

15a Conferenza Nazionale di Statistica

La statistica ufficiale nel tempo dell'Intelligenza Artificiale

Chiusura della Conferenza Nazionale di Statistica

4 luglio 2024

Prof. Francesco Maria Chelli, Presidente dell'Istat

Care colleghe, e cari colleghi dell'Istat e del Sistan,

al termine di questi due giorni pieni, ricchi, e molto stimolanti, voglio per prima cosa ringraziarvi per l'impegno, la passione e la grande competenza che avete assicurato alla XV Conferenza Nazionale di Statistica.

Quando, ormai diversi mesi fa, abbiamo scelto di dedicare la nostra Conferenza alla Statistica ufficiale nel tempo dell'Intelligenza artificiale, sapevamo bene che ci saremmo avventurati in questioni complesse e in rapidissima evoluzione. ChatGPT è stata lanciata ufficialmente il 30 novembre 2022 con una demo interattiva gratuita. Oggi si parla di Intelligenza Artificiale generativa tutti i giorni e in ogni contesto.

Siamo su una nuova frontiera tecnologica, che interesserà inevitabilmente l'intero ciclo di attività della statistica ufficiale:

dalla raccolta e l'elaborazione dei dati agli sviluppi metodologici,

dalle tecniche di indagine agli standard di classificazione,

dalla diffusione dei risultati, all'accesso ai dati e all'uso delle statistiche.

In questi due giorni, abbiamo iniziato a ragionare sulle opportunità, sulle prospettive, sulle nuove competenze e i nuovi contesti normativi, e sul modo in cui l'intero sistema Sistan possa giovare pienamente delle tecnologie più evolute, nel rispetto della nostra etica di servizio.

L'obiettivo era dare spazio alla nostra intelligenza collettiva, e alle testimonianze di chi si è avventurato oltre questa nuova frontiera.

Il Presidente di Cineca, Francesco Ubertini, ci ha parlato degli sviluppi del supercalcolo come strumento abilitante dell'Intelligenza Artificiale. Abbiamo ascoltato da Piero Cipollone come la BCE stia affrontando questa evoluzione tecnologica e oggi abbiamo qualche idea più chiara sulle sfide e le opportunità che si aprono nell'ambito della strategia europea per l'Intelligenza Artificiale, che, come sappiamo, prende le mosse dal primo quadro regolatorio adottato a livello globale con un approccio chiaro: misurare i rischi dell'Intelligenza Artificiale generativa e tutelare i diritti fondamentali di tutti i soggetti che saranno esposti ad essa.

L'utilizzo di Intelligenza Artificiale e Machine Learning è certamente un passaggio inevitabile nella statistica ufficiale. Tra i benefici attesi ci sono la possibilità di utilizzare a fini statistici fonti di dati non tradizionali, come immagini, segnali, testi, reti, in modo integrato con le fonti tradizionali, per ridurre i costi di rilevazione e il peso statistico sui rispondenti, in linea con gli indirizzi a livello europeo e internazionale.

Tuttavia, al fine di ottenere risultati affidabili è necessario affrontare alcuni problemi, di natura metodologica, epistemologica e perfino etica, che riguardano la rappresentatività e la qualità dei dati. Quando i dati non provengono da rilevazioni statistiche, ne va studiata e trattata la

rappresentatività effettiva rispetto alle popolazioni di interesse. Le nuove fonti ci impongono di sviluppare nuovi framework e misure per la documentazione della qualità dei dati e delle statistiche prodotte.

La discussione che si è avviata in questa Conferenza è stata una prima occasione utile alla definizione di una politica per la qualità per i prossimi anni, compresi gli investimenti sulla qualità dell'Istat, insieme a tutti i soggetti del Sistan. Bisognerà anche tener conto del rispetto della riservatezza e del segreto statistico: sarà necessario sviluppare metodologie per l'accesso e il trattamento sicuro dei dati, nel rispetto delle norme sulla privacy di cittadini e imprese.

