Aggiornamento 2026

AGENDA DI RICERCA E INNOVAZIONE PER LA CYBERSICUREZZA

AGENZIA PER LA CYBERSICUREZZA NAZIONALE
ACN
SAPIENTIA ADIUVAT

MUR
Ministero dell'Università e della Ricerca

Versione	Anno	Autori	Contatto
2.0	**2026**		**programmi-ricerca@acn.gov.it**

Indice

Novità di questa versione 4

01 Introduzione 5
1.1 Metodologia 9
1.2 Struttura del Documento 13

02 Panoramica dell'aggiornamento 15
2.1 Intelligenza Artificiale General-Purpose 17
2.2 Tecnologie Quantistiche 22
2.3 Tecnologie Operative 25

03 Aree di Ricerca e Innovazione 30
3.1 Area#1: Sicurezza dei Dati e Privacy 31
3.1.1 Ingegneria della Protezione dei Dati 31
3.1.2 Crittografia 33
3.1.3 Trusted Information Sharing 34
3.2 Area#2: Gestione delle Minacce Cibernetiche 35
3.2.1 Attacco e Difesa 35
3.2.2 Cyberthreat Intelligence 39
3.2.3 Gestione degli Incidenti e Operazioni di Sicurezza 41
3.2.4 Scienze Forensi Digitali 42
3.3 Area#3: Sicurezza del Software e delle Piattaforme 43
3.3.1 Sicurezza nello Sviluppo e Test del Software 43
3.3.2 Sicurezza nei Sistemi Operativi e Tecnologie di Virtualizzazione 47
3.3.3 Blockchain 48
3.4 Area#4: Sicurezza delle Infrastrutture Digitali 49
3.4.1 Hardware 49
3.4.2 Reti 51
3.5 Area#5: Aspetti della Società 53
3.5.1 Aspetti Umani 54
3.5.2 Aspetti Formativi 56
3.5.3 Aspetti Legali 56
3.6 Area#6: Aspetti di Governo 59
3.6.1 Aspetti Organizzativi 59
3.6.2 Gestione del Rischio 60
3.6.3 Standardizzazione 65
3.7. Panoramica delle Aree di R&I 68

04 Aree di R&I e EDT 76
4.1 EDT Rilevanti: Descrizioni 76
4.2 Proposta di Corrispondenza tra Subaree e EDT 82

Lista degli acronimi 85
Appendice 88

2026

Edizione aggiornata: novità di questa versione

Negli ultimi anni, tre tecnologie emergenti e dirompenti (*Emerging Disruptive Technologies*, EDT) hanno acquisito una rilevanza strategica particolare per l'Italia e l'Europa, in ragione delle loro implicazioni tecnologiche, etiche e operative, ovvero: i) l'Intelligenza Artificiale General-Purpose (*General-Purpose AI*, GPAI), introdotta come direzione emergente fondamentale dell'intelligenza artificiale, ii) le Tecnologie Quantistiche (*Quantum Technologies*, QT) e iii) le Tecnologie Operative (*Operational Technologies*, OT). Riconoscendo le sfide e le opportunità poste dalle citate EDT, questo aggiornamento arricchisce l'*Agenda di Ricerca e Innovazione 2023-2026* introducendo 1 nuovo argomento e 32 nuovi sottoargomenti, specificando inoltre ulteriormente gli aspetti rilevanti di 20 argomenti che spaziano lungo tutte e 6 le aree dell'Agenda.

Aggiornamenti delle sezioni:

1. La **Sezione 1** è stata complessivamente aggiornata; in particolare, questa versione introduce come nuovo passo metodologico la nozione di *sottoargomento* (Sezione 1.1 - Metodologia) che si affianca ad aree, subaree e argomenti, permettendo una maggiore specializzazione degli argomenti prioritari individuati. Ciò allo scopo di fornire un quadro più dettagliato che evidenzi le principali sfide aperte di ricerca e innovazione, fornendo punti focali per le indagini future. Inoltre, è stata aggiunta la Sezione 1.1.1 per evidenziare l'importanza della ricerca interdisciplinare.
2. La **Sezione 2** offre una visione d'insieme dell'aggiornamento, introducendo i nuovi sottoargomenti aggiunti per ciascuna delle tre EDT considerate. Per agevolare la lettura, questi sono raggruppati secondo classificazioni note.
3. Le **Sezioni 3 e 4** sono state riviste per riflettere gli aggiornamenti relativi ad aree, subaree, argomenti, sottoargomenti e EDT.

1 Introduzione

1 Introduzione

Vivere in un mondo digitale nel quale si interagisce con sistemi informativi e di comunicazione per ottimizzare le attività quotidiane rappresenta una straordinaria opportunità: la trasformazione digitale, infatti, abilita nuove modalità di lavoro, intrattenimento, viaggio e interazione. Di pari passo con questa positiva evoluzione delle nostre abitudini, però, sono emersi nuovi rischi di natura cibernetica, che continuano a evolversi e a moltiplicarsi. In particolare, è possibile individuare due macrocategorie di rischi che devono essere tempestivamente gestite: i rischi tecnologici e i rischi connessi ad attacchi cibernetici.

I rischi tecnologici sono legati al grado di affidabilità (*trustworthiness*) dei fornitori di prodotti hardware e software. Anche per il nostro Paese, la produzione globalizzata di componenti per la tecnologia dell'informazione ha creato un sempre maggior numero di dipendenze che, se venissero meno, causerebbero un'interruzione delle catene di approvvigionamento con conseguenze negative sulla nostra economia. Per fronteggiare questi rischi, è importante mettere in campo azioni che puntino a un rafforzamento dell'autonomia tecnologica strategica in ambito nazionale ed europeo, investendo sulla progettazione e sullo sviluppo di tecnologie affidabili, i cui anelli critici della catena di approvvigionamento si trovano su suolo italiano o in Paesi *like-minded*.

Riguardo agli attacchi cibernetici, il panorama delle minacce, sia in Italia che a livello mondiale, diventa di giorno in giorno più complesso e include svariati attori. Nello specifico, i singoli criminali cibernetici o i gruppi del cybercrimine organizzato conducono attività di hackeraggio per finalità economiche o di altro genere. Un esempio rilevante è rappresentato dagli attacchi DDoS, che continuano a costituire la minaccia più diffusa sia in Italia[1] che in Europa[2], tramite i quali i criminali cibernetici possono richiedere un riscatto in cambio dell'interruzione dell'attacco (*ransom DDoS o RDoS*), compromettere l'erogazione di servizi per danneggiare la concorrenza oppure utilizzare l'attacco come diversivo per facilitare altre attività dannose.

[1] Si veda il documento ACN - Relazione Annuale al Parlamento 2024 (acn.gov.it – PDF, 20MB).

[2] Si veda (in inglese) ENISA Threat Landscape 2025 (enisa.europa.eu).

Accanto a questi, il *ransomware* rimane lo strumento più diffuso nelle campagne malevole, caratterizzato da operatori sempre più professionalizzati e modelli *ransomware-as-a-service* che ne alimentano persistenza e diffusione. In aggiunta, le tensioni geopolitiche continuano ad avere ricadute rilevanti in campo cibernetico: attori statuali ostili possono realizzare, ad esempio, campagne di disinformazione online oppure compromettere la capacità dello Stato di garantire servizi essenziali, quali energia, acqua e trasporti, a cittadini e imprese. Tale scenario rende sempre più urgente il rafforzamento degli strumenti a difesa delle infrastrutture critiche nazionali.

Per fronteggiare efficientemente ed efficacemente i rischi cibernetici, è necessario che le organizzazioni pubbliche e private investano nello sviluppo di capacità collettive per la tutela della cybersicurezza, che rappresenta un dominio vasto, estremamente interdisciplinare e in costante evoluzione. Le diverse aree di conoscenza della sicurezza cibernetica, dalla sicurezza dei dati agli aspetti di governo, sono tra loro interconnesse e determinano uno scenario complesso che richiede sforzi coordinati da parte di tutti i soggetti coinvolti nell'ecosistema della ricerca e innovazione (R&I) per tenere il passo con i rapidi cambiamenti delle nuove minacce presenti nel cyberspazio. Inoltre, gli avanzamenti tecnologici procedono speditamente e un Paese sviluppato come l'Italia, come anche il resto d'Europa, gioca un ruolo primario nella ricerca, nello sviluppo e nell'industrializzazione di tecnologie emergenti, quali l'intelligenza artificiale, il cloud e le tecnologie quantistiche.

A questo proposito, l'Italia possiede straordinarie risorse di ricerca e sviluppo in tutti i settori: la sfida più grande è far sì che tali risorse remino nella stessa direzione, collaborando sulla base di un piano strategico che renda (e mantenga) il nostro Paese un posto *cyber*-sicuro per la società. All'interno della "Strategia Nazionale di Cybersicurezza 2022-2026 – Piano di Implementazione"[3] redatta dall'Agenzia per la Cybersicurezza Nazionale (ACN) e adottata dal Presidente del Consiglio dei Ministri, diverse misure puntano a favorire la ricerca e l'innovazione tecnologica tramite collaborazioni tra entità pubbliche e private in Italia. Le tre misure principali[4] sono:

1. **MISURA #53**

   Promuovere ogni iniziativa utile volta al rafforzamento dell'autonomia industriale e tecnologica dell'Italia, riguardo a prodotti e processi informatici di rilevanza strategica e a tutela degli interessi nazionali nel settore, anche valorizzando lo sviluppo di algoritmi proprietari nonché la ricerca e il conseguimento di nuove capacità crittografiche nazionali.

[3] Disponibile su Strategia Nazionale di Cybersicurezza 2022-2026 – Piano di Implementazione (acn.gov.it – PDF, 12MB).

[4] Inoltre, anche le **Misure #46**, **#47**, **#48** e **#49** risultano rilevanti, in quanto rafforzano il legame strategico tra ricerca, applicazione industriale, finanziamento e sviluppo delle infrastrutture.

2. MISURA #54

Favorire la ricerca e lo sviluppo, specialmente nelle nuove tecnologie, promuovendo l'inclusione dei principi di cybersicurezza e supportando, anche mediante finanziamenti, investimenti pubblici e privati e meccanismi di semplificazione, progetti di sicurezza cibernetica da parte del settore privato – con particolare riferimento alle startup e alle PMI innovative – e dei Centri di competenza e di ricerca attivi sul territorio nazionale.

3. MISURA #81

Contribuire attivamente, in ambito di Unione europea, all'individuazione delle priorità di ricerca e sviluppo per traguardare l'obiettivo dell'autonomia tecnologica UE in ambito digitale.

L'obiettivo finale è che il Governo italiano, l'industria, il mondo accademico e l'intera società collaborino su iniziative di R&I atte a consolidare ed espandere le capacità nazionali nell'ambito della cybersicurezza, creando un ecosistema che sia in grado di migliorare le linee di difesa del Paese contro i possibili attacchi di origine cibernetica perpetrati da attori ostili.

Tra le entità responsabili per l'implementazione di queste misure strategiche c'è la stessa ACN, in qualità di Autorità nazionale per la cybersicurezza a tutela degli interessi nazionali nello spazio cibernetico secondo il D.L. 82/2021[5], la quale agisce in stretta cooperazione con i Ministeri e i soggetti governativi italiani interessati, rafforzando l'impegno dell'Italia verso un ecosistema europeo coeso per l'innovazione digitale e della cybersicurezza.

In questo contesto, quale passo fondamentale per il raggiungimento degli obiettivi di Ricerca e Innovazione (R&I) stabiliti dalla Strategia[3], il presente documento costituisce la versione aggiornata dell'Agenda di Ricerca e Innovazione (di seguito anche "Agenda R&I" o "Agenda"), frutto della collaborazione tra l'ACN e il Ministero dell'Università e della Ricerca (MUR).

Dettagliando gli argomenti di ricerca prioritari per le attività di R&I nel campo della cybersicurezza, l'Agenda si propone di fornire:

- Una base di conoscenza condivisa da utilizzare per governare le attività di ricerca nazionali;
- Un fondamento per tutte le iniziative volte a promuovere e valorizzare i risultati della ricerca sulla cybersicurezza condotte dall'Agenzia[6], nonché per lo sviluppo di politiche di R&I in materia di cybersicurezza sia a favore del settore pubblico sia di quello privato.

[5] Si veda DECRETO-LEGGE 14 giugno 2021, n. 82 (normattiva.it).

[6] Tra questi, il programma annuale di finanziamento di borse di dottorato di ricerca, inaugurato nel 2024, e il Cyber Innovation Network (acn.gov.it) che supporta la creazione e la crescita delle startup italiane. Tale network è destinato a essere ampliato attraverso iniziative mirate rivolte agli enti di ricerca attivamente impegnati nella valorizzazione della ricerca applicata e nel trasferimento tecnologico dei relativi risultati.

Sebbene l'Agenda fornisca un quadro completo delle priorità di ricerca, le specifiche misure di incentivazione terranno conto anche di altri fattori, tra cui le strategie nazionali e continentali e l'evoluzione degli scenari di rischio. Ciò consente da un lato di avere una pianificazione della ricerca sul medio-lungo periodo, dall'altro di realizzare azioni di supporto mirate sulla base di esigenze concrete.

L'Agenda deve essere considerata un documento in evoluzione, soggetto a future iterazioni guidate dall'interazione con gli stakeholder dell'ecosistema R&I, nonché dall'emergere di nuovi possibili argomenti a livello nazionale e internazionale. L'edizione 2026, insieme a eventuali aggiornamenti successivi, mira a garantire che l'Agenda continui a riflettere gli aspetti e le sfide più rilevanti del panorama dinamico delle politiche, della ricerca e dell'innovazione in materia di cybersicurezza.

Il documento è principalmente rivolto a tutti gli attori che operano direttamente o beneficiano della ricerca sulla cybersicurezza, sia nel settore pubblico (ad esempio, ricercatori afferenti a università, enti pubblici di ricerca, consorzi di ricerca, nonché amministrazioni e organizzazioni pubbliche italiane, europee e internazionali operanti nel settore della cybersicurezza) sia nel settore privato (ricercatori industriali affiliati ad aziende, imprese e associazioni industriali attive nel campo della cybersicurezza).

1.1 Metodologia

La predisposizione dell'Agenda di R&I, sin dalla sua prima edizione nel 2023, è il risultato di un'attività di *desktop research* congiunta cui hanno preso parte articolazioni interne ad ACN e il MUR; a seguire sono stati raccolti riscontri da membri del Laboratorio Nazionale di Cybersecurity del CINI e dal comitato tecnico-scientifico dell'ACN, nonché da selezionati ricercatori che hanno fornito un contributo in fase di perfezionamento del documento. Come esplicitato nel seguito, sono state prese in considerazione come fonti le direzioni di ricerca e le politiche sulla cybersicurezza individuate in Italia e in Europa; sono stati, inoltre, esaminati molti riferimenti internazionali su aree specifiche.

La metodologia di costruzione dell'Agenda è essenzialmente articolata in tre passi:

1. *l'identificazione delle* ***aree*** tramite raggruppamento e indicazione delle corrispondenze con le iniziative di classificazione esistenti nel dominio di conoscenza della cybersicurezza[7];

[7] Il dettaglio della relazione tra le aree di R&I identificate e le iniziative di classificazione della conoscenza della cybersicurezza considerate si trova in appendice.

2. *la definizione delle **subaree** e dei relativi **argomenti*** tramite la valutazione sia delle priorità italiane ed europee[8] sia di riferimenti rilevanti sul piano internazionale;
3. *l'ulteriore specificazione di **sottoargomenti*** ritenuti critici o emergenti, che specializzano ulteriormente gli argomenti prioritari identificati, per evidenziare le principali sfide aperte della ricerca e fornire punti focali per l'investigazione futura.

Inoltre, l'Agenda considera gli orientamenti strategici delineati dai quadri normativi europei in materia di cybersicurezza, quali la Direttiva NIS2[9], la Direttiva CER[10] e il Cyber Resilience Act (CRA)[11], esplicitamente richiamati più avanti nel documento come riferimenti contestuali rilevanti. Questi strumenti introducono requisiti concreti in materia di cybersicurezza che influenzano la progettazione, lo sviluppo e la gestione del ciclo di vita delle tecnologie digitali[12]. Tali requisiti si riflettono sulle priorità di ricerca individuate dall'Agenda.

[8] Per quanto riguarda le priorità nazionali, un riferimento fondamentale è costituito dal Programma Nazionale per la Ricerca (PNR) 2021-2027 (mur.gov.it - PDF,13.6MB) del MUR, in particolare l'allegato esteso sulla sicurezza dei sistemi sociali (mur.gov.it - PDF,1.9MB). Per quanto riguarda le priorità europee, sono state considerate le indicazioni contenute nell'ultima versione del Research and Innovation Brief - Annual Report on Cybersecurity Research and Innovation Needs and Priorities (europa.eu) e in altri riferimenti dell'Agenzia dell'Unione Europea per la Cibersicurezza (ENISA).

[9] Direttiva (UE) 2022/2555 del Parlamento europeo e del Consiglio, del 14 dicembre 2022, relativa a misure per un livello comune elevato di cibersicurezza nell'Unione (NIS2). Gazzetta Ufficiale dell'Unione Europea – EUR-Lex (europa.eu).

[10] Direttiva (UE) 2022/2557 del Parlamento europeo e del Consiglio del 14 dicembre 2022 del Parlamento europeo e del Consiglio relativa alla resilienza dei soggetti critici (CER). Gazzetta Ufficiale dell'Unione Europea – EUR-Lex (europa.eu).

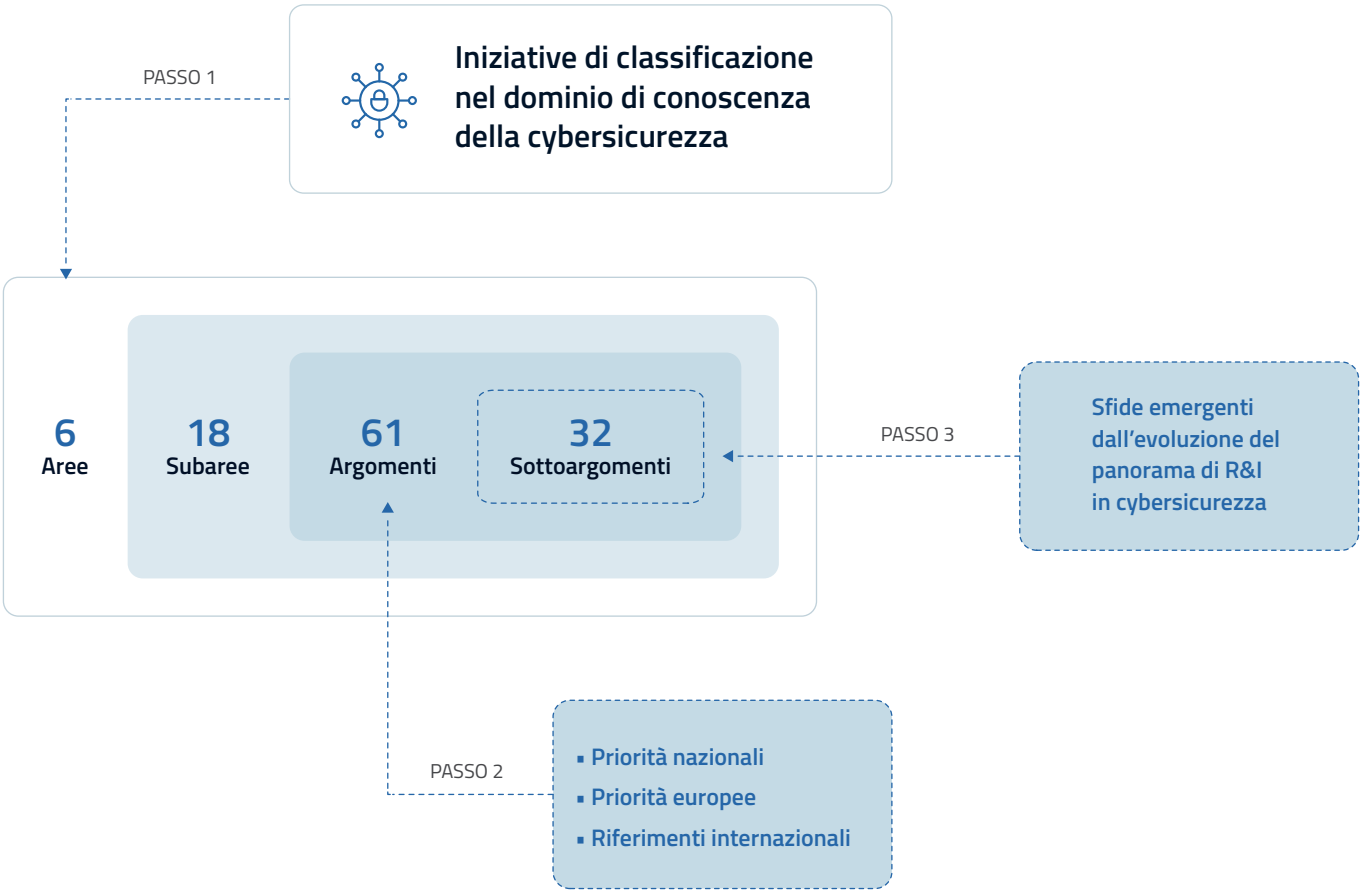
[11] Regolamento (UE) 2024/2847 del Parlamento europeo e del Consiglio, del 23 ottobre 2024, recante disposizioni orizzontali in materia di cybersicurezza per i prodotti con elementi digitali (Cyber Resilience Act, CRA). Gazzetta Ufficiale dell'Unione Europea – EUR-Lex (europa.eu).

[12] Sebbene il CRA non si concentri su tecnologie specifiche, i suoi requisiti orizzontali si applicano in modo ampio ai prodotti digitali, sottolineando l'importanza di anticipare i futuri rischi di cybersicurezza e di adottare misure allo stato dell'arte nello sviluppo dei prodotti, rafforzando così il valore dell'allineamento della ricerca con l'evoluzione delle normative.

La Figura 1-1 illustra la metodologia descritta, riportando inoltre una quantificazione delle aree, delle subaree, degli argomenti e dei sottoargomenti individuati in questa versione, nello specifico:

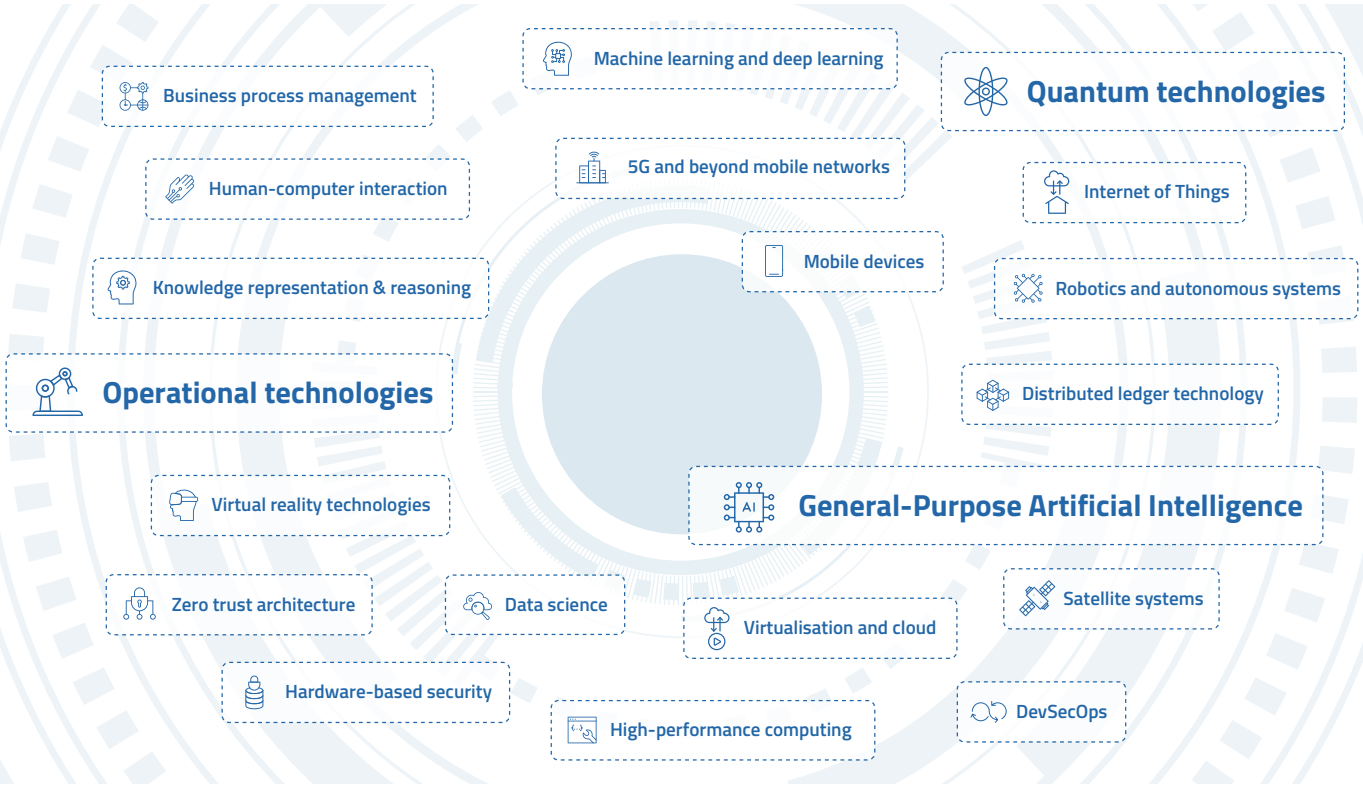
- 6 aree;
- 18 subaree;
- 61 argomenti;
- 32 sottoargomenti.

