

62. CONGRESSO NAZIONALE AIB

BIBLIOTECHE E TECNOLOGIE AL TEMPO DELL'INTELLIGENZA ARTIFICIALE



FIRENZE 16-17 NOVEMBRE 2023
BIBLIOTECA NAZIONALE CENTRALE DI FIRENZE



ABSTRACT

Giovedì 16 novembre

Intelligenza artificiale in biblioteca

Gino Roncaglia, Università di Roma Tre

L'intervento proporrà un breve itinerario che dall'intelligenza artificiale delle origini porta fino alle intelligenze artificiali generative di oggi, illustrandone sinteticamente le modalità di funzionamento. Nella seconda parte, offrirà poi una riflessione su alcuni fra gli effetti che la diffusione di sistemi di intelligenza artificiale generativa potranno avere sulle biblioteche, soffermandosi in particolare sulla possibilità di una collaborazione fra il modello architettonico tradizionalmente alla base del lavoro di organizzazione del sapere e il modello 'oracolare' proposto dalle reti neurali generative. Infine, si soffermerà sul tema dell'AI literacy, componente sempre più rilevante dell'information literacy, su cui anche le biblioteche saranno chiamate a lavorare.

Artificial Intelligence (AI) in libraries: developing a strategic response?

Andrew Cox, University of Sheffield - IFLA AI Special Group

AI has been named word of the year 2023 by the Collins (English) Dictionary, this reflects that the subject has been very much in the news both for its risks and benefits. Not surprisingly, as a profession, we have mixed feelings about AI. The presentation discusses a number of definitions of AI but I emphasize its "familiar face" in functions such as spam filtering, transcription, recommendation and search. AI has many potential applications in libraries, with descriptive AI to help increase access to collections being probably the most important. ChatGPT and other generative AI has many uses too but also illustrate all the problems of AI when it is built by BigTech companies: in terms of bias, inaccuracy, lack of explainability, use of copyright material, digital divides, exploitative use of labour and environmental impacts.

Nevertheless, AI is a strategic priority and librarians have to adopt a constructive view towards it as a profession. The talk offers a SWOT analysis of libraries and AI and suggests some important strategic options for libraries (summarised from a Jisc report I have written recently). I examine the nature of AI capability and from that suggest a number of key roles for the profession in terms of developing exemplary examples of ethical, responsible AI; contributing data skills; and through the promotion of AI literacy among users.

Quando l'informazione è sociale dentro e fuori le biblioteche

Walter Quattrocchi, Data Science and Complexity, Department of Computer Science, Sapienza, Università di Roma

Esistono le echo chamber sui social media? Concentrandoci su come gli utenti Facebook italiani e statunitensi si relazionano a due narrazioni distinte (che coinvolgono teorie del complotto e scienza), troviamo forti evidenze di polarizzazione. Gli utenti tendono a promuovere le loro storie preferite e quindi a formare gruppi polarizzati. Il pregiudizio di conferma domina il processo di diffusione creando così cascate informative all'interno di comunità ben definite. Allo stesso tempo, l'aggregazione di informazioni appositamente selezionate per coerenza narrativa all'interno di quelle comunità rafforza l'esposizione selettiva e la polarizzazione e quindi segregazione dei gruppi. Forniamo prove empiriche che gli utenti tendono ad assimilare solo informazioni che confermano la loro visione del mondo e a ignorare informazioni a contrasto. La pandemia COVID-19 è stata la tempesta perfetta per questo fenomeno e l'OMS ha coniato il termine infodemics per riferirsi alla sovrabbondanza di informazioni. Abbiamo esplorato questi processi durante la fase iniziale delle pandemie scoprendo che informazioni affidabili e discutibili si diffondevano in modo simile. Concludiamo la presentazione mostrando come diverse piattaforme di social media (Facebook, Twitter, Gab e Reddit) suscitano dinamiche di polarizzazione molto diverse.

