

SIMONE FAGIOLI¹

Archivi fotografici e testuali e intelligenza artificiale tra Accademia dei Georgofili e Archivio Elio Conti. Casi di studio

¹ Associazione di studi storici 'Elio Conti'

PREMESSA

Pur se da alcuni anni il concetto di “intelligenza artificiale” è entrato nel lessico comune – ma va osservato una lunga ascendenza che parte dagli anni Cinquanta del Novecento – si assiste ancora, sovente, a una sua narrazione come “oggetto mitologico”, derivato da una lettura disattenta e superficiale, anche da parte dei *media*, sia della letteratura scientifica sia di quella divulgativa, nella quale si pone la grande fantascienza che soprattutto tra gli anni Cinquanta e Sessanta ne ha dato una rilevante interpretazione sociale, così come il cinema in anni più tardi.

L'intelligenza artificiale (IA) non è un nuovo strumento ma incute ancora un certo timore nell'opinione pubblica, stante appunto la sommaria conoscenza che se ne ha.

L'IA in realtà si basa su modelli statistici di grandi dimensioni, che permettono l'analisi e l'elaborazione, anche creativa, dei dati, con velocità e accuratezza sempre maggiori, con la possibilità di applicazione in ogni campo, compresa la ricerca storica e sociale.

In questo senso è possibile sviluppare modelli di apprendimento e interrogazione di *corpus* come l'intera serie, dal 1791, degli «Atti della Accademia dei Georgofili», con la possibilità, come *Large Language Model* (LLM), di ricostruire ogni tipo di nesso tra i dati, restituendo un “nuovo” articolo che narra, ad esempio, la storia della coltivazione delle patate così come illustrata dal 1791 in poi, con dati e legami non sequenziali, tutto in linguaggio naturale ossia quello comunemente parlato ogni giorno.

Lo stesso vale per le immagini che possono essere descritte in modo automatico, sviluppando poi un “bot”, ossia un sistema di interrogazione in grado di re-

stituire quanto “visto” dall’IA, come nel caso dell’Archivio fotografico Elio Conti e le sue 3.200 immagini della campagna toscana tra gli anni Sessanta e Ottanta del Novecento, come ha illustrato Francesco Borghero nella sua relazione.

In questo senso l’Archivio fotografico dell’Accademia dei Georgofili, composto da circa 200.000 immagini, approfondito nella relazione di Davide Fiorino, si presta allo sviluppo di modelli di analisi e restituzione, con la possibilità di colorazione, sempre tramite algoritmi IA – e lo stesso vale per le foto di Elio Conti – aspetto che aumenta la quantità di informazioni disponibili nella fotografia e permette maggiori analisi ad esempio in campo botanico, agronomico, di storia del paesaggio, archeologico ed antropologico.

Non indico qui una bibliografia specifica, considerando che i dati esposti sono di ambito sperimentale e vengono da ricerche dell’autore, che se ne assume la responsabilità scientifica.

Per un quadro generale e storico rimando a S. FAGIOLI, *Ora! L’infinita corsa dell’Intelligenza Artificiale*, Effigi, Arcidosso, 2023.

CASI DI STUDIO

Georgofili

Grazie alla cortesia del presidente dell’Accademia Massimo Vincenzini e alla disponibilità di Davide Fiorino ho avuto la possibilità di condurre alcune prove su foto inedite dell’Archivio dei Georgofili, al fine di testare e valutare algoritmi IA sia per la colorazione automatica sia per la descrizione ai fini della catalogazione.

Quest’ultimo aspetto non è marginale se si hanno rilevanti *corpus* fotografici da rendere pubblici e fruibili tramite maschere di interrogazione per l’utente, che può essere non specialista.

Una descrizione “umana” necessita di una vasta gamma di competenze, a fronte ad esempio di archivi eterogenei, come è quello REDA facente parte dei fondi dell’Accademia e che raccoglie davvero svariate immagini sempre di ambito agricolo, ma legate alla produzione industriale del cibo, a strumenti, personaggi, alla divulgazione e così via.

Si prenda la foto 1.

L’immagine proviene dall’Archivio REDA e appunto documenta il lavoro in una industria alimentare.

