Politecnico di Milano

Dipartimento di Architettura e Studi Urbani

MASTER UNIVERSITARIO DI II LIVELLO
"DATA PROTECTION OFFICER E TRANSIZIONE DIGITALE (DPOTD)"
A.A. 2024-2025

AIoT ed Edge Computing nelle Smart Cities: opportunità e rischi dell'innovazione nella protezione dei dati

Relatore
Dott. Ing. Sergio BAREZZANI
Correlatore
Avv. Andrea REGHELIN

Tesi Master
Dott. Andrea Ravagnani

Sommario

Introduzione

Negli ultimi anni, la costante accelerazione dello sviluppo tecnologico ha profondamente trasformato il modo in cui le città affrontano le sfide legate alla crescita demografica, alla sostenibilità ambientale ed alla gestione degli spazi pubblici, risorse e sicurezza.

Il concetto di Smart City è emerso quale paradigma innovativo per il miglioramento della qualità di vita urbana attraverso una gestione basata interamente su tecnologie che vedono l'integrazione di soluzioni quali modelli di Intelligenza Artificiale (di seguito "AI"), Internet of Things (di seguito "IoT") ed Edge Computing, al fine di promuovere un'amministrazione dei sistemi urbani svolta in maniera efficiente, interconnessa ed intelligente.

Tuttavia, sebbene queste soluzioni offrano numerosi vantaggi, esse comportano anche dei rischi, non solo con riferimento alla sicurezza informatica dei sistemi utilizzati, ma anche al rispetto dei principi di Data Protection con riferimento ai dati personali (siano essi di natura comune o particolare) relativi agli interessati coinvolti nei molteplici trattamenti effettuati mediante l'implementazione di queste soluzioni.

Il presente elaborato vuole porre l'attenzione sui vantaggi e svantaggi legati all'impiego di sistemi integrati di AI e IoT (di seguito "AIoT") nelle Smart Cities attraverso un percorso che permetta di comprendere le principali sfide poste da questi sistemi con riguardo sia all'ambito legale che a quello della sicurezza informatica.

Dopo un'iniziale introduzione ai concetti di AI, IoT ed Edge Computing, esamineremo:

- la principale normativa di riferimento per la protezione dei dati raccolti mediante tali sistemi, con particolare riguardo per il GDPR e l'AI Act;
- le principali vulnerabilità e attacchi informatici che possono essere condotti verso queste soluzioni e i potenziali impatti per le persone.
- il caso pratico affrontato durante il mio tirocinio curriculare presso l'azienda Partners4Innovation (P4I), fulcro della presente tesi, il quale offrirà l'opportunità di analizzare non solo i profili strettamente normativi, ma anche quelli inerenti alla Cyber Security;
- il caso del City Brain project cinese.

Un ulteriore messaggio che questa tesi intende trasmettere è la crescente necessità di un approccio interdisciplinare da parte del Data Protection Officer (DPO), il quale deve sempre più acquisire competenze anche in ambito informatico, al fine di comprendere a fondo le complessità e i rischi connessi ai principali strumenti tecnologici impiegati nel trattamento dei dati personali.

Infine, l'elaborato si concluderà offrendo uno spunto riflessivo in merito alle opportunità e agli altrettanto reali pericoli legati alle Smart Cities, prendendo in esame il caso unico al mondo del sistema City Brain sviluppato dalla Big Tech Alibaba Cloud ed impiegato in diverse città cinesi per l'amministrazione delle aree urbane, andando ad offrire una riflessione sul bilanciamento tra Data Protection ed innovazione.

I. AIoT e Smart City: contesto

Per una maggiore chiarezza espositiva, questo capitolo fornirà una contestualizzazione dei concetti di "Smart Cities" e "AIoT" (oltre che dell'Edge Computing), esaminando alcuni esempi pratici.

In considerazione degli obiettivi dell'elaborato e dell'ampiezza degli argomenti trattati, questo si limiterà ad offrire una panoramica generale di tali concetti, ponendoli come base introduttiva per comprendere più chiaramente i rischi per gli interessati in ambito Data Protection e Cyber Security, che emergono dall'applicazione concreta di queste tecnologie innovative.

I.1. Il concetto di "Smart City"

Con il concetto di Smart City si delinea quella che è una vera e propria città "intelligente", nella quale, mediante l'integrazione di tecnologie digitali a servizio delle reti, dei servizi e delle infrastrutture in essa presenti, permette di garantire che la vita all'interno dell'area urbana sia non solo più efficiente, ma anche innovativa, sostenibile e ideata su misura per l'essere umano.

Per definire una città "intelligente" devono essere presenti alcuni elementi chiave che ne caratterizzino la sua conformazione e gestione.

Gli edifici, ad esempio, diventano progressivamente digitalizzati ed efficienti dal punto di vista energetico, riducendo l'inquinamento prodotto e abbattendo gli sprechi di risorse, e ottenendo un certo grado di automazione dei processi, grazie all'impiego dell'intelligenza artificiale a supporto della vita quotidiana delle persone.

Le infrastrutture statali e le imprese situate nella smart city riducono i propri costi e l'impatto ambientale derivante dalle proprie attività implementando soluzioni di efficienza energetica.

Anche la mobilità viene impattata, diventando integrata e condivisa per il miglioramento del trasporto sia pubblico che privato mediante vetture, autobus elettrici e mezzi di micromobilità, alimentati interamente da fonti di energia rinnovabile.

Inoltre, grande enfasi viene posta anche sulle soluzioni di sicurezza smart a sostegno dei cittadini, come le tecnologie di monitoraggio dell'ambiente urbano, di rilevamento di irregolarità e di prevenzione di eventi dannosi, le quali garantiscono un livello di sicurezza elevato tramite una rete di avvisi automatici che possono attivare le forze dell'ordine nel minor tempo possibile.

Per consentire il perfetto funzionamento di una Smart City è necessario che alla sua base ci sia un sistema integrato e interconnesso di tecnologie abilitanti che supportino e facilitino la trasformazione della città: si può parlare di un articolato framework ICT, cioè di una rete

intelligente di macchine e infrastrutture che dialogano tra di loro elaborando i dati raccolti e lavorando per restituire la migliore esperienza possibile agli utenti finali.[1]

Le tecnologie che danno vita ad una Smart City sono molteplici e ognuna di esse deve essere in grado di interagire efficientemente con le altre per garantire un elevato grado di interoperabilità. Un ruolo cruciale è sicuramente svolto dall'insieme di tecnologie che costituisce il mondo dall'IoT.

I.2. L'Internet of Things (IoT): descrizione ed esempi pratici

L'IoT, o "Internet of Things", si basa sull'impiego e interconnessione di dispositivi intelligenti (detti anche "Smart Objects") i quali non sono solo computer, smartphone o tablet, ma anche e soprattutto oggetti quotidiani dai quali siamo circondati nella vita di tutti i giorni.[2]

Queste tecnologie si sono moltiplicate e sviluppate con grande rapidità, portando al contempo ad una loro profonda evoluzione in numerosi ambiti applicativi che vanno ad impattare profondamente la vita nelle nostre città e ogni aspetto economico e amministrativo delle stesse.

Alcuni esempi di ambiti applicativi che hanno visto un crescente uso delle tecnologie IoT impattanti le Smart Cities sono:

- le Smart Car;
- le Smart Homes;
- gli Smart Buildings;
- la Smart Agricolture;
- lo Smart Metering (insieme di contatori interconnessi);
- l'Industrial IoT, nel quale rientrano la Smart Factory e la Smart Logistics.

I dispositivi IoT sono numerosi e spaziano per tipologia a seconda degli ambiti all'interno dei quali vengono impiegati, come la già citata Smart Home o la Smart Factory.

Per comprendere il livello di granularità che tali dispositivi possono raggiungere, pensiamo ai lampioni delle nostre città: l'applicazione dell'IoT a questi comuni strumenti di illuminazione urbana è in grado permettere un'autoregolazione degli stessi sulla base delle condizioni di visibilità durante l'arco della giornata.

Anche i semafori, mediante sensori IoT, possono sincronizzarsi per creare un'onda verde volta a facilitare, ad esempio, il passaggio di un mezzo di soccorso in servizio.

---

[1] V. F. Bloise, 2024, "Guida alle smart city, cosa sono e come funzionano le città "intelligenti"", *E-Motion Mag*, https://platum.com/e-motion-mag/smart-city/smart-city-cosa-e-come-funziona-citta-intelligente/.

[2] V. Osservatorio Internet of Things, 2019, "Internet of Things (IoT): significato, esempi e applicazioni", *Internet of Things*, https://blog.osservatori.net/it_it/cos-e-internet-of-things.

Consideriamo inoltre le telecamere di sorveglianza, che grazie all'IoT sono in grado di inviare segnali di allarme alle centrali di pronto intervento in caso di rilevamento di infrazioni.
Infine, un altro esempio è costituito dall'insieme di sensori che permettono il monitoraggio dei parametri microclimatici nelle coltivazioni agricole, garantendo un uso ottimizzato delle risorse e una riduzione dell'impatto ambientale derivante da tali attività.

I.3. Intelligenza Artificiale ed Internet of Things (AIoT): una sinergia innovativa

Grazie all'integrazione con un'altra tecnologia estremamente innovativa, i dispositivi IoT possono ulteriormente eccellere nello svolgimento dei loro compiti: l'Intelligenza Artificiale.
Infatti, sempre più frequentemente nelle soluzioni Internet of Things aumentano gli impieghi dell'Intelligenza Artificiale in rapporto sinergico: l'AI generativa, i Large Language Model (LLM), il Machine Learning e le altre tecniche di apprendimento alla base dei modelli di Intelligenza Artificiale svolgono e svolgeranno un ruolo sempre più cruciale all'interno della nostra società.
Le applicazioni dei sistemi AIoT sono sconfinate e avranno un impatto radicale sia sulle aziende che sulle pubbliche amministrazioni ed i consumatori: basti pensare, ad esempio, all'ipotesi dei camerini dei negozi dotati di display trasparenti e touch (nel settore che diventa così quello dello Smart Retail), in grado di fornire non solo in tempo reale tutte le informazioni alle richieste poste dall'utente finale, ma anche di poter comprendere nel corso del tempo le preferenze del consumatore generando una profilazione dello stesso che permetta, con precisione, di mostrare i prodotti che possono suscitare in lui maggior interesse.[3]
Nel caso delle Smart Cities, l'AIoT viene particolarmente utilizzata nell'ambito della pubblica sicurezza.
Ad esempio, i software di video analytics all'interno delle telecamere intelligenti sono in grado di rielaborare le immagini e riconoscere situazioni di pericolo mediante Intelligenza Artificiale e Deep Learning (tramite quella che viene definita Computer Vision).
Tali sistemi di Video Analytics, tuttavia, permettono anche di ottenere una maggiore efficienza nei processi gestionali di aree pubbliche, permettendo anche solo di comprendere con un maggiore livello di dettaglio le problematiche legate alla gestione degli spazi ad elevato tasso di affluenza, al fine di garantire una fruibilità più fluida di tali servizi per gli utenti finali a seguito dell'adozione di strategie mirate elaborate sulla base dei dati raccolti dal sistema (come vedremo nel caso pratico riportato al capitolo VII dell'elaborato).

[3] V. G. Salvadori, 2019, "IoT e AI: l'Intelligenza Artificiale incontra l'Internet of Things", *Internet of Things*, https://blog.osservatori.net/it_it/intelligenza-artificiale-e-iot.

Un'altra casistica che possiamo analizzare riguarda la possibilità che dispositivi indossabili applicati in ambito industriale, ad esempio all'interno di una fabbrica (riprendendo il concetto affrontato prima di Smart Factory), possano raccogliere informazioni relative all'ambiente di lavoro in maniera autonoma, in modo da comprendere se il lavoratore si stia esponendo o meno ad una situazione di pericolo, allertandolo tempestivamente al fine di scongiurare potenziali incidenti: un ulteriore esempio sul medesimo settore riguarda l'uso dell'AIoT per la manutenzione predittiva, mediante la rilevazione di suoni anomali dai macchinari mediante i sensori impiegati per prevenirne eventuali tempi di fermo legati a guasti o malfunzionamenti.
Pertanto, con riferimento alle classiche operazioni di gestione che verrebbero attuate negli esempi finora descritti, la qualità e la quantità dei dati raccolti risulta essere di gran lunga superiore proprio per l'integrazione dell'AI con il mondo IoT, garantendo da un lato una maggiore efficienza nella gestione dei processi aziendali, degli ambienti urbani, sicurezza pubblica e customer expirience, migliorando così la qualità di ogni potenziale servizio offerto, ma dall'altro sollevando diverse questioni sia sotto l'ambito privacy che quello della sicurezza delle informazioni stesse che transitano in tali sistemi.

I.4. Rapidità e adattività del calcolo: l'Edge Computing

Fondamentale per rendere tale integrazione ancora più sinergica e rapida risulta necessario l'impiego di ulteriori tecnologie all'avanguardia, prima fra tutte l'adozione di paradigmi di Edge Computing.
A livello macroscopico un'architettura di Edge Computing si presenta come un'infrastruttura di Information Technology distribuita e decentralizzata, in grado di elaborare grosse moli di informazioni in tempi brevi attraverso una strategica elaborazione dei pacchetti di informazioni il più vicino possibile al dispositivo di raccolta degli stessi, in base ad una gerarchia di criticità del dato.
La società di analisi di mercato IDC fornisce la seguente definizione di Edge Computing: "una rete mesh di micro data center, in grado di elaborare e memorizzare dati critici localmente, e di trasmettere tutti i dati ricevuti e/o elaborati a un data center centrale o a un repository di cloud storage".[4]
In sostanza, l'Edge Computing permette di alleggerire determinate applicazioni "time-sensitive" dalla stretta dipendenza della loro connessione con i data center remoti di riferimento,

[4] V. Redazione ZeroUno (a cura di), 2022, "Edge computing, cos'è, come funziona e come implementarlo", *Cloud Computing*, https://www.zerounoweb.it/techtarget/searchdatacenter/edge-computing-cose-come-implementarlo/.

consentendo, tramite l'impiego di risorse di calcolo locali, di elaborare i dati direttamente sul campo, senza loro spostamenti ulteriori.

Tra gli esempi d'uso dell'Edge Computing, come possiamo intuire, rientra certamente tutta quella serie di dispositivi che costituiscono il mondo IoT: i quali si trovano spesso a dover fronteggiare problematiche legate alla latenza, la mancanza di banda e l'affidabilità, non potendo quindi appoggiarsi per le operazioni di calcolo ai modelli cloud tradizionali.

La struttura stessa dell'Edge Computing è in grado di ridurre drasticamente la mole di dati che deve approdare sul cloud, mediante un'elaborazione dei dati critici sensibili alla latenza direttamente dalla sorgente della raccolta tramite uno smart device, o, in alternativa, inviando i dati ad un server intermedio, localizzato ad una distanza sensibilmente più vicina a quella che separa il dispositivo dal server centrale.

I dati che risultano essere meno 'time-sensitive' possono invece essere trasmessi direttamente all'infrastruttura cloud o al data center centrale, per permettere l'esecuzione da parte delle macchine di elaborazioni più complesse e che richiedono un maggiore tempo di calcolo computazionale.

Fra questi ultimi rientra ad esempio quella che è l'analisi dei big data, le attività di training per il fine tuning dei modelli di Machine Learning (ML), la conservazione di dati per un lungo periodo e l'analisi di dati storici conservati nei database.

L'Edge Computing è indispensabile per quelle soluzioni dove si palesa la necessità di elaborare in tempi molto rapidi una risposta a una serie indefinita e non prevedibile di stimoli esterni.

Pensiamo, ad esempio, ad un robot che si occupa di effettuare in autonomia consegne di pacchi a domicilio muovendosi nel caotico traffico cittadino (come nel caso di Amazon Scout) e ha bisogno di poter interagire con l'ecosistema circostante rielaborando il proprio percorso ogni volta che si presenti un ostacolo imprevisto sul proprio cammino. [5]

L'innovazione costituita da tale tecnologia, seppur affascinante, solleva ovviamente la discussione su importanti temi in ambito Cyber Security e Data Protection, i quali si accompagnano ad ogni innovazione tecnologica, specialmente quelle che risultano interagire con la nostra vita di tutti i giorni.

I.5. I rischi per la privacy e la Cyber Security

Negli ultimi anni, il concetto di Smart City è diventato sempre più diffuso non solo per descrivere alcune delle città più innovative dal punto di vista tecnologico e ambientale, ma

[5] V. G. Salvadori, 2019, "IoT e AI: l'Intelligenza Artificiale incontra l'Internet of Things", *Internet of Things*, https://blog.osservatori.net/it_it/intelligenza-artificiale-e-iot.

soprattutto per delineare una visione delle città del futuro: come saranno progettate, come funzioneranno e quali nuove opportunità offriranno ai cittadini.
Grazie all'uso intensivo della tecnologia e dell'innovazione, quali l'IoT, l'AI e l'Edge Computing, le aree urbane diventano così "intelligenti", portando alla trasformazione digitale e ad un continuo progresso in vari ambiti applicativi.
In questo cambiamento, tuttavia, non possiamo solamente affidarci al solo aspetto avveniristico, ma osservare e valutare con attenzione i rischi che si profilano negli ambiti della Data Protection e della Cyber Security.
La principale preoccupazione associata all'impiego estensivo dell'IoT risulta essere la mole elevatissima di dati personali che questi sistemi possono raccogliere: ogni cosa, dalla nostra posizione al nostro comportamento, alle preferenze e allo stato di salute, può essere catturata da questi dispositivi, creando una digital footprint che potrebbe permettere una profilazione molto dettagliata delle persone grazie al tracciamento delle loro attività nell'impiego dei dispositivi, oltre che alle conseguenti problematiche legate ad un loro costante ed invasivo monitoraggio.
Con la crescente diffusione di soluzioni come quelle ideate ad esempio, per le Smart Homes, le aziende si interrogano frequentemente su come sia possibile cogliere il massimo potenziale legato ai dati resi disponibili da questo vasto impiego di dispositivi intelligenti.
A prova di questo profondo interesse commerciale per i dati personali basti osservare le strategie adottate dai grandi player del mercato tecnologico come Google e Amazon, che tra le principali campagne di vendita hanno lanciato sul mercato una gamma di dispositivi smart home speaker (come Alexa e Google Home) a prezzi considerevolmente bassi, puntando a generare un business che va ben oltre la mera vendita dell'hardware in se considerato: ad esempio, in questo caso, l'interesse di queste aziende riguarda scopi ben più ampi, che vanno dalla profilazione dettagliata degli utenti, non solo con riferimento alla loro user experience nella navigazione online, ma anche attraverso assistenze granulari e personalizzate nel supporto agli acquisti effettuati da questi soggetti, i quali possono essere indirizzati alla piattaforma eCommerce del grande player (come nel caso di Amazon) oppure veicolare gli acquisti facendo leva su retailer terzi quali tramite del business diretto di queste grandi società.
Pensiamo al caso di Google Express, che permette (negli Stati Uniti) al consumatore di impiegare con dei semplici comandi vocali l'home speaker per effettuare acquisti da una rete strutturata di circa cinquanta retailer convenzionati con Google.
In questo scenario di crescita tecnologica e di nuove soluzioni IoT è necessario che la normativa sia in costante evoluzione, al fine di garantire da una parte la competitività nel mercato tra queste

aziende, e dall'altro preservare i dati personali degli interessati garantendo il rispetto dei requisiti di legittimità da parte dei titolari dei trattamenti di dati personali e mediante la definizione ed adozione, da parte degli stessi, di adeguati livelli di sicurezza, disciplinandoli non solo nei processi più organizzativi, ma anche in quelli di natura più tecnica.

Consideriamo infatti la possibilità che dei soggetti malintenzionati decidessero di attaccare questi sistemi.

Mediante attacchi informatici e lo sfruttamento delle vulnerabilità insite a questi sistemi potrebbero infatti, ad esempio, impiegare i dati raccolti per la creazione di profili dettagliati degli utenti al fine di venderli agli inserzionisti per trarne profitto, oppure utilizzarli per commettere furti di identità, frodi finanziarie e varie altre forme di criminalità informatica.

Il potenziale sconfinato costituito dai possibili impieghi dei dati personali non si limita inoltre alle sole minacce esterne quali eventuali attaccanti: i dispositivi IoT possono essere soggetti, come già accennato, a vulnerabilità tecniche che li espongano a potenziali violazioni dei dati, con il rischio che i dati personali, sia comuni che particolari, vengano diffusi ad una moltitudine indefinita e incontrollata di soggetti.

Questa tipologia di rischio risulta particolarmente accentuata nei dispositivi AIoT che presentano caratteristiche di sicurezza informatica deboli o del tutto inesistenti, in quanto non definite by design, portando ad un'elevata facilità della loro violazione da parte di criminali informatici: se consideriamo, ad esempio, i risultati dell'Annual Internet Report di Cisco, questo stima che nel solo 2023 sono stati registrati 29,3 miliardi di dispositivi IoT connessi nel Cyber-spazio, ognuno dei quali è in grado di generare un'enorme quantità di dati, rappresentando un'opportunità significativa per gli hacker e altri soggetti malintenzionati, i quali potrebbero sfruttare queste informazioni per i propri scopi personali.[6]

Anche i rischi legati alla Cyber Security sono un punto focale da tenere a mente non solo per i soggetti che lavorano più verticalmente in ambito tecnico su questi sistemi, ma anche per coloro che affrontano queste tematiche da un punto di vista legale: indispensabile deve essere l'unione dei due mondi, o perlomeno una strutturata interconnessione e collaborazione dei due, al fine di poter considerare ogni possibile problematica che costituisca una minaccia diretta o indiretta per gli interessati oggetto dei trattamenti svolti da titolari che impiegano questi sistemi complessi di raccolta dei dati.

