ACCREDIA

L'ENTE ITALIANO DI ACCREDITAMENTO

PER DEFINIRE PROCESSI DI "CONTROLLO DELLA CONFORMITÀ" DEI SISTEMI DI INTELLIGENZA ARTIFICIALE

MASSIMO DE FELICE

PRESIDENTE

OSSERVATORIO ACCREDIA

16 OTTOBRE 2024

INTRODUZIONE

La realizzazione dell'"Osservatorio" su *Intelligenza Artificiale tra valutazione del rischio e certificazione accreditata* è stata programmata da Accredia per avviare un percorso di consapevolezza.

La collaborazione col CINI (il Consorzio Interuniversitario Nazionale per l'Informatica) conferma la convinzione che la conoscenza tecnica – e di alta qualificazione – è prerequisito necessario alla *governance* dell'accreditamento.

Per individuare e collegare temi di approfondimento, proporrò un breve itinerario problematico, in quattro blocchi.

PRIMO BLOCCO: A CHE PUNTO SIAMO

Lo *status* dell'intelligenza artificiale è lontano, per concretezze, dalle genericità e dagli offuscamenti del dibattito "sociale"[1].

L'impianto tecnico parte da progetti remoti[2]. Sistemi cosiddetti di AI sono già utilizzati, e in rapido sviluppo.

Nel rapporto Draghi se ne rileva l'«uso massiccio» nel settore dell'energia, con oltre cinquanta casi che vanno dalla manutenzione della rete alla previsione del carico. E poi potenzialità di sviluppo nel settore automobilistico, nel miglioramento di efficienza della robotica industriale, nella farmaceutica e nella sanità, nello sviluppo di nuovi materiali, nell'elaborazione automatica di dati «per ridurre i costi amministrativi e di conformità per le PMI».

Altri impieghi – nei settori finanziario, assicurativo, del commercio, della pubblicità, dell'informazione, con ruolo tra le *corporate technologies* – sono ben noti, al centro del dibattito tecnico[3].

Abbiamo a disposizione ampia varietà di tecniche per la gestione dei dati, aspetto fondamentale e preliminare per l'uso e la governance dei sistemi di AI[4]. Ampio ed elegante è stato il dibattito metodologico sul se i dati "sono muti" o "parlano da soli" (si è lavorato sul confronto tra "algorithmic modeling" *vs* "data modeling", mediato se non concluso con la «*veridical data science*»)[5].

---

1 L'*Artificial Intelligence* è stata qualificata espressione dell'anno 2023 dal *Collins Dictionary*: «tema dominante di conversazione e speculazione». Per il *Cambridge Dictionary* la parola dell'anno 2023 è «hallucinate» («avere allucinazioni»). Il lemma ha un significato aggiunto, indotto dall'utilizzazione dell'AI: «When an artificial intelligence (= a computer system that has some of the qualities that the human brain has, such as the ability to produce language in a way that seems human) hallucinates, it produces false information: ▪ *LLMs are notorious for hallucinating – generating completely false answers, often supported by fictitious citations*. ▪ *The latest version of the chatbot is greatly improved but it will still hallucinate facts.».*

2 McCarthy, J., Minsky, M.L., Rochester, N., Shannon, C.E., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, C.E., August 31, 1955 (e ancora indietro si potrebbe risalire considerando alcuni lavori di Turing e di Norbert Wiener).

3 *The future of European competitiveness*. Part A | *A competitiveness strategy for Europe. The future of European competitiveness*, pagina 20. Part B | *In-depth analysis and recommendations*, September 2024.

4 Hastie, T., Tibshirani, R., Friedman, J., *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, Berlin, Springer, 2009; Efron, B., Hastie, T., *Computer Age Statistical Inference. Algorithms, Evidence, and Data Science*, Cambridge, Cambridge University Press, 2018 (in particolare *Part III – Twenty-First-Century Topics*).

5 La «*veridical* ("truthful") *data science* [...] encompasses a unified view of pure prediction and traditional regression approaches under the three principles of predictability, computability and stability (PCS). [...] Our "one culture" veridical PCS data science framework highlights the practice of veridical ("truthful") data science for reliable,

Anche le problematiche giuridiche di interrelazione dell'AI col diritto e i diritti sono ben-delineate (per dare dimensione della letteratura richiamo alcune parole-chiave e frasi caratterizzanti: responsabilità, contratti, governance, persona e privacy, diritti nelle piattaforme, proprietà intellettuale); siamo a livello di principî a disposizione per sostenete la normativa[6].

Tutto sembra pronto. La regolamentazione si trova a rincorrere: con l'obiettivo di garantire utilizzazioni a rischio controllato, con l'avvertenza di non limitare i vantaggi dell'innovazione.

