Adozione di una metodologia di valutazione di conformità etica di sistemi intelligenti governativi

Gaetano Riccio
Università degli Studi di Napoli Federico II

L'integrazione dell'Intelligenza artificiale Generativa (IAG) nella Pubblica amministrazione offre opportunità per ottimizzare decisioni, servizi e interazioni con i cittadini, ma solleva complessità normative ed etiche. Lo studio propone un framework che armonizza AI Act, GDPR e standard internazionali, adattati al settore pubblico. Governance multilivello, gestione del rischio e supervisione umana emergono come strumenti essenziali per mitigare *bias*, garantire trasparenza e tutelare diritti. L'analisi di casi europei mostra che una governance etica e robusta rafforza compliance, sicurezza e fiducia, favorendo un valore pubblico sostenibile.

The integration of Generative Artificial Intelligence (GAI) into Public Administration offers opportunities to optimize decision-making, services, and citizen interactions, while raising regulatory and ethical challenges. This study proposes a framework harmonizing the AI Act, GDPR, and international standards, tailored to the public sector. Multilevel governance, risk management, and human oversight emerge as key tools to mitigate bias, ensure transparency, and safeguard rights. Analysis of European cases shows that ethical, robust governance strengthens compliance, security, and trust, fostering sustainable public value.

DOI: 10.53223/Sinappsi_2025-02-11

Citazione
Riccio G. (2025), Adozione di una metodologia di valutazione di conformità etica di sistemi intelligenti governativi, *Sinappsi*, XV, n.2, pp.156-170

Parole chiave
Governance
Intelligenza artificiale
Pubblica amministrazione

Keywords
Governance
Artificial intelligence
Public administration

Introduzione: il contesto di adozione dell'IA generativa nella Pubblica amministrazione

L'evoluzione dell'Intelligenza artificiale ha raggiunto una fase cruciale con l'avvento dei sistemi generativi, che non si limitano ad analizzare dati esistenti ma possono creare autonomamente nuovi contenuti testuali, visivi e multimediali. Questi sistemi, basati su architetture avanzate come i *Large Language Models* (LLM) e le reti generative avversarie (GAN), offrono alla Pubblica amministrazione (PA) potenzialità trasformative senza precedenti. A differenza dei sistemi di IA tradizionali, focalizzati su compiti specifici e predeterminati, l'IA generativa permette di automatizzare la produzione di documenti amministrativi, fornire assistenza personalizzata ai cittadini attraverso chatbot avanzati, generare analisi predittive complesse e supportare processi decisionali con informazioni contestualizzate. Questa rivoluzione tecnologica si inserisce nel più ampio processo di digitalizzazione

della PA, accelerato dalla pandemia di Covid-19 e dagli investimenti previsti dal Piano nazionale di ripresa e resilienza (PNRR).

La questione di ricerca che questo articolo intende affrontare riguarda le modalità con cui le amministrazioni pubbliche possano implementare sistemi di IA generativa garantendo contemporaneamente la conformità normativa, la tutela dei diritti fondamentali e l'efficienza amministrativa. La tesi sostenuta è che un'implementazione efficace richieda un framework integrato che coordini aspetti normativi, organizzativi, tecnici ed etici in un approccio olistico, superando la frammentazione che caratterizza molte iniziative attuali. Secondo Misuraca *et al.* (2020), infatti, le iniziative di IA nella PA europea tendono a svilupparsi in maniera non coordinata, creando disallineamenti tra innovazione tecnologica e framework regolatori.

Il contesto normativo europeo si è evoluto significativamente con l'adozione dell'AI Act[1], che impone requisiti specifici per i sistemi di IA ad alto rischio, molti dei quali trovano applicazione nell'ambito pubblico. Tale Regolamento si aggiunge al GDPR, creando un complesso ecosistema normativo che le PA devono navigare nell'implementazione di soluzioni di IA generativa. Parallelamente, standard internazionali come il NIST AI Risk Management Framework (AI RMF) e l'ISO 42001 forniscono linee guida complementari sulla gestione dei rischi e sulla governance dei sistemi di IA. Questa stratificazione normativa richiede un approccio integrato che armonizzi i diversi requisiti in un framework operativo coerente.

La letteratura scientifica ha evidenziato diversi gap nella ricerca sull'implementazione dell'IA generativa nella PA. Wirtz *et al.* (2022) identificano una carenza di studi empirici sugli impatti organizzativi dell'IA nel settore pubblico, mentre Sun e Medaglia (2019) sottolineano la mancanza di framework specifici per la valutazione del rischio dei sistemi generativi. Un'analisi condotta da Cugurullo (2021) evidenzia inoltre come molte amministrazioni pubbliche implementino soluzioni di IA senza adeguate valutazioni preliminari di impatto sui diritti fondamentali, creando potenziali vulnerabilità legali e sociali. Il presente articolo intende contribuire a colmare questo gap, analizzando un framework integrato per l'implementazione dell'IA generativa nella PA che bilanci innovazione tecnologica e compliance normativa.

1. Quadro teorico e letteratura di riferimento

L'Intelligenza artificiale generativa rappresenta un sottoinsieme dell'IA caratterizzato dalla capacità di creare contenuti originali piuttosto che limitarsi ad analizzare dati esistenti. Secondo la definizione di Radford *et al.* (2019), l'IAG comprende modelli computazionali progettati per generare contenuti nuovi che imitano caratteristiche statistiche dei dati di addestramento. Questi sistemi, basati principalmente su architetture *transformer* e *deep learning*, hanno conosciuto un'accelerazione esponenziale negli ultimi anni, culminando con modelli come GPT (Generative Pre-trained Transformer), DALL-E e Midjourney, che hanno democratizzato l'accesso a potenti strumenti generativi. La loro adozione nella PA introduce potenzialità trasformative, ma solleva anche questioni uniche rispetto ai sistemi di IA tradizionali, particolarmente in termini di affidabilità, controllo dei contenuti e responsabilità giuridica.

La letteratura accademica sull'implementazione dell'IA nel settore pubblico ha iniziato a espandersi significativamente, ma presenta ancora lacune rilevanti quando si focalizza specificamente sui sistemi generativi. Wirtz e Müller (2023) hanno condotto una revisione sistematica identificando 135 articoli accademici sull'IA nella PA pubblicati tra 2000 e 2022, rilevando come solo il 12% affronti questioni di governance e il 7% aspetti etici. Ancora più esigua è la letteratura specifica sull'IA generativa nella PA, con studi pionieristici come quello di Criado e Gil-Garcia (2019) che esplorano le potenzialità delle chatbot avanzate nell'erogazione di servizi pubblici, ma sottolineano la carenza di framework implementativi completi.

Il quadro teorico per l'analisi dell'IAG nella PA può essere articolato attraverso tre prospettive complementari. La prima è la teoria del valore pubblico (Moore 1995; Twizeyimana e Andersson

1 Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n, 300/2008, (UE) n, 167/2013, (UE) n, 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (Regolamento sull'Intelligenza artificiale).

