

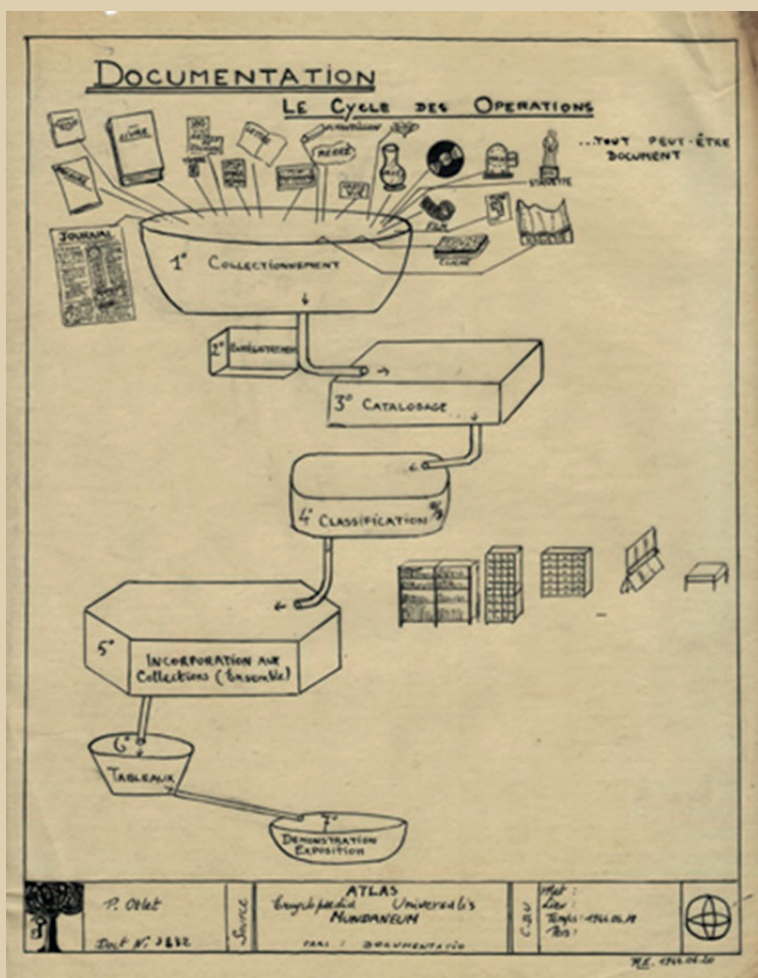
AIDainformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

NUMERO 1-2

ANNO 43

GENNAIO-GIUGNO 2025



AIDinformazioni

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 Chiaravallotti, *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 alle Scienze dell'Informazione, alla Documentazione, all'Archivistica, alla Gestione Documentale e all'Organizzazione della Conoscenza ma amplia i suoi confini in ulteriori campi di ricerca affini quali la Terminologia, la Linguistica Computazionale, la Statistica Testuale, ecc. 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 pubblicati affrontano questioni teoriche, metodologie adottate e risultati ottenuti in attività di ricerca o progettuali, definizione di approcci metodologici originali e innovativi, analisi dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Gruppo Scientifico Disciplinare 11/HIST-04 – Scienze del libro, del documento e storico-religiose e come rivista scientifica 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. È anche annoverata dall'ARES (Agence d'évaluation de la recherche et de l'enseignement supérieur) 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; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Ulrich's; 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 sono inviati in forma anonima a due referee, selezionati sulla base della loro comprovata esperienza nei topics specifici del contributo in valutazione.

AIDainformazioni

Anno 43

N. 1-2 – gennaio-giugno 2025

CACUCCI  EDITORE
BARI

PROPRIETÀ LETTERARIA RISERVATA

© 2025 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

Contributi

MONICA CONSOLANDI, SIMONE MAGNOLINI, MAURO DRAGONI, Risk Communication and Misunderstandings. The Healthcare Dominion Case	9
CAROLINE DJAMBIAN, Enhancing technoscientific heritage and introducing concepts into culture in the age of data, semantic web, and Artificial Intelligence. The ITinHeritage project	23
ERIKA PASCERI, I Knowledge Organization Systems come strumenti di sorveglianza sanitaria. L'applicazione della ICD in Italia	45
GRAZIA SERRATORE, La cartella clinica tra analogico e digitale: gestione e conservazione a norma	63
SALVATORE SPINA, <i>Schiavitù e invented archives</i> . Intelligenza Artificiale generativa nella costruzione del <i>Database on the Slave Trade</i>	85

Rubriche

ANTONIETTA FOLINO, Recensione del volume <i>ClaG – Classificazione dei giochi per ludoteche e biblioteche</i>	123
CLAUDIO GNOLI, Il macellaio pragmatista	127

Contributi

Schiavitù e invented archives. Intelligenza Artificiale generativa nella costruzione del *Database on the Slave Trade*

Salvatore Spina*

Abstract: This essay examines the construction of the *Database on the Slave Trade between the Mediterranean and the Atlantic (15th–16th centuries)* as a case study to explore the transformative role of generative Artificial Intelligence tools in historical and archival research. Positioned within the framework of *histoire sérielle*, the project addresses methodological challenges related to the datafication of sources on Mediterranean and Atlantic slavery, showing how *invented archives* function as epistemic spaces where memory, computation, and interpretation intersect. The use of LLMs (Large Language Models) for data extraction, encoding, and modeling raises critical issues concerning scientific validity, algorithmic opacity, and the indispensable role of human oversight. The essay offers a critical reflection on the integration of Digitality and Historiography, highlighting both the limitations and potentials of emerging technologies in transforming historical documents into queryable *cybertexts* and in fostering the collaborative construction of historical memory.

La ricerca si è avvalsa di un finanziamento dell'Unione europea – Next Generation EU – missione 4, componente 2, investimento 1.1, nell'ambito del programma PRIN-PNRR. Il titolo del progetto è: A Database on the Slave Trade between the Mediterranean and the Atlantic (15th-16th centuries).

Keywords: Histoire sérielle, ChatGPT, AI Models, Ollama, Chatbox.

1. Introduzione

La nascita del World Wide Web, nel 1991¹, ha segnato la “storia” della Storia e dell'Archivistica, non solo per la creazione di una nuova *heimat* comunicativa (Ciofalo e Leonzi 2013), della “Nicchia Ecologica Digitale” e della “Digitalità” (Spina 2024), ma per aver aperto un percorso in cui si faceva reale la possibilità di concretizzare progetti di valorizzazione archivistica, fino a quel momento inimmaginabili. Certamente, non tutto si è mosso con linearità; ma

* Dipartimento di Scienze Umanistiche, Università degli Studi di Catania, Catania, Italia.
salvatore.spina@unict.it. ORCID: 0000-0001-6367-8183.

¹ Il 6 agosto 1991 veniva lanciato il primo sito web della storia, dal fisico britannico Tim Berners-Lee – proprio da colui che ha pensato, progettato e realizzato il World Wide Web.

fu da subito evidente che, da quella data – che ha visto il primo sito andare online (World Wide Web, n.d.), col quale si dava libero accesso, agli utenti del Web, alle idee e alle caratteristiche proprie dell’“invenzione” del “WWW” –, tutto quello che veniva “uploadato” avrebbe avuto poco a che vedere con il concetto tradizionale di “archivio”. Si era, e lo si è ancora, di fronte a qualcosa di concettualmente nuovo: gli *invented archives*.

Grazie al Web, e alla costante implementazione di siti, è stato – quindi – possibile racchiudere documenti archivistici, con la finalità di dare corpo “digitale” a progetti dalla doppia anima, storica e archivistica; a idee di ricerca, o semplicemente ad espressioni di passioni personali – il “Mosaic Netscape” e il “Netscape Navigator” (Figg. 1, 2) (era già il 1994) diedero, infatti, a tutti la possibilità di creare le proprie collezioni, e di fare del Web una struttura informativa globale, superando le aspettative di Vannevar Bush e del suo *Memex* (Bush 1945).



Figura 1. Portale Web “Mosaic Netscape”.



Figura 2. Il browser Netscape Navigator.

Questa nuova frontiera ha segnato le prime fasi della storia “informatica” dell’Archivistica, la quale guardava attentamente alla possibilità di mettere online milioni di registri, indici e inventari, che fino a quel momento avevano guidato gli utenti alla ricerca dei documenti. Contestualmente, *a latere*, studiosi e gruppi di ricerca, grazie al Web, davano vita ad una nuova modalità di “far archivio e Storia”, che obbligò gli archivisti a fare i conti con una realtà documentaria sempre più interconnessa², che affiancò il concetto di “soggetto collettore/creatore” a quello di “soggetto produttore”, e il “vincolo tematico” al “vincolo archivistico”.

A partire dai primi anni del Ventunesimo secolo, il Web si “dota” di numerosi portali e aggregazioni tematiche, che si costruiscono intorno a svariati filoni e oggetti di ricerca (Trinkle e Merriman 2000); raccolte effettuate da singoli studiosi (professionisti e non), da comunità di ricercatori, da istituzioni accademiche o di ricerca (Rosenzweig 1997; 2001), la cui idea, di certo, ci

² Solamente per uno spunto, consiglio la lettura di Xu e Zhang (2022); Gonzalez et al. (2014); Reilly et al. (2000); Clerici e Pra (2012).

porta necessariamente, e per diversi motivi, a guardare all'*histoire sérielle*, ossia a rifocalizzarsi dentro quell'approccio che fondava, negli anni Cinquanta del Novecento, la sua metodologia sull'individuazione di determinati documenti archivistici, e il loro inserimento in una *serie* definita, allo scopo di spiegare un fenomeno storico – che diventava, per Furet (1971; 1981), condizione necessaria per il venire in essere della Storia Quantitativa (Darcy e Rohrs 1995; Dollar e Jensen 1974; Furet 1968; Itzcovich 1996).

2. *History of Slavery*. I dataset e gli archivi inventati come *histoire sérielle*?

L'introduzione dell'espressione *histoire sérielle*, in Storiografia (contemporanea), si deve alla figura di Pierre Chaunu, che ne ha elaborato e teorizzato la portata metodologica principalmente attraverso la sua opera fondamentale *Séville et l'Atlantique, 1504–1650*, pubblicata nel 1959, con la quale proponeva un radicale cambiamento epistemologico, spostando il focus della ricerca storica dall'"evento singolare e irripetibile" all'analisi di "fenomeni ripetitivi", seriali, e collocati all'interno di strutture coerenti e persistenti nel tempo, in cui la ripetizione non è più considerata un limite, ma il punto di partenza per l'interpretazione storica (Furet 1981, 6) – ed è certamente, all'interno di questa cornice teorica, che si colloca la genealogia degli *invented archives*, che hanno consentito di "materializzare" il concetto di "serialità".

Nel concreto della ricerca storica, ogni studioso "lavora nella serialità", operando inevitabilmente selezioni e organizzando le informazioni, isolando dati rilevanti e integrandoli in interpretazioni storiografiche specifiche. Questo processo, che può spaziare dall'analisi di eventi singolari, sino alla *longue durée* braudeliana (Braudel 2003; 2015; Braudel e Salsano 1974), richiede sempre una strutturazione logica delle fonti, la quale, se vero che è sempre stata frutto di una relazione "chiusa" (*close reading*) tra lo studioso e le carte, con il crescente processo di digitalizzazione e "matematizzazione" delle fonti documentarie (Spina 2021), ha richiesto strumenti adeguati, come il database – dove organizzare le sue informazioni –, che diventa indispensabile per lo studioso, il quale può, come evidenziato da Oscar Itzcovich, non solo raccogliere informazioni, ma renderle interrogabili, manipolabili e combinabili in modelli teorici dinamici (Itzcovich 1989; 1993).

"Serialità" e archivi inventati, dunque, si sovrappongono, per dare corpo ad un ambiente digitale appositamente realizzato per raccogliere e rendere accessibili documenti, originariamente disseminati in molteplici archivi fisici, e consentire indagini analitiche profonde e comparative, in grado di rendere visibile l'"invisibile" – come afferma Nowatzki (2021) –, così come nel caso del fenomeno di "lunga durata" della schiavitù, che è stato oggetto di numerose "esperienze

digitali”, quali il “North American Slave Narratives” (n.d.), ovvero i progetti “Slave Revolt in Jamaica” (Brown 2021), “Legacy of Slavery in Maryland” (The Maryland State Archives, n.d.), e l’“Early Caribbean Digital Archive” (Northeastern University Library, n. d.), “The Atlantic Slave Trade in Two Minutes” (National African-American Reparations Commission, n.d.), o, ancora, il “BlackDH Schema Project” (Lu e Pollock 2019), che propone un ripensamento radicale del *Text Encoding Initiative* (TEI), troppo neutro e distante dalle necessità di una rappresentazione dei soggetti afrodiscendenti, quindi, in chiave inclusiva.

Di rilevanza internazionale, sicuramente, il “Transatlantic Slave Trade Database” (oggi, Slave Voyages, n.d.) ha rivestito – e riveste – il ruolo di “progetto pilota” per tutte le altre esperienze successive.

Online dal 1999, il portale è frutto di anni di ricerca intrapresa da studiosi diversi, già a partire dagli anni Settanta del Novecento.

Già qualche anno prima, però, nel 1993, documenti relativi al fenomeno della schiavitù erano presenti nella rete, tra i dati raccolti nel sito “The Valley of Shadow”, ideato dallo storico Edward L. Ayers nel 1991 – progetto che nasce all’interno del Virginia Center for Digital History³ (Università della Virginia), allo scopo di condurre un’indagine storica, mai tentata prima, con l’ausilio di strumenti tecnologici (Ayers e Rubin 2000). Qui confluisce il “Free Black Registry (1803-1865)”, un apparato documentario che testimonia le dinamiche sociali di una realtà schiavista come quella della Virginia nella prima metà del XIX secolo.