È possibile certo proporre una descrizione “manuale”, senza sapere esattamente di cosa si tratta, ma allo stesso tempo è possibile ricorrere a un algoritmo per una sua analisi:



Foto 1

Questa foto ritrae un ambiente industriale in bianco e nero. Lavoratori sono concentrati attorno a macchinari, impegnati nelle loro attività. Una figura prominente si china su qualcosa, creando un'ombra contro la luce. L'atmosfera è densa di vapore o fumo, con luci soffuse che illuminano i lavoratori e le macchine. L'immagine trasmette un senso di intensità e attività frenetica. (Bing/ChatGPT 4.0, 14.2.2024)

Si tratta in qualche modo di una descrizione “creativa”, che tuttavia ci dà un sufficiente quadro della situazione, in relazione all'immagine stessa, al suo contenuto “generale”: questa foto nasce certamente con un intento “poetico”, mostrare sì la fatica ma appunto con uno sguardo più estetico.

Diverso invece il caso della foto 2.

Qui nell'analisi ci si potrebbe muovere per paradossi.

Con tutta evidenza si tratta di patate, tuttavia possiamo supporre che il descrittore “umano” non conosca, o meglio riconosca, questo ortaggio, così da dare una descrizione molto generale: “verdure”. Se la foto è messa *on line* e indicizzata come “verdure” la ricerca di *patata* darà un risultato nullo.

Questa fotografia ritrae un enorme quantità di patate che riempiono una stanza. Le patate sono ammassate fino a raggiungere quasi il soffitto in alcuni punti. L'immagine è in bianco e nero e mostra una stanza interna, presumibilmente un magazzino o un de-



Foto 2

posito, piena di patate. Le patate sono ammucciate in modo disordinato, coprendo quasi interamente il pavimento visibile. La stanza ha pareti e soffitto lisci, senza decorazioni o mobili visibili. Ci sono due aperture sulla parete del fondo che sembrano essere porte o finestre, attraverso le quali entra la luce. (*Bing/ChatGPT 4.0*, 14.2.2024)

La descrizione automatica come si vede è ben dettagliata e adeguata e una volta inserita nell'indicizzazione è in grado di restituire una vasta serie di dati. È evidente come le patate siano conosciute da tutti, per cui qui ci starebbe sempre un'etichetta con "patate" – questo è il paradosso – in ogni caso più dati ci sono più il risultato della ricerca sarà accurato.

In questo senso gli esempi possono essere molteplici, indirizzati verso una descrizione scientificamente corretta, realizzata da algoritmi di intelligenza artificiale magari con una supervisione umana.

In questo caso ho utilizzato una IA generale, *ChatGPT 4.0*, addestrata ad ampio spettro; tuttavia, è possibile addestrare in modo puntuale IA specifiche, ad esempio in grado di lavorare su temi particolari, come le macchine agricole, non di immediata riconoscibilità, mentre appunto gli ortaggi sono più di senso comune.



Foto 3

Naturalmente anche in questo caso possiamo considerare che nell'Archivio REDA ci siano casi dubbi o difficili, come nella fotografia 3.

Questa immagine è molto “estetica” e si può considerare con una certa valida approssimazione che ci siano molte persone che non sanno che questi sono cavolfiori, con il contesto che appare ambiguo, dato che i vasi sullo sfondo paiono contenere fiori recisi o piante ornamentali: superficialmente, i cavolfiori possono essere interpretati come fiori. Ma la descrizione tramite IA non lascia dubbi:

Questa è una foto in bianco e nero di un gruppo di ceste piene di cavolfiori. Le ceste sono disposte in due file, con la prima fila che ha sei ceste e la seconda fila che ha cinque ceste. Le ceste sono fatte di materiale intrecciato e hanno forma cilindrica. Le teste di cavolfiore sono grandi e strettamente impacchettate nelle ceste. Lo sfondo è un pavimento di terra e una struttura in legno, probabilmente un capanno o una stalla. La foto sembra essere stata scattata in un mercato o in un'azienda agricola. (*Bing/ChatGPT 4.0*, 28.9.2023)

Quanto riportato è del tutto attinente all'immagine e “veritiero”.