[6] V. Redazione di Talking IOT (a cura di), 2023, "How does IoT impact privacy and data security?", Talking IOT, https://talkingiot.io/how-does-iot-impact-privacy-and-data-security/.

II. Il quadro normativo

Nel panorama tecnologico appena descritto, l'innovazione portata dall'integrazione dell'AI nei sistemi di Internet of Things per una loro applicazione nelle Smart Cities costituisce un focus di grande interesse e innovazione, il quale porta con sé una pletora di sfide di profonda complessità.

Questi sistemi sono infatti caratterizzati dall'enorme quantità di dati personali che generano, raccolgono e trattano, spesso in tempo reale, mediante i molteplici dispositivi e sistemi intelligenti interconnessi tra loro, i quali interagiscono con gli utenti in modo sempre più personalizzato.

Questi scenari portano con loro una serie di interrogativi cruciali riguardo alla sicurezza informatica e alla protezione e gestione delle informazioni sensibili raccolte.

In particolare, questo capitolo si concentrerà sul Regolamento Generale sulla Protezione dei Dati 679/2016 (GDPR) e l'AI Act, oltre che ad una serie di altre normative e linee guida pertinenti, analizzando le principali problematiche legate alla protezione dei dati all'interno di questi ambienti sempre più interconnessi e automatizzati, effettuando rimandi, in particolar modo, al caso pratico da me riportato.

II.1. Una contestualizzazione sull'Internet of Things (IoT)

Con riferimento alle soluzioni Internet of Things, la raccolta dei dati mediante l'insieme di sensori e dispositivi che costituiscono questo ampio mondo pone questioni centrali dal punto di vista della sicurezza dei dati trattati, specialmente se li inseriamo all'interno di contesti di utilizzo particolarmente critici e diffusi, come quelli delle Smart Cities, dove queste reti di sensori costituiscono gli occhi e le orecchie dei complessi sistemi utili all'amministrazione della città in maniera efficiente.

Nel mondo IoT, la principale normativa di riferimento per le operazioni di trattamento svolte mediante tali soluzioni è costituita dal Regolamento UE 2023/2854 (c.d. Data Act): con esso, l'Unione Europea mira a promuovere non soltanto lo sviluppo dei servizi e prodotti IoT nuovi ed innovativi, ma anche a stimolare maggiormente lo sviluppo di ulteriori servizi che impieghino dati già raccolti provenienti da una varietà di servizi correlati o comunque connessi (tra cui IoT, auto connesse, elettrodomestici, ecc.).

In questo modo, il Data Act prevede che le imprese produttrici o fornitrici di prodotti IoT possano dotarsi di strumenti che siano adeguati alla disciplina in ambito Data Protection, al fine di garantire all'utente finale l'accesso e la portabilità dei dati che saranno generati dall'utilizzo

dei prodotti IoT stessi e dalle operazioni di trattamento correlate, ponendo le aziende nella posizione di rispondere in maniera efficace alle richieste provenienti dagli utenti volte all'accesso o alla portabilità dei propri dati, oltre che alle richieste di dati provenienti dagli organismi pubblici.[7]

Inoltre, a supporto di questo regolamento, esistono delle linee guida che risultano maggiormente verticali sulla corretta gestione e messa in sicurezza di tali soluzioni, in modo da garantire di riflesso una maggiore sicurezza dei dati trattati da questi sistemi.

Queste linee guida risultano essere quelle indicate all'interno della norma tecnica ISO 27400 la quale, nascendo nel giugno del 2022, si pone l'obiettivo di fornire delle linee guida per la sicurezza informatica e la privacy dei sistemi IoT, come anche la ISO 27402 (sui requisiti di base dei device), la ISO 27403 (linee guida sull'IoT nella domotica) e la ISO 27404 (framework di Cyber Security per l'IoT destinato ai consumatori), tutte intitolate "IoT security and privacy".

Queste norme tecniche appartengono alla famiglia della ISO 27000, la quale definisce in base ai vari settori specifici di riferimento i principali standard di sicurezza da seguire con riferimento al mondo IT.

Va tenuto a mente che essendo solamente delle linee guida, le indicazioni fornite da queste norme non sono considerabili come dei veri e propri requisiti e non trovano una loro applicazione obbligatoria, ma possono certamente aiutare i fornitori dei sistemi IoT a definire correttamente le politiche legate alla privacy dei dati trattati e alla loro sicurezza.[8]

All'interno della ISO 27400 sono innanzitutto indicati una serie di controlli a livello organizzativo necessari per le aziende volti a facilitare la gestione della sicurezza dei sistemi IoT prodotti, come, ad esempio, la definizione di politiche aziendali specifiche e la definizione di specifici ruoli e responsabilità all'interno dell'azienda, oltre che a controlli che sono generalmente raccomandati per tutti i sistemi informatici, quali il monitoraggio del funzionamento del sistema, la raccolta dei log che tengono traccia degli eventi e l'acquisizione di informazioni derivanti da eventuali incidenti.

Sono inoltre definiti i principi di ingegneria sicura dei sistemi IoT, che indirizzino la progettazione, sviluppo e implementazione di questi sistemi e la loro messa in sicurezza.

Adeguate misure di sicurezza dovrebbero essere implementate durante tutte le fasi del ciclo di vita dei sistemi IoT, a partire dalla progettazione fino alla messa in opera, manutenzione e dismissione di tali sistemi, accostandosi in questo contesto a quello che è il principio di Data Protection by Design enunciato all'articolo 25 del GDPR.[9]

---

[7] V. Regolamento UE 2023/2854, "*Data Act*".
[8] V. ISO/IEC 27400/2022, "*IoT Security and Privacy Guidelines*".
[9] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 25.

Oltre a questi controlli legati più verticalmente alla sicurezza dei sistemi IoT, viene inoltre raccomandata dalla ISO 27400 l'esecuzione di controlli più mirati sul garantire la privacy nella gestione sicura dei dati personali trattati da tali sistemi: tra i controlli che possono essere applicati da parte dei fornitori di servizi IoT e dagli sviluppatori di questi sistemi vi sono l'integrazione di funzionalità per il miglioramento della privacy all'interno di dispositivi e servizi IoT, oltre che ad un'adeguata gestione dei controlli sulla privacy, che preveda un riesame continuativo dell'efficacia dei presidi stabiliti, andando ad identificare quelli che potrebbero essere i nuovi rischi emergenti posti da tali sistemi, tenendo sempre in considerazione l'evoluzione delle esigenze degli utenti e dei requisiti normativi.[10]

Risulta fondamentale ridurre al minimo la raccolta dei dati da parte di fonti indirette, o ancora meglio, evitarla del tutto, in modo da impedire che questa avvenga senza il loro consenso durante il trattamento svolto mediante i sistemi IoT.

È infatti importante anche tutto ciò che riguarda la gestione appropriata delle misure di protezione dei dati personali, le quali devono essere comunicate solo alle parti interessate, garantendo che i dati raccolti non possano mai permettere l'identificazione dell'interessato coinvolto, al fine di impedire la raccolta sistematica di informazioni ad esso relative attraverso un monitoraggio sistematico effettuato da un dispositivo IoT.

Infine, in linea con l'obbligo stabilito all'articolo 12 del GDPR, all'utente IoT dovrebbe essere comunicata un'apposita informativa sulla privacy che indichi, prima di tutto, quali dati personali verranno raccolti dal sistema e le finalità a supporto del loro impiego.[11]

In conclusione, la ISO 27400 costituisce sicuramente una norma fondamentale per stabilire delle linee guida sulla corretta implementazione dei sistemi IoT e sulla loro messa in sicurezza, andando a disciplinare, allineandosi con il GDPR, alcuni aspetti più verticali sulla privacy, al fine di garantire un trattamento dei dati personali mediante questi sistemi il più possibile tutelante per gli interessati, indispensabile in contesti ampi e complessi di impiego quali le Smart Cities.

II.2. Sinergia tra GDPR e AI Act: una novità a livello mondiale…tutta europea!

Esaminiamo ora le sinergie tra il GDPR e l'AI Act, per comprendere al meglio come questi due regolamenti vadano a impostare una visione a 360 gradi sulla protezione dei dati degli interessati: nel secondo caso, il focus è incentrato sui sistemi di AI che come vedremo possono costituire profili particolarmente critici con riferimento all'invasività dei trattamenti messi in contesti che

[10] V. B. Ridolfi e A. Vaccarelli, 2023, "Sicurezza IoT, come evitare i rischi: le norme tecniche e la privacy", *Sicurezza Digitale*, https://www.agendadigitale.eu/sicurezza/le-norme-tecniche-e-la-privacy-per-la-sicurezza-nelliot-cosa-sapere-per-evitare-rischi/.

[11] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 12.

vedono l'impiego di sistemi integrati particolarmente complessi, quali quelli impiegati nelle Smart Cities, che, come vedremo nel caso da me affrontato, possono portare ad esempio al rischio di un controllo e monitoraggio di massa e permanente (oltre che estremamente dettagliato e invadente) delle persone.

Il Regolamento UE 1689/2024, rinominato "AI Act" seppur inizialmente concepito nel 2021 approda nel panorama normativo europeo solo a partire dal 2024, vedendo una sua concreta applicazione solo all'alba di quest'anno.

Seppur impropriamente etichettato come "Act" o "legge", a differenza della Direttiva UE (che, come sappiamo, richiede una legge nazionale di recepimento per la sua entrata in vigore), essendo un Regolamento, è direttamente applicabile a tutti gli stati membri, mirando a disciplinare le regole, garanzie e condizioni in base alle quali i sistemi di AI provenienti dall'estero potranno o meno essere immessi, distribuiti ed impiegati sul territorio dell'Unione Europea, cercando di chiarire e uniformare l'approccio europeo relativo all'adozione di tali modelli.

Il principale obiettivo posto dal Regolamento, riprendendo quanto indicano nel suo primo considerando, è quello di migliorare il funzionamento del mercato interno europeo istituendo un quadro giuridico uniforme fra gli stati membri, in particolare per quanto riguarda:

- lo sviluppo, l'immissione sul mercato, la messa in servizio e l'uso di sistemi di intelligenza artificiale all'interno dell'Unione, in conformità ai valori della stessa;
- promuovere la diffusione di un'intelligenza artificiale antropocentrica e affidabile, garantendo sempre un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea (tra cui la democrazia, lo Stato di diritto e la protezione dell'ambiente);
- proteggere contro gli effetti nocivi dei sistemi di IA nell'Unione, nonché promuovere l'innovazione.

Il Regolamento mira, inoltre, a garantire la libertà di circolazione transfrontaliera dei beni e servizi basati sull'adozione di modelli di intelligenza artificiale, impedendo così agli stati membri stessi di imporre restrizioni allo sviluppo, alla commercializzazione e all'uso di sistemi di AI (salvo su espressa autorizzazione del regolamento stesso), evitando così una frammentazione estesa tra le normative interne ad ogni entità nazionale.[12]

A rendere particolarmente interessante la sinergia tra l'AI Act e il GDPR è il comune approccio adottato da entrambi i regolamenti: si tratta infatti di un'impostazione risk based, cioè basata sulla valutazione dei potenziali impatti sulle persone coinvolte.

[12] V. Regolamento (UE) 2024/1689, "*EU AI Act*", considerando 1.

Nel primo caso con riferimento ai modelli di AI adottati, mentre nel secondo con riferimento alle attività di trattamento dei dati personali.

Nel considerare, infatti, le necessità dello sviluppo dell'AI, la visione adottata è di tipo antropocentrico, mettendo in risalto la posizione degli interessati, facendo leva sull'importanza del rispetto dei principi di trasparenza, affidabilità e resilienza.

In pratica, muovendosi a sostegno all'innovazione, grazie a un approccio basato sul rischio (contestualizzato e mutato dal GDPR) mira a proteggere gli interessati senza tuttavia "bloccare" i processi di innovazione, tenendo a mente da una parte la protezione dei dati e dall'altra l'importanza economica, tecnologica e sociale costituita dall'adozione di tali tecnologie.

L'AI Act va ad elaborare una serie di regole graduate a seconda del livello di "rischio" del sistema di AI osservato, riprendendo la formula costituita dalla combinazione tra probabilità e impatto rifacendosi a quella presente, nel panorama legislativo europeo, all'interno del GDPR.

In base a questa formula, il Regolamento distingue tra 4 categorie di sistemi di AI:

- sistemi a rischio inaccettabile;
- sistemi ad alto rischio;
- sistemi per finalità generali;
- sistemi a rischio minimo.

II.1.1. I sistemi a rischio inaccettabile

Nell'AI Act sono considerati sistemi di intelligenza artificiale a rischio inaccettabile, ai sensi del contenuto dell'articolo 5, tutte quelle soluzioni che riguardano sostanzialmente l'immissione sul mercato, la messa in servizio o l'uso di un sistema di AI per finalità o con modalità operative tali da porre un rischio per i diritti degli interessati talmente elevato da risultare non adottabili in ogni casistica.

In particolare, il Regolamento ricomprende tra queste tipologie quei sistemi che adottano i cosiddetti "dark-patterns", cioè quell'insieme di tecniche subliminali che sfruttano un intervento inconsapevole dell'uomo e quelle che utilizzano tecniche che risultino "volutamente manipolative o ingannevoli" per lo scopo perseguito o per via dell'effetto di distorsione materiale del comportamento di una o più persone, pregiudicando notevolmente la loro capacità di prendere una decisione o di prenderne una che, senza le influenze poste da tali tecniche, "non avrebbero altrimenti preso", comportando un danno significativo per i loro diritti ed interessi.[13]

[13] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 5, lettera a.

Esistono sistemi che, sfruttando le vulnerabilità di una o più persone appartenenti a particolari categorie (ad esempio per età, disabilità o altra situazione disagiata di natura sociale o economica), puntano a "distorcere" materialmente il comportamento di questi soggetti, al fine di cagionare loro un danno.[14]

Altri esempi di sistemi ricompresi in questa categoria sono:[15]

- sistemi che si occupano di compiere valutazioni o previsioni circa la probabilità di commissione di un reato "unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità";
- sistemi che costituiscono o alimentano delle banche dati volte al riconoscimento facciale mediante scraping non mirato di immagini facciali acquisite da internet o da filmati di telecamere a circuito chiuso;
- sistemi basati sul riconoscimento delle emozioni sul lavoro e nella scuola, salvo che non siano adottati "per motivi medici o di sicurezza";
- sistemi volti ad effettuare una categorizzazione biometrica tale da individuare le persone fisiche sulla sola base dei loro dati biometrici, dai quali mira a trarre deduzioni o inferenze relativamente a tutti quei dati particolari quali la razza, le opinioni politiche, l'appartenenza sindacale, le convinzioni religiose filosofiche, l'orientamento sessuale;
- sistemi utilizzati per valutare o classificare le persone fisiche o gruppi di persone per un determinato periodo di tempo sulla base del loro comportamento sociale (cd. "social scoring").

Si apre inoltre il discorso relativo all'impiego di sistemi di identificazione biometrica remota in tempo reale e in spazi accessibili al pubblico ai fini di attività di contrasto, sì ammessi ma limitatamente per confermare l'identità della persona: queste sarebbero in realtà utilizzabili purché sia tenuta in conto la natura della situazione che dà luogo al possibile uso di tale sistemi in situazioni di particolare gravità e in considerazione della probabilità ed entità del danno risultante in caso di mancato impiego di questi sistemi, nonché delle conseguenze per i diritti e le libertà di tutte le persone interessate.

L'attenzione riservata dal legislatore ai sistemi che trattano dati biometrici è un elemento di particolare rilevanza per il caso trattato.

Infatti, durante l'analisi preliminare d'impatto, è stata approfondita la valutazione riguardo all'applicabilità delle disposizioni previste rispetto al caso specifico analizzato, come verrà descritto nel capitolo IV.

[14] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 5, lettera b.
[15] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 5.

II.1.2. I sistemi ad alto rischio

Ai sensi dell'articolo 6 dell'AI Act, sono considerati sistemi ad alto rischio tutti quei sistemi che costituiscono componenti di sicurezza di prodotti che rientrano nell'ambito di applicazione, tanto in termini di regole di immissione sul mercato quanto di certificazione da parte di un ente terzo, di una serie di settori richiamati nell'allegato I del Regolamento, tra i quali vi sono:

- i sistemi volti alla raccolta di dati biometrici
- i sistemi di categorizzazione basati sulla deduzione di attributi sensibili
- i sistemi funzionali al riconoscimento delle emozioni delle persone in ambito di sicurezza e per motivi medici.

All'interno di questa tipologia di sistemi rientrano anche quelli impiegati come componente di sicurezza delle infrastrutture critiche digitali, del traffico stradale o della fornitura di acqua, riscaldamento o elettricità.

Vi sono altri ambiti quali l'istruzione e la formazione, dove questi sistemi possono essere impiegati per l'accesso, l'ammissione o l'assegnazione delle persone a scuole ed istituti (o per valutare l'apprendimento), o nell'ambito lavorativo, pensando a quei sistemi che vengono adottati per guidare il personale di HR nel processo di assunzione, tra cui la selezione e l'analisi delle candidature e la valutazione stessa dei candidati, oltre ai sistemi che gestiscono aspetti legati al percorso lavorativo delle persone, volti ad esempio a compiere decisioni nel contesto dei rapporti di lavoro o assegnare compiti basati su caratteristiche o comportamenti personali.

Infine, rientrano in tale categoria anche quei sistemi che operano in ambito pubblicistico, o meglio nella sfera di fruizione e garanzia dei servizi e diritti essenziali come la salute e la giustizia, quindi servizi pubblici essenziali.

Va precisato che la prima tipologia di sistemi citata, ovvero quelli impiegati quali componente di sicurezza di un prodotto o che sono il prodotto stesso sono classificati ad alto rischio tout court, come le categorie di prodotti che erano già considerati particolarmente impattanti anche in assenza di componenti di AI, e già soggetti a certificazioni e verifiche espressamente previsti dalle normative europee dei rispettivi settori di riferimento.[16]

Come per i sistemi a rischio inaccettabile sopra citati, il criterio adottato per determinare se un sistema AI rientra all'interno di tale categoria o meno è il medesimo, cioè la capacità, anche solo potenziale, di arrecare un danno ai diritti fondamentali delle persone, alla loro salute e alla loro sicurezza.

[16] V. L. Garbati, C. Ponti, 2024, "AI Act: che cos'è, obiettivi e sanzioni previste", *Intelligenza Artificiale*, https://www.ai4business.it/intelligenza-artificiale/ai-act-che-cose-obiettivi-e-sanzioni-previste/.

Pe rientrare in questa specifica categoria il danno non deve essere significativo come per i modelli vietati, ma deve pur sempre risultare particolarmente impattante alla luce di un suo impiego scorretto o di rischi legati alla sua sicurezza che vadano ad intaccare svariati elementi della vita delle persone.

II.1.3. I sistemi di Intelligenza Artificiale per finalità generali

All'articolo 51 dell'AI Act vengono esaminati quei sistemi che sono considerati "modelli di AI per finalità generali".

Questi vengono considerati tali sulla base di due condizioni: la prima, se presentano delle capacità di impatto elevate, valutate sulla base di strumenti tecnici e metodologie adeguate (compresi indicatori e parametri di riferimento), mentre come seconda se, sulla base di una decisione della Commissione, sono considerate aventi capacità o impatti equivalenti a quelli definiti per la prima condizione, tenendo conto dei criteri di cui all'allegato XIII.

All'allegato XIII si stabilisce infatti che la Commissione tiene a mente determinati fattori nel compiere le sue valutazioni in merito al livello di rischio del sistema di AI usato per finalità generali, osservando in particolare:[17]

- il numero di parametri del modello;
- la qualità o la dimensione del set di dati;
- la quantità di calcolo utilizzata per addestrare il modello misurata in operazioni in virgola mobile superiore a 10^25 (in termini di potenza di calcolo) o indicata da una combinazione di altre variabili (quali il costo, il tempo necessario e il consumo energetico stimati per l'addestramento);
- le modalità di input e output del modello, come da testo a testo (in caso di LLM, cioè modelli linguistici di grandi dimensioni), da testo a immagine, multimodalità e soglie di punta per determinare le capacità ad alto impatto per ciascuna modalità, nonché il tipo specifico di input e output (ad esempio sequenze biologiche);
- i parametri di riferimento e le valutazioni delle capacità del modello, anche tenendo conto del numero di compiti che non richiedono un addestramento aggiuntivo, la capacità di apprendere nuovi compiti distinti, il livello di autonomia e scalabilità e gli strumenti a cui ha accesso;

[17] V. Regolamento (UE) 2024/1689, "*EU AI Act*", allegato XIII.

- se il modello ha un alto impatto sul mercato interno in considerazione della sua portata, che viene presunta quando il modello stesso è stato messo a disposizione di almeno 10.000 utenti commerciali registrati stabiliti nell'Unione;
- il numero di utenti finali registrati.

L'AI Act pone degli obblighi in capo ai fornitori e deployer di sistemi di AI per finalità generali, tra i quali spiccano, in termini di Data Protection, quelli stabiliti dall'articolo 50, il quale dimostra più di tutti la sinergia esistente con il GDPR e con la tendenza della commissione europea di voler uniformare e mantenere unitaria la serie di indirizzi volti alla protezione dei dati personali degli interessati anche in ambiti innovativi come l'AI.