Tecnologie & biblioteche: un mezzo per un mezzo

Riccardo Ridi, Università Ca' Foscari, Venezia

Le tecnologie sono mezzi per raggiungere scopi che, di per sé, spesso non hanno niente di tecnologico. Neppure le biblioteche sono dei fini in se stesse, ma solo degli strumenti che permettono alla società di conseguire obiettivi che vengono riassunti nei codici deontologici dei bibliotecari. I bibliotecari, dunque, non devono cadere né nell'errore della tecnofobia (rinunciando a tecnologie che potrebbero permettere loro di raggiungere i propri obiettivi con maggiore efficienza ed efficacia) né in quello, opposto, della tecnomania, adottando determinate nuove tecnologie solo perché di moda, senza capire prima se e come esse potrebbero risultare utili rispetto agli scopi delle biblioteche. L'intelligenza artificiale, ad esempio, è una tecnologia che probabilmente potrà aiutare i bibliotecari, ma perché ciò avvenga in una prospettiva né tecnofoba né tecnomane sarebbe bene, prima di tutto, distinguerne le diverse forme (debole, forte, intermedia), e poi vigilare che, nei suoi vari ambiti di applicazione (indicizzazione e reference automatici, consigli di lettura, sviluppo delle raccolte, formazione professionale, ricerca bibliografica, ecc.) essa non rischi di entrare in conflitto con alcuni importanti valori della deontologia bibliotecaria, come ad esempio la privacy, la neutralità e il diritto degli utenti sia di

rendersi conto autonomamente e criticamente delle strategie di ricerca che vengono adottate per soddisfare i loro bisogni informativi sia, eventualmente, di poterle liberamente modificare.

Venerdì 17 novembre

Ore 9.30-11.00 – Sala Galileo

Panel: *Dalla tecnologia all'educazione. Accesso, conservazione, metadattazione, indicizzazione, digital literacy nell'età delle intelligenze artificiali*

Accesso all'informazione e intelligenza artificiale

Maria Cristina Mataloni, ICCU

A partire dalla recente riorganizzazione delle basi dati e dei portali gestiti dall'Istituto (mediante il progetto "Sistema di Ricerca Integrato"), l'intervento intende offrire una riflessione sulle potenzialità dei nuovi strumenti che la tecnologia mette a disposizione per migliorare l'accessibilità dei servizi e conseguentemente allargare la platea dei potenziali fruitori. Verranno dunque presentate le prime sperimentazioni, svolte in collaborazione con l'Infrastruttura Software per il Patrimonio Culturale (ISPC) della Digital Library, per l'utilizzo della IA in due distinte applicazioni: il progetto SBN sommerso e una chatbox per il portale Alphabetic. Il primo mira a sfruttare le potenzialità dell'IA per disambiguare in maniera più efficace e sicura le schede bibliografiche e di authority; la seconda è rivolta agli utenti meno abituati a utilizzare gli strumenti complessi per la ricerca nei cataloghi on line e che invece hanno maggiore familiarità con operatori virtuali. I primi risultati della sperimentazione saranno presentati al pubblico nel corso di un evento che si terrà al Collegio Romano il 30 novembre p.v.

