BIBLIO THECAE .it

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
DIPARTIMENTO DI BENI CULTURALI

Giorgia Di Marcantonio

Intelligenza artificiale, Large Language Models (LLMs) e Retrieval-Augmented Generation (RAG).[1] *Nuovi strumenti per l'accesso alle risorse archivistiche e bibliografiche*

Introduzione

La progressiva evoluzione tecnologica sembra stia trasformando il modo in cui gli utenti interagiscono con gli ambienti e gli strumenti digitali.

Grazie anche a dei «sistemi di raccomandazione»,[2] le piattaforme apprendono dalle abitudini di chi le utilizza e si adattano a restituire risultati mediamente in linea con i loro interessi. Tale grado di personalizzazione è evidente nei motori di ricerca, nei servizi di streaming o ancora nelle applicazioni sui dispositivi mobili, dove sofisticati algoritmi analizzano i nostri comportamenti

* Ultima consultazione siti web: 02/05/2024

1 Si ringraziano il prof. Tiberio Uricchio (UniMC), ingegnere informatico ed esperto di Intelligenza artificiale per il supporto nella stesura del paragrafo 2, il prof. Federico Valacchi (UniMC) per gli spunti e le riflessioni condivise durante l'elaborazione del contributo e il prof. Carlo Bianchini (UniPV) per alcuni preziosi suggerimenti bibliografici.

2 Ko et al. 2022.

Bibliothecae.it, 13 (2024), 1, 146-173 DOI <10.6092/issn.2283-9364/19982>

al fine di ottimizzare l'esperienza di interazione quasi in tempo reale.[3]

L'avvento poi di interfacce utente basate sulla voce o sui gesti aggiunge nuove modalità di accesso all'ecosistema digitale, rendendo l'interazione potenzialmente più fluida e meno dipendente da dispositivi tradizionali, come tastiere o mouse. A ciò si aggiungono la realtà virtuale e aumentata che promettono di offrire modalità immersive fino a poco tempo fa difficilmente immaginabili.

Ma non solo. In molti settori è plausibile ritenere che l'Intelligenza Artificiale (AI) possa portare dei significativi miglioramenti, considerando la possibilità di far "agire" la macchina imitando funzioni cognitive umane, come l'apprendimento, il ragionamento e la comprensione del linguaggio naturale.[4]

Tali tecnologie non solo hanno aumentato l'*engagement* degli utenti[5] – si pensi ad esempio all'impatto dei Social Network o ai *Recommendation algorithms* – ma creano anche nuove opportunità per l'apprendimento, l'intrattenimento, la ricerca e, più nel dettaglio, per l'accesso alle fonti.[6]

Sotto il profilo archivistico e biblioteconomico, l'insieme di queste potenzialità tecnologiche potrebbe avere degli impatti significativi su diversi aspetti legati alla produzione, gestione e conservazione delle risorse informative,[7] come anche sulla dimensione della restituzione della conoscenza sulla quale si focalizza questo contributo.

---

3 Gomez-Uribe - Hunt 2016; Zheng et al. 2017.

4 Nilsson 2002, p. 21.

5 O'Brien - Toms 2008.

6 E.g.: Consensus AI (https://consensus.app/) o Perplexity AI (https://www.perplexity.ai/).

7 Rispetto alla dimensione archivistica, è prova dell'attenzione che la comunità scientifica rivolge a questi temi il progetto internazionale e interdisciplinare InterPARES TrustAI (2021-2026) (https://interparestrustai.org/) che si pone tra i suoi obiettivi anche quello di valutare i rischi e i benefici dell'utilizzo dell'AI nei processi documentali e, più in generale, negli archivi.

1. Da dove partiamo, dove stiamo andando, dove potremmo andare

Gli archivi e le biblioteche sono intrinsecamente il prodotto delle trasformazioni sociali. Ogni entità conservata, da un manoscritto a un documento digitale, non solo testimonia frammenti di realtà e tendenze culturali, ma è anche parte di una rete più ampia di informazioni che, quando collegata in modo efficace, offre una comprensione più profonda e multidimensionale del passato e del presente.

Come piattaforme attive per l'apprendimento e la ricerca, gli archivi e le biblioteche sono chiamati a sfruttare tutte le potenzialità che le Tecnologie dell'Informazione e della Comunicazione (ICT) offrono, al fine di amplificare quelle finalità istituzionali che li rendono essenziali per la memoria della nostra società e la sua diffusione. Non è un caso che le comunità scientifiche di dominio si stiano dotando di nuovi standard e modelli per la descrizione e la catalogazione[8] ai quali vanno necessariamente accompagnati strumenti di restituzione della conoscenza più potenti e nuove competenze professionali.

In questo contesto, il passaggio "dal record al dato" non è indolore,[9] e implica una sapiente integrazione delle ICT nei processi di descrizione e catalogazione, come anche in quelle dinamiche volte alla creazione di ambienti di restituzione delle informazioni.

È essenziale che questi "bacini digitali di risorse" siano progettati per facilitare agli utenti la navigazione del patrimonio, evitando che si perdano tra le nuvole di dati o in un mare indistinto di informazioni che potrebbero non essere direttamente pertinenti alla loro ricerca.[10]

Senza entrare nel dettaglio, vista anche la vasta letteratura a ri-

8 Come *Records in Contexts* per il settore archivistico e *Resource Description and Access* per quello biblioteconomico.

9 Di Marcantonio 2023.

10 Prom 2004; Yakel – Shaw - Reynolds 2007; Vassallo 2023.

guardo,[11] nel rapporto tra bit e carta, le biblioteche hanno mostrato una precoce predisposizione all'adozione delle ICT, anche grazie ad una visione strategica che ha riconosciuto l'importanza della "cooperazione interbibliotecaria" e "interistituzionale". Un esempio significativo di questa proattività si rintraccia già nel 1979, con il progetto SNADOC (Servizio nazionale di accesso ai documenti, progenitore di SBN) e dopo poco nel 1980 quando:

> [...] le biblioteche e i bibliotecari italiani avevano iniziato ad avvertire la necessità sia di un servizio bibliotecario nazionale che rispondesse in modo unificato alle loro esigenze sia di un'automazione delle biblioteche per accelerare e rendere più efficace quella collaborazione. Viene così istituita una commissione per l'automazione delle biblioteche che portò alla firma nel 1984 del Protocollo d'intesa fra il Ministero per i beni e le attività culturali e le regioni per uno speciale progetto denominato Servizio Bibliotecario Nazionale. Nello stesso anno nacquero anche una Commissione paritetica di esperti per il coordinamento bibliotecario e informatico e un comitato tecnico amministrativo che si sarebbe occupato del coordinamento economico e normativo.[12]

Sempre negli anni Ottanta la realtà archivistica italiana è stata investita in maniera sporadica e parziale dai processi di informatizzazione.[13] Come confermerà la lunga realizzazione della Guida Generale degli Archivi di Stato Italiani,[14] l'assenza di una vera "cooperazione nazionale" ed adeguati investimenti portò solo negli anni Novanta all'avvio del progetto "Anagrafe",[15] al quale seguirono: il Sistema Gui-

11 Solo a titolo esemplificativo si citano alcuni contributi a riguardo senza alcuna pretesa di esaustività, escludendo la letteratura sulle Digital libraries che richiederebbe una nota distinta: Weston 2002; Santoro 2006; *Biblioteche e informazione nell'era digitale* 2007; Weston - Vassallo 2007, p. 130-67; *Online Catalogs* 2009; Willer - Dunsire 2013; Guerrini 2022; *Bibliographic Control* 2022; Tomasi 2022.