2019), che valuta l'impatto dell'IAG in termini di miglioramento dei servizi, efficienza operativa e fiducia istituzionale. La seconda prospettiva è quella dell'etica digitale (Floridi e Cowls 2019), che esamina le implicazioni morali dell'automazione decisionale nella sfera pubblica. La terza è la teoria dell'accettazione tecnologica nel settore pubblico (Venkatesh *et al.* 2016), che analizza i fattori organizzativi e culturali che influenzano l'adozione di tecnologie avanzate nelle burocrazie pubbliche.

Una revisione della letteratura evidenzia tre principali gap di ricerca. Innanzitutto, vi è una carenza di studi empirici sull'implementazione dell'IAG nelle PA, con poche analisi dei fattori di successo e delle barriere organizzative. Van der Voort *et al.* (2019) sottolineano come la maggior parte degli studi si concentri su aspetti tecnici piuttosto che su dimensioni implementative e organizzative. In secondo luogo, mancano framework integrati che armonizzino requisiti normativi, considerazioni etiche e necessità operative. Misuraca *et al.* (2020) evidenziano la frammentazione degli approcci, con modelli che tendono a focalizzarsi su singoli aspetti (conformità normativa, performance tecnica, accettazione organizzativa) senza una visione olistica. Infine, vi è una carenza di ricerche sulle specificità dell'IAG rispetto all'IA tradizionale nel contesto pubblico, con poca attenzione alle peculiari sfide poste dai sistemi generativi in termini di controllo dei contenuti, responsabilità giuridica e integrazione nei processi decisionali pubblici.

Come sottolineato da Veale e Borgesius (2021), l'efficace governance dell'IA nel settore pubblico richiede l'integrazione di molteplici framework regolatori in un approccio coerente e operativamente sostenibile, una sfida che questo articolo intende affrontare attraverso l'analisi di un modello implementativo specifico per l'IAG.

2. Metodologia della ricerca

La presente ricerca adotta un approccio metodologico misto, combinando analisi documentale, valutazione di framework e studio di casi. Questa triangolazione metodologica permette di esaminare il fenomeno dell'implementazione dell'IAG nella PA da prospettive complementari, aumentando la robustezza e la validità dei risultati. L'approccio si articola in tre fasi principali, ciascuna caratterizzata da specifiche tecniche di raccolta e analisi dei dati.

La prima fase ha comportato un'analisi sistematica della letteratura accademica sull'implementazione dell'IA nella PA, con particolare attenzione ai sistemi generativi. Seguendo la metodologia PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) proposta da Moher *et al.* (2009), sono stati selezionati 78 articoli accademici da database come Scopus, Web of Science e Google Scholar, utilizzando combinazioni di parole chiave come 'Generative Aì, 'public administration', 'implementation framework' e 'regulatory compliance'. Gli articoli sono stati sottoposti a screening e classificati secondo dimensioni analitiche predefinite (focus tecnologico, aspetti normativi, considerazioni etiche, barriere implementative). Questa fase ha permesso di identificare i principali gap nella letteratura e le dimensioni cruciali per un framework implementativo efficace.

La seconda fase ha comportato l'analisi approfondita del Framework di Intelligenza artificiale per la Pubblica amministrazione presentato nell'Allegato a questo contributo. L'analisi è stata condotta utilizzando una griglia valutativa sviluppata integrando il modello di Bannister e Connolly (2014) per la valutazione degli strumenti di e-government e l'approccio di Janowski *et al.* (2022) per l'analisi dei framework di governance dell'IA. La griglia ha permesso di valutare il framework secondo quattro dimensioni principali: completezza normativa, usabilità operativa, adeguatezza contestuale e solidità metodologica. Per ciascuna dimensione sono stati definiti indicatori specifici, valutati su una scala Likert a cinque punti.

La raccolta dati ha comportato l'analisi di documentazione progettuale, interviste semi-strutturate con 12 stakeholder chiave (dirigenti pubblici, tecnici, responsabili per la protezione dei dati) e l'osservazione diretta dei sistemi implementati. L'analisi dei casi è stata condotta utilizzando il framework sviluppato da Yin (2018) per gli studi di caso comparativi, con particolare attenzione all'identificazione di pattern ricorrenti, fattori abilitanti e barriere implementative.

I limiti metodologici includono la rappresentatività limitata dei casi studiati, che non possono catturare l'intera varietà di applicazioni dell'IAG nella PA, e la rapidità dell'evoluzione tecnologica e normativa, che può rendere alcuni risultati rapidamente obsoleti. Inoltre, essendo l'implementazione dell'IAG nella PA un fenomeno relativamente recente, non è

stato possibile valutare gli impatti a lungo termine delle soluzioni analizzate. Nonostante questi limiti, l'approccio metodologico adottato fornisce una base solida per l'analisi del framework implementativo e l'identificazione di linee guida operative per l'adozione dell'IAG nel settore pubblico.

3. Il Framework integrato per l'IA generativa nella PA

Il Framework di Intelligenza artificiale per la Pubblica amministrazione sviluppato e analizzato in questo studio rappresenta un approccio strutturato all'implementazione di sistemi di IA, inclusi quelli generativi, nel contesto del settore pubblico. La sua caratteristica distintiva risiede nell'integrazione di requisiti normativi, considerazioni etiche e necessità operative in un modello coerente e applicabile. Il framework si articola in quattro dimensioni principali: governance strategica, valutazione del rischio, sicurezza cibernetica e formazione, ciascuna adattata alle specificità dell'IAG e alle peculiarità del contesto pubblico.

La governance strategica costituisce il pilastro fondamentale del framework, ponendo le basi per un'implementazione allineata sia agli obiettivi istituzionali che ai requisiti normativi. Il modello propone una struttura di governance multilivello che comprende un Comitato etico-tecnologico interdisciplinare, un Registro centralizzato dei sistemi IA e un Piano triennale di adozione. Come evidenziato da Janssen e van der Voort (2016), un'efficace governance dell'IA richiede strutture istituzionali che bilancino competenze tecniche, giuridiche ed etiche in un approccio adattivo, principio che il framework implementa attraverso la composizione multidisciplinare del Comitato. Particolarmente rilevante per i sistemi generativi è la previsione di un AI Governance Board (AGB) che include non solo dirigenti e tecnici, ma anche esperti legali e rappresentanti etici provenienti da ONG e università, garantendo così un approccio multistakeholder alla governance.

L'allineamento agli standard internazionali costituisce un altro elemento distintivo del framework, che incorpora principi e metodologie dal NIST AI RMF, dall'ISO 42001 e dal MIT AI Risk Repository. Questa integrazione risponde alla raccomandazione di Lewis *et al.* (2022) di armonizzare standard tecnici e normativi in un approccio coerente alla governance dell'IA. Particolarmente significativa per l'IAG è l'adozione del Generative AI Profile del NIST, che introduce parametri specifici di robustezza per sistemi come chatbot e assistenti virtuali, affrontando così le peculiari sfide di controllo qualitativo poste dai sistemi generativi.

