo.

Le seconde sono destinate a cicli di vita sempre più brevi dal momento in cui le organizzazioni, private o pubbliche, sono condizionate da spasmodiche esigenze di "ottimizzazione".

Forse anche la pratica professionale della Documentazione ha scontato questo clima, anche se le sue basi teoriche, così come si stavano definendo pochi decenni fa, erano tutt'altro che deboli come opportunamente ricorda Castellucci (2003, 68-69) in questo passaggio su Bisogno:

Dal punto di vista semplicemente temporale *Teoria della documentazione* potrebbe sembrare un titolo perfino "tardivo" rispetto al trattato di Otlet che lo precede di più di quarant'anni; e pertanto potrebbe apparire solo un aggiornamento del lavoro di Otlet, magari riferito allo specifico contesto italiano e a un mutato panorama tecnologico. Ma in realtà non è così e, appunto, *Teoria della documentazione* rivela un'argomentazione precoce, perché Bisogno utilizza il termine "documentazione" per dire molto altro ancora: non solo per descrivere una nuova espressione professionale, ma per affermare una nuova modalità di categorizzazione delle conoscenze e per la diffusione delle informazioni quale vero e proprio sistema arterioso che rende viva l'intera società. Bisogno è consapevole che ormai il trattamento dell'informazione è l'elemento che caratterizza la società contemporanea. "Informazione come risorsa", è l'espressione che troviamo sempre nei suoi interventi ad esempio ai vari convegni AIDA. Bisogno è cioè testimone di un ruolo assunto dall'informazione all'interno della società che Otlet non poteva nemmeno immaginare. Bisogno è autore che descrive la postmodernità, la società che si sta trasformando da "industriale" a società "dell'informazione". Ma ancora, nel 1980, questo processo non si è

asestato e davvero arduo appare il compito di definirlo teoricamente.

Quel lavoro teorico non è andato disperso e, crediamo di poter dire, assume un valore come metodo di per sé valido e applicabile in diversi ambiti: la ridefinizione delle basi concettuali di una disciplina in presenza di fenomeni epocali è quello che Paolo Bisogno stava facendo con la Documentazione tra gli anni ottanta e novanta del secolo passato. Conservarne la memoria dal punto di vista scientifico è un elemento importante non solo per coloro che si occupano di informazione ma per tutti quelli che intendono mantenere un adeguato equilibrio tra corredo disciplinare di riferimento e definizione di pratiche professionali e applicative.

Riferimenti bibliografici

- Bisogno, Paolo. 1979. *Teoria della documentazione*. Milano: Franco Angeli.
- Bisogno, Paolo. 1985. "La politica della Documentazione in Italia." In *Le tecnologie della Documentazione nella ricerca, nei Servizi, nelle Professioni. Atti del I convegno nazionale AIDA*, a cura di Maria Pia Carosella e Marta Giorgi, 15-26. Roma.
- Carosella, Maria Pia, e Maria Valenti, a cura di. 1983. *Documentazione e biblioteconomia. Manuale per i servizi di informazione e le biblioteche speciali italiane*. Milano: Franco Angeli.
- Castellucci, Paola. 2003. "Paolo Bisogno: La precoce fondazione della disciplina della Documentazione." *AIDAinformazioni*, anno 2, no. 3 (luglio-settembre): 59-70.
- Castellucci, Paola. 2007a. "Eredità e prospettive per la documentazione in Italia." *Bollettino AIB* 47 (3): 239-55.
- Castellucci, Paola. 2007b. "Documentazione." In *Biblioteconomia. Guida classificata*, diretto da Mauro Guerrini, 186-89. Milano: Editrice Bibliografica.
- Diozzi, Ferruccio. 2017. "La cultura della Documentazione: quindici anni di AIDA in Italia." *AIDAinformazioni*, anno 35, no. 3-4 (luglio-dicembre): 47-59.
- Folino, Antonietta, e Erika Pasceri. 2017. *L'Associazione Italiana per la Documentazione Avanzata. Storia e archivio*. Canterano: Aracne editrice.
- Fukuyama, Francis. 1992. *La fine della storia e l'ultimo uomo*. Milano: Rizzoli.
- Petruciani, Alberto. 2006. "Bisogno Paolo." In *Dizionario bio-bibliografico dei bibliotecari italiani del XX secolo*. <https://www.aib.it/aib/editorial/dbbi20/bisogno.htm>.