L'Istat e il Sistan affronteranno insieme questo passaggio. Dovremo sperimentare, provare, progettare e condividere esperienze, compiere quel balzo in avanti che si prospetta sul fronte dell'interoperabilità delle rispettive banche dati. E dovremo guardare con occhi nuovi fuori dal nostro perimetro consueto. Dovremo fare anche molta esperienza, molta formazione.

Si leggono analisi di ogni tipo sul futuro prossimo che ci attende: forti incrementi di produttività determinati dall'uso dell'Intelligenza Artificiale generativa, ulteriori cambi di paradigma nell'organizzazione del lavoro, dopo quello che abbiamo vissuto e stiamo vivendo con l'esperienza dello smart working.

Secondo una statistica presentata quest'anno al Forum PA, 1,8 milioni di dipendenti pubblici saranno direttamente coinvolti nell'uso quotidiano dell'Intelligenza Artificiale, con vantaggi in termini di miglioramento dei servizi prestati e di efficientamento generale dei processi.

L'Intelligenza Artificiale mette quindi alla prova la statistica ufficiale in molte direzioni, comprese quelle della capacità di valutarne l'impatto culturale, sociale, ed economico.

Ad esempio:

Quale sarà l'impatto dell'Intelligenza artificiale sui processi educativi e formativi? Sulla domanda di lavoro e sulla organizzazione del lavoro? Sul rapporto dei cittadini con i servizi pubblici? Sulla conoscenza e sulla comunicazione? I divari si colmeranno, o si allargheranno ancora di più?

E quale sarà il futuro del lavoro, di fronte alla progressiva penetrazione delle soluzioni di Intelligenza Artificiale nei processi produttivi e nei servizi? Tutte domande con cui abbiamo iniziato a confrontarci.

Ma la nostra conferenza non è stata solo Intelligenza Artificiale. È stata anche un'occasione di dibattito, di scambio, di incontro e di confronto interdisciplinare tra soggetti del Sistan, istituzioni e utilizzatori.

I numeri di questo evento sono eloquenti: nelle circa 30 ore di lavori, gli interventi sono stati oltre 120; soprattutto in ambito Sistan, tra cui quelli dei rappresentanti dell'unione statistica dei comuni, delle province e del coordinamento statistico delle regioni; ma anche dei rappresentanti delle università, della società civile, di organismi internazionali. Sono stati proposti 232 poster scientifici. Abbiamo avuto circa 1.100 partecipanti in presenza e 4.300 collegamenti da remoto e un grande seguito sui canali social.

Le lezioni che abbiamo tratto dall'intensa attività di questi due giorni sono molte, e importanti. Non pretendo di riassumerle tutte, e mi limito a individuare alcuni messaggi chiave "dal sistema e per il sistema".

Dal sistema-Sistan, perché ci siamo confrontati su molte questioni: fonti, definizioni, misurazioni, utilizzo e organizzazione del sistema.

Per il sistema-Sistan, perché abbiamo ascoltato le voci di chi lavora con i dati (accademia, policy maker), così come delle imprese e del terzo settore; voci importanti per intercettare i temi emergenti su cui la società necessita di un quadro solido di riferimento in termini di definizioni e misurazione; ma anche per fare il punto su quei fenomeni che, nonostante i molti progressi, a volte faticano a trovare una rappresentazione adeguata nel sistema delle statistiche ufficiali.

Abbiamo visto come i nostri dati e i nostri sistemi di indicatori costituiscano ormai un punto di riferimento consolidato e rilevante per molti quadri di riferimento delle policy a livello nazionale (ad esempio le politiche di coesione) e internazionale (ad esempio gli Obiettivi di sviluppo sostenibile).

nel caso del Bes, si è generato addirittura un processo contrario in quanto un framework di misurazione statistica è stato inserito nei Documenti di economia e finanza.