Nell’ottobre del 2007, invece, presso il Nebraska Digital Workshop, tenutosi presso il Center for Digital Research in the Humanities (CDRH) dell’Università del Nebraska, Andrew Torget presenta il “Texas Slavery Project”, un portale frutto di un lavoro metodologico statistico che si proponeva di stimare i dati demografici mancanti relativi alla popolazione schiavizzata e ai possessori di schiavi nei territori della Repubblica del Texas nel periodo compreso tra il 1837 e il 1845. Operando su un insieme di 30 contee storiche texane, il progetto ha elaborato un modello predittivo basato sull’analisi comparativa delle tendenze demografiche nei contesti in cui le fonti risultano complete, applicando tali modelli ai casi lacunosi, per arrivare al numero “reale” di schiavi detenuti.

Nel 2011, trova la sua genesi il progetto “Enslaved: Peoples of the Historical Slave Trade”, che costituisce, oggi, uno degli snodi fondamentali nel panorama delle Digital Humanities applicate allo studio della schiavitù. Il progetto è una evoluzione di “Slave Biographies: The Atlantic Database Network”, finanziato dal National Endowment for the Humanities.

Nel dicembre 2014, nasce “Slavery and Remembrance”, da una iniziativa congiunta tra UNESCO Slave Route e la Colonial Williamsburg Foundation.

³ Il Virginia Center for Digital History (VCDH), uno spazio indipendente all’interno del College of Arts and Sciences dell’Università della Virginia, viene fondato nel 1998 da Edward L. Ayers e William G. Thomas III, allo scopo di sostenere progetti di ricerca in grado di sensibilizzare la comunità scientifica all’uso della tecnologia informatica in ambito umanistico.

Il progetto coinvolge più di cinquanta musei, siti storici e istituzioni in Europa, Africa e America, che hanno dato vita ad un sito web multilingue liberamente accessibile, che si concentra su musei e siti di memoria legati alla schiavitù e al commercio degli schiavi, a beneficio degli utenti in generale, degli studiosi, degli studenti e dei professionisti del patrimonio museale e culturale.

L'anno successivo, esattamente il 6 agosto 2015, appare online un nuovo progetto: "Liberated Africans", un portale scaturito da una collaborazione interdisciplinare che fa capo ad un team di sviluppo, il cui lavoro guarda alla possibilità di ripercorrere le vite di circa 225.000 schiavi africani portati via dalle navi negriere, ricostruire le attività di alcune delle prime corti internazionali del mondo, il tutto con profilo metodologico che può iscriversi nel solco della Public History, ossia nell'apertura del team di sviluppo alla comunità mondiale, garantendo la partecipazione di soggetti studiosi del fenomeno.

Contemporaneamente, ma su altri versanti, nel febbraio 2016, viene avviato il "Georgetown Slavery Archive", che costituisce, oggi, uno degli strumenti più avanzati e metodologicamente significativi per lo studio della schiavitù in relazione alle istituzioni religiose e accademiche statunitensi. Pensato e progettato nell'ambito dell'iniziativa "Slavery, Memory, and Reconciliation", promossa dalla Georgetown University, il progetto si configura come un archivio digitale in continuo aggiornamento, volto a raccogliere, conservare e rendere accessibile la documentazione relativa alla storia della schiavitù connessa ai gesuiti del Maryland e alla stessa Università di Georgetown.

Nell'aprile 2025, Internet Archive accoglie i documenti digitalizzati relativi alla Storia della Schiavitù ad Aruba, conservati presso il National Archives of Aruba (ANA) e la National Library of Aruba (BNA). La documentazione consente di percepire le condizioni di vita degli schiavi e dei loro discendenti, contribuendo a far luce sul loro doloroso passato.

Questi dati, digitalizzati, sono stati ufficialmente aggiunti al Memory of the World Program (MoW-AW) dell'UNESCO, e, da quest'anno, confluiti in Internet Archive, garantendo una ulteriore accessibilità, e una maggiore tutela contro l'obsolescenza.

3. Il "Database on the Slave Trade between the Mediterranean and the Atlantic (15th-16th centuries)"

Nel 2022, su bando PNRR, presso l'Università di Teramo e Catania prende avvio il progetto di ricerca volto alla creazione di un database dove raccogliere la documentazione utile a spiegare il fenomeno della tratta degli schiavi, tra Mediterraneo e Atlantico, nei secoli Quindicesimo e Sedicesimo.

Principal Investigators delle due unità sono Carlo Taviani (Teramo) e Lina Scalisi (Catania). Sullo sfondo del progetto, la necessità di spiegare quali di-

namiche caratterizzarono il passaggio tra il tardo Medioevo e l'Età Moderna, quando il "sistema-mondo" – così definito da Fernand Braudel (2003; 2017; Braudel e Salsano 1974) – atlantico si configurò progressivamente come uno spazio economico articolato, in cui circolavano, insieme ai metalli preziosi, alle merci e ai capitali finanziari, anche persone libere e schiavizzate. Questi flussi collegavano gli arcipelaghi atlantici, le coste occidentali dell'Africa e il Mediterraneo, inaugurando precoci percorsi del capitalismo moderno. Mentre gli studi storiografici hanno già ampiamente analizzato le dinamiche economiche⁴, politiche e sociali dei secoli XVII e XVIII, le fasi del fenomeno "schiavitù" precedenti all'impresa di Colombo sono state "offuscate" dall'interpretazione storica post-scoperta geografica del Nuovo Mondo, quindi, volta alla descrizione del fenomeno strutturato e intriso di violenza, nel consolidamento del complesso concetto "piantagione-schiavitù", nonché di sfruttamento intensivo delle risorse commerciabili, dai metalli alle fibre tessili, dalle specie animali ai prodotti agricoli. In questo quadro, i popoli africani e i territori del continente – in particolare le coste occidentali e le aree retrostanti – furono non semplicemente spettatori, ma elementi costitutivi dell'ingranaggio mercantile atlantico.

Ciò che viene posto in secondo piano, e che deve rivalutarsi come dimensione essenziale al network, è il ruolo del Mediterraneo; in particolare le sue strutture finanziarie e commerciali. Questo "continente liquido" (Braudel 2017), infatti, si innesta nel sistema, già molto tempo prima delle dinamiche atlantiche, e già prima dell'Età Moderna.

Cosa è, dunque, la figura dell'"esploratore" europeo, in Africa subsahariana, e negli arcipelaghi atlantici tra la metà del XV e l'inizio del XVI secolo?

A partire dagli anni Cinquanta del Quattrocento, si assiste a un mutamento rilevante nelle traiettorie del capitale europeo: da un sistema centrato sull'asse orientale – con il Levante e il Mar Nero come poli strategici – si passa a un orientamento occidentale che privilegia l'Iberia, il Maghreb e i margini atlantici. In questo ridirezionamento dei circuiti economici, un ruolo cruciale fu assunto dai mercanti-banchieri genovesi, i quali, attraverso una rete densa di insediamenti e partecipazioni finanziarie, si posizionarono strategicamente in tutta l'area iberica e negli arcipelaghi atlantici, finanche in Sicilia (Rau 1957; Verlinden 1970; Sicking e Wijffels 2020; Trasselli 1969; Gioffrè 1971; Bonazza 2023; Armenteros-Martinez 2022; Trasselli 1972; Gioffrè 1962; Schwartz 2004). Arcipelaghi come Madeira e le Canarie divennero poli produttivi centrati sulla coltivazione dello zucchero e sull'impiego intensivo della manodopera servile, mentre Capo Verde si affermò sin dai primissimi decenni come nodo fondamentale per la tratta atlantica degli schiavi. Pur essendo stato studiato il fenomeno dell'apertura dell'Atlantico, soltanto recentemente

⁴ Si vedano, tra la letteratura più recente: Delpiano (2021); Fiume (2012); Bono (2017); Sgroi (1979); Eltis e Richardson (2008).

la Storiografia ha iniziato a tracciare connessioni sistematiche tra mercati differenti e contesti sovranazionali, adottando approcci metodologici innovativi e sfruttando database digitali. Inoltre, sebbene la tratta transatlantica degli schiavi costituisca, oggi, un campo storiografico consolidato, le sue origini più remote – e in particolare tra Medioevo e prima età moderna – restano ancora poco investigate, soprattutto per quanto riguarda le sue premesse finanziarie, logistiche e infrastrutturali.

È in questa direzione che si inserisce il progetto per la costruzione del “Database on the Slave Trade between the Mediterranean and the Atlantic (15th-16th centuries)”³; finalizzato allo studio della tratta atlantica nelle sue fasi iniziali, con particolare attenzione ai nodi finanziari genovesi, che resero possibile la prima apertura commerciale dell’Atlantico.

L’obiettivo è quello di adottare strutture semantiche di modellazione dei dati, già sperimentate con successo nel settore dei Beni Culturali, per garantire la qualità FAIR (rintracciabili, accessibili, interoperabili e riutilizzabili) (Digital Library 2021) delle informazioni archiviate, ed assicurare la fedeltà epistemologica dei dati storici, e rendere possibile una rappresentazione accurata del fenomeno all’interno di ambienti computazionali, favorendo la condivisione delle fonti e dei risultati tra diversi gruppi di ricerca.

Alla base del database, troviamo la piattaforma open-source “Arches”, sviluppata dal Getty Conservation Institute e dal World Monuments Fund – che riunisce professionisti del patrimonio e sviluppatori di software provenienti da tutto il mondo – allo scopo di creare un sistema che consenta di creare database in cui riunire dati che possono “spiegare” un patrimonio culturale, inventari digitali, e altri tipologie di prodotti, attraverso diversi sistemi integrati di visualizzazione dei dati, finanche la georeferenziazione delle informazioni.

Il progetto ha una durata biennale, per questo, data l’enorme quantità di dati da acquisire e “tradurre” in records del database, è stata valutata – positivamente – la possibilità di utilizzare i Large Language Models, i quali, oggi, nelle “vesti” delle varie piattaforme di *Generative Artificial Intelligence*, consentono di lavorare con maggiore efficacia ed efficienza su grossi volumi di documenti.

Il progresso tecnologico, come già evidenziato nel grande filone metodologico su cui si ergono le *Digital Humanities* e la *Digital History* (Spina 2022a), ha ragionevolmente modificato la maggior parte delle metodologie di ricerca, portando anche le *Humanities* dentro la Nicchia Ecologica Digitale. Ciò si è tradotto in una rivalutazione degli approcci quantitativi, per il maggior controllo che lo studioso può operare sui dati e sulla possibilità di derivare, da essi, dei risultati epistemologicamente rilevanti.

Per la prima volta, dunque, gli storici hanno avuto la possibilità di applicare tools di Data Mining, come Weka, RapidMiner, KNIME, Orange, spaCy (Python), Stanford CoreNLP, e altri ulteriori strumenti per poter estrapolare quanti più dati possibili dalla documentazione storica, la quale,

contestualmente, veniva sottoposta ad un processo di trascrizione automatica, grazie alle IA non generative, come Transkribus (Massot et al. 2019; Milioni 2020; Muehlberger et al. 2019; Rabus 2019; Schlagdenhauffen 2020; Erwin 2020; Kahle et al. 2017; READ-COOP, s.d.; Spina 2022b; 2023a) ed eScriptorium (Gautier et al. 2022; Kiessling et al. 2019).

La documentazione archivistica esce – seppur con grossi limiti – dagli spazi di conservazione, per entrare nella rete sottoforma di “digital twins” (Banfi et al. 2023; Ahmadi-Assalemi et al. 2020; Correia et al. 2023; El Saddik 2018; Fuller et al. 2020; Gabellone 2022; Grieves 2023; Hutson et al. 2023) che cercano di tradurre in formato *machine-readable* quelli che possono diventare i Big Data della Storia (Bail 2014; Eijnatten et al. 2013; Graham et al. 2015; Franzosi 2017; Kaplan e di Lenardo 2017; Kennedy et al. 2017; Santoro 2015; Schiuma e Carlucci 2018; Spina 2020; Ruis e Shaffer 2017).

Così, tra i “ferri” del mestiere di storico, entrano i tools di Named Entity Recognition (NER), come spaCy, Flair, Transformers (Hugging Face), Stanza (Stanford NLP), OpenNLP (Apache).

Strumenti, questi tutti, che hanno realmente cambiato gli approcci quantitativi della ricerca storica, ma che richiedono una competenza informatica matura per consentire alle macchine di poter “interagire” con la documentazione storica.

La realtà della ricerca umanistica, però, è stata totalmente rivoluzionata nell’ultimo biennio, quando quello che è sempre stato il progetto “finale” della Computer Science – la creazione di una macchina “pensate” (McCarthy et al. 2006) – ha trovato una sua forma iniziale: l’Intelligenza Artificiale.

Cambiando totalmente gli schemi di interazione con i testi e i documenti digitali (e digitalizzati), le IA superano la potenza di calcolo degli altri tools – precedentemente menzionati e utilizzati –, aprendo la ricerca (sia umanistica che fisico-naturale) verso dimensioni e opportunità prima non facilmente accessibili.

Ma quali opportunità, realmente, possiamo evidenziare nel rapporto tra ricerca storica, Intelligenza Artificiale e mestiere di storico?

La creazione del nuovo database sulla schiavitù ha dato l’opportunità, al gruppo di ricerca, di testare le potenzialità di queste piattaforme, comprenderne i meccanismi e gettare le basi per la creazione di un sistema di prompt che possa tradursi in momenti di Machine Learning per l’addestramento di una IA storica, rispondendo, in tal senso, alla necessità evidenziata da Ennals (1986) di «dire alla [macchina] cosa noi [storici] facciamo».

3.1. La costruzione del dataset. *Named Entity Recognition* e modelli di IA a confronto

Il processo “digitale” si fonda sulle specifiche dell’architettura del database che si deve costruire, quindi sulla sua “ontologia”. Questo aspetto determina a quale procedimento computazionale deve essere sottoposta la documentazione archivistica per essere tradotta in *digital twin* (il gemello digitale). Oppure, optare per la sola estrapolazione dei dati, tralasciando di operare la creazione di edizioni digitali o formati che virtualizzano la consultazione – come, ad esempio, il flipbook (Spina 2023b).

Nel nostro caso, il lavoro di “digitalizzazione” si è fermato alla sola fase di estrapolazione dei dati, senza guardare alla possibilità di inserire le immagini dei documenti all’interno dei vari record creati.