Figura 1-1: Metodologia adottata.



Inoltre, è stata identificata una lista di cosiddette ***Emerging and Disruptive Technology*** **(EDT)**, selezionate come rilevanti per la cybersicurezza in generale e utili a indirizzare lo studio degli argomenti di ricerca proposti[13].

[13] Per la lista completa delle EDT e le loro definizioni, si veda la Sezione 4.1.



1.1.1 IMPORTANZA DELLA RICERCA MULTIDISCIPLINARE

Riconoscendo che la cybersicurezza si estende oltre i confini strettamente tecnici, molti dei temi trattati in questo documento[14] si concentrano sulle implicazioni etiche, giuridiche e sociali che accompagnano la pervasività delle infrastrutture digitali nella società. Uno degli obiettivi di tale inclusione è, infatti, quello di incentivare sforzi di ricerca coordinati e prospettive che provengano non solo da ambiti tecnici come l'informatica, ma anche da discipline quali l'etica, il diritto, la psicologia, l'economia e la sociologia.

ACN e MUR riconoscono il valore della ricerca interdisciplinare come fattore essenziale per bilanciare il progresso tecnologico con le implicazioni sociali, promuovendo lo sviluppo, la diffusione e l'integrazione sicura e responsabile delle tecnologie nei diversi settori.

[14] Nello specifico, tutti gli argomenti che afferiscono alle aree 5 e 6 nelle Sezioni 3.5 e 3.6.

1.2 Struttura del Documento

Il resto del documento è organizzato come segue.

- La Sezione 2 offre una visione d'insieme dell'aggiornamento, elencando i nuovi sottoargomenti introdotti per ciascuna delle tre EDT considerate. Per facilitare la lettura, tali sottoargomenti sono raggruppati secondo classificazioni note.
- La Sezione 3 descrive le aree, le subaree, gli argomenti e i sottoargomenti di R&I identificati.
- La Sezione 4 descrive le EDT rilevanti e le mette in corrispondenza con le subaree di R&I.
- Lista degli acronimi.
- Appendice.

2 Panoramica dell'aggiornamento

2

Panoramica dell'aggiornamento

Come riassunto nella Sezione 1.1, questo aggiornamento include 32 nuovi sottoargomenti[15] definiti a partire da tre *Emerging and Disruptive Technologies* (EDT) che, negli ultimi anni, hanno acquisito una rilevanza strategica particolare per l'Italia e l'Europa, in virtù delle loro implicazioni tecnologiche, etiche e operative in molteplici settori. Queste tecnologie, descritte nel seguito, sono: l'intelligenza artificiale *general-purpose*, le tecnologie quantistiche e le tecnologie operative e si distinguono per il loro impatto trasformativo che comporta l'urgenza di promuovere attività di ricerca e misure volte a gestire i rischi e le opportunità a esse associati.

1. L'**intelligenza artificiale *general-purpose* (*General-purpose Artificial Intelligence*, GPAI)** rappresenta una frontiera in rapida evoluzione dell'intelligenza artificiale (IA), che consente ai sistemi di IA di svolgere compiti eterogenei in molteplici ambiti con adattabilità e versatilità (diversamente da quella c.d. *"task-specific"*, tipicamente addestrata per svolgere un singolo compito più specifico). Basata su architetture di *deep learning*, la GPAI abilita funzionalità avanzate come la generazione di contenuti, la capacità di pianificazione e l'automazione intelligente[16]. L'emergere di questa tecnologia ha comportato importanti incrementi di efficienza in molteplici settori, incluso quello della cybersicurezza, generando al contempo nuove sfide legate all'affidabilità e alla gestione dei rischi, inclusi aspetti etici e normativi connessi alla *safety* e alla fiducia. A partire dal 2022, il rapido sviluppo e la diffusione su larga scala dei modelli GPAI, addestrati principalmente da fornitori di servizi tecnologici globali con sede esterna all'Unione europea, hanno suscitato crescenti preoccupazioni riguardo alla dipendenza tecnologica, alla frammentazione normativa e alla sovranità digitale. Tali dinamiche hanno portato la governance della GPAI al centro delle agende strategiche e geopolitiche europee. In questo scenario, l'AI Act[17] europeo ha assunto un ruolo centrale.

[15] Inoltre, 9 argomenti sono stati aggiornati.

[16] Per una definizione estesa di GPAI, si veda la Sezione 4.1 e in particolare la lista completa delle EDT riportata in Tabella 4-1.

[17] Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale (AI Act). Gazzetta Ufficiale dell'Unione Europea – EUR Lex (europa.eu).

Inizialmente concepito come quadro normativo orizzontale per la regolamentazione dei sistemi di IA, il suo sviluppo ha coinciso con la rapida affermazione delle tecnologie GPAI, rendendo necessaria l'inclusione di disposizioni specifiche per tali modelli. Tali misure mirano a gestire i rischi sistemici, promuovere un'innovazione responsabile e tutelare la sovranità digitale e la competitività dell'Europa.

2. Le **tecnologie quantistiche (*Quantum Technologies*, QT)** sono destinate a rivoluzionare i paradigmi tradizionali della computazione, introducendo al tempo stesso opportunità straordinarie e gravi minacce per la sicurezza. L'avvento del calcolo quantistico mette a rischio l'integrità degli schemi crittografici classici, rendendo necessario lo sviluppo, la standardizzazione e l'evoluzione della *Post-Quantum Cryptography* (PQC). In parallelo, la fisica quantistica può essere sfruttata per sviluppare, nell'ambito della cosiddetta crittografia quantistica, nuove tecniche per la confidenzialità, quali, ad esempio, la distribuzione di chiavi quantistiche (*Quantum Key Distribution*, QKD) oppure nuove tecniche di rilevamento di attacchi cibernetici basate su sensori quantistici ad altissima precisione. Dato il passo sempre più veloce della ricerca e degli investimenti nazionali ed europei in materia, è di primaria importanza assicurare la resilienza cibernetica nel dominio delle tecnologie quantistiche.
3. Le **tecnologie operative (*Operational Technologies*, OT)** svolgono un ruolo fondamentale nelle infrastrutture nazionali e sono cruciali per settori come l'energia, il manifatturiero, i trasporti e la difesa, nei quali garantire la sicurezza è necessario al fine di prevenire interruzioni nei servizi essenziali. A differenza delle tecnologie emergenti come l'intelligenza artificiale *general-purpose* o le tecnologie quantistiche, il dominio OT è ben consolidato, ma in rapida evoluzione. Inoltre, a causa dell'incremento della digitalizzazione e la crescente interconnessione con i dispositivi *Internet of Things (IoT)* e *Industrial Internet of Things (IIoT)*, le sfide di cybersicurezza aumentano sia in quantità che in sofisticazione. La crescente complessità degli ambienti OT, unita ai vincoli di operatività, spesso di tipo *real-time*, alla presenza di sistemi "legacy" e ai rigorosi requisiti di sicurezza fisica (*safety*), determina sfide peculiari per la sicurezza OT, rendendola una priorità strategica da affrontare con urgenza. L'importanza di queste sfide è sempre più riconosciuta nei quadri normativi europei, come le direttive NIS2 e CER, dove si affrontano i rischi specifici degli ambienti OT nel contesto della protezione delle infrastrutture critiche, nonché quelli relativi a sicurezza e *safety* industriale. La posizione di rilievo dell'Italia nella produzione ed esportazione[18] di macchine utensili, robotica e automazione rafforza ulteriormente la rilevanza strategica dei sistemi OT, rendendo la loro protezione una questione di interesse economico nazionale.

Le sezioni seguenti analizzano queste tre tecnologie in relazione alla ricerca sulla cybersicurezza, illustrando al contempo le motivazioni strategiche alla base della loro selezione e mettendone in evidenza la rilevanza rispetto alle priorità nazionali ed europee.

[18] Si veda la Sezione 2.3.

I sottoargomenti individuati approfondiscono aspetti chiave di 20 argomenti, distribuiti in 12 subaree all'interno delle 6 aree individuate, e saranno descritti in dettaglio nella Sezione 3, nell'ambito delle rispettive aree, subaree e argomenti, in coerenza con la tassonomia dell'Agenda. Sebbene questo aggiornamento consideri GPAI, QT e OT come EDT distinte, alcuni sottoargomenti possono avere implicazioni trasversali alle EDT stesse[19].

2.1 Intelligenza artificiale *general-purpose* (*General-purpose* AI, GPAI)

Il 2022 ha segnato una svolta nello sviluppo dell'IA, con il rilascio al pubblico e l'immissione sul mercato dei c.d. modelli di linguaggio di grandi dimensioni (*large language models*, LLM), consolidando il concetto di *General-Purpose Artificial Intelligence* (GPAI)[20], ovvero una categoria di modelli di IA in grado di svolgere un'ampia gamma di compiti in diversi ambiti, inclusa la generazione di contenuti testuali e video, con una versatilità senza precedenti. Da allora, la rapida evoluzione e la crescente disponibilità di tali modelli hanno avuto un impatto significativo sui mercati e sulla società, integrandosi progressivamente nelle infrastrutture digitali e aprendo nuove opportunità anche nel rafforzamento delle pratiche di cybersicurezza, ad esempio, nel miglioramento del rilevamento delle minacce, nell'automazione avanzata dei meccanismi di difesa e nell'analisi dei rischi. Tali progressi, tuttavia, comportano anche sfide complesse. Tra queste, si annovera, in particolare, la mancanza di trasparenza rispetto ai dati utilizzati per il loro addestramento, tipicamente massivi ed eterogenei, con impatti diretti sulla natura e l'affidabilità dei contenuti generati, così come la limitata comprensione di quali obiettivi i modelli stiano effettivamente ottimizzando durante l'addestramento, spesso in assenza di criteri espliciti da parte dei progettisti. Queste incertezze pongono implicazioni rilevanti in termini di sicurezza e affidabilità delle applicazioni reali, con la possibilità di introdurre rischi sistemici.

La centralità della GPAI nelle politiche pubbliche europee e nazionali riflette la crescente consapevolezza del suo impatto strategico. A livello europeo, l'AI Act, concepito inizialmente nel 2021 come quadro normativo orizzontale per garantire lo sviluppo sicuro e affidabile dei sistemi di IA immessi sul mercato nell'Unione, ha affrontato temi chiave, quali la gestione del rischio, la trasparenza e la tutela dei diritti fondamentali.

19 Ad esempio, l'uso malevolo della GPAI per sfruttare vulnerabilità di sistema può impattare altre tecnologie, tra le quali le OT.

20 Il termine è stato introdotto nell'AI Act europeo per descrivere modelli applicabili a una molteplicità di scopi distinti, che possono comportare rischi sistemici a causa della loro scala, delle loro capacità e dell'ampiezza dei contesti in cui vengono impiegati.

In fase di negoziazione, parallelamente all'ascesa delle tecnologie GPAI, il regolamento è stato esteso includendo, nella sua versione finale, disposizioni specifiche per la GPAI, con l'obiettivo di mitigare i rischi sistemici, abilitare un'innovazione responsabile e preservare la sovranità digitale dell'Europa. In Italia, la *Strategia per l'Intelligenza Artificiale 2024–2026*[21], pubblicata nel luglio 2024, promuove la creazione di un ecosistema per lo sviluppo sicuro e affidabile dell'AI, ritenuto strategico per il posizionamento del Paese all'interno dello spazio europeo della ricerca e innovazione. In tale prospettiva, nel settembre 2025, è stato approvato un provvedimento legislativo sull'intelligenza artificiale[22], volto a definire i principi per la ricerca, lo sviluppo, la sperimentazione e l'adozione responsabile[23] di sistemi e modelli di IA, includendo disposizioni per il coordinamento tra autorità competenti, obblighi di trasparenza, meccanismi di monitoraggio e sostegno all'innovazione in coerenza con l'AI Act europeo.

Con riferimento specifico ai modelli GPAI, l'AI Act prevede che i fornitori di modelli con potenziale rischio sistemico debbano effettuare valutazioni e implementare misure adeguate per mitigare tali rischi. In questo contesto, l'Ufficio europeo per l'IA (*EU AI Office*) ha avuto un ruolo centrale nel facilitare la creazione di un codice di condotta (denominato *"Code of Practice", CoP*) per la GPAI[24], frutto di uno sforzo coordinato che ha coinvolto circa 1000 partecipanti provenienti da industria, mondo accademico, società civile e organizzazioni di tutela dei diritti di proprietà intellettuale. In particolare, il CoP dedica una sezione specifica alla sicurezza delle tecnologie GPAI, descrivendo le misure e le procedure che i firmatari intendono adottare per mantenere i rischi sistemici dei propri modelli entro soglie accettabili, tra cui: i) *Valutazione dei rischi sistemici*, basata su una tassonomia dei rischi sistemici associati ai modelli GPAI, inclusi quelli derivanti da possibili usi malevoli per la realizzazione di attacchi cyber, ii) *Mitigazione tecnica dei rischi sistemici*, concernente l'adozione di misure tecniche volte a ridurre i rischi identificati e a rafforzare la cybersicurezza dei modelli GPAI[25]. Pur mirando principalmente a favorire la conformità normativa, molte delle misure previste dal *CoP* si collocano attualmente alla frontiera della ricerca, poiché lo sviluppo di soluzioni robuste in tali ambiti è tuttora in corso.

Nelle sezioni seguenti e sintetizzato nella Tabella 2-1, viene presentato l'elenco dei sottoargomenti di R&I identificati come prioritari per la GPAI, organizzati secondo i due orientamenti individuati all'interno del CoP.

[21] Sviluppata da un comitato di esperti nazionali, con il supporto di una segreteria tecnica istituita presso l'Agenzia per l'Italia Digitale (AgID). Documento disponibile su Strategia Italiana per l'IA 2024–2026 (agid.gov.it - PDF 557KB)

[22] Legge 23 settembre 2025, n. 132, "*Disposizioni e deleghe al Governo in materia di intelligenza artificiale*".

[23] Un orientamento metodologico per l'approvvigionamento, lo sviluppo e l'adozione di sistemi di intelligenza artificiale nella Pubblica Amministrazione è fornito nel documento "*Linee guida per l'adozione dell'Intelligenza Artificiale nella Pubblica Amministrazione*", redatto dai membri del gruppo di coordinamento per il Piano Triennale per l'Informatica nella Pubblica Amministrazione 2024–2026, coordinato da AgID e di cui fa parte anche l'ACN.

[24] Si veda (in inglese) General-Purpose AI Code of Practice (europa.eu)

[25] Il *Code of Practice* include anche misure su trasparenza, diritto d'autore e governo dei rischi sistemici, con riferimento all'adozione di politiche interne, strutture di governance, nonché processi per la tracciabilità degli incidenti. Tali aspetti sono stati esclusi dalla classificazione delle sfide di ricerca presentata nel seguito.

Le priorità di ricerca individuate nelle Sezioni 2.1.1 e 2.1.2 affrontano specificamente le sfide tecniche e scientifiche introdotte dai sistemi GPAI, in relazione alla loro caratteristica *general-purpose* e alla loro applicabilità trasversale ai diversi domini[26]. Alcune di queste sfide riguardano in particolare i rischi sistemici che emergono dall'impiego della GPAI in contesti reali.

Al contrario, la Sezione 2.1.3 esplora alcune applicazioni della GPAI a supporto della cybersicurezza, che possono includere implementazioni specializzate o *fine-tuned*, che tuttavia continuano a sfruttare le capacità *general-purpose* dei modelli originari.

Tabella 2-1: Sottoargomenti relativi all'intelligenza artificiale *general-purpose*.

Classificazione (CoP GPAI/Applicazioni)	Sottoargomento		Subarea
VALUTAZIONE DEI RISCHI	GAI - 6.2.2.1	Nuove metriche di *safety* e gestione del bilanciamento tra *safety* e prestazioni	6.2 Gestione del rischio
	GAI - 6.2.4.1	Migliorare la comprensione delle capacità (*capabilities*) dei modelli e la loro interpretabilità	6.2 Gestione del rischio
	GAI - 6.2.4.2	Sfide nella valutazione (*evaluation*) dei modelli GPAI	6.2 Gestione del rischio
MITIGAZIONE TECNICA DEI RISCHI	GAI - 1.1.1.1	Bilanciamento tra privacy e prestazioni dei modelli GPAI nelle fasi di addestramento e inferenza	1.1 Ingegneria della Protezione dei Dati
	GAI - 2.1.1.1	Contrasto all'uso malevolo dei modelli GPAI	2.1 Attacco e Difesa
	GAI - 2.1.3.1	Difesa contro le minacce avversarie	2.1 Attacco e Difesa
	GAI - 5.1.2.1	Miglioramento della percezione del rischio nell'uso dei modelli GPAI	5.1 Aspetti Umani

[26] Sebbene queste sezioni si concentrino sulle sfide legate alla GPAI, la necessità di solide metodologie di valutazione e mitigazione dei rischi si estende anche ai modelli di intelligenza artificiale tradizionali, inclusi quelli predittivi e *task-specific*.

Classificazione (CoP GPAI/Applicazioni)	Sottoargomento		Subarea
	GAI - 5.1.3.1	Prevenzione e mitigazione degli attacchi di ingegneria sociale e delle minacce all'identità digitale supportati da GPAI	5.1 Aspetti Umani
	GAI - 5.3.1.1	Codifica dei principi etici nell'allineamento (*alignment*) dei modelli	5.3 Aspetti Legali
	GAI - 5.3.3.1	Regolamentazione efficace e dinamica per la *governance* della GPAI	5.3 Aspetti Legali
	GAI - 6.2.4.3	Approcci robusti di *safety fine-tuning* per la GPAI	6.2 Gestione del rischio
	GAI - 6.2.4.4	Controllo dell'autonomia dei sistemi a singolo e multi-agente	6.2 Gestione del rischio
	GAI - 6.3.2.1	Interoperabilità tra gli standard di sicurezza per la GPAI e gli standard settoriali	6.3 Standardizzazione
APPLICAZIONI DI GPAI PER LA CYBERSICUREZZA	GAI - 2.2.1.1	Uso della GPAI per la rilevazione e la mitigazione delle capacità di evasione (*evasion*) del *malware*	2.2 Cyberthreat Intelligence
	GAI - 2.2.1.2	Uso della GPAI per supportare la *cyber threat intelligence* e la rilevazione degli incidenti	2.2 Cyberthreat Intelligence
	GAI - 3.1.2.1	Migliorare la sicurezza del codice generato con il supporto della GPAI	3.1 Sicurezza nello sviluppo e test del software
	GAI - 3.1.4.1	Uso della GPAI per il rilevamento e il *patching* delle vulnerabilità	3.1 Sicurezza nello sviluppo e test del software
	GAI - 4.1.1.1	Uso della GPAI per il rilevamento di *Hardware Trojan* (HT)	4.1 Hardware

2.1.1 VALUTAZIONE DEI RISCHI

I sottoargomenti di questa sezione mirano a migliorare i metodi di valutazione del rischio per la GPAI, affrontando aspetti come la comprensione delle capacità (*capabilities*) dei modelli, il perfezionamento delle valutazioni e il bilanciamento tra *safety* e prestazioni.

È centrale sviluppare metriche di *safety* più complete che, oltre alla limitazione degli output dannosi, prendano in considerazione il grado di trasparenza, adattabilità e affidabilità, e anche la capacità di gestire i compromessi tra *safety* e prestazioni (cfr. sottoargomento **GAI-6.2.2.1**). Serve inoltre rafforzare la comprensione delle capacità dei modelli e la loro interpretabilità (cfr. sottoargomento **GAI-6.2.4.1**), distinguendo tra abilità intrinseche e supporti esterni. Infine, per rendere più affidabili le valutazioni, è fondamentale affrontare limiti quali l'instabilità rispetto alle variazioni dei *prompt* (*prompt-sensitivity*), la contaminazione dei test set e i *bias* delle valutazioni umane (cfr. sottoargomento **GAI-6.2.4.2**).

2.1.2 MITIGAZIONE TECNICA DEI RISCHI

Le misure tecniche sono fondamentali per proteggere la GPAI dalle minacce emergenti. In questa direzione, i sottoargomenti individuati riguardano la sicurezza delle informazioni utilizzate dai modelli, la difesa di questi da attacchi avversari, l'allineamento dei comportamenti dei modelli ai valori umani e la prevenzione dell'uso improprio a fini malevoli. È inoltre essenziale affrontare le sfide regolatorie e garantire l'integrazione tra futuri standard orizzontali di sicurezza per la GPAI e le normative settoriali.

Una priorità chiave è lo sviluppo di tecniche che bilancino privacy ed efficienza dei modelli durante l'addestramento e l'inferenza (cfr. sottoargomento **GAI-1.1.1.1**). È altrettanto importante rafforzare la resilienza dei modelli, con particolare attenzione alla robustezza contro manipolazioni e trasferibilità degli attacchi tra un modello e l'altro (cfr. sottoargomento **GAI-2.1.3.1**) e alla robustezza delle tecniche di *safety fine-tuning*, ovvero quelle tecniche che mirano a rimuovere i comportamenti potenzialmente insicuri dei modelli (cfr. sottoargomento **GAI-6.2.4.3**). Un ulteriore ambito riguarda l'allineamento dei modelli ai principi etici umani (cfr. sottoargomento **GAI-5.3.1.1**), nonché lo sviluppo di misure per arginare i rischi legati all'autonomia, a oggi sempre crescente, di questi sistemi, ovvero il disallineamento rispetto agli obiettivi, i comportamenti ingannevoli e le dinamiche multi-agente (cfr. sottoargomento **GAI-6.2.4.4**). Sul versante dell'uso da parte degli utenti, è cruciale mitigare l'eccessiva fiducia da parte degli utenti negli output dei modelli, promuovendo un uso critico dei risultati generati (cfr. sottoargomento **GAI-5.1.2.1**). Occorre inoltre contrastare l'uso malevolo della GPAI, incluse le applicazioni in attacchi con *Hardware Trojan* (HT), canali laterali (*side channels*) e malware (cfr. sottoargomento **GAI-2.1.1.1**), così come la disinformazione, gli attacchi di ingegneria sociale e le minacce all'identità digitale (cfr. sottoargomento **GAI-5.1.3.1**). Infine, sono fondamentali le attività di ricerca verso una regolamentazione efficace e dinamica per la *governance* della GPAI (cfr. sottoargomento **GAI-5.3.3.1**), insieme all'integrazione tra standard orizzontali di sicurezza e normative settoriali per garantirne l'interoperabilità (cfr. sottoargomento **GAI-6.3.2.1**).

