

Chi “e-controlla” i “BIAS controllori”? Raccomandazioni e suggerimenti per un audit algoritmico efficiente ed efficace

di Gianluigi Delle Cave - pubblicato su “www.irpa.eu” - Osservatorio sullo Stato digitale, 21 giugno 2023

I c.d. “audit dell’I.A.” rappresentano uno strumento sempre più popolare per la corretta allocazione delle “responsabilità” algoritmiche. Tuttavia, tali meccanismi di controllo e verifica sono, allo stato, molto poco definiti, sia da un punto di vista tecnico che, soprattutto giuridico. Va da sé, quindi, che senza una chiara comprensione dello spettro di azione di tali pratiche di audit, e in assenza di standard o linee guida tecniche e normative sul punto, il controllo da parte dei “revisori del digitale” risulta poco incisivo e sicuramente non idoneo a mitigare tutti quei potenziali pregiudizi e danni che i sistemi di I.A. possono causare. In tale contesto, si innesta un recente scritto di un gruppo di ricercatori del digitale, che tenta di identificare alcune best practice per procedere con gli audit sopra richiamati, pure enucleandone gli ostacoli applicativi e i potenziali impatti pratici.

In principio fu Giovenale con “Quis custodiet ipsos custodes?”, locuzione latina tratta dalla VI Satira (“tracce” di tale espressione si rinvencono, ben prima, anche in Platone, serafico nel sostenere che i custodi dello Stato devono guardarsi dall’ubriachezza, per non avere essi stessi bisogno di esser sorvegliati). Oggi, invece, e con riferimento a tutt’altri scenari e circostanze, si parla di “Who Audits the Auditors?” per un audit algoritmico sicuro, imparziale ed efficace. Così, infatti, esordisce un recente scritto di tre ricercatori del Algorithmic Justice League (“AJL”), organizzazione no-profit la cui mission è quella di sensibilizzare l’opinione pubblica sugli impatti dell’I.A., stimolando ricercatori, responsabili politici e operatori del settore a mitigare i danni e i pregiudizi dell’intelligenza algoritmica.

Sintetizzando, senza pretese di esaustività, i risultati del lavoro di approfondimento dei paladini dell’AJL (per l’appunto incentrato sulle potenzialità e i limiti dell’audit algoritmico), questi offrono al lettore/studioso/operatore di settore pubblico-privato del tema alcune raccomandazioni “pratiche” per migliorare e implementare il sistema di verifica in oggetto, ossia: (i) imporre a proprietari e a sviluppatori di sistemi di intelligenza artificiale audit algoritmici indipendenti rispetto a standard tecnici chiaramente definiti e individuati; (ii) informare – in maniera chiara e trasparente – i singoli individui quando sono soggetti a sistemi decisionali algoritmici; (iii) imporre la divulgazione di componenti chiave dei risultati degli audit per la revisione paritaria; (iv) considerare maggiormente i “danni” producibili dai sistemi di I.A. nel processo di revisione, prevedendo meccanismi standardizzati di segnalazione e risposta agli incidenti dannosi; (v) coinvolgere direttamente, nel processo di revisione algoritmica, le parti interessate che con maggiore probabilità possono essere danneggiate dai sistemi di I.A.; (vi) formalizzare la valutazione e, potenzialmente, l’accreditamento degli auditor algoritmici.

Muovendo lo sguardo, ora, ai profili giuridici (in specie nazionali), appaiono certamente interessanti i potenziali risvolti di tali “suggerimenti d’audit” laddove, ad esempio, si consideri l’applicazione delle decisioni algoritmiche in seno alle “scelte” amministrative. L’algoritmo, come più volte

evidenziato, porta sempre ad un risultato imparziale, senza che alcun elemento soggettivo possa intervenire a alterare o mutare il risultato. Applicando pure gli audit digitali, si potrebbe quindi maggiormente rafforzare quelle esigenze – in contro bilanciamento delle istanze di semplificazione e accelerazione procedimentali da tanti anni predicate – volte ad assicurare un “controllo umano” del procedimento, in funzione di garanzia (c.d. “human in the loop”). In sostanza, da un lato e a monte, vi sarebbe il controllo del funzionario, che può intervenire in qualsiasi momento per compiere interlocuzioni con il privato, per verificare l’esattezza dei dati da elaborare e mantenendo il costante controllo del procedimento; dall’altro e a valle, vi sarebbe invece l’immediato e rigido controllo dell’auditor, pronto a formalizzare le proprie valutazioni anche coinvolgendo direttamente i soggetti interessati dal sistema algoritmico.

A raccogliere, quindi, gli spunti di riflessione offerti dall’AJL, e declinando gli stessi nel contesto procedimentale amministrativo (ad esempio), il ricorso all’algoritmo, in funzione integrativa e servente della decisione umana, non comporterebbe alcun “abbassamento” del livello delle tutele, essendo sempre rispettato – sia a monte che a valle – il principio di trasparenza (che, come noto, trova un immediato corollario nell’obbligo di motivazione degli atti amministrativi ex art. 3, l. n. 241/1990 e che certamente non può essere soppresso né ridotto solo per la presenza di un algoritmo all’interno dell’iter procedimentale).

Vero è che l’attenzione principale del dibattito si concentra nella fase “a monte”, sulla costruzione dell’algoritmo, ossia su come i parametri dell’algoritmo vengono scelti (operazione di per sé soggettiva), e come si combinano tra loro; e – ancor prima – su come i termini assunti quale parametro siano stati realizzati (momento in cui si opera la scelta caratterizzata da discrezionalità). E però, non vi è chi non veda come il momento dell’audit (inteso, anche qui, come passaggio “a valle” dell’individuazione di tali parametri) possa certamente contribuire ad assicurare quella conoscibilità della costruzione dell’algoritmo, anche, eventualmente, in funzione del sindacato sull’atto adottato sulla base dello stesso.

Peraltro, non può tacersi del fatto che le “raccomandazioni” suggerite dal AJL richiamano, di fatto, anche la tutela del principio di conoscibilità, vale a dire che ognuno ha diritto sia a conoscere l’esistenza di processi decisionali automatizzati che lo riguardano, sia a ricevere informazioni significative sulla logica utilizzata, così come previsto dagli artt. 13 e 14 del GDPR (Regolamento 2016/679) che risultano formulati in maniera generale e, perciò, applicabili sia a decisioni prese da soggetti privati che da soggetti pubblici. La conoscibilità dell’algoritmo (e del sistema in cui esso si inserisce), del resto, deve essere garantita in tutti gli aspetti: dai suoi autori al procedimento usato per la sua elaborazione, al meccanismo di decisione, comprensivo delle priorità assegnate nella procedura valutativa e decisionale e dei dati selezionati come rilevanti.

Ed è proprio quella “caratterizzazione multidisciplinare” dell’algoritmo a imporre che, per la sua comprensione e verifica, siano coinvolte non solo competenze giuridiche, ma anche tecniche, informatiche, statistiche, amministrative, di modo che la “formula tecnica”, che di fatto rappresenta l’algoritmo, sia corredata da spiegazioni che la traducano nella “regola giuridica” (verificabile, conoscibile e comprensibile) ad essa sottesa.

Alla luce di tutto quanto sopra detto, dunque, se anche la revisione contabile deve evolversi in un meccanismo chiave per la responsabilità algoritmica, è importante comprendere e dare una effettiva “forma” all’ecosistema emergente dei “revisori dell’I.A.”, ecosistema che necessita di una regolamentazione quantomeno chiara e inequivocabile.