Quindi, il lavoro ha previsto l’uso di strumenti di NER, per l’estrapolazione delle entità nominate (soggetti, luoghi, oggetti), la successiva compilazione di un file CSV con i dati ricavati, e la sua strutturazione per rispondere alle necessità dell’ontologia che sottostà al database.

Tralasciando, fin da subito, l’opzione di utilizzare gli strumenti classici, il ragionamento è stato tutto rivolto su quale LLM utilizzare per l’estrapolazione dei dati, tenendo ben chiaro che le IA generative, pur essendo strumenti potentissimi, in grado di svolgere operazioni diverse – per cui prima si utilizzava un solo tool per ognuna di esse –, senza uscire fuori dalla piattaforma, molte volte producono *outcomes* non esatti, o meglio: allucinati (Maleki et al. 2024; Salvagno et al. 2023).

Questo “bug”, soprattutto quando si applicano gli LLMs a complessi documentari molto vasti, scritti in lingue diverse dalle parlate – come il latino, per cui si richiede una specifica competenza –, rischia di creare set di dati non corretti, inficiando il lavoro di ricerca e la messa online di un database le cui informazioni non hanno corrispondenza archivistica e storica con quanto realmente scritto nel documento. Tale fenomeno, ad esempio, si è palesato in una fase di estrapolazione dei dati in cui si è chiesto a ChatGPT-3.5 di effettuare una operazione di NER su un testo latino, tratto da Carlo Taviani (2024, 206), di cui si riporta, qui, solamente la parte interessata: «[...] in dicto loco Orani, in quo loco eum vendidit dictus Dominicus cuidam mauro speciariorum qui multum frequentat in *terra nigrorum* mercatorio modo; quam venditionem fecit dictus Dominicus mediante ipso teste, ex quo causam habuit videndi dictum papirru, ex quo etiam scit que testificatur. Interrogatus quatenis est, respondit etatis annorum triginta quinque vel circa».

L'outcome dell'IA è il seguente (Fig. 3):

Il testo latino presentato è una documentazione notarile che risale al 18 e al 23 ottobre 1452, realizzata a Genova. Si tratta di una testimonianza legale riguardante una transazione commerciale e le relative contestazioni. Ecco un'analisi dettagliata dei soggetti e dei luoghi menzionati nel documento:

Soggetti

1. **Gabriel Pardus de Saona** - Navigatore di Saona, chiamato come testimone nella causa.
2. **Raffarelis Buroni** - Cittadino di Genova, padre e rappresentante di Dominicus Buroni.
3. **Dominicus Buroni** - Figlio di Raffarelis, diretto interessato nella questione della vendita di carta.
4. **Raffael Barixonus** - Cittadino di Genova, venditore delle ballas (rotoli) di carta a Dominicus.
5. **Thadeus Spinula** - Capitano o proprietario della nave su cui è stato trasportato il papirus (carta).
6. **Testimoni aggiuntivi e notari** - Includono Bartholomeo de Viacava, Quilico de Serravalle, Benedicto de Franchis de Pagano, e altri.

Luoghi

1. **Genova (Ianue)** - Luogo di redazione del documento, uno dei principali centri commerciali e marittimi del Mediterraneo nel XV secolo. Il documento stesso è stato creato nella piazza del Palazzo Ducale, luogo centrale per gli affari legali e politici della città.
2. **Marica** - Probabilmente un luogo di transito o di scambio commerciale. Non è specificato dove si trovi esattamente, ma era il luogo in cui Dominicus ha acquistato la carta.
3. **Orani** - Il luogo in cui Dominicus ha tentato di vendere la carta. È indicato che ha subito perdite economiche a causa della qualità inferiore di parte della carta venduta qui.
4. **Navi e viaggi** - Le navi sono un elemento centrale in questa narrazione, fungendo da mezzi per il trasporto delle merci (carta in questo caso) tra diverse parti del Mediterraneo. La nave di Thadeus Spinula è specificatamente menzionata come il mezzo su cui il papirus è stato caricato e trasportato.

Figura 3. Outcome ChatGPT 3.5.

Seppur corretti entrambi gli elenchi proposti dall'IA, quello che salta all'attenzione – palesando una “difficoltà” del modello – è l'assenza di una entità nominata, la cui formulazione nel documento non è specificatamente toponomastica, ossia l'espressione *terra nigrorum*, la quale si riferisce all'Africa subsahariana.

Questa è una informazione essenziale, non solo per la valenza storica, quanto per la correttezza del network che deve creare attraverso i dati del dataset, e la corretta esecuzione della funzione di NER.

La considerazione logica, quindi, è che la “black box”, che sottostà all'IA, non contiene informazioni storiche complesse, ma soprattutto, complete. ChatGPT, così come le altre IA, è stata allenata su dataset costruiti sulla base delle informazioni storiche presenti su Wikipedia e su altri siti internet “accessibili” alle macchine. Non esiste, dentro questi modelli, alcun dato storico ricavato dall'enorme e complesso patrimonio storiografico, esegetico, interpretativo, su cui si fonda il “mestiere di storico” (Bloch 1963).

Tale assenza può essere, dunque, colmata solamente con un corretto *training* che coinvolga gli esperti – in questo caso, gli storici –, al fine di “formare” efficacemente l'IA. Nel nostro caso, la semplice istruzione (prompt) «Terra nigrorum = Africa» ha consentito all'IA di effettuare correttamente l'estrapolazione e correggere l'elenco delle entità (Fig. 4).

Terra nigrorum = Africa

Hai ragione, la locuzione "terra nigrorum" si traduce letteralmente come "terra dei neri" e storicamente era utilizzata per riferirsi all'Africa. Nel contesto medievale e rinascimentale, questa espressione veniva comunemente adoperata per indicare varie parti dell'Africa, soprattutto quelle note ai commercianti e navigatori europei, come l'Africa del Nord e l'Africa Subsahariana, in base alle rotte commerciali e agli interessi coloniali dell'epoca.

Figura 4. Training attraverso prompt.

Questo aspetto, se vero che mette in evidenza i limiti dei modelli, è ancor più vero che sottolinea la necessità di un controllo costante sul flusso di "ragionamento" dell'IA. Ad oggi, non esistono strumenti per accertare il lavoro della macchina. Agli operatori e agli studiosi è dato l'onere di esaminare continuamente i risultati a cui perviene un LLM. Tutto quello che esso produce, in termini di *outcome*, infatti, non è "costantemente" corretto, *in primis* per il fatto che l'IA ed i modelli si basano su *black boxes* che possono essere descritte come "ambienti stocastici", i quali cercano di portare dentro questa tecnologia un principio basilare dell'"esistenza", ossia quel «dynamical behavior of each process (biological, industrial, economical, cultural) that is not stationary but evolves with the passage of time» (Wang et al. 2023) – si vedano, inoltre, anche Galić et al. (2023); El Haoud e Bachiri (2019). Questo procedimento punterebbe a rendere il processo computazionale sempre più simile a quello umano – progetto a cui OpenAI guarda con i suoi modelli di "reasoning" (OpenAI o3 e o4-mini), i quali vengono allenati per produrre *outcomes* dettagliati e per risolvere problemi complessi –, 'contemplando' la possibilità di migliorare, ad ogni step, le risposte fornite dall'IA. Tutto questo si traduce in un ambiente il cui meccanismo è incomprensibile, e che nei modelli più piccoli comporta che, date specifiche istruzioni, l'IA non le esegue in *routine*. Se al primo prompt, infatti, il sistema risponde efficacemente, la sua capacità computazionale cade di prestazione sui nuovi processi di NER. Per superare tale limite, occorre, infatti, dare alla macchina il prompt di partenza, per evitare che essa "dimentichi" cosa fare. Tutto questo perché il lavoro di elaborazione comporta, per i modelli (sia in remoto che *online*), un elevatissimo consumo di energia e dati, obbligando le aziende a porre dei limiti specifici sul numero di token che l'utente può utilizzare per far eseguire un procedimento computazionale. Nel caso dei modelli gratuiti di OpenAI, ad esempio, il limite è di 15.000 token, superati i quali, tutto quello che è stato inserito come informazione e/o prompt, viene progressivamente cancellato per far spazio ad un nuovo procedimento, oppure il modello restituisce l'errore di bloccare per alcune ore l'inserimento di nuovi prompt.

A questo si aggiungano – per quello che concerne la ricerca storica – due aspetti essenziali: (1) dentro i grandi modelli, l'informazione storica è pressoché nulla. I dati su cui vengono allenate le IA sono ricavati da enciclopedie online, e da qualche sito ad accesso aperto (senza la creazione di profili utente); tutto quello che riguarda la grande letteratura e la Storiografia, dunque, non è stato tradotto in dati per il training – a causa di violazione di copyright sulle pubblicazioni, accesso negato ai file delle pubblicazioni digitali, e altri aspetti che impediscono all'IA di “leggere” testi e documenti storiografici. (2) Le fonti archivistiche (su cui si fonda il lavoro di ricerca dello storico), per il loro “logico” *status* cartaceo, non sono consultabili dalle IA; e anche i grandi progetti di digitalizzazione di materiale d'archivio non guardano ad una adeguata “traduzione” dei testi in linguaggio comprensibile dalla macchina, e quindi non utilizzabile da quest'ultima per rispondere adeguatamente ai prompt che le vengono dati.

Tutto questo “obbliga” gli utenti/ricercatori/storici/archivisti a lavorare a fasi, in quanto, al raggiungimento del limite di token, il modello smette di elaborare, chiedendo all'utente di non immettere nuovi comandi, per un certo lasso di tempo – aumentando gli intervalli nel flusso di lavoro, che mostra tutti i limiti di un approccio digitale *tout court*, per cui, ad esempio, non è possibile effettuare una operazione di NER su un lungo testo, con un unico prompt.

Si è deciso, così, di effettuare dei test per capire quale modello avrebbe consentito di diminuire i tempi delle operazioni di riconoscimento delle entità nominate e compilare un file CSV scaricabile con i dati estratti.

Per operare efficacemente, i vari modelli vengono scaricati in locale, attraverso la piattaforma Ollama, che consente di effettuare il download dei modelli open-access, allo scopo di creare una IA “personalizzata”. I test vengono effettuati su macchina MAC STUDIO, con chip Apple M2 Max, con sistema operativo macOS Sequoia 15.4.1, 12 Core, 64 GB di Ram, GPU integrata con supporto Metal 3.

Configurata la piattaforma Ollama (Github, n.d.), e la Chatbox AI (n.d.) per l'interazione con gli LLMs, sono stati scaricati i modelli deepseek-v2, openai/hello, openchat, gemma, mistral, olmo2, deepseek-r1, llama3.1 (Figg. 5-12).

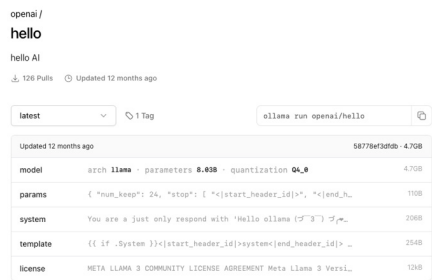


Figura 5. Modello Hello.

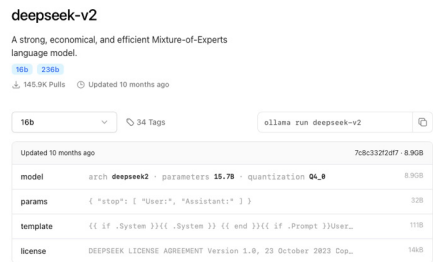


Figura 6. Modello Deepseek-V2.

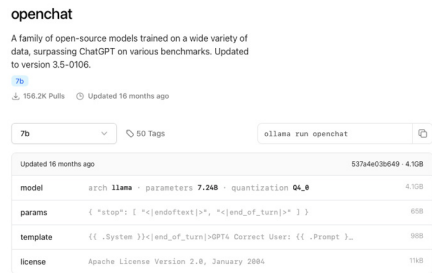


Figura 7. Modello Openchat.

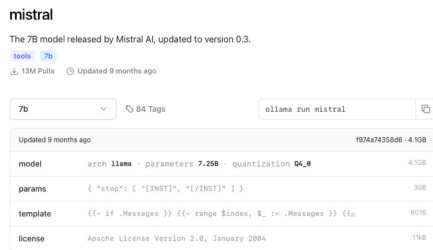


Figura 8. Modello Mistral.

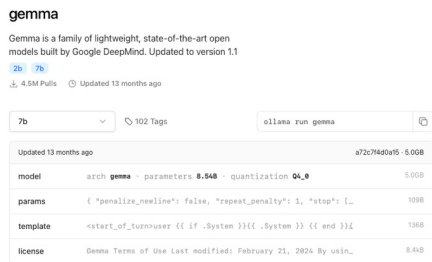


Figura 9. Modello Gemma.

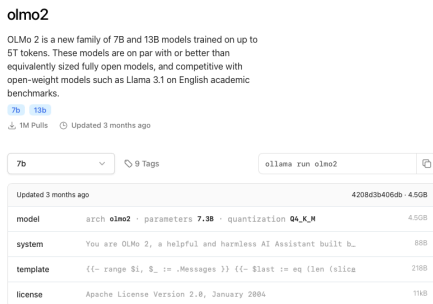


Figura 10. Modello Olmo2.

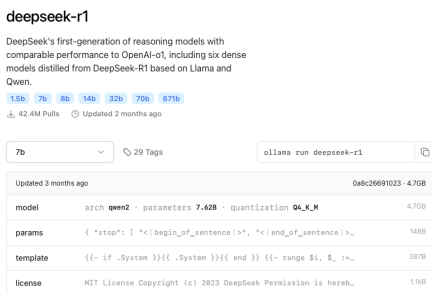


Figura 11. Modello Deepseek-R1.

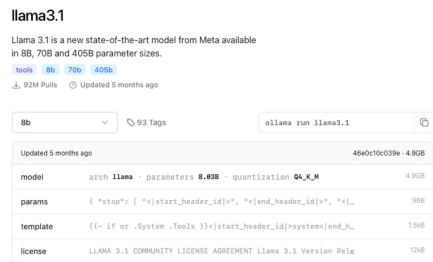


Figura 12. Modello Llama 3.1.