2.1.3. APPLICAZIONI DI GPAI PER LA CYBERSICUREZZA

I sottoargomenti di questa sezione esplorano alcune direzioni promettenti in cui la GPAI può rafforzare la cybersicurezza. In particolare, è rilevante indagare il loro uso per la rilevazione e mitigazione delle capacità di evasione (*evasion*) del malware (cfr. sottoargomento **GAI-2.2.1.1**), così come il potenziale di questa nel supportare i flussi operativi di *Cyber Threat Intelligence* (CTI), migliorando le capacità di rilevamento e risposta automatizzata (cfr. sottoargomento **GAI-2.2.1.2**). Ulteriori direzioni di rilievo includono l'impiego della GPAI per l'analisi della sicurezza di codice, binari e *firmware*, automatizzando il rilevamento delle vulnerabilità, l'identificazione di *exploit* e la generazione di *patch* (cfr. sottoargomento **GAI-3.1.4.1**), così come per la rilevazione di potenziali minacce nascoste nei progetti hardware, in particolare, nei componenti di terze parti (cfr. sottoargomento **GAI-4.1.1.1**). Inoltre, per affrontare le sfide legate alla generazione di codice assistita da GPAI, è necessario sviluppare strumenti che guidino i modelli verso la produzione di codice sicuro (cfr. sottoargomento **GAI-3.1.2.1**). I rischi di cybersicurezza indotti dall'uso della GPAI, come il possibile impiego per creare malware o supportare attacchi di social engineering automatizzati, o l'introduzione di nuove vulnerabilità, quali le minacce avversarie, sono state affrontate nelle Sezioni 2.1.1 e 2.1.2.

2.2 Tecnologie quantistiche (*Quantum technologies*, QT)

I principi della meccanica quantistica, quali la sovrapposizione e l'*entanglement*, possono essere sfruttati per migliorare le funzionalità e le prestazioni delle infrastrutture digitali classiche, dando origine alle cosiddette tecnologie quantistiche. La Commissione europea ha adottato la "Strategia per l'Europa Quantistica"[27], finalizzata a trasformare l'Europa in una potenza quantistica, promuovendo un ecosistema resiliente e sovrano, capace di favorire la crescita delle start-up e trasformare la scienza pionieristica in applicazioni pronte per il mercato, mantenendo nel contempo la propria leadership scientifica. Inoltre, il Ministero dell'Università e della Ricerca, insieme al Ministero delle Imprese e del Made in Italy, al Ministero della Difesa, l'ACN, il Ministero degli Affari Esteri e della Cooperazione Internazionale e la Presidenza del Consiglio dei Ministri (Dipartimento per la trasformazione digitale) ha pubblicato la "Strategia italiana per le tecnologie quantistiche"[28], allo scopo di identificare i punti di forza e favorire la crescita dell'ecosistema italiano della R&I nel campo, in linea con le priorità strategiche europee. Dal punto di vista dell'ACN, le tecnologie quantistiche rappresentano sia una minaccia che un'opportunità per la cybersicurezza e pongono sfide di ricerca di estrema rilevanza.

[27] Disponibile su Strategia per l'Europa quantistica (europa.eu).

[28] Si veda il documento Strategia Italiana per le tecnologie quantistiche (mur.gov.it – PDF, 8.8MB).

In Tabella 2-2 è riportata la lista degli argomenti e sottoargomenti di R&I associati alle tecnologie quantistiche. In particolare, traendo spunto dalla strategia continentale e da quella nazionale, ogni argomento e sottoargomento è stato classificato rispetto a uno dei tre principali filoni sulle tecnologie quantistiche, ovvero *calcolo quantistico, comunicazioni quantistiche e sensoristica quantistica*.

Tabella 2-2: Argomenti e sottoargomenti relativi alle tecnologie quantistiche.

Classificazione (Pilastri QT)	Argomento/sottoargomento		Subarea
CALCOLO QUANTISTICO	**Argomento#1.2.1**	Primitive, algoritmi e protocolli per PQC	**1.2 Crittografia**
	QT - 3.1.2.1	Messa in sicurezza dell'ingegneria del software quantistico	**3.1 Sicurezza nello sviluppo e test del software**
	QT - 6.1.2.1	Metodologie per la transizione alla crittografia basata su QT	**6.1 Aspetti organizzativi**
COMUNICAZIONE QUANTISTICA	**Argomento#1.2.2**	Crittografia quantistica	**1.2 Crittografia**
	QT - 4.1.1.1	Messa in sicurezza delle installazioni di QKD	**4.1 Hardware**
	QT - 4.2.4.1	Evoluzione della QCI verso l'Internet quantistico	**4.2 Reti**
	QT - 6.3.2.1	Favorire la standardizzazione dei protocolli di QKD	**6.3 Standardizzazione**
	QT - 6.3.2.2	Ampliamento della standardizzazione dei protocolli di PQC	**6.3 Standardizzazione**
SENSORISTICA QUANTISTICA	**QT - 2.1.1.1**	Sensoristica quantistica per migliorare il rilevamento degli attacchi	**2.1 Attacco e Difesa**
	QT - 4.1.3.1	Modelli per il rilevamento quantistico sicuro	**4.1 Hardware**

2.2.1 CALCOLO QUANTISTICO

Il calcolo quantistico è la disciplina che sfrutta i principi della meccanica quantistica per lo sviluppo e l'implementazione di calcolatori estremamente più potenti di quelli classici attuali. Al momento della stesura di questo documento, i computer quantistici sono ancora a uno stadio iniziale, con poche installazioni presenti, rendendo complicato testarne a fondo le capacità. La notevole efficienza attesa dalle future generazioni di computer quantistici mette a rischio gli schemi crittografici asimmetrici classici, basati su determinate assunzioni di complessità computazionale. Diventa prioritario, quindi, creare primitive *post-quantum* che possano resistere agli attacchi di un computer quantistico, così come l'esigenza di pianificare la transizione delle infrastrutture digitali dalle *suite* crittografiche attuali a quelle *post-quantum* (cfr. **Argomento#1.2.1** e sottoargomento **QT-6.1.2.1**). Inoltre, con l'avvento di questa nuova generazione di computer, la sicurezza dei programmi ingegnerizzati per processori quantistici deve essere attentamente considerata (cfr. sottoargomento **QT-3.1.2.1**).

2.2.2 COMUNICAZIONI QUANTISTICHE

Le comunicazioni quantistiche sfruttano stati e risorse quantistiche per sviluppare nuovi protocolli di comunicazione in grado di abilitare livelli di confidenzialità e traffico scambiato mai raggiunti prima. A tale riguardo, alla data di scrittura del presente documento, la generazione e distribuzione di chiavi quantistiche (cfr. **Argomento#1.2.2**) rappresenta lo stato dell'arte delle comunicazioni quantistiche, nonostante la necessità di un maggiore studio sulla standardizzazione e sulla validazione del livello di sicurezza di installazioni su ampia scala (cfr. sottoargomenti **QT-4.1.1.1** e **QT-6.3.2.1**). Inoltre, è importante favorire la ricerca sui passi evolutivi necessari allo sviluppo futuro dell'Internet quantistico, che connetterà computer e sensori quantistici (cfr. sottoargomento **QT-4.2.4.1**).

2.2.3 SENSORISTICA QUANTISTICA

La sensoristica quantistica sfrutta proprietà quantistiche avanzate per implementare sensori ad altissima precisione. Dal punto di vista della cybersicurezza, da un lato questi sensori potrebbero essere utili a migliorare le misure di difesa contro gli attacchi cibernetici (cfr. sottoargomento **QT-2.1.1.1**). D'altro canto, dal momento che diversi sensori potrebbero presentare diversi stadi di sviluppo, è necessario assicurare che sistemi complessi di sensori quantistici siano dispiegati in maniera sicura (cfr. sottoargomento **QT-4.1.3.1**).

2.3 Tecnologie operative *(Operational Technologies, OT)*

Con l'aumento dell'utilizzo delle tecnologie OT, sia nel settore manifatturiero che nelle infrastrutture critiche, la necessità di proteggerle parallelamente ai sistemi IT è diventata essenziale. Questa urgenza deriva dall'aumento delle minacce informatiche che prendono di mira i settori che dipendono fortemente da queste tecnologie, come l'energia, le telecomunicazioni, i trasporti e il settore manifatturiero. A livello europeo, ciò si riflette nei recenti sviluppi normativi. In particolare, sia la Direttiva NIS2 sia la Direttiva CER rafforzano la protezione delle infrastrutture critiche, riconoscendo il ruolo dei sistemi OT nel garantire la continuità e la sicurezza dei servizi da esse forniti. Inoltre, il Regolamento Macchine dell'UE del 2023[29] introduce requisiti vincolanti in materia di cybersicurezza per le apparecchiature industriali, comprese misure volte a salvaguardare i software che operano in ambienti critici sia da danni fortuiti che da azioni malevole.

A partire dal 2024, l'Italia si colloca al quinto posto a livello mondiale per la produzione e il consumo di macchine utensili e al quarto posto per le esportazioni[30]. Utilizzate negli impianti industriali più avanzati, queste apparecchiature sono parte integrante di sistemi cyber-fisici fondati su componenti OT per il controllo, il monitoraggio e il coordinamento dei processi.

Sebbene esistano soluzioni di cybersicurezza consolidate per i sistemi IT, queste sono molto più complesse da implementare nei contesti OT, dove la rapida evoluzione tecnologica e la natura multivettoriale degli attacchi pongono sfide aggiuntive, combinate con la specificità e la complessità intrinseca dei sistemi OT. In questo contesto, un esempio di sfida particolarmente rilevante è lo sviluppo di metodi e modelli capaci di garantire la cybersicurezza senza compromettere la continuità e la sicurezza dei processi operativi e produttivi. In questa direzione, l'aggiornamento dell'Agenda è basato sullo stato dell'arte della letteratura scientifica sull'argomento, nonché degli standard[31], delle linee guida e delle raccomandazioni di enti istituzionali rilevanti, come la pubblicazione speciale NIST 800-82 relativa alla cybersicurezza delle OT[32], che adatta le ben note funzioni del *NIST Cybersecurity Framework (CSF)* al dominio OT.

[29] Regolamento (UE) 2023/1230 del Parlamento europeo e del Consiglio, del 14 giugno 2023, relativo alle macchine. Gazzetta Ufficiale dell'Unione Europea – EUR-Lex (europa.eu).

[30] Source: UCIMU-SISTEMI PER PRODURRE, associazione ufficiale dei produttori italiani di macchine utensili, robot, sistemi di automazione e prodotti affini. Maggiori informazioni disponibili su Settore macchine utensili (ucimu.it).

[31] Per il contesto dei sistemi di automazione e controllo industriale (Industrial Automation and Control Systems), si veda in particolare la serie ISA/IEC 62443 (isa.org)

[32] Documento disponibile (in inglese) su SP 800-82 Rev. 3, Guide to Operational Technology (OT) Security (csrc.nist.gov)

In particolare, per le finalità di questo aggiornamento, sono state identificate alcune funzioni rilevanti come descritto nel seguito del paragrafo. Per *Identificazione (Identify, ID)*, sono state ritenute d'interesse le sottocategorie *Governance* e *Strategia di gestione del rischio*, per *Protezione (Protect, PR)* le sottocategorie *Sicurezza dei dati* e *Processi e procedure per la protezione delle informazioni*. Infine, per *Rilevamento (Detection, DE)* e *Risposta (Response, RS)*, sono state prese in considerazione le sottocategorie *Monitoraggio continuo* e *Mitigazione durante la risposta*.

Le sottocategorie menzionate[33], alle quali sono associati i sottotemi di ricerca identificati, sono riportate nella Tabella 2-3.

Tabella 2-3: Sottoargomenti relativi alle tecnologie operative.

Classificazione secondo NIST SP 800-82 Rev. 3	Sottoargomento		Subarea
GOVERNANCE	**OT - 6.3.2.1**	Progettazione di standard di cybersicurezza specifici per OT	**6.3 Standardizzazione**
STRATEGIA DI GESTIONE DEL RISCHIO	**OT - 6.2.2.1**	KPI per la gestione del rischio di cybersicurezza nei sistemi OT	**6.2 Gestione del rischio**
SICUREZZA DEI DATI	**OT - 4.2.2.1**	Miglioramento della robustezza del sistema di gestione delle chiavi crittografiche per OT integrate con sistemi IIoT	**4.2 Reti**
PROCESSI E PROCEDURE PER LA PROTEZIONE DELLE INFORMAZIONI	**OT - 2.3.1.1**	*Digital twins* delle infrastrutture critiche comprendenti OT interconnesse e settoriali	**2.3 Gestione degli incidenti e delle operazioni di sicurezza**

[33] Le sottocategorie elencate sono descritte rispettivamente nelle Sezioni 6.1.2 (*Governance*), 6.1.4 (*Risk Management Strategy*), 6.2.3 (*Data Security*), 6.2.4 (*Information Protection Processes and Procedures*), 6.3.2.1 (*Continuous Monitoring – Network Monitoring*), e 6.4.4 (*Response Mitigation*) della pubblicazione speciale citata.

Classificazione secondo NIST SP 800-82 Rev. 3	Sottoargomento		Subarea
MONITORAGGIO CONTINUO -MONITORAGGIO DI RETE	OT - 2.1.1.1	Rilevamento delle intrusioni di rete per sistemi OT	2.1 Attacco e Difesa
MITIGAZIONE DURANTE LA RISPOSTA	OT - 3.1.4.1	Individuazione e correzione delle vulnerabilità di sicurezza nel firmware dei dispositivi per OT	3.1 Sicurezza nello sviluppo e test del software

2.3.1 GOVERNANCE

La *governance* coinvolge la dirigenza di una organizzazione, integrando gli obiettivi di gestione del rischio nel processo di pianificazione strategica, congiuntamente agli obiettivi di resilienza, privacy e cybersicurezza. Sebbene esistano standard rilevanti che forniscono una base per strutturare la governance della cybersicurezza nei sistemi OT, questi non sempre tengono in considerazione aspetti quali l'interoperabilità, necessaria a gestire la complessità intrinseca dei moderni sistemi OT, nonché la gestione della complessità di interconnessione o la frammentazione dei sistemi OT moderni. In questo contesto, quindi, è importante affrontare la carenza di meccanismi per la cybersicurezza OT che siano condivisi, standardizzati e interoperabili (cfr. sottoargomento **OT-6.3.2.1**).

2.3.2 STRATEGIA DI GESTIONE DEL RISCHIO

La strategia di gestione del rischio, parte della fase di identificazione, determina il modo in cui il rischio viene inquadrato, valutato, affrontato e monitorato, fornendo un valido supporto alle decisioni. In questo contesto, è particolarmente importante garantire che la strategia tenga conto dei rischi sistemici, degli effetti a cascata e della necessità di framework di *governance* integrati ed efficienti, in particolare, per ambienti OT altamente interconnessi (cfr. sottoargomento **OT-6.2.2.1**).

2.3.3 SICUREZZA DEI DATI

La sicurezza dei dati include sia la protezione della riservatezza, dell'integrità e della disponibilità (paradigma RID) dei dati sia la protezione dei dispositivi che li gestiscono, oltre a metodi per la prevenzione dalla perdita di dati. Di conseguenza, garantire la sicurezza dei dati nei sistemi OT presenta sfide principalmente legate alle caratteristiche degli ambienti che li gestiscono; tali sfide includono la compatibilità con i sistemi *legacy*, oppure i vincoli di esercizio sia temporali sia di risorse (cfr. sottoargomento **OT-4.2.2.1**).

2.3.4 PROCESSI E PROCEDURE PER LA PROTEZIONE DELLE INFORMAZIONI

Il ciclo di vita delle politiche, dei processi e delle procedure atte alla protezione delle informazioni nei sistemi OT fa parte della fase di *protezione*. In questo contesto, è particolarmente critica la raccolta e l'analisi efficace delle informazioni relative alle minacce informatiche OT, anche al fine di rafforzare la cybersicurezza delle infrastrutture critiche. In questo contesto, è importante sviluppare metodi e strumenti avanzati, come ad esempio *digital twins* e *cyber-range*, che generano rappresentazioni simulate di oggetti fisici e processi, riflettendo comportamenti specifici e dipendenze dei sistemi OT (cfr. sottoargomento **OT-2.3.1.1**).

2.3.5 MONITORAGGIO CONTINUO - MONITORAGGIO DI RETE

Le organizzazioni dovrebbero adottare il monitoraggio continuo come componente chiave della propria strategia di gestione del rischio, al fine di garantire l'efficacia costante delle misure di protezione. Il monitoraggio continuo comprende diverse attività, tra cui la gestione degli asset, la definizione di parametri di riferimento per il traffico di rete tipico, nonché dei flussi di dati e delle comunicazioni tra dispositivi, la diagnosi di problemi di prestazioni della rete e l'identificazione di configurazioni errate o malfunzionamenti nei dispositivi connessi. Il monitoraggio delle reti OT pone sfide di ricerca uniche legate al funzionamento affidabile in presenza di vincoli di tempestività operativa, alla presenza di tecnologie legacy e alla necessità di mantenere l'integrità fisica dei processi (cfr. sottoargomento **OT-2.1.1.1**).

2.3.6 MITIGAZIONE DURANTE LA RISPOSTA

I componenti OT sono spesso collocati in luoghi remoti e non costantemente presidiati; quindi, una sfida risiede nel loro ripristino in caso di attacco. Rileva inoltre notare che un sistema OT dovrebbe essere progettato con la capacità di minimizzare gli impatti in caso di incidenti. In questa direzione, quindi, occorre prestare attenzione allo sviluppo di metodi di *hot-patching* per i sistemi OT (cfr. sottoargomento **OT-3.1.4.1**).

3 Aree di Ricerca e Innovazione

3

Aree di Ricerca e Innovazione

Le aree di R&I sulla cybersicurezza identificate sono:

- Area#1 – Sicurezza dei Dati e Privacy;
- Area#2 – Gestione delle Minacce Cibernetiche;
- Area#3 – Sicurezza del Software e delle Piattaforme;
- Area#4 – Sicurezza delle Infrastrutture Digitali;
- Area#5 – Aspetti della Società;
- Area#6 – Aspetti di Governo.

Le aree di R&I sono raffigurate in Figura 3-1, la quale evidenzia anche il ruolo trasversale che le EDT rivestono rispetto alle stesse.

Figura 3-1: relazione tra aree di R&I e EDT.

Di seguito, ogni area di R&I è presentata nel dettaglio.

3.1 Area#1: Sicurezza dei Dati e Privacy

La sicurezza dei dati è il processo che consente di preservare la riservatezza, l'integrità e la disponibilità (la cosiddetta triade RID) dei dati nel mondo digitale, in aggiunta ad altre caratteristiche, quali l'autenticità, la responsabilità (*accountability*) e il non ripudio (*non-repudiation*). In un numero sempre maggiore di casi, i servizi digitali che usiamo tutti i giorni, come ad esempio i servizi di acquisto online, i servizi di localizzazione per la navigazione o le reti sociali, utilizzano dati personali[34] – che includono *personally identifiable information* (PII)[35]. Sebbene forniscano molti vantaggi ai loro utenti, questi servizi introducono rischi legati al trattamento dei dati, siano essi personali o meno. Questi rischi non possono essere trascurati, ma vanno accuratamente gestiti, soprattutto quando legati a servizi che sono parte di processi articolati in cui monitorare tutti i soggetti coinvolti diventa molto complesso.

L'area *Sicurezza dei Dati e Privacy* include tre subaree, descritte a seguire.

3.1.1 SUBAREA#1.1: INGEGNERIA DELLA PROTEZIONE DEI DATI

La protezione dei dati *by design* è un obbligo di legge in Europa a partire dalla piena operatività del GDPR nel 2018. Dal punto di vista tecnologico, sono adottate sempre più frequentemente nuove soluzioni per preservare la privacy, comunemente indicate come *Privacy-Enhancing Technology* (PET). Tuttavia, garantire un'efficace protezione dei dati *by design* richiede un approccio globale che contempli l'intero processo di elaborazione dei dati e sfrutti le PET in maniera corretta e tempestiva.

In questa subarea, gli argomenti prioritari sono i seguenti:

[34] La definizione di dati personali secondo il GDPR è disponibile (in inglese) al sito Art. 4 GDPR - Definitions (gdpr.eu).

[35] Si veda (in inglese) personally identifiable information - Glossary (csrc.nist. gov).

- **[Argomento#1.1.1] l'elaborazione *privacy-preserving* dei dati**, basata su tecniche come, ad esempio, la crittografia omomorfica[36], la *Secure Multi-party Computation* (SMC), la *differential privacy*, la generazione di dati sintetici e soluzioni basate su sicurezza hardware come *Trusted Execution Environment* (TEE) (cfr. Sezione 3.4.1);

#1.1.1 - Sottoargomento

GAI - 1.1.1.1

Bilanciamento tra privacy e prestazioni dei modelli GPAI nelle fasi di addestramento e inferenza

Le tecniche *privacy-preserving* e la crittografia sono sempre più impiegate per proteggere i dati sensibili all'interno dei modelli GPAI da attacchi di inferenza. Durante la fase di addestramento, queste tecniche vengono adattate per perturbare i dati o i parametri del modello, in fase di *pre-processing* o durante l'ottimizzazione, ad esempio, mediante l'aggiunta di rumore, al fine di ridurre il rischio di *data leakage*, ovvero la riproduzione da parte del modello di informazioni sensibili viste in fase di addestramento. In scenari in cui la condivisione dei dati è preclusa da vincoli di riservatezza, tecniche come il *collaborative/federated learning* permettono a più entità di addestrare congiuntamente modelli mantenendo la confidenzialità dei dati. Nella fase di inferenza, invece, si ricorre a metodi crittografici per proteggere sia i parametri del modello sia i dati. Entrambi gli approcci possono influire negativamente sulle prestazioni del modello. **Per la fase di addestramento, la ricerca dovrebbe concentrarsi su meccanismi adattivi, possibilmente limitati ai soli dati più sensibili, così da ridurre l'impatto sulle prestazioni. Per l'inferenza, è essenziale migliorare l'efficienza dei protocolli crittografici.**

- **[Argomento#1.1.2] la memorizzazione privacy-preserving dei dati**, che prevede tecniche crittografiche applicate all'hardware (cfr. Sezione 3.4.1) e al software che gestisce i supporti di memorizzazione (cfr. Sezione 3.3.1);
- **[Argomento#1.1.3] l'autenticazione, l'autorizzazione e il controllo di accesso con garanzie di privacy aggiuntive**, sfruttando tecniche che aumentano il livello di protezione dei dati come, ad esempio, le *zero-knowledge proofs*.

[36] Si veda anche l'Argomento#1.2.3 incentrato sulla *Fully Homomorphic Encryption (FHE)*.

3.1.2 SUBAREA#1.2: CRITTOGRAFIA

La promozione dell'uso della crittografia, così come lo sviluppo di sistemi di crittografia nazionali in ambito non classificato, sono esplicitamente menzionati nella "Strategia Nazionale di Cybersicurezza 2022-2026 – Piano di Implementazione"[37]. Di recente, il campo della crittografia sta affrontando sfide importanti che derivano dai progressi del *quantum computing*. In quest'ambito, gli argomenti prioritari sono:

[Argomento#1.2.1] lo studio di nuove primitive, algoritmi e protocolli per la *Post-Quantum Cryptography* (PQC)[38], che si focalizza sulla progettazione e lo sviluppo di soluzioni adottabili dagli attuali elaboratori e dispositivi elettronici, che siano tuttavia robuste nei confronti della crittoanalisi sia tradizionale sia quantistica. Nello specifico, è incluso l'approfondimento di nuovi problemi matematici adatti alla PQC, ovvero che siano difficili da risolvere anche con un computer quantistico – tali problemi sono tipicamente basati su reticoli, funzioni di *hashing*, codici e isogenie su curve ellittiche[39, 40]. Inoltre, è d'interesse lo studio di tecniche crittografiche ibride, che combinino approcci classici e post-quantum. Si incoraggia, altresì, la crittoanalisi e la validazione di algoritmi e protocolli di PQC allo stato dell'arte, al fine di individuare nuovi attacchi e garantirne l'effettiva sicurezza.