IL SECONDO BLOCCO: LA REGOLAMENTAZIONE EUROPEA E IL "CHE FARE"

Con l'*AI Act* è stata avviata – ufficialmente a luglio – la definizione del quadro normativo: definita la catena dei controlli per garantire qualità e sicurezze: le Autorità preposte, gli enti di accreditamento, gli organismi notificati (possibilmente accreditati) che valutano la conformità dei sistemi[7].

I sistemi di intelligenza artificiale sono caratterizzati in termini molto generali; rilevante – per le responsabilità di controllo e di governance – è la classificazione per livelli di rischio.

Particolari attenzioni richiedono i sistemi qualificati ad "alto rischio". Sono ad alto rischio – a esempio – i sistemi di AI utilizzati per la biometria, per le infrastrutture digitali critiche, nell'istruzione e formazione, nell'amministrazione della giustizia, per giudicare dell'affidabilità creditizia delle persone, nel pricing delle polizze di assicurazione sulla vita e sanitarie, … (allegato III, dell'*AI Act*).

Il "come fare" è tutto da definire.

Per l'accreditamento e per i giudizi di conformità è necessario definire protocolli d'azione, su: processi di prova; per il giudizio sui dati (*data quality*, adeguatezza rispetto all'obiettivo, verifica delle proprietà statistiche); sulle modalità di documentazione tecnica (a tutela della trasparenza); sulla qualità dell'interfaccia uomo-macchina (a tutela del ruolo dello *human-in-the-loop*); sull'adeguatezza delle risorse computazionali e di hardware; sul sistema di monitoraggio (strutturazione dell'*event log*, ciclo di vita del sistema); sulle garanzie di cibersicurezza; sui criteri di verifica dei livelli di "consapevolezza" degli utilizzatori, e quindi sui piani di formazione.

Oltre al giudizio sui "macchinari" e sull'algoritmica, rilevante è l'impegno a giudicare il processo d'uso dei sistemi. Molto lavoro verso la concretezza c'è da fare.

La "*Standardization Request"* rivolta dalla Commissione europea ai Comitati [Europei] di Normazione dovrebbe produrre schemi che diano modalità operative (prassi d'azione) al "testo regolamentare"[8].

---

reproducible and transparent datadriven decision making and knowledge generation from data for particular domain problem» [Yu, B., Barter, R., *The Data Science Process: One Culture*, (discussion all'articolo di Efron), Journal of the American Statistical Association, 2020, vol. 115, n. 530, pagina 673].; altre utili considerazioni in Yu, B., Kumbier, K., *Veridical Science*, PNAS, February 25, 2020 (vol.117, no. 8).

6 Ruffolo, U., (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Milano, Giuffrè Francis Lefebvre, 2020; Ruffolo, U., (a cura di), *XXVI Lezioni di Diritto dell'Intelligenza Artificiale*, Giappichelli, Torino, 2021; Pinelli, C., Ruffolo, U., *I diritti nelle piattaforme*, Giappichelli, Torino, 2023.

7 Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio *che stabilisce regole armonizzate sull'intelligenza artificiale*, Gazzetta ufficiale dell'Unione europea, 12.7.2024.

8 La "*Standardization Request"* è rivolta al Comitato Europeo di Normazione (CEN), al Comitato Europeo di Normazione Elettrotecnica (CENELEC), all'Istituto europeo per le norme di telecomunicazione (ETSI). Sono dieci gli ambiti individuati (tutti ovviamente riferiti ai sistemi di AI): *1*. sistemi di gestione del rischio; *2*. governance e qualità dei data set utilizzati; *3*. tracciamento e monitoraggio delle attività; *4*. obblighi di trasparenza e informazioni agli utenti; *5*. requisiti di supervisione umana; *6*. accuratezza; *7*. specifiche di robustezza; *8*. requisiti in tema di cybersicurezza; *9*. gestione del monitoraggio della qualità, incluso il monitoraggio post commercializzazione; *10*. valutazione della conformità.

Non può essere sottovalutata la difficoltà indotta dalla dipendenza dai dettagli: non basta definire i modi per giudicare della conformità dei "macchinari" (hardware, software, algoritmica), l'esito dell'utilizzazione dei sistemi dipende in modo rilevante dai dati utilizzati (qualità e completezza rispetto al metodo e allo scopo) e dal controllo di "stabilità" (tecnicamente entra in gioco l'inferenza)[9].

Sarà difficile prescindere, per normare, dai riferimenti agli ambiti specifici (a classi "omogenee" di casi concreti). Nell'"Osservatorio" tre "casi di applicazione" sono proposti per avviare – come si diceva – il percorso.

Un "Osservatorio" perciò (questo che presentiamo) non-conclusivo. Gli sviluppi tecnologici e l'ampliamento degli ambiti di applicazione richiederanno "osservazione" continua e attenta per specificare i protocolli di conformità, sempre sostenuta da quella conoscenza tecnica di alta qualificazione.

PENULTIMO BLOCCO: TRE OSSERVAZIONI SUL DISEGNO DI LEGGE

Il Disegno di legge *in materia di intelligenza artificiale* è in discussione al Parlamento[10].

