AI, Data Protection & Cybersecurity

Avv. Giuseppe Serafini Law Firm - 06012 - Città Di Castello - (PG) - Italy - Via S. Antonio Nr. 7

Avv. Giuseppe Serafini

giuseppe.serafini@ordineavvocati.perugia.it

Donna, che stando a sedere tenga una torcia accesa, ed appresso avrà un libro aperto, che col dito indice della destra mano l'accenni.
La Torcia accesa significa, che come a' nostri occhi corporali, fa bisogno della luce per vedere, così all'occhio nostro interno, che è l'intelletto, per ricevere la cognizione delle specie intelligibili, fa mestiero dell'istrumento eltrinseco de' sensi, e particolarmente di quello del vedere, che dimostrasi col lume della torcia, perciocchè, come dice Aristotele, *Nihil est in intellectu, quod prius non fuerit in sensu*; ciò mostrando ancora il libro aperto, perchè o per vederlo, o per udirlo leggere, si fa in noi la cognizione delle cose.

Avvocato Cassazionista del Foro di Perugia.

Master II - LUISS in Cybersecurity Politiche pubbliche, normative e gestione.

Master II - Roma Tre Data Protection Officer & Privacy Expert.

Perf. UNIMI - Cyber Crime, Cloud, Data Protection, Digital Forensics.

DPO - Alta Form. CNF - CNI.

ISO / IEC 27001:2022
27701:2019 - Lead Auditor.

DPO UNI 11697:2017.

ECCE - European Certificate on Cybercrime Evidence.

Digital Forensics Expert Witness.
Docente Master Cybersecurity Uni-PG / Unilink.

CSIG - Perugia.
D.F.A. - Digital Forensics Alumni.
CLUSIT - Associazione Italiana per la Sicurezza Informatica.

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

II. - Artificial Intelligence.

- Basics
- Transformers
- Prompting

III. - DPIA

- Risk Management
- Fundamental Rights Impact Assessment
- Neural data privacy

IV. - CyberSecurity

- Data Poisoning
- Bias
- Adversarial Attack
- Deep Fake

Artificial Intelligence is computer systems that exhibit human like intelligence. It is a group of science fields and technologies concerned with creating machines take intelligent actions based on inputs.

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

Eu Cybersecurity Act
Art. 2.1. - Cybersicurezza.

L'insieme delle attività necessarie per proteggere la rete e i sistemi informativi, gli utenti di tali sistemi e altre persone interessate dalle minacce informatiche.

La complessità genera incertezza, l'incertezza genera rischio.

Grazie alla miniaturizzazione dei dispositivi, alla loro aumentata capacità di elaborazione, allo sviluppo di linguaggi di programmazione sempre più friendly, alla semplicità di impiego di GUI (Graphical User Interface) e NUI (Natural User Interface) sempre più evolute, ed alla sempre maggiore disponibilità di interconnessione (5G / IoT), ed all'impiego di strumenti di AI, sono stati digitalizzati, sia molti beni e servizi, sia molti dei rapporti (blockchain) in grado di crearli e di costituire prova della loro esistenza legale,

Il patrimonio immateriale digitalizzato (corpo elettronico) costituisce oggi, nella maggior parte dei casi, nelle varie forme, asset primario degli individui, delle organizzazioni e delle Istituzioni e degli Stati, tutelato da disposizioni normative nazionali ed internazionali.

Rischio: effetto dell'incertezza sugli obiettivi.

Quadro strategico nazionale per la sicurezza dello spazio cibernetico.

Lo spazio cibernetico è l'insieme delle infrastrutture informatiche interconnesse, comprensivo di hardware, software, dati ed utenti, nonché delle relazioni logiche, comunque stabilite, tra di essi.

Esso dunque comprende internet, le reti di comunicazione, i sistemi su cui poggiano i processi informatici di elaborazione dati e le apparecchiature mobili dotate di connessioni di rete.

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

Carta dei diritti fondamentali dell'Unione Europea (Roma 2000)

Articolo 7 - Rispetto della vita privata e della vita familiare.