12 Sabba - Plachesi 2017, p. 494.

13 Grana 2005, p. 92.

14 Pavone 1995.

15 Gruppo di lavoro per la revisione e la reingegnerizzazione del sistema informa-

da Generale,[16] il Sistema informativo unificato per le Soprintendenze archivistiche,[17] e il Sistema informativo degli Archivi di Stato.[18] La vasta, e a tratti frammentaria, offerta sul web delle risorse archivistiche,[19] rese necessario elaborare un punto di accesso unificato ai tanti bacini informativi (nazionali e locali) che si tradusse nell'elaborazione del Sistema Archivistico Nazionale.[20]

In tutti questi ambienti è possibile avviare delle ricerche per parole chiave o farsi guidare da alcune maschere predefinite.

Le homepage dell'OPAC-SBN[21] e di Alphabetica[22] presentano subito dei box per effettuare la ricerca, seguendo un'architettura *Google-like* a cui gli utenti sono mediamente abituati (Fig.1 e 2).[23]

Fig. 1. Homepage OPAC SBN

tivo nazionale «Anagrafe informatizzata degli archivi italiani 2000.

16 Sistema Guida generale degli Archivi di Stato italiani, <http://www.guidageneralearchivistato.beniculturali.it>. Cfr. Carucci 2004.

17 Sistema informativo Unificato per le Soprintendenze Archivistiche (SIUSA), <https://siusa.archivi.beniculturali.it>. Cfr. Bondielli 2001; Pastura 2006.

18 Sistema informativo degli Archivi di Stato (SIAS), <https://sias.archivi.beniculturali.it>. Cfr. Grana 2004; Feliciati - Grana 2005.

19 Valacchi 2008.

20 Sistema Archivistico Nazionale (SAN), <https://www.san.beniculturali.it>.

21 OPAC SBN, <https://opac.sbn.it/>.

22 Alphabetica, <https://alphabetica.it/web/alphabetica/>.

23 Bianchini 2017.

Fig. 2. Homepage Alphabetica

Internet Culturale[24] e il Sistema Archivistico Nazionale (SAN) hanno invece delle pagine di presentazione ricche di contenuti (a prima vista quasi affollate), nelle quali il box di ricerca è inserito ai lati della Homepage (Fig. 3 e 4).

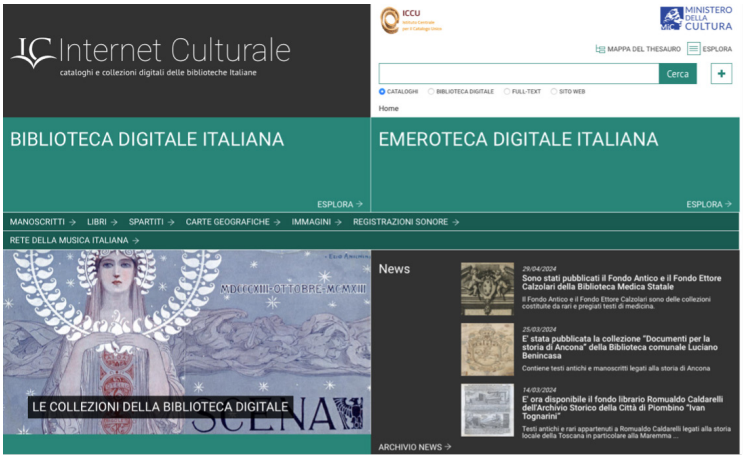

Fig. 3. Homepage Internet Culturale

24 Internet Culturale, <https://www.internetculturale.it>.

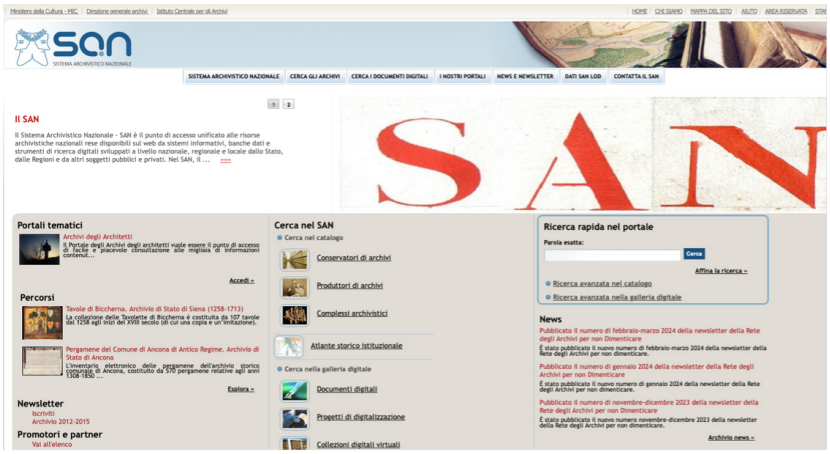

Fig. 4. Homepage Sistema Archivistico Nazionale (SAN)

Sul Sistema informativo degli Archivi di Stato (SIAS) è stato operato un recente *restyling* che lo ha reso abbastanza speculare all'impostazione di quello dedicato alle Soprintendenze archivistiche (SIUSA) e le homepage sono orientate a descrivere gli obiettivi progettuali, come anche vari possibili percorsi di ricerca (Fig.5 e 6).

Fig. 5. Homepage Sistema informativo degli Archivi di Stato (SIAS)

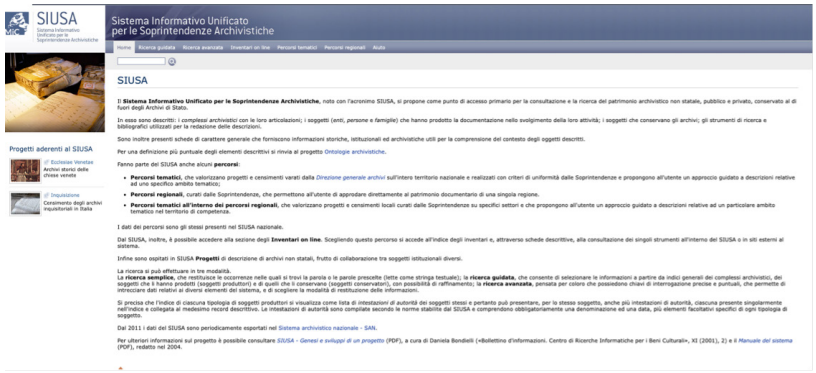

Fig. 6. Homepage Sistema informativo unificato per le Soprintendenze Archivistiche (SIUSA)[25]

Se è preferibile orientare questi ambienti verso una logica *Less is more*,[26] è di fatto comprensibile che la insita complessità delle risorse non faciliti le architetture dei contenuti. È altrettanto vero che alcune categorie di ricerca potrebbero essere semplificate, usando termini di uso comune, o comunque spiegate con un pop-up a fianco a diciture come "Date di esistenza" o "Intestazione di autorità" (Fig.7).

