Ancora sulla Governance dell’AI

Come è stato già più volte rimarcato anche da fonti autorevoli la governance dell’AI si basa su una visione e un approccio Olistico al problema. E’ opportuno considerare tutte le variabili in gioco e gli aspetti che impattano sull’organizzazione, sui processi e sui progetti. In particolare l’utilizzo della metodologia G.R.C: Governance, Risk e Compliance può aiutare e migliorarne la gestione.

In particolare i rischi relativi all’AI possono essere così definiti:

R1: Perdita di autonomia personale

Lo sviluppo di assistenti intelligenti può aiutare a gestire le attività amministrative e comuni di tutti i giorni, ma potrebbe comportare una diminuzione dell’intelligenza e dell’autonomia dell’utente. Tale fenomeno viene descritto dagli esperti come: demenza digitale cioè una possibile conseguenza dell’uso improprio della tecnologia digitale, che potrebbe portare al deterioramento o alla rottura delle capacità cognitive. A causa della crescente dipendenza dall’assistenza digitale, le capacità umane potrebbero andare perdute.

R2: Responsabilità per i danni causati dai sistemi di IA

I sistemi di intelligenza artificiale sono sempre di più utilizzati in sostituzione dell’attività svolta dagli esseri umani. Questo solleva il problema dell’attribuzione delle responsabilità il che solleva la questione di chi dovrebbe essere nel caso di potenziali danni causati dal loro funzionamento.

Il problema, anche dal punto di vista etico è se un essere umano è responsabile dei danni causati da un sistema di IA che ha preso una decisione autonoma e ha causato danni alla vita, alla salute o alla proprietà di un essere umano, questo aspetto crescerà come criticità aumentando la capacità di prendere sempre più decisioni autonome da parte di sistemi di intelligenza artificiale. L'aspetto giuridico non è assolutamente da sottovalutare, considerando magari il concetto di responsabilità in caso di un prodotto difettoso. E' quindi opportuno definire correttamente i ruoli e le responsabilità.

R3: Rischio di errori e manipolazione dell'IA

Utilizzare e implementare l'intelligenza artificiale significa utilizzare e implementare software e sistemi informatici. E' possibile che questi sistemi, per natura intrinseca dell'informatica, possano non funzionare correttamente o che si possano verificare diversi alcuni errori.

Le decisioni e i risultati dell'IA si basano su algoritmi che utilizzano modelli statistici che analizzano determinate quantità e qualità di dati, infatti l'uso di modelli statistici può generare decisioni e risultati errati, sia perché i dati analizzati non riflettono le singole circostanze del rispettivo scenario o campione, sia perché i dati analizzati sono parziali o errati, sia perché il modello statistico è incompleto o incorretto.

I sistemi AI possono inoltre essere manipolati e possono contenere vulnerabilità o carenze dei sistemi di sicurezza.

R4: Manipolazione, sorveglianza e comportamento illecito

Infine, il rischio per manipolazione, sorveglianza o altri

scopi illegali è elevato. Ad esempio, le tecniche biometriche e di sorveglianza potrebbero essere abusate per controllare i cittadini, le organizzazioni potrebbero utilizzare algoritmi di estrema profilazione sia dei consumatori che dei dipendenti, oppure effettuare un controllo massimo della popolazione.

R5: Perdita della privacy

Solo attraverso una accurata raccolta ed elaborazione dei dati i sistemi di AI possono prendere decisioni.

L'accesso ai dati è di fondamentale importanza per lo sviluppo delle tecnologie digitali in generale, e dell'IA.

Una delle principali preoccupazioni è quella di garantire che la privacy dei dati sia garantita, protetta e mantenuta.

Il GDPR, regolamento generale sulla protezione dei dati, introduce degli articoli che permettono una accurata tutela dei dati personali, in particolare gli articoli 25 (Privacy by Design/Default), 32 (Misure di Sicurezza) e 35 (Valutazione d'Impatto).

R6: Mancanza di trasparenza dell'IA

I sistemi di intelligenza artificiale hanno subito innumerevoli critiche a causa della loro mancanza di trasparenza.

La trasparenza è necessaria per costruire la fiducia nell'uso di un sistema di intelligenza artificiale, inoltre per comprendere cosa farà un sistema di intelligenza artificiale in diverse circostanze.