Domenico (Ingo) Bogliolo

Profilo di un navigatore nell'*Information Science*

Paola Gargiulo*, Lucia Maffei**

Tratteggiare il profilo di Domenico (Ingo) Bogliolo, significa disporsi a farsi accompagnare nella vicenda professionale di questa figura insolita e determinante per la vita dell'Associazione Italiana Documentazione Avanzata (AIDA), della rivista «AIDAinformazioni» e più in generale del mondo professionale che si riconosceva nella riflessione e nella prassi delle professioni bibliotecarie e documentarie. Seguiremo il suo cammino attraverso i mutamenti della professione di documentalista che lui, come vedremo, aveva la capacità di intuire prima di altri. Era un “visionario concreto”, un ossimoro per dire che, grazie alla sua curiosità e profonda cultura che spaziavano ambedue ben oltre i limiti professionali, nel tempo ha colto i segni dei mutamenti e si è impegnato a costruire strumenti e riflessioni che aiutassero la professione di documentalista ad attecchire in Italia e ad evolvere nella teoria e nella prassi.

Seguiremo la sua biografia che è stata simile a quella di tanti bibliotecari “speciali” come si chiamavano negli anni Settanta coloro che nella professione operavano in ambienti di ricerca pubblici e privati, affiancando spesso gli studiosi con il supporto delle proprie capacità di organizzare e diffondere l'informazione.

Era nato ad Alassio il 3 ottobre 1945 e dalla sua origine aveva conservato certi caratteri riservati di chi vive stretto fra il mare e la montagna e l'amore per i prodotti e la cucina della sua terra.

Dopo la laurea in Scienze politiche presso l'Università di Genova nel 1975, si trasferì a Roma ed entrò nel mondo delle biblioteche e dell'informazione nel 1977, iniziando una carriera professionale più che trentennale, caratterizzata da grande vivacità intellettuale, intelligenza critica e spirito visionario e dal-

* Già socia AIDA (Associazione Italiana Documentazione Avanzata) e già tecnologa presso CASPUR (Consorzio interuniversitario per le Applicazioni di Supercalcolo). paola.gargiulo@gmail.com.

** Già Presidente AIDA. luciamaffei52@gmail.com.

la capacità di coniugare biblioteconomia e *information science* con approcci e prospettive provenienti da altri ambiti, creando insolite e originali connessioni.

Nel 1977 vinse il concorso pubblico per funzionari di biblioteca dell'Università, ed assunse la direzione della Biblioteca Centrale dell'Istituto Navale di Napoli, oggi Università Parthenope, dove si impegnò con leggerezza e giocosità alla trasformazione della biblioteca in una struttura efficiente di servizio a supporto alla didattica e alla ricerca. Si distinse per un approccio manageriale efficace, antiburocratico e fortemente orientato alla risoluzione dei problemi. Nel 1980 venne assegnato alla direzione della Biblioteca dell'Istituto di Matematica dell'Università degli Studi di Roma "La Sapienza" che guidò fino al 1993. La sua idea guida è stata coniugare la centralità e la soddisfazione dei bisogni degli utenti con la gestione efficiente della struttura. Ha diretto il suo impegno in varie direzioni: valutare l'efficacia dei servizi, applicare le emergenti tecnologie dell'informazione, promuovere la formazione e l'auto-formazione dei bibliotecari, valorizzare le competenze acquisite.

Dal 1983 al 1988 è stato responsabile del coordinamento della segreteria del progetto di cooperazione e di informatizzazione bibliografica tra le biblioteche delle università del Lazio, COBBUL - Cooperazione Biblioteche Università del Lazio. Il progetto, nato dal basso, innovativo, pionieristico, era finalizzato alla creazione di un sistema bibliotecario universitario su base regionale. In questi stessi anni partecipò a progetti finanziati dal Consiglio Nazionale delle Ricerche (CNR) per la creazione di infrastrutture e sistemi informativi per le biblioteche di matematica ed è stato tra i co-fondatori del progetto *SINM-Sistema informativo Nazionale della Matematica* (1980-2010).