I vari modelli presentano specifiche caratteristiche, le quali sono determinanti nella scelta della versione da scaricare. I modelli più performanti, dal punto di vista dei risultati, sono quelli più grandi (671b, 405b, etc.). Ma più grande è il modello, maggiori sono le prestazioni hardware richieste. Ad esempio, un prompt di NER, nel caso di un modello di 236bv (deepseek-v2), espleta il processo in circa 40 minuti. I tempi aumentano quando ci spostiamo su modelli ancora più grandi. Invece, quelli più piccoli vengono gestiti senza difficoltà dall'hardware, ma i risultati che offrono sono meno dettagliati, e i processi di estrazione delle entità non portano i risultati sperati. Di seguito, nelle figure 13-18, si riportano gli *outcomes* di una NER eseguita su un documento trascritto (e tradotto) in inglese, pubblicato – e a cui si rimanda – in Aldo Agosto (1998, 92), in cui sono presenti queste entità nominate (persone): Susanna, Giacomo Fontanarossa, Domenico Colombo, Giuliano Caprile, Stampino Caprile, Goagnino Fontanarossa, Pietro Fazio, Antonio Amico, Paolo, Antonio Fontanarossa, Guglielmo Fontanarossa, Giovanni Fontanarossa, Bernardo Fontanarossa, Forestino Bavero, Antonio Bavero, Nicolò Bovero, Desidero Pitto, Nicolò, Lord Vicar, Lord Podestà of Genoa, Giovanni De Benedetto, Antonio, Antonio Garelo, Gerolamo.

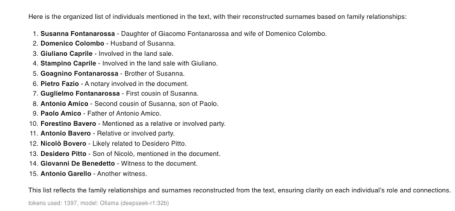


Figura 13. Outcome modello Deepseek-R1: 32b.

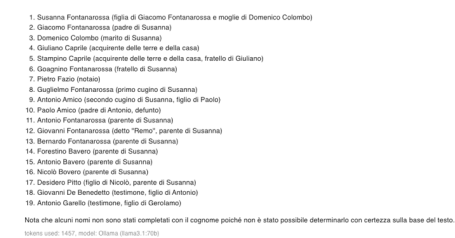


Figura 15. Outcome modello Llama 3.1: 70b.

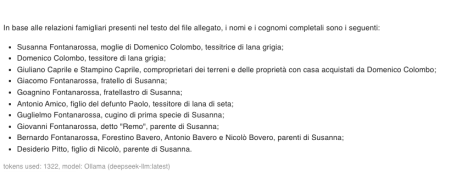
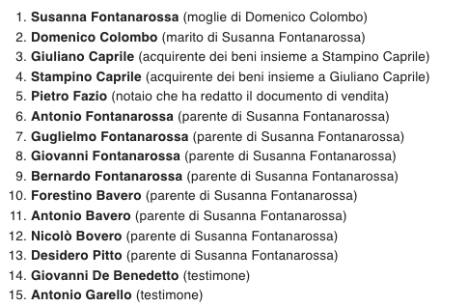


Figura 14. Outcome modello Deepseek-Llm: latest.



Spero che questo ti sia stato utile!

tokens used: 5382, model: Ollama (llama3.1:70b)

Figura 16. Outcome modello Llama 3.1: 70b.

Nomi di persone:

- Susanna
- Giacomo Fontanarossa
- Domenico Colombo
- Giuliano Caprile
- Stampino Caprile
- Goagnino Fontanarossa
- Pietro Fazio
- Guglielmo Fontanarossa
- Antonio Amico
- Paolo

Nomi di luoghi:

- Genoa

Organizzazioni:

- Notario

Altri nomi:

- Senatusconsultum Veleianum
- Lex Iulia de fundo dotali
- Epistole divi Adriani
- Authentica si qua mulier

tokens used: 1553, model: Ollama (gemma:latest)

Figura 17. Outcome modello Gemma: latest.

- Susanna Fontanarossa
- Domenico Colombo
- Giuliano Caprile
- Stampino Caprile
- Goagnino Fontanarossa
- Pietro Fazio
- Guglielmo Fontanarossa
- Antonio Amico
- Paolo
- Antonio Fontanarossa
- Guglielmo Fontanarossa
- Giovanni Fontanarossa ("Remo")
- Bernardo Fontanarossa
- Forestino Bavero
- Antonio Bavero
- Nicolò Bovero
- Desidero Pitto

Tabella dei Nominativi:

Nominativo	Descrizione in inglese
Susanna Fontanarossa	Daughter of the deceased Giacomo Fontanarossa
Domenico Colombo	Weaver of woolen cloth
Giuliano Caprile	Purchaser of certain lands and possessions
Stampino Caprile	Purchaser of certain lands and possessions
Goagnino Fontanarossa	Brother of Susanna Fontanarossa
Pietro Fazio	Notary
Guglielmo Fontanarossa	First cousin of Susanna Fontanarossa
Antonio Amico	Second cousin of Susanna Fontanarossa
Giovanni Fontanarossa ("Remo")	Son of the deceased Paolo
Bernardo Fontanarossa	Son of Giacomo Fontanarossa
Forestino Bavero	Son of Giacomo Fontanarossa
Antonio Bavero	Son of Giacomo Fontanarossa
Nicolò Bovero	Son of Giacomo Fontanarossa
Desidero Pitto	Son of Nicolò

tokens used: 1653, model: Ollama (gemma:latest)

Figura 18. Outcome tabella con modello Gemma: latest.

I modelli testati, al di là delle diverse prestazioni in ordine al tempo impiegato per l’elaborazione del prompt, compongono liste di nomi non complete, e la maggior parte di loro non riesce a ricostruire correttamente il quadro delle relazioni famigliari, partendo dalle descrizioni dei legami presenti nel testo del documento.

L’unico modello in grado di soddisfare il prompt è stato il “deepseek-r1:671b”, il quale, però, ha impiegato 57 minuti per comporre la lista e le relazioni famigliari.

Tutti i modelli, inoltre, in remoto, non sono in grado di organizzare i dati in un file CSV e consentire il download di quest’ultimo, per la costruzione dei records del database.

Questo quadro, così come descritto, si traduce in limiti reali nel lavoro dello studioso. Tempi lunghi di elaborazione e computazione non corretta creano un divario tra il mondo della Storia – e dell’Archivistica – e quello delle tecnologie di IA, le quali, seppur promettenti nelle loro prospettive di impiego, non presentano ancora un addestramento adeguato nei due settori di ricerca.

Un ulteriore test è stato effettuato mediante costruzione di un codice Python, con piattaforma Jupyter Notebook (n.d.), attraverso la quale è stato elaborato un codice per l’estrazione delle entità nominate, che ha restituito dettagliatamente le varie entità, ma mostrando un “bug” che necessitava un ul-

teriore lavoro (in termini di tempi di programmazione) di modifica del codice. Nel PDF, ogni documento è diviso da quello precedente, attraverso un sistema tradizionale di numerazione araba (Fig. 19). Tale separazione non è “individuata” dal codice, conseguentemente, l’elenco delle entità estratte comprende tutto quello che viene nominato nella pagina del PDF, creando una lista che, seppur corretta, non consente di estrapolare i dati per singolo documento archivistico, costruire il singolo regesto e compilare il singolo file CSV.

A questo si aggiunga la necessità di una configurazione costante del codice, relativamente alle API Key dei singoli modelli, per cui, pur valutando positivamente il risultato, non è stato implementato alcun codice Python per continuare il lavoro in questa direzione.

Di seguito, l’esempio del layout del PDF, il codice costruito e i risultati del test (Figg. 20-22).

the Genoese reckoning, the sixth day of September; present were Brothers Guglielmo di Tridino, Antonio Ratto di Monterosso and Simone di Mongiardino, citizens of Genoa, witnesses to this act, called and invited.

6
In the name of the Lord, amen. Pietro Ghirardetto, voluntarily and from certain knowledge and not influenced by any misunderstanding of law or fact, has acknowledged and has publicly recognized that Antonio Musante, called Fogliazzo, should have and should receive, either from himself, or from another for him, fifty-two lire and ten soldi di genovesi, for which sum Pietro was bound and obliged by the force of a public instrument written by the my hand, the below-written notary, in the year. . . . [Both men are] residents of the village of Quinto, in the jurisdiction of the podestà of Bisagno.

Renouncing, etc.
On which account the aforesaid Pietro gives quitclaim in every fashion, way, law, and form to the said Antonio, present, stipulating and receiving for himself and his heirs, and his goods, etc.; Making, etc.; Promising, etc.; That all, etc.; He promises, etc.; Under penalty of double, etc.; By calculation, etc.; And equally, etc.

Done at Genoa in the Borgo Santo Stefano on the door-step of the house inhabited by Lodisio di Pavia, wool-worker, in the year of the Lord's Nativity 1444, sixth indiction according to Genoese reckoning. Wednesday, the twenty-first day of January. Present as witnesses called and invited: Antonino, son of Colombo di Moconesi, resident of the village of Quinto and Lodisio di Pavia, wool-worker, citizen of Genoa.

7
Domenico di Terrarossa, resident of the village of Quinto in the jurisdiction of the podestà of Bisagno, voluntarily and from his certain knowledge, not influenced by any misunderstanding of law or fact, nor in any way deceived, sold and gave, granted, delivered, and consigned title of sale, or just as though he had, to Benedetto di Moconesi, resident of the village of Quinto, of the said jurisdiction of the podestà of Bisagno, present, stipulating, and purchasing for himself, his heirs and his successors that from him or from them may have and shall have a legal claim, a plot of land from the same Domenico, part in chestnut trees and wooded, and part in meadow, site in the said village of Quinto, locality called “le Fasciole”; bounded above and below by the public street, on the one side by the property of Antonio Bagagna, and on the other by the property of Stefano Beaccia, and in part by the property of Deserino Garro, and if there are others, usage will determine the true boundaries.

To have, to hold, to enjoy, to possess, to benefit from, and to do anything else they please, the purchaser and his heirs and successors, from him, or from them, shall have a legal claim, with all their rights, etc.; Free, etc.; Except that from dam-

6
In the name of the Lord, amen. Pietro Ghirardetto, voluntarily and from certain knowledge and not influenced by any misunderstanding of law or fact, has acknowledged and has publicly recognized that Antonio Musante, called Fogliazzo, should have and should receive, either from himself, or from another for him, fifty-two lire and ten soldi di genovesi, for which sum Pietro was bound and obliged by the force of a public instrument written by the my hand, the below-written notary, in the year. . . . [Both men are] residents of the village of Quinto, in the jurisdiction of the podestà of Bisagno.

Renouncing, etc.

On which account the aforesaid Pietro gives quitclaim in every fashion, way, law, and form to the said Antonio, present, stipulating and receiving for himself and his heirs, and his goods, etc.; Making, etc.; Promising, etc.; That all, etc.; He promises, etc.; Under penalty of double, etc.; By calculation, etc.; And equally, etc.

Done at Genoa in the Borgo Santo Stefano on the door-step of the house inhabited by Lodisio di Pavia, wool-worker, in the year of the Lord's Nativity 1444, sixth indiction according to Genoese reckoning. Wednesday, the twenty-first day of January. Present as witnesses called and invited: Antonino, son of Colombo di Moconesi, resident of the village of Quinto and Lodisio di Pavia, wool-worker, citizen of Genoa.

Figura 20. Testo tratto da Aldo Agostò (1998, 32).

Figura 19. Pagina PDF dei vari documenti, senza separazione tra le parti.

```

import os
import sys
import sys
import sys
import sys

from urllib.request import urlopen
from bs4 import BeautifulSoup

import requests

from langchain.prompts.chat import ChatPromptTemplate
from langchain_openai import ChatOpenAI
from langchain.memory import ConversationBufferMemory
from langchain.llms.openai import OpenAIDefault
from langchain.callbacks.base import BaseCallbackHandler
from langchain.callbacks.streaming_stdout import StreamingStdOutCallbackHandler
from typing import Callable


##### CONFIG #####

# Load configuration from .env file (contains openai, api_key, anthropic, google)
OPENAI_API_KEY = "config=api-key=sk-xxx,config=openai=xxx"
ANTHROPIC_API_KEY = "config=api-key=sk-xxx,config=anthropic=xxx"
GOOGLE_API_KEY = "config=api-key=sk-xxx,config=google=xxx"

##### CONTEXT #####

def load_context_config_file():
    try:
        with open(config_file, "r") as f:
            return json.load(f)
    except FileNotFoundError:
        print(f"Configuration file {config_file} not found.")
        return {}
    except json.JSONDecodeError as e:
        print(f"Error decoding JSON: {e}")
        return {}

def extract_text_from_pdf(pdf_path):
    """
    Extract e restituisce il testo dal file PDF specificato.
    """
    text = ""

    try:
        with open(pdf_path, "rb") as f:
            reader = PyPDF2.PdfReader(f)
            for page in reader.pages:
                page_text = page.extract_text()
                if page_text:
                    text += page_text

    except Exception as e:
        print(f"Errore nell'estrazione del testo da {pdf_path}: {e}")
        return text

def main():
    # Load config file
    with open(config_file, "r") as f:
        config = json.load(f)

    # Load context config file
    context_config = load_context_config_file()

    # Create LLM
    llm = OpenAI(
        api_key=config.get("openai_api_key"),
        model_name=config.get("openai_model"),
        temperature=config.get("openai_temperature"),
    )

    # Create Memory
    memory = ConversationBufferMemory(
        chat_history=[],
        output_key="response",
    )

    # Create Prompt Template
    prompt_template = ChatPromptTemplate.from_messages(
        [
            ("system", config.get("system_prompt")),
            ("human", "{text}"),
        ]
    )

    # Create Chain
    chain = LLMChain(llm=llm, memory=memory, prompt=prompt_template)

    # Run Chain
    response = chain.run(text)

    # Print Response
    print(response)

if __name__ == "__main__":
    main()
```

Per eseguire la **Named Entity Recognition (NER)** sui testi in inglese e latino, identificheremo le **entità** nominate come persone, luoghi (inclusi borghi, navi, carceri, mercati e luoghi simili), date, oggetti e altre **entità** rilevanti. Ecco l'analisi:

Testo in Inglese

Persone:

- Guglielmo di Triclinio
- Antonio Ratto di Monterosato
- Simone di Mongiardino
- Pietro Ghisardotto
- Antonio Musante (Fogliazzo)
- Lodovico di Pavia
- Antonio
- Celso de' Moconesi
- Domenico di Terrarossa
- Benedetto de' Moconesi
- Antonio Sagnora
- Stefano Resacchia
- Desiderio Garro

Luoghi:

- Genoa
- Borgo Santo Stefano
- Quinto
- Bisagno
- ~~Maconesi~~
- Terrarossa

```
### Date:
- 1444
- Wednesday, the twenty-first day of January
- the sixth day of September
```

~~see~~ Oggetti:

- fifty-two lire and ten soldi di **genovini**
- plot of land
- chestnut trees
- meadow
- public street
- property
- door-step of the house
- house
- wool-worker
- village
- jurisdiction of the **podestà**

Figura 22. Outcome analisi riconoscimento entità nominate.