In particolare, la norma dispone che questi soggetti garantiscano che i sistemi di AI destinati a interagire direttamente con le persone fisiche vengano progettati e sviluppati in modo tale che gli interessati oggetto del trattamento siano informati del fatto di stare interagendo con un modello di AI (a meno che ciò non risulti evidente dal punto di vista di una persona fisica ragionevolmente informata, attenta e avveduta, tenendo conto delle circostanze e del contesto di utilizzo).[18]

Questo disposto si allinea perfettamente con il principio di Data Protection by Design di cui all'articolo 25 del GDPR, dimostrando come il legislatore europeo volesse riprendere, anche con riferimento ai fornitori di modelli di AI, l'obbligo di garantire un corretto livello di privacy e di protezione dei dati personali fin dalla fase di progettazione di qualunque sistema, servizio, prodotto o processo così come durante il loro ciclo di vita.[19]

I fornitori e deployer di sistemi di AI, compresi i sistemi adibiti a finalità generali, che generano contenuti audio, immagine, video o testuali sintetici, sono tenuti a garantire in particolar modo che gli output di tali sistemi siano strutturati in un formato leggibile e che sia possibile comprendere che questi siano frutto di una generazione o manipolazione artificiale. [20]

Quest'obbligo di trasparenza, richiamato anche in numerose sezioni dell'AI Act stesso, tra cui il considerando 67, si allinea con gli obblighi di trasparenza imposti anche dal disposto dell"articolo 12 del GDPR, compiendo un rimando diretto al Regolamento 679/2016 affermando che "*Al fine di agevolare il rispetto del diritto dell'Unione in materia di protezione dei dati, come il regolamento (UE) 2016/679, le pratiche di governance e di gestione dei dati dovrebbero includere, nel caso dei dati personali, la trasparenza in merito alla finalità originaria della raccolta dei dati*".[21]

[18] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 50.
[19] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 25.
[20] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 50.
[21] V. Regolamento (UE) 2024/1689, "*EU AI Act*", considerando 67.

In fondo, gli stessi sette principi fondamentali alla base dell'AI Act comprendono proprio quello di trasparenza, assieme a quelli di intervento e sorveglianza umani (c.d. "Human In The Loop"), robustezza tecnica, sicurezza, vita privata, governance dei dati, diversità, non discriminazione, equità, benessere sociale, ambientale e responsabilità, dove anche quest'ultimo è espressione pura del principio di accountability di cui all'articolo 24 del GDPR.[22]

Inoltre, secondo il dettato dell'articolo 50, tali fornitori sono tenuti a garantire che le loro soluzioni tecniche siano efficaci, interoperabili, solide e affidabili nella misura in cui ciò sia tecnicamente possibile, tenendo conto delle specificità e dei limiti dei vari tipi di contenuti, dei costi di attuazione e dello stato dell'arte generalmente riconosciuto, come eventualmente indicato nelle pertinenti norme tecniche.[23]

Questa disposizione riprende non solo i contenuti, ma le stesse parole dell'articolo 32 del GDPR, il quale richiede ai titolari del trattamento di dotarsi di adeguate misure tecniche ed organizzative a protezione dei dati personali oggetto del trattamento, sempre ponendo anche l'attenzione allo stato dell'arte e ai costi attuativi necessari per l'adozione di tali misure a seconda dell'attività svolta.[24]

Infine, i deployer e fornitori dei modelli di AI mirati al riconoscimento delle emozioni o alla categorizzazione biometrica sono espressamente tenuti ad informare le persone fisiche che sono esposte a tali tecnologie relativamente al funzionamento effettivo del sistema e che il trattamento dei dati personali avviene in conformità dei Regolamenti UE 2016/679 e 2018/1725 e della Direttiva 2016/680, a seconda dei casi.[25]

II.1.4. I sistemi a rischio minimo

Tutti quei modelli che invece non rientrano, per le loro caratteristiche, all'interno di una delle categorie sopra citate ricadono automaticamente tra i sistemi di AI ad impatto minimo, per i quali non vi sono particolari previsioni da adottare in considerazione del loro impatto minimo sugli interessati, se non un generale suggerimento di adottare adeguate misure volte a garantire una corretta sicurezza delle informazioni: alcuni esempi di questi sistemi possono essere le Intelligenze Artificiali impiegate nel mondo videoludico, o ancora i filtri antispam.

[22] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 24.
[23] V. Regolamento (UE) 2024/1689, "AI Act", art. 50.
[24] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 32.
[25] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 50.

II.1.5. Accenno ad alcune norme dell'AI Act e GDPR affrontate nel caso pratico

Esaminiamo ora alcuni articoli e disposizioni che risultano essere particolarmente importanti per quella che sarà l'analisi del caso pratico da me riportato in questo elaborato, nel quale presenterò una soluzione che, tra le varie tecnologie innovative che adotta, quali sensori IoT e paradigmi di Edge Computing, impiega anche un modello di AI di Video Analytics, basato sul Machine Learning e Deep Learning.

Come vedremo, tuttavia, vi sono diverse precauzioni che sono state prese per ridurre il rischio fino ad un livello addirittura limitato, nonché una diffusa conformità ai principi e disposizioni stabiliti sia dal GDPR che dall'AI Act.

Con particolare riferimento a quest'ultimo, molte delle sue previsioni riguardano misure obbligatorie volte a mitigare il rischio costituito da questi sistemi di AI, che seppure apparentemente riferite solo ai sistemi ad alto rischio, sono in realtà misure proprie per una buona ed efficiente gestione di impresa e di gestione dei dati trattati.

L'articolo 9 prevede che i sistemi di AI ad alto rischio implementino un sistema di gestione del rischio, inteso quale processo continuativo e ben pianificato eseguito nel corso dell'intero ciclo di vita del sistema di AI, il quale impone un periodico e sistematico svolgimento di attività di riesame e aggiornamento.[26]

Fornendone una sintesi, deve essere integrato e adottato un processo che permetta di valutare sia i rischi connessi alla finalità originariamente immaginata per l'adozione del sistema, sia con riferimento ad eventuali usi impropri che possono essere fatti dello stesso, in modo tale da fare in modo che i rischi residui rientrino in una fascia accettabile per il fornitore o deployer del sistema.

Proprio per questa ragione prima della sua immissione sul mercato, questi sistemi devono essere obbligatoriamente sottoposti a test e prove eseguite, riprendendo quanto indicato nell'articolo 9 stesso: "in un qualsiasi momento dell'intero processo di sviluppo […] ed effettuate sulla base di metriche e soglie probabilistiche […] e adeguate alla finalità prevista perseguita dal sistema di AI ad alto rischio".[27]

Tale disposizione si fonda su alcuni importantissimi principi stabiliti all'interno delle linee guida contenenti le best practice da adottare in materia di sviluppo, prima fra tutte la best practice, consigliata anche dall'OWASP, della definizione di un Secure Software Development Life-Cycle (o SSDLC) strutturato da parte del fornitore del modello per garantire l'adozione di misure

[26] V. Regolamento (UE) 2024/1689, "*EU AI Act*", paragrafo 2 art. 9.
[27] V. Regolamento (UE) 2024/1689, "*EU AI Act*", paragrafo 8 art. 9.

adeguate di sicurezza durante tutta la fase di sviluppo, test, lancio in ambiente di produzione e continuo monitoraggio del sistema.[28]

L'articolo 10 contiene invece una delle raccolte più importanti di disposizioni dell'AI Act, e cioè quelle relative alla Governance dei dati raccolti ed elaborati dai sistemi di AI.

A ben vedere, l'intero mondo dell'AI risulta essere una questione fondata sulla raccolta, rielaborazione e produzione di dati, spesso anche personali: ben si comprende come l'adozione di un corretto sistema di governance dei dati costituisce, prima ancora di un obbligo, una vera e propria necessità per l'efficacia ed efficienza di qualsiasi fornitore e deployer di sistemi di AI, che questi siano modelli ad alto rischio o meno.

Con il termine Governance possiamo riferirci, oltre al rispetto delle previsioni contenute nel GDPR, anche alla qualità e integrità stessa dei dati: nessun sistema di AI, infatti, per quanto evoluto e potente, può produrre risultati validi se nutrito mediante dati inesatti o non pertinenti e "inquinati".[29]

Per i fornitori di sistemi di AI ad alto rischio , tale Governance del dato richiede in prima istanza un attento tracciamento e documentazione secondo alti standard di qualità (previsti ai paragrafi dal 2 al 5 dell'articolo 10) dei dataset impiegati per l'addestramento, convalida e prova del modello: ciò implica che i dati dovranno necessariamente essere gestiti in modo tale da garantire la conformità alle finalità per le quali è stato ideato il sistema in termini progettuali, tenendo sempre in conto l'origine dei dati raccolti e, se personali, della finalità a supporto del loro trattamento, incluse tutte le valutazioni del caso sulla loro adeguatezza e aggiornamento.[30]

Queste valutazioni dovranno inoltre concentrarsi sull'ipotizzare e strutturare una risposta alle possibili distorsioni (c.d. bias) che tali dataset potrebbero generare, mettendo successivamente in campo idonee misure correttive o di mitigazione.

Un'altra fondamentale disposizione risulta essere quella contenuta all'articolo 14 dell'AI Act, il quale prevede che i sistemi di AI ad alto rischio debbano essere progettati e sviluppati anche mediante l'impiego di strumenti di interfaccia uomo-macchina che risultino adeguati: in questo modo, il modello può essere efficacemente supervisionato da persone fisiche durante tutto l'arco di utilizzo dello stesso.

---

[28] V. Fondazione OWASP, 2005, "Secure development and integration", OWASP Developer Guide, https://devguide.owasp.org/02-foundations/02-secure-development/.

[29] V. V. L. Garbati, C. Ponti, 2024, "AI Act: che cos'è, obiettivi e sanzioni previste", Intelligenza Artificiale, https://www.ai4business.it/intelligenza-artificiale/ai-act-che-cose-obiettivi-e-sanzioni-previste/.

[30] V. Regolamento (UE) 2024/1689, "*EU AI Act*", paragrafi da 2 a 5 art. 10.

Questa misura sarà chiaramente proporzionata al rischio e alla struttura del modello di AI di caso in caso, in modo da poter permettere ai sorveglianti di raccogliere la necessaria quantità di informazioni per gestire con attenzione e consapevolezza il sistema nel suo complesso.[31]

In termini di Data Protection, il fornitore dovrà implementare processi "by design" che siano attivati sin dal momento della progettazione del sistema, andando ad effettuare un monitoraggio continuo del sistema (specialmente attraverso la tracciatura dei log di sistema), che permetta sempre la presenza di una sorveglianza umana, garantendo il cosiddetto approccio "Human In The Loop" (o HITL).

Al fine di mitigare il rischio sotteso al sistema di AI risulta vitale quale misura organizzativa, quella di creare modelli che siano resilienti e robusti, secondo il disposto dell'articolo 15 dell'AI Act, in modo tale da poter reagire al meglio non solo in risposta a errori, guasti e incongruenze, ma anche con riferimento a tentativi di soggetti malintenzionati di danneggiare o violare il sistema stesso.

Tali misure, le quali sono sia di natura tecnica che organizzativa, sono implementabili in diverse modalità: ad esempio, queste possono essere soluzioni di ridondanza del sistema presso server situati in una regione geografica diversa da quella del datacenter originario quale piano di Disaster Recovery, oppure possono essere piani di backup o fail-safe (cioè meccanismi di blocco del sistema) volti a garantire (specialmente con riferimento ai sistemi ad apprendimento continuo) l'eliminazione o riduzione del rischio della produzione di output potenzialmente distorti che vadano ad influenzare gli input impiegati in operazioni future (c.d. "feedback loops") o ancora dei sistemi di anonimizzazione dei dati in real-time che permettano ad esempio operazioni di offuscamento delle immagini raccolte da un sistema di Computer Vision (come vedremo nel caso pratico di cui al capitolo IV).[32]

Concludendo questo esame legislativo della materia, l'Unione Europea più di tutto mirava mediante la sinergia tra queste normative a realizzare un apparato documentale omogeneo e ben strutturato, in grado di tenere conto del contesto e della relativa complessità di volta in volta riscontrata nelle realtà aziendali e pubbliche, a prova dell'evidenza di come le organizzazioni gestiscono, organizzano e rispettano l'intero impianto normativo, e quindi la compliance che è integrata in quelle più virtuose in ottica di conformità al principio di accountability: tale compliance (intesa quale atteggiamento mentale dinamico di conformità e non certo di mera acquiescenza, spinto da un animus inerte e passivo) richiede che i Titolari che adottano tali tecnologie, così come i fornitori che le offrono, adempiano ad una serie di attività, quali

[31] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 14.
[32] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 15.

l'adozione di un sistema di gestione della qualità, l'implementazione di adeguate misure di sicurezza e l'esecuzione di attività di formazione e sensibilizzazione dei propri dipendenti e collaboratori, in ottica di permettere una cooperazione attiva con le Autorità di controllo, in ottica di trasparenza e collaborazione proficua nell'interesse di tutte le parti in gioco.

III. Le principali vulnerabilità e minacce cibernetiche

La rapida evoluzione delle tecnologie odierne, ed in particolare dei modelli di Intelligenza Artificiale e dei sistemi che adottano soluzioni di Internet of Things e di Edge Computing, l'innovazione e l'efficienza delle operazioni all'interno di numerosi settori hanno innegabilmente avuto una spinta sempre maggiore, fornendo vantaggi non solo nella vita di tutti i giorni, ma anche all'economia, alla gestione delle aziende, delle città e delle pubbliche amministrazioni.

Tuttavia, contestualmente a questi numerosi benefici, queste tecnologie hanno anche introdotto tutta una serie di nuove sfide in ambito Cyber Security: infatti, la crescente interconnessione dei dispositivi e la decentralizzazione dell'elaborazione dei dati, che caratterizzano l'IoT e l'Edge Computing, se da un lato hanno permesso un maggior efficientamento dei processi, dall'altro hanno anche notevolmente ampliato la superficie di attacco alla quale sono esposte le reti e le infrastrutture con riferimento a minacce sempre più sofisticate e a malintenzionati sempre più esperti e abili.

Inoltre, numerose sono le problematiche legate alle potenziali vulnerabilità e bias presenti o inseriti malevolmente in fase di apprendimento delle reti neurali insite nei modelli di AI: queste soluzioni infatti risultano essere delle armi a doppio taglio, in quanto se da un lato possono certamente considerarsi indispensabili per l'efficientamento dei processi di calcolo, efficienza ed affidabilità degli output prodotti, dall'altro lato possono creare rischi considerevoli per gli interessati, specialmente con riferimento ai modelli implementati all'interno di sistemi di videosorveglianza o necessari ad altre attività di trattamento particolarmente invasive.

Questo capitolo tratterà alcune delle principali vulnerabilità e minacce cibernetiche associate alle tecnologie di cui sopra, con particolare attenzione agli impatti che possono avere non solo sulla sicurezza informatica, ma anche sulle persone fisiche.

L'introduzione di tali tecnologie implica, in risposta alla continua evoluzione di queste minacce, una ridefinizione continua delle strategie di difesa poiché le soluzioni di sicurezza tradizionali risultano essere sempre più inadeguate a proteggere un ambiente dotato di una dinamicità e distribuzione elevata.

Passando dall'esame di queste minacce e vulnerabilità, il capitolo si svilupperà esaminando le misure di sicurezza volte a mitigare i rischi legati all'adozione di tali soluzioni, focalizzandosi sulle best practice del settore e sulle soluzioni che possono contribuire a costruire un ecosistema più sicuro e resiliente in un mondo sempre più interconnesso ed automatizzato, specialmente alla luce di impieghi massicci come quelli previsti all'interno delle Smart Cities.

III.1. Internet of Things (IoT)

Con riferimento all'insieme dei dispositivi IoT, vi sono diversi livelli operazionali che devono essere preservati mediante l'adozione di apposite misure di sicurezza che vadano a prevenire l'efficacia di un attacco informatico o ne vadano a mitigare gli effetti.

Può giovare in quest'ottica partire da una tassonomia dei principali requisiti di sicurezza che, secondo le best practice del settore e gli standard definiti nelle linee guida della ISO 27402, devono essere presenti in un sistema IoT.

A livello di gestione dei dati e delle informazioni, la sicurezza del sistema IoT dovrebbe garantire i seguenti requisiti:

- Confidenzialità, la quale implica che i dati non possano essere letti da terzi non autorizzati, in modo da instaurare un rapporto di fiducia tra i dispositivi IoT comunicanti per lo scambio di informazioni che siano sicure;
- Integrità, che consiste nel fatto che i dati ricevuti non devono essere in alcun modo alterati durante la trasmissione;
- Disponibilità, cioè la garanzia che siano implementate misure atte ad evitare o mitigare il rischio della perdita dei dati in fase di trasmissione;
- Protezione dei dati personali, che indica la necessità che l'identità dei soggetti che sono la fonte dei dati dovrebbe rimanere ignota a terzi (che raggruppa tecniche quali, a titolo esemplificativo, la pseudonimizzazione, l'anonimizzazione, k-anonymity, differential privacy, ecc.);

Per garantire tali requisiti, è necessario stabilire un insieme di meccanismi di sicurezza volti a controllare innanzitutto l'accesso alla rete: ad esempio, l'implementazione di un adeguato sistema di controllo degli accessi permetterebbe di garantire che solo gli utenti legittimi possano accedere ai dispositivi e alla rete per attività amministrative, quali la riprogrammazione o il controllo remoto dei dispositivi IoT e della rete.

Il controllo degli accessi andrebbe integrato con meccanismi robusti di autenticazione, i quali consentono che vi sia un'attenta verifica che un determinato dispositivo sia in possesso dei diritti di accesso ad una rete e, viceversa, che una rete abbia il diritto di connettersi al dispositivo.

Va tenuto a mente che i dispositivi IoT devono fornire procedure di autenticazione avanzate per evitare eventuali minacce alla propria sicurezza: infatti, se tutti i dispositivi IoT prodotti dallo stesso vendor sono configurati con le stesse credenziali di autenticazione, l'hacking di uno solo

di questi comprometterebbe la sicurezza dell'intera rete di dispositivi, portando a conseguenze catastrofiche.[33]

A livello funzionale risulta di vitale importanza che i requisiti di sicurezza stabiliti si fondino su due principi fondamentali: il primo è costituito dalla resilienza del sistema, la quale si riferisce alla capacità di quest'ultimo di essere in grado di garantire la sicurezza per tutti i dispositivi connessi alla rete, specialmente quando questi siano sotto attacco o colpiti da guasti, bug o malfunzionamenti.

Il secondo criterio è invece quello di "self organization": questo principio si riferisce alla capacità di un sistema IoT di gestirsi autonomamente e regolarsi per rimanere operativo anche in caso di guasto di alcune parti a causa di occasionali malfunzionamenti o attacchi da parte di malintenzionati.

Proprio con riferimento a questi attacchi, esaminiamo ora le principali tipologie di attacco partendo dal relativo layer del sistema del quale vanno a sfruttare le vulnerabilità.

Generalmente, l'architettura della comunicazione di un sistema di dispositivi IoT può essere suddivisa in tre layer principali:

- l'Edge layer, il quale si occupa prevalentemente di fornire funzionalità del layer PHY (Physical layer, il mezzo fisico di trasmissione in sé, quale l'ethernet) e funzionalità MAC (necessario a trasferire i pacchetti dalla rete al mezzo fisico di trasmissione) per le comunicazioni locali;
- l'Access layer, che permette la connessione verso l'esterno dei dispositivi, di solito mediante l'impiego di un Gateway (punto di accesso ad una rete esterna) ed uno strato di Middleware che fungono da intermediari tra il sistema IoT ed Internet.
- l'Application layer, che si occupa della gestione e indirizzamento della comunicazione dei dati in modo da fornire servizi di rete direttamente agli end users e gli applicativi mediante i quali viene gestito il sistema IoT.[34]

Con riferimento all'Edge layer, una delle principali minacce è costituita dagli attacchi cosiddetti "side channel": il principale scopo di questi attacchi è quello di recuperare determinate informazioni dall'analisi dei segnali laterali dell'infrastruttura sulla quale poggia il sistema (come il consumo energetico, i tempi di comunicazione e le emissioni elettromagnetiche) mentre i vari nodi eseguono le procedure di crittografia dei dati.

[33] V. F. Meneghello, M. Calore, D. Zuccheto, M. Polese e A. Zanella, "IoT: Internet of Threats? A survey of practical security vulnerabilities in real IoT devices.", *IEEE Internet of Things Journal*, Vol.6, 2019.

[34] V. C. Pielli, D. Zucchetto, A. Zanella, L. Vangelista, e M. Zorzi, "*Platforms and protocols for the Internet of Things*,", *EAI Endorsed Transactions on Internet of Things*, vol. 15, n. 1, 2015.

Il consumo energetico dei dispositivi, ad esempio, è largamente impiegato dagli attaccanti per riuscire ad indovinare e risalire alle chiavi segrete crittografiche impiegate dal sistema, in quanto ad ogni operazione di crittografia eseguita da quest'ultimo è possibile risalire ad una specifica traccia di alimentazione: infatti, i dati di potenza sono generalmente calcolati dalla differenza di voltaggio tra un resistore inserito in serie con la relativa fonte di alimentazione.

Gli attacchi side channel possono variare da attacchi di semplice analisi dell'alimentazione (nei quali l'attaccante tenta di interpretare direttamente le tracce di alimentazione relative ad un numero ristretto di cicli di crittografia), oppure essere attacchi di analisi della potenza differenziale (approccio ben più efficace e avanzato, in quanto una maggiore quantità di tracce viene analizzata statisticamente al fine di estrarre ulteriori informazioni relative ai cicli crittografici).

Altri attacchi che colpiscono questo livello perimetrale dei dispositivi IoT sono i Trojan che vanno a colpire direttamente la componente hardware del sistema (c.d. "Hardware Trojan") e gli attacchi di tipo Denial of Service (DoS) i quali causano un malfunzionamento del sistema dovuto ad un esaurimento delle risorse di un sistema informatico che fornisce un servizio ai client (ad esempio un sito web su un server), fino a impedire l'erogazione del servizio agli utenti che lo richiedono.