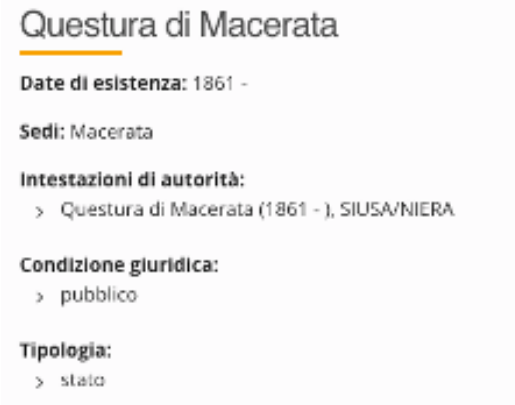

Fig. 7. Ricerca elaborata nel SIAS sulla Questura di Macerata (Soggetto produttore)

25 L'immagine risulta poco leggibile in quanto si è volutamente riportata la homepage nel suo complesso.

26 Beaird 2007, p. 67-98; Djamasbi - Siegel - Tullis 2010.

Se l'utente ha a disposizione tanti bacini di ricerca che, per ragioni diverse, presentano percorsi di navigazione differenti, grazie ad un uso consapevole e accorto dell'Intelligenza artificiale si potrebbero semplificare i sistemi di accesso agli ambienti di restituzione delle risorse,[27] o quanto meno rendere più intuitive le prime esplorazioni del patrimonio archivistico e bibliografico agli utenti.

2. Intelligenza Artificiale, LLM e RAG. Tecnologie utili per archivi e biblioteche?

Non è semplice fornire una definizione di Intelligenza artificiale. Luciano Floridi nel suo volume *Etica dell'Intelligenza artificiale* afferma che l'assenza di una vera e propria definizione dell'AI dimostra che «l'espressione non è un termine scientifico ma un'utile scorciatoia per far riferimento a una famiglia di scienze, metodi, paradigmi, tecnologie, prodotti e servizi».[28]

Ai fini di questo contributo, si circoscrive l'AI ad un sistema di tecnologie applicabile in ambienti di restituzione delle risorse culturali. È quindi utile richiamare la ISO/IEC 42001:2023 all'interno della quale l'AI è definita come quella capacità di un sistema informativo di mostrare capacità umane quali il ragionamento o l'apprendimento.[29] O ancora riportare la definizione che viene data nel recente AI ACT emanato dalla Commissione Europea, nel quale un:

> AI system' means a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives,

[27] Per un'interessante riflessione applicata sul fronte della didattica e dell'apprendimento si veda Kasneci et al. 2023

[28] Floridi 2002, p. 39.

[29] International Organization for Standardization 2023.

> how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.[30]

Un sistema di AI può agire come un agente, imitando processi di ragionamento, pensiero e comportamenti propri dell'essere umano,[31] utilizzando funzioni quali:

1. il riconoscimento di *pattern*;
2. l'apprendimento, anche da grandi basi di dati;
3. l'interpretazione semantica del linguaggio naturale o di altre risorse multimediali;
4. il ragionamento, come motore di *reasoning*;
5. l'esecuzione di comandi, sfruttando funzioni quali *Application Programming Interfaces* (APIs).

Tutte queste funzioni possono essere agilmente collocate all'interno degli ambienti di restituzione del patrimonio culturale.

Rispetto al primo punto, il "Design Pattern" è quella soluzione progettuale volta a risolvere delle richieste ricorrenti - *Pattern name; Problem; Solution; Consequences*.[32] Un sistema come Internet culturale o SAN ha l'obiettivo di organizzare le informazioni secondo architetture definite, al fine di restituire delle risposte agli utenti partendo da determinati input (chiavi di ricerca). L'AI potrebbe analizzare le interazioni come pattern di chi usa il sistema, identificare le ricerche ricorrenti e ottimizzare le risposte in base al contesto specifico. Questo si tradurrebbe in un ambiente più reattivo, anticipando le esigenze attraverso un processo di apprendimento continuo o comunque migliorando le prestazioni in fase di risposta.

Per quanto concerne la "grande base di dati", sia per gli archivi che per le biblioteche c'è solo l'imbarazzo della scelta. Forse troppa. Dagli anni Novanta ad oggi è difficile quantificare quanti progetti siano stati finanziati per la produzione di risorse digitalizzate o per la

30 European Commission 2024, Art. 3.
31 Russell - Norvig 2005, p. 4.
32 Gamma et al. 2011, p. 4; Heer - Agrawala 2006.

restituzione di dati catalografici o archivistici. Limitando la riflessione ai soli casi nazionali si distinguono: Opac-SBN, Alphabetica, Internet Culturale, Manus online, Cultura Italia, SAN (con tutti i suoi progetti affiliati come "Strumenti di ricerca online"), SIAS, SIUSA, Sistema Guida generale, Archivio Digitale,[33] etc. Alcuni di questi riportano anche le stesse informazioni perché creati come punti di raccordo di altri progetti (come nel caso di SAN o Alphabetica).

In questa miriade di risorse si collegano il terzo, il quarto e il quinto elemento di raccordo tra i sistemi informativi culturali e l'AI, ossia i *Large Language Models* (LLMs) integrati anche con sistemi di *Retrieval - Augmented Generation* (RAG).

I LLMs sono una classe di modelli di Intelligenza artificiale, addestrati per comprendere e generare testo in maniera naturale, come GPT3.[34] Questi sistemi si basano su architetture neurali – le più recenti sono le *transformer architectures*[35] – che consentono loro di apprendere rappresentazioni linguistiche anche molto complesse, attraverso l'elaborazione di grandi quantità di dati. Una volta addestrati, sono in grado di leggere ed interpretare del testo, rispondere a domande (chiamate *prompt*) ed interagire con le basi di dati eseguendo ricerche per termini chiave per conto dell'utente. Per realizzare queste funzioni sono necessarie principalmente tre fasi: una di *pre-training*, una di *instruction tuning* e infine una di utilizzo.[36] Nella prima fase il modello sfrutta una vasta gamma di fonti come libri, articoli, documenti, pagine web, etc. e viene addestrato tramite un apprendimento di tipo auto-supervisionato, dove l'obiettivo è quello di predire automaticamente la parola successiva dopo una determinata serie di vocaboli. Ad esempio, al modello viene sottoposta la sequenza "Il libro è sul" e ci si aspetta che questo predica la parola successiva "tavolo". Quando il modello predice qualcosa di improbabile (e.g. "zebra"), i suoi

33 Archivio Digitale, <https://www.archiviodigitale.icar.beniculturali.it>.
34 Floridi - Chiriatti 2020.
35 Min et al. 2022.
36 Hoffmann et al. 2022; Kaddour et al. 2023.

parametri vengono modificati per correggerne l'errore. Il testo viene quindi via via presentato al modello in sequenze dette *batch* ed eventualmente ripetuto fino al raggiungimento delle prestazioni massime ottenibili.