Figura21. Prompt training NER file PDF.

Il processo, invece, è totalmente diverso, e realmente promettente, quando si passa alle versioni online delle IA, soprattutto nei modelli a pagamento (in abbonamento). In questo caso, infatti, non si riscontrano errori negli *outcomes* (Figg. 23-24).

In the name of the Lord, amen **Susanna**, daughter of the deceased **Giacomo Fontanarossa** and wife of **Domenico Colombo**, weaver of woolen cloth, present and authorizing all and each of the conditions written below, knowing and having specific knowledge concerning a sale by the same **Susanna**, made to **Giuliano Caprile** and **Stampino Caprile**, of certain lands and possessions with a house, as contained in a document of sale written by my hand, the notary inscribed below, and protested by **Goagnino Fontanarossa**, brother of the same **Susanna**, as shown by the document prepared by the hand of **Pietro Fazio**, notary; And knowing that the said **Domenico** promised that the same **Susanna** would have made the present renunciation, etc.; And knowing that the said lands and house had been pledged and placed under lien by her dowry and dower, and knowing that the said sale had been made for every advantage and benefit of the same consort, etc.; Therefore, voluntarily and from his certain knowledge, induced by no misunderstanding, either of law or fact, or in any other way misled, he consented and consents to the said sale, and renounced and renounces in regard to the same **Susanna**, his absence notwithstanding, and advised by me, the notary inscribed below stipulating as a public person and receiving by public office in the name of, and in the place of the same **Goagnino**, and of his heirs and successors, any right and lien whatsoever, and of whatever kind, due to, and that may be due to, and that the same **Susanna** may demand, in the house and lands aforesaid, as much because of her dowry and dower as for any other reason or cause that may be declared or devised.

The same **Susanna**, renouncing in the aforesaid the exception that the present renunciation, as above, was not made for the said reason, and matter, as above, that it is not, or was not done, either in that way, or otherwise, not finding fraud, negligence, fear in fact, actions, conditions, without cause, or from unjust cause, and all other rights; and renounced and renounces benefit of the *senatusconsultum Veleianum*, *lex Julia de fundo dotali*, right of lien, *epistole divi Adriani*, and the *authentica si quis maritus* and any other right, and the aid of rights and laws with which she may oppose any of the aforesaid, or protect herself in any way.

Informed, etc.; That all, etc. And she swore, etc.; Under penalty of double, etc.; Ratified, etc.; And therefore, etc.

The said **Susanna**, doing all the aforesaid in the presence, with and on the advice of **Goagnino Fontanarossa**, her first cousin, and **Antonio Amico**, son of the deceased **Paolo**, weaver of silk cloth, her second cousin, who are among her most competent and nearest relatives, as is affirmed, etc.; Swearing, etc.

Antonio Fontanarossa, **Goagnino Fontanarossa**, **Giovanni Fontanarossa**, nicknamed **Remo**, **Bernardo Fontanarossa**, **Forastino Bovero**, **Antonio Bovero**, **Niccolò Bovero**, and **Domenico Fazio**, son of **Niccolò**, are excepted because they were all cited by the distinguished **Pietro Fazio** of the upper hall of the magnificent Lord **Podestà di Genova** at the insistence of the said **Susanna**, in his capacity as husband and person married to the said **Susanna**, because they ought to consent and give their consent to the present document; yet they have not consented, nor given their consent; the said **Susanna** was pronounced and declared to be able to make the said document, with the consent of other relatives of the said **Susanna**, and therefore it shall be valid as if the said **Antonio** and his associates had consented and given their consent, as established in my records, the notary inscribed below.

Done at Genoa, in the upper hall of the Palace of the Podestà of Genoa, namely at my established bench, the notary inscribed below, in the year of the Nativity of the Lord, 1471, third indiction according to the Genoese reckoning, Saturday, the twenty-fifth day of May, after Terce, in the presence of the witnesses **Giovanni De Benedetto**, son of **Antonio**, and **Antonio Garello**, son of **Gerolamo**, Genoese citizens, invited and called to witness this.

Ecco la lista riorganizzata con i cognomi attribuiti sulla base delle relazioni familiari:

1. Susanna Fontanarossa – figlia del defunto Giacomo Fontanarossa e moglie di Domenico Colombo.
2. Giacomo Fontanarossa (defunto) – padre di Susanna.
3. Domenico Colombo – marito di Susanna, tessitore di panni di lana.
4. Giuliano Caprile – acquirente delle terre e possedimenti.
5. Stampino Caprile – acquirente delle terre e possedimenti.
6. Goagnino Fontanarossa – fratello di Susanna, protestante della vendita.
7. Pietro Fazio – notaio che redasse un documento riguardante la protesta di Goagnino.
8. Guglielmo Fontanarossa – primo cugino di Susanna, presente per consigliarla.
9. Antonio Amico – figlio del defunto Paolo Amico, secondo cugino di Susanna, presente per consigliarla.
10. Paolo Amico (defunto) – padre di Antonio Amico.
11. Antonio Fontanarossa – parente di Susanna, citato ma non consenziente.
12. Giovanni Fontanarossa, soprannominato "Remo" – parente di Susanna, citato ma non consenziente.
13. Bernardo Fontanarossa – parente di Susanna, citato ma non consenziente.
14. Forestino Bovero – parente di Susanna, citato ma non consenziente.
15. Antonio Bovero – parente di Susanna, citato ma non consenziente.
16. Niccolò Bovero – parente di Susanna, citato ma non consenziente.
17. Desidero Pitto – figlio di Niccolò Pitto, parente di Susanna, citato ma non consenziente.
18. Niccolò Pitto – padre di Desidero Pitto, parente di Susanna.
19. Signore Vicario della sala superiore del Podestà di Genova – funzionario che gestì la questione.
20. Magnifico Signore Podestà di Genova – autorità sotto cui si svolse l'atto.
21. Giovanni De Benedetto – figlio di Antonio De Benedetto, testimone.
22. Antonio De Benedetto – padre di Giovanni De Benedetto.
23. Antonio Garello – figlio di Gerolamo Garello, testimone.
24. Gerolamo Garello – padre di Antonio Garello.
25. Notaio (autore del documento) – non nominato esplicitamente, ma identificato come colui che ha redatto l'atto.

Figura 24. Risultato della NER con ChatGPT-4.5.

Figura 23. Testo tratto da Aldo Agostò (1998, 92).

Testate le varie piattaforme, dunque, il lavoro di estrazione dati e compilazione del CSV viene affidato completamente a ChatGPT-4 (successivamente 4.5).

Entrambi i modelli soddisfano le aspettative, ed i risultati sono eccellenti in tutte le operazioni richieste, consentendo di procedere all'analisi dei documenti utili alla creazione del database.

I documenti sono trascrizioni digitali tratti dalla relazione storica di Emilio Marengo, *Genova e Tunisi 1388-1515*, scritta nel 1901 per gli *Atti della Società Ligure di Storia Patria, Volume XXXII*, in cui sono confluiti 64 documenti, conservati presso l'Archivio di Stato di Genova, relativi all'oggetto di studio del progetto PNRR.

Il PDF del testo è stato trattato con software OCR, Adobe PDF Pro, per ottenere la trascrizione dei documenti, in formato DOC. Ottenute le trascrizioni, i testi sono stati processati, come già detto, con i modelli ChatGPT-4 e 4.5.

La scelta, oltre che motivata dagli errori negli *outcomes* dei modelli in remoto, è stata puntualizzata anche dal fatto che il modello di OpenAI riesce a

produrre file CSV correttamente compilati e scaricabili direttamente dall'interfaccia di dialogo. Inoltre, i due modelli riescono, con specifici comandi, a realizzare processi di Network Analysis che possono essere tradotti in chiave “visual” (Fig. 25) (Bleichmar e Schwartz 2019; Bradley et al. 2018; Ebbrecht-Hartmann et al. 2023; Staley 2015; Theibault 2016).

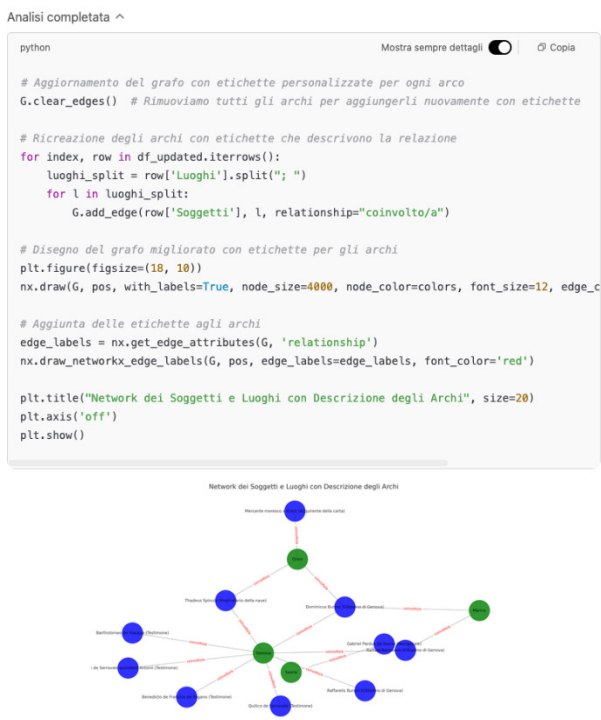


Figura 25. Esempio di network analysis.

Si aggiunga l'enorme capacità, di tale IA, di sviluppare codici in diversi linguaggi (es. Python), che possono essere utilizzati in processi di training – ad esempio, con TensorFlow (TF) –, al fine di potenziare il modello con i dati che esso stesso ha ricavato, in un processo che, se da un lato estrae dati, da un altro lato, li utilizza in una successiva fase di training, che riesce a tradurre quei dati in informazioni storiche.

Nel nostro caso, il codice TF compilato ha consentito di “allenare” il modello sui significati delle espressioni toponomastiche relative ai termini *Terra nigrorum* e su altre che sono state individuate in ulteriori documenti (Fig. 26).



```

Tensorflow Storia
1 import tensorflow as tf
2 from tensorflow.keras.preprocessing.text import Tokenizer
3 from tensorflow.keras.preprocessing.sequence import pad_sequences
4 import numpy as np
5
6 # Dataset di esempio (frasi e loro significati)
7 dataset = {
8     "Terra Nigrorum": "Africa Subsahariana",
9     "Regnum Guineae": "Africa Occidentale",
10    "Aethiopia": "Africa Orientale",
11    "Sudan Nigritarum": "Africa Centrale"
12 }
13
14 # Creazione del tokenizer
15 tokenizer = Tokenizer()

```

Figura 26. Training con TensorFlow per entità nominate (toponomi non specifici).

Tutto questo, però, richiede un controllo specifico da parte dell'utente e la compilazione di specifici prompt, che devono istruire la macchina a non creare *outcomes* "allucinati".

Il lavoro di configurazione dei prompt, infatti, è lungo e capzioso, in quanto l'utente deve predisporre le istruzioni in una vera e propria sequenza, in cui l'IA non deve essere lasciata libera. Ogni singolo step computazionale deve essere correttamente eseguito dal modello, e, per questo, il prompt deve essere esplicito, logico e discreto.

Dopo varie fasi e diversi test sulla struttura del prompt, che andava dal semplice «Estrai entità nominate» – che lasciava all'IA la libertà di individuare *entities* non esatte, o di crearne altre –, si addiviene ad un comando esplicito: «Estrai entità nominate, nello specifico: persone, luoghi, oggetti, funzioni, lasciando le date».

In questo caso, ChatGPT è riuscita ad estrarre correttamente tutte le entità e a dividerle per categoria (Fig. 27).

Soggetti

- **G. B. di Monteburgo:** la persona a cui è indirizzata l'istruzione.
- **Mercanti di Genova:** riferimento generico ai commercianti coinvolti.
- **Serenissimo re di Tunisi:** il re a cui si riferisce il documento.
- **Baptista de Grimaldis, Jacobus Lercarius, Joannes Palmarius, Paulus de Rapallo:** mercanti specificamente nominati.
- **Johanne Baptista de Monteburgo:** riferimento ripetuto al destinatario principale.
- **Augustino Petito:** citato in connessione con il galleone.
- **Procuratori del archayto Gemoxi:** menzionati in relazione a una transazione.
- **Jo. Baptista Sixto, Benedecto Joanne Pinello, Petro Paulo de Vivaldo:** individui coinvolti in una presunta cospirazione a Roma.

Luoghi

- **Tunisi:** la destinazione di G. B. di Monteburgo.
- **Bona e altri luoghi subditi del re:** luoghi sotto il controllo del re menzionato.
- **Roma:** luogo dove alcuni individui tentano di organizzare un insulto.
- **Tunice:** ulteriore riferimento, presumibilmente una variante del nome di Tunisi.

Oggetti

- **Lettere di credenziali:** documenti che G. B. di Monteburgo porterà con sé.
- **Galea e altri brigantini:** tipi di imbarcazioni menzionate.
- **Inventario:** menzionato in connessione con l'archayto Gemoxi.
- **Lettere cifrate:** documenti menzionati come prova di una cospirazione.

