

Progetto di Ricerca e Sviluppo

ONCOLOGIA

Applicazione delle tecnologie
AI e **ML** per un'analisi integrata
e predittiva in campo oncologico



INDICE

1. Il contesto di riferimento	3
2. Il progetto Oncologia	7
3. Le aree di innovazione di Oncologia	9
Il modello di erogazione in cloud	9
Il Data Lake oncologico	10
Gli algoritmi predittivi in ambito clinico	11
Le Identità Digitali Sanitarie (IDS)	13
Gli algoritmi predittivi per la sanità territoriale	15
4. Conclusioni	18
Gli attori di Oncologia	20

1. Il contesto di riferimento

Negli ultimi anni, le informazioni digitali di tutto il mondo sono più che raddoppiate e questa tendenza è destinata ad aumentare in modo esponenziale, generando la disponibilità di enormi volumi di dati elettronici, i.c.d. Big Data. Parallelamente, è in atto una rivoluzione culturale, tecnologica e scientifica centrata sull'applicazione dell'apprendimento automatico (Machine Learning - ML) che utilizza banche dati di imponenti dimensioni. Contestualmente, l'Intelligenza Artificiale (IA) e i sistemi di apprendimento automatico stanno dimostrando, ancora una volta, che le macchine sono in grado di analizzare grandi quantità di dati più velocemente ed efficacemente di quanto potrebbe fare un essere umano e, se opportunamente programmate, di generare dei modelli predittivi.

La ricerca e lo sviluppo delle tecnologie afferenti alla IA stanno ampliando il paradigma culturale anche nell'ambito della medicina "di precisione", ossia la medicina personalizzata per singolo paziente. Le relative applicazioni potrebbero diventare in breve tempo indispensabili per fornire rapide risposte in contesti a elevata complessità e incertezza, consentendo ai medici di liberarsi del lavoro mnemonico e ripetitivo e di poter disporre di elementi valutativi e decisionali atti a fare le scelte migliori per il paziente, anche in termini di bisogni assistenziali.

In questo senso, i sistemi di IA devono essere considerati veri e propri strumenti, come il microscopio, il fonendoscopio o l'elettrocardiografo, sviluppati nel tempo per amplificare la capacità percettiva dei medici con rilevazioni biologiche più approfondite. L'obiettivo sarà raggiunto quando l'IA costituirà un reale e costante supporto a tutto il personale sanitario, fornendo un "secondo set di occhi" in una modalità di integrazione culturale tra umani e macchine smart basata sulla consapevolezza che il sistema cognitivo umano resta sempre più "intelligente" di quello artificiale.

L'integrazione della IA con l'intelligenza umana sarà completa quando i medici, dotati di rigorosi strumenti predittivi e di guida nelle scelte cliniche, potranno delegare alle macchine i calcoli e le operazioni sui dati, focalizzandosi



sull'interpretazione dei fenomeni complessi e le conseguenti possibili soluzioni.

I medici continueranno a svolgere un ruolo di guida, supervisione e monitoraggio, utilizzando la propria intelligenza e le capacità tipicamente umane quali astrazione, intuizione, flessibilità ed empatia per esercitare un approccio conservativo e costruttivamente critico utilizzando, allo stesso tempo, tutte le enormi potenzialità rappresentate da un sistema predittivo della IA in ambito prognostico, attraverso la disponibilità dei Big Data e dell'apprendimento automatico.

A titolo di esempio, la Figura 1 riporta l'integrazione tra gli studi con metodologia tradizionale (in blu) e quelli condotti con la tecnologia di IA e apprendimento automatico (in verde).

In campo medico, l'ambito oncologico è sicuramente tra le principali discipline a trarre notevoli vantaggi da un approccio

che integri le attività mediche con strumenti di IA in grado di supportare e accelerare il passaggio verso una medicina di precisione e una personalizzazione delle cure.