Anche i pacchetti trasferiti dal dispositivo IoT possono essere manomessi per comprendere i protocolli crittografici impiegati: alcuni metodi possono essere quelli di modificarne la componente software mascherando un nodo dannoso in modo che il sistema lo riconosca erroneamente quale un nodo legacy (c.d. "camuffamento"), oppure tentare il reverse engineering per comprendere meglio i dettagli dei protocolli di comunicazione impiegati al fine di estrarre eventuali informazioni riservate (come algoritmi coperti da brevetti).

Passando all'Access layer i principali attacchi condotti da soggetti malintenzionati sono quelli volti all'intercettazione o manomissione del contenuto dei pacchetti che passano dal middleware, come il cosiddetto "sniffing" dei pacchetti (dove si intercettano i dati in transito su di una rete) e gli attacchi di Injection (costituiti dall'"iniezione" da parte dell'attaccante di stringhe di codice malevolo o comandi dannosi per il sistema).

Un altro tipo di attacco che va preso in considerazione in questo layer è quello di "routing" dei pacchetti, mediante il quale gli attaccanti possono ridirezionare i pacchetti mandandoli ad un destinatario da loro definito o generare errori di comunicazione che portino ad una perdita dei pacchetti (c.d. "dropping") in modo che non raggiungano mai il destinatario originale.

Con riferimento infine all'Application layer, gli attacchi che vanno a colpire questo strato della comunicazione risultano essere piuttosto diversi da quelli che abbiamo potuto osservare nei

precedenti layer: infatti, per colpire questo strato è necessario che l'attaccante si rivolga direttamente al software in fase di esecuzione sui dispositivi piuttosto che durante la comunicazione di questi con il sistema.

Alcuni attacchi, come vedremo meglio parlando di quelli che mirano a colpire la componente AI dei sistemi IoT, possono riguardare l'integrità stessa dei dati raccolti ed elaborati, andando ad esempio a colpire gli algoritmi sottostanti all'apprendimento automatico del sistema, in modo da andare a manipolare il processo di addestramento stesso del modello al fine di indurre in errore il sistema e fargli compiere elaborazioni incorrette o restituire risultati del tutto fuorvianti. Possono verificarsi anche attacchi per ottenere l'accesso indebito al sistema concentrandosi sulla fase di autenticazione.[35]

Indipendentemente dal layer di riferimento, esistono una serie di attacchi che mirano a sfruttare il sistema IoT per intenti malevoli o sabotarlo per ridurne o manometterne le funzionalità.

Ad esempio, alcuni attacchi possono essere mirati a creare un disservizio per gli utenti creando dei malfunzionamenti o blocchi delle funzionalità dei dispositivi, come smart TV o elettrodomestici smart, magari mediante un ransomware volto a limitare l'impiego dell'apparecchio fino al pagamento di un riscatto.

Altri attacchi vanno invece ad impiegare le funzionalità del dispositivo IoT per raggiungere scopi completamente diversi da quelli originariamente pensati, come il violare un sensore di presenza parte di un sistema di allarme o delle telecamere smart per monitorare a distanza e conoscere in ogni momento la posizione delle vittime nel loro ambiente di vita, anche quando questo sistema di allarme sia apparentemente disattivato.

Infine, vi sono una serie di attacchi che vanno del tutto ad ignorare le funzionalità del sistema IoT, concentrandosi invece sullo sfruttare la capacità di tale insieme di dispositivi di connettersi alle reti Locali (c.d. LAN, Local Area Network) e ad Internet: attacchi di questo tipo possono essere, ad esempio, quelli che mirano a trasformare i dispositivi IoT in una componente di una "bot-net" (ovvero una rete di dispositivi, definiti bot, controllata interamente da remoto dall'attaccante), o ancora attacchi volti a violare la rete domestica degli interessati per violare a loro volta altri dispositivi della vittima, quali PC, laptop o smartphone collegati alla rete: quest'ultima tipologia di attacchi risulta particolarmente pericolosa in quanto permette agli attaccanti di sfruttare il sistema per lanciare ulteriori attacchi, come nel caso delle bot-net, le quali, una volta create, possono essere impiegate per lanciare un ulteriore attacco ad un altro

[35] V. F. Meneghello, M. Calore, D. Zuccheto, M. Polese e A. Zanella, "IoT: Internet of Threats? A survey of practical security vulnerabilities in real IoT devices.", *IEEE Internet of Things Journal*, Vol.6, 2019.

obiettivo, come un attacco Denial of Service, facendo connettere simultaneamente alla rete del target tutti i dispositivi bot dietro un comando dell'hacker.

Per rendere sicuri i sistemi IoT da questa serie di attacchi, possono essere previste diverse tipologie di misure di sicurezza volte a irrobustire tali soluzioni e a renderle più resilienti ad eventuali tentativi di attacco da parte di soggetti malintenzionati.

Una prima misura imprescindibile (anche per sistemi non IoT) risulta essere l'impiego di protocolli crittografici standard che rendano sicura non solo la conservazione dei dati quando questi si trovano *at rest* all'interno dei dispositivi (come AES 256), ma anche protocolli che vadano a proteggere i dati nella fase di trasferimento durante la comunicazione con altre componenti del sistema (come protocolli HTTPS over TLS 1.2 o superiori se la comunicazione avviene mediante Internet).

Esiste una serie di protocolli di cifratura sviluppati recentemente che sono definiti "lightweight" e che possono essere di estrema valenza in quanto non solo sfruttano nuovi blocchi e flussi di crittografia, ma anche nuovi cifrari combinati con codici di autenticazione dei messaggi e funzioni di hash, permettendo inoltre di essere eseguiti dai dispositivi impiegando limitate risorse di calcolo, comunicazione e archiviazione.

Per quanto l'applicazione di questi protocolli sia abbastanza recente, l'idea sottostante circolava già da alcuni anni nel settore dell'IoT, tanto che nel 2012 fu pubblicato lo standard ISO/IEC 29192, che andava proprio a specificare la struttura ed il funzionamento di alcuni di questi protocolli, quali CLEIFA e PRESENT.[36]

Un'altra misura consigliata è l'adozione di soluzioni di Intrusion Detection System (IDS) e Intrusion Prevetion System (IPS) per la prevenzione e il rilevamento delle intrusioni che, ad esempio, mediante il monitoraggio di anomalie nei parametri di sistema, come un utilizzo eccessivo della Central Processing Unit (CPU), della memoria di lavoro o del throughput della rete possono permettere di indicare la presenza di un attacco in corso o di un intruso all'interno del sistema.[37]

Particolarmente utile per evitare attacchi veicolati dalle manipolazioni dei dispositivi IoT stessi in termini di hardware (questi possono essere infatti installati in aree remote con scarse se non assenti misure di sicurezza fisica) sono le cosiddette PUF, o Physical Unclonable Functions: il concetto alla base delle PUF consiste nell'impiegare le piccole differenze introdotte durante il processo di fabbricazione dei chipset interni ai dispositivi per generare una firma univoca per ciascuno di essi.

---

[36] V. ISO/IEC 29192-2012, "*Information technology – Security techniques – Lightweight cryptography*".

[37] V. F. Li, A. Shinde, Y. Shi, J. Ye, X. Li, e W. Z. Song, "System statistics learning-based IoT security: Feasibility and suitability.", *IEEE Internet of Things Journal*, 2019.

Un circuito PUF fornisce una risposta univoca ad un dato input e, a causa delle differenze hardware intrinseche ai dispositivi, fornisce specifiche per ogni chipset: prendendo ad esempio una PUF "Arbiter", questa prevede che il circuito sia composto da due percorsi solo apparentemente identici, ma che, per ogni input, fa dipendere l'output dal tempo impiegato dall'input per viaggiare all'interno del chipset, prendendo come riferimento quello che tra i due percorsi risulta il più veloce.[38]

Infine, un'importante misura di sicurezza organizzativa suggerita dall'European Network and Information Security Agency (ENISA) risulta essere la strutturazione di un attento controllo dei processi di update dei dispositivi: il corretto aggiornamento dei dispositivi IoT per risolvere bug e vulnerabilità insite nel sistema risulta essere un compito complesso per via del fatto che i prodotti stessi si basano solitamente su vari pacchetti provenienti da sviluppatori e vendor diversi, i quali impiegano a loro volta diversi strumenti e componenti di terze parti.[39]

La pianificazione e la gestione di questi aggiornamenti è imprescindibile per garantire la sicurezza di questi dispositivi.

III.2. Intelligenza Artificiale (AI)

Con la crescente integrazione dei sistemi IoT con i modelli di AI, sempre più minacce specifiche e attacchi vengono ideati da parte di malintenzionati e criminali al fine di sfruttare le vulnerabilità insite in questi sistemi.

L'uso o sfruttamento improprio di questi sistemi da parte di un attaccante comporta rischi significativi non solo con riferimento al sistema informatico in sé considerato, ma anche per le persone fisiche che nella vita di tutti i giorni si trovano ad interagire con tali tecnologie innovative.

Una prima minaccia è costituita dagli "Adversarial Attacks", dove gli attaccanti generano input malevoli o manomessi medianti modifiche ai dati difficilmente percettibili che possono invece indurre i sistemi AI a compiere decisioni o strutturare previsioni del tutto errate: ad esempio, gli aggressori possono manipolare leggermente un'immagine per ingannare un sistema di riconoscimento facciale basato sull'intelligenza artificiale al fine di indurlo a identificare erroneamente le persone o errare del tutto a riconoscere un essere umano quale tale.

Un'altra minaccia è costituita dalla possibilità che un attaccante a conoscenza del funzionamento del modello di AI possa decidere di sfruttare i Bias insiti allo stesso o generarne mediante tecniche di manipolazione.

---

[38] V. M. Roel, "Physically unclonable functions: Constructions, properties and applications.", *Springer*, 2012.
[39] V. ENISA, "Guidelines for securing the Internet of Things: Security supply chain", 2020.

I sistemi di intelligenza artificiale addestrati su dati distorti o manipolati infatti tendono a produrre risultati errati o del tutto discostati da quelli attesi.

Tale problema è particolarmente rilevante nel caso dei Large Language Models (LLM), dove si manifesta negli attacchi di data poisoning, prompt injection (inserimento di prompt malevoli) e manipulative framing (manipolazione del modello scegliendo particolari combinazioni di parole).

Una nuova minaccia nasce dagli attacchi informatici basati essi stessi sull'impiego dell'intelligenza artificiale a supporto del malware principale: questi sistemi di intelligenza artificiale dannosi possono automatizzare gli attacchi informatici identificando e sfruttando le vulnerabilità dei sistemi target a una velocità senza precedenti e con una precisione elevata, oltre che evolversi e adattarsi di bersaglio in bersaglio aggiornando le proprie tattiche in tempo reale, eludendo le difese tradizionali e amplificando la portata del loro impatto, sfruttando ogni angolo della superficie d'attacco sfruttabile nel tentativo di violare il sistema.

La combinazione tra i modelli di AI e i dispositivi IoT crea sistemi potenti ed estremamente efficienti, ma introduce anche una serie di vulnerabilità ibride particolarmente insidiose che gli aggressori più attenti possono cercare di sfruttare a proprio favore: ad esempio, i sistemi IoT basati sull'intelligenza artificiale, quali i veicoli intelligenti, i droni e i robot industriali, sono suscettibili ad attacchi di tipo "hijacking", cioè di dirottamento, nei quali l'aggressore che ottiene il controllo del sistema può comandarlo da remoto al fine di causare danni fisici, interrompere operazioni critiche o accedere ad informazioni sensibili. [40]

Consideriamo il caso di un'automobile a guida autonoma che viene colpita da un attacco hacker: il veicolo target potrebbe essere impiegato dall'attaccante per mettere in grave pericolo i passeggeri o altre persone, ad esempio facendola schiantare contro altri veicoli, ostacoli o gruppi di pedoni.

Allo stesso modo, potrebbero essere violati sistemi di telecamere intelligenti al fine di non permettergli di riconoscere determinate minacce o problematiche, portando a gravi conseguenze non solo in termini di operatività, ma anche di pubblica sicurezza.

Inoltre, tramite gli attacchi di data poisoning, l'iniezione di dati dannosi o fuorvianti all'interno delle reti IoT e in particolare della loro componente AI possono portare ad una corruzione dei dataset di addestramento del modello, causando di riflesso la completa compromissione dell'intero processo decisionale.

[40] V. A. McCall, 2024, "Cybersecurity in the Age of AI and IoT: Emerging Threats and Defense Strategies", *Research Gate*, https://www.researchgate.net/publication/386050391_Cybersecurity_in_the_Age_of_AI_and_IoT_Emerging_Threats_and_Defense_Strategies.

All'interno delle Smart City e delle infrastrutture critiche l'integrazione tra AI e il mondo IoT genera nelle città intelligenti un generale miglioramento dell'efficienza operativa e organizzativa della vita quotidiana in vari processi, ma rende allo stesso tempo questi servizi critici altamente suscettibili alle minacce informatiche, rendendoli potenzialmente proprio l'anello debole del corretto funzionamento di queste città intelligenti, costituendo anche un pericolo per le persone che vivono al loro interno.

Pensiamo alle difficoltà e disagi creati da un'improvvisa interruzione dei servizi essenziali, un malfunzionamento causato da un attacco informatico ad un sistema IoT che gestisce i semafori, o ancora a dei disservizi e interruzioni relativi alle reti energetiche o di approvvigionamento idrico.

Un attacco informatico ben assestato ad uno di questi sistemi potrebbe causare ingorghi nel traffico, diffuse interruzioni della distribuzione di corrente agli edifici o contaminazione dell'acqua per malfunzionamenti legati al processo di depurazione, che comportano gravi rischi per la sicurezza e stabilità economica delle Smart Cities.

Oramai le infrastrutture alla base delle Smart Cities si affidano sempre più all'intelligenza artificiale per compiere i processi decisionali sottostanti al sistema, ed un avvelenamento dei dati che vada a creare contraddizioni o risultati errati potrebbe generare malfunzionamenti critici del sistema, come segnali stradali errati nel caso dei semafori che portino ad incidenti o, come detto precedentemente, una serie di interruzioni dei servizi essenziali.

Per mitigare o evitare gli effetti di questi attacchi, i modelli di AI devono essere adeguatamente protetti da misure di sicurezza tecniche ed organizzative che vadano a far parte della struttura del sistema stesso fin dalla sua progettazione.

Una prima misura indispensabile risulta essere l'esecuzione del cosiddetto "Adversarial Training", cioè una strategia di addestramento del modello che si basi proprio sull'impiego di input malevoli durante il processo di apprendimento in modo da rendere il sistema più resiliente e performante nell'individuare questa tipologia di input in modo da neutralizzarne o mitigarne gli effetti una volta che questi si manifestino durante l'impiego operativo non appena il sistema viene rilasciato in ambiente di produzione.[41]

Importantissima risulta la comprensione e il monitoraggio della capacità di ricostruzione delle decisioni compiute dai modelli di AI: la rete neurale, in fase di training, apprende dai dati che gli

[41] V. A. McCall, 2024, "Cybersecurity in the Age of AI and IoT: Emerging Threats and Defense Strategies", *Research Gate*, https://www.researchgate.net/publication/386050391_Cybersecurity_in_the_Age_of_AI_and_IoT_Emerging_Threats_and_Defense_Strategies.

vengono forniti la creazione di un modello di Deep Learning volto a risolvere il problema assegnato, massimizzando la performance e la velocità di elaborazione.

Tuttavia, la complessità della struttura di queste reti neurali porta spesso ad una sempre più complessa ricostruzione delle decisioni compiute dal sistema nel produrre un determinato output, portando ad un'opacità delle reti di DL e a potenziali rischi di bias o vulnerabilità insite nelle stesse ma di difficile individuazione.

Risulta indispensabile, per evitare il proliferare di vulnerabilità, bias o bug incontrollati, che gli sviluppatori dell'algoritmo sottostante al modello siano in grado di effettuare un monitoraggio continuato del sistema di AI per avere sotto controllo l'intero processo decisionale, in modo non solo da garantirne la trasparenza, ma anche per evitare l'insorgere di vulnerabilità "Zero Day" , cioè vulnerabilità ignote persino allo sviluppatore del modello, che un eventuale attaccante potrebbe rilevare e sfruttare a proprio vantaggio.

Per attuare questa misura risulta imprescindibile adeguarsi ad un principio cardine dello sviluppo e monitoraggio di un sistema di AI, cioè quello dello "Human in the Loop": tale approccio consiste nella supervisione del modello durante tutto il suo ciclo di vita e nel monitoraggio continuato dello stesso anche una volta che questo viene impiegato attivamente, oltre che ad un processo di decision making condiviso che permetta all'uomo di poter intervenire nel processo decisionale del modello producendo scelte che possano andare a modificare l'output finale laddove questo potesse essere inizialmente errato, consentendo una collaborazione uomo-macchina indispensabile per evitare i rischi e le criticità legati alla piena automazione dei processi.

Infine, per contrastare eventuali attacchi condotti con malware accompagnati dall'uso di AI, risulterebbe utile anche l'impiego di sistemi di IDS e IPS a loro volta basati sull'impiego di un modello di AI, in modo da poter rilevare ed andare a contrastare efficacemente gli attacchi adeguando le risposte ed interventi in tempo reale seguendo e adattandosi come fa tale tipologia avanzata di malware.

III.3. Edge Computing

Secondo il report di Red Hat intitolato "Global Tech Outlook" del 2022, il 28% delle aziende oggetto dello studio avevano indicato l'adozione dei paradigmi di Edge Computing come una tecnologia emergente sulla quale avrebbero investito entro l'anno stesso, non solo per beneficiare dell'alta flessibilità offerta dall'impiego di un sistema ibrido di cloud computing, ma anche per fornire servizi più veloci e affidabili a costi totali notevolmente inferiori.[42]

[42] V. Red Hat, "2022 Global Tech Outlook: A Red Hat report", https://www.redhat.com/en/resources/global-tech-outlook-overview-2022.

Tuttavia, l'impiego di soluzioni tecnologiche di questa tipologia può risultare particolarmente complesso sia in fase di implementazione che di gestione, e anche laddove l'integrazione con il sistema che si vuole andare a migliorare risultasse un successo, può risultare difficile riuscire a mantenere sicura da eventuali attaccanti non solo la soluzione di Edge Computing in sé, ma anche l'intera infrastruttura che ne prevede l'impiego: ciò è dovuto al fatto che, se da una parte si ottiene un notevole vantaggio in termini di costi e qualità dei processi che queste tecnologie mirano ad automatizzare ed efficientare, dall'altra vi è un'inevitabile incremento della superficie d'attacco disponibile per gli avversari intenzionati a violare o compromettere l'integrità di tali sistemi.

Una particolare minaccia è posta, specialmente se pensiamo ad infrastrutture critiche e alle Smart Cities, dagli attacchi ATP (Advanced Persistent Threat), cioè minacce portate avanti da collettivi di attaccanti dotati di notevoli expertise tecniche e vaste risorse economiche, in grado di effettuare attacchi su vasta scala mediante l'impiego di molteplici vettori per periodi di tempo talvolta molto estesi.[43]

Un esempio è costituito dal collettivo cinese "Evasive Panda", il quale ha sviluppato specificatamente per i dispositivi Edge una backdoor nel protocollo crittografico Secure Shell (SSH) avanzata, la quale permette all'attaccante di ottenere il completo accesso al dispositivo compromesso: l'attacco in sé si basa su di una injection di codice malevolo all'interno del daemon SSH (una componente vitale per un sistema basato su Linux, che si occupa di gestire le connessioni sicure dei dispositivi ai server), il quale, successivamente all'aver ottenuto l'accesso al sistema, va a verificare se il dispositivo è già stato infettato e se è in possesso di privilegi root. Se tutte le condizioni sono soddisfatte, il malware permette agli attaccanti di ottenere il controllo totale sul sistema, consentendogli di compiere operazioni quali caricare, modificare ed eliminare file sui dispositivi compromessi, intercettare le comunicazioni che passano dai dispositivi ed eseguire comandi da remoto con privilegi elevati, producendo effetti devastanti se calato su potenziali ambienti urbani in cui questi sistemi di Edge Computing potrebbero essere integrati con componenti IoT e modelli di AI.[44]

Per scongiurare queste minacce, è necessario irrobustire la sicurezza delle reti mediante le quali questi sistemi di dispositivi comunicano, adottando, come per i dispositivi IoT, sistemi atti ad

[43] V. F. Ferrazza, 2019, "Minacce APT: cosa sono le Advanced Persistent Threat, come funzionano e come difendersi", Malware e Attacchi Hacker, https://www.cybersecurity360.it/nuove-minacce/minacce-apt-cosa-sono-le-advanced-persistent-threat-come-funzionano-e-come-difendersi/.

[44] V. L. Varriale, 2025, "CISA: nuove linee guida per i dispositivi edge mentre APT cinesi utilizzano backdoor SSH", *Sicurezza Informatica*, https://www.matricedigitale.it/sicurezza-informatica/cisa-nuove-linee-guida-per-i-dispositivi-edge-mentre-apt-cinesi-utilizzano-backdoor-ssh/.

individuare e prevenire potenziali attacchi, nonché prevedere delle politiche di controllo degli accessi e delle autorizzazioni efficaci, seguendo un approccio "Zero Trust".

Indispensabile è inoltre un'attenta progettazione di procedure di Patch management e di aggiornamento dei dispositivi di rete, in modo da evitare l'impiego di software o firmware obsoleti e scongiurare l'insorgere di vulnerabilità sfruttabili dagli attaccanti.

IV. AIoT: un caso di studio reale

Nel panorama dei concetti affrontati nei capitoli precedenti reputo indispensabile, al fine di dare uno stampo più pratico fornendo un esempio concreto di quanto finora discusso, riportare un caso nel quale mi sono imbattuto, durante la stesura di documenti o nella partecipazione agli incontri relativi alle trattative tra Titolare e fornitore, durante lo svolgimento del mio tirocinio curriculare.