Proprietà

- **Prudenza, capacità di parlare e tacere, e di governare:** qualità attribuite a G. B. di Monteburgo.
- **Sospetto e dolore:** sentimenti espressi riguardo alle azioni dei Genovesi.
- **Sicurezza e integrità:** proprietà delle relazioni e della pace tra Genova e il regno di Tunisi.
- **Lealtà e deviazione:** caratteristiche attribuite ai Genovesi rispetto al loro comportamento.

Figura 27. Divisione delle entità nominate per classi (Soggetti, Luoghi, Oggetti, Proprietà).

Successivamente, è stato costruito il prompt relativo alla creazione del file CSV. Anche in questo caso, dal semplice comando «Crea file CSV con i dati estratti dal testo analizzato», è stato necessario costruire un prompt specifico, in quanto l'IA organizzava il foglio in maniera non funzionale alla compilazione del record per il database finale. La maggior parte degli *outcomes*, infatti, si limitava alla semplice lista dei nomi e la relativa categoria, oppure – a causa del fenomeno di “allucinazione” – alla creazione di una vera e propria immagine, che rappresentava un foglio CSV (Fig. 28).

python Mostra sempre dettagli Copia

```
# Creating the DataFrame with the specified columns for the new document content
data_1498 = {
    "Nome Entità Italiano Corrente": [
        "Jacopo Centurione", "Consiglio di Genova", "Pietro Paolo dal Fiesco", "Re di T
    ],
    "Nome nel Documento": [
        "Jacopo Centuriono", "Augustinus et Consilium", "Pietro Paolo dal Fiesco", "Ser
    ],
    "Tipo di Entità": [
        "person", "person", "person", "person", "place", "place", "physical object"
    ],
    "Professione": [
        "Console", "Membri del consiglio", "", "", "", ""
    ],
    "Ruolo e Funzione": [
        "Console a Tunisi", "Governo di Genova", "Mercante detenuto", "Capo di stato",
    ],
    "Luogo di Origine": [
        "Genova", "Genova", "Genova", "", "", ""
    ],
    "Luogo dove si Trova": [
        "Tunisi", "Genova", "Tunisi", "Tunisi", "", ""
    ]
}

# Create DataFrame
df_1498 = pd.DataFrame(data_1498)

# Save DataFrame to Excel file
file_path_1498 = "/mnt/data/Entita_Storiche_1498.xlsx"
df_1498.to_excel(file_path_1498, index=False)

file_path_1498
```

Risultato

'/mnt/data/Entita_Storiche_1498.xlsx'

Ho creato il file Excel con le entità menzionate nel documento del 1498, classificate e dettagliate secondo le tue istruzioni. Puoi scaricare il file da questo link:
[Scarica il file Excel delle Entità Storiche 1498 \[-\]](#)

Figura 30. Esecuzione in codice del prompt per la costruzione e compilazione del file CSV.

Il lavoro si completa con il comando in cui si istruisce l’IA alla compilazione di un “registro”, partendo da un breve training che obbligava il modello a tenere fermi dei limiti di carattere/parole, e normalizzare il testo, al fine di costruire un registro in linguaggio moderno (Fig. 31).

cosa è un registro

Il termine "registro" non è comunemente usato o riconosciuto come parte del vocabolario standard italiano. Potresti aver inteso "registro", che si riferisce a un elenco sistematico di informazioni o eventi, tenuto per scopi ufficiali o amministrativi. Se "registro" si riferisce a qualcosa di specifico che non ho considerato, potresti fornire ulteriori dettagli o contesto?

il registro è il riassunto del contenuto di un documento archivistico o di un atto giuridico che assume la valenza di sintesi più o meno ampia.

Figura 31. Training per la compilazione del registro.

Costruito il prompt, quest'ultimo è stato testato sul testo che viene di seguito riportato – in formato immagine, per garantire la riproducibilità del test –, in cui sono presenti anche errori di trascrizione, per verificare, contestualmente, la capacità del modello ChatGPT-4 di normalizzare il testo e correggere i vari errori (Figg. 32, 33):

«In eterni Dei nomine amen. Magnificus dominus Anthoniotus Adurnus pro serenissimo domino rege Francorum gubernator Januensis et communis et populi defensor. Et suum venerabile consilium dominorum decem octo sapientium antianorum. In pleno, integro et totali numero congregatorum, quorum nomina sunt hec: Dom inus Antonius Justinianus miles, prior; Gentilis de Grimaldis; Raffus Lecavellum; Antonius de Lacastanea; Leonel de M ari; Janinus de Serra de Pulciferia; Raffael de Facio; Abraymus Pilavicinus; Johannes Ususm aris Petri; Lucianus Spinula Cepriani; Andreas M am illus; Francus de Francis; Jacobus de Auria; Dom inicus Bosonus de Struppa; Cataneus Cigala; Jacobus de S alv o; Inoffinos de Solario de Cogoletto et Petrus de Vivaldis. Agentes nomine et vice communis fanue et pro ipso communi. Confisi de discrezione et probitate nobilis viri Caroli G rilli civis Januensis. Omni via, iure, modo et forma, quibus m elius potuerunt et possunt, fecerunt, constituerunt, creaverunt et ordinaverunt eorum dicto nomine et dicti communis Janue certum verum legitimum et indubitatum actorem, syndicum, procuratorem et ambaxatorem ac nuncium specialem, et quidquid et prout de jure melius fieri et esse potesi loco ipsorum dicto nom ine et dicti communis posuerunt et ponunt dictum Carolum G rillum, licet absentem, duraturum usque ad annum unum proxime venturum. A d eundum et se personaliter conferendum ad sublimem conspectum serenissimi et illustrissimi principis et domini domini Mulev Buffers (i), Dei gratia regis Tunexis et Buzee etc., et coram quibuscumque consiliariis, auditoribus vel officialibus ipsius constitutis vel constituendis. Et ad confirmandum, ratificandum et approbandum ac de novo, si expedit, faciendum pacem, compositionem seu conventionem et pacta, vigentem et regentia inter dictum serenissimum regem si/e digne memorie quondam dominum... (2) olim regem patrem suum, ex una parte, et dictos constituentes dicto nomine sive dictum commune Janue, ex altera, sub illis pactis, promissionibus, obligationibus et cautellis, sub quibus fuit ultimo dicta pax et conventio com posita et firmata, existentibus ambaxatoribus pro dicto communi Janue ad predictam nobilibus et discretis viris Gentile de G rim aldis et Luchino de Bonavey. E t ad petendum, requirendum et postulandum restitutionem seu satisfactionem et emendam omnium et singulorum dampnorum, robariarum et depredacionum Januensibus seu subditis communis Janue illatorum seu commissarum contra aliquos Januenses vel eorum bona per quosvis subditos prefati serenissimi domini regis Tunexis a tempore dicte alias firmate pacis per supradictos Gentilem et Luchinum ambaxatores dicti communis Janue citra. Nec non ad requirendum, instandum et prosequendum liberationem et relaxationem quorumcumque in regno ipsius captivorum Januer.sium vel subditorum communis Janue, tam qui restarent dicto tempore firmate dicte pacis captivi, quam qui postea captivati fuissent. Item ad quitandum, liberandum et absolvendum, nomine dictorum constituentium et dicti communis, prefatum serenissimum dominum regem, heredes et successores eiusdem, de omni eo et toto quod pro emendea, restitutione vel alia satisfactione dictorum dampnorum Januensibus illatorum receperit vel habuerit, seque de eo bene quietum et solum vocandum nomine sepedicto. Iit ad unum et plura instrumentum et instrumenta de et super predictis omnibus et singulis. Cum illis omnibus confessionibus, renuntiationibus, promissionibus, obligationibus, clausulis et cautellis, de quibus dicto sindaco ambaxatori et procuratori dicto nomine videbitur et placuerit faciendum et fieri seu confici faciendum. Et demum generaliter ad omnia et singula gerendum, faciendum, procurandum et administrandum in predictis omnibus et singulis, et in dependentibus, accessoriis, annexis et connexis predictis et a predictis, et cuilibet et a quolibet predictorum, et circa predicta que opportuna necessaria vel utilia videantur, queque per quemcumque verum, legitimum et indubitatum procuratorem et syndicum plena et omnimoda potestate sussultum fieri possent, et que ipsimet constituentes possent facere si personaliter interessent, etiam si talia forent que mandatum exigenter speciale. Dantes et concedentes dicto nomine eidem Carolo, sindaco et procuratori predicto, in predictis omnibus et singulis et circa predicta, ac in dependentibus, accessoriis, annexis et connexis predictis et a predictis, et cuilibet et a quolibet predictorum, plenum largum libicum et generale mandatum cum plena larga libera et generali administratione. Promittentes michi Antonio de Credentia notario et dicti com munis Janue cancellario infrascripto, tamquam publice persone officio publico stipulanti et recipienti nomine et vice prefati sci diissim i domini regis Tunexis et heredum et successorum suorum et cuius et quorum interest, intererit vel interesse poterit, sese dicto nomine habituros et quod dictum commune habebit et tenebit perpetuo ratum gratum et firmum quidquid per dictum eorum syndicum, ambaxatorem et procuratorem, in predictis om nibus et singulis actum, gestum, vel factum fuerit, seu etiam procuratum. Sub ypotheca et obligatione bonorum ipsorum constituentium dicto nomine seu dicti communis Janue habitorum et habendorum. Mandantes etiam de predictis omnibus confici debere publicum instrumentum per me supradictum Antonium de Credentia not-irium et dictorum constituentium ac dicti communis Janui cancellarium infrascriptum. Actum Janue in palacio communis, millesimet in camera cubiculari dicti magnifici domini gubernatoris iuxta turrim, anno dominice nativitatis millesimo trecentesimo nonagesimo septimo, indictione quarta secundum cursum Janue, die mercurii X III mensis Marci, in sero, presentibus Conrado Mazurro notario et communis Janue cancellario, Antonio de Montaldo d^ Gavio quondam Priani, et A ntonio Pellerano de Camulio dicto longo, macerio dicti magnifici dom ini gubernatoris, testibus ad predicta vocatis specialiter et rogatis».

Figura 32. Documento usato per testare il prompt per la compilazione dei registi, tratto da Marengo (1901, 139).

Ecco il regesto del documento fornito:

Il documento, redatto il 13 marzo 1397 a Genova, sotto l'autorità di Anthoniotus Adurnus, governatore per il re di Francia e difensore del comune e del popolo di Genova, insieme al consiglio di diciotto saggi, dettaglia la nomina di Carlo Grilli come attore legale, sindaco, procuratore e ambasciatore speciale del Comune di Genova. Questo incarico è valido per un anno e conferisce a Grilli il mandato di rappresentare il Comune di Genova nelle trattative per la conferma, la ratifica o la rinnovazione di trattati di pace e altri accordi precedentemente stabiliti con il re di Tunexis, Mulev Buffers, e di negoziare la liberazione di prigionieri genovesi e la restituzione di beni danneggiati o depredati. Il documento stabilisce inoltre che Grilli ha piena autorità e un mandato generale per gestire tutte le questioni necessarie, utili o opportune relative a tali trattative. Antonio de Credentia, notario e cancelliere del Comune di Genova, è incaricato di redigere un atto pubblico relativo a queste nomine e mandati.

bene, adesso riscrivilo in massimo 40 parole

Documento del 13 marzo 1397, Genova: Carlo Grilli nominato procuratore e ambasciatore per trattative di pace e liberazione prigionieri con il re di Tunexis, Mulev Buffers, da Anthoniotus Adurnus, governatore per il re di Francia.

normalizza tutti i nomi

Documento del 13 marzo 1397, Genova: Carlo Grilli nominato procuratore e ambasciatore per trattative di pace e liberazione prigionieri con il re di Tunisi, Mulev Buffers, da Anthoniotus Adurno, governatore per il re di Francia.

Figura 33. *Outcome* della compilazione del regesto.

Ultimati i vari test, il prompt finale, utilizzato nelle varie fasi di estrazione dei dati e compilazione dei records è il seguente (Fig. 34):

crea un breve regesto (di massimo 50 parole) dei documenti che hai analizzato, individua la data e trascrivila separatamente dal regesto, nel formato "anno mese (a lettere in inglese) giorno", ma senza interpunzioni tra anno mese e giorno. Inoltre, non inserire le etichette "Regesto" e "Data del documento", in quanto irrilevanti.

Figura 34. Prompt dettagliato per la compilazione dei regesti.

Il complessivo lavoro di estrazione e compilazione ha consentito di individuare – in un workflow di 450 ore complessive –, su 64 documenti d'archivio trascritti in digitale, 754 *entities*:

- 406 persone
- 8 gruppi
- 3 istituzioni
- 2 organizzazioni
- 144 oggetti
- 198 luoghi

- 20 proprietà
- 2 n.c.

Tali entità, nel corpo del file CSV finale (Figg. 35, 36), vengono organizzate sulla base di un ID univoco, che consente, da un lato, di individuare correttamente l'info richiesta, e, da un altro lato, permettere al sistema “Arches” di disporre la visualizzazione delle informazioni, per i web users, in maniera univocamente correlata.