Deposito legale digitale e intelligenza artificiale

Elda Merenda, Biblioteca nazionale centrale di Roma

Il deposito legale è attualmente normato dalla legge 106/2004 e dal D.P.R. 252/2006 suo regolamento attuativo. Sebbene tra i documenti oggetto di deposito siano previsti anche i documenti diffusi tramite rete informatica, ad oggi il deposito legale del digitale non è obbligatorio poiché non è ancora stato emanato lo specifico regolamento attuativo a cui si riferisce l'art. 37 del D.P.R. 252/2006, che comunque prevede forme volontarie di sperimentazione da attuarsi presso le due Biblioteche nazionali centrali di Roma e Firenze. Da qui l'esperienza ormai più che decennale di Magazzini Digitali che da progetto si è trasformata in vero e proprio servizio. Nel frattempo il legislatore italiano ha recepito la Direttiva EU 790/2019 sul diritto d'autore e i diritti connessi nel mercato unico digitale attraverso il D.Lgs 177/2021 modificando la legge 633/1941 sul diritto d'autore e consentendo, tra le altre cose, l'estrazione di dati e la comunicazione al pubblico degli esiti delle ricerche basate sui dati estratti tramite il *text and data mining* compiuta da parte di organismi di ricerca e da istituti di tutela del patrimonio culturale per scopi di ricerca scientifica. Diventa quindi sempre più urgente avviare un servizio di deposito legale digitale stabile, che consentirebbe la creazione di una grande archivio istituzionale di dati autorevoli, autentici e fruibili, fertile terreno per le molteplici applicazioni dell'Intelligenza Artificiale al fine di migliorare l'efficienza e

l'accessibilità dei dati ed i servizi offerti agli utenti, oltre che di semplificare la gestione dei documenti da parte del bibliotecario.

ANNIF per la lingua italiana: i requisiti per l'addestramento degli strumenti di indicizzazione automatica e il progetto V.I.CO.

Chiara Storti, Biblioteca nazionale centrale di Firenze

Nel 2022, la Biblioteca Nazionale Centrale di Firenze ha avviato una sperimentazione per l'utilizzo per la lingua italiana del toolkit per l'indicizzazione automatica "ANNIF", sviluppato e mantenuto dalla Biblioteca Nazionale di Finlandia.

Il progetto ha avuto come esito la realizzazione di un prototipo funzionante, in grado cioè di restituire termini semanticamente rilevanti per testi in italiano, almeno nei domini del sapere per cui si dispone di sufficienti dati per l'addestramento. È infatti contemporaneamente emersa la mancanza, nel contesto nazionale, dei prerequisiti essenziali per l'addestramento di strumenti riconducibili alla categoria del *Machine learning* e per una loro ampia applicabilità in ambito GLAM/MAB.

Sarà quindi presentato il progetto "V.I.CO. - Vocabolario delle Identità Controllate", che ha lo scopo di mettere a disposizione un vocabolario partecipato, in formato RDF, dei termini relativi alle entità utili alla descrizione del contenuto dei testi in lingua italiana.

Si tenterà, infine, di collocare il progetto all'interno di una prospettiva più ampia riguardante l'impatto delle *Intelligenze Artificiali* sugli strumenti di recupero delle informazioni.

Educare alla Digital literacy in biblioteca

Luigi Catalani, Biblioteca nazionale di Potenza

Come applicare concretamente in biblioteca il Quadro delle competenze digitali per i cittadini (DigComp 2.2)? L'intervento prende le mosse dall'offerta formativa e dalla giovane esperienza del FabLab "La fabbrica delle idee", il laboratorio didattico e digitale incastonato nel Polo bibliotecario di Potenza, in cui sono confluite la Biblioteca nazionale e la Biblioteca provinciale del capoluogo lucano. Il primo anno di attività svolte nell'unico spazio permanente dedicato al coding, alla robotica educativa e alla fabbricazione digitale allestito in una biblioteca statale italiana è l'occasione per riflettere su potenzialità e limiti della coesistenza tra servizi bibliotecari, innovazione e apprendimento creativo. Combinando educazione alla lettura, information literacy e cultura open, la biblioteca può riaffermare il suo ruolo di hub culturale, spazio pubblico di produzione collaborativa, reinterprestando la sua missione di luogo deputato alla tutela e alla valorizzazione del patrimonio culturale alla luce delle nuove opportunità offerte dall'ecosistema digitale.