Da tempo, infatti, la medicina moderna cerca di fornire ai pazienti terapie specifiche e un piano dei trattamenti che tengano conto delle esigenze individuali delle persone, nel continuo tentativo di conciliare gli aspetti clinici con la qualità della vita e di fornire risposte anche alle esigenze socio-sanitarie e socio-assistenziali durante tutto il percorso di cura.

Le possibilità offerte dalla digitalizzazione delle informazioni e dall'utilizzo dei Big Data come base per l'impiego delle tecnologie di IA/ML consentono di connettere tra loro le diverse fasi di gestione del paziente e, con le opportune forme e trasformazioni dei dati disponibili, di alimentare la ricerca clinica e la pianificazione amministrativa in capo agli Enti sanitari locali, regionali e nazionali (Regioni/ASL/...).

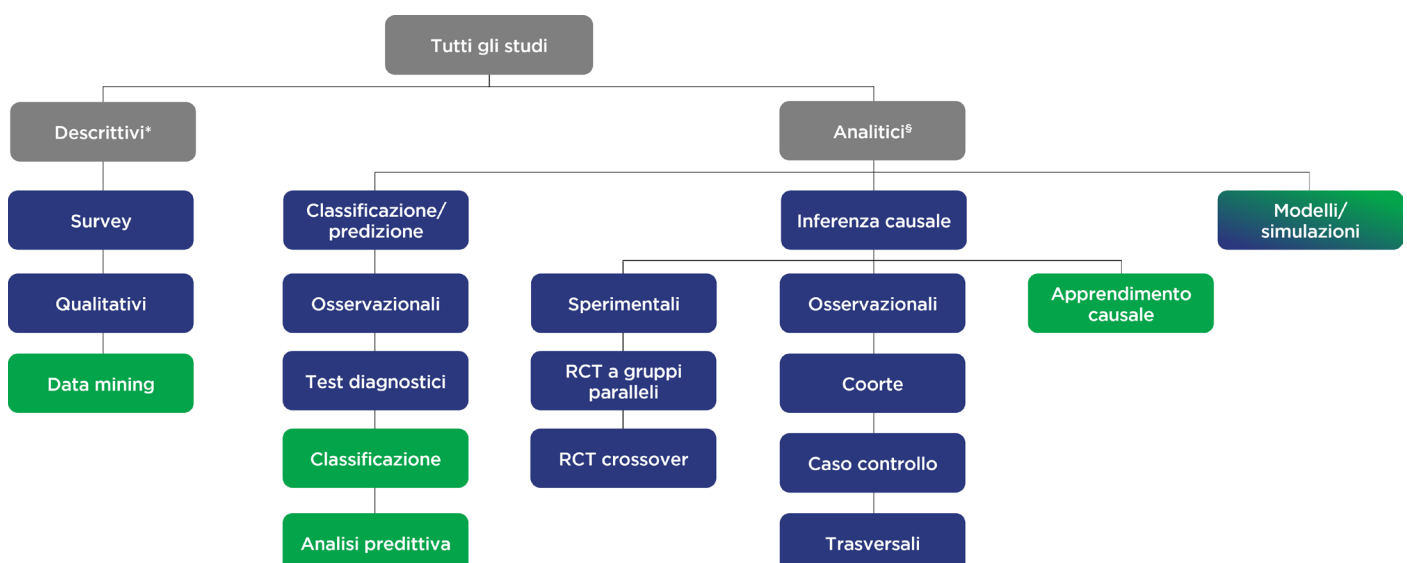


Figura 1: Le metodologie di studio tradizionali e quelle supportate da AI

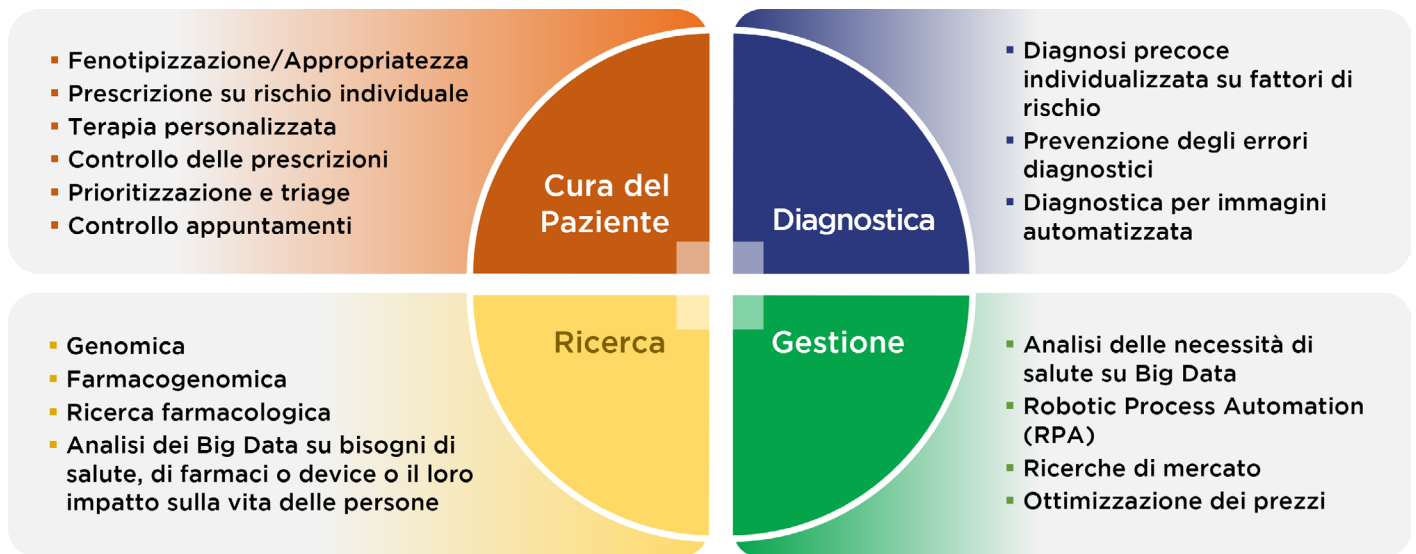


Figura 2: Le metodologie di studio tradizionali e quelle supportate da AI

L'organizzazione e l'elaborazione di grandi quantità di dati e variabili può essere valorizzata dai sistemi di supporto alle decisioni (Decision Support Systems - DSS). Questi ultimi sono strumenti informatici che, sulla base di dati e modelli matematici, restituiscono al cosiddetto "decisore" - dal policy maker al medico fino al