La valutazione del rischio rappresenta una componente cruciale del framework, particolarmente rilevante per l'IAG dato il suo carattere potenzialmente imprevedibile. Il modello propone un'innovativa Valutazione integrata che combina la Fundamental Rights Impact Assessment (FRIA) richiesta dall'AI Act con la Data Protection Impact Assessment (DPIA) prevista dal GDPR. Questa integrazione risponde alla necessità, evidenziata da Kaminski e Malgieri (2021), di superare la frammentazione delle valutazioni d'impatto in un approccio olistico alla gestione del rischio. Il framework prevede metodologie specifiche per l'identificazione del *bias* nei dataset, utilizzando strumenti come AI Fairness 360, e protocolli per test di robustezza contro attacchi adversarial, particolarmente rilevanti per i sistemi generativi esposti a tentativi di manipolazione.

Un elemento distintivo del framework è l'enfasi sui meccanismi '*human-in-the-loop*' per garantire supervisione umana e spiegabilità delle decisioni automatizzate. Come sottolineato da Veale e Borgesius (2021), i sistemi generativi richiedono modelli di supervisione adattivi che combinino controllo preventivo e monitoraggio continuo. Il framework implementa questo principio attraverso dashboard di monitoraggio in tempo reale che segnalano anomalie statistiche, soglie di incertezza per l'attivazione di canali umani e meccanismi di spiegabilità 'a gradiente' che forniscono diversi livelli di dettaglio esplicativo a seconda della criticità della decisione.

La sicurezza cibernetica costituisce un altro pilastro del framework, con particolare attenzione ai rischi specifici dei sistemi generativi. Il modello prevede un approccio di 'security by design' con particolare enfasi sulla protezione contro attacchi mirati come il *data poisoning*, particolarmente pericolosi per i sistemi basati su LLM. Significativa è l'introduzione di tecniche di crittografia omomorfica che permettono di elaborare dati sensibili senza esporli in chiaro, rispondendo così alla raccomandazione di Floridi *et al.* (2020) di implementare protezioni tecniche

che garantiscano privacy by design nei sistemi di IA pubblica. Il framework prevede inoltre protocolli specifici per la gestione degli incidenti che integrano i requisiti del GDPR con quelli dell'AI Act, garantendo notifiche tempestive e analisi post-mortem strutturate.

Infine, il framework dedica ampio spazio alla formazione e al cambiamento culturale, riconoscendo che l'implementazione efficace dell'IAG richiede non solo strumenti tecnici ma anche adeguate competenze organizzative. Il modello prevede corsi certificati su *AI Ethics*, simulazioni di caso su scenari reali e toolkit dipartimentali con linee guida specifiche, ad esempio sul prompt engineering per sistemi generativi. Questo approccio risponde alla raccomandazione di Mergel *et al.* (2019) di investire nello sviluppo di competenze trasversali che combinino comprensione tecnologica ed etica applicata.

4. Sfide implementative e fattori critici di successo

L'implementazione di sistemi di Intelligenza artificiale generativa nella Pubblica amministrazione presenta sfide distintive che richiedono strategie mirate. L'analisi dei casi studio e della letteratura ha permesso di identificare cinque aree critiche di sfida: resistenza organizzativa, carenza di competenze, integrazione nei processi esistenti, gestione dell'incertezza algoritmica e sostenibilità economica. Ciascuna di queste dimensioni presenta specificità legate all'IAG e richiede approcci gestionali differenziati.

La resistenza organizzativa emerge come una barriera significativa, manifestandosi in diverse forme: dalla diffidenza verso l'automazione di funzioni tradizionalmente umane al timore di perdita di controllo decisionale. Come evidenziato da Mergel *et al.* (2019**)**, l'introduzione di sistemi di IA nella PA si scontra con culture organizzative consolidate e strutture burocratiche rigide. Nel caso dell'IAG, questa resistenza assume caratteristiche peculiari legate alla percezione di imprevedibilità dei sistemi generativi. L'analisi del caso dell'amministrazione comunale italiana ha evidenziato come la resistenza fosse particolarmente accentuata per funzioni percepite come 'creative' o richiedenti discrezionalità amministrativa. Strategie efficaci di mitigazione hanno incluso approcci di co-design che hanno coinvolto i funzionari nella definizione dei parametri operativi del sistema generativo e la graduale introduzione dell'automazione, mantenendo significativi spazi di controllo umano.

La carenza di competenze rappresenta una sfida particolarmente acuta per l'IAG, che richiede non solo conoscenze tecniche ma anche capacità di valutazione critica degli output generati. Wirtz e Müller (2023) sottolineano come il digital divide interno alle PA costituisca un ostacolo maggiore all'adozione dell'IA rispetto a vincoli tecnologici o economici. I casi analizzati hanno evidenziato la necessità di un triplice livello di competenze: tecniche (comprendere i meccanismi di funzionamento), operative (saper interagire efficacemente con i sistemi) e valutative (saper giudicare criticamente gli output). Il framework affronta questa sfida attraverso programmi formativi multilivello, comunità di pratica tra enti e toolkit operativi specifici per diverse funzioni amministrative.

L'integrazione nei processi esistenti costituisce una sfida cruciale, particolarmente complessa per i sistemi generativi che possono produrre output non standardizzati. L'analisi del caso dell'agenzia governativa spagnola ha evidenziato difficoltà di integrazione della chatbot avanzata con i sistemi di gestione documentale esistenti e con i workflow amministrativi consolidati. Janssen e van der Voort (2016) sottolineano l'importanza di approcci di integrazione adattivi che bilancino innovazione e continuità operativa. Strategie efficaci hanno incluso l'adozione di interfacce standardizzate, lo sviluppo di API dedicate e la riprogettazione incrementale dei processi, mantenendo inizialmente percorsi paralleli (automatizzati e tradizionali) per garantire la continuità operativa durante la transizione.

La gestione dell'incertezza algoritmica rappresenta una sfida distintiva dell'IAG, i cui output possono presentare gradi variabili di accuratezza, rilevanza e conformità normativa. Questa sfida assume particolare rilevanza nel contesto pubblico, dove le decisioni amministrative devono rispettare stringenti requisiti di correttezza, imparzialità e tracciabilità. Il caso del Ministero olandese ha evidenziato come la gestione dell'incertezza algoritmica richieda un'articolata strategia di controllo qualità, combinando filtri preventivi (prompt engineering, controlli preventivi), monitoraggio in tempo reale (*confidence scores*, alert su anomalie) e revisione umana post-generazione. Come sottolineato da Floridi *et al.* (2020), i sistemi

generativi richiedono un continuum di supervisione umana calibrato sul livello di criticità della funzione.

La sostenibilità economica, infine, rappresenta una sfida significativa, considerando i costi di implementazione, mantenimento e aggiornamento dei sistemi generativi. L'analisi dei casi ha evidenziato come i costi iniziali di acquisizione tecnologica rappresentino solo una frazione del costo totale di possesso, con significativi investimenti richiesti per formazione, personalizzazione, integrazione e monitoraggio continuo. Il framework propone modelli di condivisione delle risorse tra amministrazioni, approcci incrementali all'implementazione e metodologie di valutazione costi-benefici che considerino non solo l'efficienza operativa, ma anche dimensioni di valore pubblico più ampie come trasparenza, accessibilità e fiducia istituzionale.