Al termine di questa prima fase il modello è in grado di predire il testo conseguente ma non di rispondere a domande o prompt di richieste. Nella seconda fase il processo di addestramento viene ripetuto in modo simile alla prima, ma sostituendo il testo generico con delle domande e risposte preparate da esperti, istruendo quindi il modello a quello che poi sarà il processo finale (Fig.8).

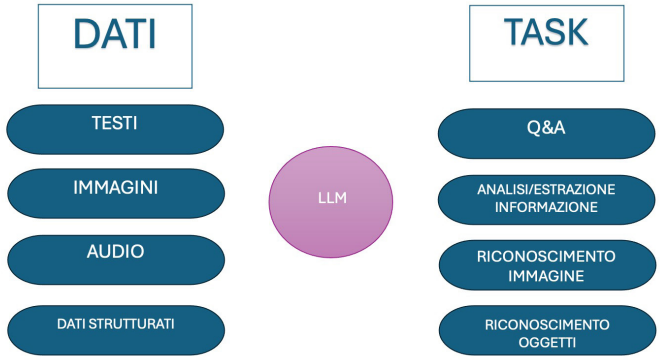

Fig. 8. Schema semplificato di un LLM generico dove sono evidenti le possibili tipologie di dati per l'addestramento e alcuni esempi di task

Una volta completato l'addestramento è possibile utilizzare l'LLM per tutti quei compiti che prevedono la generazione di testo, la traduzione automatica o l'analisi di altri dati. È da sottolineare che l'implementazione di queste tecnologie comporta costi significativi, sia in termini di investimento monetario che di risorse computazionali necessarie per l'addestramento dei modelli. Tuttavia, sono disponibili degli "Open LLMs"[37] che possono rappresentare un'opportunità per

37 Open LLMs, <https://github.com/eugeneyan/open-llms>.

ridurre i costi iniziali di implementazione e uno studio recente mostra come l'adozione di alcune strategie potrebbe ridurre l'investimento fino al 98%.[38]

I LLMs hanno dimostrato di essere dei modelli in grado di apprendere rappresentazioni linguistiche molto complesse e, al di là delle questioni tecniche, devono il loro successo anche alla possibilità di poter interagire con essi in modo naturale, quasi umano.

Grazie a OPEN AI[39] e al rilascio di Chat GPT[40] è possibile utilizzare un LLM di dimensioni considerevoli per svolgere una serie di attività, come ad esempio generare del testo a partire da delle domande o tradurre ed elaborare delle sintesi di corpus testuali anche molto estesi. *Facebook AI research* ha deciso di sviluppare e rilasciare apertamente un suo modello "LLAMA" (Large Language Model Meta AI), recentemente giunto alla terza versione, ma ne esistono anche molti altri frutto del contributo di società ed università.[41]

Questi sistemi hanno sì delle grandi potenzialità, ma non sono scevri da limitazioni e rischi. Possono infatti produrre dei risultati indesiderati, o completamente falsi detti anche allucinazioni,[42] riprodurre stereotipi a causa di *bias* presenti durante l'addestramento,[43] e sollevare questioni etiche rilevanti, vista la loro capacità di manipolare il linguaggio e influenzare potenzialmente le opinioni di chi li usa.[44] Tali limitazioni sarebbero del tutto inaccettabili all'interno di un ambiente di restituzione delle risorse culturali, proprio per lo scopo che essi hanno, ossia di offrire informazioni affidabili, scientificamente quali-

38 Chen, Zaharia - Zou 2023.
39 Open AI, <https://openai.com/about>.
40 ChatGPT, <https://openai.com/chatgpt>.
41 T5 – Text to Transfer Trasformer sviluppato da Google, XLNet, Bloom, Mee gatron, GPT-neo, etc. Cfr. Dao 2023.
42 Zhang et. al. 2023; Ji et al. 2023.
43 Ferrara 2023, p. 3-7.
44 Si vedano in tal senso alcune ricerche svolte su Minerva: Pezzali 2024, online: <https://www.dday.it/redazione/49301/abbiamo-provato-minerva-lia-italiana-della-sapienza-di-roma-e-fissata-con-il-sesso-e-spesso-risponde-senza-senso>.

ficate e quindi riutilizzabili per ricerche personali o per l'autoapprendimento.

Anche se le allucinazioni sono al momento inevitabili negli LLMs,[45] si possono apportare dei correttivi. Per ridurre le possibilità di allucinazioni su risposte dove le fonti sono essenziali è possibile abbinare i LLMs con dei sistemi per il *Retrieval - Augmented Generation* (RAG).[46]

Queste tecnologie rappresentano un'area di ricerca emergente molto interessante per il contesto applicativo nell'ambito culturale, poiché combinano in maniera sinergica tecniche di recupero dell'informazione e di generazione automatica di contenuti. In sostanza, durante i processi di generazione del testo, un sistema RAG integra un meccanismo di recupero dell'informazione che consente di estrarre dati rilevanti da una fonte di conoscenza predisposta, come un database o un insieme di risorse selezionate a monte per gli scopi dell'LLM. Questo corpus informativo viene quindi utilizzato per arricchire e guidare il processo di generazione del testo stesso, migliorando la pertinenza e l'accuratezza delle risposte generate.[47] Grazie alla sinergia tra il recupero dell'informazione e la generazione automatica di contenuti, i sistemi RAG possono restituire risultati più pertinenti e informativi, riducendo al contempo il rischio di generare contenuti distorti o del tutto falsi. Inoltre, l'integrazione di un meccanismo di recupero dell'informazione consente ai sistemi RAG di affrontare in modo più efficace il problema dei *bias* nei dati di addestramento, in quanto si utilizzano database accuratamente selezionati per gli scopi progettuali. Per far si che questo "archivio" creato appositamente per il sistema sia comprensibile alla macchina, di quel database se ne crea un altro speculare ma di tipo vettoriale, nel quale i testi sono rappresentati in sequenze numeriche attraverso un algoritmo di conversione (modello linguistico basato su *embedding*). Il sistema, quindi, confronta il vettore di input con i vettori di rappresentazione delle

45 Xu, Jain - Kankanhalli 2024.
46 Shuster et al. 2021.
47 Lewis et al. 2020.

informazioni presenti sul database e, utilizzando tecniche di similarità vettoriale (prevalentemente la similarità coseno), identifica le risorse più rilevanti per rispondere alla richiesta.

Al fine di comprendere come questi RAG funzionino è necessario concedersi delle semplificazioni formulando degli esempi:

- LLM stand-alone (i.e. ChatGPT): l'utente pone una domanda sotto forma di un prompt che viene sottoposta al LLM. Questo emette una risposta basata sul principio di predizione della parola successiva, cercando quindi di selezionare solo quella che sembra più corretta data una conoscenza del linguaggio naturale sviluppata durante l'apprendimento (database generico). Nel caso in cui debba presentare la data di nascita di un personaggio storico, se la conosce ne emette quella corretta, altrimenti potrebbe "allucinare" l'informazione, cercando di mantenere solo la sequenza sintattica corretta del testo (giorno, mese, anno) (Fig.9).