Ogni individuo ha diritto al rispetto della propria vita privata e familiare, del proprio domicilio e delle sue comunicazioni.

Articolo 8 - Protezione dei dati di carattere personale

1. Ogni individuo ha diritto alla protezione dei dati di carattere personale che lo riguardano.

2. Tali dati devono essere trattati secondo il principio di lealtà, per finalità determinate e in base al consenso della persona interessata o a un altro fondamento legittimo previsto dalla legge. Ogni individuo ha il diritto di accedere ai dati raccolti che lo riguardano e di ottenerne la rettifica.

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

Reg. (UE) 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (Regolamento Generale sulla Protezione dei Dati).

D.Lgs. 101/2018 - Disposizioni per l'adeguamento della normativa nazionale alle disposizioni del regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personal

Principi Fondamentali

I. Liceità, correttezza e trasparenza
II. Limitazione della finalità
III. Minimizzazione dei dati
IV. Esattezza e aggiornamento
V. Limitazione della conservazione:
VI. Integrità e riservatezza

I. Accountability (Art. 24)
II. Data protection by design & by default (Art. 25)
III. Approccio basato sul rischio

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

I. Base giuridica
II. Valutazione d'Impatto
III. Informazioni obbligatorie
IV. Data Processing Agreement
V. Trasferimenti Internazionali
VI. Consenso
VII. Misure di sicurezza (Cfr. IV)

Articolo 24 - Responsabilità del titolare del trattamento.

1. Tenuto conto della natura, dell'ambito di applicazione, del contesto e delle finalità del trattamento, nonché dei rischi aventi probabilità e gravità diverse per i diritti e le libertà delle persone fisiche, il titolare del trattamento mette in atto misure tecniche e organizzative adeguate per garantire, ed essere in grado di dimostrare, che il trattamento è effettuato conformemente al presente regolamento. Dette misure sono riesaminate e aggiornate qualora necessario.

2. Se ciò è proporzionato rispetto alle attività di trattamento, le misure di cui al paragrafo 1 includono l'attuazione di politiche adeguate in materia di protezione dei dati da parte del titolare del trattamento.

Trattamento: Qualsiasi operazione o insieme di operazioni, compiute con o senza l'ausilio di processi automatizzati e applicate a dati personali o insiemi di dati personali, come la raccolta, la registrazione, l'organizzazione, la strutturazione, la conservazione, l'adattamento o la modifica, l'estrazione, la consultazione, l'uso, la comunicazione mediante trasmissione, diffusione o qualsiasi altra forma di messa a disposizione, il raffronto o l'interconnessione, la limitazione, la cancellazione o la distruzione;

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

Prompt: *"Icona di donna in stile vittoriano ispirata alla Privacy"*

Dato personale

Qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un id online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale;

Esempi

- E-mail (Meta Dati)
- Identificativi
- Carte di credito
- Numero IP
- User ID, Mac Address;
- Numeri telefonici

Dati particolari

L'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona.

Dati relativi alla salute

I dati personali attinenti alla salute fisica o mentale di una persona fisica, compresa la prestazione di servizi di assistenza sanitaria, che rivelano informazioni relative al suo stato di salute;

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

Titolare del trattamento - La persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali

Responsabile del trattamento - La persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che tratta dati personali per conto del titolare del trattamento;

Contitolari - Determinano congiuntamente le finalità e i mezzi del trattamento,

Autorizzati - Persone fisiche che operano sotto la diretta autorità del Titolare o del Responsabile del trattamento.

DPO - Persona fisica incaricato di vigilare sulla conformità alle normative in materia di protezione dei dati personali all'interno di un'organizzazione.

Gruppo imprenditoriale - un gruppo costituito da un'impresa controllante e dalle imprese da questa controllate;

Autorità di controllo - l'autorità pubblica indipendente istituita da uno Stato membro ai sensi dell'articolo 51;

Destinatario - la persona fisica o giuridica, l'autorità pubblica, il servizio o un altro organismo che riceve comunicazione di dati personali, che si tratti o meno di terzi.