5. Raccomandazioni di policy per l'implementazione efficace

Sulla base dell'analisi del framework e dei gap implementativi identificati, è possibile formulare raccomandazioni di policy strutturate per guidare l'efficace adozione dell'IAG nella Pubblica amministrazione. Queste raccomandazioni sono organizzate su tre livelli complementari: strategico, tattico e operativo, fornendo così una roadmap completa per decisori pubblici e implementatori.

A livello strategico, è essenziale sviluppare una visione nazionale coordinata per l'adozione dell'IAG nella PA, superando l'attuale frammentazione delle iniziative. Questa visione dovrebbe tradursi in una Strategia nazionale per l'IA generativa nella PA, integrata nel più ampio Piano triennale per l'Informatica nella Pubblica amministrazione. La strategia dovrebbe definire priorità di investimento, casi d'uso prioritari e meccanismi di coordinamento inter-istituzionale. Come evidenziato da Misuraca *et al.* (2020), l'assenza di una visione strategica coerente rappresenta il principale ostacolo alla scalabilità delle iniziative di IA nel settore pubblico. Un elemento cruciale di questa strategia dovrebbe essere l'istituzione di un Centro di Competenza nazionale sull'IA generativa per la PA, con il mandato di sviluppare linee guida, fornire supporto tecnico e facilitare la condivisione di esperienze tra amministrazioni. Il centro potrebbe essere istituito presso AgID o altra struttura esistente, con un modello di governance che assicuri rappresentanza delle diverse tipologie di amministrazioni e adeguato coinvolgimento di stakeholder accademici e della società civile.

È inoltre raccomandabile lo sviluppo di un Regulatory Sandbox specifico per l'IAG nella PA, che permetta la sperimentazione controllata di applicazioni innovative in ambienti a ridotto rischio regolatorio. Questo meccanismo, già sperimentato con successo in domini come fintech e health-tech, permetterebbe di testare applicazioni generative in ambiti delimitati prima di una più ampia implementazione, facilitando l'innovazione senza compromettere la tutela dei diritti fondamentali. Come sottolineato da Lodge e Mennicken (2019), i *regulatory sandboxes* creano spazi protetti per l'innovazione responsabile, particolarmente preziosi in domini ad alta regolamentazione come la PA.

A livello tattico, è raccomandabile lo sviluppo di Toolkit implementativi settoriali che operazionalizzino il framework generale per specifici domini applicativi (es. servizi sociali, urbanistica, gestione documentale), fornendo linee guida contestualizzate, template per valutazioni d'impatto e checklist operative. L'approccio '*one size fits all*' risulta infatti inadeguato considerando l'eterogeneità delle funzioni amministrative e dei relativi requisiti. Tali toolkit dovrebbero includere modelli di governance adattati alle dimensioni organizzative dell'ente, semplificando le strutture proposte dal framework per piccole amministrazioni senza rinunciare ai principi fondamentali di supervisione multi-stakeholder e di valutazione di impatto.

È inoltre cruciale lo sviluppo di un AI Evaluation Framework standardizzato che permetta alle PA di valutare sistemi di IAG prima dell'acquisizione, verificandone non solo le performance tecniche, ma anche conformità normativa, spiegabilità e robustezza. Tale framework dovrebbe essere incorporato nelle linee guida per gli appalti pubblici, richiedendo ai fornitori evidenze documentate di conformità. Janssen *et al.* (2020) sottolineano come la fase di procurement rappresenti un'opportunità cruciale per incorporare requisiti di responsible AI nelle implementazioni pubbliche, un principio che dovrebbe guidare l'aggiornamento delle procedure di appalto per soluzioni di IAG.

A livello operativo, è raccomandabile l'implementazione di un sistema di Certificazione delle competenze sull'IAG nella PA, che definisca

profili professionali (es. AI Ethics Officer, LLM Prompt Engineer, AI Data Steward) e percorsi formativi certificati. Tale sistema dovrebbe essere integrato nei piani di fabbisogno del personale e nei percorsi di sviluppo professionale, creando incentivi per l'acquisizione di competenze specialistiche. Il sistema di certificazione potrebbe essere sviluppato in collaborazione con università ed enti formativi, garantendo allineamento con competenze richieste dal mercato e specificità del contesto pubblico.

È inoltre essenziale lo sviluppo di una Piattaforma collaborativa che faciliti la condivisione di risorse (modelli, dataset, prompt *libraries*, valutazioni di impatto) tra amministrazioni, riducendo duplicazioni e facilitando l'adozione di soluzioni già testate. Tale piattaforma dovrebbe implementare principi di 'governo come piattaforma' (ÒReilly 2011), facilitando ecosistemi collaborativi piuttosto che approcci top-down. Un esempio di successo è rappresentato dal programma *Developers* Italia, che potrebbe essere esteso con sezioni dedicate a risorse per implementazioni di IAG.

Infine, è raccomandabile l'implementazione di un sistema di monitoraggio e apprendimento continuo che raccolga dati sull'implementazione dell'IAG nelle diverse amministrazioni, identificando pattern di successo, barriere ricorrenti e impatti su diversi stakeholder. Tale sistema dovrebbe adottare metodologie miste (quantitative e qualitative) e prevedere meccanismi di feedback da cittadini e funzionari pubblici. Come sottolineato da de Bruijn *et al.* (2020), i sistemi di performance management per l'IA pubblica devono superare metriche puramente tecniche per incorporare dimensioni di valore pubblico più ampie, un principio che dovrebbe guidare lo sviluppo di framework valutativi per l'IAG nella PA.

Conclusioni e direzioni future di ricerca

L'implementazione dell'Intelligenza artificiale generativa nella Pubblica amministrazione rappresenta una frontiera di innovazione con potenzialità trasformative, ma caratterizzata da sfide complesse che richiedono approcci strutturati e contestualmente sensibili. L'analisi del Framework di Intelligenza artificiale per la Pubblica amministrazione ha evidenziato come un'efficace adozione dell'IAG richieda l'integrazione di molteplici dimensioni: governance strategica, valutazione del rischio, sicurezza cibernetica e formazione, ciascuna adattata alle specificità dei sistemi generativi e alle peculiarità del contesto pubblico. Il framework analizzato offre un approccio olistico che bilancia innovazione tecnologica, compliance normativa e tutela dei diritti fondamentali, rispondendo alla necessità di armonizzare requisiti dell'AI Act, del GDPR e degli standard internazionali in un modello operativamente sostenibile.

L'analisi comparativa tra il framework ideale e le concrete esperienze implementative ha evidenziato significativi gap da colmare, particolarmente nelle aree della governance multi-stakeholder, delle metodologie di valutazione del rischio, dei meccanismi '*human-in-the-loop*' e delle pratiche di sicurezza cibernetica. Questi gap non derivano necessariamente da carenze del framework, ma piuttosto dalle complessità organizzative, tecniche e culturali che caratterizzano i contesti applicativi reali. La loro identificazione offre opportunità di affinamento del framework e sviluppo di strategie implementative più efficaci, in un processo iterativo di apprendimento e adattamento.

Le raccomandazioni di policy proposte offrono una roadmap strutturata per decisori pubblici e implementatori, articolata su tre livelli complementari: strategico (Strategia nazionale, Centro di competenza, Regulatory Sandbox), tattico (Toolkit implementativi settoriali, AI Evaluation Framework) e operativo (Certificazione delle competenze, Piattaforma collaborativa, sistema di monitoraggio). Queste raccomandazioni non rappresentano soluzioni definitive, ma piuttosto orientamenti per un percorso di trasformazione che richiederà continuo adattamento alle evoluzioni tecnologiche, normative e organizzative.