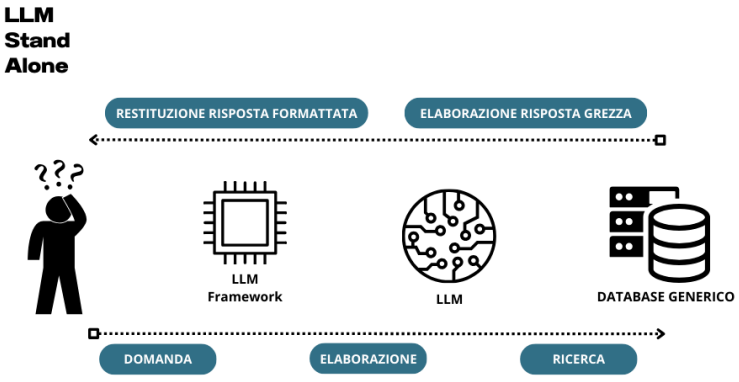

Fig. 9. Schema semplificato di un LLM Stand Alone

- LLM con sistema RAG integrato: l'utente pone una domanda. Il framework RAG interviene per primo eseguendo una ricerca semantica nel database vettoriale (che si basa su quella serie di

risorse che sono state precedentemente selezionate). I risultati ottenuti dalla ricerca vengono inviati come prompt al LLM insieme alla domanda dell'utente. Rispetto al sistema precedente, la risposta dell'LLM si può quindi basare non solo sulle conoscenze interne, ma anche su quelle che sono state inserite come input nel prompt, e quindi può rispondere correttamente anche quando il database interno risulta insufficiente. In questo caso il risultato della richiesta non sarà quindi solo una sequenza di parole ma la risposta verrà corredata anche da link o altre fonti che provengono direttamente da quel database selezionato in precedenza (Fig.10).

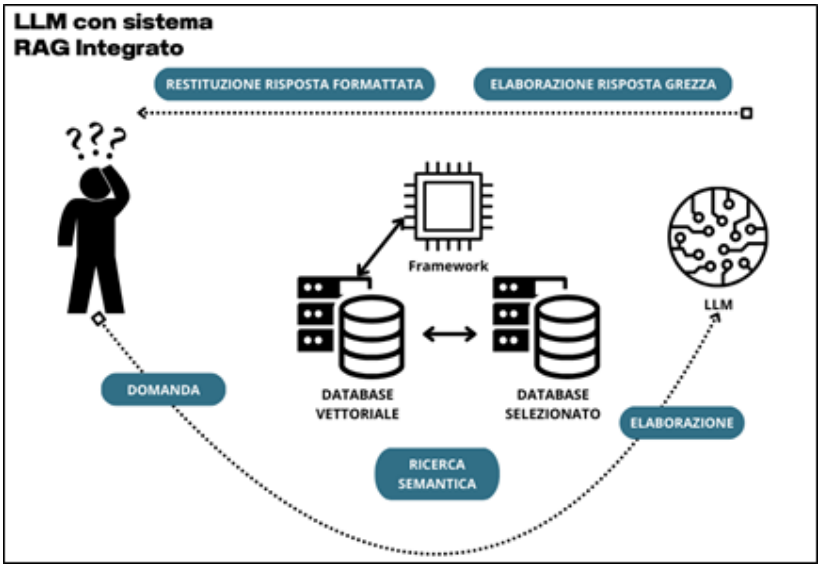

Fig. 10. Schema semplificato di un LLM con sistema RAG integrato

I RAG quindi si inseriscono direttamente nel prompt (*prompt before the prompt*) permettendo a chi utilizza questi sistemi di avere dei risultati più puntuali e affidabili rispetto ad un LLM addestrato su un database generico.

Inoltre, utilizzando i RAG è possibile aggiornare agilmente le risor-

se nel proprio database, inserirne di più pertinenti, identificare le fonti delle informazioni ed eventualmente correggere gli errori. Anche in questo caso permangono delle problematiche, soprattutto legate alla rigidità del database vettoriale e sulle possibili implicazioni nel medio lungo termine dell'utilizzo dei RAG, sui quali però si sta lavorando.[48]

In un'ottica di sistemi informativi integrati, anche ambienti come le Digital libraries o piattaforme orientate alla restituzione integrata di risorse diverse (archeologiche, archivistiche, bibliografiche, demoetnoantropologiche, storico-artistiche, etc.), i RAG possono svolgere un ruolo cruciale nell'interpretazione e nella sintesi di informazioni complesse, agevolando l'accesso a conoscenze altrimenti frammentarie o difficilmente consultabili. Per esempio, un sistema RAG integrato in una Digital library potrebbe gestire query molto complesse formulate in linguaggio naturale e restituire risorse contestualizzate e approfondimenti su documenti, opere d'arte, manoscritti o reperti archeologici, correlando automaticamente le richieste degli utenti con le informazioni contenute nei database collegati.

In definitiva, l'adozione di tecnologie che sfruttano una potenza computazionale fino a poco tempo fa impensabile e che permette di aggiungere nuove funzioni ai sistemi di restituzione, potrebbe ottimizzare i sistemi di accesso alle risorse culturali e aprire nuove prospettive per un'esplorazione del patrimonio, non solo nazionale ma, in potenza, anche globale.

2.1 E quindi?

I LMMs stand alone o con integrazione RAG non stravolgono gli ambienti di restituzione che attualmente vengono utilizzati. Le schede catalografiche, i cataloghi, le descrizioni archivistiche e gli inventari dovrebbero, anzi devono continuare ad essere disponibili per gli utenti. Quello che queste tecnologie potenziano sono i punti di ac-

48 Chen et al. 2024; Desai et al. 2024.

cesso all'informazione, ossia agevolano quella prima esplorazione/navigazione che si svolge quando non è ancora chiaro quali risorse esistono su ciò che si sta cercando. I RAG aumentano la serendipità e rendono la riposta, a partire dalla domanda dell'utente, più strutturata. Aggiungere queste tecnologie negli ambienti di restituzione dalla conoscenza può ridurre il tempo necessario per ottenere ed utilizzare le risorse, ma vanno considerate come delle funzionalità aggiuntive e non sostitutive a quello che già si dispone.

Si pensi ad un utente che sta cercando dei documenti, dei volumi, degli articoli, delle immagini, su un particolare evento storico. Con i LLMs/RAG si potrebbero ottenere non solo dei riferimenti a delle risorse pertinenti, ma anche una sintesi testuale che collega i vari documenti insieme.