Gli anni tra il 1993 e il 2002 sono stati i più fecondi per sua attività professionale. Ha coordinato presso il CICS - Centro Interdipartimentale per il Calcolo Scientifico dell'Università "La Sapienza" di Roma, in qualità di web master, la creazione e la gestione dei siti web di diversi dipartimenti della Sapienza. In quegli anni, ha inoltre progettato, realizzato e curato diversi siti web, tra cui il primo al mondo, dedicato a Ranganathan nel 1995, il sito di INFER, l'osservatorio sul mercato editoriale delle risorse informative elettroniche, sulle acquisizioni consortili e le licenze di accesso a periodici elettronici ed e-book nel 1999, il VRD (Virtual Reference Desk) della Sapienza. Socio di AIB (Associazione Italiana Biblioteche), AIDA e Ass.I.Term (Associazione Italiana per la Terminologia). Nell'AIB ha presieduto la prima Commissione Nazionale dell'Università e Ricerca AIB-CNUR (1987-1988) occupandosi di profili professionali, sistemi bibliotecari ed automazione dei servizi bibliotecari; partecipa nel corso degli anni successivi a convegni e seminari AIB con interventi su tematiche diverse: dall'architettura dei sistemi informativi, all'automazione, ai profili professionali, e al *Knowledge Management (KM)*.

Stesso fervore di iniziative ha riversato su AIDA e la frequenza dei suoi scritti su «AIDAinformazioni» ce lo testimonia anche oggi. Tali contributi ci

aiutano a snodare il filo dei suoi interessi professionali nel corso di questi anni. Nell'Associazione la sua figura travalicava i confini del suo ruolo formale di segretario-tesoriere (dalla fondazione 1982 al 2013) per assumere quella di punto di riferimento costante, al di là dell'alternarsi dei membri del Consiglio Direttivo e dei Presidenti che hanno sempre trovato un ancoraggio solido in Bogliolo. Dai primi scritti in cui si appassionava ai temi di come strutturare i servizi bibliotecari intorno ai bisogni dell'utente, sperimentando progetti nella biblioteca di Matematica prima e poi come webmaster dell'Università della Sapienza di Roma, ha voluto sempre con ostinazione affiancare alla riflessione teorica la creazione di strumenti che organizzassero e facilitassero l'accesso all'informazione.

Nel 1998 cominciò a seguire uno degli argomenti che più lo hanno appassionato: il *Knowledge Management*. Data la sua storia professionale non poteva non rimanere affascinato da un argomento che in quegli anni attrasse molte riflessioni. In un mondo del lavoro in piena evoluzione tecnologica, organizzare la conoscenza come una vera risorsa strategica per le imprese era vitale. Dedicò molti anni della sua riflessione a questo tema riversandola su «AIDAinformazioni», formandosi sui testi fondamentali disponibili all'epoca, a cominciare da uno dei più importanti teorici: Laurence Prusak. Coerente con i suoi interessi divise il suo saggio su «AIDAinformazioni» pubblicato nel 1998 in tre distinti articoli in cui l'autore affrontava diversi aspetti del tema: un'introduzione al *KM* (Bogliolo 1998a, 18), i risvolti attinenti alla terminologia (Bogliolo 1998b, 14), e il percorso dall'*Information Retrieval* al *KM* e i possibili mutamenti nella natura delle professioni collegate (Bogliolo 1998c, 16). Quest'ultimo aspetto sarà poi approfondito dall'autore in un ulteriore intervento (Bogliolo 1999, 8). Il suo interesse per l'argomento *KM* si prolungò negli anni, almeno fino a tutto il 2001 e con la rubrica *Schegge* su «AIDAinformazioni»: analizzava casi di studio, discussioni teoriche e sul piano applicativo che si svolgevano prevalentemente su liste di discussione internazionali dedicate. È sintomatico, per capire il suo modo di agganciare sempre la teoria con le applicazioni nella professione, la puntata di *Schegge* che dedicò al *KM* – *Appunti, 4: Anti-terrorism KM task force* (Bogliolo 2001, 4) all'indomani dell'attentato alle Torri gemelle a New York, il suo interesse professionale era seguire e riportare le discussioni che si accendevano fra i professionisti di tutto il mondo, sull'opportunità di creare una *task force* di esperti del *KM*, nella consapevolezza che le guerre sempre più sarebbero state vinte non tanto da un armamento superiore quanto da una conoscenza superiore.