Si sono avviate azioni; sono tre gli ambiti principali.

1. Con l'articolo 18, si stabilisce che è l'Agenzia per l'Italia Digitale – l'AgID – «a definire le procedure e a esercitare le funzioni e i compiti in materia di notifica, valutazione, accreditamento e monitoraggio dei soggetti incaricati di verificare la conformità dei sistemi di intelligenza artificiale».

   Accredia e AgID hanno avviato una collaborazione – un "tavolo di lavoro" – per specificare attività da svolgere e modalità di formale collaborazione.

2. Anche con l'Agenzia per la Cybersicurezza Nazionale andrà estesa la collaborazione già in-essere con Accredia. L'ACN si prevede (sempre nell'articolo 18) sia nell'ambito "intelligenza artificiale, sicurezza informatica" «responsabile per la vigilanza, ivi incluse le attività ispettive e sanzionatorie».

3. Nell'articolo 22 si pone rilevante l'esigenza di «formazione in materia di utilizzo dei sistemi di intelligenza artificiale»: con esplicito impegno «da parte degli ordini professionali, […] per i professionisti e per gli operatori», e poi anche «nei corsi universitari», «negli istituti tecnologici superiori».

   Riguardo alla formazione (prerequisito per poter tutelare il ruolo dello *human-in-the-loop*) l'*Accredia Academy* potrà portare – ai fini dell'accreditamento responsabile – contributi di adeguata qualità, magari estendendo la collaborazione con unità di ricerca universitarie a alta specializzazione.

[9] Il tema non è nuovo, e ampiamente sondato; per una partenza: Yu, B., *Stability*, Bernoulli 19(2013), 4.

[10] Senato della Repubblica, *Disposizioni e delega al Governo in materia di intelligenza artificiale* [Disegno di legge N. 1146, 20 maggio 2024].

PER ULTIMO: DUE TEMI DA CONSIDERARE

Altri due temi rilevanti, su cui sarà necessario lavorare.

Il primo riguarda il rapporto tra accreditamento, verifica di conformità, e ruolo delle Autorità di vigilanza (Banca d'Italia, Ivass, Consob): interazione con i loro regolamenti.

È un tema all'attenzione[11], segnalato pochi giorni fa anche dal Segretario generale dell'Ivass (il dottor De Polis)[12]; richiederà efficace coordinamento (tecnico e normativo).

Il secondo tema è sul come regolare i giudizi di conformità (tempi e modi) per sistemi già in uso, che per ottemperare a nuovi principî potrebbero richiedere impegnativi programmi di cambiamento nelle tecniche e nell'organizzazione d'impresa.

11 Per inquadrare la problematica, tre recenti documenti (con i loro rimandi): Ivass, *Indagine sull'utilizzo degli algoritmi di 'Machine Learning' da parte delle imprese assicurative nei rapporti con gli assicurati*, Febbraio 2023; Siani, G., *AI-driven bank: Opportunità e sfide strategiche per il sistema finanziario e la vigilanza*, Banca d'Italia, 3 ottobre 2023; Perrazzelli, A., *Le interconnessioni tra Intelligenza Artificiale, Cloud e Cyber nel settore finanziario*, Banca d'Italia, 5 giugno 2024 [relazione al *Cetif Summit. Innovation Trends in Finance 2024: Journey to AI, Cloud e Cyber*]. A conferma della rilevanza primaria di tecnica e tecnologia nella gestione dei sistemi di AI, significativo che nella riorganizzazione interna della Consob tra gli obiettivi ci sia «maggiore efficienza e più spazio […] alla gestione dei dati», e quindi la costituzione della «Divisione Informatica e Intelligenza artificiale» [CONSOB, *Comunicato stampa*, 9 settembre 2024].

12 De Polis, S., *L'intelligenza artificiale nel settore assicurativo*, [relazione al Convegno Banca d'Italia Milano e ANSPC, su "Intelligenza artificiale e mondo finanziario: quali applicazioni, quali implicazioni"], Ivass, 9 ottobre 2024, in particolare pagina 5.

ACCREDIA

L'ENTE ITALIANO DI ACCREDITAMENTO

ACCREDIA

Via Guglielmo Saliceto, 7/9 - 00161 Roma
T +39 06 8440991 / F +39 06 8841199
info@accredia.it

Dipartimento Certificazione e Ispezione

Via Tonale, 26 - 20125 Milano
T +39 02 2100961 / F +39 02 21009637
milano@accredia.it

Dipartimento Laboratori di prova

Via Guglielmo Saliceto, 7/9 - 00161 Roma
T +39 06 8440991 / F +39 06 8841199
info@accredia.it

Dipartimento Laboratori di taratura

Strada delle Cacce, 91 - 10135 Torino
T +39 011 328461 / F +39 011 3284630
segreteriadt@accredia.it