Solo a titolo esemplificativo, si immagini un sistema informativo di una biblioteca che utilizza LLMs/RAG per assistere gli studenti nella ricerca di fonti su "La crisi del 1929 e le sue conseguenze in Italia". In questo contesto, quali parole chiave potrebbero essere inserite nel box di ricerca del sistema per ottenere risultati pertinenti? "Crisi del 1929"? "Conseguenze economiche in Italia"?. Invece di navigare attraverso un vasto database di documenti partendo da quelle parole, gli studenti potrebbero formulare la domanda in maniera naturale e ottenere come risposta un riassunto che sintetizza le teorie principali, corredate di fonti puntuali collegate a quelle più pertinenti alla ricerca. Un altro esempio potrebbe essere un sistema informativo archivistico che utilizza i LLMs/RAG per fornire agli utenti cronologie dettagliate e multimediali di eventi storici, integrando automaticamente testi, immagini e fonti documentali.

Per evitare ambiguità e permettere agli utenti di contestualizzare le risorse, è necessario che queste siano sempre accompagnate da quel corredo di metadati derivanti dai processi di catalogazione e descrizione, fornendo agli utenti la possibilità di approfondire autonomamente la ricerca a partire da quella prima risposta ricevuta. D'altro canto, i ricercatori più esperti potrebbero scegliere di non affidarsi a

dei LLMs con sistemi RAG ma esplorare il patrimonio direttamente tramite gli strumenti di ricerca tradizionali. In sostanza queste tecnologie aumentano le possibilità di esplorazione del patrimonio ma non possono sostituire (al momento) uno strumento di ricerca come un inventario o un catalogo.

Conclusioni

In conclusione, i RAG rappresentano un avanzamento significativo nell'ambito dell'accesso alle informazioni, poiché combinano l'efficienza dei LLMs nell'elaborazione del linguaggio naturale con potenti meccanismi di recupero delle risorse. Questo connubio permette di generare risposte più accurate e contestualmente rilevanti, riducendo il rischio di fornire informazioni inesatte o fuorvianti, un aspetto cruciale quando si tratta di restituire risorse culturali complesse. Per massimizzare i benefici delle tecnologie AI nel settore archivistico e bibliotecario, è essenziale che le istituzioni adottino una strategia olistica che includa la formazione dei professionisti, l'aggiornamento continuo delle infrastrutture tecnologiche e una collaborazione attiva sia a livello nazionale che internazionale.

Tutto questo potrebbe essere vantaggioso per l'Istituto centrale per la digitalizzazione del Patrimonio culturale. Infatti, tra le linee di azione del Piano Nazionale di Ripresa e Resilienza è stato previsto uno specifico investimento «finalizzato alla creazione di un patrimonio digitale della cultura attraverso la digitalizzazione dei beni culturali custoditi nei musei, negli archivi, nelle biblioteche e in tutti i luoghi della cultura».[49] L'Istituto centrale per la digitalizzazione del patrimonio culturale ha previsto un'infrastruttura software (ISPC) volta ad abilitare una serie di servizi al fine di consentire la creazione di un vasto ecosistema digitale nazionale per il patrimonio culturale. Tra i

[49] Piano nazionale di ripresa e resilienza 2021, in particolare 1.1 Strategie e piattaforme digitali per il patrimonio culturale.

servizi di accesso alle risorse, l'Istituto ha programmato la creazione di una Digital library grazie ad un investimento di 36 milioni di euro. Stando alla documentazione disponibile, l'ISPC garantirà la possibilità di creare grafi di conoscenza cross-dominio anche attraverso l'implementazione di motori semantici basati su tecnologie NPL, ontologie e Machine Learning (Cerullo e Negri 2023). Ciò implica che il database integrato dell'infrastruttura potrebbe potenzialmente già essere pronto per implementare una tecnologia RAG. Teoricamente questo "ecosistema" ospiterà milioni di oggetti digitali (con i relativi metadati) e pensare ad un servizio di accesso integrato che permetta di gestire query complesse in una Digital library particolarmente ricca, non potrà che essere vantaggioso alle macro-finalità dell'investimento del PNRR. Se poi la Digital library ospiterà anche i dati provenienti da quegli ambienti già esistenti è essenziale utilizzare delle tecnologie di accesso più performanti e meno rigide di come sono quelle attualmente utilizzate.

Bibliografia

Beaird 2007 = Jason Beaird, *The Principles of Beautiful Web Design*, Collingwood (Australia), SitePoint, 2007.

Bianchini 2017 = Carlo Bianchini, *«Funziona come Google, vero?» Prima indagine sull'interazione utente-catalogo nella biblioteca del Dipartimento di musicologia e beni culturali (Cremona) dell'Università di Pavia*, «AIB Studi», 57 (2017), 1, p. 23-49, <https://doi.org/10.2426/aibstudi-11557>.

Bibliographic Control 2022 = *Bibliographic Control in the Digital Ecosystem*, a cura di Giovanni Bergamin, Mauro Guerrini e Carlotta Alpigiano, vol. 7. Biblioteche & Bibliotecari / Libraries & Librarians, Roma, Macerata, Firenze, AIB, EUM, FUP, 2022.

Biblioteche e informazione nell'era digitale 2007 = *Biblioteche e informazione nell'era digitale: atti del convegno della 4a Giornata delle biblioteche siciliane, Ragusa, 26 maggio 2006*, a cura di Renato Meli, Palermo, AIB Sezione Sicilia, 2007.

Bondielli 2001 = Daniela Bondielli, *SIUSA - Sistema Informativo Unificato per le Soprintendenze Archivistiche Genesi e sviluppi di un progetto*, Pisa, Bollettino d'Informazioni. Centro di ricerche informatiche per i Beni Culturali, XI (2001), 2, p. 43-73.

Carucci 2004 = Paola Carucci, *Sistema Guida Generale degli Archivi di Stato italiani*, «Archivi & Computer», XIV (2004), 2, p. 52-63.

Cerullo - Negri 2023 = Luigi Cerullo, Antonella Negri, *L'infrastruttura software per il patrimonio culturale (ISPC) come abilitatore di un Ecosistema digitale nazionale del patrimonio culturale*, «DigItalia», 18 (2023), 1, p. 38-50, <https://doi.org/10.36181/digitalia-00059>.

Chen - Zaharia - Zou 2023 = Lingjiao Chen, Matei Zaharia, James Zou, *FrugalGPT: How to Use Large Language Models While Reducing Cost and Improving Performance,* «ArXiv», 2023, p. 1-13. Preprint: <https://doi.org/10.48550/ARXIV.2305.05176>.

Chen et al. 2024 = Xiaoyang Chen, Ben He, Hongyu Lin, Xianpei Han, Tianshu Wang, Boxi Cao, Le Sun, Yingfei Sun, *Spiral of Silences: How is Large Language Model Killing Information Retrieval? A Case Study on Open Domain Question Answering*, «ArXiv», 2024, p. 1-20. Preprint: <https://doi.org/10.48550/ARXIV.2404.10496>.

Dao 2023 = Xuan-Quy Dao, *Performance Comparison of Large Language Models on VNHSGE English Dataset: OpenAI ChatGPT, Microsoft Bing Chat, and Google Bard*, «ArXiv», 2023, p. 1-12. Preprint: <https://doi.org/10.48550/ARXIV.2307.02288>.

Desai et al. 2024 = Meera A. Desai, Irene V. Pasquetto, Abigail Z. Jacobs, Dallas Card, *An archival perspective on pretraining data*, «Patterns», 5 (2024), 4, p. 1-11, <https://doi.org/10.1016/j.patter.2024.100966>.