Lo scoppio della guerra in Iraq fu la tragica occasione di dare concretezza, almeno in parte, a quelle premesse. Il portale *AIDAPace*, da lui proposto e accettato con tanto entusiasmo da tutto il Consiglio direttivo, era un esempio di rispetto della tradizione pacifista, insita nella disciplina della documentazione che trovava in Paul Otlet uno dei suoi capisaldi e insieme un tentativo generoso

di contribuire con gli strumenti della professione all'informazione su tutto ciò che si muoveva nel vasto mondo del pacifismo e della promozione della pace. Il sito, creato e gestito da Bogliolo, si definiva una risorsa documentaria per la coscienza critica e si articolava in varie sezioni: quotidiani d'informazione, notiziari e basi di dati, siti ufficiali, centri di ricerca e documentazione, petizioni e raccolta di fondi, voci di donne, iniziative contro la guerra. Crediamo sia stato un *unicum* nel panorama informativo italiano: un sito non commerciale che provava ad applicare criteri professionali di gestione della conoscenza al settore dell'informazione di attualità e di approfondimento.

Bogliolo in quegli stessi anni, ha creato e gestito il sito web di AIDA arricchendolo tra il 2000 e 2010 di diverse sezioni, *AIDANews*, *AIDAJobs*, per offrire opportunità di incontro fra domanda e offerta di lavoro per le professioni legate alle scienze dell'informazione, *AIDALampi*, supplemento elettronico di «AIDAinformazioni» che raccoglieva notizie brevi e aggiornate sul mondo della documentazione avanzata, *AIDACornici* che aggregava materiale informativo e documentale in costante aggiornamento su temi quali Ambiente, Pace, Servizi di Reference e infine *AIDALavori in corso*, Bibliografie sul *Knowledge Management* e i servizi di *reference* online. Bogliolo fu anche responsabile dell'edizione elettronica della rivista «AIDAinformazioni» dal 2000 al 2013.

Il sito web di AIDA poi, arricchito delle diverse sezioni che, come abbiamo visto, intendevano raggiungere chi voleva avvicinarsi alla professione, contribuirono non poco alla crescita dell'Associazione nei primi anni 2000. Con la crescita degli iscritti e la collaborazione con associazioni europee dei professionisti dell'informazione, si affermò in AIDA la convinzione che gli interessi che si erano già manifestati in passato per percorsi formativi e di certificazione della professione a livello europeo, potessero concretizzarsi, grazie ai rapporti di antica data e robuste radici che AIDA coltivava con associazioni di professionisti europei collaborazioni e scambio con le altre associazioni esistenti (IAML-Italia, GDIF-RBM, AIDB)¹ e con il *chapter* italiano di ISKO, International Society for Knowledge Organization, in cui Bogliolo fu attivo tramite.

È lo stesso Bogliolo che di questo processo fu appassionato protagonista con Maria Pia Carosella e Ferruccio Diozzi che ci racconta in un suo articolo su «AIDAinformazioni» del 2005:

Il primo incontro formale di AIDA con il movimento della certificazione professionale data dal 24 ottobre 1995, quando l'*European Council of Information Associations (ECIA)*, ospite di AIDA per la sua assemblea semestrale, sottoscrisse la "Dichiarazione di Roma" per l'impegno interassociativo a instaurare dei sistemi di certificazione, facilitarne il mutuo riconoscimento e contribuire alla loro compatibilità nel contesto europeo. Il testo della "Dichiarazione di

¹ Associazione Italiana delle Biblioteche, Archivi e Centri di documentazione musicali; Gruppo Italiano dei Documentalisti dell'Industria Farmaceutica e della Ricerca Biomedica; Associazione Italiana Documentalisti Brevettuali.