ID	COLLOCAZIONE	DATA ROGITO	LUOGO ROGITO	OSSGITO DELL'ATTO	REPERITO
1	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 A	1397 marzo 14	Genova	Mandatam Carlo Gridi Tondet proferitur pro concordando patto cum Regi.	Il documento emanato a Genova sotto Anthonianus Adunius, governatore per il Re di Francia, designa Carlo Gridi come attore legale, sindaco, e La lettera, scritta dal Cardinale e al Consiglio, il Re di Tunisi tratta questioni commerciali e diplomatiche, segnalando il disprezzo per la sua
2	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 B	1427 maggio 24	Genova	Lettera del Governatore di Genova al Re di Tunisi.	La lettera, scritta dal Cardinale e al Consiglio, il Re di Tunisi tratta questioni commerciali e diplomatiche, segnalando il disprezzo per la sua
3	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 C	1432 ottobre 15	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento emanato a Genova sotto Anthonianus Adunius, governatore per il Re di Francia, designa Carlo Gridi come attore legale, sindaco, e La lettera, scritta dal Cardinale e al Consiglio, il Re di Tunisi tratta questioni commerciali e diplomatiche, segnalando il disprezzo per la sua
4	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 D	1435 ottobre 14	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento emanato a Genova sotto Anthonianus Adunius, governatore per il Re di Francia, designa Carlo Gridi come attore legale, sindaco, e La lettera, scritta dal Cardinale e al Consiglio, il Re di Tunisi tratta questioni commerciali e diplomatiche, segnalando il disprezzo per la sua
5	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 E	1436 marzo 3	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Lettera scritta dal Capitano generale della libertà e del Consiglio degli Anciani, inviata al Re di Tunisi, espone la recente liberazione di Genova al
6	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 F	1447 dicembre 12	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
7	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 G	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
8	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 H	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
9	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 I	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
10	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 J	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
11	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 K	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
12	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 L	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
13	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 M	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
14	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 N	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
15	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 O	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
16	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 P	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
17	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 Q	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
18	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 R	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
19	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 S	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese
20	Archivio di Stato di Genova, Sala 58, busta 10/10, n. 2707 T	1453 gennaio 30	Genova	Lettera del Governatore di Genova al Re di Tunisi.	Il documento riguarda la cessione di un titolo a Tunisi, dettagliando la procedura e le azioni per la gestione e la liberazione del suo paese

Figura 35. Struttura finale del file CSV del dataset (foglio relativo alla “Collocazione”).

ID	has name	has name from source	has type	has occupation	has social category	has provenance	has place from source	has role	in contract
1	Agostino Lombardi	notari Agostino Lombardi	person	notario	artefice	notario	notario	artefice	A001
2	Simone Caddini	Simone Caddini fabris	person	fabbrico	artefice	artefice	artefice	artefice	A001
3	taoletta	una tetta parva	physical object				marco	marco	A001
4	taoletta	carta map minuscule pauci valoris	physical object				marco	marco	A001
5	taoletta	una tetta parva	physical object				marco	marco	A001
6	taoletta	carta una e navigandi	physical object				marco	marco	A001
7	manuale	manuale due scriptis	physical object				marco	marco	A001
8	Genova	ciuitate lanus	place				lungo	lungo	A001
9	Lodovico de Boccio	Lodovico de Boccio notari	person	notario	artefice	artefice	artefice	artefice	A002
10	Bartholomeo Riccio	Bartholomeo Riccio notari	person	notario	artefice	artefice	artefice	artefice	A002
11	Angelo de Negro	Angelo de Negro	person	notario	artefice	artefice	artefice	artefice	A002
12	Casa di Angelo de Negro	scab domo Angelo De Negro et strata	property				lungo di emanazione	lungo	A002
13	the auto-dictated (conestaglie) di Branci	lanus, in branci	place				lungo	lungo	A002
14	Pietro de Torsani	Pietro de Torsani iacobis, lanus	person				lungo	lungo	A002
15	Giuseppe de Torsani	de Torsani iacobis	person				lungo	lungo	A002
16	Infante Don Fernando	illustratissimi principis domini infantis doni Ferendis	person				lungo	lungo	A002
17	Genova	lanus	place				lungo	lungo	A002

Figura 36. Struttura finale del file CSV del dataset (foglio relativo alle entità nominate).

4. Conclusioni

La costruzione del “Database on the Slave Trade between the Mediterranean and the Atlantic (15th-16th centuries)” si colloca al crocevia tra innovazione tecnologica e riflessione epistemologica, restituendo centralità all'*histoire sérielle* come metodo e rilanciando la funzione dell'archivio come dispositivo cognitivo, interpretativo e politico. L'introduzione di strumenti di Intelligenza Artificiale generativa nei processi di estrazione, codifica e strutturazione dei dati documentari non rappresenta semplicemente un salto tecnologico, ma impone una ridefinizione profonda della semantica archivistica, delle pratiche storiografiche e delle modalità di costruzione della memoria storica. In questa prospettiva, gli *invented archives* non sono solo ambienti digitali, ma spazi di negoziazione epistemica, in cui i dati sono selezionati, modellati e interrogati in relazione a specifici bisogni interpretativi, teorici e politici.

L'esperienza condotta ha mostrato come i modelli di IA, se opportunamente addestrati e guidati da prompt costruiti con rigore metodologico, siano oggi in grado di affrontare le complessità linguistiche, strutturali e semantiche della

documentazione storica. Tuttavia, le criticità emerse – dalla possibilità di generare output allucinati, all’opacità della logica inferenziale dei modelli – rendono evidente la necessità di mantenere un controllo umano costante, capace di orientare l’interazione uomo-macchina entro un quadro di responsabilità epistemica. L’intervento dello storico, in tal senso, non può essere surrogato dall’automazione: è nella progettazione ontologica, nella calibrazione dei criteri di estrazione, nella verifica contestuale e nell’analisi qualitativa che si gioca la scientificità del risultato.

Al tempo stesso, il progetto ha rivelato le ambivalenze di un’archivistica digitale ancora parzialmente ancorata a logiche indicizzative tradizionali, incapaci di attivare pienamente le potenzialità computazionali offerte dalla Digitalità. La trasformazione del documento in *cybertesto* (Aarseth 1997), la costruzione di ambienti interrogabili e relazionali, l’integrazione tra approccio seriale e Machine Learning, sono obiettivi ancora da consolidare, soprattutto in una prospettiva disciplinare in cui Storia e Archivistica, spesso, procedono su binari paralleli. In questo contesto, l’Intelligenza Artificiale si configura come agente tecnico, ma anche come oggetto critico, capace di rimettere al centro il dibattito su autorità archivistica, gerarchie semantiche, invisibilità strutturale e *agency* dei soggetti rappresentati.

Il database, pensato come piattaforma aperta, scalabile e FAIR-oriented, intende contribuire anche a tale ripensamento, offrendo uno strumento per la riconfigurazione storica e computazionale dei legami tra Africa, Mediterraneo e Atlantico nella lunga durata del primo capitalismo globale. Restituire visibilità a soggetti subalterni, riconoscere il ruolo strategico delle reti finanziarie e mercantili premoderne, valorizzare la complessità documentaria delle fasi arcaiche della tratta schiavista: sono questi i compiti di una Digital History che, lungi dall’essere ancella della tecnica, si pone come frontiera metodologica ed etica, orientata alla costruzione critica della memoria e alla trasformazione consapevole degli strumenti della conoscenza storica.

Bibliografia

- Aarseth, Espen J. J. 1997. *Cybertext: Perspectives on Ergodic Literature*. Baltimore: Johns Hopkins University Press.
- Agosto, Aldo. 1998. *Christopher Columbus and His Family: The Genoese and Ligurian Documents*. Turnhout, Belgium: Brepols Publishers.

- Ahmadi-Assalemi, Gabriela, Haider Al-Khateeb, Carsten Maple, et al. 2020. "Digital Twins for Precision Healthcare." In *Cyber Defence in the Age of AI, Smart Societies and Augmented Humanity*, edited by Hamid Jahankhani, Stefan Kendzierskyj, Nishan Chelvachandran, and Jaime Ibarra, 133-58. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-35746-7_8.
- Armenteros-Martinez, Ivan. 2022. "Enslaved Africans in Medieval and Early Modern Iberia." In *Oxford Research Encyclopedia of African History*. <https://doi.org/10.1093/acrefore/9780190277734.013.864>.
- Ayers, Edward L., and Anne S. Rubin. 2000. *The Valley of the Shadow. Two Communities in the American Civil War - The Eve of War*. Har/Cdr edizione. New York: WW Norton & Co. Inc.
- Bail, Christopher A. 2014. "The Cultural Environment: Measuring Culture with Big Data." *Theory and Society* 43 (3-4): 465-82.
- Banfi, Fabrizio, Elena Dellù, Chiara Stanga, et al. 2023. "Representing Intangible Cultural Heritage of Humanity: From the Deep Abyss of the Past to Digital Twin and XR of the Neanderthal Man and Lamalunga Cave (Altamura, Apulia)." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLVIII-M-2-2023 (giugno):171-81. <https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-171-2023>.
- Bleichmar, Daniela, and Vanessa R. Schwartz. 2019. "Visual History. The Past in Pictures." *Representations* 145 (1): 1-31. <https://doi.org/10.1525/rep.2019.145.1.1>.
- Bloch, Marc. 1963. *The Historian's Craft: Ou metier d'historien*. Tradotto da Peter Putnam. New York: Knopf.
- Bonazza, Giulia. 2023. "Slavery in the Mediterranean." In *The Palgrave Handbook of Global Slavery throughout History*, edited by Damian A. Pargas and Juliane Schiel, 227-42. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-13260-5_13.
- Bono, Salvatore. 2017. "Schiavitù mediterranea." *Nuova informazione bibliografica* 56 (4): 675-98. <https://doi.org/10.1448/87956>.
- Bradley, Adam James, Mennatallah El-Assady, Katharine Coles, et al. 2018. "Visualization and the Digital Humanities." *IEEE Computer Graphics and Applications* 38 (6): 26-38. <https://doi.org/10.1109/MCG.2018.2878900>.
- Braudel, Fernand. 2003. *Scritti sulla storia*. Milano: Bompiani.
- Braudel, Fernand. 2015. *Storia, misura del mondo*. Il Mulino.
- Braudel, Fernand. 2017. *La Méditerranée*. Paris: FLAMMARION.
- Braudel, Fernand, e Alfredo Salsano. 1974. *La storia e le altre scienze sociali*. Roma: Laterza.

- Brown, Vincent. 2021. "Slave Revolt in Jamaica, 1760-1761." <http://revolt.axismaps.com/>.
- Bush, Vannevar. 1945. "As We May Think." *Atlantic* 176 (1).
- Chatbox AI. n.d. Consultato il 19 maggio 2025. <https://chatboxai.app/it-IT>.
- Chaunu, Pierre. 1959. *Séville et l'Atlantique, 1504-1650: Structures et conjoncture de l'Atlantique espagnol et hispano-américain (1504-1650). Tome II, volume 1 : La conjoncture (1504-1592)*. 12 voll. Travaux et mémoires. Paris: SEVPEN. <https://books.openedition.org/iheal/5807>.
- Ciofalo, Giovanni, e Silvia Leonzi. 2013. *Homo communicans. Una specie dilin evoluzione*. Roma: Armando Editore.
- Clerici, Alberto, e Maurizio De Pra. 2012. *Informatica e web*. EGEA spa.
- Correia, Luiz Felipe Ribeiro, Roberto dos Santos Bartholo, Aline Winckler Brufato, and Edney Christian Thomé Sanchez. 2023. "Toward a Digital Twin for Cultural Heritage." In *Advances in Tourism, Technology and Systems: Selected Papers from ICOTTS 2022, Volume 1*, edited by João Vidal Carvalho, António Abreu, Pedro Liberato, and Alejandro Peña, 419-30. Singapore: Springer Nature. https://doi.org/10.1007/978-981-99-0337-5_35.
- Darcy, Robert, e Richard C. Rohrs. 1995. *A Guide to Quantitative History*. Westport, Conn.: Praeger.
- Delpiano, Patrizia. 2021. *La schiavitù in età moderna*. Gius. Laterza & Figli Spa.
- Digital Library. 2021. "Principi F.A.I.R. – Interoperabilità." 28 Ottobre. <https://digitallibrary.cultura.gov.it/notizie/principi-f-a-i-r-interoperabilita/>.
- Dollar, Charles M., and Richard J. Jensen. 1974. *Historian's Guide to Statistics; Quantitative Analysis and Historical Research*. Huntington, N.Y.: R.E. Krieger Pub. Co.
- Ebbrecht-Hartmann, Tobias, Noga Stiassny, and Lital Henig. 2023. "Digital visual history: historiographic curation using digital technologies." *Rethinking History* 27 (2): 159-86. <https://doi.org/10.1080/13642529.2023.2181534>.
- Eijnatten, Joris van, Toine Pieters, and Jaap Verheul. 2013. "Big Data for Global History: The Transformative Promise of Digital Humanities." *BMGN - Low Countries Historical Review* 128 (4): 55-77. <https://doi.org/10.18352/bmgn-lchr.9350>.

- El Haoud, Naima, and Zineb Bachiri. 2019. "Stochastic Artificial Intelligence benefits and Supply Chain Management inventory prediction." In *2019 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA)*, 1-5. <https://doi.org/10.1109/LOGISTIQUA.2019.8907271>.
- El Saddik, Abdulmotaleb. 2018. "Digital Twins: The Convergence of Multimedia Technologies." *IEEE MultiMedia* 25 (2): 87-92. <https://doi.org/10.1109/MMUL.2018.023121167>.
- Eltis, David, and David Richardson. 2008. *Extending the Frontiers: Essays on the New Transatlantic Slave Trade Database*. Yale University Press.
- Ennals, Richard. 1985. *Artificial Intelligence: Applications to Logical Reasoning and Historical Research*. Ellis Horwood.
- Erwin, Brittany. 2020. "Digital Tools for Studying Empire: Transcription and Text Analysis with Transkribus." *Not Even Past*, 6 November 2020. <https://notevenpast.org/digital-tools-for-studying-empire-transcription-and-text-analysis-with-transkribus/>.
- Fiume, Giovanna. 2012. *Schiavitù Mediterranee. Corsari, Rinnegati e Santi Di Età Moderna*. Bruno Mondadori.
- Franzosi, Roberto. 2017. "A third road to the past? Historical scholarship in the age of big data." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 50 (4): 227-44. <https://doi.org/10.1080/01615440.2017.1361879>.
- Fuller, Aidan, Zhong Fan, Charles Day, and Chris Barlow. 2020. "Digital Twin: Enabling Technologies, Challenges and Open Research." *IEEE Access* 8:108952-71. <https://doi.org/10.1109/ACCESS.2020.2998358>.
- Furet, François. 1968. «Sur quelques problèmes posés par le développement de l'histoire quantitative.» *Social Science Information* 7 (1): 71-82.
- Furet, François. 1971. «L'histoire quantitative et la construction du fait historique.» *Annales. Histoire, Sciences Sociales* 26 (1): 63-75.
- Furet, François. 1981. «Il quantitativo in Storia.» In *Fare storia. Temi e metodi della nuova storiografia*, a cura di Jacques Le Goff e Pierre Nora. Torino: Einaudi.
- Gabellone, Francesco. 2022. "Digital Twin: A New Perspective for Cultural Heritage Management and Fruition." *Acta IMEKO* 11 (1): 1-7. https://doi.org/10.21014/acta_imeko.v11i1.1085.
- Galić, Radoslav, Elvir Čajić, Zvezdan Stojanovic, and Dario Galić. 2023. "Stochastic Methods in Artificial Intelligence." *Research Square*. <https://doi.org/10.21203/rs.3.rs-3597781/v1>.