Il caso è costituito da un insieme di soluzioni adottate da una struttura soggetta ad un elevato transito di persone in quanto area pubblica (di seguito "Struttura" o "Titolare"), al fine di accelerare od efficientare i diversi processi necessari alla gestione della stessa e della sua corretta operatività.

Mediante questo caso sarà possibile osservare non solo la concreta applicazione dei dispositivi IoT, ma anche le diverse integrazioni delle quali sono oggetto, sia con l'intelligenza artificiale che con diverse altre tecnologie, ognuna delle quali apporta un proprio essenziale contributo alla sicurezza, solidità ed efficacia del sistema di cui fanno parte.

Entrando maggiormente nel dettaglio del caso, questo è costituito da una soluzione di video-analytics adottata dalla Struttura, su fornitura della società X, per la misurazione dei livelli di occupazione delle aree ad alta affluenza, in maniera tale da poter gestire al meglio i flussi sia dei veicoli che delle persone, consentendo così di rilevare e risolvere prontamente eventuali criticità che possano causare disagi verso gli stessi.

Il Titolare, valutato il rischio per i diritti e le libertà degli interessati costituito dal trattamento (considerato "Alto" seguendo la metodologia proposta dal framework ENISA), ha previsto la stesura di una DPIA, alla cui redazione ho personalmente contribuito concentrandomi specialmente sugli aspetti relativi alla Cyber Security: questo capitolo vuole pertanto offrire una visione di quanto emerso dall'analisi condotta dal team DPO, offrendo inoltre degli spunti non solo sull'attenzione posta dal Titolare all'implementazione della soluzione in ottica Data Protection by Design, ma anche Security by Design, con un particolare riferimento alle misure di sicurezza tecniche ed organizzative adottate e a quelle che si potrebbero ulteriormente adottare.

IV.1 L'architettura della soluzione

L'architettura della soluzione prevede l'impiego di una rete distribuita di sensori audio e video, i quali vengono installati direttamente in prossimità delle aree da monitorare: a supporto di questi sensori, il fornitore ha messo a disposizione del personale della Struttura preposto al

monitoraggio delle aree sopra citate un'applicazione software per la configurazione dei sensori, costituita da un'interfaccia grafica (GUI) in grado di consentire agli operatori di impostare i parametri di funzionamento dei sensori (tra i quali la calibrazione, la messa a punto dei parametri di analisi della scena e la configurazione dei dispositivi stessi) ed un software di gestione centralizzato per il controllo e la telegestione del sistema, che permette l'acquisizione e la visualizzazione dei dati elaborati dalla rete di sensori, l'esportazione dei dati da essi raccolti e la gestione degli eventi di allerta generati dal sistema stesso.

La soluzione prevede inoltre l'adozione di un'applicazione mobile progettata per monitorare e ricevere istantaneamente allarmi di sicurezza sui dispositivi dei dipendenti del Titolare preposti all'attività di monitoraggio del corretto andamento dell'infrastruttura.

Infine, nonostante la possibilità prevista dal fornitore di impostare un layer di comunicazione basato su canali wireless (3G/LTE e/o WiFi), per una maggiore sicurezza, il Titolare ha preferito che la comunicazione delle informazioni raccolte dal sistema nella rete locale fosse configurato mediante l'adozione di canali wired di tipo LAN IP: il canale LAN IP permette non solo di ottimizzare il carico computazionale e minimizzare il più possibile l'utilizzo della banda, ma anche di evitare potenziali azioni di tampering o sniffing in fase di trasmissione dei pacchetti nella rete locale della Struttura.

Per ogni intervento manutentivo o di accesso al software di configurazione, al fine di garantire anche in questo caso una più robusta postura di sicurezza del sistema, l'accesso da parte dei tecnici del fornitore alle impostazioni dei sensori può avvenire solo on premise, senza possibilità di connessione da remoto.

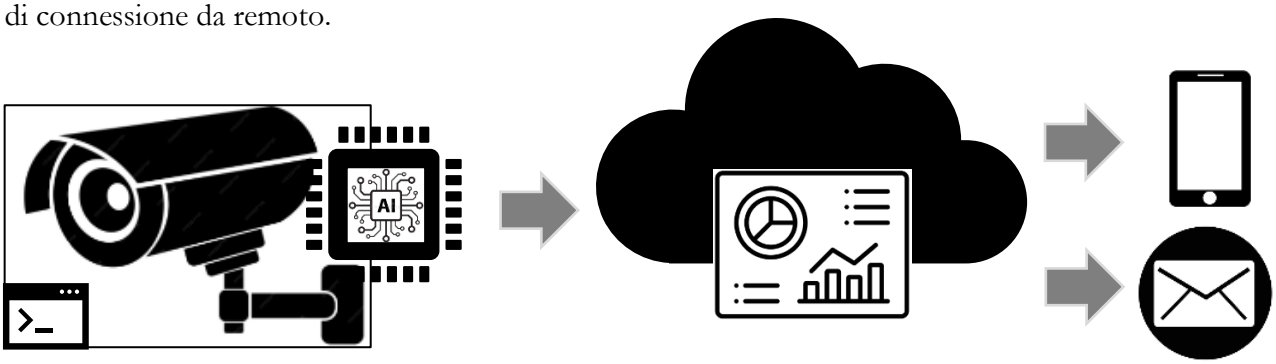

Figura 1 - Architettura della soluzione ricostruita dall'autore dell'elaborato: i sensori dotati di display dal quale può essere impiegato il software di configurazione passano le informazioni raccolte al dispositivo di Edge Computing posto di fianco ai sensori e ospitante il modello di AI. Una volta avvenuta l'elaborazione, le heatmap vengono mandate alla piattaforma gestionale situata in cloud, la quale invia eventualmente alert in caso di livelli elevati di occupazione degli spazi al personale della struttura (sull'app apposita o via mail).

IV.1.1. Le specifiche dei sensori

I sensori sviluppati e prodotti dal fornitore sono dotati di una processing unit dedicata al calcolo di algoritmi di intelligenza artificiale basati sulla visione e sull'analisi delle immagini.

In abbinamento alla visualizzazione e registrazione video e audio, infatti, i sensori permettono di elaborare la rilevazione e la classificazione dell'oggetto o dell'evento di interesse, il tutto senza la necessità di appoggiarsi ad unità di calcolo estranee grazie all'implementazione di dispositivi di Edge Computing (visibili nella figura 1).

Entrando ulteriormente nel dettaglio, le specifiche dei sensori sviluppati dal fornitore ed impiegati dalla struttura sono le seguenti:

- Un angolo di visione di 360°;
- Una struttura composta dalla telecamera principale e fino a due camere satelliti orientabili indipendentemente l'una dall'altra;
- Un doppio flusso video interamente configurabile con audio integrato;
- Un'unità di elaborazione AI incorporata;
- Una registrazione video e audio con buffering circolare;
- Un sistema di visione notturna con interruttore infrarossi automatico;
- Un emettitore infrarossi integrato ad alta potenza (fino a 40 metri di distanza);
- Una lunghezza focale configurabile;
- Una connessione di rete via wireless o cablata;
- La conformità al protocollo Message Queuing Telemetry Transport (c.d. MQTT, impiegato per situazioni in cui è richiesto un basso impatto energetico e dove la banda è limitata) per il trasferimento dei dati raccolti dai sensori ai software citati facenti parte della soluzione.

Rispettivamente alla componente hardware, i video sensori sono dotati di piattaforme multi-core INTEL, le quali sfruttano le capacità computazionali aggiuntive dei processori specializzati di intelligenza artificiale TPU Coral al fine di delegare su di questi il carico di video inferenza, liberando così risorse sui processori adibiti ad uso generale.

Il firmware sviluppato dal fornitore adotta best practices e paradigmi di programmazione altamente ottimizzati al fine di sfruttare al meglio tutte le risorse hardware disponibili: questi paradigmi includono la parallelizzazione del flusso logico degli algoritmi in job separati per diversi processori, al fine di ridurre il carico computazionale complessivo degli stessi, aumentandone il throughput.

Tale ottimizzazione garantisce, inoltre, la minimizzazione della banda di trasmissione, attraverso la gestione di formati dei dati compressi, la cui codifica e decodifica non è gestita dalla CPU stessa, ma è invece delegata a routine interne della GPU, consentendo così ai sensori di elaborare flussi video ad alta risoluzione, con frame di rete elevati e con bassa latenza.

IV.1.2. L'algoritmo di Intelligenza Artificiale

Il fornitore della soluzione di video analytics è specializzato nello sviluppo di reti neurali convoluzionali, mirate alla classificazione, segmentazione, identificazione e riconoscimento di oggetti, pattern ed eventi che si verificano nella visuale dei sensori.

L'intero processo di ideazione, creazione, addestramento e testing della rete neurale implementata nella soluzione è realizzato da parte del fornitore stesso: la formazione del modello di machine learning prevede l'uso di un framework di deep learning sviluppato mediante l'impiego di librerie di codice open source quali Tensor Flow e Keras, accompagnato da un set di dati di training completo, stabilendo un processo di apprendimento supervisionato dell'algoritmo.

L'uso di tecniche di visione artificiale per l'elaborazione di immagini o video da parte dei sensori è molto oneroso in termini computazionali, e per questa ragione il fornitore ha ritenuto opportuno, per alcuni componenti, svilupparle da zero, mentre per altre reingegnerizzandole da soluzioni di altri vendor, in modo da poter garantire elevate prestazioni anche su piattaforme embedded.

L'unità di processing dedicata all'esecuzione del modello di intelligenza artificiale può inoltre essere collegata ad un sistema di videosorveglianza già presente nel perimetro della Struttura, permettendogli

di acquisire il flusso video direttamente dalle telecamere di videosorveglianza già installate, inserendosi così in parallelo al sistema complessivo di videosorveglianza: tale possibilità, tuttavia, non è stata presa in considerazione da parte del Titolare, che ha voluto limitare l'impiego della soluzione per la sola finalità di gestione delle affluenze nelle aree target.

Risulta necessaria una precisazione con riferimento al modello di AI impiegato dalla soluzione: questo viene esclusivamente utilizzato per convertire le immagini che riceve dai sensori video (input) in una heatmap (output), la qual viene poi inviata alla piattaforma gestionale in cloud.

Pertanto, la componente di AI funge esclusivamente da unità di elaborazione, rimanendo del tutto separata dal resto del sistema, non essendo in alcun modo interrogabile o oggetto di interazione diretta né da parte degli interessati, né da parte del personale della Struttura.

IV.1.3. Il paradigma di Edge Computing

Su indicazione del Titolare, il fornitore della soluzione, al fine di efficientare maggiormente l'elaborazione da parte del sistema delle immagini raccolte e per garantire i requisiti di minimizzazione dei dati carpiti dai sensori, ha previsto l'installazione a bordo dei dispositivi di

unità di calcolo contenenti il modello di AI impiegato per la conversione delle immagini in mappe di calore (heatmap) delle aree monitorate.
Grazie al modello di AI e all'acceleratore di calcolo TPU Coral, il sistema riesce ad elaborare le immagini in soli 50 millisecondi, non solo efficientando la velocità di elaborazione delle immagini, ma anche riducendo al minimo la conservazione di immagini che possano contenere dati personalizzati, trasformandoli velocemente in meri dati statistici.
Dall'adozione di questo paradigma di Edge Computing si evita così del tutto la trasmissione al cloud per l'elaborazione delle immagini, effettuandole direttamente a livello del dispositivo preposto alla raccolta delle informazioni, passando solo nella fase successiva all'invio dei dati alla piattaforma, la quale tuttavia non avrà al suo interno altre informazioni se non quelle relative alle statistiche complessive di occupazione degli spazi, del tutto anonimizzate.

IV.1.4. Il Dataflow della soluzione

A partire dalla descrizione più tecnica delle varie componenti sopra riportata, ritengo sia utile ora riassumere, per una maggiore chiarezza, quello che è il flusso dei dati raccolti dalla soluzione.
Le persone presenti nelle aree di interesse vengono riprese dal video sensore, il quale, una volta raccolte le immagini, le invia quasi istantaneamente al dispositivo di Edge Computing apposto in sua prossimità: tale dispositivo, grazie al suo modello di AI e all'acceleratore di calcolo TPU Coral, riesce ad elaborare le immagini in soli 50 millisecondi, trasformando le informazioni da un'immagine effettiva dell'area ad una mappa di calore (heatmap) della stessa.

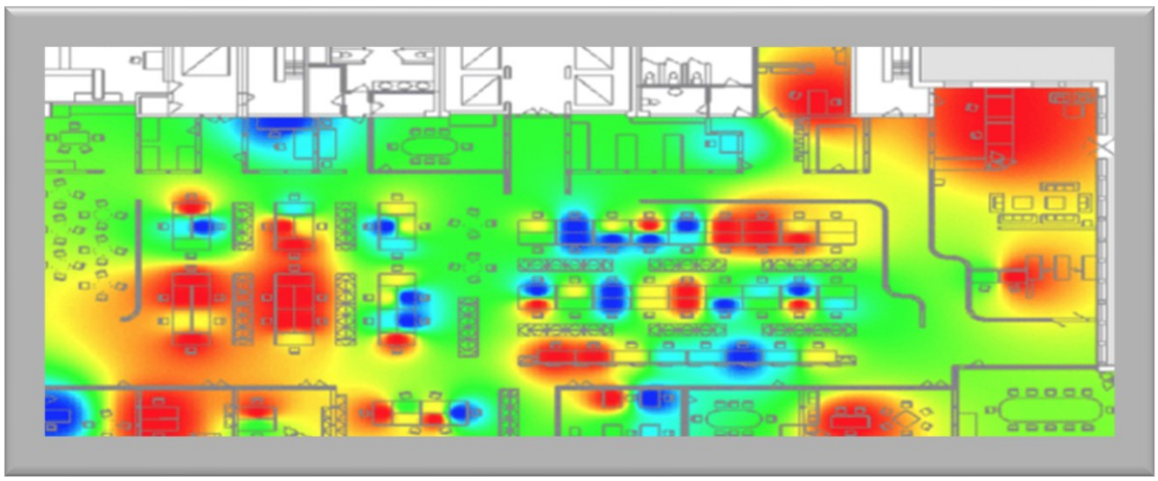

Figura 2 – Esempio di heatmap generata dal sistema: la scala di colore dal blu al rosso va ad indicare quella che è l'intensità della concentrazione di persone in un'area (dove il blu indica un afflusso minimo, mentre il rosso un elevato afflusso).[45]

[45] V. S. Ghayyur, X. He, D. Ghosh, S. Mehrotra, "Towards Accuracy Aware Minimally Invasive Monitoring (MiM)", ACM Conference on Computer and Communications Security (Theory and Practice of Differential Privacy), 2019

Questa mappa di calore viene inviata alla piattaforma gestionale passando dal canale sicuro costituito dal tunnel IPSec (progettato per ottenere connessioni sicure a livello di reti di IP, diversamente da quanto fatto da protocolli quali HTTPS) della Virtual Private Network (VPN), permettendo la protezione di queste informazioni da azioni di sniffing del traffico di rete.
Raggiunta la piattaforma gestionale, i dati sono visibili anche sottoforma di statistiche utili per comprendere al meglio l'utilizzo degli spazi.
In seguito all'elaborazione del contenuto video in tempo reale, la piattaforma genera delle allerte automatiche non appena avviene il superamento dei tempi di attesa definiti dai parametri della Struttura, oppure al superamento del numero massimo di persone in attesa che l'area di accodamento può contenere, senza intralciare in alcun modo gli altri servizi erogati.
Il sistema, passando dalla piattaforma gestionale, invierà al personale del Titolare le allerte (accompagnate da uno snapshot immagine dell'evento di rilevato) tramite messaggistica e-mail, in modo che questo possa intervenire prontamente per la risoluzione del problema.
I soggetti autorizzati riceveranno inoltre una notifica tramite l'apposita applicazione per dispositivi mobili impiegata dal personale della Struttura.
La comunicazione tra la piattaforma gestionale e l'app, sempre per proteggere tali informazioni da possibili intercettazioni, prevede l'impiego del protocollo Transport Layer Security 1.2 (TLS 1.2) per la cifratura della comunicazione end-to end (cioè tra la sorgente ed il destinatario).

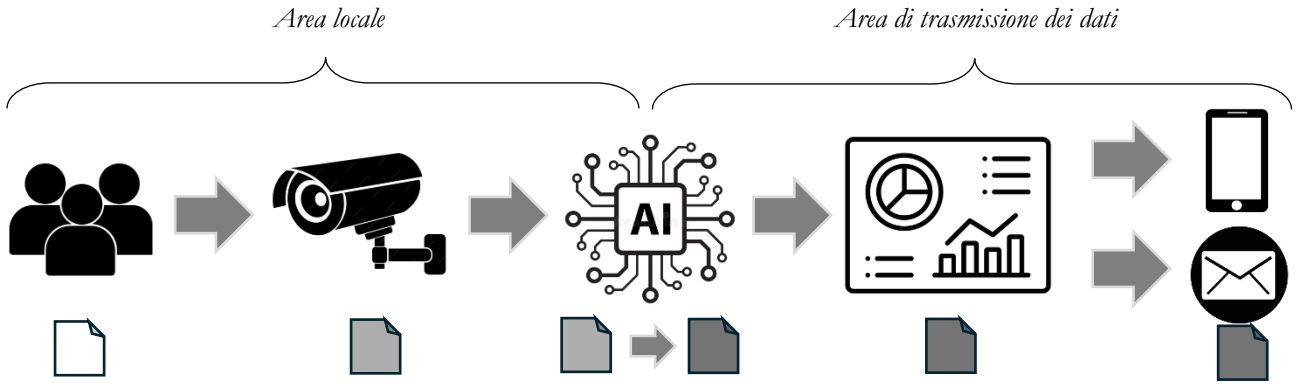

Figura 3 - Schema ricostruito dall'autore dell'elaborato del Dataflow della soluzione: in bianco è possibile vedere i dati biometrici dei volti delle persone, in grigio chiaro i dati personali (che per via della risoluzione non sono di natura biometrica), mentre in grigio scuro i dati anonimizzati.

IV.2 La conformità della soluzione alle previsioni del GDPR e dell'AI Act

Esaminiamo ora il grado di conformità alle disposizioni del GDPR della soluzione presa in esame.
I dati personali oggetto del trattamento preso in considerazione sono prevalentemente quelli relativi alle riprese video ed alle immagini raccolte dall'insieme di telecamere e sensori impiegati dalla soluzione. Insieme a questi, vengono inoltre raccolti anche i dati anagrafici e di contatto

dei dipendenti e dei collaboratori del Titolare e del fornitore: questi soggetti infatti, per funzioni di monitoraggio, configurazione e manutenzione, accedono al sistema, necessitando, ai fini di garantire l'integrità dello stesso e delle informazioni ivi contenute, una raccolta di informazioni utili a poter ricostruire le cause del disservizio in caso di malfunzionamenti inaspettati, risalendo al responsabile laddove questi eventi siano dovuti ad azioni interne malevole o abusive.

La base giuridica individuata dal Titolare quale condizione di liceità per i trattamenti effettuati è quella prevista all'articolo 6, punto 2 lettera f) del GDPR, cioè il legittimo interesse del Titolare: per mezzo della soluzione, infatti, il Titolare si propone di fornire un adeguato supporto agli utenti rilevando il numero delle persone in coda nelle aree della Struttura presso le quali, ad essere raccolti, non sono solo le immagini delle persone ivi presenti, ma anche, potenzialmente, le targhe dei veicoli in accodamento.[46]

Il Titolare intende analizzare i dati raccolti al fine di ottimizzare, nel loro complesso, i servizi offerti agli utenti, garantendo un'esperienza di viaggio fluida ed efficiente mediante una gestione attenta degli spazi sopra citati.

Il Titolare ha ritenuto la sussistenza del legittimo interesse quale base giuridica del trattamento in oggetto in quanto la finalità perseguita dall'installazione del sistema prevale sulle legittime aspettative di riservatezza degli individui, anche in considerazione delle misure tecniche che riducono il trattamento.

A supporto della sussistenza di tale base giuridica, il Titolare ha inoltre provveduto alla stesura di un Legitimate Interest Assessment (LIA), la quale prova il bilanciamento di interessi effettuato dalla Struttura prima dell'avvio delle operazioni di trattamento.

Il Titolare ha previsto la nomina, mediante apposito atto, dei soggetti autorizzati al trattamento dei dati, stabilendo ruoli e responsabilità legate al trattamento dei dati e garantendo le competenze e adeguati livelli di formazione del personale che impiega il sistema di Video Analytics, rispettando quanto sancito dall'articolo 4 del GDPR[47].

Essendo la risoluzione dei sensori video non idonea a poter ricostruire con dettaglio quelle che sono le caratteristiche dei volti delle persone riprese, non si applica al caso in esame l'articolo 9 del GDPR, in quanto, non venendo trattati dati biometrici, non risulta necessaria l'individuazione di un'ulteriore base giuridica a supporto del trattamento di queste categorie particolari di dati.

[46] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 6, punto 2, lettera f).
[47] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 4.

In questo caso, infatti, il sistema è stato configurato da parte degli operatori del Titolare in modo da raccogliere solo quelle immagini e video del tutto generali, le quali non possono essere ingrandite in alcun modo.

Inoltre, le immagini non vengono registrate se non a bordo dei dispositivi per un brevissimo lasso di tempo, e successivamente alla loro elaborazione i video sensori sono impostati in modo tale da cancellare le immagini tramite sovrascrittura.