- **[Argomento#1.2.2] lo studio della crittografia quantistica**, la quale, al contrario della PQC, sfrutta le proprietà della fisica dei quanti per realizzare procedure crittografiche. Una delle direzioni più promettenti è la distribuzione di chiavi quantistiche (*Quantum Key Distribution*, QKD), la quale punta a fornire un approccio sicuro dal punto di vista della teoria dell'informazione per lo scambio di chiavi crittografiche. Attualmente, i sistemi di QKD sono soggetti a limitazioni tecniche che la rendono applicabile solo in ambiti specifici, quali, ad esempio, la protezione delle comunicazioni su collegamenti punto-punto, sfruttando un'autenticazione iniziale di tipo tradizionale. Il superamento di tali limitazioni è necessario per garantirne l'implementazione su larga scala[41].

In aggiunta, meritano attenzione le soluzioni di crittografia omomorfica (*homomorphic encryption*), che rappresentano una forma di crittografia che permette di effettuare elaborazioni direttamente sui dati

[37] In particolare: **Misura #22** (Promuovere l'uso della crittografia in ambito non classificato, quale impostazione predefinita e comunque fin dalla fase di progettazione di reti, applicazioni e servizi, in conformità ai principi della sicurezza e della tutela della vita privata, nel rispetto dei principi stabiliti dalla normativa nazionale ed europea) e **Misura #23** (Sviluppo di tecnologie/sistemi di cifratura nazionale in ambito non classificato. A sostegno di tale iniziativa è prevista la creazione di un ecosistema nazionale per il suo mantenimento ed evoluzione).

[38] Si veda il sito (in inglese) Post-Quantum Cryptography (csrc.nist.gov).

[39] Si veda il documento introduttivo sul tema, a cura dell'ACN, Crittografia Post-quantistica e quantistica (acn.gov.it).

[40] A tal proposito, è opportuno menzionare il processo di standardizzazione avviato dal NIST (csrc.nist.gov), basato su una competizione internazionale avviata nel 2016 e attualmente ancora in corso. La competizione gioca un ruolo fondamentale nell'identificare e valutare algoritmi *quantum-resistant*, assicurandone la robustezza e l'adozione su diverse piattaforme. In tale contesto, il ruolo della standardizzazione è fondamentale – cfr. sottoargomento **QT-6.3.2.2**.

[41] Nello specifico, la QKD si affida a metodi di autenticazione classici, quali, ad esempio, segreti condivisi a latere o certificati digitali – si veda il sito (in inglese) ENISA Briefing on QKD (europa.eu). Lo sviluppo di protocolli di comunicazione quantistica efficaci costituisce un'area di ricerca attiva. Anche in questo contesto, il ruolo della standardizzazione è fondamentale – cfr. sottoargomento **QT-6.3.2.1**.

crittografati, senza la necessità di effettuare una preventiva decrittazione.

In questo campo, l'argomento prioritario è:

- **[Argomento#1.2.3] la ricerca di schemi per la *Fully Homomorphic Encryption* (FHE)**, e altre tecniche crittografiche per *data-in-use*, applicabili in scenari concreti per abilitare l'elaborazione dei dati nel rispetto della privacy anche in ambienti non nativamente affidabili (ad es. *cloud computing*), con applicazioni in ambiti critici come la sanità, la finanza e l'analisi forense.

Infine, come parte della continua evoluzione dello stato dell'arte degli algoritmi crittografici, si propone un ulteriore argomento prioritario riguardante:

- **[Argomento#1.2.4] nuove funzioni crittografiche di *hashing***, le quali, tra i possibili utilizzi, possono essere elementi costitutivi di alcuni schemi di crittografia omomorfica e sono impiegate nelle *blockchain* (cfr. Sezione 3.3.3).

3.1.3 SUBAREA#1.3: TRUSTED INFORMATION SHARING

Il *Data Act*[42], come parte della strategia europea complessiva sui dati, è un importante passo in avanti per utilizzare pienamente il potenziale dei moderni ecosistemi di dati, in particolare facilitando la condivisione orizzontale tra diversi settori. Il regolamento rende obbligatoria la condivisione di dati per scopi ben definiti di pubblica utilità, vincolandola allo stesso tempo a determinate garanzie di diritti e interessi degli individui a cui i dati si riferiscono. Inoltre, nell'ambito della strategia europea sui dati, il "*Data Governance Act*"[43] introduce i *common European data space* come soluzioni specifiche per abilitare la condivisione e il riuso dei dati.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#1.3.1] l'arricchimento delle soluzioni standard di interoperabilità semantica con controlli di sicurezza**, ad esempio per garantire la confidenzialità e l'integrità dei dati;
- **[Argomento#1.3.2] l'analisi di soluzioni architetturali per *trusted information sharing* e *storage***, sia per un settore specifico che per la condivisione dei dati tra settori, ad esempio tramite l'uso di PET che permettono di accedere ai dati in maniera *privacy-preserving* (cfr. Sezione 3.1.1). In particolare, lo sviluppo di soluzioni di *trusted information storage* sono di estremo interesse con riferimento a infrastrutture su scala nazionale che memorizzino copie sicure di dati pubblici e privati per garantire il ripristino da condizioni di disastro causate da incidenti cibernetici (cfr. Sezione 3.2.3).

[42] Regolamento (UE) 2023/2854 del Parlamento europeo e del Consiglio, del 13 dicembre 2023, riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo (Data Act). Gazzetta Ufficiale dell'Unione Europea EUR-Lex (europa.eu).

[43] Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022 relativo alla *governance* europea dei dati (Data Governance Act). Gazzetta Ufficiale dell'Unione Europea EUR-Lex (europa.eu).

3.2 Area#2: Gestione delle Minacce Cibernetiche

Nel *Threat Landscape Report*[2], ENISA ha evidenziato come le tensioni geopolitiche e i conflitti regionali, tuttora in corso, abbiano prodotto effetti di ampia portata sul panorama della cybersicurezza, contribuendo a una crescita costante e capillare dell'*hacktivism*, che oggi rappresenta quasi l'80% degli incidenti segnalati. Tra i settori maggiormente colpiti si confermano la pubblica amministrazione (38,2%), riconosciuta come settore chiave dalla Direttiva NIS2, insieme a trasporti (7,5%) e infrastrutture e servizi digitali (4,8%), a testimonianza della loro particolare vulnerabilità nel contesto attuale. In tale scenario, il ruolo dei *Security Operation Centers* (SOC) risulta cruciale e necessita di un ulteriore rafforzamento, come indicato anche nella *Strategia Nazionale di Cybersicurezza 2022-2026 – Piano di Implementazione*[3,44].

L'area *Gestione delle Minacce Cibernetiche* include le quattro subaree descritte nel seguito.

3.2.1 SUBAREA#2.1: ATTACCO E DIFESA

Questa subarea tratta le tecniche che consentono il rilevamento degli attacchi cibernetici e le reazioni a essi mirate a preservare l'operatività dei sistemi impattati.

In questa subarea, gli argomenti prioritari sono i seguenti:

[44] Nello specifico, si veda la **Misura #30** (*Realizzare un sistema di raccolta e analisi HyperSOC per aggregare, correlare e analizzare eventi di sicurezza di interesse, al fine di individuare precocemente eventuali "pattern" di attacco complessi, nonché abilitare una gestione del rischio cyber in chiave preventiva e integrata tra molteplici sorgenti dati, sfruttando anche infrastrutture di High Performance Computing e tecnologie di Intelligenza Artificiale e il machine learning*).

- **[Argomento#2.1.1] le nuove tecniche di attacco e misure di difesa**, che possono essere applicate nelle diverse fasi di un attacco cibernetico, quali, ad esempio, nuove misure di difesa contro attacchi di tipo *Distributed Denial of Service* (DDoS);

#2.1.1 - Sottoargomenti

GAI - 2.1.1.1

Contrasto all'uso malevolo dei modelli GPAI

I modelli GPAI sono oggetto di studio crescente, oltre che per le loro numerose applicazioni positive, anche per i possibili usi in scenari offensivi. Ad esempio, possono essere impiegati per analizzare e inserire componenti malevoli nascosti (*Hardware Trojan*, HT) in progetti hardware complessi, interrogare lo stato dei sistemi e suggerire sequenze di attacco. I modelli possono inoltre essere sfruttati per realizzare attacchi *side-channel*[45] per ottenere dati sensibili, come chiavi crittografiche. Un ulteriore rischio rilevante risiede nella possibilità di impiego dei modelli per sviluppare nuove minacce, in grado di sfruttare vulnerabilità sconosciute e sfuggire ai meccanismi di rilevamento. **Contrastare l'uso malevolo della GPAI richiede un approccio articolato che combini strategie preventive e difensive. Da un lato, occorre mettere in sicurezza i modelli (tramite addestramento avversario, il *safety fine-tuning* e l'adozione di meccanismi di controllo, c.d. *guardrailing*, – cfr. anche il sottoargomento GAI-2.1.3.1) al fine di ridurre la possibilità che modelli presenti sul mercato vengano facilmente sfruttati da attori malevoli. Tuttavia, tali misure non sono sufficienti a impedire che attaccanti dotati di capacità e risorse elevate eseguano attività di addestramento o *fine-tuning* mirato. Per questo motivo, è di grande interesse la ricerca sull'impiego della GPAI, in particolare gli approcci c.d. ad agenti (*agentic*) per generare scenari di attacco[46] utili alle attività di difesa, supportando i *blue team* nell'anticipazione e nel contrasto delle azioni ostili, rafforzando così le strategie difensive.**

[45] Ad es. su un dispositivo crittografico o un'implementazione software.

[46] L'uso della GPAI per l'istanziazione controllata di tecniche di attacco permette altresì di generare dati sintetici utili a migliorare i modelli di rilevamento e a fornire set di dati robusti per il test, la validazione e l'addestramento di difese che restino efficaci nel tempo.

#2.1.1 - Sottoargomenti

QT - 2.1.1.1

Sensoristica quantistica per migliorare il rilevamento degli attacchi

L'estrema precisione e sensibilità dei sensori quantistici, ad esempio per l'analisi dei campi elettromagnetici, potrebbe essere sfruttata per migliorare i meccanismi di difesa contro gli attacchi cibernetici. **Pertanto, è incoraggiato lo studio di nuovi scenari in cui tale tecnologia può essere sfruttata per migliorare il rilevamento di tali attacchi**.

OT - 2.1.1.1

Rilevamento delle intrusioni di rete per sistemi OT

Il rilevamento delle minacce informatiche negli ambienti OT presenta sfide uniche a causa della criticità, dell'eterogeneità e dei vincoli operativi dei sistemi industriali. Gli strumenti e gli approcci di monitoraggio tradizionali, come il rilevamento delle anomalie comportamentali (*Behaviour Anomaly Detection* - BAD), i sistemi di gestione delle informazioni e degli eventi di sicurezza (*Security Information and Event Management* - SIEM) e i sistemi di rilevamento delle intrusioni *(Intrusion Detection Systems* - IDS), devono essere adattati per funzionare in modo affidabile in presenza di vincoli di tempestività operativa, tecnologie obsolete e dalla necessità di mantenere l'integrità fisica dei processi, essenziale per la sicurezza delle operazioni critiche. **La ricerca dovrebbe concentrarsi sulla progettazione di soluzioni di rilevamento delle intrusioni specifiche per l'OT, che combinino dati di rete e di processo, modellando attacchi furtivi e a bassa intensità che sfuggono al rilevamento tradizionale, come le minacce persistenti avanzate (*Advanced Persistent Threats* - APT), la cui furtività si basa tipicamente sull'uso di comandi e operazioni autorizzati, aspetto che li rende meno visibili agli approcci di rilevamento standard. Un'ulteriore direzione promettente è l'integrazione del rilevamento delle intrusioni nella progettazione di *Safety Instrumented Systems* – SIS per le infrastrutture OT, arricchendoli con aspetti di cybersicurezza**[47].

[47] I SIS sono meccanismi di sicurezza ingegnerizzati utilizzati negli ambienti OT industriali per portare automaticamente i processi in uno stato sicuro quando vengono rilevate condizioni pericolose. Si veda (in inglese) safety instrumented system - Glossary (csrc.nist.gov)

- **[Argomento#2.1.2] rilevamento e risposta contro i *malware***, incluse tecniche specifiche per offuscare, de-offuscare e analizzare i *malware*, così come quadri procedurali per eseguire tali tecniche in maniera non bloccante sui sistemi impattati. Inoltre, in generale, sono da investigare nuove tecniche per l'identificazione di attività maliziose, al fine di riconoscere prontamente tipologie di *malware* non ancora note, con l'inclusione di modelli di analisi statistica.
- **[Argomento#2.1.3] la messa in sicurezza di algoritmi e modelli per il *machine learning***, studiando tecniche per migliorarne la robustezza rispetto a minacce quali gli attacchi avversari (*adversarial attacks*), ad esempio l'*adversarial model inference*, o la manipolazione dei dati utilizzati da tali sistemi come, ad esempio, il *data poisoning* e la *data exfiltration*[48];

#2.1.3 - Sottoargomento

GAI - 2.1.3.1

Difesa contro le minacce avversarie

Gli attacchi ai modelli GPAI, come *jailbreak e prompt injection*, possono comportare manipolazioni dirette o indirette volte a sovvertire le istruzioni definite dagli sviluppatori, sfruttando fonti esterne per infiltrare o esfiltrare informazioni o per innescare comportamenti malevoli. Le tecniche più avanzate intensificano tali rischi facendo leva anche sulla conoscenza dell'architettura del modello. Gli agenti autonomi (cfr. sottoargomento **GAI-6.2.4.1**) risultano particolarmente vulnerabili, diventando frequentemente obiettivo di attacchi avversari con impatti sugli utilizzatori. Gli attacchi di tipo *data poisoning* e *backdoor* compromettono l'integrità dell'addestramento della GPAI attraverso l'inserimento nei dataset di dati malevoli, inclusi schemi attivabili dagli attaccanti in condizioni specifiche. L'uso di dati non verificati, spesso provenienti da fonti aperte, può compromettere l'integrità dei modelli e ridurre l'efficacia dei meccanismi di difesa. I rischi legati a queste minacce non sono ancora stati analizzati in modo sistematico durante le diverse fasi di sviluppo dei modelli, né è stato chiarito quanto le scelte progettuali possano amplificarne le vulnerabilità. **Per affrontare queste sfide, è necessario sviluppare *benchmark* standardizzati, migliorare le metodologie di *adversarial testing* e *red-teaming*, e approfondire la trasferibilità degli attacchi tra diversi modelli. È inoltre essenziale esplorare più a fondo i vettori di attacco multimodali**[49].

[48] Si veda (in inglese) Securing Machine Learning Algorithms (enisa.europa.eu).

[49] Ovvero, attacchi che combinano input di diverse tipologie (es. testo, immagini, audio) per eludere i sistemi di difesa, aumentando la complessità e l'efficacia degli attacchi.

- **[Argomento#2.1.4] la ricerca sul cosiddetto *adversarial behaviour***, al fine di analizzare le motivazioni e le capacità degli attaccanti con particolare riferimento a tecniche e tecnologie utilizzate nelle attività ostili da loro condotte, delegando invece gli aspetti umani all'Area#5 (cfr. Sezione 3.5.1).

3.2.2 SUBAREA#2.2: CYBERTHREAT INTELLIGENCE

Questa subarea si focalizza sulla raccolta e l'analisi informativa, al fine di caratterizzare possibili minacce cibernetiche. La creazione di un servizio nazionale di monitoraggio delle minacce cibernetiche è esplicitamente menzionata nella *"Strategia Nazionale di Cybersicurezza 2022-2026 – Piano di Implementazione"*[3, 50].

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#2.2.1] l'utilizzo del *machine learning* per la *cyberthreat intelligence***, dove il *machine learning* è da intendersi in generale, a ricomprendere anche il c.d. *deep learning*. Sono incluse tecniche di *machine learning* per l'individuazione di *malware*, spam e tentativi di intrusione nei sistemi, oltre che altri utilizzi finalizzati in generale al contrasto degli attacchi cibernetici. In aggiunta, le tecniche di *machine learning* possono essere usate anche per la costruzione e l'aggiornamento di *knowledge graph* che, ad esempio, rappresentino e organizzino la conoscenza relativa agli attori ostili o alla cosiddetta impronta digitale *(digital footprint)* di un'organizzazione. Le soluzioni basate su *machine learning* necessitano di adeguati controlli di qualità e correttezza che devono essere applicati sia nelle fasi di addestramento che in quelle di identificazione delle feature dei modelli. Il machine learning può essere anche usato per potenziare strategie di attacco e sviluppare strategie di difesa: in questa accezione, sarà trattato nella Sezione 3.2.3. Questo argomento è esplicitamente menzionato nella *"Strategia Nazionale di Cybersicurezza 2022-2026 – Piano di Implementazione"*[3, 51];

[50] In particolare, si veda la **Misura #13** (Realizzare un servizio di monitoraggio del rischio cyber nazionale a favore delle organizzazioni e del pubblico in generale, al fine di comunicare l'effettivo livello della minaccia, nonché di informare adeguatamente i processi decisionali).

[51] In particolare, si veda la **Misura #32** (Creare un'infrastruttura di High Performance Computing dedicata alla cybersecurity nazionale per il potenziamento dei servizi cyber nazionali dell'Agenzia, nonché lo sviluppo di strumenti di simulazione, basati sull'Intelligenza Artificiale e il machine learning, per supportare le fasi di prevenzione, scoperta, risposta e predizione degli impatti di attacchi cyber di natura sistemica).

#2.2.1 - Sottoargomenti

GAI - 2.2.1.1

Uso della GPAI per la rilevazione e la mitigazione delle capacità di evasione (*evasion*) del *malware*

L'impiego della GPAI per la rilevazione di *malware* si basa su capacità avanzate di comprensione contestuale e di elaborazione del linguaggio naturale, che permettono a questi modelli di individuare, nell'analisi del software malevolo, schemi complessi e anomalie sottili che i metodi tradizionali, basati su firme o euristiche, potrebbero non rilevare. Tuttavia, queste soluzioni sono particolarmente prone a falsi positivi (ovvero la classificazione errata di software benigno come malevolo) e a falsi negativi (mancata individuazione di *malware* sofisticati). Inoltre, gli attaccanti possono ricorrere a tecniche di elusione relativamente semplici, come offuscamento del codice, polimorfismo o perturbazioni avversarie, al fine di aggirare i meccanismi di rilevamento. **Gli approcci emergenti sfruttano la capacità della GPAI di elaborare e codificare relazioni semantiche e sintattiche all'interno del codice, permettendo l'identificazione di *malware* offuscati o metamorfici. Rimane tuttavia fondamentale migliorare la robustezza degli approcci, anche considerando sistemi ibridi che combinino le capacità della GPAI con i metodi tradizionali di analisi statica e dinamica**.

GAI - 2.2.1.2

Uso della GPAI per supportare la *cyber threat intelligence* e la rilevazione degli incidenti

I modelli GPAI vengono sempre di più impiegati per potenziare i processi di *cyber threat intelligence* (CTI), automatizzando la generazione di report e l'estrazione di dati da fonti eterogenee. Contribuiscono inoltre all'analisi delle *Tactics, Techniques and Procedures* (TTP) e alla costruzione di *knowledge graph* delle minacce. Nel contesto della gestione degli incidenti, la GPAI accelera i processi analizzando metriche, eventi, *log* e tracce di esecuzione, e assistendo gli operatori nell'identificazione di possibili cause (*root-cause analysis*), inclusi potenziali sfruttamenti di vulnerabilità *zero-day*, e contromisure. Tuttavia, l'impiego della GPAI in attività come la rilevazione in tempo reale di anomalie *(real-time anomaly detection)* presenta criticità in termini di efficienza computazionale delle fasi di addestramento, soprattutto in ambienti che richiedono bassa latenza come i motori di correlazione utilizzati nei sistemi SIEM o SOAR. **La ricerca dovrebbe concentrarsi sul miglioramento della robustezza, dell'efficienza computazionale e della spiegabilità di queste applicazioni. In ambito di rilevazione, approcci ibridi che combinano l'uso della GPAI con metodi di analisi classici potrebbero rappresentare una via promettente per aumentare l'interpretabilità senza compromettere le capacità predittive**.

- **[Argomento#2.2.2] la ricerca di nuove metodologie per monitorare la disinformazione**, come parte di una strategia globale per mitigare l'impatto delle *fake news* o di altri tipi di contenuti falsi, anche di tipo audio, video e fotografico (ad esempio, i cosiddetti *deepfakes*), creati e condivisi da attori ostili per influenzare l'opinione pubblica. Sono incluse come metodologie di interesse le recenti proposte basate su GPAI, in particolare, sui modelli generativi del linguaggio;
- **[Argomento#2.2.3] le metodologie per l'attribuzione di attacchi alla sicurezza cibernetica,** finalizzate al tracciamento e all'identificazione delle attività dell'attaccante, anche nell'ottica di supportare operazioni di profilazione nell'ambito di indagini investigative (cfr. Sezioni 3.2.1 e 3.5.1).

3.2.3 SUBAREA#2.3: GESTIONE DEGLI INCIDENTI E OPERAZIONI DI SICUREZZA

Secondo il recente "*Threat Landscape*" dell'ENISA[2], i beni ICT, e in generale la popolazione europea, sono ormai oggetto di un numero di incidenti cibernetici rilevati stabilmente elevato. In particolare, sono state registrate operazioni cibernetiche ostili contro infrastrutture critiche, al fine primario di ottenere informazioni confidenziali, ma anche di diffondere nuovi *malware* e interrompere servizi e funzioni essenziali. Pertanto, in questo contesto, è importante stabilire e perfezionare approcci destinati a (i) rispondere agli incidenti e (ii) condurre operazioni efficaci di rafforzamento della sicurezza su sistemi cibernetici e cyber-fisici.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#2.3.1] la resilienza dei sistemi cyber-fisici** include, (i) il potenziamento della consapevolezza situazionale attraverso il miglioramento delle metodologie che permettono una percezione aumentata del contesto da parte degli operatori umani e (ii) la progettazione di framework per la gestione della risposta agli incidenti di sicurezza contro i sistemi cyber-fisici, comprensivi dei sistemi OT utilizzati nelle infrastrutture critiche[52], che raccolgano e analizzino efficacemente le informazioni sulle minacce informatiche, tenendo esplicitamente conto dell'impatto intersettoriale degli attacchi e della loro propagazione attraverso infrastrutture interconnesse.

[52] Quali, ad esempio, i sistemi robotici, nonché i dispositivi medicali avanzati, come quelli impiantabili.

#2.3.1 - Sottoargomento

OT - 2.3.1.1 **Digital Twin delle infrastrutture critiche comprendenti OT interconnesse e settoriali**	I sistemi OT possono far parte di infrastrutture critiche nazionali o transfrontaliere e, in tal caso, i dati trattati (ad es. misurazioni e comandi remoti) sono particolarmente sensibili. In questo scenario, è fondamentale sviluppare metodi e strumenti avanzati, come i *digital twin*[53] per generare rappresentazioni sintetiche di oggetti fisici e processi, riflettendone le dipendenze. Questi strumenti permetterebbero di eseguire simulazioni di cyber-resilienza su reti di sistemi OT in ambiti critici come i trasporti, la difesa e lo spazio. **Si incoraggia a investigare come implementare, manutenere e aggiornare modelli di digital twin accurati, che riflettano comportamenti e vulnerabilità specifiche di sistemi OT settoriali**.

- **[Argomento#2.3.2] l'automazione via *machine learning* dei sistemi tradizionali**, inclusi, ad esempio, i SIEM e i sistemi antintrusione (*Intrusion Prevention Systems* – IPS), al fine di migliorarne sia l'efficacia (ad esempio, rimuovendo le dipendenze da controlli statici secondo regole predefinite) che l'efficienza, grazie a funzioni più automatizzate;
- **[Argomento#2.3.3] le nuove strategie di difesa** da adottare a valle delle fasi di progettazione già condotte o in contesti ibridi, quali ad esempio il telelavoro, utilizzando sistemi di gestione delle infrastrutture digitali resilienti e adattativi. Questo argomento include anche le strategie di difesa in risposta ad attacchi potenziati dall'IA, che presentano una combinazione di velocità, scala, automazione e sofisticazione tale da richiedere risposte innovative.

3.2.4 SUBAREA#2.4: SCIENZE FORENSI DIGITALI

A seguito di un incidente di sicurezza, si attua dapprima una fase di reazione che comporta decisioni e azioni di contenimento e, successivamente, un'ulteriore fase che consiste nell'analisi a posteriori di ciò che si è verificato, per scopi investigativi e al fine di migliorare la postura cibernetica dell'infrastruttura impattata dall'incidente.