- Gautier, Dassonneville, Adèle Huguet, Marie-Laure Massot, et al. 2022. «Compte-rendu de la journée d'étude 'Point HTR 2022' Transkribus / eScriptorium : Transcrire, annoter et éditer numériquement des documents d'archives.» Report, CAPHES - UMS 3610 CNRS/ENS; AOROC. <https://hal.science/hal-03692413>.
- Gioffrè, Domenico. 1962. "Il commercio d'importazione genovese alla luce dei registri del dazio, 1495-1537." In *Studi in onore di Amintore Fanfani*, vol. 5. Giuffrè.
- Gioffrè, Domenico. 1971. *Il mercato degli schiavi a Genova nel secolo XV*. Bozzi.
- Github. n.d. "Ollama." Consultato il 19 maggio 2025. <https://ollama.com/>.
- Gonzalez, Teofilo, Jorge Diaz-Herrera, and Allen Tucker. 2014. *Computing Handbook, Third Edition: Computer Science and Software Engineering*. CRC Press.
- Graham, Shawn, Ian Milligan, and Scott Weingart. 2015. *Exploring Big Historical Data. The Historian's Macroscopic*. Imperial College Press. <https://www.perlego.com/book/839949/exploring-big-historical-data-the-historians-macroscopic-pdf>.
- Grievies, Michael W. 2023. "Digital Twins: Past, Present, and Future." In *The Digital Twin*, edited by Noel Crespi, Adam T. Drobot, and Roberto Minerva, 97-121. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-21343-4_4.
- Hutson, James, Joe Weber, and Angela Russo. 2023. "Digital Twins and Cultural Heritage Preservation: A Case Study of Best Practices and Reproducibility in Chiesa dei SS Apostoli e Biagio." *Art and Design Review* 11 (February). <https://doi.org/10.4236/adr.2023.111003>.
- Itzcovich, Oscar. 1989. "Lo storico e il database." *Quaderni storici* 24 (70 (1)): 321-25.
- Itzcovich, Oscar. 1993. *L'uso del calcolatore in storiografia*. Milano: FrancoAngeli.
- Itzcovich, Oscar. 1996. "Dal mainframe al personal, il computer nella storiografia quantitativa." *Storia & computer*.
- Jupyter. n.d. Consultato il 19 maggio 2025. <https://jupyter.org>.
- Kahle, Philip, Sebastian Colutto, Günter Hackl, and Günter Mühlberger. 2017. "Transkribus. A Service Platform for Transcription, Recognition and Retrieval of Historical Documents." In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, 19-24. <https://doi.org/10.1109/ICDAR.2017.307>.
- Kaplan, Frédéric, and Isabella di Lenardo. 2017. "Big Data of the Past." *Frontiers in Digital Humanities* 4. <https://doi.org/10.3389/fdigh.2017.00012>.

- Kennedy, S. Wright, Jessica C. Kuzmin, and Benjamin Jones. 2017. "New Methods in the History of Medicine: Streamlining Workflows to Enable Big-Data History Projects." *Medical History* 61 (3): 477-80. <https://doi.org/10.1017/mdh.2017.54>.
- Kiessling, Benjamin, Robin Tissot, Peter Stokes, and Daniel Stökl Ben Ezra. 2019. "eScriptorium: An Open Source Platform for Historical Document Analysis." In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, 2:19-19. <https://doi.org/10.1109/ICDARW.2019.10032>.
- Lu, Jessica H., and Caitlin Pollock. "Design, Development, and Documentation: Hacking TEI for Black Digital Humanities." Maryland Institute for Technology in the Humanities (MITH). <https://mith.umd.edu/digital-dialogues/dd-fall-2019-jessica-h-lu-caitlin-pollock/>.
- Maleki, Negar, Balaji Padmanabhan, and Kaushik Dutta. 2024. "AI Hallucinations: A Misnomer Worth Clarifying." In *2024 IEEE Conference on Artificial Intelligence (CAI)*, 133-38. <https://doi.org/10.1109/CAI59869.2024.00033>.
- Marengo, Emilio. 1901. "Genova e Tunisi (1388-1515). Relazione storica del socio avv. Emilio Marengo, sotto-archivista nel R. Archivio di Stato di Genova." In *Atti della Società Ligure di Storia Patria*, XXXII. Genova.
- Massot, Marie-Laure, Arianna Sforzini, and Vincent Ventresque. 2019. "Transcribing Foucault's handwriting with Transkribus." *Journal of Data Mining and Digital Humanities Atelier Digit_Hum* (marzo). <https://hal.archives-ouvertes.fr/hal-01913435>.
- McCarthy, John, Nathaniel Rochester, Marvin L. Minsky, and Claude E. Shannon. 2006. "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955." *AI Magazine* 27 (4). <https://doi.org/10.1609/aimag.v27i4.1904>.
- Milioni, Nikolina. 2020. "Automatic Transcription of Historical Documents. Transkribus as a Tool for Libraries, Archives and Scholars." Student Thesis, Master's Programme in Digital Humanities. <http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-412565>.
- Muehlberger, Guenter, Louise Seaward, Melissa Terras, et al. 2019. "Transforming scholarship in the archives through handwritten text recognition. Transkribus as a case study." *Journal of Documentation* 75 (5): 954-76. <https://doi.org/10.1108/JD-07-2018-0114>.
- National African-American Reparations Commission, n.d. "The Atlantic Slave Trade in Two Minutes." Consultato il 19 maggio 2025. <https://reparationscomm.org/reparations-news/atlantic-slave-trade-in-two-minutes/>.

- North American Slave Narratives. n.d. Consultato il 19 maggio 2025. <https://docsouth.unc.edu/neh/index.html>.
- Northeastern University Library. n. d. "Early Caribbean Digital Archive." Consultato il 19 maggio 2025. <https://ecda.northeastern.edu/#fl-main-content>.
- Nowatzki, Robert. 2021. "From Datum to Databases: Digital Humanities, Slavery, and Archival Reparations." *The American Archivist* 83 (2): 429-48. <https://doi.org/10.17723/0360-9081-83.2.429>.
- Rabus, Achim. 2019. "Recognizing Handwritten Text in Slavic Manuscripts, a Neural-Network Approach Using Transkribus." *Scripta & E-Scripta*, fasc. 19, 9-32.
- Rau, Virginia. 1957. "A family of Italian merchants in Portugal in the XVth century: the Lomellini." In *Studi Armando Saporì*, 1:715-26.
- READ-COOP. s.d. "Transkribus." Consultato 30 agosto 2022. <https://readcoop.eu/it/transkribus/>.
- Reilly, Edwin D., Anthony Ralston, and David Hemmendinger. 2000. *Encyclopedia of Computer Science*. Nature Publishing Group.
- Rosenzweig, Roy. 1997. "Brave New World or Blind Alley? American History on the World Wide Web." *Journal of American History* 84 (1): 132-55.
- Rosenzweig, Roy. 2001. "The Road to Xanadu: Public and Private Pathways on the History Web." *Journal of American History* 88 (2): 548-79. <https://doi.org/10.2307/2675105>.
- Ruis, Andrew R., and David Williamson Shaffer. 2017. "Annals and Analytics: The Practice of History in the Age of Big Data." *Medical History* 61 (1).
- Salvagno, Michele, Fabio Silvio Taccone, and Alberto Giovanni Gerli. 2023. "Artificial Intelligence Hallucinations." *Critical Care* 27 (1): 180. <https://doi.org/10.1186/s13054-023-04473-y>.
- Santoro, Raffaele. 2015. "I grandi archivi: un patrimonio di Big Data." In *La grande bellezza dei grandi dati. Frontiere di fruizione del patrimonio archivistico per studiosi e società civile*. Roma: Archivio di Stato di Roma.
- Schiuma, Giovanni, and Daniela Carlucci. 2018. *Big Data in the Arts and Humanities. Theory and Practice*. Auerbach Publishers.
- Schlagdenhauffen, Régis. 2020. "Optical Recognition Assisted Transcription with Transkribus: The Experiment concerning Eugène Wilhelm's Personal Diary (1885-1951)." *Journal of Data Mining & Digital Humanities Atelier Digit_Hum*. <https://doi.org/10.46298/jdmdh.6249>.
- Schwartz, Stuart B. 2004. *Tropical Babel: Sugar and the Making of the Atlantic World, 1450-1680*. Univ of North Carolina Press.

- Sgroi, Salvatore. 1979. Recensione di *La schiavitù [domestica] in Sicilia nel XVI secolo*, Corrado Avolio. Catania, Bacina 1978.
- Sicking, Louis, and Alain Wijffels. 2020. "Introduction Flotsam and Jetsam in the Historiography of Maritime Trade and Conflicts." In *Conflict Management in the Mediterranean and the Atlantic, 1000-1800: Actors, Institutions and Strategies of Dispute Settlement*. Brill Nijhoff. <https://brill.com/display/title/39048>.
- Slave Voyages. n.d. "Explore the Origins and Forced Relocations of Enslaved Africans Across the Atlantic World." Consultato il 19 maggio 2025. <https://www.slavevoyages.org/>.
- Spina, Salvatore. 2020. *Archivi nell'era delle Digital Humanities, dei Big Data e della Genetica*. Viagrande (Catania): Algra.
- Spina, Salvatore. 2021. "Datificazione delle fonti storiche per la digital history delle pandemie." *Umanistica Digitale* 10.
- Spina, Salvatore. 2022a. *Digital History. Metodologie informatiche per la ricerca storica*. Napoli: Edizioni Scientifiche Italiane.
- Spina, Salvatore. 2022b. "Historical Network Analysis & Htr Tool. Per un approccio storico metodologico digitale all'archivio Biscari di Catania." *Umanistica Digitale*, fasc. 14, 163-81. <https://doi.org/10.6092/issn.2532-8816/15159>.
- Spina, Salvatore. 2023a. "Artificial Intelligence in archival and historical scholarship workflow: HTR and ChatGPT." *Umanistica Digitale*. <https://doi.org/10.48550/arXiv.2308.02044>.
- Spina, Salvatore. 2023b. "Biscari Epistology. From Archive to the Website." *DigItalia* 2.
- Spina, Salvatore. 2024. "Digitality as a longue durée historical phenomenon." *Umanistica Digitale* 18:1-25.
- Staley, David J. 2015. *Computers, visualization, and History. How New Technology Will Transform Our Understanding of the Past*. 1 voll. London and New York: Routledge, Taylor & Francis Group.
- Taviani, Carlo. 2024. "The Genoese Merchant Network and Gold (Ea.1450-1530)." In *Global Gold: Aesthetics, Material Desires, Economies in the Late Medieval and Early Modern World*, edited by Thomas B. F. Cummins. Florence: Villa I Tatti.
- The Maryland State Archives, n.d. "Legacy of Slavery in Maryland." Consultato il 19 maggio 2025. <https://slavery.msa.maryland.gov/>.
- Theibault, John. 2016. "Visualizations and Historical Arguments." In *Writing History in the Digital Age*, edited by Jack Dougherty and Kristen Nawrotzki. University of Michigan Press.

- Trasselli, Carmelo. 1969. "Genovesi in Sicilia." In *Atti della Società Ligure di Storia Patria, nuova serie*, 153-78. Genova: Società Ligure di Storia Patria. https://www.storiapatriagenova.it/Scheda_vs_info.aspx?Id_Scheda_Bibliografica=711.
- Trasselli, Carmelo. 1972. "Considerazioni sulla schiavitù in Sicilia alla fine del Medioevo." *Clio* 8:67-90.
- Trinkle, Dennis A, and Scott A. Merriman. 2000. *The History Highway 2000, a Guide to Internet Resources*. Armonk, N.Y.: M.E. Sharpe.
- Verlinden, Charles. 1970. *The Beginnings of Modern Colonization: Eleven Essays*. Ithaca: Cornell Univ Pr.
- Wang, Fuzhang, Ayesha Sohail, Wing-Keung Wong, Qurat Ul Ain Azim, Shabieh Farwa, and Maria Sajad. 2023. "Artificial intelligence and stochastic optimization algorithms for the chaotic datasets." *Fractals* 31 (06): 2240175. <https://doi.org/10.1142/S0218348X22401752>.
- World Wide Web. n.d. Consultato il 19 maggio 2025. <http://info.cern.ch/hypertext/WWW/TheProject.html>.
- Xu, Zhiwei, and Jialin Zhang. 2022. *Computational Thinking: A Perspective on Computer Science*. Springer Nature.

AIDainformazioni

Rivista semestrale di Scienze dell'Informazione

Anno 42

N. 3-4 – luglio-dicembre 2024

Contributi

MONICA CONSOLANDI, SIMONE
MAGNOLINI, MAURO DRAGONI

Risk Communication and
Misunderstandings. The Healthcare
Dominion Case

CAROLINE DJAMBIAN

Enhancing technoscientific heritage
and introducing concepts into
culture in the age of data, semantic
web, and Artificial Intelligence. The
ITinHeritage project

ERIKA PASCERI

I Knowledge Organization Systems
come strumenti di sorveglianza
sanitaria. L'applicazione della ICD in Italia

GRAZIA SERRATORE

La cartella clinica tra analogico e
digitale: gestione e conservazione a norma

SALVATORE SPINA

Schiavitù e invented archives.

Intelligenza Artificiale generativa nella
costruzione del *Database on the Slave Trade*

Rubriche

ANTONIETTA FOLINO

Recensione del volume *ClaG –
Classificazione dei giochi per ludoteche e
biblioteche*

CLAUDIO GNOLI

Il macellaio pragmatista



mundaneum

In copertina

Disegno di Paul Otlet, Collections Mundaneum, centre d'Archives, Mons (Belgique).

ISBN 979-12-5965-549-3



9 791259 655493

ISSN 1121-0095



9 770112 100950