Roma” è pubblicato sul fascicolo invernale di *AIDAinformazioni* di quell’anno. Di conseguenza, nel 1997 viene creato il progetto *DECIDoc* [*Développer les Eurocompétences en Information et Documentation*], finanziato dall’Unione Europea nel quadro del programma “Leonardo da Vinci”, per un inventario delle competenze europee del settore, come sviluppo, su un’area regionale più ampia, del lavoro che già da tre anni l’associazione francese e quella spagnola proseguivano per raccogliere esempi di profili professionali presenti nei rispettivi Paesi. L’esito del *DECIDoc* è costituito dall’*Euroréférentiel I&D* (1999), che AIDA ha versato in italiano e proposto alla riunione semestrale italiana dell’ECIA del 4 ottobre 1999 a Pontignano. Questa, pubblicata da Casalini libri nel 2000, è la prima edizione italiana della Euroguida I&D, che AIDA ha presentato nel corso di un apposito seminario al Bibliocom2000 (Bogliolo 2005, 70).

La novità assoluta consisteva nella volontà di utilizzare l’*Euroguida I&D* come base per la successiva evoluzione del progetto di certificazione professionale e la sua innovativa impostazione, diversa da altre iniziative maggiormente orientate verso strutture di tipo ordinistico o da Albo professionale. Il modello di certificato europeo si ispirava infatti, ad alcuni capisaldi che possiamo così sintetizzare: indipendenza dal titolo di studio, indipendenza dall’iscrizione a un’associazione professionale, dipendenza dalle esperienze acquisite, dipendenza dalla formazione ricevuta.

Un’impostazione così inusuale in Italia, che puntava a certificare le competenze e non i percorsi formativi, non poteva non interessare ed entusiasmare Bogliolo che infatti si dedicò prima con Maria Pia Carosella alla cura della prima edizione ufficiale italiana (l’originale era stato pubblicato dall’ECIA) dell’*EuroGuida I&D. Competenze dei professionisti europei dell’informazione e della documentazione* nel 2000 (Bogliolo e Carosella 2000), curandone poi, come abbiamo visto, anche le edizioni successive, ma impegnandosi anche con Ferruccio Diozzi e Augusta Franco a concretizzare il percorso certificativo in Italia.

L’ultimo interesse che vogliamo evidenziare nel profilo professionale di Bogliolo è quello relativo alla terminologia, intesa come scienza della definizione dei linguaggi settoriali. È un interesse che si è manifestato in lui già dagli anni novanta e lo ha accompagnato anche quando ha affrontato nel tempo le diverse tematiche di interesse, dal *KM* ai processi di definizione delle varie professioni legate al mondo del trattamento delle informazioni, fino alla creazione e organizzazione dei flussi documentali amministrativi. Fu infatti nel 2002 che fu chiamato dall’Amministrazione centrale dell’Università degli Studi “la Sapienza” a guidare il progetto per la creazione del protocollo elettronico di cui si è occupato fino al suo pensionamento nel novembre del 2010. Fu un impegno, con cui si cimentò con la consueta passione e nel quale riversò tutto il bagaglio professionale accumulato in tanti anni di navigazione nell’*Infor-*

mation Science. Si può dire che questa parte della sua carriera diede a lui e a tutti i documentalisti la prova che, quando affermavano che le competenze variegata, da quelle manageriali, a quelle tassonomiche, a quelle biblioteconomiche, potevano essere la base di nuove professioni, avevano visto giusto. Non furono infatti pochi i documentalisti, i bibliotecari e anche gli archivisti chiamati dalle amministrazioni pubbliche a impiantare i protocolli elettronici che, per loro natura, prevedevano una completa revisione delle procedure e del flusso documentale.