[53] Tra le applicazioni dei *cyber-range* (oltre a quelle approfondite nelle Sezioni 3.3.1 e 3.5.2) vi è la generazione di dati sintetici da impiegare nei *digital twin*.

Questa importante fase di analisi è effettuata sfruttando tecniche di scienze forensi digitali, ossia procedure informatiche e investigative che esaminano prove digitali garantendo autorità, catena di custodia, utilizzo di strumenti validati, ripetibilità e sfruttando reportistica, verifiche matematiche e testimonianze di esperti[54]. Le tecniche a garanzia delle prove digitali da utilizzare nell'ambito di procedure legali sono trattate anche dalla comunità di standardizzazione[55].

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#2.4.1] i nuovi approcci per gestire la complessità dei processi di scienze forensi digitali**, i quali debbono far fronte all'acquisizione, all'archiviazione e all'elaborazione di una grande quantità di prove digitali, nonché alla dispersione delle informazioni in diversi luoghi fisici o virtuali;
- **[Argomento#2.4.2] lo sviluppo di contromisure per aggirare le tecniche antiforensi** che impediscono alle forze dell'ordine di indagare sui criminali cibernetici.

3.3 Area#3: Sicurezza del Software e delle Piattaforme

Il software sta progressivamente permeando tutti gli aspetti della vita quotidiana e, specialmente in alcuni contesti importanti, quali, ad esempio, la guida autonoma, la sua affidabilità è sempre più essenziale. Infatti, le vulnerabilità che affliggono il codice sorgente o che possono essere introdotte tramite dipendenze da terze parti potrebbero danneggiare severamente gli utilizzatori finali, al punto tale da metterne a rischio l'incolumità fisica. A tal riguardo, è necessario ricordare che la disponibilità di software e piattaforme digitali nativamente nazionali o europee risulta ad oggi fortemente limitata.

L'area *Sicurezza del Software e delle Piattaforme* include tre subaree, descritte di seguito.

3.3.1 SUBAREA#3.1: SICUREZZA NELLO SVILUPPO E TEST DEL SOFTWARE

Ad oggi, il paradigma *DevOps*, garantisce un aumento del grado di efficienza e automazione dei rilasci delle applicazioni software, che avvengono direttamente in ambiente di produzione, a valle di ogni iterazione di sviluppo.

[54] Si veda digital forensics - Glossary | CSRC (nist.gov) (in inglese).

[55] Si veda, ad esempio, TS 103 643 - V1.2.1 - Techniques for assurance of digital material used in legal proceedings (etsi.org) (PDF, 140 KB) (in inglese).

Per questo motivo, la presenza eventuale di vulnerabilità all'interno del software, di natura volontaria o involontaria, può essere estremamente dannosa, specialmente se parte di infrastrutture critiche digitali (cfr. Sezione 3.4). Standard internazionali quali l'ISO/IEC 27001[56] e la *Special Publication* 800-160[57] del NIST affrontano questo problema.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#3.1.1] i nuovi linguaggi di programmazione *security-aware***, basati su modelli semantici che consentano di formalizzare, collaudare e certificare i meccanismi di sicurezza adottati *by design*;
- **[Argomento#3.1.2] le pratiche innovative di gestione sicura del ciclo di vita del software (*secure software development lifecycle*)**, che comprendono strumenti basati su *machine learning* per l'automazione dei controlli di sicurezza sul codice sorgente (cfr. Sezione 3.2.2), miglioramenti della sicurezza di applicazioni web e browser, anche tramite nuove tecniche per il rilevamento di bug a livello logico, l'applicazione di tecniche di *hardening* del codice e del *software bill of materials* (SBOM)[58]. In aggiunta, è di interesse anche l'adozione di *cyber-range* come ambiente sicuro per lo sviluppo e la verifica della postura di sicurezza di nuovi programmi, oltre all'uso primario dei *cyber-range* che è quello in ambito formativo (cfr. Sezione 3.5.2). L'argomento comprende soluzioni di automazione della gestione sicura di prodotti ampiamente distribuiti e con risorse limitate, per uso domestico o da ufficio, e i dispositivi IoT di nuova generazione[59] (cfr. anche la Sezione 3.4.2 per gli approcci specifici per la sicurezza IoT).

#3.1.2 - Sottoargomenti

GAI - 3.1.2.1 **Migliorare la sicurezza del codice generato con il supporto della GPAI**	La generazione di codice assistita da GPAI è in rapida diffusione, sollevando preoccupazioni in merito a sicurezza, robustezza e verificabilità del codice prodotto. Gli approcci attuali, basati principalmente su *prompt engineering* o euristiche, non garantiscono l'applicazione sistematica di principi *secure-by-design*. Con l'avvento dei sistemi GPAI ad agenti, dotati di capacità di pianificazione, invocazione di strumenti esterni e raffinamento iterativo, la sfida di ricerca si sposta verso la costruzione di pipeline controllabili e verificabili.

[56] Per approfondimenti si veda (in inglese) ISO/IEC 27001 and related standards — Information security management (iso.org).

[57] Disponibile (in inglese) al sito SP 800-160 Vol. 1 Rev. 1 - Engineering Trustworthy Secure Systems (csrc.nist.gov).

[58] Questo argomento è fortemente allineato con i requisiti del Cyber Resilience Act (CRA) relativi ai processi di sviluppo sicuro, inclusa l'adozione obbligatoria del SBOM, la gestione delle vulnerabilità e il supporto post-commercializzazione dei prodotti digitali.

[59] Questa direzione comporta sfide rilevanti a causa delle capacità limitate dei dispositivi, dell'elevata eterogeneità e della possibile assenza di sistemi centralizzati di controllo.

#3.1.2 - Sottoargomenti

Le direzioni future riguardano l'integrazione dei principi del *Secure Software Development* Lifecycle (SSDLC) nei workflow generativi, la formalizzazione di vincoli per ridurre configurazioni o comportamenti insicuri e la definizione di benchmark per individuare errori logici e costrutti vulnerabili. I controlli di sicurezza dovranno sempre più estendersi all'intero spettro degli artefatti software, inclusi componenti infrastrutturali e di c.d. *Policy as a Code*, richiedendo ai sistemi GPAI di ragionare su rappresentazioni eterogenee e supportare l'individuazione di deviazioni di configurazioni, politiche di controllo di accesso o di esposizione dell'ambiente di esecuzione, considerando diversi contesti e domini applicativi.

QT - 3.1.2.1

Messa in sicurezza dell'ingegneria del software quantistico

Con l'incremento delle capacità dei computer e delle comunicazioni quantistiche, è necessario che anche i linguaggi di programmazione e, più in generale, l'ingegneria del software evolvano, al fine di sfruttare appieno le capacità dell'infrastruttura sottostante. A tal riguardo, è probabile che le tecniche tradizionali di analisi del codice a garanzia della sicurezza dell'implementazione software non siano generalizzabili. Il "malware quantistico" potrebbe infatti eludere il rilevamento o disturbare il processo di calcolo sfruttando proprietà quantistiche come la sovrapposizione e l'*entanglement*. **Pertanto, a valle di un'analisi comparativa dello stato dell'arte sui linguaggi di programmazione quantistici finalizzato a valutarne la postura di sicurezza, si incoraggia la progettazione di modelli di sicurezza per il software quantistico che prevengano e mitighino l'impatto di minacce relative a *malware* quantistico. È interessante notare che, in futuro, questi modelli potrebbero essere applicati anche all'intelligenza artificiale quantistica, una disciplina emergente che ruota attorno all'accelerazione di specifiche subroutine, come quelle per l'ottimizzazione, il campionamento e l'elaborazione di dati ad alta dimensionalità, che risultano computazionalmente onerose per i calcolatori classici. L'obiettivo è contribuire alla mitigazione delle minacce informatiche contro il *machine learning* (cfr. Argomento#2.1.3), nonché migliorare le prestazioni del *machine learning* per scopi di difesa (cfr. Argomento#2.2.1).**

- **[Argomento#3.1.3] lo studio del paradigma dell'*open cloud***, fattore abilitante del trasferimento del software (*lift-and-shift*) tra cloud privati e pubblici e tra diversi fornitori di cloud pubblico, riducendo le dipendenze da fornitori specifici (*vendor lock-in*) e promuovendo l'interoperabilità.
- **[Argomento#3.1.4] le tecniche per rilevare vulnerabilità di sicurezza nel *firmware* e nei file binari**, i quali dovrebbero essere approfonditamente vagliati nel momento in cui il codice sorgente non è accessibile, al fine di rilevare *bug* e *backdoor*, assicurandone così la conformità al comportamento dichiarato o atteso dal fornitore. In particolare, tra le tecniche di interesse si menzionano quelle atte a favorire (i) il rilevamento di vulnerabilità di sicurezza, *bug* o *backdoor* utilizzando analisi statica e dinamica di file binari e *firmware*, (ii) l'analisi dei file binari da parte di uno sviluppatore terzo e (iii) l'hardening dei file binari stessi.

#3.1.4 - Sottoargomenti

GAI - 3.1.4.1

Uso della GPAI per il rilevamento e il *patching* delle vulnerabilità

I modelli GPAI sono sempre più impiegati nell'analisi della sicurezza di codice sorgente, dei file binari e del *firmware*, grazie alla loro capacità di automatizzare il rilevamento delle vulnerabilità. Tra le tecniche utilizzate rientrano la *taint analysis* statica, l'identificazione di flussi di dati non previsti, *bug* e modifiche non autorizzate. Inoltre, i modelli supportano attività quali generazione di possibili *exploit* e l'analisi di similitudini nel codice, utile per individuare varianti di *malware*. Per quanto riguarda il *patching*, i modelli GPAI sono utilizzati per correggere vulnerabilità comuni, come i *buffer overflow* e di errori nella validazione dell'input. Il loro pieno potenziale è tuttavia limitato da sfide significative, tra cui l'elevata incidenza nel rilevamento di falsi positivi, problemi di scalabilità e difficoltà di integrazione con gli strumenti di sicurezza esistenti. Un'ulteriore criticità è la difficoltà, da parte dei modelli, nel comprendere logiche complesse del codice, con il rischio di generare *patch* errate o incomplete, spesso eccessivamente attagliate a scenari specifici e quindi poco efficaci nel caso generale. **La ricerca dovrebbe concentrarsi sul miglioramento della comprensione del contesto, ridurre i falsi positivi e scalabilità, oltre all'integrazione efficace nei framework di sviluppo e sicurezza esistenti**.

#3.1.4 - Sottoargomenti

OT - 3.1.4.1

Individuazione e correzione delle vulnerabilità di sicurezza nel *firmware* dei dispositivi per OT

Il *firmware* è essenziale per il corretto funzionamento delle apparecchiature da campo negli impianti industriali OT, come centrali nucleari, stazioni della rete elettrica o gasdotti. Un aspetto fondamentale nella gestione di questi dispositivi è mantenere aggiornato il firmware per correggere eventuali vulnerabilità. Nei casi in cui il sistema OT funzioni come una "scatola nera", diventa necessario sviluppare meccanismi e approcci che permettano il *virtual patching* delle vulnerabilità, tramite API fornite dai produttori. Se le vulnerabilità nel firmware non vengono individuate, tali difetti possono propagarsi rapidamente negli impianti industriali o lungo la catena di approvvigionamento globale. Tuttavia, il ciclo operativo tipico dei sistemi OT è estremamente lungo, e l'interruzione del servizio comporta un costo non trascurabile per il proprietario dell'impianto. **Pertanto, la ricerca dovrebbe essere orientata allo sviluppo di metodi per applicare le patch di sicurezza identificate mentre il sistema è in funzione (*hot-patching*). La ricerca potrebbe anche esplorare strategie alternative di mitigazione, tra cui controlli compensativi, tecniche di segmentazione o metodi di *retrofitting*, per affrontare vulnerabilità persistenti negli asset industriali.**

3.3.2 SUBAREA#3.2: SICUREZZA NEI SISTEMI OPERATIVI E TECNOLOGIE DI VIRTUALIZZAZIONE

Oltre ad essere eseguiti su un sistema operativo *host* (installazione su *bare metal*), oggi i programmi sfruttano sempre più il paradigma del *cloud computing*, appoggiandosi a servizi sistemistici infrastrutturali (*Infrastructure o Containers-as-a-Service* – IaaS o CaaS, rispettivamente), a servizi di piattaforme computazionali (*Platform-as-a-Service* – PaaS), oppure a servizi applicativi (*Software-as-a-Service* – *SaaS*)[60].

[60] Quest'area di ricerca è strategicamente rilevante non solo per la cybersecurity, ma anche per abilitare la progettazione, l'implementazione e la gestione delle tecnologie di virtualizzazione in condizioni di sovranità tecnologica.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#3.2.1] la sicurezza dei nuovi approcci alla virtualizzazione**, in grado di orchestrare il carico computazionale dei programmi in maniera sempre più trasparente. Dal punto di vista della sicurezza cibernetica, tali paradigmi divengono di massima importanza nel momento in cui sono utilizzati per orchestrare programmi per il calcolo e la comunicazione nel *continuum cloud-edge*. Pertanto, essi richiedono studi approfonditi su approcci di sicurezza *by design* che includano, ad esempio, la progettazione e l'applicazione di politiche di controllo degli accessi basate sui ruoli. Si sottolinea che, ai fini della resilienza, è importante focalizzare l'attenzione sull'*edge cloud*, che rappresenta un'infrastruttura vitale per il Paese nell'assicurare la continuità delle applicazioni a bassa latenza a dispetto di potenziali guasti o attacchi alle dorsali di rete internazionali;
- **[Argomento#3.2.2] la progettazione di nuovi modelli di protezione per la sicurezza delle piattaforme e l'aggiornamento continuo dei modelli esistenti**, dove quest'ultimo obiettivo costituisce un'attività cruciale data la già menzionata rapida evoluzione dei sistemi operativi e delle piattaforme di virtualizzazione. Pertanto, sono richiesti ulteriori sforzi nell'adozione di approcci di *hardening* basati su linee guida di implementazione sicura (*Security Technical Implementation Guidelines* – STIG) e nella progettazione di nuovi ed efficaci meccanismi di *observability* per la sicurezza, eventualmente accoppiati con l'*hardware* sottostante (cfr. Sezione 3.4.1).

3.3.3 SUBAREA#3.3: BLOCKCHAIN

La *blockchain* costituisce il primo e più diffuso esempio di tecnologia a registro distribuito (*Distributed Ledger Technology* – DLT), ovvero un sistema distribuito che consente ai partecipanti (definiti *nodi*) di inviare, validare e conservare informazioni su una base di dati condivisa, senza dipendere da un nodo centrale. A seconda del modello di autorizzazione adottato per la rete di nodi, una DLT può prevedere libera ammissione (*permissionless DLT*) o restrizioni all'ammissione (*permissioned DLT*) alla rete di nodi. In particolare, la *blockchain* implementa una DLT in cui gli insiemi di dati conservati formano una catena di blocchi collegati tra loro tramite *hash* crittografici, i quali dimostrano in modo inequivocabile lo storico delle transazioni che avvengono tra i nodi della rete. Dati i numerosi possibili casi d'uso che potrebbero beneficiare della *blockchain*, la comunità di R&I è estremamente attiva in questo campo.

In questa subarea, oltre alle questioni menzionate in precedenza riguardo lo sviluppo sicuro di software e piattaforme, gli argomenti prioritari sono i seguenti:

- **[Argomento#3.3.1] le soluzioni ai problemi di sicurezza degli algoritmi di consenso e *mining***, ad esempio, *peer flooding*, *network partitioning*, *block reorganization* e *Byzantine faults*;
- **[Argomento#3.3.2] le tecnologie per *smart contract*** che comprendano la progettazione e lo sviluppo di nuovi linguaggi di programmazione sicuri (cfr. Sezione 3.3.1) in grado di assicurare l'efficacia dell'implementazione di *smart contract* nei nodi della blockchain;

- **[Argomento#3.3.3] le tecniche di crittografia e anonimizzazione per *blockchain***, al fine di proteggere, dove applicabile, le transazioni pubblicate sulla *blockchain* e le identità delle parti coinvolte, specialmente in relazione alla minaccia posta dal calcolo quantistico alle suite crittografiche asimmetriche tradizionali (**cfr. Argomento#1.2.1**);
- **[Argomento#3.3.4] lo studio dell'economia delle blockchain**, che comprenda l'identificazione di casi d'uso appropriati per la *blockchain*, anche a seconda delle politiche di autorizzazione sopra menzionate.

3.4 Area#4: Sicurezza delle Infrastrutture Digitali

Le infrastrutture digitali ricoprono un ruolo importante per la società dal momento che abilitano lo scambio di informazioni (siano caratterizzate dalla presenza di dati personali o meno, (cfr. Sezione 3.1) tra diverse istanze di software e piattaforme (cfr. Sezione 3.3). Per questo motivo, è fondamentale tenere in dovuta considerazione la sicurezza degli apparati ICT che compongono le infrastrutture digitali per assicurare la resilienza cibernetica dei sistemi informativi nel loro complesso.

L'area *Sicurezza delle Infrastrutture Digitali* include due subaree, descritte in seguito.

3.4.1 SUBAREA#4.1: HARDWARE

Le infrastrutture digitali sono primariamente costituite da componenti hardware, quali dispositivi per gli utenti e apparati per centri di elaborazione dati che offrono servizi *on-premise* oppure in cloud.

Mentre la sicurezza del software e delle piattaforme che operano su tale hardware sono trattate in Sezione 3.3, gli argomenti prioritari riguardanti specificamente la sicurezza hardware sono i seguenti:

- **[Argomento#4.1.1] lo studio di vulnerabilità e attacchi indotti dall'hardware**, che comprende, ad esempio, la progettazione e lo sviluppo di metodi di test a scatola chiusa (*black box testing*) o aperta (*white box testing*), inclusa la verifica formale dell'hardware, ai fini di rilevare hardware malevolo e *hardware trojan*, e l'identificazione di nuovi attacchi a livello hardware che sfruttino canali laterali (*side channels*) per sottrarre informazioni sensibili.

#4.1.1 - Sottoargomenti

GAI - 4.1.1.1

Uso della GPAI per il rilevamento di *Hardware Trojan* - HT

La crescente complessità dei *System-on-Chip* - SoC, in particolare per i sistemi *embedded* e i dispositivi IoT, ha spinto i produttori a fare affidamento su strumenti di *Electronic Design Automation* e su *Intellectual Property cores* di terze parti (3P-EDA/IP) per rispettare vincoli stringenti di tempo e costo. Tuttavia, questa dipendenza introduce rischi significativi, tra cui il possibile inserimento malevolo di *Hardware Trojan* - HT nella fase di progettazione, ovvero nel momento in cui le vulnerabilità sono più facilmente sfruttabili e meno onerose da introdurre. Con l'avanzare del design verso l'implementazione fisica, l'individuazione e la rimozione degli HT diventa più complessa e costosa, rendendo fondamentali contromisure nelle fasi iniziali. **I modelli GPAI offrono un approccio promettente per affrontare questo problema, confrontando la documentazione di progetto a livello *Register Transfer Level* - RTL con schemi e descrizioni in linguaggio naturale degli HT, al fine di rilevarne l'eventuale presenza. L'applicazione in questo ambito è tuttavia attualmente limitata dalla scarsa capacità di generalizzazione dei modelli rispetto a schemi progettuali o pattern di attacco non presenti nei dati di addestramento.**

QT - 4.1.1.1

Messa in sicurezza delle installazioni di QKD

Mentre a livello teorico i protocolli di QKD sono sicuri anche contro attaccanti con infinita capacità computazionale o nel caso di future svolte innovative in termini di algoritmi di crittoanalisi, la loro implementazione pratica, la quale richiede hardware dedicato, introduce non idealità che portano i protocolli a deviare dal modello teorico. **Pertanto, si incoraggia la ricerca di nuove vulnerabilità e attacchi all'hardware nelle installazioni di QKD**.

- **[Argomento#4.1.2] lo studio di approcci di sicurezza cibernetica ancorati all'hardware (*hardware-anchored*)**, che possano sostanzialmente rafforzare il senso di fiducia (*trust*) verso algoritmi di autenticazione. I TEE rientrano anch'essi nell'ambito di interesse. È altresì incoraggiato lo sviluppo di primitive di sicurezza hardware, tra cui le funzioni fisiche non clonabili (*physically unclonable functions* – PUF) per la generazione di firme univoche per i circuiti integrati;
- **[Argomento#4.1.3] la ricerca su nuovi modelli per il rilevamento sicuro dei dati (*safe sensing*)**, ai fini di assicurare la generazione sicura di dati da parte di nodi sensori e la loro integrità durante la fase di trasmissione grazie a tecnologie IoT affidabili.

#4.1.3 - Sottoargomento

QT - 4.1.3.1 **Modelli per il rilevamento quantistico sicuro**	La sensoristica quantistica ha un grande potenziale per applicazioni ancora da esplorare, grazie ai miglioramenti che garantisce in termini di precisione e sensibilità rispetto ai sensori tradizionali. Le tecnologie di sensoristica quantistica si trovano a diversi stadi di sviluppo, talvolta anche nelle fasi iniziali di validazione. **In tale contesto, si incoraggia lo sviluppo di sistemi di sensori quantistici sicuri *by design*, al fine di prevenire e mitigare l'impatto di attacchi cibernetici**.

- **[Argomento#4.1.4] la progettazione di architetture hardware aperte** ai fini di rafforzare il senso di fiducia nell'hardware e supportare l'ecosistema industriale nazionale e, più in generale, europeo, prevenendo la contraffazione dell'hardware e la dipendenza tecnologica da fornitori non affidabili (*vendor lock-in*).

3.4.2 SUBAREA#4.2: RETI

Le reti di telecomunicazioni consentono lo scambio sicuro di informazioni tra sistemi. Mentre i nodi di una rete consistono sia di componenti hardware (cfr. Sezione 3.4.1) che software (cfr. Sezione 3.3), le reti meritano una considerazione a parte, con una subarea dedicata, a motivo della peculiarità e criticità delle funzioni che svolgono e che le rendono obiettivo di minacce specifiche (cfr. Sezione 3.2.2).

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#4.2.1] il miglioramento della sicurezza e della resilienza di rete**, perseguendo diversi obiettivi. Anzitutto, assicurare l'evoluzione della sicurezza nelle tecnologie *wireless* più diffuse come, ad esempio, identificazione a radiofrequenza (*Radio-Frequency Identification* – RFID), *Near-Field Communication* (NFC), sistemi radar, sistemi satellitari globali di navigazione (*Global Navigation Satellite Systems* – GNSS) e reti sottomarine. In secondo luogo, accelerare l'adozione dei principali standard Internet e la progettazione di nuovi meccanismi per il rilevamento degli attacchi cibernetici, compresi approcci basati su capacità di machine learning, al fine di incorporarli in nuovi protocolli di rete *by design*. Infine, favorire la creazione di servizi di rete diversificati, quali, ad esempio, il *Domain Name System* (DNS), per il rafforzamento dell'autonomia tecnologica;

- **[Argomento#4.2.2] nuovi approcci per la sicurezza dei sistemi IoT**, che comprendono, ad esempio, la sicurezza degli aggiornamenti firmware *over-the-air* (OTA), il rilevamento di attacchi di *jamming* e relative contromisure, approcci per la sicurezza allo strato fisico (*physical layer security*)[61], la diversificazione delle piattaforme IoT, il tutto, al fine di stabilire canali di comunicazione autenticati e resilienti considerando i tipici vincoli IoT in termini di energia, *storage*, latenza e larghezza di banda.