Voglio approfittare di questo passaggio per ricordarvi un importante appuntamento, che si terrà a Roma dal 4 al 6 novembre di quest'anno, su un tema a noi caro. Si tratta del settimo Forum mondiale dell'OCSE sul benessere, che riunisce esperti, professionisti della materia e leader mondiali per mettere al centro il benessere multidimensionale nelle politiche, nella misurazione e nell'azione sociale.

L'Istat collabora all'organizzazione di questa edizione insieme al Ministero dell'Economia e delle Finanze e con il sostegno della Banca d'Italia.

Di dati e indicatori si è discusso molto nel corso della Conferenza. Il dibattito ha offerto numerosi spunti, messo in luce criticità e fornito preziose indicazioni su come proseguire nel nostro lavoro. Ad esempio, è emerso come nell'analisi del nostro sistema produttivo lo studio delle filiere o delle aree di specializzazione sia fondamentale per avere un quadro interpretativo che superi la tradizionale classificazione per settori di attività economica. Oppure, si è resa evidente la necessità di rafforzare le statistiche sull'istruzione e l'analisi dell'impatto dell'invecchiamento sulla nostra società e i nostri sistemi di welfare; o ancora l'assoluta importanza di quadri informativi dettagliati a livello territoriale che consentano di misurare i divari.

Il confronto sugli indicatori congiunturali ha testimoniato poi un rinnovato interesse verso la produzione statistica infrannuale, valorizzando la presenza di nuove fonti e nuovi attori rispetto alla statistica ufficiale. Questa maggiore ricchezza informativa ci sollecita, come Sistan, a porci, da un lato, come punto di riferimento sul tema della qualità per gli altri soggetti produttori di statistiche, dall'altro come elemento fondamentale per la formulazione di un solido quadro informativo sulla congiuntura.

Abbiamo fatto il punto sul sistema delle statistiche della popolazione, dopo cinque edizioni di censimento permanente della popolazione (i risultati della sesta sono in corso di validazione).

Siamo ormai in una fase in cui l'Istat, oltre a farsi carico della restituzione al pubblico della ricca e dettagliata informazione statistica, può proporre, insieme agli altri enti del Sistan, temi che fino a qualche anno fa erano fuori

dal nostro campo di osservazione. Ne costituiscono un esempio le analisi strutturali condotte sulla popolazione straniera o le analisi sul capitale umano a disposizione del Paese, condotte anche a un livello territoriale particolarmente fine, quale la sezione di censimento.

Una menzione particolare merita poi il lavoro che stiamo facendo per diffondere, a partire da settembre 2024, i risultati della revisione generale dei conti nazionali, coordinata a livello europeo (il cosiddetto nuovo benchmark), e che vedrà l'inserimento di alcune fonti e il miglioramento di metodi di stima. La Conferenza è stata l'occasione per riflettere su altri temi di rilievo per la Contabilità nazionale, quali l'importanza dei Conti satellite come strumento a supporto dell'analisi economica e sociale e le implicazioni di finanza pubblica del nuovo Patto di stabilità e crescita.

Ma il messaggio più forte che a mio avviso è emerso da tutti questi incontri è che la ricerca sui nostri dati costituisce un feedback insostituibile per innescare un continuo miglioramento della nostra produzione. Per questo motivo, l'Istat dovrà potenziare le attività di ricerca nell'ambito del Sistan e con enti esterni, e all'interno dell'Istituto, in modo inclusivo efficace e proficuo per i nostri obiettivi; continuando a perseguire la strada già intrapresa di una sempre maggiore valorizzazione dei microdati per la ricerca.