R7: Perdita degli aspetti umani nelle

relazioni sociali e mancanza di protezione della vita umana e della dignità umana

L’IA riesce a creare cambiamenti fondamentali sia alle attività che all’umanità in generale.

Una maggiore automazione decisionale e dipendenza dall’IA nella nostra vita quotidiana può portare ad una diminuzione dei contatti sociali.

Le applicazioni di intelligenza artificiale, come i robot sanitari, quelli di servizio per anziani o quelli utilizzati nel campo del turismo o addirittura i giocattoli, possono comportare un aumento dell’interazione uomo-machina importante. I nuovi sviluppi tecnologici relativi all’attuazione e all’uso dell’IA, compresa l’analisi dei Big Data, porranno questioni su cosa sia la vita umana, l’etica, la dignità e le relazioni che devono incorrere tra uomo e macchina.

R8: Restrizione della pluralità delle opinioni e della concorrenza

Un altro rischio è che le applicazioni di AI riflettono il contesto e la distorsione della fonte che le ha programmate.

L’uso crescente degli algoritmi può ridurre la pluralità delle opinioni livellando i risultati o creando/alimentando dei pregiudizi, influenzando le decisioni o distorcendo le varie opinioni.

Ad esempio strategie di comunicazione potrebbero avere un’ingerenza sulle elezioni e su altre procedure decisionali democratiche, causando in tal modo notevoli problemi.

R9: Perdita dei posti di lavoro

Il cambiamento tecnologico è stato generalmente accompagnato da cambiamenti sociali spesso con implicazioni in campo lavorativo. Questo solleva preoccupazioni ed incertezze con la paura che la AI possa sostituire il lavoro umano sia da un punto di vista dell'analisi che da un punto di vista operativo.

I rischi sopra riportati sono solo un esempio delle criticità alle quali si va incontro implementando un sistema di AI. Un'analisi dei rischi, attraverso una metodologia consolidata (definendo la probabilità, la gravità e di conseguenza il rischio con le relative soglie), dovrebbe permettere alle organizzazioni di individuare i rischi, analizzarli e definire i relativi piani d'azione per la riduzione degli stessi considerando anche l'eventuale rischio residuo.

Effettuando un'autovalutazione sui rischi elencati sopra, si potrebbe ottenere un grafico come quello della figura sottostante (ad esempio), dove la scala per la probabilità P e la gravita G è da 1 a 4, conseguentemente il rischio R va da 1 a 16.

Dal grafico emergono immediatamente le aree forti e le aree deboli dove è necessario definire e attuare al più presto dei piani d'azione.

Il risultato si potrebbe integrare con il risultato ottenuto dal tool di autovalutazione proposto nell'articolo *"Modello di autovalutazione relativo a un sistema di Intelligenza Artificiale" pubblicato su ICT Security Magazine* ottenendo un posizionamento aziendale come illustrato nel grafico seguente.

Dal grafico si può notare come la curva riporta che, come un

aumento dei valori dell’autovalutazione corrisponda una diminuzione dei rischi. Si nota inoltre un andamento quadratico della curva rappresentata.

Per una corretta governance è quindi indispensabile approcciare in modo olistico il quale contempli tutti gli aspetti del G.R.C. integrandoli nei processi e nei progetti relativi. Il supporto dell’autovalutazione permette di effettuare una radiografia dell’organizzazione implementando i relativi piani d’azione. Solo attraverso una metodologia rigorosa che analizzi tutte le problematiche e i rischi relativi all’AI possiamo fornire fiducia e sicurezza alle applicazioni e all’utilizzatore.

Articolo a cura di **Stefano Gorla**

Stefano Gorla è nato a Milano, si è laureato in fisica nucleare presso l’Università di Padova. Ha insegnato per qualche anno Matematica e Fisica presso alcuni licei.

Da anni si occupa di privacy e sicurezza dei dati attraverso un approccio sistemico e olistico.
Consulente e formatore in ambito sicurezza e tutela dei dati e delle informazioni e relatore in numerosi convegni.

Auditor ISO 27001, DPO Certificato FAC Certifica, NIST Specialist Certificato Fac Certifica, Certificato COBIT5 ISACA, Master EQFM.
Ha maturato notevoli esperienze come Referente Privacy (DPO) e consulente in tali ambiti.
È autore di varie pubblicazioni sul tema.

https://www.ictsecuritymagazine.com/articoli/ancora-sulla-governance-della-ai/