manager, a seconda dei contesti - ulteriori informazioni utili a prendere decisioni nello svolgimento delle proprie attività. Queste informazioni possono essere di tipo sintetico (indicatori provenienti dall'elaborazione di più variabili) o esteso, in forma di analytics in cui vengono riportate tutte le logiche e i modelli significativi che sono alla base dei risultati restituiti.

Gli ambiti applicativi sono tra i più vari: da quelli di pianificazione territoriale - in cui le informazioni restituite attraverso l'elaborazione di dati sono di tipo cartografico, permettendo al pianificatore

di aver contezza di specifici fenomeni per aree territoriali e quindi essere supportato nella definizione di nuove policy - a quelli di tipo clinico, in cui il medico potrà visualizzare i dati per specifici indicatori relativi alla salute del paziente e, sulla base di questi, decidere se e come apportare integrazioni a quanto previsto dal piano di cura.

Attraverso i DSS, i dati organizzati ed elaborati attraverso tecnologie di Big Data e Intelligenza Artificiale vengono trasformati e valorizzati in informazioni e conoscenza utili per prendere decisioni.

In conclusione, con riferimento alle possibilità offerte dallo sviluppo tecnologico, osserviamo che la disponibilità di nuovi strumenti a un livello di maturità adeguato abilita nuovi scenari e nuovi paradigmi.