European Data protection Board.

Opinion 7/2020 - Sui concetti di titolare del trattamento e di responsabile del trattamento ai sensi del GDPR (Pag. 25)

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti

(26). - I principi di protezione dei dati non dovrebbero applicarsi a informazioni anonime, vale a dire informazioni che non si riferiscono a una persona fisica identificata o identificabile o a dati personali resi sufficientemente anonimi da impedire o da non consentire più l'identificazione dell'interessato. Il presente regolamento non si applica pertanto al trattamento di tali informazioni anonime, anche per finalità statistiche o di ricerca.

Pseudonimizzazione - Il trattamento dei dati personali in modo tale che i dati personali non possano più essere attribuiti a un interessato specifico senza l'utilizzo di informazioni aggiuntive, a condizione che tali informazioni aggiuntive siano conservate separatamente e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile;

Profilazione - qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica;

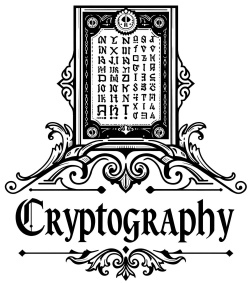
EDPB - Guidelines on **Pseudonymisation** Adopted on 16 January 2025

WP 251 Guidelines on Automated individual decision-making and Profiling

WP29 216 - Opinion on **Anonymisation** Techniques - 10 April 2014

I. - Laws & Standards

- Cyberspace
- Privacy Vs Data Protection
- GDPR
- Definizioni
- Altre norme rilevanti



- Cybersecurity Act
- AI Act
- Cyber Resilience Act

- Digital Operational Resilience Act
- Digital Service Act
- NIS 2 Directive

- Commission Delegated Regulation (EU) 2023/66 of 21 October 2022 amending Regulation (EU) 2021/821 of the European Parliament and of the Council as regards the list of dual-use items.

ISO/IEC TR 29119-11:2020 - Software and systems engineering - Software testing - Part 11: Guidelines on the testing of AI-based systems.

ISO/IEC 27001:2022 - Information security, cybersecurity and privacy protection - Information security management systems - Requirements;

ISO/IEC 27701:2019 - Security techniques - Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management - Requirements and guidelines

ISO/IEC 42001:2023(en) Information technology - Artificial intelligence - Management system.

II. - Artificial Intelligence

- Basics
- Transformers
- Prompting

NEURAL NETWORKS

Computing systems that organises the computing elements in a layered way that is loosely modelled on the human brain. Enables Deep Learning.

Sistema di IA.

un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali; (Infers / déduit / infiere / infere)

Deployer: una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità (....)

Rischio: *la combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso.*

Processi logici per trarre conclusioni a partire da informazioni o premesse.

Deduzione: parte da premesse generali e universali per arrivare a una conclusione specifica. La deduzione è rigorosa e segue regole formali di logica.

Inferenza: termine più ampio che indica il processo mentale di trarre conclusioni a partire da informazioni disponibili, può includere deduzione, induzione e abduzione.

Non garantisce necessariamente la verità della conclusione; può essere probabilistica o plausibile, basandosi su osservazioni incomplete o indizi.

II. - Artificial Intelligence

- Basics
- Transformers
- Prompting

- Convolutional Neural Networks (CNN: Ottimizzate per dati visivi, come immagini e video).
- Recurrent Neural Networks (RNN): Utilizzate per dati sequenziali come testo e serie temporali.