Ci sarebbe molto da discutere su quest'ultimo termine. In sintesi possiamo definirlo nel *range* in cui statisticamente ci si avvicina al mondo reale, considerando per certi aspetti la fotografia *tout court* una "rappresentazione", che ha appunto necessità di una specifica descrizione.

La descrizione ampia, collocata in un *Large Language Model* (LLM), il "sistema" che l'IA legge, analizza e apprende, ci permette poi di fare ricerche ben più articolate che quelle date da una semplice etichetta (#tag) come può essere "verdura".

Se la foto è etichettata appunto solo con "verdura" solo quel termine scritto esattamente in una barra di ricerca mi porterà a quella immagine. Se per errore scrivo "virdura" il risultato sarà nullo.

Se la descrizione è parte di un LLM l'IA nell'interrogazione mi può sempre aiutare a ricerche più articolate, ad esempio segnalandomi che quel "virdura" è un errore e che con la chiave "verdura" le foto disponibili sono una certa quantità, così come se nel caso della foto sopra scrivo "cavolfiore" invece di "cavolfiore".

Ulteriormente un LLM è interrogabile in linguaggio naturale («Puoi cercarmi per favore tutte le foto di cavolfiori?») in ogni lingua disponibile, con risultati che saranno appunto nella lingua in cui è posta la domanda: italiano:italiano; inglese:inglese, cinese:cinese e così via.

Se quanto visto sin qui sta nel campo appunto della descrizione testuale è possibile applicare a queste foto un ulteriore aspetto, quello della colorazione automatica tramite algoritmi IA che mette in evidenza nell'immagine informazioni che il bianco e nero poteva mascherare o ridurre (foto 4).

Come si osserva la foto dei cavolfiori opportunamente colorata ci presenta molte più informazioni, che ampliano in modo netto il *range* di comprensione, pure a non specialisti. Come abbiamo visto l'IA riconosce il tipo di ortaggio anche nella foto in bianco nero, ma in generale l'occhio umano, non più "abituato" a vedere e interpretare il mondo in bianco e nero – dai libri, le riviste, i quotidiani, le foto e la televisione – mostra qui una maggior capacità di analisi. Non si tratta di approfondire gli aspetti estetici, il colore non necessariamente è "bello", qui ci si indirizza più verso una utilità dell'interpretazione.

Quanto detto per le immagini va in parallelo per i testi, in questo caso quelli conservati in Accademia, siano essi a stampa o manoscritti.

È possibile, ad esempio, addestrare una IA con tutti i testi degli «Atti della Accademia» dal primo numero del 1791 ad oggi e interrogarla con ogni genere di domanda, ottenendo risposte coordinate e ben strutturate.

Ciò che segue è una simulazione, non avendo potuto addestrare una IA sui documenti dell'Accademia; tuttavia, è quanto si potrebbe ottenere in termini di risposta dalla "macchina":



Foto 4

Dialogo (simulato) tra un cittadino e l'AI Georgofila.

Ciao. Puoi cercarmi negli «Atti» dei Georgofili, dal primo numero del 1791 all'ultimo del 2023 dove si parla della coltivazione delle patate?

Certo, rispondo ben volentieri alla tua richiesta. Negli «Atti» dell'Accademia dei Georgofili di Firenze non ho trovato nulla di rilevante in merito alla coltivazione della patata, ma nel numero 2 dell'anno 1828 del «Giornale agrario toscano» è presente un articolo a firma F. D. Gallizioli dal titolo *Sopra una varietà di patata*. È molto breve, vuoi che ne faccia comunque una sintesi o vuoi leggerlo per intero?

Archivio fotografico Elio Conti

Date le altre relazioni qui presenti non mi dilungo sulla storia dell'Archivio fotografico di Elio Conti, conservato digitalmente in copia dall'Associazione di studi storici 'Elio Conti' di Firenze e posto *on line*.

Quanto visto per l'Accademia dei Georgofili ha valore anche per questo archivio: la descrizione automatica, la colorazione, lo sviluppo di un LLM con i testi di Conti sono operazioni che possono essere condotte su questo *corpus* di immagini, integrato appunto con gli scritti dello storico.