#4.2.2 - Sottoargomento

OT - 4.2.2.1 **Miglioramento della robustezza del sistema di gestione delle chiavi crittografiche per OT integrate con IIoT**	Implementare un sistema di gestione delle chiavi crittografiche (*Key Management System* - KMS) che possa abilitare la cifratura in ambito OT e, nello specifico, in sistemi che integrano reti di sensori industriali (IIoT), pone importanti sfide a causa della specificità di tali sistemi. **È incoraggiata la ricerca su soluzioni di KMS scalabili e resilienti, che siano adatte al contesto OT, indirizzando sfide quali la generazione sicura delle chiavi, la loro distribuzione, rotazione e memorizzazione. Tali soluzioni dovranno, inoltre, considerare la natura spesso geograficamente distribuita delle reti di sistemi OT, i vincoli di disponibilità e la conformità alla normativa.**

- **[Argomento#4.2.3] il rafforzamento della sicurezza delle reti di telecomunicazioni cellulari**, favorendo l'adozione della sicurezza *by design* a livello di specifiche tecniche (cfr. Sezione 3.6.3) e a livello implementativo, sviluppando, ad esempio, sistemi di collaudo atti a rilevare potenziali vulnerabilità legate alle caratteristiche della tecnologia di rete mobile;
- **[Argomento#4.2.4] lo sviluppo di infrastrutture per la comunicazione quantistica *(quantum communication infrastructures* - QCI)**, allo scopo di costruire una rete con tecnologia nazionale e, in generale, europea, la quale consenta di trasmettere informazioni tramite crittografia con garanzie di sicurezza fondate sulla teoria dell'informazione (QKD, cfr. Sezione 3.1.2) come scudo contro gli attacchi cibernetici.

[61] La sicurezza allo strato fisico sfrutta l'intrinseca aleatorietà del canale radio per mettere in sicurezza le trasmissioni senza affidarsi solamente ai metodi crittografici tradizionali.

#4.2.4 - Sottoargomento

QT - 4.2.4.1 **Evoluzione della QCI verso l'Internet quantistico**	A livello nazionale ed europeo sono state avviate diverse iniziative per la progettazione e lo sviluppo di una infrastruttura QCI diffusa. **Oltre a incoraggiare il mondo della ricerca a partecipare a tali iniziative, si invita a focalizzarsi sui passi evolutivi che seguiranno, andando nella direzione della creazione di un Internet quantistico, sicuro *by design*, che consentirà lo scambio di informazione tramite stati quantistici. Nello specifico, una delle principali sfide tecniche relative alle installazioni attuali di QCI, dedicate interamente alla QKD, riguarda l'estensione del raggio di trasmissione e la convivenza su mezzi trasmissivi non dedicati.**

3.5 Area#5: Aspetti della Società

Secondo i più recenti report sulla minaccia cibernetica[2, 62], la compromissione dell'identità e il furto di credenziali sono sempre più diffusi. L'ingegneria sociale continua a rappresentare un fattore determinante in questo tipo di incidenti, con i criminali informatici che sfruttano campagne di *phishing*, attacchi *brute-force* e *credential stuffing* per ottenere accessi non autorizzati e trarre profitto economico. L'aumento degli attacchi incentrati sull'identità è correlato alla diffusione di tecniche di ingegneria sociale, come l'uso di *link staging*, il *traffic filtering* e lo sfruttamento di servizi cloud legittimi per occultare attività malevole. Le piattaforme di social media vengono frequentemente sfruttate per creare esche mirate, mentre anche i canali di comunicazione a minore visibilità sono sempre più presi di mira.

L'area *Aspetti della Società* include tre subaree, descritte in seguito.

62 Si veda (in inglese) 2025 Data Breach Investigations Report (verizon.com).

3.5.1 SUBAREA#5.1: ASPETTI UMANI

Questa subarea si focalizza principalmente sulle persone come soggetti che rischiano di subire un attacco cibernetico, cercando modi per migliorare la loro postura di cybersicurezza. D'altro canto, questo non può prescindere da un'analisi delle motivazioni e degli incentivi che si celano dietro il comportamento degli attaccanti, richiedendo un approccio interdisciplinare con un forte accoppiamento con la tecnologia (cfr. Sezione 3.2.1).

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#5.1.1] la ridefinizione dei confini dell'interazione uomo-macchina e i relativi rischi di sicurezza**, anche a seguito dell'adozione delle tecnologie dirompenti[63], che richiedono l'integrazione della sicurezza all'intersezione tra percezione, cognizione e attuazione;
- **[Argomento#5.1.2] il miglioramento della percezione del rischio di attacchi cibernetici** da parte dei cittadini, sfruttando, ad esempio, modelli psicologici;

#5.1.2 - Sottoargomento

GAI - 5.1.2.1 **Miglioramento della percezione del rischio nell'uso dei modelli GPAI**	Garantire che gli utenti percepiscano e comprendano in modo accurato i rischi associati agli output della GPAI è fondamentale per una sua adozione responsabile. A causa di *bias* cognitivi, infatti, molti utenti tendono a sovrastimare l'affidabilità delle risposte dell'IA, soprattutto quando queste sono formulate in tono persuasivo o autorevole, con il rischio di sviluppare una dipendenza e di farne un uso improprio. Questa dipendenza può tradursi nel c.d. *automation bias*, che si manifesta come errore di omissione, quando l'utente decide di non verificare la validità di una risposta, oppure di esecuzione, quando accetta e agisce sulla base della risposta dell'IA anche in contrasto con le proprie conoscenze. Tali problematiche risultano particolarmente critiche laddove l'utente non possieda le competenze necessarie per valutare criticamente gli output.

[63] Quali, ad es., i modelli di IA utilizzati in sistemi robotici, che integrano sempre di più capacità di ragionamento multimodale e comprensione dell'ambiente circostante. Tale trasformazione espande la superficiedi attacco e richiede ricerca dedicata su strategie di sicurezza integrate che considerino l'interazione fisica e i requisiti di tempestività operativa.

#5.1.2 - Sottoargomento

Per affrontare queste sfide sono necessari sforzi multidisciplinari che spaziano dallo studio dell'interazione uomo-macchina, finalizzato a progettare interfacce che favoriscano un uso critico, alla psicologia cognitiva, volta ad analizzare e mitigare i *bias*, fino all'etica, per orientare l'implementazione dell'IA secondo principi morali. Infine, rivestono un ruolo cruciale gli approcci educativi, essenziali per formare gli utenti sui limiti della GPAI e incoraggiarli a verificarne le risposte.

- **[Argomento#5.1.3] le metodologie per la prevenzione degli attacchi di ingegneria sociale** e riduzione dei rischi legati all'uomo, anche considerando i recenti avanzamenti nei modelli generativi del linguaggio, che, ad esempio, permettono di creare messaggi e-mail maliziosi in maniera automatica (cfr. Sezione 3.2.2);

#5.1.3 - Sottoargomento

GAI - 5.1.3.1

Prevenzione e mitigazione degli attacchi di ingegneria sociale e delle minacce all'identità digitale supportati da GPAI

La GPAI abilita tecniche di frode sempre più sofisticate che minacciano la sicurezza dell'identità digitale e la fiducia nei servizi digitali. Oltre a generare contenuti ingannevoli e manipolatori (quali ad es. messaggi di *phishing*), è possibile realizzare *deepfake*, clonazioni vocali e identità sintetiche per impersonare individui o organizzazioni, anche grazie alla capacità di inferire, a partire da contenuti presenti online, dati personali quali posizione, reddito e genere. Queste tecniche rafforzano gli attacchi di *social engineering*, rendendoli più difficili da rilevare, con impatti economici e reputazionali rilevanti, incluse violazioni dei sistemi biometrici in settori come *e-government*, sanità e finanza. **Sono necessari approcci innovativi per sviluppare contromisure efficaci, come soluzioni per la verifica sicura dell'identità digitale, tecniche avanzate di rilevamento dei *deepfake*, strumenti adattivi per prevenire le frodi e assicurare la gestione sicura delle credenziali digitali. È inoltre essenziale la collaborazione interdisciplinare con sociologi, psicologi e linguisti, per creare metodologie e strumenti di sensibilizzazione mirati a diverse tipologie di utenti.**

- **[Argomento#5.1.4] favorire l'accettazione di politiche e tecnologie di sicurezza da parte degli utenti finali**, prendendo in considerazione il paradigma della "sicurezza utilizzabile" (*usable security*) ed estendendo l'approccio basato sulla tecnologia con aspetti psicologici;
- **[Argomento#5.1.5] la profilazione dell'attaccante**, identificando le categorie di attaccanti e utilizzando tale informazione in fase sia di prevenzione che di risposta.

3.5.2 SUBAREA#5.2: ASPETTI FORMATIVI

La forza lavoro nel dominio della cybersicurezza manca di un sufficiente numero di professionisti, che siano dotati delle competenze e della formazione necessarie ad affrontare le sfide correnti. Ciò comporta un enorme rischio per le imprese, il Governo italiano e l'intera società.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#5.2.1] la definizione di metodologie per la valutazione e lo sviluppo di competenze di cybersicurezza e per l'identificazione delle competenze per le professioni legate alla cybersicurezza**;
- **[Argomento#5.2.2] lo studio di nuovi modelli e strumenti per la consapevolezza (*awareness*) della cybersicurezza**, quali, ad esempio, *serious games*, che possano sostenere dal punto di vista tecnologico campagne informative atte a sensibilizzare i cittadini nelle loro esperienze quotidiane di sperimentazione dello spazio cibernetico, andando verso una cultura dello *zero trust*;
- **[Argomento#5.2.3] lo studio di nuovi modelli e strumenti per la formazione sulla cybersicurezza**, in considerazione della complessità di tale mondo, che sempre più richiede strumenti didattici dedicati come, ad esempio, i *cyber-range*, che forniscano un ambiente sicuro e legale per acquisire competenze pratiche sulla sicurezza cibernetica. Un esempio di contesto che richiede nuovi modelli e strumenti formativi è quello relativo alle OT, in ragione della presenza di tecnologie dedicate e di figure professionali, quali operatori OT con una formazione di base altamente specializzata su impianti e sistemi di controllo, che però non copre necessariamente in modo adeguato gli aspetti di cybersicurezza.

3.5.3 SUBAREA#5.3: ASPETTI LEGALI

Sta emergendo con forza la necessità di iniziative di carattere normativo che contrastino l'utilizzo improprio della tecnologia, la distribuzione illecita di materiale sottratto durante una violazione dei dati (*data breach*), potenzialmente coperto da diritti di proprietà intellettuale, e il cybercrimine.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#5.3.1] regole e principi etici per uno spazio cibernetico sicuro**, che riguardino, ad esempio, i confini nell'utilizzo di applicazioni di intelligenza artificiale[64];

#5.3.1 - Sottoargomento

GAI - 5.3.1.1

Codifica dei principi etici nell'allineamento (*alignment*) dei modelli

Garantire che i modelli GPAI siano allineati a principi etici e valori umani è una sfida complessa che richiede una collaborazione interdisciplinare, in particolare, tra esperti di sociologia, etica e governance. Gli attuali tentativi di codificare valori come utilità, innocuità e onestà nei modelli GPAI incontrano difficoltà legate a una tendenza a privilegiare sistemi valoriali tecnicamente più trattabili, rischiando di trascurare principi più complessi da codificare ma altrettanto rilevanti. Affrontare questo problema richiede sia un miglioramento della fattibilità tecnica della codifica di principi eterogenei, sia un'analisi critica e inclusiva dei principi effettivamente rappresentati. Inoltre, la gestione dei compromessi, ad esempio, tra riduzione dei rischi (intesi, principalmente nei riguardi della *safety*) e utilità (prestazioni) del modello, impone la definizione di quadri solidi di prioritizzazione e la mediazione. **La ricerca futura dovrà concentrarsi sull'elaborazione di framework etici capaci di integrare prospettive diversificate, riducendo al contempo i *bias* derivanti dalla semplificazione tecnica. Ciò include lo sviluppo di meccanismi di governance che garantiscano una selezione dei principi equa e sistematicamente giustificabile, nonché l'ideazione di approcci dinamici in grado di adattarsi all'evoluzione delle norme sociali.**

[64] Come previsto dall'AI Act europeo, si enfatizzano principi chiave quali la trasparenza, la responsabilità e la tutela dei diritti fondamentali. Tali principi sono essenziali per mitigare i rischi associati ai sistemi di IA, in particolare nei settori ad alto rischio, e per promuovere uno spazio digitale che rispetti la privacy, eviti forme di discriminazione e garantisca la sicurezza digitale per tutti gli utenti.

- **[Argomento#5.3.2] modelli e regole per favorire l'autonomia strategica**, quali, ad esempio, lo "*European Chips Act*"[65], le "*US reviews of critical supply chains*"[66] e la normativa "*Golden Power*" italiana[67];
- **[Argomento#5.3.3] modelli e regole per il governo del cyberspazio** inclusivi, ad esempio, di una disciplina normativa del rischio, di disposizioni sulla disinformazione nello spazio cibernetico, di regolamenti internazionali ed europei che garantiscano una *leadership* del cyberspazio sicuro e regole per meglio gestire scenari transfrontalieri (cfr. Sezione 3.2.4).

#5.3.3 - Sottoargomento

GAI - 5.3.3.1

Regolamentazione efficace e dinamica per la governance della GPAI

La corretta governance della GPAI è essenziale per una sua integrazione sicura nella società. Tuttavia, gli approcci normativi e di governance devono adattarsi al rapido avanzamento delle capacità della GPAI e alla loro diffusione sociale. Se da un lato una *governance* efficace richiede di scoraggiare pratiche irresponsabili, garantire l'attribuzione delle responsabilità (*accountability*) in caso di danni e superare inefficienze burocratiche, dall'altro rimane aperta la questione relativa a se e in che modo gli organismi di regolazione debbano incidere sul ritmo del progresso dell'IA, evitando di gravare in modo sproporzionato sugli attori più piccoli e consentendo al contempo l'innovazione. **Ciò implica lo sviluppo di modelli di *governance* adattivi e proporzionati, quali *framework* di compliance progressivi[68] e *regulatory sandboxes* orientati alla sostanza della conformità basata sul rischio, piuttosto che a obblighi puramente formali. Di importanza fondamentale è anche affrontare l'influenza delle grandi imprese tecnologiche sull'accademia, sulle politiche e sulla ricerca. Serve infine ulteriore ricerca multidisciplinare per comprendere l'impatto dei diversi meccanismi di *governance*, inclusi gli approcci basati sulle risorse computazionali (*compute governance*).**

[65] Regolamento (UE) 2023/1781 del Parlamento europeo e del Consiglio del 13 settembre 2023 che istituisce un quadro di misure per rafforzare l'ecosistema europeo dei semiconduttori (Chips Act). Gazzetta Ufficiale dell'Unione Europea – EUR-Lex (europa.eu)

[66] Si faccia riferimento, ad esempio, al rapporto (in inglese) Quadriennal Supply Chain Review 2021–2024 (bidenwhitehouse.archives.gov – PDF, 8.7MB).

[67] Il Decreto-Legge 21/2012 è volto alla tutela degli asset strategici dell'Italia in settori quali difesa, sicurezza, energia, trasporti, comunicazioni e tecnologie critiche (queste ultime definite dal Decreto-Legge 23/2020, in conformità a quanto previsto dall'articolo 4, paragrafo 1, del Regolamento (UE) 2019/452 (europa.eu). La normativa attribuisce al governo poteri speciali di monitoraggio e intervento nelle operazioni societarie che coinvolgono questi settori. L'obiettivo è salvaguardare la sicurezza nazionale, garantire la continuità operativa delle infrastrutture critiche e proteggere l'autonomia strategica del Paese, in particolare da investimenti esteri non trasparenti o potenzialmente dannosi.

[68] È particolarmente importante considerare le sfide specifiche che le PMI affrontano nel conformarsi ai framework normativi, ad esempio, nella direzione dello sviluppo di protocolli di conformità semplificati.

3.6 Area#6: Aspetti di Governo

In considerazione della varietà di minacce alla sicurezza cibernetica e della complessità del loro trattamento (cfr. Sezione 3.2), sono stati proposti diversi sistemi di gestione della sicurezza delle informazioni tra i quali si segnalano lo standard ISO/IEC 27001[53] (e correlati), il *Cybersecurity Framework* del NIST[69] e il Framework Nazionale per la Cybersecurity e la Data Protection[70].
Si evidenzia come l'efficacia di tali sistemi di gestione è massima solo se sono applicati a tutti i processi di un'organizzazione e dal numero maggiore possibile di organizzazioni sia del settore pubblico che di quello privato.

L'area Aspetti di Governo include tre subaree, descritte di seguito.

3.6.1 SUBAREA#6.1: ASPETTI ORGANIZZATIVI

Questa subarea tratta le politiche di indirizzo, i processi e le procedure che, tramite appropriate metodologie e strumenti, garantiscono che un sistema di gestione della sicurezza delle informazioni persegua obiettivi di sicurezza fissati.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#6.1.1] la ricerca sull'economia dell'ecosistema della cybersicurezza** e, in particolare, delle catene di approvvigionamento (*supply chain*);
- **[Argomento#6.1.2] l'innovazione nelle metodologie per il raggiungimento degli obiettivi di cybersicurezza**, al fine di implementare politiche di governo che includano la possibilità di difesa attiva contro gli attacchi cibernetici, nonchè lo sviluppo di strategie di *capacity building* per rafforzare la resilienza cibernetica a livello nazionale e organizzativo.

[69] Si veda il sito (in inglese) Cybersecurity Framework (csrc.nist.gov).

[70] Si veda il sito Framework Nazionale per la Cybersecurity e la Data Protection | Framework Nazionale per la Cyber Security e la Data Protection (cybersecurityframework.it).

#6.1.2 - Sottoargomento

QT - 6.1.2.1

Metodologie per la transizione alla crittografia basata su QT

Secondo importanti istituzioni internazionali[71,72], la preparazione alla minaccia quantistica dovrebbe essere considerata un aspetto integrante della gestione del rischio cibernetico. **Pertanto, si incoraggia la progettazione di nuovi processi organizzativi finalizzati all'individuazione degli schemi crittografici non robusti contro la minaccia quantistica e l'implementazione di strategie per la migrazione a schemi basati sulle tecnologie quantistiche, sia in termini di PQC sia di QKD a seconda degli specifici casi d'uso[73]. Tali metodologie di transizione dovrebbero considerare l'impatto di potenziali nuovi vettori d'attacco contro gli algoritmi crittografici, sia post-quantum sia quantistici (cfr. sottoargomenti QT-6.3.2.1 e QT-6.3.2.2), nonché l'allineamento con i vincoli operativi, sfruttando, ad esempio, schemi crittografici ibridi (cfr. Argomento#1.2.1). Dovrebbe, inoltre, essere tenuta in considerazione l'adozione di implementazioni nazionali o, almeno, provenienti da Paesi *like-minded.***

- **[Argomento#6.1.3] l'innovazione nelle metodologie per gli audit di cybersicurezza**;
- **[Argomento#6.1.4] l'innovazione nei modelli a supporto della continuità operativa (*business continuity*) e del recupero dal disastro (*disaster recovery*)** per rafforzare la resilienza cibernetica.

3.6.2 SUBAREA#6.2: GESTIONE DEL RISCHIO

Come parte degli aspetti di governo, una subarea cruciale è la gestione dei rischi associati alla possibilità che una o più minacce alla sicurezza cibernetica (cfr. Sezione 3.2) sfruttino vulnerabilità presenti su uno o più apparati ICT, causando un danno al singolo individuo o all'intera organizzazione, sia essa parte del settore pubblico o privato.

[71] La legge statunitense riguardo la migrazione dei sistemi informativi delle agenzie federali alla PQC - si veda (in inglese) H.R.7535 - 117th Congress (2021-2022): Quantum Computing Cybersecurity Preparedness Act (congress.gov)- è diventata Public Law il 21 dicembre 2022.

[72] Nell'aprile 2024, la Commissione Europea ha rilasciato la Raccomandazione (EU) 2024/1101(europa.eu), a sostegno di un approccio coordinato tra gli Stati Membri per la migrazione delle infrastrutture e dei servizi digitali alla PQC, enfatizzando la necessità di sincronizzare i piani nazionale e garantendo l'interoperabilità transfrontaliera. A giugno 2025, gli Stati Membri dell'Unione Europea, con il supporto della Commissione, hanno pubblicato una tabella di marcia (europa.eu) con tempistiche di dettaglio per attuare una transizione coordinata alla PQC.

[73] Le strategie di transizione alla crittografia basata sulle tecnologie quantistiche riflettono le raccomandazioni del CRA sulla gestione del rischio nel ciclo di vita delle tecnologie digitali e sulla mitigazione proattiva delle minacce di cybersicurezza che si possono prevedere, incluse quelle che derivano dagli avanzamenti tecnologici in campo quantistico.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#6.2.1] lo studio e il miglioramento continuo dei quadri procedurali per la valutazione del rischio**[74], al fine di facilitare l'identificazione, l'analisi e la quantificazione del rischio rendendo quindi più efficace il relativo piano di trattamento (e l'accettazione del rischio ove ragionevole). In particolare, rientrano in questo argomento di R&I i modelli per la valutazione d'impatto;
- **[Argomento#6.2.2] lo studio e il miglioramento continuo della misurazione di *key performance indicator* (KPI) sulla sicurezza cibernetica**, inclusi strumenti per la loro raccolta e visualizzazione, al fine di facilitare il monitoraggio e la revisione del rischio[75];

#6.2.2 - Sottoargomenti

GAI - 6.2.2.1

Nuove metriche di *safety* e gestione del bilanciamento tra *safety* e prestazioni

Le metriche di *safety* attualmente utilizzate per la GPAI si concentrano spesso sull'evitare *output* ritenuti dannosi da valutatori umani, generalmente tramite approcci semi-automatizzati di *red-teaming*. Sebbene utili, tali metriche non coprono in modo adeguato dimensioni più ampie, come la *corrigibility* (disponibilità del modello ad accettare correzioni), la trasparenza (comprensibilità del comportamento del modello) e la robustezza rispetto a input che esulano dalla distribuzione dei dati di addestramento (*out-of-distribution robustness*). Emergono inoltre compromessi tra sicurezza e prestazioni: ad esempio, vincolare capacità potenzialmente rischiose del modello può ridurne l'utilità, e il *safety fine-tuning* può limitare la diversità delle risposte. **La ricerca futura dovrà dare priorità allo sviluppo di metriche robuste e sensibili al contesto, capaci di misurare in modo affidabile questi compromessi e di supportare meccanismi per l'adattamento dinamico del bilanciamento tra safety e prestazioni in base al contesto d'impiego. L'elaborazione di metodologie valutative alternative, come i *framework* di analisi comportamentale, potrà migliorare la comparabilità dei risultati e la comprensione dei compromessi nelle diverse applicazioni. Infine, sarà importante indagare come questi *trade-off* varino tra diverse tipologie di sistemi GPAI (ad esempio, agenti autonomi rispetto ad assistenti conversazionali), per supportare interventi di sicurezza più mirati.**

[74] Importanti quadri procedurali esistenti sono (in inglese), le Special Publication del NIST 800-30, 800-53, 800-82 (nist.org) e lo standard ISA 62443 (3-2) (isa.org)

[75] Inoltre, è di particolare interesse lo sviluppo di KPI a supporto della valutazione dei rischi sistemici e delle dipendenze della *supply chain*, come ulteriormente dettagliato nell'**Argomento#6.2.3**.

#6.2.2 - Sottoargomenti

OT - 6.2.2.1 **KPI per la gestione del rischio di cybersicurezza nei sistemi OT**	Per affrontare le minacce cibernetiche nei sistemi OT, è necessario sviluppare modelli integrati capaci di identificare i rischi relativi all'interconnessione tra diverse infrastrutture, nonché agli effetti a cascata delle interruzioni di servizio. **Un'importante sfida di ricerca consiste nell'integrare un sistema di misurazione di KPI specifici per le componenti OT all'interno del sistema di gestione del rischio di un'organizzazione.**

- **[Argomento#6.2.3] la gestione del rischio e l'analisi delle reti di catene di approvvigionamento (*supply chain network*)**, al fine di sviluppare modelli di rischio che considerino i possibili effetti a cascata derivanti da eventi avversi cibernetici che coinvolgono tali reti. Questi modelli dovrebbero evidenziare rilevanti dipendenze funzionali tra soggetti e prodotti, sfruttando, ad esempio, le SBOM dei prodotti ICT (cfr. Sezione 3.3.1) e, in questo modo, essere applicabili anche a livello nazionale e internazionale (cfr. Sezione 3.2.3). Tali studi completano le analisi strategiche menzionate in precedenza, dipingendo un quadro esaustivo del rischio per settore economico[76], incluso quello legato alla catena d'approvvigionamento (cfr. Sezioni 3.5.3 e 3.6.1);
- **[Argomento#6.2.4] la gestione dei rischi sistemici derivanti dalle tecnologie pervasive**, derivanti dall'adozione su larga scala di tecnologie pervasive e dirompenti, quali, ad es. la GPAI o biosensori, con enfasi sull'identificazione, valutazione e mitigazione delle problematiche legate all'adozione massiva di tali tecnologie a livello sociale. Ciò include lo sviluppo di quadri di governo degli impatti sulla sicurezza fisica (*safety*) e sull'incolumità di individui e comunità, nonché la definizione di metodologie e strumenti per monitorare, gestire e mitigare danni potenziali e rischi a cascata, garantendo un impiego sicuro e un comportamento affidabile dei sistemi e degli artefatti coinvolti.