Bogliolo ha partecipato a convegni nazionali, seminari, workshop e corsi di formazione in Italia e all'estero, curatore di atti di convegno e seminari di AIDA, di CIBER (Coordinamento Interuniversitario Basi dati ed Editoria in Rete). Autore di numerosi articoli pubblicati su «AIDAinformazioni», «Bollettino AIB», «AIB Notizie», «Bibliotime», «Il Bibliotecario». La vastità degli argomenti trattati con arguzia, espressi con chiarezza e semplicità pure nella loro complessità, attesta la sua versatilità, una cultura multidisciplinare, un'apertura al nuovo, una curiosità illimitata accompagnata da senso di humour, leggerezza e dalla capacità di creare inesplorate connessioni.

Negli anni tra il 2008 e il 2011 tiene una rubrica *De nuptiis Mercurii et philologiae* sulla terza serie della rivista «Il Bibliotecario» dove prende spunto da notizie, articoli, da libri scandagliati in rete per fulminee, argute, colte riflessioni sul rapporto uomo-macchina, info-scientometria, infologia, gli impatti di Internet sulla lettura, l'evoluzione del web 2.0, il cloud computing, l'interoperabilità ed altro ancora.

La passione per la scrittura lo portò ad essere autore di racconti, di cui il primo è stato pubblicato in *Umor di libro*, l'antologia dei migliori racconti del premio letterario per inediti umoristici sul mondo bibliotecario, bandito nel 1991 dall'AIB – Sezione Emilia-Romagna. La scrittura di racconti, fantasiosi e immaginifici ha accompagnato come una presenza costante l'ultima parte della sua vita, conclusasi prematuramente il 2 luglio 2017. Condivideva e sollecitava riflessioni e reazioni da parte dei lettori: era il suo modo, crediamo, di continuare ad esplorare sé stesso, le relazioni con gli altri e con il mondo che ci circonda.

Speriamo, con questo breve e incompleto profilo, di aver dato il senso di una vita professionale segnata dalla curiosità, dall'entusiasmo per il nuovo, ma anche dalla capacità di inserirlo in categorie mentali rigorose, per capirne le potenzialità e i rischi, dalla rara dote di riportare contaminazioni ideali e frequentazioni senza pregiudizi, al risultato di proporre soluzioni innovative, praticarle e diffonderle.

Riferimenti bibliografici

- Bogliolo, Domenico. 1998a. "KM, Knowledge Management. 1/3." *AIDAinformazioni*, anno 16, no. 2 (aprile-giugno): 18-22.
- Bogliolo, Domenico. 1998b. "KM, Knowledge Management. 2/3." *AIDAinformazioni*, anno 16, no. 3 (luglio-settembre): 8-14.
- Bogliolo, Domenico. 1998c. "KM, Knowledge Management. 3/3." *AIDAinformazioni*, anno 16, no. 4 (ottobre-dicembre): 16-24.
- Bogliolo, Domenico. 1999. "KM, Knowledge Management. Aggiunta." *AIDAinformazioni*, anno 17, no. 1 (gennaio-marzo): 8-15.
- Bogliolo, Domenico, e Maria Pia Carosella. a cura di. 2000. *EuroGuida I&D. Competenze dei professionisti europei dell'informazione e della documentazione*. Fiesole: Casalini libri.
- Bogliolo, Domenico. 2001. "KM-Appunti. 4. Anti-terrorism KM task force." *AIDAinformazioni*, anno 19, no. 4 (ottobre-dicembre): 36-41.
- Bogliolo, Domenico. 2005. "Certificazione degli specialisti dell'informazione: opinioni ed esperienze dall'Europa - Le novità dell'Euroguida I&D, seconda edizione." *AIDAinformazioni*, anno 23, no. 4 (ottobre-dicembre): 70-78.

Bibliografia

Si acclude una bibliografia molto selettiva delle oltre cento pubblicazioni di Domenico (Ingo) Bogliolo, basata sui contributi, in ordine cronologico, più rappresentativi delle tematiche, degli ambiti di studio e di interessi maturati nel corso della sua vita professionale.