Le immagini in sé vengono impiegate ai fini della generazione di una mappa di calore (heatmap) all'interno della dashboard della piattaforma gestionale, dove è possibile visionare la rappresentazione grafica della numerosità delle persone presenti nelle aree di interesse: la variazione di colore si baserà sul raggiungimento di soglie di numero di persone e verrà rappresentata sulla planimetria statica dell'area stessa: in questo modo, non solo viene rispettato il principio di Data Protection by Design di cui all'articolo 25 del GDPR[48], ma anche quello di necessità, il quale comporta un obbligo di attenta configurazione di sistemi informativi e di programmi informatici per ridurre al minimo l´utilizzazione di dati personali, come previsto dallo stesso punto 2 lettera b) del provvedimento in materia di videosorveglianza, e al rimando fatto dallo stesso all'ex articolo 3 del Codice Privacy.[49]

L'impianto, non venendo impiegato specificatamente per funzioni di videosorveglianza, non prevede che sia richiesta la previa stipula di un accordo sindacale per la tutela dei dipendenti che entrano, durante lo svolgimento delle loro mansioni, nel raggio d'azione dei sensori.

Per quanto concerne l'informazione degli interessati del trattamento posto in essere, il rispetto dell'obbligo di informativa di cui all'articolo 13 del GDPR, nonché, di riflesso, del più generale principio di trasparenza (il quale, riprendendo il considerando 37 del GDPR, afferma che "*le modalità con cui i dati sono raccolti, utilizzati e consultati grazie ad informazioni e comunicazioni facilmente accessibili e comprensibili, utilizzando un linguaggio semplice e chiaro*") è rispettato grazie alla reperibilità dell'informativa breve sulla cartellonistica presente nelle aree oggetto del trattamento, dotata di un rimando, mediante scansione del QR o inserimento in un browser dell'URL dedicato, all'informativa estesa sul sito istituzionale del Titolare, disponibile anche cliccando l'apposito link nel footer della pagina stessa del sito web.[50]

Gli interessati possono esercitare il diritto di accesso di cui all'articolo 15 del GDPR inviando un'apposita richiesta all'indirizzo di posta elettronica reso disponibile dal Titolare all'interno dell'informativa.

---

[48] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 25.
[49] V. D.L. 30 giugno 2003, n. 196, art. 3.
[50] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", considerando 37.

Va tuttavia precisato che nel caso in esame il diritto di portabilità non è esercitabile da parte degli interessati poiché il trattamento, come abbiamo avuto modo di vedere in precedenza, non si basa sulle basi giuridiche del consenso o del contratto, ma su quella del legittimo interesse del Titolare.

Nel rispetto dell'articolo 28 del GDPR è stato stipulato un apposito atto di nomina tra il fornitore, in qualità di Responsabile del trattamento, e la struttura in qualità di Titolare, il quale garantisce, all'interno delle sue clausole, tutti gli obblighi previsti dall'articolo stesso.[51]

All'interno dell'atto, sono riportate anche le certificazioni conseguite (in questo caso ISO27001), in modo da dimostrare l'adozione di adeguate misure volte a garantire il rispetto delle previsioni di cui ai commi 1 e 4 dell'articolo 28.

Per quanto riguarda il modello di Intelligenza Artificiale di Computer Vision, il fornitore lo ha sviluppato non solo impiegando il framework sullo sviluppo sicuro Software Assurance Maturity Model (SAMM) dell'OWASP[52], ma anche nel pieno rispetto dell'articolo 10 dell'AI Act, garantendo che, in termini di governance dei dati raccolti, la formazione del modello fosse strutturata eliminando potenziali bias o lacune di dati all'interno del dataset di training.[53]

Inoltre, l'AI Act pone all'articolo 5, come abbiamo avuto modo di vedere nell'apposito capitolo, un generale divieto all'uso di sistemi di identificazione biometrica remota «in tempo reale» in spazi accessibili al pubblico o per la categorizzazione, mediante tale identificazione biometrica, degli individui al fine di trarre deduzioni in merito alla loro razza, opinioni politiche, appartenenza sindacale, ed altri dati ritenuti particolari.[54]

Il modello di Intelligenza Artificiale impiegato dalla soluzione in esame, tuttavia, esce dal campo di applicazione di tale divieto, in quanto viene impiegato solo al fine di facilitare un processo di gestione dello spazio complessivo delle aree di interesse, e non è dotato né strutturato per compiere operazioni di riconoscimento dei volti.

Non potendo quindi raccogliere dati biometrici dalle riprese che effettua, il modello contenuto nella soluzione è perfettamente conforme alle previsioni dell'AI Act anche sotto questo punto di vista.

Come citato nell'introduzione del capitolo, in ragione dell'analisi del rischio (considerato alto) costituito dal trattamento e dei potenziali impatti sugli interessati in caso di perdita di riservatezza, integrità o disponibilità dei dati, il Titolare ha previsto, in conformità all'articolo 35

[51] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 28.

[52] V. Fondazione OWASP, 2009, "Software Assurance Maturity Model (OWASP SAMM)", *SAMM*, https://owaspsamm.org/model/.

[53] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 1.

[54] V. Regolamento (UE) 2024/1689, "*EU AI Act*", art. 5, lettere g) e h).

del GDPR, la stesura di una DPIA: dei 9 criteri stabiliti dal working party, infatti, sono stati considerati trattati dati relativi a soggetti vulnerabili (le telecamere potrebbero infatti riprendere soggetti con disabilità o minori), il monitoraggio sistematico e l'impiego di soluzioni e tecnologie innovative, oltre che al trattamento di dati su larga scala in considerazione della mole elevata di persone che vengono riprese negli spazi della Struttura.

Infine, il trattamento svolto dal Titolare rispetta quanto previsto dall'articolo 32 del GDPR in merito all'adozione di idonee misure tecniche ed organizzative a supporto dello stesso: infatti, nell'impiego della soluzione, sia il titolare che il responsabile hanno individuato e disposto una serie di misure a tutela dei sistemi impiegati, in modo da rendere il più possibile sicuri i dati personali oggetto del trattamento da accessi non autorizzati, manomissioni o attacchi condotti da soggetti malintenzionati.[55]

Esaminiamo ora quali misure sono state adottate per mitigare il livello di rischio al quale i dati sono esposti.

IV.3 Le misure tecniche ed organizzative adottate

In un panorama dove il progresso delle tecnologie e misure di sicurezza implementate per proteggerle va di pari passo con lo sviluppo di nuove metodologie d'attacco, più che mai si fa concreta la necessità di rendere la propria infrastruttura robusta e resiliente ad ogni tipo di insidia, che sia questa esterna o interna alla rete aziendale.

Proprio con quest'ottica, e con lo scopo di conformarsi alla normativa nel pieno rispetto della previsione di cui all'articolo 32 del GDPR, la Struttura ha implementato una serie di misure tecniche ed organizzative atte a mitigare il rischio che eventi di Data Breach possano concretizzarsi.

Il mondo dell'IoT, come abbiamo potuto osservare all'interno del capitolo dedicato a tale questione, rappresenta un ambiente nel quale la Cyber Security trova ancora una serie di quesiti sulla messa in sicurezza di questi sistemi, per i quali le risposte difficilmente si presentano unitarie nel fornire una soluzione al problema.

Non si può negare l'essenzialità di determinate misure trasversali impiegate, come in questo caso, da parte del titolare del trattamento, quali la definizione di una politica interna per la protezione dei dati personali, l'attenta definizione dei ruoli e responsabilità legate al trattamento e la gestione del personale mediante controlli sul loro operato e attività di formazione e sensibilizzazione degli stessi.

[55] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 32.

Proprio la formazione, infatti, riveste un ruolo fondamentale nella creazione di una vera e propria cultura aziendale della sicurezza, che permette di rendere edotti i dipendenti sugli accorgimenti da adottare nel trattamento dei dati personali e nell'impiego sicuro e cosciente dei sistemi informatici impiegati nell'attività lavorativa.
Tuttavia, anche i sistemi stessi devono essere protetti mediante misure di sicurezza verticali sulla singola soluzione implementata, in modo da colmare quelle fallacie che potrebbero essere sfruttate da un avversario nei suoi tentativi di violazione dell'infrastruttura.

IV.3.1 L'elaborazione delle immagini e l'Edge Computing

Rispetto alle misure di sicurezza tecniche ed organizzative fin qui analizzate, il cuore della sicurezza del sistema implementato dal Titolare sta proprio nel paradigma di Edge Computing adottato: mediante infatti i dispositivi di elaborazione installati adiacentemente ai sensori video, questi consentono il processamento delle immagini solamente per il tempo strettamente necessario, evitando la loro permanenza nei dispositivi, la quale comporterebbe dei rischi in caso di attacchi mirati alla compromissione delle componenti hardware del sistema.
In particolare, le immagini raccolte dal sensore video vengono elaborate direttamente a bordo del dispositivo in tempo reale in soli 50 millisecondi.
Il frame video acquisito viene memorizzato temporaneamente (per una parte del tempo di elaborazione) nella cache della CPU e, successivamente, cancellato e sovrascritto, impedendo il recupero dei dati originali dai dispositivi stessi, nei quali, per l'appunto, non permane nulla oltre il tempo sopracitato.
Grazie a questa tecnica, il Titolare garantisce il rispetto dei principi di minimizzazione e di limitazione della conservazione dei dati, sanciti al paragrafo 1 dell'articolo 5 del GDPR, trattando solo quei dati personali strettamente necessari al perseguimento della finalità predeterminata dal Titolare e solamente per il tempo necessario al suo conseguimento.

IV.3.2 L'autenticazione ed il controllo degli accessi

Nel caso in esame, una prima misura adottata da parte del Titolare consiste nella predisposizione di un efficace controllo degli accessi al sistema effettuati dal proprio personale e da quello del fornitore: questo prevede l'implementazione di un sistema di ruoli applicativi basato sul principio del Least Privilege, garantendo la sicurezza ed il controllo dell'accesso alle informazioni e alle funzionalità della soluzione mediante la definizione delle autorizzazioni e dei permessi che un utente possiede all'interno della stessa.

L'accesso, ad esempio, alla piattaforma gestionale deve necessariamente avvenire mediante l'inserimento di username e password, dove quest'ultima garantisce i requisiti di complessità e robustezza richiesti dalle best practice del settore mediante la definizione di una Password Policy, indicante elementi quali il numero minimo di caratteri, l'impiego di numeri e simboli ed il tempo di validità delle stesse.

In aggiunta, per l'accesso al software gestionale, è stato previsto l'impiego di una Two-Factor Authentication, per la quale il fornitore ha deciso di utilizzare il servizio Google Authenticator.

Per quanto riguarda invece il software di configurazione, l'accesso può avvenire in locale, con inserimento di credenziali nominali nell'apposita centralina installata sui sensori video, oppure da remoto, sempre mediante l'inserimento delle proprie credenziali da parte degli operatori: l'accesso con piene funzionalità tuttavia è stato concesso, al fine di garantire un alto livello di sicurezza del sistema e per prevenire eventuali manomissioni (anche involontarie), solamente al personale tecnico del fornitore, mentre il personale autorizzato della Struttura accede esclusivamente mediante un'utenza in modalità sola lettura.

Infine, per quanto riguarda l'applicazione mobile sulla quale vengono generati gli alert, essa prevede l'impiego del protocollo MQTT con un livello Quality of Service 2 (QoS 2), il quale fornisce un identificatore client e credenziali di tipo username e password, al fine di permettere l'autenticazione dei dispositivi dai quali avviene l'accesso all'applicazione.

Il provisioning delle utenze finora descritte è interamente affidato al fornitore, il quale si assicura di mantenerle aggiornate e monitorate, curandosi di dismettere quelle associate ad utenti che non appartengono più al proprio assetto organizzativo (o a quello del titolare) o quelli che, in ragione, ad esempio, di un cambio di mansione, non necessitano più dei privilegi dei quali erano originariamente in possesso: in questo modo, viene evitato il fenomeno del cosiddetto "Access Creep" (o "Privilege Creep"), nel quale un lento accumulo non monitorato di autorizzazioni, diritti di accesso e privilegi definitivi non necessari in capo ai singoli utenti porta al rischio di accessi indesiderati, manomissioni o veri e propri sabotaggi interni ed attacchi, nel caso in cui queste credenziali finiscano nelle mani di un soggetto malintenzionato.

Un adeguato controllo degli accessi risulta indispensabile per assicurare su base permanente la riservatezza, l'integrità, la disponibilità e la resilienza dei sistemi adottati per il trattamento dei dati personali, adeguandosi a quanto previsto dalla lettera b) del primo paragrafo dell'articolo 32 del GDPR.

IV.3.3 La sicurezza della rete

Oltre al controllo degli accessi, la sicurezza della rete costituisce una misura imprescindibile per la protezione dei dati di tipo testuale e multimediale elaborati dai dispositivi nella loro comunicazione da un dispositivo all'altro.

Tale comunicazione avviene infatti, tra le varie piattaforme software, esclusivamente in modalità protetta: la connessione di rete è resa sicura ed affidabile tramite l'impiego di una rete VPN per tutte le comunicazioni tra i vari client e tra i client e l'unità centrale.

La VPN permette di crittografare la comunicazione tra i client grazie all'impiego di un tunnel IPsec: combinata alla Two-Factor Authentication prima descritta, anche laddove le password vengano sottratte agli utenti, i malintenzionati non potrebbero comunque avere accesso a tale tunnel, rendendo così ancora più complesso ogni tentativo di penetrazione della comunicazione tra gli applicativi.

Infine, nella comunicazione con protocollo MQTT utilizzata per trasferire i dati dai sensori video al software gestionale (nonché l'app mobile per il monitoraggio degli alert) è previsto l'impiego del protocollo TLS 1.2, al fine di evitare possibili azioni di tampering, falsificazione o intercettazione dei dati.

IV.3.4 La cifratura dei dati

Oltre ai protocolli di cifratura visti al paragrafo precedente per quanto riguarda la protezione dei dati nella fase *in transit,* è stata prevista da parte del Titolare la cifratura dei dati *at rest*: questa prevede l'impiego dell'algoritmo AES di cifratura a blocchi, a chiave simmetrica della lunghezza di 256 bit.

In questo modo, i database locali situati all'interno dei sensori video vengono resi impenetrabili da attacchi di forza bruta: combinato con l'impiego di password complesse e robuste, di una Multi Factor Authentication e la presenza di Firewall e Antivirus sui dispositivi impiegati sia dal fornitore che dal Titolare, la chiave viene protetta da possibili attacchi del canale laterale volti a recuperarla.

Per quanto riguarda il database remoto, il fornitore impiega Google Cloud Platform, e pertanto la crittografia dello stesso è gestita dal provider, il quale impiega sempre l'algoritmo AES 256.[56]

Mediante l'adozione della cifratura dei dati sia nella fase di trasmissione che in quella di riposo, il Titolare si è conformato a quanto previsto dall'articolo 32 del GDPR, ed in particolare alla

[56] V. Google Cloud, "Crittografia at-rest predefinita", https://cloud.google.com/docs/security/encryption/default-encryption?hl=it.

lettera a) del paragrafo 1, il quale richiede espressamente l'adozione di protocolli di cifratura e di tecniche atte a pseudonimizzare i dati.

IV.3.5 Il Logging

Al fine di rendere ulteriormente sicuro il sistema e facilitare le operazioni di ricostruzione delle anomalie all'interno del sistema e degli eventi legati ad un eventuale incidente di sicurezza, è stata prevista la raccolta ed il monitoraggio dei log generati dal sistema da parte del Titolare.
In particolare, per tutti gli accessi ai vari software che compongono la soluzione da parte degli utenti e degli amministratori di sistema, viene prevista la raccolta dei loro login, logout e numero di tentativi di accesso falliti.
In aggiunta a questi log di accesso, tuttavia, è stata prevista la registrazione ed il monitoraggio di alcuni log aggiuntivi: tra questi vi sono, per quanto riguarda il software gestionale e l'app mobile, i dettagli relativi alle sezioni del software visitate e le pagine visualizzate, mentre per quanto riguarda il software di configurazione, si prevede anche la raccolta dei log relativi agli eventi relativi alle modifiche apportate ed i relativi esiti.
Tutti questi log sono corredati dai relativi timestamp, che permettono di ricostruire data, ora, minuto e secondo dell'evento generato, consentendo agli operatori del fornitore di rilevare eventuali anomalie o ricostruire a quale utente corrispondono determinate azioni.
La data retention prevista di default dal fornitore della soluzione per la conservazione di questi log è di quindici giorni per quanto riguarda il software di configurazione e quello gestionale, e di sette giorni relativamente all'app mobile, ma questi termini sono liberamente modificabili da parte del Titolare.
L'accesso ai log è stato ristretto esclusivamente ai soggetti con privilegi di amministratore di sistema.

IV.3.6 Il Secure Development Lifecycle

Un elemento di vitale importanza per la sicurezza del sistema che si traduce anche in un maggior grado di sicurezza per i dati trattati dalla Struttura è stata la definizione, da parte del fornitore della soluzione, di un processo di Secure Development Lifecycle (SDLC), la quale permette di ridurre esponenzialmente la quantità e la criticità delle vulnerabilità dalla quale potrebbe essere afflitto il sistema da esso sviluppato.
Tale processo inizia con la fase di pianificazione e progettazione, dove vengono identificati i requisiti di sicurezza e le potenziali minacce per la soluzione, oltre che alla definizione delle misure di controllo adeguate volte a rilevare e prevenire tali minacce.

Segue una fase di sviluppo vero e proprio, che prevede l'effettuazione di un'analisi statica volta a identificare le vulnerabilità presenti nel codice e l'esecuzione di test di sicurezza volti a verificare l'efficacia delle misure di controllo implementate: i test di sicurezza svolti, laddove restituiscano un esito positivo, vengono integrati all'interno del processo di test e collaudo generale sui sistemi del cliente, il quale una volta ultimato con successo porta alla fase finale di rilascio e distribuzione.

In questa fase vengono implementate le procedure di rilascio sicure, volte a garantire che solamente il software che sia privo di vulnerabilità note venga distribuito agli utenti finali (come in questo caso la Struttura).

Tuttavia, tale procedura non si conclude successivamente al rilascio: è infatti prevista una fase di manutenzione e supporto che ricomprende il monitoraggio continuo del software per identificare e correggere eventuali nuove vulnerabilità insorte dopo il rilascio, seguendo così, come dice il nome stesso, il software lungo tutto il suo ciclo di vita, prevenendo l'insorgere di eventuali problematiche future legate all'iniziale sviluppo della soluzione.

È stato inoltre previsto un meccanismo di versioning, dove il software viene rilasciato sotto forma di versioni e release per monitorarne l'andamento e correggere eventuali problematiche insorgenti.

Tale meccanismo di versioning viene eseguito mediante le seguenti fasi:

- Alpha, costituita dal rilascio della soluzione finalizzato al testing interno;
- Beta, costituita da un rilascio finalizzato al testing esterno;
- Candidate Release, costituito da un rilascio per l'effettuazione di test da parte del Titolare;
- General Available Release (rilascio in produzione);
- Maintenance Release;
- Aggiornamento in produzione.

La gestione del processo di versioning avviene mediante l'impiego di un repository GitHub privato, il quale permette di salvare, conservare e tenere traccia di tutte le versioni del codice sottostante ai software sviluppati dal fornitore.

Al fine di irrobustire ulteriormente la sicurezza ed affidabilità della sua soluzione, il fornitore ha organizzato le attività di hardening e patch management a livello di sistema operativo, web server e database sottostanti, permettendo un continuo processo di risoluzione di vulnerabilità e bug volto a migliorare le prestazioni generali del sistema e ridurre la superficie d'attacco disponibile per i soggetti malintenzionati.

Queste attività non sono gestite direttamente dal fornitore, ma da Google, provider dell'ambiente Google App Engine (di seguito anche "GAE") impiegato dal fornitore per lo sviluppo del codice sottostante alla propria soluzione.
A livello di sistema operativo, Google gestisce le attività di patch management e hardening, inclusi i suoi regolari aggiornamenti di sicurezza e la sua configurazione, anche laddove l'accesso diretto al sistema operativo stesso sia possibile: pertanto, non è consentito applicare patch o modificare configurazioni manualmente, ma solamente in modalità automatica.
Viene utilizzato un web server configurato nel rispetto degli standard di settore, e sullo stesso è stata prevista l'esecuzione di aggiornamenti automatici contenenti le patch di sicurezza.
A livello di database, infine, viene utilizzato un database cloud Sequel (SQL), aggiornato automaticamente con le patch di sicurezza necessarie: il database ospita prevalentemente i dati statistici relativi ai livelli di occupazione delle aree della struttura ricavati dalle heatmap generate dal sistema.
Solo i dati relativi ai log di sicurezza della piattaforma presentano dati personali quale il nome utente, ma questi dati (login, logout, tentativi falliti di accesso) vengono esclusivamente impiegati per garantire la sicurezza dell'applicativo da potenziali tentativi di attacco.
Inoltre, GAE è conforme a diversi standard di sicurezza e conformità, tra cui ISO 27001 e Security Operation Center 2 (SOC 2), offrendo diversi livelli di protezione integrati per mitigare i rischi di attacchi applicativi, tra i quali l'impiego di un Web Application Firewall (WAF) integrato, che analizza il traffico in entrata e in uscita per rilevare e bloccare potenziali minacce basate su regole predefinite e personalizzabili, ed un ambiente sandbox, tramite il quale GAE isola ogni applicazione impedendo alle stesse di interagire tra loro o con il sistema sottostante, aiutando a prevenire la propagazione di malware e attacchi tra le diverse applicazioni.
Pertanto, alla luce dell'analisi del rischio compiuta tenendo conto anche del coinvolgimento di Google quale sub-responsabile per il trattamento, in via residuale, dei dati relativi ai log di connessione alla piattaforma (con dati residui quale user ID), questo comporta impatti trascurabili, specialmente in ragione delle efficaci misure tecniche ed organizzative adottate dal provider per la difesa della piattaforma e dei dati ivi contenuti.
Infine, il fornitore ha stabilito una politica di Change Management atta al monitoraggio delle modifiche apportate al sistema, in modo da poter rilevare eventuali malfunzionamenti dovuti ad anomalie o cambiamenti apportati da soggetti non autorizzati.
A tale scopo, ha inoltre individuato un apposito gruppo di Application Management, il quale esegue, sul sistema del fornitore, test funzionali ogni 30 giorni, rilasciando un report dove

vengono segnalati eventuali bug rilevati e predisponendo un piano di remediation per risolvere le eventuali anomalie riscontrate.
In caso di importanti evoluzioni del sistema, viene inoltre valutata, assieme al personale tecnico del Titolare, la necessità di svolgere attività di formazione ed affiancamento, per permettere l'aggiornamento degli stessi relativamente alle nuove funzionalità implementate nella soluzione.