Foto 5

Ci sono tuttavia alcune differenze, che vado a sottolineare.

Le oltre 3000 foto di Elio Conti per l'area di Firenze tra gli anni Sessanta e Ottanta del Novecento sono state poste *on line* sfruttando ventisette #tag proposti dallo studioso durante le sue campagne fotografiche: *panoramica; terreno vitato; terreno olivato; coltura promiscua; terreno lavorativo; boscato; specializzato; sodo-pascolo; cipressi in filare; terrazzamenti-cigliani; corso d'acqua; casa colonica; casa padronale; annessi rurali; abitato; edificio abitativo; edificio religioso; struttura fortificata; resti di interesse archeologico; edificio in rovina; mulino; strada campestre; stemma-epigrafe; strumenti agricoli; forno; interno; altro.*

Queste categorie permettono di identificare “tutte” le immagini di Conti. Tuttavia sono necessario dei distinguo.

Per restare nel tema dell'Archivio REDA se si volessero ricercare nelle foto di Conti mezzi agricoli o automobili, nell'ottica di una ricerca appunto della “modernizzazione” dell'agricoltura toscana e le ultime famiglie mezzadri, i #tag sopra riportati ci darebbero un risultato nullo, perché indirizzati a una ricerca più di ambito storico e paesaggistico.

Si prenda ad esempio la foto 5.

Si tratta della località Montessassi, Casenuove (MSS0009) nel comune di Vicchio tra la fine degli anni Sessanta e i primi Settanta.

La fotografia riveste un ampio interesse per molti temi: architettura, storia sociale, storia dell'agricoltura, del paesaggio ecc., tuttavia la categoria derivata da Conti per questa immagine è solo "abitato". Si comprende come questo #tag sia riduttivo e non in grado di descrivere compiutamente il "paesaggio sociale" che qui appare in tutta la sua estensione.

La descrizione IA è articolata, pur se non perfetta e molto estetica, rispetto alle immagini viste in precedenza:

Questa fotografia in bianco e nero mostra un edificio antico con muri di pietra, circondato da piante e alberi. L'edificio ha finestre quadrate e una porta ad arco, ed è a due piani. Un'auto d'epoca è parcheggiata davanti all'edificio, aggiungendo un tocco vintage alla scena. Piante giovani sono visibili in primo piano, e l'erba e altre piante circondano l'area. In alto a sinistra, c'è una piccola illustrazione o simbolo che sembra rappresentare un edificio o una struttura simile a quella nell'immagine. L'atmosfera è tranquilla e pacifica. (*Bing/ChatGPT 4.0*, 16.2.2024)

Il simbolo al quale fa riferimento è il logo dell'Associazione 'Elio Conti'. Anche se questa descrizione appare poco scientifica aggiunge alcuni elementi in più, che possono essere approfonditi stilando una descrizione "umana" più ampia. Lo stesso vale per tutte le fotografie Conti, che possono essere appunto integrate dall'IA.

La colorazione di questa foto poi aggiunge molti altri dati (*Reimagine*, algoritmo *DeOldify*, 23.9.2023), analizzabili dall'occhio umano (foto 6).

È possibile che la colorazione delle automobili non sia, relativamente alla marca e modello, corretta, tuttavia questo aspetto potrebbe essere superabile con un addestramento specifico di un algoritmo dedicato, se questo aspetto viene considerato rilevante.

Nello stesso contesto è possibile sviluppare un LLM con gli studi di Elio Conti che fanno direttamente riferimento ai luoghi di queste foto nel suo sviluppo storico, luoghi analizzate in *La formazione della struttura agraria moderna nel contado fiorentino* (Roma, Istituto storico italiano per il Medio Evo, 1965).

CONCLUSIONI

Possiamo trarre ad esempio e paradigma di quanto visto sin qui una frase tratta dalle scene finali del film *Terminator 2. Il giorno del giudizio* (James Cameron, 1991), pronunciata dalla protagonista Sarah Connor:



Foto 6

Il futuro, di nuovo ignoto, scorre verso di noi, e io lo affronto per la prima volta con un senso di speranza, perché se un robot, un *Terminator*, può capire il valore della vita umana, forse potremo capirlo anche noi.