Ore 09.30-11.30 - Sala Cataloghi

Panel: Leggere con le macchine. Modelli, esperienze, prospettive

Reading(&)Machine: prime esperienze di valutazione

Sara Dinotola e Anna Maria Marras, Dipartimento di Studi storici, Università di Torino

Reading(&)Machine, sviluppato dai centri SmartData@PoliTO e VR@PoliTO del Politecnico di Torino con la collaborazione dal Dipartimento di Studi storici dell'Università di Torino e delle Biblioteche Civiche Torinesi (BCT), consiste in un ambiente immersivo in cui sono resi disponibili i suggerimenti di lettura generati da un sistema di raccomandazione. Il prototipo è stato messo a disposizione del pubblico, in diversi contesti, tra aprile e settembre 2023.

Dopo una breve presentazione delle caratteristiche e delle finalità del progetto, verrà descritto dove e come sono avvenute le sperimentazioni e in base a quali metodologie è avvenuta la valutazione dell'uso. Essa si è soffermata da un lato sull'accessibilità e sull'usabilità del sistema, dall'altro sull'utilità dei consigli di lettura. I dati raccolti hanno permesso di portare alla luce i punti di forza e le criticità del prototipo e hanno offerto interessanti spunti di riflessione per i possibili sviluppi futuri di Reading(&)Machine.

L'IA nello spazio della biblioteca

Federico Cesareo, Dipartimento di Architettura e Design, Politecnico di Torino

L'attuale era digitale sta ridefinendo radicalmente la fruizione e l'erogazione di servizi a livello societario, ponendo la necessità di ripensare e riconfigurare gli spazi fisici di molte istituzioni, tra cui quelli delle biblioteche italiane. In questa presentazione si tenterà di esplorare la tipologia di configurazione dello spazio necessario all'ibridazione dell'IA nelle sale delle biblioteche: come cambiano forma le biblioteche in funzione di un loro adeguamento funzionale e infrastrutturale? Per rispondere, si riconoscerà un'eterogeneità dello stato dell'arte, riscontrando la necessità di definire scenari di modalità di avvicinamento alla biblioteca digitale: le caratteristiche che lo spazio deve avere dipendono direttamente da come immaginiamo che vengano svolte funzioni e relative modalità di accesso allo spazio digitale. Si presenteranno, pertanto, quattro strategie di progettazione atte a prefigurare scenari facilmente futuribili, ma capaci di dare luogo a spazi fisici molto differenti, con significati solo in parte condivisi. Sul piano economico, invece, è cruciale riconoscere che le risorse finanziarie delle biblioteche italiane variano notevolmente. Per poter predisporre ambienti ibridi, permeabili all'IA, occorre riconoscere, da un lato, una tassonomia minima degli elementi d'arredo comune a tutte le sedi bibliotecarie e, dall'altro, la dotazione necessaria a garantire l'immersività del punto di accesso ai sistemi bibliotecari digitali. La combinatoria di questi due gruppi condurrà l'indagine a definire quattro soglie di investimento, ossia livelli di dotazioni tecnologiche variamente modulabili all'interno degli elementi d'arredo che caratterizzano gli spazi delle biblioteche, in funzione delle risorse economiche a disposizione. Ciò permetterà una prima imbastitura di un campo semiotico nuovo, utile a produrre prefigurazioni della biblioteca virtuale nello spazio fisico, al fine di favorire la comunicazione e la valorizzazione delle collezioni e aiutare le biblioteche a rimanere faro di conoscenza per le future generazioni.

Come (si) scrive e (si) legge (con) ChatGPT

Maurizio Lana, Dipartimento di Studi umanistici, Università del Piemonte Orientale

Scrivere testi con ChatGPT è attività ormai molto popolare, come mostra bene una qualsiasi ricerca in Amazon, dove sono molto numerosi sia i libri di tutti i generi (romantico, giallo,

fantascienza, ...) che dichiarano ChatGPT come (co)autore, sia i libri che insegnano come usare ChatGPT per scrivere, anche con connotazioni specifiche: scrivere con ChatGPT per guadagnare, scrivere con ChatGPT nella scuola o nell'università, ...