Questa ricerca contribuisce alla letteratura sull'IA nella PA colmando parzialmente i gap identificati, particolarmente riguardo ai framework implementativi specifici per sistemi generativi e all'integrazione di requisiti normativi in modelli operativamente sostenibili. Tuttavia, diversi aspetti richiedono ulteriore approfondimento attraverso ricerche future. Innanzitutto, è necessario sviluppare metodologie più robuste per la valutazione empirica dell'impatto dell'IAG nella PA, superando approcci aneddotici per misurare sistematicamente effetti su efficienza amministrativa, qualità dei servizi e fiducia istituzionale. Come sottolineato da Wirtz e Müller (2023), la carenza di studi empirici longitudinali limita la comprensione degli effetti trasformativi

dell'IA nel settore pubblico, una lacuna che future ricerche dovrebbero colmare.

È inoltre cruciale approfondire le dinamiche di accettazione organizzativa dell'IAG nella PA, identificando fattori culturali, strutturali e individuali che influenzano resistenze e adozione. Particolarmente rilevante sarebbe l'analisi delle interazioni tra sistemi generativi e lavoratori pubblici in contesti reali, esaminando come si riconfigurino ruoli professionali, competenze e identità lavorative. Mergel *et al.* (2019) sottolineano come la trasformazione dei ruoli professionali rappresenti una delle dimensioni meno esplorate dell'adozione dell'IA nella PA, un gap che richiederebbe approcci etnografici e longitudinali.

Un terzo filone di ricerca riguarda i modelli di governance dell'IAG nella PA, con particolare attenzione a meccanismi di accountability democratica e partecipazione civica. Come evidenziato da Veale e Borgesius (2021), l'automazione decisionale nel settore pubblico solleva questioni fondamentali di responsabilità democratica che trascendono aspetti puramente tecnici o legali, richiedendo ricerche interdisciplinari che integrino prospettive giuridiche, politologiche e sociologiche.

Infine, è necessario sviluppare approcci valutativi che misurino l'impatto dell'IAG non solo in termini di efficienza, ma rispetto a valori pubblici più ampi come equità, trasparenza e legittimità. Tali ricerche dovrebbero adottare metodologie miste, combinando analisi quantitative di performance con indagini qualitative su percezioni ed esperienze di diversi stakeholder, in un approccio valutativo multidimensionale e partecipativo.

In conclusione, l'IAG rappresenta per la PA non solo una sfida tecnologica, ma una più ampia opportunità di ripensamento istituzionale, in cui innovazione digitale e valori pubblici fondamentali devono integrarsi in un nuovo paradigma di governance. Come sottolineato da Janowski *et al.* (2022), il successo dell'IA nel settore pubblico si misurerà non solo in termini di efficienza operativa, ma nella capacità di rafforzare legittimità democratica e fiducia istituzionale, una prospettiva che dovrebbe guidare sia l'implementazione pratica, che la ricerca futura in questo campo.

Allegato
Framework di Intelligenza artificiale per la Pubblica amministrazione: integrazione normativa e strumenti operativi

1. Governance strategica e allineamento normativo

1.1 Quadro Giuridico Integrato

L'AI Act e il GDPR costituiscono il fondamento giuridico per l'uso dell'IA nella PA. L'AI Act classifica i sistemi in quattro categorie di rischio (inaccettabile, alto, limitato, minimo), imponendo obblighi specifici per quelli ad alto rischio, come i sistemi di valutazione creditizia o di reclutamento pubblico. Il GDPR, invece, regola il trattamento dei dati personali, richiedendo valutazioni d'impatto (DPIA) e garanzie per i diritti degli interessati.

Per allinearsi a questi requisiti, la PA deve adottare un **modello di governance multilivello**:

- **Comitato etico-tecnologico**: organo interdisciplinare per supervisionare progetti IA, verificando la conformità all'AI Act e ai principi del GDPR. Deve includere rappresentanti legali, esperti di cybersecurity, e delegati per la protezione dei dati (DPO).
- **Registro centralizzato dei sistemi IA**: catalogo aggiornato di tutti i sistemi in uso, con metadata su finalità, rischi, basi giuridiche e cicli di audit. Il registro deve essere accessibile alle autorità di vigilanza e aggiornato trimestralmente.
- **Piano triennale di adozione IA**: allineato alle linee guida AGID e al Piano triennale per l'Informatica 2024-2026, con priorità definite tramite analisi costi-benefici e impatto sociale. Il Piano deve identificare almeno tre progetti pilota a basso rischio (es. ottimizzazione flussi documentali) entro il primo anno.

Checklist operativa:

- Designare un responsabile IA per ogni ente, con competenze in diritto digitale e gestione del rischio.
- Integrare il registro IA con il sistema di gestione documentale (es. Protocollo informatico nazionale).

- Sottoporre il Piano triennale a consultazione pubblica per 60 giorni prima dell'adozione.

1.2 Allineamento agli standard internazionali

Il framework incorpora elementi da:

- **NIST AI RMF**: gestione del rischio tramite categorie di threat (sicurezza, *bias*, privacy) e mitigazioni basate su livelli di maturità. Le PA devono mappare i controlli del NIST AI RMF su quattro funzioni: Govern, Map, Measure, Manage.
- **ISO 42001**: certificazione per i sistemi di gestione dell'IA, con focus su accountability e tracciabilità delle decisioni algoritmiche. Richiede l'identificazione di 39 controlli nell'Annex A, tra cui auditabilità dei modelli e monitoraggio dei *bias*.
- **MIT AI Risk Repository**: catalogazione dei rischi specifici per contesti pubblici (es. discriminazione in servizi sociali). Da utilizzare per scenari di stress-testing.

Checklist operativa:

- Condurre un gap analysis tra i controlli ISO 42001 e i processi esistenti, con priorità sulle lacune ad alto impatto.
- Adottare il Generative AI Profile del NIST per sistemi come chatbot, con test su 15 parametri di robustezza.
- Integrare nel registro IA un modulo per la tracciabilità dei dataset, includendo provenienza, licenze e potenziali *bias*.

1.3 Creazione di un AI Governance Board

La creazione di un **AI Governance Board** (AGB) rappresenta un passo cruciale per garantire che i sistemi di IA siano progettati e implementati rispettando i principi di trasparenza, equità e responsabilità. L'AGB dovrebbe includere:

- **Dirigenti della PA**: per assicurare che il progetto sia allineato agli obiettivi strategici dell'amministrazione.
- **Esperti legali**: per monitorare la conformità normativa, in particolare riguardo al GDPR e all'AI Act.
- **Rappresentanti etici**: provenienti da ONG, università e associazioni civiche, per supervisionare l'aderenza ai principi etici.
- **Stakeholder tecnici**: ingegneri e data scientist per fornire supporto operativo e garantire la robustezza tecnica del sistema.

