L'aplicació de la Intel·ligència Artificial a les tesis de l'Institut Nova Història

Autor: Jordi Berenguer
Data de publicació: 17-11-2025

En Jordi Berenguer ens presenta la possibilitat d'obrir una nova etapa en el camp de la recerca històrica crítica mitjançant l'aplicació de mètodes d'intel·ligència artificial a les tesis de l'Institut Nova Història. Aquesta proposta ha estat compartida recentment durant la 12a Universitat Nova Història, en el marc d'un Llevant de taula titulat "Com introduir la IA a la recerca de la Història", moderat per Jordi Manchon i amb la participació de l'autor d'aquest text (que n'és una síntesi), Jordi Berenguer, de Josep Lluís Larriba i de Frederic Monràs.

Des de fa anys, l'Institut Nova Història (INH) ha publicat desenes de llibres i centenars d'articles de recerca històrica, amb hipòtesis innovadores i una gran quantitat de dades. En Jordi Bilbeny ha fet escola i des de fa molts anys són nombrosos els investigadors que s'hi han anat sumant des d'una aproximació multidisciplinar, aportant noves troballes i tesis que han enriquit tot el gruix de recerca de l'Institut. Fins ara, cada investigador ha fet la seva contribució i l'hem anat compartint a través d'articles al web, de la publicació de llibres, de documentals, de conferències, de simposis, i a través d'aquesta universitat d'estiu que enguany ja celebra la seva dotzena edició. Però el repte i la necessitat d'avui és gestionar aquest volum d'informació i generar un diàleg crític tant intern com amb interlocutors externs. Ara, amb l'ajuda de la intel·ligència artificial (IA), aquest corpus documental podrà ser explorat, consultat i visualitzat de manera interactiva, la qual cosa facilitarà la preservació, l'estudi comparatiu i la transmissió de tot aquests materials.

Aquest immens bagatge entra, doncs, en una nova etapa: la integració amb tecnologies d'intel·ligència artificial. No es tracta, però, d'una substitució de la tasca humana, sinó d'un reforç. D'un instrument que ens permet millorar l'organització, l'accés i l'exploració crítica del coneixement acumulat. Gràcies a aquestes tecnologies podrem avançar la feina de molts anys en poc temps. Tanmateix, cal advertir que aquesta és una cursa de fons, una evolució progressiva que exigeix un tractament acurat dels documents, un afinament constant, la supervisió humana i molta feina. El cert és que els fruits que podem obtenir superen en escreix l'esforç i recursos que hi puguem dedicar.

En aquest context, hem iniciat un projecte —actualment en fase experimental— per posar a prova l'aplicació d'eines d'IA sobre els continguts publicats per l'Institut. Aquest projecte parteix de la creació d'un corpus digital estructurat a partir d'articles ja publicats, que han estat ingerits i preparats per a ser entesos pels models de llenguatge. Perquè aquests models puguin generar resultats significatius, cal una feina prèvia de neteja, jerarquització i estandardització del contingut. Un cop estructurat, aquest corpus es transforma en una base de dades indexada amb vectors semàntics —anomenats embeddings— que permeten realitzar cerques ràpides i contextuals. La tecnologia emprada per a aquest fi és FAISS, un sistema d'indexació robust i escalable que garanteix la cerca ràpida i contextual sobre grans volums de dades. Gràcies a això, no només podem identificar les tesis que defensa l'INH, sinó entendre com es connecten les unes amb les altres, detectar arguments recurrents, veure com es desenvolupen al llarg del temps o identificar buits en la recerca.

Un dels aspectes clau per garantir resultats útils és el procés de preparació del corpus abans de ser indexat. Els textos han estat sotmesos a una neteja acurada per eliminar elements no rellevants, com ara codis que interfereixen amb la lectura semàntica, duplicats o seccions redundants, i s'han segmentat per fragments per facilitar la seva indexació semàntica. Allà on ha estat possible, també s'hi ha afegit anotació estructural: identificació de la tesi, arguments, subarguments i altres metadades com autor, títol, data i referències bibliogràfiques. Aquesta estructuració permet que els models entenguin millor les relacions internes dels documents i responguin de manera més precisa i contextualitzada.

La base tecnològica del projecte és un sistema RAG (Retrieval-Augmented Generation), que permet aportar context rellevant als models de llenguatge a partir del corpus digital, sense que aquest corpus s'hagi de transferir ni exposar públicament de forma completa. El sistema consulta el coneixement existent en la base indexada i el presenta al model com a context per generar respostes acurades i justificades. Això garanteix la preservació de la privacitat documental i la sobirania sobre les dades, ja que la IA actua com a generador controlat d'informació a partir de fonts verificables i pròpies, per la qual cosa evita improvisacions sense fonament o respostes desconnectades del treball de recerca acumulat.