- Bogliolo, Domenico. 1983. "Finanziamento, produttività, progetti di ricerca: due sistemi informativi per la matematica." *Bollettino d'informazioni AIB* 23(3): 304-07.
- Bogliolo, Domenico. 1983. "Il vigente regolamento delle biblioteche delle università." *Bollettino d'informazioni AIB* 23(3): 73-282.
- Bogliolo, Domenico. 1983. "La valutazione d'efficacia e le biblioteche speciali." Con la collaborazione del Sottogruppo per la valutazione delle biblioteche speciali dell'AIB-Lazio. *Bollettino d'informazioni AIB* 23(1): 5-19.
- Bogliolo, Domenico, Giulia Corvino, Anna Maria Tammaro, e Lucilla Vespucci. 1984. "Progetto COBBUL: un esempio di cooperazione tra biblioteche universitarie." *Biblioteche oggi* 2(4): 89-96.

- Bogliolo, Domenico. 1985. "COBBUL: a real environment experience on automated versus manual catalogue usage education." In *User education in the online age II. IATUL proceedings, 30th July-2nd August 1984, Delft, The Netherlands*, vol. 17, edited by Nancy Fjällbrant, 81-89.
- Bogliolo, Domenico. 1987. "Le biblioteche tra geografia e articolazione disciplinare: l'architettura del sistema." *Università Progetto* 25: 27-35.
- Bogliolo, Domenico. a cura di. 1989. "Questo "speciale". Documenti e repertori. La Commissione nazionale Università ricerca: sintesi dei resoconti di un anno di lavoro 1989." *Bollettino d'informazioni AIB* 29(2-3): 167- 415.
- Bogliolo, Domenico. 1991. "Le attività scientifiche del bibliotecario fra "Università ricerca" e "Informazione e documentazione." In *AIB. Commissione nazionale Informazione e documentazione: atti del seminario della Commissione nazionale "Informazione e documentazione", Roma, 4 giugno 1990*, a cura di Vilma Alberani. Roma: AIB.
- Bogliolo Domenico. 1996. "Profilo sperimentale del documentalista nell'università italiana." In *Documentazione: professione trasversale. Atti del 5° Convegno Nazionale AIDA. Fermo, Palazzo dei Priori, 23-25 ottobre 1996*, a cura di Carla Basili, 247-53. Roma: AIDA; CNR-ISRDS.
- Bogliolo, Domenico. 1998. "KM, Knowledge Management. 1/3." *AIDAinformazioni*, anno 16, no. 2 (aprile-giugno): 18-22.
- Bogliolo, Domenico. 1998. "KM, Knowledge Management. 2/3." *AIDAinformazioni*, anno 16, no. 3 (luglio-settembre): 8-14.
- Bogliolo, Domenico. 1998. "KM, Knowledge Management. 3/3." *AIDAinformazioni*, anno 16, no. 4 (ottobre-dicembre): 16-24.
- Bogliolo, Domenico. 1999. "KM, Knowledge Management. Aggiunta." *AIDAinformazioni*, anno 17, no. 1 (gennaio-marzo): 8-15.
- Carosella, Maria Pia, e Domenico Bogliolo. a cura di. 2000. *Euroguida I&D dei professionisti europei dell'informazione e della documentazione*. Fiesole: Casalini libri.
- Bogliolo, Domenico. 2000. "Libri, editori e pubblico nell'era di Internet." *Bibliotime*, anno 3, no. 2 (luglio). <https://www.aib.it/aib/sezioni/emr/bibtme/num-iii-2/bogliolo.htm>.
- Bogliolo, Domenico. 2000. "Lo spazio "non fisico" della biblioteca." *Bibliotime*, anno 3, no. 3 (novembre). <https://www.aib.it/aib/sezioni/emr/bibtme/num-iii-3/bogliolo.htm>.
- Bogliolo, Domenico. 2000. "KM - Appunti. 1: Knowledge vs Information." *AIDAinformazioni*, anno 18, no. 1 (gennaio-marzo): 24-28. <http://eprints.rclis.org/4362/2/2000-1Schegge.pdf>.