IV.3.7 La formazione

Una misura organizzativa imprescindibile per la sicurezza nel trattamento dei dati personali e nell'impiego della soluzione è costituita dalla formazione del personale: la Struttura, in qualità di Titolare del trattamento, ha infatti previsto l'erogazione di corsi di formazione generali verso i propri dipendenti, mirati a sensibilizzarli sulla materia della protezione dei dati e della sicurezza dei sistemi informativi, adempiendo all'obbligo previsto dall'articolo 29 del GDPR.
Allo stesso modo, anche il fornitore ha stabilito che i propri dipendenti e appaltatori siano tenuti a partecipare a specifici corsi di formazione inerenti alla sicurezza delle informazioni.
Grazie alla formazione su tali tematiche, il personale risulterà essere in grado di riconoscere ed affrontare potenziali tentativi di phishing, sapendo come comportarsi ed evitando di esporre il sistema agli attaccanti mediante comportamenti attenti e consci.
Inoltre, grazie alla formazione in ambito Data Protection, i dipendenti autorizzati al trattamento dei dati saranno in grado di svolgere le loro mansioni nel rispetto dei principi e previsioni del GDPR, consci delle principali criticità poste dal trattamento e in un'ottica di accesso ai dati per le sole finalità previste.

IV.4 Misure di sicurezza: spunti di miglioramento

Sebbene le misure abbiano portato a mitigare i rischi censiti in fase di DPIA fino ad un rischio residuo considerato accettabile, sono stati tuttavia individuati degli spunti di miglioramento, che verranno di seguito descritti.

IV.4.1 Autenticazione: adozione di un sistema di Identity and Access Management (IAM)

Un primo spunto di miglioramento potrebbe essere l'integrazione della soluzione, per quanto concerne il controllo degli accessi e la gestione delle autenticazioni, con il sistema di Identity and Access Management (IAM) della Struttura, in modo che il suo reparto IT sia in grado di automatizzare il controllo dei permessi e dei privilegi in capo agli utenti, affinché le funzioni e i dati siano limitati solo ai soggetti autorizzati e solo per le operazioni strettamente necessarie: l'IAM, infatti, permette all'organizzazione di verificare in modo rapido e preciso l'identità di una

persona, assicurandosi che tale utente abbia le autorizzazioni necessarie per usare la risorsa richiesta durante ogni suo tentativo di accesso.

IV.4.2 Sicurezza della rete: TLS 1.2 vs TLS 1.3

Per quanto riguarda invece la sicurezza della rete, è consigliabile l'adozione del protocollo TLS 1.3 rispetto a quello 1.2, il quale viene superato dal primo rispetto a vari elementi.
Innanzitutto, TLS 1.3 offre una velocità di handshake maggiore rispetto al suo predecessore, il che significa che le applicazioni in tempo reale e i dispositivi IoT possono interagire con i server con una latenza minore: il risultato è una maggiore efficienza, rapidità e fluidità del servizio, il tutto senza dover compromettere la sicurezza della rete.
Proprio rispetto a quest'ultimo punto, TLS 1.3 offre un livello di sicurezza migliore rispetto a TLS 1.2: infatti, rispetto alla versione precedente, è in grado di risolve le vulnerabilità note che si presentano nel processo di handshake (come, ad esempio, quegli algoritmi che gli attaccanti hanno decifrato con successo) ed evitare le vulnerabilità sfruttate in TLS 1.2 relativamente alla compressione dei pacchetti, la quale non viene impiegata in TLS 1.3.

IV.4.3 Offuscamento delle immagini raccolte dai sensori

Passando invece alla considerazione di un'eventuale estrazione delle immagini raccolte dai sensori video, si presenterebbe la necessità di ridurre il più possibile l'eventualità che un malintenzionato possa ricostruire l'identità delle persone recuperando ed operando sulle immagini raccolte dalle telecamere.
Prendendo in considerazione il tempo di elaborazione estremamente ridotto dato dall'impiego del paradigma di Edge Computing, la permanenza delle immagini per solo 50 millisecondi è stata ritenuta dal Titolare una misura di per sé adeguata a evitare possibili visualizzazioni non autorizzate delle immagini.
Tuttavia, per irrobustire ulteriormente la postura di sicurezza del sistema, l'adozione di tecniche di offuscamento delle immagini potrebbe costituire una soluzione al problema: il fornitore stesso ha sviluppato la soluzione prevedendo la possibilità di configurare un meccanismo di blurring dei volti e delle targhe che entrano nella visione dei sensori, in modo tale da rendere del tutto inservibili le immagini ad un potenziale malintenzionato che voglia impiegarle per identificare le persone riprese.

Figura 4 – Esempio di operazione di blurring dei volti compiuta da una telecamera.[57]

Questo meccanismo non è stato tuttavia attivato, in quanto il Titolare non ha riscontrato una criticità in considerazione di due fattori.

Il primo è il passaggio delle informazioni dai sensori video al dispositivo di elaborazione, il quale avviene quasi immediatamente, come menzionato in precedenza.

Questa considerazione è sicuramente corretta con riferimento alla sicurezza del trasferimento delle informazioni al dispositivo di elaborazione menzionato (proprio a tale scopo è stata prevista l'adozione del paradigma di Edge Computing), ma risulta inesatta in considerazione delle riprese fatte direttamente dal sensore, le quali potrebbero essere intercettate.

Immaginiamo infatti che un attaccante riesca ad avere accesso all'hardware del sensore: una volta collegato un dispositivo che si interponga nel flusso di raccolta delle immagini sfruttando una delle porte di cui è dotato il sensore, potrebbe iniziare ad intercettare le immagini in tempo reale e registrarle da remoto, potendo, in un secondo momento, cercare di ricostruire i volti delle persone che sono passate nel raggio d'azione delle telecamere.

La seconda valutazione per la quale il Titolare ha ritenuto non critica la mancata attivazione del meccanismo di blurring sono invece le misure di sicurezza fisica alle quali sono sottoposti i sensori.

Questi sono infatti non solo posti ad un'altezza pari ad una decina di metri, ma rientrano anche nel campo di visione di quelle che sono le telecamere impiegate per scopi di sicurezza, le quali

[57] V. Facit, 2024, "A complete guide to face blurring software", https://facit.ai/insights/face-blurring-software-guide.

sono collegate ad una centrale dove il personale di security monitora quanto avviene negli spazi della struttura: un attaccante avrebbe perciò un elevato grado di difficoltà nell'avvicinarsi ai sensori senza essere rilevato.

La valutazione compiuta dal Titolare risulterebbe rispettare quanto disposto dall'articolo 32 del GDPR, il quale afferma che "*Tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'oggetto, del contesto e delle finalità del trattamento, come anche del rischio di varia probabilità e gravità per i diritti e le libertà delle persone fisiche, il titolare del trattamento e il responsabile del trattamento mettono in atto misure tecniche e organizzative adeguate per garantire un livello di sicurezza adeguato al rischio*".[58]

Sebbene l'attivazione della funzionalità di blurring per il flusso streaming di immagini risulti non strettamente necessaria, al fine di non lasciare le informazioni in chiaro in nessuno dei passaggi compiuti dal sistema per la raccolta dei dati e irrobustire ulteriormente la sicurezza del trattamento in essere, sarebbe consigliabile l'attivazione di tale meccanismo, in modo da ridurre al massimo la superficie d'attacco.

IV.4.4 Logging: conformità al Provvedimento dell'Autorità Garante del 27 novembre 2008

Una misura che va certamente corretta risulta essere quella del logging, in particolare, nel caso in esame, per quanto concerne i tempi di conservazione definiti: questi, infatti, per quanto riguarda gli user, sono generalmente valutati in base alle necessità di sicurezza, ma lo stesso non si può dire per quanto riguarda i log degli amministratori di sistema.

Questi soggetti, avendo un accesso con privilegi massimi o estremamente estesi al sistema, espongono la soluzione a possibili sabotaggi o accessi indiscriminati ed illegittimi alle informazioni ivi contenute, laddove gli amministratori stessi decidessero di abusare dei propri permessi: è con tale *ratio,* infatti, che si è mossa l'Autorità Garante italiana nella creazione del Provvedimento del 27 novembre 2008, considerando proprio la centralità di queste persone nel trattamento dei dati.

Il Provvedimento, intitolato "*Misure e accorgimenti prescritti ai titolari dei trattamenti effettuati con strumenti elettronici relativamente alle attribuzioni delle funzioni di amministratore di sistema"*, prevede infatti, al punto 4.5, che, al fine di controllare l'operato dei soggetti nominati amministratori di sistema, sia prevista la raccolta dei loro access log e la loro conservazione per un minimo di 6 mesi: ciò in funzione della necessità di avere informazioni abbastanza risalenti da poter ricostruire, in caso di data breach o di sospetto sul corretto comportamento degli admin stessi,

[58] V. Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*", art. 32.

eventuali azioni malevole o illegittime che possano aver contribuito ad arrecare danno al sistema o ai dati ivi contenuti.[59]

Perciò, dovrebbe essere corretta da parte del fornitore la data retention stabilita per i log relativi a tali soggetti, in quanto essa risulta del tutto inadeguata, garantendo così, anche su questo aspetto, la conformità al Provvedimento: altre previsioni contenute nel Provvedimento (quali la pubblicazione e comunicazione dell'elenco degli amministratori di sistema al titolare del trattamento e l'immodificabilità e correttezza dei log) sono già rispettate dal fornitore, il quale prevede anche un controllo annuale sull'operato degli amministratori di sistema.

IV.4.5 Protezione del modello di Intelligenza Artificiale

Per quanto riguarda l'algoritmo di Intelligenza Artificiale di Computer Vision impiegato dai sensori IoT, il fornitore non sembra accennare a particolari misure di sicurezza: tuttavia, al fine di irrobustire la certezza dei risultati restituiti, garantendo così un'efficacia maggiore del sistema, sarebbe consigliabile l'adozione di tecniche volte a ridurre l'impatto causato da potenziali Adversarial Attacks.

Una possibile soluzione potrebbe essere quella di introdurre dei campioni nel modello di training che contengano Adversarial Attacks, oppure simulare degli Adversarial Attacks durante la fase stessa di training: tuttavia, in questo modo, si va ad intaccare la procedura di training stessa, e si aumentano i costi dovendo consumare ulteriori risorse.

Una soluzione che non comporterebbe tali problematiche potrebbe essere quella di non andare ad intaccare il modello di training, mantenendo quello definito, ma andare ad agire introducendo nuove fasi prima del passaggio delle immagini al modello di Computer Vision.[60]

Tale modello ottenuto alle fine del training regolare e, che in fase di test, ha soddisfatto tutte le KPI di performance, viene utilizzato così com'è: quello che cambia è il non passare le immagini in input direttamente al modello, ma di anteporre allo stesso due layer di randomizzazione disposti in sequenza.

Il primo layer esegue un ridimensionamento random della immagine in input (con intervallo di valori non troppo grande intorno alle dimensioni previste per il layer di input della rete neurale), mentre il secondo, prendendo quale input l'output del primo layer, introduce un cosiddetto "*padding"* (cioè un'aggiunta di spazi a uno dei lati dell'immagine) in maniera casuale.

---

[59] V. Provvedimento del Garante per la Protezione dei Dati Personali del 27 novembre 2008, "*Misure e accorgimenti prescritti ai titolari dei trattamenti effettuati con strumenti elettronici relativamente alle attribuzioni delle funzioni di amministratore di sistema",* punto 4.5.

[60] V. G. Iozzia, 2021, "*Adversarial Attacks in Computer vision: una strategia difensiva per mitigarne gli effetti"*, *Intelligenza Artificiale*, https://www.ai4business.it/intelligenza-artificiale/adversarial-attacks-in-computer-vision-una-strategia-difensiva-per-mitigarne-gli-effetti/.

Il passaggio attraverso i due nuovi layer va a rimuovere gran parte delle perturbazioni che verrebbero introdotte dall'attaccante mediante l'Adversarial Attack, a prescindere dalla tecnica da esso usata.

Pensiamo ad un attacco che vada a disturbare la frequenza del segnale delle telecamere, andando a disturbare la capacità delle stesse di raccogliere un'immagine dettagliata, o ancora quegli attacchi dove, mediante alcune immagini contenenti pattern particolari stampate, ad esempio, sulla maglietta dell'attaccante, si va a confondere la capacità del modello di riconoscere la presenza di una persona, andando a nasconderla, ad esempio, facendola considerare dal modello quale un oggetto appartenente all'area (cd. "adversarial patching").

Figura 5 – Esempio di un attacco "adversarial patches": la persona sulla destra, grazie all'immagine stampata apposta alla base della maglietta è in grado di ingannare il modello di computer vision.[61]

L'immagine, laddove non completamente alterata, viene "ripulita" e data in ingresso al modello, il quale, nella maggioranza dei casi, produrrà il risultato corretto o rileverà la problematica, anche se con un confidence score un po' più basso rispetto al caso in cui l'immagine originale fosse stata passata direttamente al modello stesso, evitando così di intaccare il dataset di training.

Il compromesso nell'adozione di questa misura sarebbe, quindi, quella di avere risultati più sicuri a leggero discapito della precisione del modello.

[61] V. B. Yrka, "Using a printed adversarial patch to fool an AI system", 2019, https://techxplore.com/news/2019-04-adversarial-patch-ai.html.

Il motivo per cui il Titolare non ha voluto implementare ulteriori misure a protezione del modello risiede prevalentemente nell'uso che ne viene fatto: gli attacchi precedentemente ipotizzati infatti sarebbero facilmente rilevabili o aventi un impatto nullo in considerazione del fatto che il modello si concentra sulla raccolta di una mera immagine generale delle affluenze nelle aree di interesse.
Anche in questo caso, in funzione del bilanciamento tra costi attuativi, dell'effettiva probabilità della manifestazione di un danno in capo agli interessati e delle misure di sicurezza già adottate, viene garantito, da parte del Titolare, il rispetto dell'articolo 32 del GDPR.

IV.4.6 Procedura di smaltimento dei dispositivi

Infine, un'ulteriore misura di sicurezza di tipo organizzativo dovrebbe essere la definizione, da parte del fornitore, di una procedura accurata di smaltimento dei dispositivi: la fase di End-of-Life (EOL) di un dispositivo e la dismissione del suo hardware risulta infatti essere un momento critico nella protezione dei dati raccolti tramite lo stesso, in quanto uno smaltimento incorretto e svolto senza l'esame dell'effettiva assenza di dati ancora presenti all'interno degli asset può risultare un errore problematico, portando potenzialmente ad una violazione di dati personali laddove un soggetto riesca ad entrare in possesso di questi dispositivi dismessi.
La definizione di una precisa procedura di smaltimento permette la rimozione dei dati sovrascritti prima della dismissione dei sensori, garantendo la non permanenza al loro interno di dati personali che potrebbero eventualmente essere utilizzati da soggetti non autorizzati che ne entrino in possesso.

V. Il progetto City Brain: uno spunto di riflessione

Il recentissimo sviluppo di sempre più innovative tecnologie digitali ha promosso la proliferazione, in numerosi stati del mondo, dell'idea di adottare nelle proprie aree urbane un modello di gestione basato sulle Smart Cities, al fine di ottimizzare ulteriormente i servizi pubblici, efficientare la gestione delle risorse e la qualità complessiva della vita dei cittadini che vi abitano.

Il progetto con maggiore visibilità nel settore è costituito dal "City Brain" sviluppato dalla nota tech company Alibaba Cloud: tale soluzione si propone come una delle più avanzate tecnologie volte alla gestione delle aree urbane intelligenti, andando a sfruttare a pieno l'integrazione tra modelli di AI, dispositivi IoT, Edge computing e cloud computing ibrido già presenti nelle Smart Cities, ma rendendo questi sistemi comandati centralmente da un unico insieme di modelli AI che, in maniera multitasking e senza separazioni infrastrutturali per i vari ambiti applicativi, va a raccogliere e processare enormi quantità di dati in tempo reale per rendere la città un'estensione vivente di questo sistema in grado di controllarne ogni aspetto, aumentando la qualità complessiva della vita dei suoi abitanti.

Questo progetto innovativo è diventato realtà in Cina, dove è stato implementato in più di 20 città, e ha riscosso così tanto successo che per la prima volta un'applicazione è stata vista anche al di fuori della Repubblica Popolare Cinese, nella capitale della Malesia, Kuala Lumpur.

L'implementazione di un sistema così centralizzato e strutturato porta tuttavia, assieme a grandi opportunità, anche enormi sfide, specialmente con riguardo all'ambito della protezione dei dati personali.

Questo sistema, essendo costituito da un insieme elevato di tecnologie sempre in ascolto, porta ad una raccolta continua di dati comuni e particolari da telecamere, sensori ambientali e altri dispositivi IoT alla base di queste estese infrastrutture digitali, che possono inevitabilmente portare a conseguenze disastrose laddove vi sia un'interruzione del servizio o l'accesso al sistema da parte di attaccanti con intenzioni malevole, tenendo a mente che, seppur le tecnologie siano le medesime alla base di ogni Smart City, particolare attenzione vada posta anche a come i principali attori decidono di impiegare tali sistemi.

Il capitolo si concentrerà sull'esame del progetto "City Brain", osservandone le applicazioni concrete nelle realtà cinesi, fornendo uno spunto di riflessione sulle criticità ed opportunità offerte da tale tecnologia con particolare riferimento alla protezione dei dati, esaminando il bilanciamento tra innovazione e Data Protection effettuati nella realtà asiatica e rapportandola a quella europea alla luce dei diversi approcci geopolitici al tema, indicando inoltre delle possibili strategie volte ad adattare tale soluzione al contesto comunitario.

V.1. City Brain Project: nuova frontiera per le Smart Cities?

Il progetto City Brain venne presentato per la prima volta da Alibaba Cloud nel 2016, proponendosi quale nuova frontiera tecnologica per l'efficientamento della gestione delle Smart Cities grazie all'integrazione di diversi sistemi per creare una soluzione del tutto innovativa: il City Brain si basa infatti su di una piattaforma di raccolta ed elaborazione dei dati su larga scala mediante tecnologie proprietarie di Alibaba Cloud, integrata con tecnologie avanzate quali modelli AI di Computer Vision, calcolo di rete topologica su larga scala mediante creazione di un Digital Twin della città target ed una mappatura predittiva del flusso del traffico stradale.

Il City Brain è così in grado di fornire enormi quantità di dati raccolti da un sistema multi-fonte che permette la raccolta, elaborazione in tempo reale e calcolo intelligente dei diversi eventi e delle gigantesche moli di informazioni generate quotidianamente nell'ambiente urbano.[62]

Un ruolo da protagonista nel City Brain è svolto dai modelli di AI ML e DL (Machine Learning e Deep Learning) sottostanti al sistema decisionale dell'intera infrastruttura, i quali sono in grado di coordinare le principali infrastrutture, servizi, dispositivi e tecnologie necessarie tanto per la gestione del traffico che per motivi di pubblica sicurezza, arrivando ad occuparsi persino dell'efficienza energetica ed ecologica degli ambienti urbani coinvolti.

I vantaggi e le opportunità offerte dall'adozione di questo paradigma sembrerebbero essere numerosi, proprio grazie all'integrazione con le diverse tecnologie già disponibili all'interno degli ecosistemi digitali costituiti dalle Smart Cities, quali i dispositivi IoT, le reti 5G, i Big Data e l'estesa presenza di sensoristica negli spazi pubblici.[63]

Attualmente la Cina ha deciso di implementare tale sistema in ben 22 città, tra le quali Shangai, Beijing e Macao, e perfino la capitale della Malesia, Kuala Lumpur, ha previsto l'impiego di questo sistema per la gestione del traffico cittadino, e con risultati che non hanno certo tardato a dimostrare l'efficacia di tale soluzione.

Basti pensare che nell'arco di un solo anno dall'adozione del sistema la velocità del traffico in queste città vide un incremento superiore al 15% e tutti gli incidenti stradali nelle zone coperte dal sistema vennero automaticamente rilevati con una tale precisione che persino i tempi medi di intervento dei servizi emergenziali furono ridotti di ben 3 minuti.

Anche le soste ed i parcheggi illegali vennero costantemente monitorati in diretta, permettendo addirittura di assegnare automaticamente le multe senza la necessità di dover inviare sul posto

---

[62] V. J. Zhang, X. Hua, J, Huang, X. Shen, J. Chen, Q. Zhou, Z. Fu, Y. Zhao, "*City brain: practice of large-scale artificial intelligence in the real world*", *IET Journals*, 2019.