[76] I rischi cyber possono avere caratteristiche differenti nei diversi settori, in particolare, in relazione all'*outsourcing* dei servizi digitali. Ad esempio, nel settore finanziario, il Digital Operational Resilience Act – DORA (europa.eu), il regolamento UE sulla resilienza digitale del sistema finanziario, affronta esplicitamente le minacce legate alle dipendenze da terze parti tipicamente presenti in questo settore, quali ad es. FinTech (servizi finanziari digitali) e RegTech (tecnologie per la conformità normativa).

#6.2.4 - Sottoargomenti

GAI - 6.2.4.1

Migliorare la comprensione delle capacità (*capabilities*) dei modelli e la loro interpretabilità

Garantire l'affidabilità della GPAI richiede lo sviluppo di strumenti formali per comprendere le capacità dei modelli e l'interpretabilità del loro comportamento. Una sfida centrale è distinguere tra capacità intrinseche (*capabilities*) e comportamenti indotti da strumenti esterni (es. il c.d. *scaffolding*), che possono portare a sovrastimare le reali capacità dei modelli e complicare la valutazione della loro sicurezza. Anche i metodi di interpretabilità esistenti, basati su rappresentazioni interne o su spiegazioni in linguaggio naturale, sono spesso limitati da astrazioni deboli e disallineamenti concettuali che comportano una mancanza di fedeltà alle reali modalità di ragionamento del modello.

È incoraggiato lo sviluppo di metodologie solide per separare capacità intrinseche e indotte dal contesto, benchmark per distinguere le cause di fallimento (*failure*) relative a limitazioni di capacità rispetto a problemi di allineamento e metodi di interpretabilità scalabili e con garanzie di consistenza. È inoltre rilevante approfondire lo studio delle modalità di ragionamento non-deduttive (comuni nei GPAI ma ancora poco comprese) e l'utilizzo di approcci neuro-simbolici[77]. Infine, il miglioramento degli approcci di allineamento delle rappresentazioni (*representation alignment*) e di sanitizzazione degli spazi latenti interni (*latent spaces*) dei modelli è cruciale per rafforzare sia interpretabilità che sicurezza.

GAI - 6.2.4.2

Sfide nella valutazione (*evaluation*) dei modelli GPAI

La valutazione delle prestazioni dei modelli GPAI, presenta criticità che si manifestano lungo tutte le fasi dello sviluppo dei sistemi, dalla raccolta dei dati e l'addestramento del modello fino alla fase di *deployment*, legate a fenomeni come la *prompt sensitivity* (variazione dei risultati al variare della formulazione dell'input) e la *test-set contamination*, che si verifica quando il modello è esposto inavvertitamente ai dati di valutazione durante l'addestramento, che compromettono l'affidabilità delle metriche. A queste si aggiungono i *bias* nelle valutazioni umane, dovuti a soggettività e incoerenza tra annotatori.

[77] Tali approcci rappresentano una prospettiva promettente, ma la loro integrazione con i modelli GPAI resta una questione aperta.

#6.2.4 - Sottoargomenti

Queste problematiche evidenziano vulnerabilità sistemiche nei processi di sviluppo e valutazione. Linee di ricerca rilevanti includono la progettazione di benchmark robusti, tecniche di *data attribution* e metodi ibridi per combinare valutazioni umane e automatiche. È inoltre essenziale definire protocolli valutativi più rigorosi, rappresentativi e replicabili, per garantire misurazioni affidabili in contesti d'uso reali.

GAI - 6.2.4.3

Approcci robusti di *safety fine-tuning* per la GPAI

Uno degli obiettivi principali del *safety fine-tuning* è rimuovere capacità potenzialmente indesiderate da un modello GPAI, indirizzandone il comportamento verso risultati conformi a quanto atteso. Tuttavia, i metodi di *fine-tuning* attualmente disponibili faticano spesso a eliminare in modo efficace tali comportamenti, che possono essere riattivati tramite *fine-tuning* malevolo o tecniche di *jailbreaking*. Inoltre, il *safety fine-tuning* viene generalmente effettuato su una distribuzione molto più ristretta rispetto a quella del pre-addestramento (*pretraining*), lasciando i modelli vulnerabili ad attacchi che sfruttano lacune nella generalizzazione. Ad esempio, un avversario può utilizzare testo codificato o lingue poco rappresentate per aggirare i meccanismi di sicurezza. **La ricerca futura dovrebbe esplorare nuove tecniche, come gli approcci meccanicistici (*mechanistic approaches*), per ottenere modifiche mirate dello stato interno del modello e abilitare processi di *unlearning* selettivo di conoscenze dannose o obsolete[78]. Vi è inoltre urgente bisogno di metodi capaci di generalizzare in modo efficace e di evitare il ricorso a strategie di allineamento superficiali che falliscono in presenza di condizioni avversarie. Infine, è incoraggiato lo sviluppo di benchmark per valutare e confrontare in modo rigoroso i metodi di *fine-tuning*, così da identificare le tecniche più efficaci per garantire la sicurezza dei modelli GPAI, assicurandone la robustezza rispetto a minacce note ed emergenti.**

[78] Vedi anche il sottoargomento GAI-6.2.2.1 sull'aspetto correlato della gestione del compromesso tra sicurezza e prestazioni.

#6.2.4 - Sottoargomenti

GAI - 6.2.4.4 **Controllo dell'autonomia dei sistemi a singolo e multi-agente**	I modelli GPAI possono essere potenziati per dare origine ad agenti in grado di pianificare e agire autonomamente nel mondo reale, svolgendo compiti in modo proattivo. Tuttavia, l'aumento dell'autonomia, la supervisione limitata e gli orizzonti d'azione più estesi introducono nuovi rischi, con potenziali implicazioni sulla sicurezza. Nei sistemi GPAI multi-agente, la complessità di garantire la *safety* aumenta significativamente. I rischi includono sia fallimenti correlati, dovuti a vulnerabilità condivise tra agenti, sia forme di collusione che possono portare a effetti indesiderati o dannosi. Sebbene la cooperazione tra agenti possa generare benefici, obiettivi divergenti possono causare disallineamenti reciproci o rispetto ai valori umani. **La ricerca futura dovrà concentrarsi sulla quantificazione e il controllo delle capacità di apprendimento continuo (*lifelong learning*), migliorando i meccanismi di *safety assurance* al manifestarsi di eventuali capacità non previste. Sarà inoltre cruciale comprendere come *pretraining*, *prompt engineering* e *fine-tuning* influenzino i comportamenti emergenti in contesti multi-agente. In questa direzione, assumono particolare rilievo strategie come il *diversity-enhanced training* e l'impiego di meccanismi automatizzati per il rilevamento di fenomeni di collusione.**

3.6.3 SUBAREA#6.3: STANDARDIZZAZIONE

Molti sistemi complessi del mondo digitale richiedono che valga il requisito di interoperabilità tra diversi produttori. Al fine di assicurare il raggiungimento dell'interoperabilità, è necessario promuovere e supportare politiche e standard aperti internazionali, quali, ad esempio, le specifiche tecniche emanate dal 3GPP per le reti di telecomunicazioni cellulari.

In questa subarea, gli argomenti prioritari sono i seguenti:

- **[Argomento#6.3.1] l'innovazione nella standardizzazione dei processi aziendali per la cybersicurezza**, sfruttando, ad esempio, i *Capability Maturity Model* (CMM) e i modelli di sicurezza *zero trust*;

- **[Argomento#6.3.2] lo sviluppo di standard di sicurezza cibernetica specifici per le singole tecnologie**[79], al fine di monitorare e rilevare proattivamente possibili vulnerabilità *by design* introdotte nelle specifiche tecniche da attori ostili[80];

#6.3.2 - Sottoargomenti

GAI - 6.3.2.1

Interoperabilità tra gli standard di sicurezza per la GPAI e gli standard settoriali

Sono in corso sforzi rilevanti per lo sviluppo di nuovi standard di sicurezza per i sistemi di intelligenza artificiale, in particolare, per i GPAI, con un'attenzione specifica ad approcci orizzontali e trasversali. Tali *framework* puntano a requisiti generali, come robustezza, misure di cybersicurezza e tracciabilità, con l'obiettivo di garantire un'ampia applicabilità intersettoriale. Tuttavia, molti ambiti applicativi critici si basano su standard tecnici settoriali consolidati, profondamente integrati nei rispettivi contesti operativi. **La ricerca dovrebbe indagare metodologie per allineare i requisiti trasversali ai diversi settori con le specifiche proprie di ciascun ambito, tenendo conto delle aspettative settoriali in termini di metriche di performance e pratiche operative (come ad es. i requisiti di *logging*) nell'integrare gli standard di sicurezza GPAI nei framework operativi. Ciò implica lo studio di metodologie per l'allineamento tra requisiti trasversali e specifiche settoriali, la gestione di eventuali conflitti o lacune, e la promozione di una convergenza fluida senza compromettere l'integrità dei *framework* esistenti.**

OT - 6.3.2.1

Progettazione di standard di cybersicurezza specifici per OT

L'assenza di meccanismi di sicurezza standard per le OT, specialmente se operanti in contesti critici, pone rischi significativi. Ad esempio, gli attuali metodi di autenticazione, come quelli implementati nei *Programmable Logic Controllers* (PLC), sono spesso soluzioni proprietarie che non garantiscono l'interoperabilità tra diversi costruttori. Sebbene gli standard come l'ISA/IEC 62443[29] forniscano fondamenta solide per talune tecnologie OT, essi non indirizzano le sfide chiave di interoperabilità, i vincoli dei sistemi in tempo reale o l'integrazione con *framework* per la cybersicurezza a livello aziendale generale.

[79] In questa direzione, le iniziative includono la mappatura degli standard e dei *framework* di cybersecurity, con l'obiettivo di razionalizzare le best practice internazionali, facilitarne l'adozione da parte delle organizzazioni e sostenere approcci di conformità armonizzati.

[80] Si veda il rapporto del CISA (in inglese) Potential threat vectors to 5G Infrastructure (cisa.gov - PDF, 5MB).

#6.3.2 - Sottoargomenti

	La sfida di ricerca principale in questo contesto consiste nello sviluppare un quadro unificato ed efficace di standardizzazione dei processi di cybersicurezza che sia *vendor-agnostic*.
QT - 6.3.2.1 **Favorire la standardizzazione dei protocolli di QKD**	Dal momento che uno dei principali vantaggi della standardizzazione è l'introduzione di un rigoroso scrutinio dei meccanismi crittografici, **si incoraggia lo sviluppo di standard aperti per i protocolli di QKD, in maniera analoga a quanto già avviene per la PQC, al fine di migliorarne la postura di sicurezza (**cfr. Argomento#1.2.2**)**.
QT - 6.3.2.2 **Ampliamento della standardizzazione dei protocolli di PQC**	La standardizzazione dei protocolli di PQC è stata avviata dalla competizione internazionale organizzata dal NIST. D'altro canto, altre iniziative di standardizzazione sono state avviate in parallelo a quella del NIST. In tale contesto, **si incoraggia la partecipazione a iniziative di standardizzazione, al fine di favorire l'adozione pervasiva della PQC, consentendo, allo stesso tempo, uno scrutinio continuo della sicurezza dei protocolli sia contro futuri attacchi quantistici sia contro potenziali attacchi classici non ancora scoperti a causa della loro recente introduzione**.

- **[Argomento#6.3.3] la formalizzazione delle specifiche tecniche** scritte in linguaggio naturale attraverso linguaggi formali.

3.7 Panoramica delle Aree di R&I

La Tabella 3-1 fornisce una panoramica strutturata della tassonomia individuata nell'Agenda, dettagliando ciascuna area con le relative subaree, argomenti di ricerca e sottoargomenti.
Questa rappresentazione offre una visione gerarchica che facilita la comprensione dell'ambito tematico e delle specifiche direttrici di ricerca nel dominio della R&I per la cybersicurezza.

Tabella 3-1: Panoramica delle Aree di R&I.

Aree	Subaree	Argomenti	Sottoargomenti NEW
1. SICUREZZA DEI DATI E PRIVACY	1.1 Ingegneria della protezione dei dati	1.1.1 Elaborazione *privacy-preserving* dei dati 1.1.2 Memorizzazione *privacy-preserving* dei dati 1.1.3 Autenticazione, autorizzazione e controllo dell'accesso con garanzie di privacy aggiuntive	GAI - 1.1.1.1 Bilanciamento tra privacy e prestazioni dei modelli GPAI nelle fasi di addestramento e inferenza
	1.2 Crittografia	1.2.1 UPDATED Primitive, algoritmi e protocolli per PQC 1.2.2 UPDATED Crittografia quantistica 1.2.3 UPDATED Schemi di FHE 1.2.4 Funzioni crittografiche di hashing	
	1.3 Trusted information sharing	1.3.1 Arricchimento delle soluzioni standard di interoperabilità semantica con controlli di sicurezza 1.3.2 Soluzioni architetturali per *trusted information sharing e storage*	

Aree	Subaree	Argomenti	Sottoargomenti NEW
2. GESTIONE DELLE MINACCE CIBERNETICHE	**2.1 Attacco e difesa**	**2.1.1** Tecniche di attacco e misure di difesa **2.1.2** Rilevamento e risposta contro i *malware* **2.1.3** Messa in sicurezza di algoritmi e modelli per il *machine learning* **2.1.4.** *Adversarial behaviour*	**GAI - 2.1.1.1** Contrasto all'uso malevolo dei modelli GPAI **QT - 2.1.1.1** Rilevamento delle intrusioni di rete per sistemi OT **OT - 2.1.1.1** Sensoristica quantistica per migliorare il rilevamento degli attacchi **GAI - 2.1.3.1** Difesa contro le minacce avversarie
	2.2 Cyberthreat intelligence	**2.2.1** Machine learning per la *cyberthreat intelligence* **2.2.2** Metodologie per monitorare la disinformazione **2.2.3** Attribuzione di attacchi alla sicurezza cibernetica	**GAI - 2.2.1.1** Uso della GPAI per la rilevazione e la mitigazione delle capacità di evasione (*evasion*) del malware **GAI - 2.2.1.2** Uso della GPAI per supportare la *cyber threat intelligence* e la rilevazione degli incidenti
	2.3 Gestione degli incidenti e operazioni di sicurezza	**2.3.1** UPDATED Resilienza dei sistemi cyber-fisici **2.3.2** Automazione via *machine learning* dei sistemi tradizionali **2.3.3** Strategie di difesa	**OT - 2.3.1.1** *Digital Twin* delle infrastrutture critiche comprendenti OT interconnesse e settoriali
	2.4 Scienze forensi digitali	**2.4.1** Processi di scienze forensi digitali **2.4.2** Contromisure per aggirare le tecniche antiforensi	

Aree	Subaree	Argomenti	Sottoargomenti NEW
3. SICUREZZA DEL SOFTWARE E DELLE PIATTAFORME	**3.1 Sicurezza nello sviluppo e test del software**	**3.1.1** Linguaggi di programmazione *security-aware* **3.1.2** Gestione sicura del ciclo di vita del software **3.1.3** UPDATED Paradigma dell'*open cloud* **3.1.4** Tecniche per rilevare vulnerabilità di sicurezza nel *firmware* e nei file binari	**GAI - 3.1.2.1** Migliorare la sicurezza del codice generato con il supporto della GPAI **QT - 3.1.2.1** Messa in sicurezza dell'ingegneria del software quantistico **GAI - 3.1.4.1** Uso della GPAI per il rilevamento e il *patching* delle vulnerabilità **OT - 3.1.4.1** Individuazione e correzione delle vulnerabilità di sicurezza nel *firmware* dei dispositivi per OT
	3.2 Sicurezza nei sistemi operativi e tecnologie di virtualizzazione	**3.2.1** Sicurezza dei nuovi approcci alla virtualizzazione **3.2.2** Modelli di protezione per la sicurezza delle piattaforme	
	3.3 Blockchain	**3.3.1** Problemi di sicurezza degli algoritmi di consenso e *mining* **3.3.2** Tecnologie per *smart contract* **3.3.3** Tecniche di crittografia e anonimizzazione **3.3.4** Economia delle *blockchain*	

Aree	Subaree	Argomenti	Sottoargomenti NEW
4. SICUREZZA DELLE INFRASTRUTTURE DIGITALI	4.1 Hardware	**4.1.1** Vulnerabilità e attacchi indotti dall'hardware **4.1.2** Approcci di sicurezza cibernetica ancorati all'hardware **4.1.3** Modelli per il rilevamento sicuro dei dati **4.1.4** Architetture hardware aperte	**GAI - 4.1.1.1** Uso della GPAI per il rilevamento di *Hardware Trojan* **QT - 4.1.1.1** Messa in sicurezza delle installazioni di QKD **QT - 4.1.3.1** Modelli per il rilevamento quantistico sicuro
	4.2 Reti	**4.2.1** Miglioramento della sicurezza e della resilienza di rete **4.2.2** Approcci per la sicurezza dei sistemi IoT **4.2.3** Rafforzamento della sicurezza delle reti di telecomunicazioni cellulari **4.2.4** Sviluppo di infrastrutture per la QCI	**OT - 4.2.2.1** Miglioramento della robustezza del sistema di gestione delle chiavi crittografiche per OT integrate con sistemi IIoT **QT - 4.2.4.1** Evoluzione della QCI verso l'Internet quantistico

Aree	Subaree	Argomenti	Sottoargomenti NEW
5. ASPETTI DELLA SOCIETÀ	**5.1 Aspetti umani**	**5.1.1** UPDATED Ridefinizione dei confini dell'interazione uomo-macchina **5.1.2** Miglioramento della percezione del rischio **5.1.3** Metodologie per la prevenzione degli attacchi di ingegneria sociale **5.1.4** Favorire l'accettazione di politiche e tecnologie di sicurezza da parte degli utenti finali **5.1.5** Profilazione dell'attaccante	**GAI - 5.1.2.1** Miglioramento della percezione del rischio nell'uso dei modelli GPAI **GAI - 5.1.3.1** Prevenzione e mitigazione degli attacchi di ingegneria sociale e delle minacce all'identità digitale supportati da GPAI
	5.2 Aspetti formativi	**5.2.1** Metodologie per la valutazione e lo sviluppo di competenze di cybersicurezza e per l'identificazione delle competenze per le professioni legate alla cybersicurezza **5.2.2** Modelli e strumenti per la consapevolezza della cybersicurezza **5.2.3** UPDATED Modelli e strumenti per la formazione sulla cybersicurezza	

Aree	Subaree	Argomenti	Sottoargomenti NEW
	5.3 Aspetti legali	5.3.1 Regole e principi etici per uno spazio cibernetico sicuro 5.3.2 UPDATED Modelli e regole per favorire l'autonomia strategica 5.3.3 Modelli e regole per il governo del cyberspazio	GAI - 5.3.1.1 Codifica dei principi etici nell'allineamento (*alignment*) dei modelli GAI - 5.3.3.1 Regolamentazione efficace e dinamica per la governance della GPAI
6. ASPETTI DI GOVERNO	6.1 Aspetti organizzativi	6.1.1 Economia dell'ecosistema della cybersicurezza e delle catene di approvvigionamento 6.1.2 UPDATED Metodologie per il raggiungimento degli obiettivi di cybersicurezza 6.1.3 Metodologie per gli audit di cybersicurezza 6.1.4 Continuità operativa e recupero dal disastro	QT - 6.1.2.1 Metodologie per la transizione alla crittografia basata su QT

Aree	Subaree	Argomenti	Sottoargomenti NEW
6. ASPETTI DI GOVERNO	**6.2 Gestione del rischio**	**6.2.1** Quadri procedurali per la valutazione del rischio **6.2.2** KPI sulla sicurezza cibernetica **6.2.3** Gestione del rischio e analisi delle reti di catene di approvvigionamento **6.2.4** NEW Gestione dei rischi sistemici derivanti dalle tecnologie pervasive	**GAI - 6.2.2.1** Nuove metriche di safety e gestione del bilanciamento tra *safety* e prestazioni **OT - 6.2.2.1** KPI per la gestione del rischio di cybersicurezza nei sistemi OT **GAI - 6.2.4.1** Migliorare la comprensione delle capacità (*capabilities*) dei modelli e la loro interpretabilità **GAI - 6.2.4.2** Sfide nella valutazione (*evaluation*) dei modelli GPAI **GAI - 6.2.4.3** Approcci robusti di *safety fine-tuning* per la GPAI **GAI - 6.2.4.4** Controllo dell'autonomia dei sistemi a singolo e multi-agente
	6.3 Standardizzazione	**6.3.1** Standardizzazione dei processi aziendali per la cybersicurezza **6.3.2** Sviluppo di standard di sicurezza cibernetica specifici per le singole tecnologie **6.3.3** Formalizzazione delle specifiche tecniche	**GAI - 6.3.2.1** Interoperabilità tra gli standard di sicurezza per la GPAI e gli standard settoriali **QT - 6.3.2.1** Favorire la standardizzazione dei protocolli di QKD **OT - 6.3.2.1** Progettazione di standard di cybersicurezza specifici per OT **QT - 6.3.2.2** Ampliamento della standardizzazione dei protocolli di PQC

4 Aree di R&I e EDT

4

Aree di R&I e EDT

A valle dell'illustrazione delle aree di R&I, identifichiamo a seguire la lista delle EDT che sono rilevanti per lo studio dei vari argomenti e sottoargomenti facenti parte delle aree. Le EDT considerate sono per la maggior parte attribuibili al contesto dell'ICT e appartengono a diversi livelli di astrazione, spaziando da generici paradigmi a tecniche di dettaglio. Si sottolinea che alcune EDT potrebbero essere considerate come argomenti di ricerca a sé stanti, come, ad esempio, il *machine learning*; tuttavia, è stato ritenuto d'interesse trattarle indipendentemente dagli argomenti di R&I identificati, per rendere esplicito come esse contribuiscono alla ricerca sulla cybersicurezza. Inoltre, alcune EDT sono parte di domini più ampi: ad esempio, insieme all'intelligenza artificiale *general-purpose* (GPAI), contribuiscono al dominio più ampio dell'intelligenza artificiale (AI) anche data science, machine learning, rappresentazione della conoscenza, robotica e sistemi autonomi[81].

In questa sezione è disponibile una lista di EDT ritenute rilevanti, assieme a una descrizione per ciascuna di esse. Quindi, le EDT sono messe in relazione con le subaree di R&I, evidenziando come esse contribuiscono a indirizzare (o supportando o vincolando) gli argomenti all'interno delle subaree.

4.1 EDT Rilevanti: Descrizioni

La Tabella 4-1 contiene la lista delle EDT con la rispettiva descrizione.

[81] La *data science* consente l'analisi esplorativa, la preparazione e la visualizzazione dei dati, principalmente attraverso approcci statistici; la rappresentazione della conoscenza permette all'AI di modellare e ragionare sulle informazioni; il *machine learning* e il *deep learning* costituiscono il motore dell'adattamento e dell'apprendimento dai dati; mentre la robotica e i sistemi autonomi portano l'intelligenza artificiale nel mondo fisico, permettendole di svolgere compiti in modo indipendente.