ASI - AGI - ANI - GAN - NLP - LLM

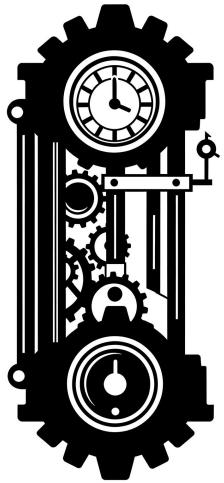
I. 1950 - A. Turing - "Computer Machinery and Intelligence"
II. 1956 - John McCarthy (Dartmouth) Artificial Intelligence
III. Anni 60/70
IV. Anni 70/80 - Winter
V. Anni 90/2000

Potenza di calcolo + Big Data

Machine Learning

Neural Network
RNN - CNN
Deep Learning

Alberi decisionali
(If Then)
Modelli Bayesiani



Turing Test

II. - Artificial Intelligence

- Basics
- Transformers
- Prompting

- GAN (Generative Adversarial Networks): Due reti neurali che competono tra loro per generare dati realistici (es. immagini).
- Transformers: Modelli avanzati come BERT e GPT per il trattamento del linguaggio naturale (NLP).

2017 "Attention is all you need"

Transformers

General Pre-trained Transformer

MACHINE LEARNING

Algorithms that can learn from and make predictions on data. Overlaps with Computational Statistics. Overlaps with Bayesian Statistics. Underpins Predictive Analytics. Underpins Data-Mining.

DEEP LEARNING

A high powered type of Machine learning algorithms that uses a cascade of many computing layers. Each layer uses the input from the previous later as input.

PATTERN RECOGNITION

A branch of Machine Learning and Deep Learning which focusses on recognition of patterns in data.

II. - Artificial Intelligence

- Basics
- Transformers
- Prompting

Prompt (Role-Play, con scenario):

Sei un consulente legale specializzato in cybersecurity.
Un'azienda tua cliente ha subito un attacco ransomware che ha causato la crittografia dei dati personali dei propri clienti.
L'azienda non ha ancora notificato la violazione all'autorità garante né agli interessati.
Quali sono i tuoi consigli immediati all'azienda?
Quali sono i rischi legali e le possibili sanzioni?
Quali azioni deve intraprendere l'azienda per mitigare i danni e ripristinare la conformità?

I. Tecnica di interazione con LLM che consiste nella formulazione di un input testuale, detto prompt, che specifica il compito desiderato, il contesto rilevante, il formato di output e, opzionalmente, esempi di input-output.

II. Funge da istruzione per il modello, guidandolo nella generazione di una risposta coerente e pertinente.

III. L'efficacia dipende dalla chiarezza, specificità e completezza del prompt, nonché dalla capacità del modello di comprendere e seguire le istruzioni implicite ed esplicite contenute nel prompt stesso.

IV. Processo iterativo, in cui il prompt viene raffinato in base alle risposte generate dal modello, al fine di ottimizzare la qualità e la rilevanza dell'output.

Prompt (Instruction Prompting, con checklist):

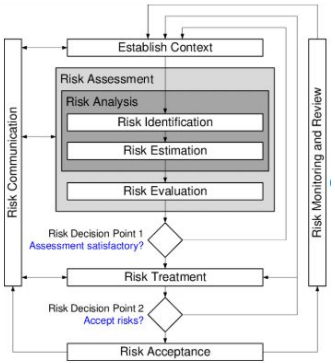
Crea una check list dettagliata delle azioni che un'azienda deve intraprendere in caso di data breach, in conformità al GDPR.

La checklist deve includere:

I. Azioni immediate di contenimento e mitigazione
II. Valutazione della gravità della violazione
III. Notifica all'autorità garante (se necessaria)
IV. Notifica agli interessati (se necessaria)
V. Documentazione dell'incidente
VI. Misure correttive e preventive

III. - DPIA

- Risk Management
- Fundamental Rights Impact Assessment
- Neural data Privacy



RISK MANAGEMENT

Rischio = Probabilità x Impatto

Il processo di gestione del rischio secondo ISO 27005 è un ciclo continuo che comprende diverse fasi.

La prima fase è la **definizione del contesto**, che stabilisce i criteri per l'identificazione dei rischi, l'attribuzione della responsabilità e il calcolo dell'effetto e della probabilità del rischio

Segue l'**identificazione dei rischi,** che comporta il riconoscimento di potenziali fonti di danno per gli asset informativi e l'individuazione di vulnerabilità e minacce

L'**analisi del rischio** consiste nel valutare la probabilità e l'impatto dei rischi identificati

Mentre la **valutazione del rischio** comporta il confronto dei risultati dell'analisi con i criteri di accettabilità del rischio definiti dall'organizzazione.