Mitjançant l'ús de models de llenguatge de grans dimensions, com ara Mistral-7B o DeepSeek, o els models d'OpenAI, podem identificar automàticament les tesis centrals de cada article, els arguments i subarguments que les sustenten, així com les notes bibliogràfiques o proves documentals citades. Tot aquest coneixement es converteix en metadades essencials per a l'anàlisi posterior. Cal destacar que aquest projecte es fonamenta en la sobirania tecnològica: no volem dependre de cap gran empresa ni de models entrenats amb fonts oficialistes. Fem servir models lliures, que podem executar localment i afinar amb els nostres propis continguts, la qual cosa garanteix un entorn segur, autònom i respectuós amb el coneixement que generem. Per a la representació semàntica dels documents, s'han fet servir embeddings generats mitjançant models de transformadors adaptats a textos en català, optimitzats per a cerques de similitud i recuperació contextual. Actualment, estem utilitzant embeddings basats en Sentence Transformers del repositori HuggingFace, amb entrenament multilingüe i bon rendiment en català. No obstant això, el sistema està preparat per acceptar embeddings més específics si en el futur es fa un afinament (fine-tuning) sobre el corpus propi. Aquesta flexibilitat ens permet adaptar el motor de recuperació segons el creixement i la complexitat del projecte.

Tot el coneixement que es genera a través del sistema d'IA és objecte d'una supervisió humana rigorosa. En cap cas es considerarà la resposta d'un model com a definitiva si no ha estat validada per investigadors amb coneixement del context històric. Les funcionalitats desenvolupades permeten auditar cada resposta generada, i es traçarà l'origen exacte del contingut utilitzat per generar-la. Això facilita una revisió transparent i col·laborativa de les conclusions provisionals, a partir d'evidències documentades. A més, es preveu la incorporació d'un sistema de valoració o correcció per part d'usuaris experts que permeti anar afinant el rendiment i fiabilitat de les respostes generades.

A més de facilitar la consulta i l'exploració del coneixement acumulat, el sistema ens ofereix l'oportunitat de valorar les tesis des d'una nova perspectiva: mitjançant l'anàlisi automatitzada del grau de fonamentació documental i l'estructura argumentativa de cada treball, podrem establir indicadors qualitatius i quantitatius que diferenciïn les tesis més robustes —aquelles que es basen en múltiples raonaments lògics i fonts documentades— de les hipòtesis o intuïcions que encara no han estat suficientment provades. Aquesta classificació no té una funció excloent ni jeràrquica, sinó que serveix per reconèixer les diverses etapes del procés de recerca i contribuir a una millor comprensió col·lectiva del coneixement en evolució.

Un dels riscos coneguts dels models de llenguatge és la generació de respostes inventades o no fonamentades, conegudes com a al·lucinacions. Per mitigar-ho, el sistema força l'ús de context específic recuperat del corpus digital com a única base per generar respostes. A més, les respostes inclouen sempre referències explícites a les fonts documentals utilitzades, cosa que facilita la revisió humana i evita errors de confiança no justificats.

Som conscients que qualsevol corpus històric pot contenir biaixos, omissions o interpretacions parcials pròpies del moment i l'autor. Per això, un dels objectius del sistema és fer visibles aquestes tendències a través de l'anàlisi global i transversal de les tesis, identificant recurrències, contradiccions i buits. La tecnologia no es fa servir per imposar cap veritat, sinó per obrir el debat i facilitar noves aproximacions crítiques. A més, el sistema es desenvolupa amb criteris d'ètica de la IA: transparència, traçabilitat, sobirania de les dades i respecte pels límits epistemològics del coneixement històric.

A partir d'aquesta base, s'obre una nova manera d'explorar i fer créixer el coneixement històric. Gràfics interactius, mapes conceptuals i visualitzacions dinàmiques permeten navegar pel corpus segons criteris diversos: per autor, per temàtica, per projecte, per període històric o per paraules clau. En aquest entorn digital, les tesis apareixen agrupades en clústers temàtics, i es poden consultar amb un sol clic, desplegant-ne els arguments i les fonts associades. Aquesta nova visualització facilita no només l'exploració, sinó també la revisió col·laborativa, la detecció de mancances i la identificació de relacions fins ara invisibles. És, en definitiva, una eina per prendre consciència del que ja tenim, posar-ho tot damunt la taula i generar noves preguntes de recerca.

Amb aquesta infraestructura al darrere, les possibilitats s'amplien enormement. Es poden desenvolupar aplicacions per a la redacció automatitzada de resums de tesis, articles o materials pedagògics; eines col·laboratives que permetin als investigadors comentar, validar i complementar tesis ja existents; traducció automàtica a altres idiomes per projectar la recerca internacionalment; o fins i tot agents conversacionals —chatbots— capaços de respondre consultes sobre qualsevol aspecte del corpus, mantenint-se en la nostra llengua, el català, i aportant referències verificables en cada resposta.

Aquest projecte vol demostrar que la intel·ligència artificial, lluny de suposar una amenaça per al rigor o la veritat, pot ser una aliada fonamental per a una recerca històrica crítica, col·laborativa i basada en evidències. El que proposem no és un model tancat, sinó un sistema viu, auditable, revisable, i sempre obert a nous coneixements i matisos. L'Institut Nova Història, fidel al seu esperit fundacional, assumeix així el repte de posar-se al dia tecnològicament sense renunciar a la profunditat, al rigor ni a la llibertat intel·lectual. En un món cada cop més saturat d'informació, poder estructurar, contrastar i compartir coneixement de qualitat esdevé una tasca imprescindible a favor de la revisió de la nostra Història.

Jordi Berenguer
Juliol, 2025