La Figura 3 sintetizza le diverse fasi evolutive e i fattori tecnologici abilitanti che hanno consentito la trasformazione dalla medicina tradizionale alla medicina distribuita per arrivare oggi alla medicina di precisione, in cui la variabilità genetica, ambientale e degli stili di vita del singolo paziente sono elementi da valutare nel trattamento e nella prevenzione delle diverse patologie.

Progetti di ricerca industriale e di sviluppo sperimentale in questo campo incrementano esponenzialmente l'innovazione applicata alla medicina e consentono di progredire più velocemente verso obiettivi di cura e assistenza personalizzati per il paziente, contribuendo a migliorare efficienza ed efficacia del sistema assistenziale e sanitario nel suo insieme.

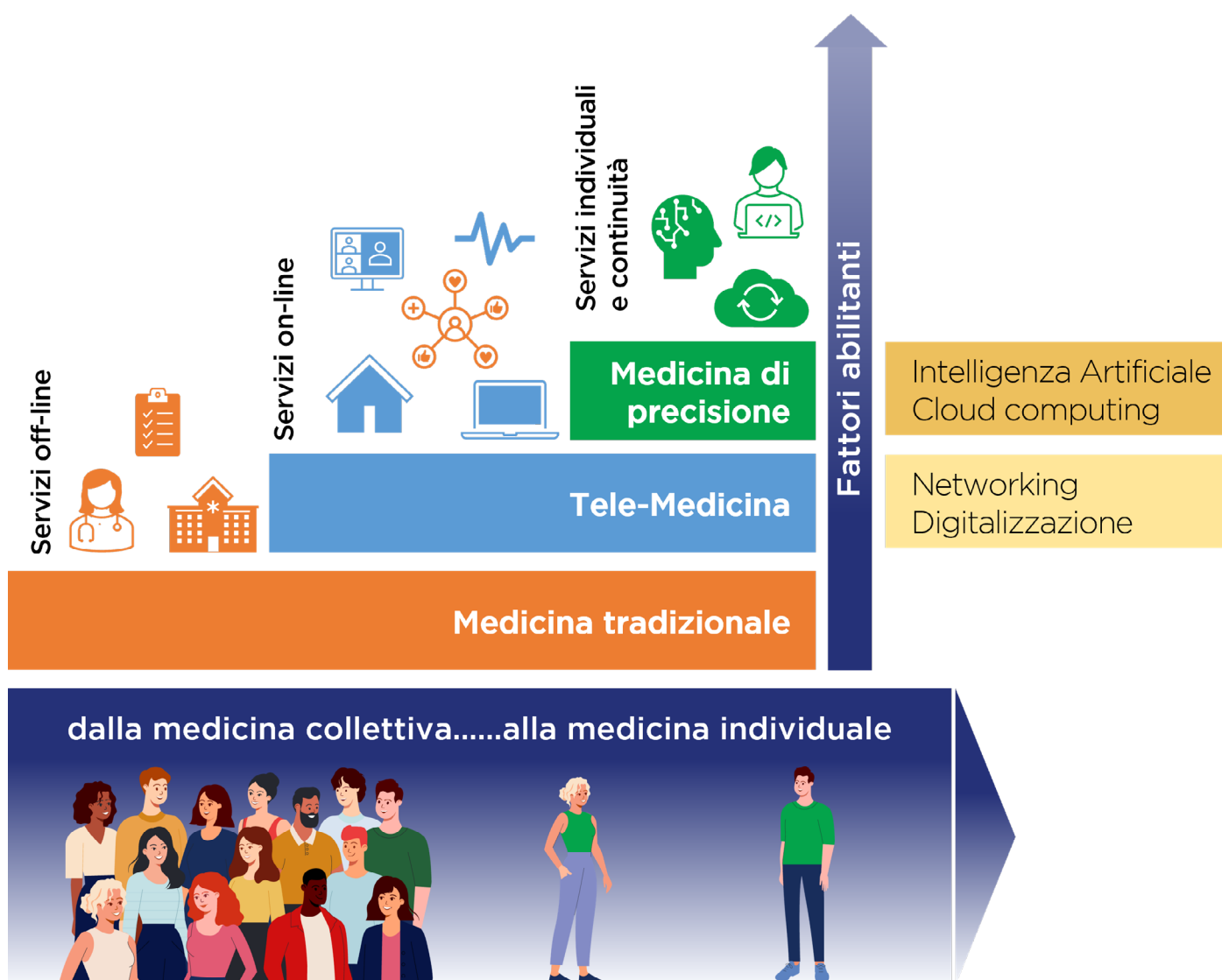


Figura 3: L'iter evolutivo della medicina, da tradizionale a quella di precisione

2. Il progetto Oncologia

Oncologia è un progetto di investimento di ricerca industriale e sviluppo sperimentale, co-finanziato dalla Regione Puglia attraverso il Fondo Europeo di Sviluppo Regionale P.O. Puglia FESR 2014 – 2020, per l'applicazione di tecnologie innovative IT in ambito medico oncologico. Quest'ultimo rappresenta una delle specializzazioni a oggi più avanzate e con una rete di servizi significativa in cui l'approccio "di precisione" può garantire differenze di esito determinanti. Inoltre, l'ambito oncologico necessita di disporre di volumi e qualità di dati importanti, nonché un numero elevato di possibili utilizzatori.

Tali aspetti costituiscono un banco di prova impegnativo per testare concetti e implementazioni che siano valide sul piano tecnico-scientifico e risultino anche semplici da adottare sia dai professionisti che dai pazienti stessi.

Il progetto Oncologia mira all'abilitazione di un modus operandi di precisione, attraverso un utilizzo "sostenibile" dell'intelligenza artificiale e dei dati, nel rispetto delle normative che riguardano queste applicazioni e delle implicazioni etiche, anch'esse vincolanti.