Il trattamento del rischio include la selezione di opzioni per affrontare i rischi inaccettabili, come la mitigazione, l'evitamento, il trasferimento o l'accettazione

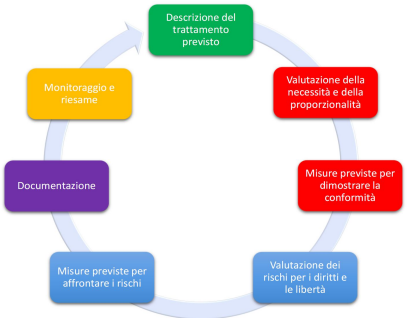
L'accettazione del rischio si verifica quando il livello di rischio rientra nei criteri definiti dall'organizzazione

La comunicazione e la consultazione sul rischio sono essenziali per garantire che le decisioni e le azioni intraprese siano comprese e accettate dalle parti interessate

Infine, il monitoraggio e la revisione del rischio sono cruciali per garantire che il processo di gestione del rischio rimanga efficace nel tempo e si adatti ai cambiamenti del contesto

III. - DPIA

- Risk Management
- Fundamental Rights Impact Assessment
- Neural data Privacy



Art. 35 - Valutazione d'impatto sulla protezione dei dati

1. Quando un tipo di trattamento, allorché prevede in particolare l'uso di nuove tecnologie, considerati la natura, l'oggetto, il contesto e le finalità del trattamento, può presentare un rischio elevato per i diritti e le libertà delle persone fisiche, il Titolare effettua, prima di procedere al trattamento, una valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali. Una singola valutazione può esaminare un insieme di trattamenti simili che presentano rischi elevati analoghi.

2. Il Titolare del, allorquando svolge una valutazione d'impatto sulla protezione dei dati, si consulta con il DPO, qualora ne sia designato uno.

3. La valutazione d'impatto sulla protezione dei dati di cui al paragrafo 1 è richiesta in particolare nei casi seguenti:

b) il trattamento, su larga scala, di categorie particolari di dati personali di cui all'articolo 9, paragrafo 1, o di dati relativi a condanne penali e a reati di cui all'articolo 10; o

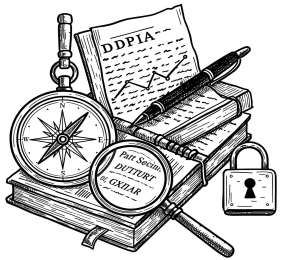
- Minaccia
- Vulnerabilità
- Impatto
- Probabilità

- Discriminazione
- Esclusione
- Perdita di opportunità

.

III. - DPIA

- Risk Management
- Fundamental Rights Impact Assessment
- Neural Data Privacy



DPIA

7. La valutazione contiene almeno:

a) una descrizione sistematica dei trattamenti previsti e delle finalità del trattamento, compreso, ove applicabile, l'interesse legittimo perseguito dal titolare del trattamento;

b) una valutazione della necessità e proporzionalità dei trattamenti in relazione alle finalità;

c) una valutazione dei rischi per i diritti e le libertà degli interessati di cui al paragrafo 1; e

d) le misure previste per affrontare i rischi, includendo le garanzie, le misure di sicurezza e i meccanismi per garantire la protezione dei dati personali e dimostrare la conformità al presente regolamento, tenuto conto dei diritti e degli interessi legittimi degli interessati e delle altre persone in questione.

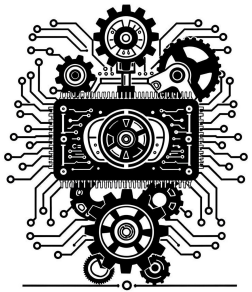
9. Se del caso, il titolare raccoglie le opinioni degli interessati o dei loro rappresentanti sul trattamento previsto, (...)