Compiti principali dell'AI Governance Board

1. Definizione delle policy sull'uso dell'IA

- Stabilire linee guida per l'utilizzo responsabile dell'IA, come la necessità di rendere spiegabili tutte le decisioni automatizzate.
- Adottare criteri di supervisione umana per le decisioni ad alto impatto.

2. Monitoraggio degli impatti sociali ed economici

- Analizzare come le decisioni automatizzate influenzano gruppi vulnerabili e verificare che non si creino discriminazioni sistemiche.
- Valutare i benefici economici derivanti dall'automazione, come la riduzione del carico amministrativo.

3. Supervisione dei KPI

- Stabilire indicatori di performance come:
  - **Accuratezza delle decisioni**: Maggiore del 97%.
  - **Percentuale di decisioni contestate**: Inferiore al 5%.
  - **Tempo medio di risposta**: Inferiore a 24 ore.

1.4 Policy e linee guida operative

Le policy operative devono essere documentate e aggiornate regolarmente per garantire che il sistema di IA rispetti le normative e i principi etici. Le linee guida includono:

- **Trasparenza**: ogni decisione automatizzata deve essere accompagnata da una spiegazione comprensibile per l'utente, utilizzando strumenti di Explainable AI (XAI).
- **Equità**: i dataset utilizzati devono essere rappresentativi della popolazione servita per prevenire discriminazioni indirette.
- **Protezione dei dati**: devono essere implementate misure di sicurezza avanzate, come l'anonimizzazione e la crittografia dei dati personali.
- **Supervisione umana**: le decisioni che superano soglie di rischio predefinite devono essere esaminate manualmente da operatori esperti.

2. Valutazione del Rischio e Compliance Operativa

2.1 Modello di Valutazione Integrata (DPIA + FRIA)

L'AI Act richiede una **Fundamental Rights Impact Assessment (FRIA)** per i sistemi ad alto rischio, da condurre in parallelo alla DPIA del GDPR. Per la PA, questo si traduce in:

1. **Mappatura degli stakeholder**: cittadini, dipendenti, autorità di vigilanza. Utilizzare matrici di coinvolgimento basate sul livello di impatto (es. alto per sistemi di welfare).
2. **Analisi dei dati**: verifica della liceità del trattamento (base giuridica, minimizzazione dati, consenso esplicito per categorie speciali). Per dataset storici, applicare tecniche di *debiasing* come *reweighting* o *adversarial debiasing*.
3. **Valutazione del *bias***: utilizzo di tool come AI Fairness 360 (IBM) per rilevare discriminazioni in dataset storici. Soglia di disparità accettabile: <5% tra gruppi protetti.
4. **Test di robustezza**: simulazione di attacchi *adversarial* per verificare la resilienza dei modelli. Per sistemi NLP, testare con prompt injection e dati *out-of-distribution*.

Checklist operativa:

- Documentare la base giuridica per ogni trattamento (GDPR Art.6), specificando se si applicano eccezioni per 'interesse pubblico' (Art. 6(1)(e)).
- Implementare meccanismi di *anonymization* o *pseudonymization* per dataset sensibili, con valutazione del rischio di re-identificazione tramite k-anonymity (k≥3).
- Redigere report di conformità AI Act (Allegato III) per sistemi ad alto rischio, includendo piani di mitigazione per almeno tre scenari di *failure*.

2.2 Human-in-the-loop e spiegabilità

L'art. 22 GDPR e l'AI Act impongono il diritto a un intervento umano su decisioni automatizzate. Nella PA, ciò si applica a:

- **Processi amministrativi**: es. assegnazione di benefici sociali, con dashboard di monitoraggio in tempo reale che segnalano anomalie statistiche (es. deviazione >2σ dalla media storica).
- **Comunicazione con i cittadini**: chatbot devono chiarire quando non sono in grado di rispondere e attivare canali umani. Soglia di incertezza: confidence score <85%.

Checklist operativa:

- Integrare dashboard di monitoraggio con metriche di *fairness* (disparità di trattamento per gruppi demografici) aggiornate ogni 24 ore.
- Fornire spiegazioni 'a gradiente' (*layer-wise relevance propagation*) per decisioni critiche, con report generati in formato XAI (XML for AI Explanations).
- Addestrare dipendenti su protocolli di *override* dei sistemi IA, con simulazioni semestrali su casi reali (es. correzione manuale di assegnazioni errate di alloggi popolari).

3. Sicurezza Cibernetica e Resilienza

3.1 Protezione dei Sistemi IA

L'AI Act richiede che i sistemi ad alto rischio siano progettati con sicurezza by design. Per la PA:

- **Modello di threat intelligence**: collaborazione con CERT-AgID per identificare attacchi mirati (es. *data poisoning*). Implementare *honeypot* con dati sintetici per rilevare tentativi di manipolazione.
- **Crittografia omomorfica**: per elaborare dati sensibili senza esporli in chiaro. Utilizzare librerie verificabili (es. Microsoft SEAL) con audit di sicurezza semestrali.
- **Controlli di accesso RBAC (Role-Based)**: limitazione dei privilegi in base al principio del minimo privilegio. Ruoli predefiniti: Data Scientist (sola lettura modelli), Auditor (accesso completo), Operatore (solo inferenza).

Checklist operativa:

- Eseguire penetration test semestrali su API e modelli esposti, utilizzando framework come Adversarial Robustness Toolbox.
- Utilizzare hardware *trusted execution environment* (TEE) per modelli critici, con certificazione Common criteria EAL4+.
- Implementare SIEM (Security Information and Event Management) con regole specifiche per *anomaly detection* IA, come variazioni >30% nell'*accuracy* di modello.

3.2 Gestione degli incidenti

In caso di *data breach* o malfunzionamenti algoritmici, la PA deve attivare protocolli conformi al GDPR Art. 33-34 e all'AI Act Art. 17:

- **Notifica all'Autorità garante**: entro 72 ore dalla rilevazione, con descrizione del sistema IA coinvolto, impatto stimato su diritti fondamentali, e misure correttive immediate.
- **Comunicazione agli interessati**: se il *breach* comporta rischi per diritti e libertà. Utilizzare canali prioritari (SMS, e-mail certificata) con linguaggio chiaro (livello A2 QCER).
- **Post-mortem analitico**: identificazione delle root cause tecniche e organizzative. Per incidenti gravi (>10.000 persone coinvolte), coinvolgere esperti terzi indipendenti 4.

Checklist operativa:

- Mantenere un inventario aggiornato dei contatti delle autorità competenti (Garante privacy, AGID) in formato *machine-readable* (es. JSON-LD).
- Predisporre template precompilati per le notifiche, con campi obbligatori indicati dall'AI Act Annex V[2].
- Integrare nel SIEM un modulo di *forensic automation* per la raccolta automatica di log rilevanti (es. modifiche ai modelli nelle ultime 72 ore).

4. Formazione e Cambiamento culturale

La PA deve superare il digital divide interno attraverso:

- **Corsi certificati su AI Ethics**: in collaborazione con università (es. Università di Genova) e organismi come AgID. Durata minima: 40 ore, con esame finale su casi reali.
- **Simulazioni di caso**: esercitazioni su scenari reali (es. correzione di bias in algoritmi di assegnazione alloggi) utilizzando piattaforme sandbox con dati sintetici.
- **Toolkit per Dipartimenti**: linee guida su prompt engineering per sistemi generativi (es. ChatGPT per redazione atti), con whitelist di termini vietati (es. "cittadino onorevole").