Sono movimenti di fronte ai quali anche se ci si reputa estranei non si può rimanere indifferenti perché stanno dando avvio a trasformazioni che operano nel profondo dell'universo librario e bibliotecario.

Questo contributo propone un approccio curioso, critico e studioso alla produzione di testi con ChatGPT, mostrandone esempi reali; ma suggerisce anche che attraverso la loro lettura trapassano sia rischi sia opportunità, di fronte ai quali il primo passo è essere informati, conoscere.

Tagging e sistemi di raccomandazione in ambiente documentario

Antonio Santangelo, Dipartimento di Filosofia e Scienze dell'educazione, Università di Torino

Partendo dall'analisi del funzionamento di Reading(&)Machine, il dispositivo basato sulla tecnologia dell'intelligenza artificiale e progettato per le biblioteche civiche torinesi al fine di favorire l'accesso dei cittadini al patrimonio culturale custodito all'interno di queste ultime, mi interrogo sul contributo che può essere fornito dalla teoria della narrazione di matrice sociosemiotica allo sviluppo dei sistemi di raccomandazione informatici di libri. Sulla scorta dei consigli di lettura proposti da Reading(&)Machine a chi si dimostra interessato a *Il fu Mattia Pascal* (Pirandello, 1904), mi propongo innanzitutto di verificare, appunto per mezzo di un metodo di indagine sociosemiotico, se le opere collegate automaticamente da questa macchina al noto romanzo di Pirandello dimostrino di avere un significato simile a quest'ultimo, o quantomeno a esso comparabile. Inoltre, mettendo alla prova tale metodo, tento di dimostrare che esso può venire utilizzato – innanzitutto nell'ambito bibliotecario, ma in generale da chiunque se ne voglia servire – per produrre parole chiave e metadati efficaci al fine di accomunare testi narrativi in maniera sensata. Il mio scopo, infatti, è di concentrarmi sul tema dell'indicizzazione semantica o per soggetto, al fine di ragionare su come questa disciplina così centrale della biblioteconomia contemporanea possa essere fatta dialogare, per l'appunto, con la sociosemiotica.

Ore 11.15-13.30 – Sala Galileo

Panel: *Tecnologie e ricerca di informazione in Biblioteca. Nuove prospettive*

Using generative AI for discovering library content

Christine Stohn, Exlibris-Clarivate

Generative AI has taken the world by storm with the launch of ChatGPT. This opens up new opportunities but also challenges for our community. In this session I'm going to discuss a prototype project combining the convenience of conversational discovery with the reliability of academic sources. Based on a subset of the ProQuest One Literature, the prototype is using Large Language Models (LLM) to generate answers from a corpus of academic literature including references to the source of information. I will discuss the goals, the project itself, the status and the experiences so far.

Un apprendista bibliotecario: l'Intelligenza Artificiale generativa al servizio delle persone

Mitchell B. Squire, DM cultura

Le potenzialità dell'Intelligenza Artificiale (IA) generativa nel settore bibliotecario delineano nuove prospettive per l'accesso, la gestione e la fruizione delle informazioni. L'uso dell'IA generativa potrà fornire supporto bibliotecario ininterrotto, personalizzato, disponibile 24/7, e assicurare risposte personali, informate e contestualizzate per ciascun utente. In quanto tecnologia abilitante e trasversale, l'IA nel contesto bibliotecario interessa vari aspetti, dalla gestione dinamica delle richieste degli utenti, all'ottimizzazione delle ricerche bibliografiche, dalla personalizzazione e "umanizzazione" dell'esperienza utente, all'approccio ludico-narrativo nella restituzione dei contenuti e dei servizi bibliotecari. Non meno importante è il supporto che l'IA può dare al bibliotecario nel suo lavoro. Massima attenzione, domani come ieri, tuttavia, andrà all'accuratezza scientifica dei contenuti.