I fornitori e al tempo stesso gli utilizzatori più importanti di statistiche ufficiali rimangono però i cittadini, le imprese, le istituzioni. La ricchezza e la qualità dell'informazione statistica ufficiale prodotta dall'Istat e dal Sistan si fondano sul rapporto che abbiamo saputo costruire con i rispondenti, basato su una solida e storica tradizione di ascolto e fiducia

reciproca. Se in quasi cento anni di attività abbiamo creato e diffuso un patrimonio informativo di inestimabile valore è grazie alle innumerevoli persone che hanno risposto con pazienza ai nostri questionari. È stato così in passato e sarà così anche domani: né gli archivi amministrativi né gli algoritmi dell'intelligenza artificiale, infatti, potranno fornirci dati sufficienti a conoscere e descrivere in maniera accurata bisogni e comportamenti di una società in costante e rapido mutamento. Ecco perché l'Istat continua a impegnarsi per facilitare la partecipazione alle rilevazioni, innovando le tecniche di indagine e adottando le tecnologie più avanzate per andare incontro alle esigenze dei rispondenti, che sono e resteranno la risorsa più preziosa per la statistica ufficiale.

Il rilascio del nuovo sito è un tassello che si aggiunge al mosaico che stiamo componendo. Progettato come un giornale online, il nuovo sito è il risultato di un complesso lavoro, ampiamente partecipato, in cui l'uso della componente tecnologica è stato un fattore chiave abilitante per il rinnovo del web design, la razionalizzazione dei contenuti, la facilità di ricerca delle informazioni. Il nuovo sito pone al centro le esigenze dei diversi utenti, ed è reso totalmente accessibile attraverso l'introduzione di un motore di ricerca semantico, pronto alle applicazioni di Intelligenza Artificiale.

L'ultima riflessione che vi propongo riguarda il Sistema. Siamo sistema, e dobbiamo impegnarci a rafforzare questa nostra natura, nel rispetto delle diversità. La frammentazione condanna alla irrilevanza, come è stato ricordato in più interventi.

Sul fronte delle amministrazioni comunali, bisogna investire di più sulle Unioni e le Associazioni tra i Comuni, al fine di rafforzarne la capacità

statistica, anche come unità rispondenti alle nostre indagini periodiche, ma non solo. Mentre con gli altri enti nazionali del Sistema - penso a quelli di maggiori dimensioni, come le Ona, le Regioni, le Authority di settore, le unità statistiche di Poste.it, Ferrovie, Terna, Gse, per citarne solo alcune - bisogna consolidare le collaborazioni già in corso, sperimentarne di nuove, rendere più efficienti gli scambi di dati, le interoperabilità, condividere i calendari delle pubblicazioni, insomma fare Sistema con un approccio nuovo, magari investendo insieme su modelli di Intelligenza Artificiale.

Rimane comunque la necessità di una legge quadro per la statistica ufficiale e l'attività dell'Istat e del Sistan che risponda al tempo attuale. E non solo perché lo ha auspicato Eurostat, ancora una volta, al termine del terzo ciclo della peer review che si è conclusa all'inizio del 2023. È un tema di cui s'è discusso anche qui in questa Conferenza. La norma base che regola la vita dell'Istituto e del Sistema Statistico nazionale, fortemente innovativa quando fu promulgata, è stata concepita quando non esisteva internet, la posta elettronica, figuriamoci la Pec o lo Spid...e l'Intelligenza artificiale era ancora solo un promettente campo di ricerca.

Care Colleghe e cari Colleghi, avviandomi alla conclusione, ricordo che la prossima Conferenza Nazionale di Statistica si terrà nel 2026, e coinciderà con il centenario dalla fondazione dell'Istat. Penso a quante trasformazioni sociali, culturali, economiche, politiche, e ambientali profondissime, in Italia e nel mondo intero, in questi nostri primi cento anni, si sono rapidamente susseguite. È per me affascinante e straordinario osservare come la statistica ufficiale sia stata capace di rimanere, rinnovandosi, sempre al suo posto, e di servire il Paese, accompagnandolo, passo dopo passo.

E ora, non mi resta che ringraziare coloro che hanno contribuito alla vivacità e all'intelligenza delle discussioni di questi due giorni, e, soprattutto, le colleghe e i colleghi che, con il loro lavoro, la loro dedizione e la loro cura, hanno reso possibile la realizzazione di questa XV Conferenza. Un successo di tutti.