[63] V. G. Mussi, 2025, "City Brain: l'intelligenza artificiale al servizio delle Smart City", *Smart*, https://elettricomagazine.it/smart-tech-tecnologie-intelligenti/city-brain-intelligenza-artificiale-trasforma-smart-city/.

un agente, efficientando così anche l'impiego del personale delle forze dell'ordine per i soli eventi strettamente necessari.[64]

Per comprenderne meglio il funzionamento, il City Brain si basa sui seguenti step:

Figura 6 – Overview dei layer di funzionamento del City Brain[65]

Tramite l'elaborazione e l'analisi massiva dei video e delle immagini raccolte mediante i sistemi di telecamere diffusi nell'ambiente urbano, il primo layer cognitivo ottiene non solo dati relativi allo stato del traffico e delle emergenze all'interno delle aree pubbliche in tempo reale, ma anche agli eventi anomali in aree specifiche nel tempo, quali crimini o assembramenti non autorizzati. Il sistema di cognizione, mediante l'accesso visivo ai dati e all'elaborazione multimediale degli stessi da parte dell'algoritmo visivo prevedendo che le risorse video provenienti dalle diverse fonti siano accessibili tramite protocolli video standard, permettendo di elaborare su larga scala enormi moli di informazioni basandosi sulla piattaforma cloud proprietaria di Alibaba, al fine di produrre una prima opera di differenziazione e di riconoscimento degli scenari monitorati dal sistema.

Mediante il secondo layer, preposto all'ottimizzazione dei processi ed al compimento di decisioni basate sulla base della precedente acquisizione dei dati compiuta dal sistema di cognizione, viene effettuata un'integrazione e analisi delle grandi moli di informazioni eterogenee generate dalle varie fonti quali dispositivi IoT e sensori installati nelle aree urbane.

Grazie a questo layer, il City Brain è in grado di ottimizzare il flusso dei veicoli e del traffico urbano, migliorando al contempo la governance della città anche in termini di sicurezza compiendo processi decisionali non solo con riferimento alla gestione del traffico, ma anche delle costruzioni, del dispendio energetico e della sicurezza degli spazi pubblici, chiamando in automatico soccorsi in caso di incidenti o le forze dell'ordine nel caso del compimento di crimini.

---

[64] V. M. Ferrera, 2025, "City Brain, come l'Intelligenza Artificiale rivoluziona un brand – e una città", *E-Campus Digital School*, https://www.digitalschool.com/blog/city-brain-alibaba-come-intelligenza-artificiale-rivoluziona-un-brand-e-una-citta/.

[65] V. J. Zhang, X. Hua, J, Huang, X. Shen, J. Chen, Q. Zhou, Z. Fu, Y. Zhao, "*City brain: practice of large-scale artificial intelligence in the real world*", *IET Journals*, 2019.

Nel layer di ricerca e mining, il sistema inserisce l'intero flusso dei dati raccolti mediante le telecamere, sensori ed altri dispositivi IoT all'interno di un database in cloud, effettuando operazioni di ricerca su questi dati indicizzandoli e interrogando modelli di AI richiamati mediante apposite API (Application Programming Interface).

Alibaba mette a disposizione uno specifico motore di ricerca per analizzare i video e localizzare rapidamente oggetti ed individui, come persone scomparse, ricercate o veicoli fuggiti dalla scena di un incidente o crimine, selezionando gli scenari ed immagini pertinenti fra l'enorme volume di video generato in maniera rapida ed efficiente attraverso diverse tecnologie quali sistemi di riconoscimento facciale e biometrico, sistemi di analisi intelligente dei video e strategie di re identificazione degli individui.

Infine, è negli ultimi due layer, quello predittivo e quello di intervento, che entra in gioco il Digital Twin della città: infatti, mediante le analisi ed elaborazioni compiute dai layer precedenti, vengono creati degli accurati modelli tridimensionali delle aree urbane, i quali vengono prodotti da una trasposizione in 3D delle immagini e disegni di progettazione 2D mediante tecniche di ricostruzione e modellazione degli scenari. [66]

In particolare, mediante un modello AI di Computer Vision, le immagini e video vengono sottoposti ad un processo di analisi intelligente comprensivo del rilevamento, tracciamento, conteggio delle folle e rilevamento delle anomalie rilevate dai sistemi di sorveglianza installati nelle aree urbane, integrando tali dati con ulteriori informazioni carpite dai numerosi sensori e dispositivi IoT (come temperatura, umidità, fumo, ecc.).

In questo modo, il sistema può compiere accurate previsioni di eventuali eventi atmosferici o umani che possano mettere in pericolo l'ordine pubblico o il normale funzionamento della città, rilasciando allerte meteo, ridirezionando il traffico o ancora, inviando pompieri sul luogo di un possibile principio di incendio o pattuglie di polizia in caso di situazioni sospette prima di una loro potenziale escalation.

V.2. I rischi legati alla protezione dei dati

I vantaggi, come abbiamo potuto vedere, sono evidenti con riferimento alla gestione delle aree urbane.

Tuttavia, dobbiamo anche volgere l'attenzione sulle problematiche in ambito Data Protection legate ad un uso così smodato di sistemi di rilevamento sempre attivi e che permettono il

[66] V. J. Zhang, X. Hua, J, Huang, X. Shen, J. Chen, Q. Zhou, Z. Fu, Y. Zhao, "City brain: practice of large-scale artificial intelligence in the real world", *IET Journals*, 2019.

riconoscimento ed il monitoraggio continuato degli individui, i quali risultano estremamente impattanti per gli interessati di questi trattamenti di dati personali.

Fra le numerose problematiche vi è innanzitutto il fatto che le infrastrutture che implementano tale soluzione, che si tratti di trasporto pubblico, viabilità, gestione degli spazi pubblici o persino i semplici sistemi di gestione dei rifiuti, raccolgono dati su ogni individuo che interagisce con tali elementi.

In assenza di una normativa specifica come quella europea, principi cardine, come quello della trasparenza dei trattamenti dei dati, vengono inevitabilmente meno, portando ad una mancanza di conoscenza da parte degli interessati non solo delle finalità a supporto di questi trattamenti, ma anche dei rischi ai quali sono esposti.

Inoltre, anche laddove sia definita una specifica finalità, difficilmente questa verrebbe rispettata se prendiamo in considerazione l'estensione dei dati raccolti da questi sistemi, la quale talvolta può sfuggire anche agli stessi soggetti che ne hanno previsto l'implementazione considerata l'ampiezza della superficie ricoperta da questa infrastruttura.

Il trattamento risulterebbe in questo caso invasivo, discostato da principi del GDPR quali quello della minimizzazione nella raccolta dei dati che permettono di garantire che questi vengano raccolti solo nella quantità necessaria per il raggiungimento della finalità prevista.[67]

Un altro enorme problema legato alle infrastrutture digitali delle Smart Cities e di riflesso anche del City Brain quale loro evoluzione è garantire la sicurezza informatica di questi complessi sistemi: pensiamo a come anche un singolo punto di vulnerabilità sfruttata da un gruppo di attaccanti potrebbe portare ad un data breach di proporzioni elevate, consentendo agli hacker di avere accesso ad un volume di informazioni personali elevatissimo tra cui indirizzi, dettagli di pagamento e persino dati biometrici utilizzati per il riconoscimento facciale.

Prendiamo un esempio pratico: nel 2019 un singolo attacco informatico condotto contro la rete elettrica di Johannesburg ha causato interruzioni della distribuzione di elettricità in diverse aree della città lasciando sfornite le persone di questo bene essenziale per ore, dimostrando in concreto quanto un anello debole nella catena di sicurezza di queste infrastrutture possa portare ad esiti disastrosi se non addirittura al parziale collasso delle reti urbane, creando non solo meri disservizi e disagi, ma anche vere e proprie minacce per la pubblica sicurezza o ritardi negli interventi dei servizi di soccorso o delle forze dell'ordine.[68]

---

[67] V. BigDataDissent (a cura di), 2024, "Smart Cities and Privacy Concerns: 7 Hidden Risks of Digital Infrastructure You Should Know", *Tech Society*, https://bigdatadissent.com/smart-cities-and-privacy-concerns-7-hidden/.

[68] V. BigDataDissent (a cura di), 2024, "Smart Cities and Privacy Concerns: 7 Hidden Risks of Digital Infrastructure You Should Know", Tech Society, https://bigdatadissent.com/smart-cities-and-privacy-concerns-7-hidden/.

V.3. I diversi bilanciamenti tra realtà cinese ed europea

L'applicazione di un sistema di questa tipologia all'interno dello spazio europeo, in ragione del GDPR e dell'AI Act, parrebbe molto problematico in quanto in contrasto con i principi alla base di questi regolamenti.

Ad esempio, nel caso di studio illustrato nel capitolo IV è stata evidenziata l'attenzione posta alla minimizzazione dei dati personali raccolti dai sensori (ad esempio la ridotta risoluzione dell'immagine, le heatmap, blurring) pur ottemperando alle finalità del trattamento.

Nella realtà cinese il focus principale parrebbe invece focalizzarsi sull'efficienza dei processi relativi alla gestione della sicurezza, andando a concentrarsi maggiormente proprio sulla raccolta di dati biometrici mediante telecamere intelligenti all'interno di spazi pubblici, o andando addirittura ad implementare dei meccanismi che vanno a ripulire e rendere più nitide anche quelle immagini raccolte da telecamere a bassa risoluzione, al fine di garantire il riconoscimento dei soggetti ripresi.

Per applicare questa tipologia di soluzione all'interno dell'UE, sarebbe innanzitutto necessario che non vengano raccolti negli spazi pubblici i dati biometrici (categorizzati quali dati particolari ai sensi dell'art.9 del GDPR) oltre che ad una loro elaborazione da parte di modelli di AI, in modo da poter fuoriuscire dalla categoria dei sistemi a rischio inaccettabile previsti dall'AI Act all'articolo 5.

Sarebbe necessario informare attentamente gli interessati, mediante opportune affissioni e rimandi ad informative estese, quali siano le finalità dei trattamenti posti in essere e quali siano le tipologie di dati raccolte da parte di questi sistemi, adeguandosi al principio di trasparenza.

Inoltre, dato che tale soluzione è interamente sviluppata e gestita da un'azienda Big Tech, seppur applicata in spazi la cui gestione dovrebbe essere prevalentemente affidata all'amministrazione dello Stato, andrebbe stipulato da parte del Titolare del trattamento un apposito accordo ai sensi dell'articolo 28 del GDPR che definisca in modo chiaro i ruoli e responsabilità legati a tale trattamento dei dati, stabilendo inoltre gli obblighi in capo alle diverse parti e le misure adottate per garantire la sicurezza del trattamento.

Infine, per garantire il principio di minimizzazione, il Titolare deve definire chiaramente e raccogliere solo i dati strettamente necessari alla gestione della Smart City, senza sforare in impieghi che costituiscano finalità a sé stanti o illecite.

Pertanto, perché il City Brain sia applicabile nell'UE, il bilanciamento degli interessi dovrà pendere maggiormente dalla parte del rispetto della protezione dei dati degli interessati, rinunciando ad alcuni potenziamenti ed efficientamenti dei servizi resi alla popolazione in cambio del rispetto dei loro diritti fondamentali.

V.4. L'attenzione ai possibili abusi del sistema

Nell'attenzione posta dal pubblico e dagli esperti con riferimento a soluzioni tecnologiche di questa portata, un'importante riflessione viene rivolta anche ai potenziali impieghi abusivi che possono essere fatti di questi sistemi da parte di enti governativi, i quali potrebbero decidere di impiegarli per fini ulteriori rispetto a quelli di mera gestione ed amministrazione intelligente delle città: considerando ad esempio la Cina, basti osservare il loro sistema di "credito sociale", mediante il quale il comportamento pubblico e gli allineamenti politici della popolazione viene assiduamente monitorato e giudicato da parte del governo e delle forze dell'ordine.

Combinando la videosorveglianza dotata di meccanismi di riconoscimento biometrico con questo sistema di credito sociale (nonché le decisioni prese dai modelli di intelligenza artificiale in autonomia), il rischio che tale tecnologia venga impiegata per fini ulteriori diventa inevitabilmente una realtà.

Una problematica in termini di Data Protection è legata ai ruoli privacy di questi grandi provider: il titolare può liberamente decidere di avvalersi di un responsabile del trattamento, ma le responsabilità indicate dal GDPR restano pur sempre in capo ad egli in via esclusiva.

Con la progressiva dematerializzazione delle strutture informatiche e l'affidamento sempre più frequente ai grandi cloud provider, è il fornitore della soluzione che detiene le competenze ed i mezzi per gestire la soluzione, in via quasi del tutto esclusiva, rischiando di ribaltare il ruolo che normalmente si configura tra responsabile e titolare.

Il City Brain, infine, porta con sé un ulteriore problematica legata agli impieghi di queste soluzioni: il controllo prevalentemente in capo alle Big Tech, nonché la pervasività del trattamento e l'esclusività in capo a queste aziende del possesso di competenze e mezzi potrebbero portare ad abusi da parte di queste grandi aziende, detentrici, all'interno di un mercato in costante accentramento, di porzioni di spazio digitale sempre maggiori, le quali portano a loro volta ad una sempre maggiore incisività nelle vite delle persone.

Perciò, seppure i vantaggi costituiti dalle tecnologie impiegate nelle Smart Cities siano innegabili, un loro impiego improprio potrebbe facilmente portare a forme di controllo governativo e corporativo estremamente invasive, sollevando non pochi dubbi in merito al trade-off tra libertà fondamentali dei cittadini e democrazia con l'innovazione ed il progresso.

Conclusioni

In conclusione, a seguito delle riflessioni fin qui condotte, abbiamo potuto osservare come nel panorama sempre in evoluzione degli ecosistemi digitali costituiti dalle Smart Cities, a giocare un ruolo sempre più importante è l'integrazione sempre più diffusa tra le innovative tecnologie che stanno rivoluzionando il nostro mondo, come l'AI, l'IoT e l'Edge Computing, le quali risultano i cardini essenziali per un efficientamento sempre maggiore dei processi non solamente gestionali, ma anche strategici e decisionali.

Pur riconoscendo i notevoli vantaggi offerti da queste soluzioni, non abbiamo potuto trascurare i rischi connessi al loro utilizzo.

Tali rischi riguardano infatti non solo le potenziali minacce di natura cibernetica che possono compromettere la sicurezza di tali sistemi, ma anche gli impatti che i trattamenti di dati possono avere su un numero elevatissimo di soggetti, con conseguenze potenzialmente gravi in caso di violazione o uso improprio delle informazioni.

In particolare, con il caso del City Brain, si evidenziano i potenziali impieghi impropri delle funzionalità che tali tecnologie offrono, permettendo di dare uno sguardo, in prospettiva, all'esigenza di un approccio fortemente focalizzato sul bilanciamento tra libertà, diritti fondamentali e progresso tecnologico, dove i potenziali abusi di potere si fanno tristemente sempre più concreti e possibili.

Il framework regolatorio vigente in Europa risulta essere fortemente incentrato sulla protezione dei dati: il legislatore europeo vuole infatti porre cautela nell'impiego di tecnologie che potrebbero portare a rischi elevati per gli interessati, tracciando dei limiti per quelle soluzioni tecnologiche e modelli di Intelligenza Artificiale che potrebbero costituire un pericolo per tutti i cittadini dell'Unione.

Risulta inoltre vitale, per proteggere con maggiore attenzione e consapevolezza i diritti degli interessati dei trattamenti, che la figura del DPO ampli le proprie competenze e si evolva al passo con l'innovazione tecnologica, ottenendo sempre più un'interdisciplinarietà che contempli uno spazio sempre maggiore per quelle che sono le conoscenze informatiche e di Cyber Security, in modo da poter comprendere al meglio il funzionamento delle tecnologie digitali e, di riflesso, le minacce e i rischi specifici legati alle attività di trattamento dei dati personali.

Bibliografia

C. Pielli, D. Zucchetto, A. Zanella, L. Vangelista, e M. Zorzi, "Platforms and protocols for the Internet of Things,", *EAI Endorsed Transactions on Internet of Things*, vol. 15, n. 1, 2015

F. Li, A. Shinde, Y. Shi, J. Ye, X. Li, e W. Z. Song, "System statistics learning-based IoT security: Feasibility and suitability.", *IEEE Internet of Things Journal*, 2019

F. Meneghello, M. Calore, D. Zuccheto, M. Polese e A. Zanella, "IoT: Internet of Threats? A survey of practical security vulnerabilities in real IoT devices.", *IEEE Internet of Things Journal*, Vol.6, 2019

J. Zhang, X. Hua, J, Huang, X. Shen, J. Chen, Q. Zhou, Z. Fu, Y. Zhao, "City brain: practice of large-scale artificial intelligence in the real world", *IET Journals*, 2019

S. Ghayyur, X. He, D. Ghosh, S. Mehrotra, "Towards Accuracy Aware Minimally Invasive Monitoring (MiM)", ACM Conference on Computer and Communications Security (Theory and Practice of Differential Privacy), 2019

Sitografia

A. McCall, 2024, "Cyber Security in the Age of AI and IoT: Emerging Threats and Defense Strategies", Research Gate, https://www.researchgate.net/publication/386050391_Cyber Security_in_the_Age_of_AI_and_IoT_Emerging_Threats_and_Defense_Strategies

BigDataDissent (a cura di), 2024, "Smart Cities and Privacy Concerns: 7 Hidden Risks of Digital Infrastructure You Should Know", Tech Society, https://bigdatadissent.com/smart-cities-and-privacy-concerns-7-hidden/.

B. Ridolfi e A. Vaccarelli, 2023, "Sicurezza IoT, come evitare i rischi: le norme tecniche e la privacy", Sicurezza Digitale, https://www.agendadigitale.eu/sicurezza/le-norme-tecniche-e-la-privacy-per-la-sicurezza-nelliot-cosa-sapere-per-evitare-rischi/

B. Yrka, 2019, "Using a printed adversarial patch to fool an AI system", https://techxplore.com/news/2019-04-adversarial-patch-ai.html.

V. Facit, 2024, "A complete guide to face blurring software", https://facit.ai/insights/face-blurring-software-guide

F. Bloise, 2024, "Guida alle smart city, cosa sono e come funzionano le città "intelligenti"", *E-Motion Mag*, https://platum.com/e-motion-mag/smart-city/smart-city-cosa-e-come-funziona-citta-intelligente/

F. Ferrazza, 2019, "Minacce APT: cosa sono le Advanced Persistent Threat, come funzionano e come difendersi", Malware e Attacchi Hacker, https://www.Cyber Security360.it/nuove-minacce/minacce-apt-cosa-sono-le-advanced-persistent-threat-come-funzionano-e-come-difendersi/

Fondazione OWASP, 2005, "Secure development and integration", OWASP Developer Guide, https://devguide.owasp.org/02-foundations/02-secure-development/

Fondazione OWASP, 2009, "Software Assurance Maturity Model (OWASP SAMM)", SAMM, https://owaspsamm.org/model/

G. Iozzia, 2021, "Adversarial Attacks in Computer vision: una strategia difensiva per mitigarne gli effetti", Intelligenza Artificiale, https://www.ai4business.it/intelligenza-artificiale/adversarial-attacks-in-computer-vision-una-strategia-difensiva-per-mitigarne-gli-effetti/

G. Mussi, 2025, "City Brain: l'intelligenza artificiale al servizio delle Smart City", *Smart*, https://elettricomagazine.it/smart-tech-tecnologie-intelligenti/city-brain-intelligenza-artificiale-trasforma-smart-city/

G. Salvadori, 2019, "IoT e AI: l'Intelligenza Artificiale incontra l'Internet of Things", *Internet of Things*, https://blog.osservatori.net/it_it/intelligenza-artificiale-e-iot

Google Cloud, "Crittografia at-rest predefinita", https://cloud.google.com/docs/security/encryption/default-encryption?hl=it

L. Garbati e C. Ponti, 2024, "AI Act: che cos'è, obiettivi e sanzioni previste", Intelligenza Artificiale, https://www.ai4business.it/intelligenza-artificiale/ai-act-che-cose-obiettivi-e-sanzioni-previste/

L. Varriale, 2025, "CISA: nuove linee guida per i dispositivi edge mentre APT cinesi utilizzano backdoor SSH", Sicurezza Informatica, https://www.matricedigitale.it/sicurezza-informatica/cisa-nuove-linee-guida-per-i-dispositivi-edge-mentre-apt-cinesi-utilizzano-backdoor-ssh/

M. Ferrera, 2025, "City Brain, come l'Intelligenza Artificiale rivoluziona un brand – e una città", E-Campus Digital School, https://www.digitalschool.com/blog/city-brain-alibaba-come-intelligenza-artificiale-rivoluziona-un-brand-e-una-citta/

Osservatorio Internet of Things (a cura di), 2019, "Internet of Things (IoT): significato, esempi e applicazioni", *Internet of Things*, https://blog.osservatori.net/it_it/cos-e-internet-of-things

Redazione di Talking IOT (a cura di), 2023, "How does IoT impact privacy and data security?", Talking IOT, https://talkingiot.io/how-does-iot-impact-privacy-and-data-security/

Redazione ZeroUno (a cura di), 2022, "Edge computing, cos'è, come funziona e come implementarlo", *Cloud Computing*, https://www.zerounoweb.it/techtarget/searchdatacenter/edge-computing-cose-come-implementarlo/

Red Hat, "2022 Global Tech Outlook: A Red Hat report", https://www.redhat.com/en/resources/global-tech-outlook-overview-2022

Fonti legislative

Decreto-legge 30 giugno 2003, n. 196, "*Codice in materia di protezione dei dati personali*"

ENISA, "Guidelines for securing the Internet of Things: Security supply chain", 2020

ISO/IEC 27400/2022, "*IoT Security and Privacy Guidelines*"

ISO/IEC 29192-2012, "Information technology – Security techniques – Lightweight cryptography"

Provvedimento del Garante per la Protezione dei Dati Personali del 27 novembre 2008, "*Misure e accorgimenti prescritti ai titolari dei trattamenti effettuati con strumenti elettronici relativamente alle attribuzioni delle funzioni di amministratore di sistema"*

Provvedimento generale del Garante per la Protezione dei Dati Personali dell'8 aprile 2010, "*Provvedimento in materia di videosorveglianza*"

Regolamento UE n.679/2016, "*Regolamento generale sulla protezione dei dati*"

Regolamento UE 2023/2854, *"Data Act"*

Regolamento (UE) 2024/1689, "*EU AI Act*"