Tabella 4-1: descrizioni delle EDT

EDT (ordine alfabetico in inglese)	Descrizione
***5G and beyond mobile networks* - Reti mobili 5G e successive**	La quinta generazione delle reti wireless a lungo raggio per terminali mobili ha migliorato la qualità del servizio a banda larga e introdotto il supporto nativo per l'IoT industriale (si veda descrizione dell'IoT come EDT a sé stante) e l'utilizzo delle reti mobili a scopo non pubblico (*non-public networks*). Le successive generazioni (ovvero 5G-Advanced e sesta generazione) amplieranno ulteriormente le capacità della tecnologia, al fine di mantenere la promessa dell'*Internet of Everything*[82].
***Business Process Management (BPM)* - Gestione dei processi aziendali**	Insieme di principi, metodi e strumenti per progettare, analizzare e monitorare i processi aziendali, ovvero raccolte di eventi, attività e decisioni che coinvolgono un gran numero di attori e risorse e che portano collettivamente a un risultato di valore per l'organizzazione o i suoi clienti[83].
Data science	Approcci di analisi dei dati, inclusi modelli statistici e di machine learning (quest'ultimo anche trattato come EDT a sé stante); metodi di preparazione dei dati (*data preparation*), di qualità e integrazione dei dati, nonché di memorizzazione degli stessi.
DevSecOps	Sebbene introdotta recentemente, la pratica *DevOps* è stata ampiamente adottata da industrie e governi, grazie all'agilità che offre per mantenere il ritmo dell'innovazione. In particolare, l'approccio *DevSecOps* garantisce che la sicurezza sia affrontata in tutte le parti di tale pratica, incorporando controlli di sicurezza e generando automaticamente artefatti conformi alle regole di sicurezza.

[82] Per approfondimenti, si veda la pagina (in inglese) 3GPP 5G standard (3gpp.org - PDF. 750KB).

[83] Definizione tratta dal libro intitolato "*Fundamentals of Business Process Management*" di M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers. Per approfondimenti, si faccia riferimento alla pagina (in inglese): Lectures | Fundamentals of BPM (bpm.org).

EDT (ordine alfabetico in inglese)	Descrizione
***Distributed Ledger Technology (DLT)* - Tecnologia a registro distribuito**	Sistema distribuito che consente ai partecipanti (chiamati *nodi*) di inviare, validare e conservare informazioni su una base di dati condivisa, senza dipendere da un nodo centrale. A seconda del modello di autorizzazione adottato per la rete di nodi, una DLT può prevedere libera ammissione (*permissionless DLT*) o restrizioni all'ammissione (*permissioned DLT*)[58] alla rete di nodi.
***General-purpose Artificial Intelligence (GPAI)* - Intelligenza artificiale general-purpose**	Tecnologia alla base di una categoria di sistemi di intelligenza artificiale che eseguono un'ampia gamma di compiti in diversi ambiti, dimostrando adattabilità, versatilità e ampia applicabilità (a differenza della c.d. *task-specific* AI). Questa categoria include i modelli generativi, come i *Large Language Models* (LLM), tipicamente addestrati su vaste moli di dati e specializzati nell'elaborazione e generazione del linguaggio naturale, e i *Large Multimodal Models* (LMM), che combinano testo, immagini, audio e video. Basati su architetture di tipo *deep learning* questi sistemi vengono utilizzati in applicazioni che spaziano dalla creazione di contenuti all'automazione e all'elaborazione avanzata delle informazioni[84, 85].
***Hardware-based security* - Sicurezza basata sull'hardware**	Dominio tecnologico che mira a preservare la sicurezza di un bene ICT e dell'informazione ivi contenuta tramite componenti hardware ad hoc. Specifiche tecnologie che appartengono a questo dominio sono, ad esempio, *Hardware Security Module* (HSM) e *secure enclave*.
***High-Performance Computing (HPC)* - Calcolo a elevate prestazioni**	Pratica che consiste nell'aggregare potenza di calcolo in modo tale da raggiungere prestazioni molto superiori rispetto a quelle che si potrebbero ottenere da un computer tradizionale o una *workstation*, al fine di risolvere problemi complessi in ambito scientifico, ingegneristico o di business.

[84] Descrizione parzialmente tratta dal Considerando 99 dell'AI Act.

[85] L'EDT GPAI è stata introdotta come voce autonoma in questo aggiornamento, mentre nella precedente versione dell'Agenda era inclusa all'interno della EDT "*Machine Learning and Deep Learning*".

EDT (ordine alfabetico in inglese)	Descrizione
***Human-Computer Interaction (HCI)* - Interazione uomo-computer**	Studio dell'interazione tra persone e computer. Comprende: (i) *tecnologie di visualizzazione dei dati e delle informazioni*, ovvero, rappresentazioni computazionali, visuali e interattive dei dati, allo scopo di dare loro un significato, acquisire conoscenza, scoprire nuove informazioni e/o presentare efficacemente i risultati; (ii) *interfacce uomo-macchina*, ovvero, componenti hardware o software che consentono all'operatore umano di monitorare lo stato di un processo, modificare le configurazioni di controllo per cambiarne l'obiettivo ed escludere manualmente le operazioni di controllo automatico in caso d'emergenza[86].
***Internet of Things (IoT)* - Internet delle Cose**	Interconnessione di oggetti e dispositivi in grado di trasmettere e ricevere dati su una rete, offrendo così un nuovo livello di interazione e di controllo a distanza che fornisca soluzioni a richieste socioeconomiche, presenti o future, di servizi di rilevamento e monitoraggio, nonché di nuove applicazioni, modelli di business e settori industriali. Quando i dispositivi sono impiegati in processi industriali che richiedono una qualità del servizio molto alta, il paradigma viene definito come Industrial IoT (IIoT).
***Knowledge representation and reasoning* - Rappresentazione della conoscenza**	Campo di studio pertinente il dominio dell'intelligenza artificiale che si occupa di utilizzare simboli formali per rappresentare una raccolta di proposizioni ritenute vere da un agente[87].

[86] Descrizione tratta in parte da Human-Machine Interface (HMI) - Glossary (csrc.nist.gov) (in inglese).

[87] Descrizione tratta in parte dal libro intitolato "Knowledge Representation and Reasoning" di R. Brachman, and H. Levesque, The Morgan Kaufmann Series in Artificial Intelligence, Morgan Kaufmann Publishers, 2004 (in inglese).

EDT (ordine alfabetico in inglese)	Descrizione
Machine learning and deep learning	Il *machine learning* è il sottocampo dell'intelligenza artificiale definito come la capacità di una macchina di imitare un comportamento umano intelligente. Al momento della stesura di questo documento, una delle direzioni più promettenti consiste nelle reti neurali profonde (*deep neural network*) e i modelli generativi del linguaggio basati su di esse (esplicitamente ricompresi nella EDT GPAI).
***Mobile devices* - Dispositivi mobili**	Un dispositivo mobile è da intendersi come un dispositivo informatico portatile che: (i) ha un ridotto fattore di forma che permette che il dispositivo possa essere trasportato da un singolo individuo; (ii) è progettato per operare senza una connessione fisica (ad esempio, è in grado di trasmettere o ricevere informazioni wireless); (iii) possiede supporti di archiviazione dati locali, non rimovibili; (iv) è in grado di rimanere acceso per lunghi periodi di tempo, grazie a una fonte di alimentazione autonoma[88].
***Operational Technologies (OT)* - Tecnologie operative**	Sistemi programmabili che interagiscono con l'ambiente fisico monitorando processi ed eventi e/o introducendo dei cambiamenti nell'ambiente tramite nodi sensori e attuatori. Esempi di sistemi OT sono i sistemi di controllo industriale (ICS, che comprendono tecnologie, quali, ad esempio, sistemi SCADA - *Supervisory Control And Data Acquisition* e sistemi di controllo distribuiti (DCS - *Distributed Control Systems*), utilizzati per il controllo dei processi industriali)[89].
***Quantum technologies* - Tecnologie quantistiche**	Dominio tecnologico che sfrutta le proprietà della meccanica quantistica, quali, ad esempio, entanglement quantistico, sovrapposizione quantistica ed effetto tunnel, per innovare campi quali il calcolo, il monitoraggio e la crittografia. I tre pilastri principali delle tecnologie quantistiche sono il calcolo quantistico (*quantum computing*), la comunicazione quantistica (*quantum communication*) e la sensoristica quantistica (*quantum sensing*).

88 Descrizione tratta da mobile device - Glossary (csrc.nist.gov) (in inglese).

89 Descrizione tratta in parte da operational technology (OT) - Glossary (csrc.nist.gov) (in inglese).

EDT (ordine alfabetico in inglese)	Descrizione
Robotics and autonomous systems **- Robotica e sistemi autonomi**	Comprende architetture di sistema, algoritmi e strumenti software che mirano a realizzare procedure avanzate di automazione e agenti autonomi che siano in grado di adattarsi al cambiamento di condizioni, della conoscenza e dei vincoli che li circondano[90]. In quanto dominio tecnologico abilitante chiave per l'intelligenza artificiale, la robotica si sta evolvendo verso implementazioni su larga scala e distribuite di sistemi autonomi operanti in ambienti dinamici[91].
Satellite systems **- Sistemi satellitari**	Classe di sistemi di comunicazione senza fili che sfrutta costellazioni di satelliti artificiali al fine di creare canali di comunicazione tra trasmettitore e ricevitore posizionati in diversi punti sulla Terra, altrimenti impossibilitati a comunicare tra loro tramite tecnologie di rete terrestri. Recentemente, i sistemi satellitari stanno passando dalla tradizionale condizione di integrazione verticale a un aggregato di segmenti di proprietà e gestione indipendenti che determinano un sistema satellitare che include anche reti terrestri[92].
Virtual reality technologies **- Tecnologie per la realtà virtuale**	Classe di tecnologie che includono elementi visuali, uditivi e sensoriali, combinati per creare un'interfaccia verso un ambiente cibernetico completamente separato dal mondo reale.
Virtualisation and cloud **- Virtualizzazione e cloud**	La virtualizzazione è una tecnologia che consente agli utenti di creare molteplici ambienti simulati o di dedicare risorse a partire da un singolo sistema hardware fisico. Il paradigma del *cloud computing* si riferisce a un ambiente informatico che consente l'astrazione, il raggruppamento e la scalabilità delle risorse fisiche attraverso una rete. A seconda del modello di risorse computazionali offerte, l'ambiente cloud fornisce diversi livelli di virtualizzazione quali IaaS, CaaS, PaaS e SaaS.

[90] Descrizione tratta in parte dalla pagina (in inglese) Autonomous Systems & Robotics (nasa.gov).

[91] La diffusione su larga scala di sistemi robotici autonomi introduce sfide critiche in materia di sicurezza. A differenza delle piattaforme digitali, i robot compromessi rappresentano una minaccia fisica per le persone e per l'ambiente in cui operano.

[92] Si veda (in inglese) Hybrid Satellite Networks Cybersecurity (nccoe.nist.gov).

EDT (ordine alfabetico in inglese)	Descrizione
	Si parla di *edge computing* quando l'elaborazione è eseguita da un ambiente cloud in prossimità di una particolare origine dati per minimizzare la latenza.
Zero trust architecture	La proliferazione di alcune delle precedenti EDT, quali, ad esempio, *cloud computing*, dispositivi mobili e IoT, ha comportato il superamento del tradizionale modello di sicurezza basato sul perimetro di rete. Il paradigma di cybersicurezza della *zero trust* prevede di spostare le linee di difesa dal perimetro statico della rete, ai beni ICT e alle risorse, assumendo che la fiducia non possa essere concessa sulla base della sola posizione fisica o in rete[93].

4.2 Proposta di Corrispondenza tra Subaree e EDT

Questa sezione presenta la mappatura tra le EDT identificate e le sotto-aree di Ricerca e Innovazione (R&I) definite nell'Agenda. In particolare, la Tabella 4-2 offre una prospettiva guidata dalla tecnologia, in cui ogni EDT è mappata rispetto alle aree e sotto-aree di R&I rispetto alle quali è maggiormente rilevante, evidenziando i domini di ricerca specifici in cui ciascuna tecnologia emergente è destinata ad avere un impatto o nei quali richiederà sforzi dedicati di ricerca e innovazione nel contesto della cybersicurezza.

[93] Descrizione tratta in parte dalla pagina (in inglese) Zero Trust Architecture (nist.gov).

Tabella 4-2: relazioni tra ogni EDT (in inglese) e le subaree di R&I.

	Area#1 - Sicurezza dei dati e privacy			Area#2 - Gestione delle minacce cibernetiche				Area#3 - Sicurezza del software e delle piattaforme		
	Ing. della prot. dei dati	Crittografia	Trusted info. sharing	Attacco e difesa	CTI	Gestione degli incidenti e operazioni di sicurezza	Scienze forensi digitali	Sicurezza SW	Sicurezza SO e virtualiz.	Block-chain
5G and beyond mobile networks				X		X	X	X	X	
BPM	X		X	X		X	X	X	X	X
Data science	X		X	X	X	X	X			
DevSecOps		X				X		X	X	
DLT		X				X	X	X		X
General-purpose AI (GPAI)	X			X	X	X	X	X		
Hardware-based security	X	X							X	
HPC		X		X	X					X
HCI			X	X	X	X				
IoT	X			X	X	X		X	X	
Knowledge representation and reasoning	X		X	X	X	X				
Machine learning and deep learning	X		X	X	X	X	X	X		
Mobile devices	X	X	X	X		X	X	X	X	
Operational Technologies	X				X	X		X		
Quantum technologies		X		X				X		
Robotics and autonomous systems				X	X	X		X	X	
Satellite systems	X	X								
Virtual reality technologies								X	X	
Virtualisation and cloud	X	X	X	X		X		X	X	
Zero trust architecture	X		X	X		X			X	

	Area#4 - Sicurezza delle infrastrutture digitali		Area#5 - Aspetti della società			Area#6 - Aspetti di governo		
	HW	Reti	Aspetti umani	Aspetti form.	Aspetti legali	Aspetti organiz.	Gest. rischio	Std
5G and beyond mobile networks	X	X			X		X	X
BPM			X	X	X	X	X	X
Data science			X		X		X	X
DevSecOps	X	X						X
DLT					X	X	X	X
General-purpose AI (GPAI)	X		X	X	X	X	X	X
Hardware-based security	X	X						X
HPC	X	X						
HCI			X	X	X		X	
IoT	X	X	X				X	X
Knowledge representation and reasoning							X	X
Machine learning and deep learning	X	X	X		X		X	X
Mobile devices	X	X	X	X			X	X
Operational Technologies	X	X					X	X
Quantum technologies	X	X				X		X
Robotics and autonomous systems	X	X	X		X		X	
Satellite systems		X						
Virtual reality technologies	X	X	X	X			X	
Virtualisation and cloud		X			X	X	X	X
Zero trust architecture	X	X	X	X	X	X	X	X

LISTA DEGLI ACRONIMI

3GPP	Third Generation Partnership Project
5G	Fifth Generation (Quinta Generazione di reti mobili)
ACN	Agenzia per la Cybersicurezza Nazionale
APT	Advanced Persistent Threat
BPM	Business Process Management
CaaS	Containers-as-a-Service
CIA	Confidentiality, Integrity, Availability
CINI	Consorzio Interuniversitario Nazionale per l'Informatica
CISA	(US) Cybersecurity and Infrastructure Security Agency
CyBOK	The Cyber Security Body of Knowledge
CMM	Capability Maturity Model
DDoS	Distributed Denial of Service
D.L.	Decreto Legge
DCS	Distributed Control Systems
DLT	Distributed Ledger Technology
DNS	Domain Name System
EDT	Emerging and Disruptive Technology
ENISA	European Union Agency for Cybersecurity
EU	European Union
FHE	Fully Homomorphic Encryption
GDPR	(EU) General Data Protection Regulation
GNSS	Global Navigation Satellite Systems
GPAI	General-Purpose Artificial Intelligence
HCI	Human-Computer Interaction
HPC	High-Performance Computing

HSM	Hardware Security Module
HT	Hardware Trojan
IA	Intelligenza Artificiale
IaaS	Infrastructure-as-a-Service
ICS	Industrial Control Systems
ICT	Information and Communication Technology
IIoT	Industrial Internet of Things
IoT	Internet of Things
IPS	Intrusion Prevention System
JRC	(EU) Joint Research Centre
KPI	Key Performance Indicator
MUR	Ministero dell'Università e della Ricerca
NATO	North Atlantic Treaty Organization
NFC	Near-Field Communication
NIST	(US) National Institute of Standards and Technology
OS	Operating System
OT	Operational Technology
OTA	Over-The-Air
PaaS	Platform-as-a-Service
PET	Privacy-Enhancing Technologies
PII	Personally Identifiable Information
PLC	Programmable Logic Controllers
PNR	Programma Nazionale per la Ricerca
PQC	Post-Quantum Cryptography
PUF	Physically Unclonable Function
QCI	Quantum Communication Infrastructure
QKD	Quantum Key Distribution
QT	Quantum Technologies

R&I	Research and Innovation
RDDos	Ransom Distributed Denial of Service
RFID	Radio-Frequency Identification
SaaS	Software-as-a-Service
SBOM	Software Bill of Materials
SCADA	Supervisory Control and Data Acquisition
SIEM	Security Information and Event Management
SMC	Secure Multi-party Computation
SME	Small-Medium Enterprise
SOAR	Security Orchestration, Automation and Response
SoC	System-on-Chip
SOC	Security Operations Center
SSDLC	Secure Software Development LifeCycle
STIG	Security Technical Implementation Guideline
TEE	Trusted Execution Environment
TTP	Tactics, Techniques and Procedures
UK	United Kingdom
US	United States of America

Appendice

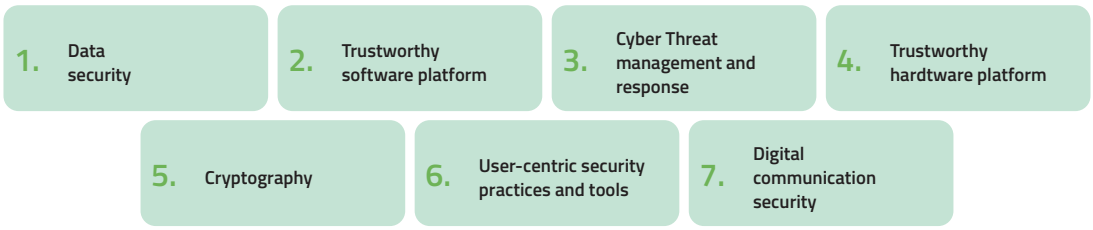
La classificazione delle aree principali del dominio di conoscenza della cybersicurezza si basa su quattro riferimenti chiave:

- le attività in corso a livello nazionale da parte dei partenariati pubblico-privati che stanno implementando la Componente 2 ("Dalla Ricerca all'Impresa") della Missione 4 ("Istruzione e Ricerca") di *Italia Domani* - Piano Nazionale di Ripresa e Resilienza[94], in particolare, il progetto SERICS - *Security and Rights in the CyberSpace*[95].
- *Cybersecurity Research Directions for the EU's Digital Strategic Autonomy*[96] dell'ENISA;
- *European Cybersecurity Centres of Expertise Map – Definitions and Taxonomy*[97] del JRC;
- *CyBOK – The Cyber Security Body of Knowledge*[98], un'iniziativa internazionale coordinata dal Regno Unito.

In particolare, sono state considerate:

- **la classificazione di SERICS:**

- **la classificazione di ENISA:**



94 Per approfondimento, si veda il sito Istruzione e ricerca - Italia Domani (gov.it).

95 Si veda il sito Fondazione SERICS – SEcurity and RIghts in the CyberSpace (serics.eu).

96 Disponibile (in inglese) al sito Cybersecurity Research Directions for the EU's Digital Strategic Autonomy - ENISA (europa.eu).

97 Disponibile (in inglese) al sito JRC Publications Repository - European Cybersecurity Centres of Expertise Map - Definitions and Taxonomy (europa.eu).

98 Disponibile (in inglese) al sito (cybok.org).

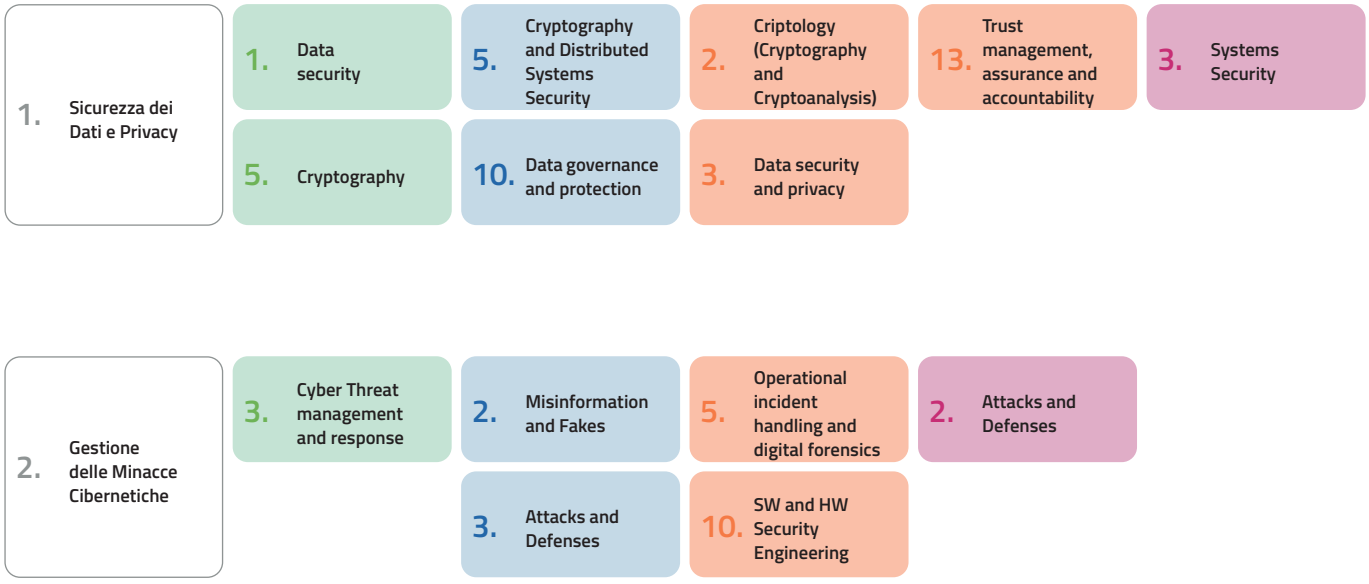
la classificazione del JRC:

1. Assurance, Audit and Certification
2. Criptology (Cryptography and Cryptoanalysis)
3. Data security and privacy
4. Education and Training
5. Operational incident handling and digital forensics
6. Human aspects
7. Identity and Access Management (IAM)
8. Security Management and governance
9. Network and Distributes Systems
10. SW and HW Security Engineering
11. Legal Aspects
12. Theoretical foundations
13. Trust management, assurance and accountability

la classificazione del CyBOK:

1. Human Organisational & Regulatory Aspects
2. Attacks and Defenses
3. Systems Security
4. Software and Platform Security
5. Infrastructure Security

La corrispondenza tra le iniziative di classificazione menzionate in precedenza e le aree di R&I identificate in questo documento è mostrata nelle figure seguenti.



Legenda SERICS ENISA JRC CyBOK

3. Sicurezza del Software e delle Piattaforme	2. Trustworthy software platform	4. Operating Sistems and Virtualization Security	7. Identity and Access Management (IAM)	12. Theoretical foundations	3. Systems Security	
		5. Cryptography and Distributed Systems Security	10. SW and HW Security Engineering	13. Trust management, assurance and accountability	4. Software and Platform Security	
		6. Software and Platform Security				
4. Sicurezza delle Infrastrutture Digitali	4. Trustworthy hardware platform	7. Infrastructure Security	9. Network and Distributes Systems	5. Infrastructure Security		
	7. Digital communication security	9. Securing Digital Transformation	10. SW and HW Security Engineering			
5. Aspetti della Società	6. User-centric security practices and tools	1. Human, Social and Legal Aspects	4. Education and Training	6. Human aspects	1. Human Organisational & Regulatory Aspects	
			5. Operational incident handling and digital forensics	11. Legal Aspects		
6. Aspetti di Governo	8. Risk Management and Governance	1. Assurance, Audit and Certification	1. Human Organisational & Regulatory Aspects			
		8. Security Management and governance				

Legenda SERICS ENISA JRC CyBOK


REPUBBLICA ITALIANA
AGENZIA PER LA CYBERSICUREZZA NAZIONALE
ACN
SAPIENTIA ADIUVAT
MUR
Ministero dell'Università e della Ricerca