La “macchina” Terminator nel primo film della serie era stata inviata dal futuro per uccidere quello che sarebbe poi diventato il capo della resistenza John Connor, ma nel secondo episodio si assiste a una profonda trasformazione. Il *Terminator*, la macchina, si dimostra verso John protettivo, fino ad arrivare a immolarsi per lui. Il *cyborg* si era sforzato di comprendere un mondo non suo, sino a sviluppare un istinto tutelare e “umano”.

Senza approfondire le complesse riflessioni che almeno sin dal Settecento gli illuministi legano verso uomini e macchine, si può considerare come l’intelligenza artificiale non ci renda meno “umani” bensì si indirizzi a una proficua integrazione tra noi e le macchine, allo sviluppo di strumenti, di idee, che rafforzano la conoscenza, *tout court*.

Gli strumenti, le possibilità, che ho qui illustrato in merito agli archivi dell’Accademia dei Georgofili e a quello di Elio Conti ci permettono sia sulle immagini sia sui testi analisi profonde, in grado di mettere in evidenza nessi che a una pur attenta lettura umana possono sfuggire.

La lettura analitica da parte della IA di tutti gli «Atti» dell'Accademia, ma anche di altre pubblicazioni correlate di ambito agricolo, come il «Giornale agrario toscano», e le nostre possibilità di interrogazione sono un'alleanza formidabile per il campo della ricerca non solo storica, ma allo stesso tempo di comprensione e tutela del territorio.

Gli studi di Elio Conti, integrati con le sue fotografie, possono acquisire una rinnovata profondità di analisi, mostrando davvero per un periodo lunghissimo, dal Medioevo al nostro presente, come aree cruciali della Toscana, gran parte del territorio fiorentino, siano mutate, con la possibilità, integrando sempre più dati, di una rinnovata e corretta pianificazione dello spazio.

L'intelligenza artificiale diviene quindi una nuova e forte alleata non solo degli studi umanistici ma pure una grande risorsa per una più avanzata progettazione del territorio.

RIASSUNTO

Lo sviluppo di algoritmi di intelligenza artificiale (IA), addestrati all'analisi di testi e fotografie con l'ulteriore possibilità di generazione di immagini "reali", ci pone di fronte alla possibilità di utilizzo di questi strumenti per la ricerca storica, sociale, antropologica, stante l'opportunità di processare previo addestramento rilevanti quantitativi di dati sovente "dormienti" in archivi e biblioteche. La gestione analitica di archivi fotografici di migliaia o anche di milioni di immagini necessita di strumenti aggressivi, in grado di muoversi in profondità e restituire in brevissimo tempo e con l'uso del linguaggio naturale le informazioni richieste. In un archivio fotografico un'IA addestrata può essere utilizzata per la descrizione automatizzata delle immagini, con la restituzione di #tag complessi, poi analizzati da una ulteriore IA interfacciata con gli utenti per l'estrazione delle informazioni. Lo stesso vale per archivi testuali, sia a stampa sia manoscritti, dove l'IA può essere utilizzata per una lettura automatica e per la costruzione di modelli (*Large language model*, LLM) in grado di restituire nel modo più ampio possibile le informazioni apprese nell'addestramento.

ABSTRACT

The development of *Artificial Intelligence* (AI) algorithms, trained to analyze texts and photographs with the additional possibility of generating "real" images, confronts us with the challenge of using these tools for historical, social, anthropological research, giving the opportunity to process significant amounts of data after training, often "dormant" in archives and libraries. The analytical management of photographic archives of thousands or even millions of images requires aggressive tools, capable of deep moving and returning the required information in a very short time and with the use of natural language. In a

photo archive, a trained AI can be used for the automated description of images, giving back complex #tag, then analyzed by an additional AI interfaced with users for information extraction. The same applies to textual archives, printed as well as manuscript, where AI can be used for automatic reading and for the construction of models (*Large language models*, LLMs) capable of returning the information learned in the training as widely as possible.