11. Se necessario, il Titolare procede a un riesame per valutare se il trattamento sia effettuato conformemente alla valutazione d'impatto almeno quando insorgono variazioni del rischio rappresentato dalle attività relative al trattamento.

enisa
EUROPEAN UNION AGENCY FOR CYBERSECURITY

III. - DPIA

- Risk Management
- Fundamental Rights Impact Assessment
- Neural Data Privacy



Articolo 27 - Valutazione d'impatto sui diritti fondamentali per i sistemi di IA ad alto rischio

1. Prima di utilizzare un sistema di IA ad alto rischio di cui all'articolo 6, paragrafo 2, ad eccezione (...)i deployer che sono organismi di diritto pubblico o sono enti privati che forniscono servizi pubblici e i deployer di sistemi di IA ad alto rischio di cui all'allegato III, punto 5, lettere b) e c), effettuano una valutazione dell'impatto sui diritti fondamentali che l'uso di tale sistema può produrre. A tal fine, i deployer effettuano una valutazione che comprende gli elementi seguenti:

a) una descrizione dei processi del deployer in cui il sistema di IA ad alto rischio sarà utilizzato in linea con la sua finalità prevista;

b) una descrizione del periodo di tempo entro il quale ciascun sistema di IA ad alto rischio è destinato a essere utilizzato e con che frequenza;

c) le categorie di persone fisiche e gruppi verosimilmente interessati dal suo uso nel contesto specifico;

d) i rischi specifici di danno che possono incidere sulle categorie di persone fisiche o sui gruppi di persone individuati a norma della lettera c), del presente paragrafo tenendo conto delle informazioni trasmesse dal fornitore a norma dell'articolo 13;

e) una descrizione dell'attuazione delle misure di sorveglianza umana, secondo le istruzioni per l'uso;

f) le misure da adottare qualora tali rischi si concretizzino, comprese le disposizioni relative alla governance interna e ai meccanismi di reclamo.

.

III. - DPIA

- Risk Management
- Fundamental Rights Impact Assessment
- Neural Data Privacy

Dark Patterns

- Guidelines 3/2022 on Dark patterns in social media platform interfaces: How to recognise and avoid them.
- Interfaces and user experiences implemented on social media platforms that lead users into making unintended, unwilling and potentially harmful decisions regarding the processing of their personal data.
- Dark patterns aim to influence users' behaviour and can hinder their ability to effectively protect their personal data and make conscious choices. Data protection authorities are responsible for sanctioning the use of dark patterns if these breach GDPR requirements.

Neural Data

Qualsiasi informazione derivata dalla misurazione dell'attività del sistema nervoso, in particolare del cervello.

Fonti:

- Elettroencefalografia (EEG),
- Magnetoencefalografia (MEG),
- Risonanza magnetica funzionale (fMRI),
- Tomografia ad emissione di positroni (PET)
- Interfacce cervello-computer (BCI).

La natura di questi dati è intrinsecamente legata alla sfera più intima e personale dell'individuo.

IV. - Cybersecurity

- Basics
- Data Poisoning
- Bias
- Adversarial Attack

Funzione di perdita (o loss function).

Funzione matematica, utilizzata per l'addestramento e la classificazione dei dati nei sistemi di Machine Learning che quantifica quanto le predizioni di un modello sono sbagliate rispetto ai valori reali (o "verità"). In altre parole, misura l'errore del modello.

- **Rilevamento delle minacce:** analisi di dati per identificare modelli e anomalie che indicano attività dannose.
- **Classificazione delle minacce:** in base alla gravità e al tipo, consentendo una risposta più rapida ed efficace.
- **Previsione delle minacce:** prevedere potenziali attacchi futuri analizzando i dati storici e identificando tendenze emergenti.
- **Analisi comportamentale degli utenti:** rilevare comportamenti insoliti che potrebbero segnalare un account compromesso o un insider threat.
- **Sandboxing automatizzato:** creazione di ambienti sicuri e isolati per analizzare in modo sicuro il comportamento di potenziali minacce.