- Bogliolo, Domenico. 2000. "KM - Appunti. 2: Ontologie." *AIDAinformazioni*, anno 18, no. 2 (aprile-giugno): 13-16.
- Bogliolo, Domenico. 2000. "KM - Appunti. 3: DBMS vs KBMS." *AIDAinformazioni*, anno 18, no. 3-4 (luglio-dicembre): 26-29.
- Bogliolo, Domenico. 2001. "KM - Appunti. 4: Anti- terrorism KM task force." *AIDAinformazioni*, anno 19, no. 4 (ottobre-dicembre): 36-41. <http://eprints.rclis.org/4726/2/2001-4Scegge.pdf>.
- Bogliolo, Domenico. 2001. "Dal reference al KM: il caso Lisa Guedea Carreño." *Bibliotime*, anno 4 no. 1 (marzo). <https://www.aib.it/aib/sezioni/emr/bibtime/num-iv-1/bogliolo.htm>.
- Bogliolo, Domenico. 2002. "KM - Appunti. 5: KM e "nuova logica." *AIDAinformazioni*, anno 20, no. 2/3 (aprile-settembre): 101-12. <http://eprints.rclis.org/4991/2/2002-23Scegge.pdf>.
- Bogliolo, Domenico. 2003. "L'informazione di guerra al femminile." *AIDAinformazioni*, anno 21, no. 1 (gennaio-marzo): 267-72.
- Bogliolo, Domenico. 2003. "KM - Appunti. 6. Il Ba." *AIDAinformazioni*, anno 21, no. 4 (ottobre-dicembre): 33-37.
- Bogliolo, Domenico. a cura di. 2005. *Euroguida I&D*, seconda edizione. Fiesole: Casalini Libri.
- Bogliolo, Domenico. 2005. "Certificazione degli specialisti dell'informazione: opinioni ed esperienze dall'Europa - Le novità dell'Euroguida I&D, seconda edizione." *AIDAinformazioni*, anno 23, no. 4 (ottobre-dicembre): 70-78.
- Bogliolo, Domenico. 2008. "Linee di problema per il trattamento terminologico di documenti amministrativi elettronici." In *I-TerAnDo: "Terminologia, analisi testuale e documentazione nella città digitale": Atti del convegno Ass.I.Term, Università della Calabria, 5-7 giugno 2008*. *AIDAinformazioni*, anno 26, no. 1-2: 143-47.
- Bogliolo, Domenico. 2008. "De nuptiis Mercurii et philologiae: Knowware. Liquefazione. Scientografia. Turn. Web librarianship." *Il bibliotecario*, III serie, no. 3 (settembre-dicembre): 135-43.
- Bogliolo, Domenico. 2009. "De nuptiis Mercurii et philologiae: Comunità di prassi. E-ISSN. Esposizione/conoscenza. Informologia. Social software. Webliografia. Webologia. Webology." *Il bibliotecario*, III serie, no. 1 (gennaio-aprile): 117-27.
- Bogliolo, Domenico. 2009. "De nuptiis Mercurii et philologiae; Archivist d'assalto. Catene cognitive. Rappresentazione. Information architecture." *Il bibliotecario*, III serie, no. 2-3 (maggio-dicembre): 177-83.

- Bogliolo, Domenico. 2010. "De nuptiis Mercurii et Philologiae: Biblioterapia. Crisi. Falso documentale. Informazionismo. Mapping." *Il bibliotecario*, III serie, no. 1-2 (gennaio-agosto): 243-49.
- Bogliolo, Domenico. 2010. "De nuptiis Mercurii et Philologiae: Cinecarta. Diogene. Informivori. Inter-operabilità. Istruzione digitale. Muraglie." *Il bibliotecario*, III serie, no. 3 (settembre-dicembre): 159-66.
- Bogliolo, Domenico. 2011. "De nuptiis Mercurii et Philologiae: Culturomica. Griglie e nuvole (e folla). Hita-hita. Peer Web. Ovo Qwiki. Vook. Zotero e gli altri." *Il bibliotecario*, III serie, no. 1-2 (gennaio-agosto): 195-200.
- Bogliolo, Domenico. 2015. "Lo Zen e l'arte della manutenzione del Knowledge Management". In *Noetica versus informatica: le nuove strutture della comunicazione scientifica: atti del Convegno internazionale, Roma, Tempio di Adriano, 19-20 novembre 2013*, a cura di Fiammetta Sabba, 59-70. Firenze: Olschki.