Checklist operativa:

- Definire un piano formativo annuale con KPI di completamento (minimo 80% del personale entro Q4).
- Istituire comunità di pratica tra enti locali per condividere best practice.

5. Processo di Gestione dei rischi

5.1. Identificazione e Classificazione dei Rischi

La gestione dei rischi è un elemento cruciale nella progettazione e implementazione di sistemi di Intelligenza artificiale (IA) nelle Pubbliche Amministrazioni (PA). L'identificazione e la classificazione dei

2 Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024.

rischi mirano a prevenire problematiche etiche, tecniche e operative, garantendo che il sistema sia equo e conforme alle normative.

Rischi etici

- **Discriminazione indiretta**: la presenza di bias nei dati di addestramento può causare trattamenti ingiusti per gruppi specifici di utenti.
- **Mancanza di trasparenza**: decisioni non spiegabili possono ridurre la fiducia degli utenti e compromettere la responsabilità amministrativa.

Rischi tecnici

- **Attacchi informatici**: manipolazione intenzionale dei dati di input (data poisoning) per influenzare i risultati.
- **Errori sistematici**: correlazioni spurie nei dataset che portano a decisioni errate.

Rischi operativi

- **Scarsa supervisione**: mancanza di competenze tecniche tra il personale amministrativo per gestire le complessità del sistema.
- **Incapacità di adattamento**: difficoltà nel rispondere rapidamente a cambiamenti normativi o sociali.

5.2. Piano di Mitigazione

Un piano di mitigazione strutturato è essenziale per affrontare i rischi identificati, riducendo l'impatto negativo e aumentando la resilienza del sistema. Le principali attività includono:

Validazione dei dataset

- **Obiettivo**: garantire che i dati utilizzati siano rappresentativi e bilanciati.
- **Esempio pratico**: una PA ha riscontrato che un modello per distribuire fondi alle imprese favoriva le grandi aziende urbane. Per correggere questa disparità, è stato introdotto un filtro per normalizzare i dati in base a settore e dimensione aziendale, migliorando l'equità nelle assegnazioni.

Simulazioni e test di robustezza

- **Obiettivo**: verificare il comportamento del sistema in scenari critici.
- **Esempio pratico**: test di simulazione hanno evidenziato che il sistema era vulnerabile a manipolazioni dei dati di input. Sono stati quindi implementati controlli di integrità e meccanismi di rilevamento per prevenire tali manipolazioni.

Strumenti utilizzati

- **Fairlearn**: per identificare e mitigare i bias nei dataset.
- **What-If Tool**: per simulare scenari ipotetici e analizzare l'impatto delle decisioni del modello.

5.3. Monitoraggio continuo

Il monitoraggio continuo garantisce che il sistema mantenga performance ottimali e conformità alle normative, adattandosi alle esigenze emergenti.

Le linee guida AgID introducono un **Maturity Model** su cinque livelli (Initial, Managed, Defined, Measured, Optimizing), aiutando gli enti a autovalutarsi. Parametri chiave includono:

- **Allineamento Strategico**: integrazione dell'IA negli obiettivi istituzionali.
- **Gestione del Rischio**: frequenza di aggiornamento delle FRIA.
- **Impatto Sociale**: metriche di soddisfazione cittadina post-implementazione IA.

Le modalità tramite cui possono essere individuati i punti da migliorare possono essere:

1. Audit Trimestrali

- **Obiettivo**: valutare regolarmente l'equità, l'accuratezza e la trasparenza del sistema.
- **Esempio pratico**: un sistema per l'assegnazione di borse di studio viene sottoposto ad audit trimestrali per identificare variazioni nei criteri di selezione che potrebbero indicare la presenza di bias.

2. Raccolta di Feedback dai Cittadini

- **Obiettivo**: coinvolgere gli utenti finali per migliorare il sistema sulla base delle loro esperienze e segnalazioni.
- **Esempio pratico**: una PA ha implementato un'app dedicata che consente agli utenti di contestare decisioni errate o richiedere chiarimenti. Le segnalazioni raccolte sono state utilizzate per migliorare l'accuratezza del modello.

3. Revisione iterativa dei Modelli

- **Obiettivo**: aggiornare periodicamente i modelli per rispondere a cambiamenti nei contesti normativi e sociali.
- **Esempio pratico**: una PA aggiorna il modello utilizzato per l'assegnazione di contributi agricoli per includere le nuove normative europee sui sussidi verdi.

5.4. Processo di Audit

Un elemento cruciale per garantire la conformità e il miglioramento continuo dei sistemi di Intelligenza artificiale (IA) nella Pubblica amministrazione (PA) è rappresentato dal processo di audit. Questo capitolo descrive le diverse fasi dell'audit, che includono verifiche iniziali, periodiche e straordinarie, illustrando casi concreti e approcci metodologici per una corretta implementazione.

5.4.1. Audit Iniziale

Obiettivo

L'audit iniziale mira a garantire che il sistema di IA sia conforme ai requisiti normativi, tecnici ed etici prima della sua implementazione operativa.

Attività principali

1. Validazione della qualità dei dati

- **Descrizione**: i dataset utilizzati per addestrare il modello devono essere rappresentativi, completi e privi di bias significativi.
- **Caso pratico**: una PA che sviluppa un sistema per il calcolo delle tasse ha verificato che i dati storici includessero correttamente tutte le fasce di reddito e non presentassero anomalie che potessero influire sull'equità delle decisioni.

2. Test di robustezza del modello

- **Descrizione**: il modello deve essere sottoposto a test per valutare la sua capacità di gestire input inattesi o incompleti.
- **Caso pratico**: la PA ha eseguito test simulando errori nei dati di input, come valori fuori scala o incompleti, per garantire che il sistema rispondesse con risultati affidabili o generasse avvisi appropriati.

3. Valutazione dei rischi etici

- **Descrizione**: un'analisi dettagliata deve identificare potenziali impatti etici, come discriminazioni indirette o mancanza di trasparenza.
- **Caso pratico**: durante l'audit di un sistema di assegnazione di borse di studio, è stato verificato che il modello non penalizzasse inconsapevolmente studenti provenienti da scuole situate in aree svantaggiate.

5.4.2. Audit Periodico

Obiettivo

Gli audit periodici garantiscono che il sistema di IA continui a operare in modo conforme, identificando eventuali nuove problematiche emerse durante il suo utilizzo.

Attività Principali

1. Analisi Campionata delle Decisioni

- **Descrizione**: viene analizzato un campione rappresentativo delle decisioni del sistema per verificare l'equità e l'aderenza alle policy operative.
- **Caso pratico:** una PA che gestisce un sistema per l'assegnazione di contributi culturali controlla trimestralmente che le decisioni siano distribuite equamente tra enti pubblici e privati, evitando concentrazioni eccessive.

2. Aggiornamento del Registro dei Rischi

- **Descrizione**: le valutazioni periodiche dei rischi vengono integrate con nuove informazioni derivanti dall'uso reale del sistema.