Di Marcantonio 2023 = Giorgia Di Marcantonio, *From Record to Data. New purposes for Archival Description processes*, «*JLIS.it*», 14 (2023), 2, p. 1-11, <https://doi.org/10.36253/jlis.it-549>.

Djamasbi - Siegel - Tullis 2010 = Soussan Djamasbi, Marisa Siegel, Tom Tullis, *Generation Y, Web Design, and Eye Tracking*, «International Journal of Human-Computer Studies», 68 (2010), 5, p. 307-323, <https://doi.org/10.1016/j.ijhcs.2009.12.006>.

European Commission 2024 = European Commission, «Regulation 2021/206 of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts», 2024.

Feliciati - Grana 2005 = Pierluigi Feliciati, Daniela Grana, *Dal labirinto alla piazza. Il progetto "Sistema Informativo degli Archivi di Stato"*, «Scrinia», II (2005), 2-3, p. 9-18.

Ferrara 2023 = Emilio Ferrara, *Should ChatGPT be Biased? Challenges and Risks of Bias in Large Language Models*, «FirstMonday», 28 (2023), 11. <https://doi.org/10.5210/fm.v28i11.13346>.

Floridi - Chiriatti 2020 = Luciano Floridi, Massimo Chiriatti, *GPT-3: Its Nature, Scope, Limits, and Consequences*, «Minds and Machines», 30 (2020), 4, p. 681-694, <https://doi.org/10.1007/s11023-020-09548-1>.

Floridi 2022 = Luciano Floridi, *Etica dell'intelligenza artificiale: sviluppi,*

opportunità, sfide, Milano, Raffaello Cortina, 2022.

Gamma et al. 2011 = Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides, *Design Patterns: Elements of Reusable Object-Oriented Software,* Boston, Addison Wesley, 2003.

Gomez-Uribe - Hunt 2016 = Carlos A. Gomez-Uribe, Neil Hunt, *The Netflix Recommender System: Algorithms, Business Value, and Innovation,* «ACM Transactions on Management Information Systems», 6 (2016), 4, p. 1-19, <https://doi.org/10.1145/2843948>.

Grana 2004 = Daniela Grana, *Il Sistema informative degli Archivi di Stato*, «Archivi & Computer», 2 (2004), p. 78-84.

Grana 2005 = Daniela Grana, *Le attività e i progetti di digitalizzazione nell'amministrazione archivistica*, «DigItalia», 1 (2005), p. 92-96.

Gruppo di lavoro per la revisione e la reingegnerizzazione del Sistema Informativo Nazionale "Anagrafe informatizzata degli archivi italiani 2000 = Gruppo di lavoro per la revisione e la reingegnerizzazione del Sistema Informativo Nazionale "Anagrafe informatizzata degli archivi italiani, *Riprogettare "Anagrafe". Elementi per un nuovo sistema archivistico nazionale*, «Rassegna degli Archivi di Stato», LX (2000), 2, p. 373-454.

Guerrini 2022 = Mauro Guerrini, *Dalla catalogazione alla metadatazione: tracce di un percorso*, a cura di Denise Biagiotti e Laura Manzoni, 2. ed., Roma, Associazione italiana biblioteche, 2022.

Heer - Agrawala 2006 = Jeffrey Heer e Maneesh Agrawala, *Software Design Patterns for Information Visualization*, «IEEE Transactions on Visualization and Computer Graphics», 12 (2006), 5, p. 853-860, <https://doi.org/10.1109/TVCG.2006.178>.

Hoffmann et al. 2022 = Jordan Hoffmann, Sebastian Borgeaud, Arthur Mensch Elena Buchatskaya, Trevor Cai, Eliza Rutherford, Diego de Las Casas, Lisa Anne Hendricks, Johannes Welbl, Aidan Clark, Tom Hennigan, Eric Noland, Katherine Millican, George van den Driessche, Bogdan Damoc, Aurelia Guy, Simon Osindero, Karen Simonyan, Erich Elsen, Oriol Vinyals, Jack William Rae, Laurent Sifre, *An empirical analysis of compute-optimal large language model training*, in *Advances in Neural Information Processing Systems*, edited by Sanmi Koyejo, Shakir Mohamed,

Alekh Agarwal, Danielle Belgrave, Kyunghyun Cho, Alice Oh, New York, Curran Associates Inc., 2022. Preprint: <https://proceedings.neurips.cc/paper_files/paper/2022/file/c1e2faff6f588870935f114ebe04a3e5-Paper-Conference.pdf>.

International Organization for Standardization 2023 = International Organization for Standardization, «ISO/IEC 42001:2023. Information technology, Artificial intelligence, Management system», 2023.

Lewis et al. 2020 = Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, Douwe Kiela, *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*, in *NIPS '20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, edited by Hugo Larochelle, Marc'Aurelio Renzato, Raia Hadsell, Maria-Florina Balcan, Hsuan-Tien Lin, New York, Curran Associates Inc., 2020, p. 9459-9474.

Ji et al. 2023 = Ziwei Ji, Niyeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishil, Ye Jin Bang, Andrea Madotto, Pascale Fung, *Survey of Hallucination in Natural Language Generation*, «ACM Computing Surveys», 55 (2023), 12, p. 1-38, <https://doi.org/10.1145/3571730>.

Kaddour et al. 2023 = Jean Kaddour, Oscar Key, Piotr Nawrot, Pasquale Minervini, Matt J Kusner, *No Train No Gain: Revisiting Efficient Training Algorithms For Transformer-based Language Models*, in *Advances in Neural Information Processing Systems*, edited by Alice Oh, Tristan Neumann, Amir Globerson, Kate Saenko, Moritz Hardt, Sergey Levine, New York, Curran Associates Inc., 2023. Preprint: <https://proceedings.neurips.cc/paper_files/paper/2023/file/51f3d6252706100325ddc435ba0ade0e-Paper-Conference.pdf>.

Kasneci et al., 2023 = Enkelejda Kasneci, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, Stephan Krusche, Gitta Kutyniok, Tilman Michaeli, Claudia Nerdel, Jürgen Pfeffer, Oleksandra Poquet, Michael Sailer, Albrecht Schmidt, Tina Seidel, Matthias Stadler, Jochen Weller, Jochen Kuhn, Gjergji Kasneci, *ChatGPT for*

good? On opportunities and challenges of large language models for education, «Learning and Individual Differences», 103 (2023), <https://doi.org/10.1016/j.lindif.2023.102274>.

Ko et. al. 2022 = Hyeyoung Ko, Suyeon Lee, Yoonseo Park e Anna Choi, *A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields*, «Electronics», 141 (2022), 11, p. 1-48, <https://doi.org/10.3390/electronics11010141>.

Min et al. 2022 = Erxue Min, Runfa Chen, Yatao Bian, Tingyang Xu, Kangfei Zhao, Wenbing Huang, Peilin Zhao, Junzhou Huang, Sophia Ananiadou, e Yu Rong, *Transformer for Graphs: An Overview from Architecture Perspective*, «ArXiv», 2022, p. 1-8. Preprint: <https://doi.org/10.48550/ARXIV.2202.08455>.

