PuntoSicuro

ARTICOLO DI PUNTOSICURO

Anno 25 - numero 5462 di Venerdì 15 settembre 2023

Come mettere sotto controllo gli applicativi di intelligenza artificiale?

ETSI ha appena pubblicato tre rapporti, sviluppati dal gruppo che controlla i livelli di sicurezza dell'intelligenza artificiale, che rappresentano un prezioso strumento di valutazione e messa sotto controllo di questi applicativi.

Siamo lieti di offrire a tutti i lettori tre documenti, che rappresentano un prezioso strumento di valutazione e messa sotto controllo degli applicativi di intelligenza artificiale. Questi rapporti sono stati sviluppati da ETSI (European telecommunication standard istitute) un prestigioso istituto normativo con sede ad Antibes, in Francia, che ha fondato uno specifico gruppo di lavoro per la valutazione degli applicativi AI.

Davanti a tutti i dubbi che stanno nascendo, in ogni parte del mondo, circa le modalità di utilizzo di questi applicativi, uno attento studio di questi rapporti è di estremo interesse. Questi rapporti prendono sotto controllo i seguenti aspetti:

- la comprensibilità e la trasparenza dell'intelligenza artificiale,
- lo sviluppo di una piattaforma informatica sicura,
- un quadro di riferimento degli elementi progettuali.

Ricordiamo ai lettori che il ruolo di questo gruppo di lavoro ETSI ISG SAI è quello di sviluppare delle linee guida per la messa a punto di normative, che possono inquadrare le minacce e le vulnerabilità legate ad applicativi di intelligenza artificiale.

L'attività di questo gruppo di lavoro è mirata soprattutto ad analizzare gli aspetti sociali dell'intelligenza artificiale in modo da garantire che le norme, che a breve verranno pubblicate, possano dare adeguate garanzie sugli aspetti di sicurezza e protezione dati personali delle tecnologie informatiche, sulle cui basi tali applicativi sono sviluppati.

Passiamo ora ad analizzare in modo sintetico i tre documenti sopra menzionati.

Pubblicità
<#? QUI-PUBBLICITA-MIM-[ALDIG02] ?#>

ETSI GR SAI 007

È un rapporto afferente alla comprensibilità e trasparenza dei trattamenti di intelligenza artificiale, che identifica i passi che devono essere attuati dai progettisti e realizzatori di piattaforme di intelligenza artificiale, per dare sufficienti garanzie di comprensibilità e trasparenza del trattamento. Il trattamento di intelligenza artificiale include i processi decisionali, nonché

l'elaborazione vera e propria dei dati. Il rapporto prende in considerazione sia gli aspetti statici, sia gli aspetti dinamici. Questo approccio permette di mettere in evidenza dei possibili sbilanciamenti, presenti nell'applicativo, che potrebbero portare a valutazioni positive o negative di un soggetto, ad esempio un candidato ad un posto di lavoro, basando la valutazione su caratteristiche personali, invece che su criteri esclusivamente meritocratici. L'utilizzo delle indicazioni di questo rapporto permette non solo di mettere in evidenza possibili sbilanciamenti della valutazione, ma anche di indicare come questi bilanciamenti possono essere messi sotto controllo.

ETSI GR SAI 009

È un rapporto che offre un quadro di riferimento per progettare una piattaforma informatica di sviluppo di intelligenza artificiale, con adeguate caratteristiche di sicurezza. Appare evidente come la adozione di una piattaforma sicura rappresenti un aspetto essenziale nella progettazione di un sistema di intelligenza artificiale, per essere certi che l'hardware e software di base proteggano aspetti critici dell'applicativo, come ad esempio i modelli ed i dati utilizzati, sia in fase di calcolo, sia in condizioni di riposo. Il rapporto specifica il quadro di riferimento di sicurezza della piattaforma di calcolo, che viene così protetta da possibili attacchi e offre idonee garanzie di sicurezza a tutti i soggetti, che sono coinvolti nello sviluppo nell'utilizzo di sistemi di intelligenza artificiale.

ETSI GR SAI 013

Questo rapporto offre un quadro di riferimento come strumento per dimostrare l'applicabilità di idee e tecnologie, utilizzate nello sviluppo di applicativi di intelligenza artificiale. I risultati di questo studio permettono al gruppo di lavoro specializzato di elevare il livello di visibilità dei problemi e introdurre dei concetti di sensibilizzazione sui problemi di sicurezza dell'intelligenza artificiale. Il quadro di riferimento è stato progettato per includere il maggior numero possibile di soluzioni applicative, comprese quelle che svolgono funzioni critiche, collegate all'analisi di dati, alla gestione delle infrastrutture ed alla sicurezza informatica. Infatti è noto che, almeno in teoria, un sistema basato su intelligenza artificiale può diventare il bersaglio di un attacco informatico e la tempestiva individuazione di questi tipi di attacchi può rappresentare una sfida difficile da affrontare. Una buona conoscenza degli aspetti pratici di sicurezza informatica, da un lato analizzando lo sviluppo di un attacco contro sistemi basati su intelligenza artificiale, e dall'altro attuando tecniche di protezione e mitigazione di queste minacce, rappresenta un aspetto fondamentale per garantire la sicurezza degli applicativi.

ETSI GR SAI 007 (pdf)

ETSI GR SAI 013 (pdf)

ETSI GR SAI 013 (pdf)

Adalberto Biasiotti

Licenza Creative Commons

www.puntosicuro.it