- **Caso pratico:** durante un audit trimestrale, la PA ha rilevato che il sistema stava privilegiando imprese con maggiori risorse tecniche per completare correttamente la domanda online. Questo ha portato all'introduzione di una guida interattiva per supportare le imprese meno tecnologicamente preparate.

3. Monitoraggio dei KPI

- **Descrizione**: gli indicatori di performance vengono confrontati con le soglie target per garantire che il sistema operi entro i parametri stabiliti.
- **Caso pratico**: un sistema di gestione dei trasporti pubblici utilizza KPI come il tempo di risposta alle segnalazioni degli utenti e la precisione nelle previsioni di affluenza.

5.4.3. Audit Straordinario

Obiettivo

L'audit straordinario viene attivato in risposta a incidenti critici o segnalazioni significative, con l'obiettivo di identificare e risolvere le cause profonde dei problemi.

Attività Principali

1. Analisi Forense dei Log

- **Descrizione**: vengono analizzati i registri delle attività del sistema per tracciare le decisioni problematiche e identificare eventuali anomalie.
- **Caso pratico:** in seguito a una segnalazione di errori nell'assegnazione delle borse di studio, la PA ha analizzato i log e scoperto che un parametro mal configurato aveva penalizzato alcune richieste legittime.

2. Revisione dei Processi Operativi

- **Descrizione**: l'intero processo decisionale del sistema viene riesaminato per individuare eventuali lacune nelle policy o nelle procedure.
- **Caso pratico:** dopo un audit straordinario, una PA ha introdotto controlli aggiuntivi per garantire che le modifiche ai parametri del modello fossero approvate da un team multidisciplinare.

3. Proposta di Miglioramenti

- **Descrizione**: sulla base dei risultati dell'audit, vengono sviluppate raccomandazioni per prevenire il ripetersi degli incidenti e migliorare la resilienza del sistema.
- **Caso pratico:** una PA ha implementato una dashboard di monitoraggio in tempo reale per rilevare automaticamente eventuali discrepanze nelle decisioni del sistema.

Bibliografia

Bannister F., Connolly R. (2014), ICT, public values and transformative government. A framework and programme for research, *Government Information Quarterly*, 31, n.1, pp.119-128

Criado J.I., Gil-Garcia J.R. (2019), Creating public value through smart technologies and strategies: From digital services to artificial intelligence and beyond, *International Journal of Public Sector Management*, 32, n.5, pp.438-450

Cugurullo F. (2021), Urban artificial intelligence: From automation to autonomy in the smart city, *Frontiers in Sustainable Cities*, n.3, Article 38

de Bruijn H., Janssen M., Warnier M. (2020), Accountability challenges in the transnational algorithmic governance: The case of public sector algorithms, in Ebers M., Cantero Gamito M. (eds), *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges,* Berlin, Springer, pp.137-156

Floridi L., Cowls J. (2019), A unified framework of five principles for AI in society, *Harvard Data Science Review*, 1, n.1 <DOI:10.1162/99608f92.8cd550d1>

Floridi L., Cowls J., King T. C., Taddeo M. (2020), How to design AI for social good: Seven essential factors, *Science and Engineering Ethics*, 26, n.3, pp.1771-1796

Janowski T., Estevez E., Ojo A. (2022), Digital government for sustainable development, *Government Information Quarterly*, 39, n.2, 101667

Janssen M., Brous P., Estevez E., Barbosa L.S., Janowski T. (2020), Data governance: Organizing data for trustworthy Artificial Intelligence, *Government Information Quarterly*, 37, n.3, 101493

Janssen M., van der Voort H. (2016), Adaptive governance: Towards a stable, accountable and responsive government, *Government Information Quarterly*, 37, n.1, 101411

Kaminski M.E., Malgieri G. (2021), Algorithmic impact assessments under the GDPR: Producing multi-layered explanations, *International Data Privacy Law*, 11, n.2, pp.125-144

Lewis D., Mooney J., Brady M., Sánchez-Ortiz V., Litchfield I. (2022), The adoption of artificial intelligence in public services: A systematic literature review, *Public Administration Review*, 82, n.6, pp.1049-1064

Lodge M., Mennicken A. (2019), The importance of regulation of and by algorithm, in Yeung M., Lodge K. (eds.), *Algorithmic Regulation*, Oxford, Oxford University Press, pp.1-12

Mergel I., Edelmann N., Haug N. (2019), Defining digital transformation: Results from expert interviews, *Government Information Quarterly*, 36, n.4, 101385

Misuraca G., van Noordt C., Boukli A. (2020), The use of AI in public services: Results from a preliminary mapping across the EU, in *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2020)*, New York, ACM, pp.90-99

Moher D., Liberati A., Tetzlaff J., Altman D.G. (2009), Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement, *British Medical Journal*, n.339, b2535

Moore M.H. (1995), *Creating public value: Strategic management in government*, Harvard, Harvard University Press

ÒReilly T. (2011), Government as a platform. Innovations: Technology, *Governance, Globalization*, 6, n.1, pp.13-40

Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I. (2019), Language models are unsupervised multitask learners, *OpenAI Blog*, 1, n.8, p.9

Sun T.Q., Medaglia R. (2019), Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare, *Government Information Quarterly*, 36, n.2, pp.368-383

Twizeyimana J.D., Andersson A. (2019), The public value of E-Government. A literature review, Government Information Quarterly, 36, n.2, pp.167-178

van der Voort H.G., Klievink A.J., Arnaboldi M., Meijer A.J. (2019), Rationality and politics of algorithms, Will the promise of big data survive the dynamics of public decision making?, *Government Information Quarterly*, 36, n.1, pp.27-38

Veale M., Borgesius F.Z. (2021), Demystifying the Draft EU Artificial Intelligence Act, *Computer Law Review International*, 22, n.4, pp.97-112

Venkatesh V., Thong J.Y., Xu X. (2016), Unified theory of acceptance and use of technology: A synthesis and the road ahead, *Journal of the Association for Information Systems*, 17, n.5, pp.328-376

Wirtz B.W., Müller W.M. (2023), Artificial Intelligence and public sector management: A systematic literature review, *International Journal of Public Administration*, 46, n.1, pp.34-51

Wirtz B. W., Weyerer J. C., Geyer C. (2022), Artificial intelligence in the public sector: Foundations, applications, consequences, and challenges, *International Journal of Public Administration*, 45, n.1, pp.80-89

Yin R.K. (2018), *Case study research and applications: Design and methods*, Los Angeles, Sage Publications

Gaetano Riccio
gaetano.riccio84@gmail.com
È Docente a contratto presso l'Università degli Studi di Napoli Federico II, corso di Sistemi di elaborazione delle informazioni, e presso l'Università degli Studi della Campania Luigi Vanvitelli, corso Sistemi informativi per la PA. È Dirigente dell'Area Informatica della Giunta della Regione Campania, Direttore del Dipartimento Intelligenza artificiale di supporto alle Pubbliche amministrazioni dell'Ente nazionale Intelligenza artificiale (ENIA) e Componente esterno del Comitato d'indirizzo dei Corsi di studio dell'Area economica del Dipartimento di Economia, management e metodi quantitativi.