Navigating the Future of Academic Libraries: Harnessing Generative AI for Research Enhancement and Information Integrity

Meghan Tylec, EBSCO

In an era defined by rapid progress in AI and machine learning, academic libraries remain pivotal in shaping research and the future of scholarship. Generative AI provides opportunities to dramatically improve access to information; from semantic search to conversational chat, and helping deliver insights to users, AI can improve the effectiveness of research. Yet challenges remain. While generative AI holds remarkable promise, it's crucial to understand its vulnerability to potential misinformation and factual inaccuracies. Furthermore, when implementing generative AI, we must consider how best to approach the user experience, as well as the quality and trustworthiness of the underlying metadata, knowledge graph and search technology. This presentation will look at how AI can improve research while ensuring the continued delivery of trustworthy and meaningful information to our users.

In un'epoca definita dai rapidi progressi dell'IA e dell'apprendimento automatico, le biblioteche accademiche continuano a svolgere un ruolo centrale nel plasmare la ricerca e il futuro dell'attività accademica. L'IA generativa offre l'opportunità di migliorare drasticamente l'accesso alle informazioni; dalla ricerca semantica alla chat conversazionale, e aiutando a fornire approfondimenti agli utenti, l'IA può migliorare l'efficacia della ricerca. Tuttavia, le sfide rimangono. Sebbene l'IA generativa sia molto promettente, è fondamentale comprendere la sua vulnerabilità a potenziali disinformazioni e imprecisioni. Inoltre, nell'implementazione dell'IA generativa, dobbiamo considerare il modo migliore per affrontare l'esperienza dell'utente, nonché la qualità e l'affidabilità dei metadati, del knowledge graph e della tecnologia di ricerca sottostanti. Questa presentazione analizzerà come l'IA possa migliorare la ricerca, garantendo al contempo la fornitura continua di informazioni affidabili e significative ai nostri utenti.

Surprise and Coincidence. Il machine learning di Mloll al servizio della coda lunga

Marco Goldin, Horizons Unlimited

Uno dei problemi legati alle biblioteche digitali è spesso la ratio tra l'offerta di grandi quantità di oggetti digitali e la consultazione degli stessi da parte degli utenti. Ne derivano inevitabilmente matrici sparse che non riescono a valorizzare la natura e l'offerta delle collezioni. Questo peraltro è ugualmente valido per i documenti fisici delle biblioteche e i loro cataloghi, ma è possibile dimostrare i vantaggi dell'accorciare i percorsi di trovabilità degli oggetti, secondo il paradigma *finding without seeking*.

Il servizio "Suggerimenti di Lettura" Mlol nasce così con lo scopo di ottimizzare il catalogo e di personalizzare la ricerca e la trovabilità di libri, audiolibri e risorse ad accesso aperto per l'utente, attraverso l'uso di un sistema di raccomandazioni creato appositamente per la lettura e basato sull'intelligenza artificiale.

L'innesto nella costruzione del modello di machine learning di algoritmi come il *Correlated Cross-Occurrence* (CCO) e la *Log-Likelihood Ratio* (LLR), ha permesso di sviluppare in autonomia, e di renderlo adatto al dominio dei libri, un sistema che va oltre i limiti dell'approccio unimodale che ha contraddistinto per lungo tempo i modelli di raccomandazione.

Data l'impossibilità di prescindere dalle "cascate informative" e la natura stessa della lettura e dell'editoria per quel che riguarda la "coda lunga" (che Chris Anderson nel 2006, allora direttore di Wired USA, sosteneva erroneamente sarebbe diventata l'ambito privilegiato dei consumi culturali e dell'industria dell'intrattenimento), il sistema di raccomandazioni Mlol nasce con lo scopo di valorizzare le sotto-comunità di lettori in un sistema informativo complesso, facendo leva sui segnali sociali e sulle intersezioni di preferenze nel grafo dei lettori.