Nilsson 2002 = Nils J. Nilsson, *Intelligenza artificiale*, a cura di Salvatore Gaglio, Milano, APOGEO, 2002.

O'Brien - Toms 2008 = Heather O'Brien, Elaine G. Toms, *What Is User Engagement? A Conceptual Framework for Defining User Engagement with Technology*, «Journal of the American Society for Information Science and Technology», 59 (2008), 6, p. 938-955, <https://doi.org/10.1002/asi.20801>.

Online Catalogs 2009 = *Online Catalogs: What Users and Librarians Want: An OCLC Report,* a cura di Karen Calhoun, Diane Cellentani, OCLC, Dublin, Ohio, OCLC, 2009.

Pastura 2006 = Maria Grazia Pastura, *Il Sistema informatico unificato delle soprintendenze archivistiche (SIUSA)*, «Archivi & Computer», XVI (2006), 3, p. 12-18.

Pavone 1995 = Claudio Pavone, *La Guida generale degli Archivi di Stato, riflessioni su un'esperienza*, «Le carte e la storia», 1 (1995), p. 10-12.

Pezzali 2024 = Roberto Pezzali, *Abbiamo provato Minerva, l'AI italiana della Sapienza di Roma: è fissata con il sesso e risponde senza senso*, DDAY.it, 10/05/2024, online: <https://www.dday.it/redazione/49301/abbiamo-provato-minerva-lia-italiana-della-sapienza-di-roma-e-fissata-con-il-sesso-e-spesso-risponde-senza-senso>.

Piano Nazionale di Ripresa e Resilienza 2021 = Piano Nazionale di Ripre-

sa e Resilienza, PNRR, 2021, <https://www.governo.it/sites/governo.it/files/PNRR.pdf>.

Prom 2004 = Christopher Prom, *User Interactions with Electronic Finding Aids in a Controlled Setting*, «The American Archivist», 67 (2004), 2, p. 234-268, <https://doi.org/10.17723/aarc.67.2.7317671548328620>.

Russell - Norvig 2005 = Stuart J. Russell, Peter Norvig, *Intelligenza artificiale: un approccio moderno*, Milano, Pearson Prentice Hall, 2005.

Sabba - Plachesi 2017 = Fiammetta Sabba, Giorgia Plachesi, *Origini e prospettive del progetto SBN*, «AIB Studi», 57 (2017), 3, p. 493-514, <https://doi.org/10.2426/aibstudi-11711>.

Santoro 2006 = Michele Santoro, *Biblioteche e innovazione. Le sfide del nuovo millennio*, Milano, Editrice Bibliografica, 2006.

Shuster et al. 2021 = Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela, e Jason Weston, *Retrieval Augmentation Reduces Hallucination in Conversation*, in *Findings of the Association for Computational Linguistics: EMNLP 2021,* edited by Marie-Francine Moens, Xuanjing Huang, Lucia Specia, Scott Wentau Yih, Punta Cana, Dominican Republic., Association for Computational Linguistics, 2021, p. 3784-3803, <https://doi.org/10.18653/v1/2021.findings-emnlp.320>.

Tomasi 2022 = Francesca Tomasi, *Organizzare la conoscenza: digital humanities e web semantico*, Milano, Editrice Bibliografica, 2022.

Valacchi 2008 = Federico Valacchi, *Contenitori e contenuti: ancora sull'offerta archivistica nel web*, «Archivi», IV (2008), 1, p. 33-72.

Vassallo 2023 = Salvatore Vassallo, *From typewriter to bit: how finding aids evolve*, « JLIS.it», 14 (2023), 3, p. 83-104, <https://doi.org/10.36253/jlis.it-559>.

Weston - Vassallo 2007 = Paul Gabriele Weston, Salvatore Vassallo., *"… e il navigar m'è dolce in questo mare": linee di sviluppo e personalizzazione dei cataloghi*, in *La biblioteca su misura: verso la personalizzazione del servizio*, a cura di Claudio Gamba e Maria Laura Trapletti, Milano, Editrice Bibliografica, 2007, p. 130-67.

Weston 2002 = Paul Gabriele Weston, *Il catalogo elettronico. Dalla biblioteca cartacea alla biblioteca digitale*, Roma, Carocci, 2002.

Willer - Dunsire 2013 = Mirna Willer, Gordon Dunsire, *Bibliographic Information Organization in the Semantic Web*, Oxford, Chandos, 2013.

Xu, Jain - Kankanhalli 2024 = Ziwei Xu, Sanjay Jain, Mohan Kankanhalli, *Hallucination is Inevitable: An Innate Limitation of Large Language Models*, «ArXiv», 2024, p. 1-26. Preprint: <https://doi.org/10.48550/arXiv.2401.11817>.

Yakel - Shaw - Reynolds 2007 = Elizabeth Yakel, Seth Shaw, Polly Reynolds, *Creating the Next Generation of Archival Finding Aids*, «D-Lib Magazine», 13 (2007), 5-6, <http://dx.doi.org/10.1045/may2007-yakel>.

Zhang et. al. 2023 = Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, Longyue Wang, Anh Tuan Luu, Wei Bi, Freda Shi, Shuming Shi, *Siren's song in the AI ocean: a survey on hallucination in large language models*, «ArXiv», 2023. Preprint: <https://doi.org/10.48550/arXiv.2309.01219>.

Zheng et al. 2017 = Pai Zheng, Shiqiang Yu, Yuanbin Wang, Ray Y. Zhong, e Xun Xu, *User-Experience Based Product Development for Mass Personalization: A Case Study*, «Procedia CIRP», 63 (2017), p. 2-7, <https://doi.org/10.1016/j.procir.2017.03.122>.

Abstract

I Large Language Models (LLMs) e i Retrieval-Augmented Generation (RAG) offrono un nuovo paradigma per l'interrogazione e la restituzione di informazioni, rendendo i processi di recupero delle risorse più efficienti e accurati grazie alla loro capacità di apprendere e generare delle risposte basate su vasti database di conoscenza. Il contributo prova ad illustrare questi sistemi in una forma semplificata al fine di aprire una riflessione scientifica sulla possibilità di integrare queste tecnologie nei sistemi di restituzione delle risorse archivistiche e bibliografiche, e più in generale del patrimonio culturale.

Intelligenza Artificiale (AI); Large Language Models (LLMs); Retrieval-Augmented Generation (RAG); Archivi; Biblioteche; Restituzione della conoscenza.

Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) systems offer a new paradigm for querying and retrieving information, making the resource recovery processes more efficient and accurate due to their ability to learn and generate responses based on vast knowledge databases. This paper aims to demonstrate these systems in a simplified form to initiate a scientific discussion on the possibility of integrating these technologies into archival and bibliographic resource retrieval systems, and more broadly, into cultural heritage management.

Artificial Intelligence (AI); Large Language Models (LLMs); Retrieval-Augmented Generation (RAG); Archives; Libraries; Retrieval Knowledge.