I. Furto d'identità (phishing, social engineering)
II. Malware (virus, ransomware, spyware)
III. Attacchi DDoS (interruzione di servizi)
IV. Violazione di dati (furto di informazioni sensibili)
V. SQL injection (manipolazione di database)
VI. Cross-Site Scripting (XSS)
VII. Man-in-the-middle attack

IV. - Cybersecurity

- Basics
- Data Poisoning
- Bias
- Adversarial Attack

Sistema di Sicurezza	Obiettivo dell'Attacco di Data Poisoning	Metodo di Iniezione dei Dati Dannosi	Impatto sulla Sicurezza
Rilevamento Intrusioni	Indurre falsi negativi	Aggiunta di traffico malevolo etichettato come benigno	Il traffico dannoso non viene rilevato, consentendo l'intrusione nel sistema.
Filtro Antispam	Indurre falsi negativi	Aggiunta di email di spam etichettate come non spam	Le email di spam raggiungono la casella di posta dell'utente.
Riconoscimento Biometrico (Facciale)	Compromissione dell'accuratezza	Aggiunta di immagini manipolate per associare volti a identità errate	Possibile accesso non autorizzato a sistemi protetti da riconoscimento facciale.
Rilevamento Malware	Indurre falsi negativi	Modifica di firme di malware per farle apparire come software benigno	Il malware non viene rilevato e può infettare il sistema.
Analisi del Comportamento Utente	Indurre falsi negativi	Iniezione di dati che mascherano attività malevole come normali	Attività sospette non vengono identificate, consentendo agli attaccanti di operare indisturbati.

IV. - Cybersecurity

- Basics
- Data Poisoning
- Bias
- Adversarial Attack

AI BIAS

Sistema di Sicurezza	Fonte del Bias	Come il Bias Può Essere Sfruttato	Potenziali Conseguenze
Riconoscimento Facciale	Dati di training non rappresentativi per etnia	Minore accuratezza nel riconoscimento di volti di determinate etnie, consentendo a individui di questi gruppi di impersonare altri.	Accesso non autorizzato, falsi positivi o negativi con impatto sproporzionato su specifici gruppi.
Rilevamento Intrusioni	Sotto-rappresentazione di specifici tipi di attacchi	Il modello potrebbe non rilevare attacchi nuovi o che utilizzano tecniche non presenti nei dati di training.	Mancata rilevazione di intrusioni, compromissione della sicurezza del sistema.
Analisi del Comportamento Utente	Pregiudizi impliciti nei dati relativi a specifici ruoli	Attività malevole da parte di utenti con profili diversi da quelli predominanti nei dati di training potrebbero non essere rilevate.	Gli attaccanti potrebbero mascherare le proprie azioni comportandosi in modo simile a utenti meno rappresentati nei dati di training.
Filtro Antispam	Dati di training prevalentemente in una sola lingua	Potrebbe essere meno efficace nel filtrare spam in altre lingue.	Gli utenti potrebbero ricevere un elevato volume di spam in lingue diverse da quella predominante nei dati di training.

IV. - Cybersecurity

- Basics
- Data Poisoning
- Bias
- Adversarial Attack

Adversarial Attack

Sistema di Sicurezza Attaccato	Tipo di Attacco Adversarial	Modifica Adversarial	Risultato dell'Attacco
Classificatore di Malware	Fast Gradient Sign Method	Aggiunta di un piccolo rumore calcolato al file eseguibile.	Il malware viene classificato come benigno.
Sistema di Rilevamento Intrusioni	Projected Gradient Descent	Modifica di pacchetti di rete in modo impercettibile.	Il traffico di rete malevolo non viene rilevato.
Riconoscimento di Immagini	DeepFool	Alterazione dei pixel di un'immagine in modo da spostarla oltre il confine di decisione.	Un'immagine di un segnale di stop viene classificata come segnale di limite di velocità.

GRAZIE PER L'ATTENZIONE

Avv. Giuseppe Serafini Law Firm - 06012 - Città Di Castello - (PG) - Italy - Via S. Antonio Nr. 7