In particolare, il progetto intende porre le basi per la creazione di un framework di riferimento, gestito in cloud, all'interno del quale sarà creato un ecosistema in grado di ospitare diversi livelli di utilizzo delle

tecnologie IA e ML: dagli algoritmi predittivi e di supporto decisionale in ambito clinico e logistico alla costituzione di Identità Digitali Sanitarie (IDS). Queste rappresentano un modello evoluto e digitale del paziente reale, connesso attraverso specifici sensori e dispositivi, in grado di offrire una "vista" in near-real time del paziente al medico e/o alle persone che lo hanno in cura.

Le aree su cui si concentrerà la sperimentazione applicativa di Oncologia sono il mieloma – tumore maligno del sangue – e il tumore al seno, per le quali il progetto effettuerà sia delle implementazioni tecnologiche verticali su casi studio relativi alle patologie citate, sia di tipo "general purpose", in un'ottica di estensione del campo applicativo ad altre patologie o ambiti anche non appartenenti a quello oncologico. In tal senso, il perimetro di sperimentazione tecnologica potrà essere esteso alla verifica dei risultati di addestramento degli algoritmi su dataset a tematica clinica differente, verificando la misura in cui la capacità di questi algoritmi si conservi al mutare della tematica clinica specifica.

Oncologia intende integrare tutti gli aspetti summenzionati – dalla sicurezza dei dati alla necessità di far riferimento a più fonti di dati, fino alla costruzione di sistemi di supporto alle decisioni e predizioni in ambito clinico – attraverso attività di ricerca industriale e sviluppo sperimentale focalizzate nei

seguenti stream tecnologici:

- Studio e definizione del modello di erogazione in cloud:
OncologIA intende approfondire gli aspetti relativi alla sicurezza del dato nell'erogazione dei servizi in cloud, e per ognuno sarà sviluppato un set specifico di funzionalità per garantire riservatezza e rispetto delle normative in materia di Privacy (GDPR). In particolare, per i diversi soggetti (amministrativi / clinici / pazienti / caregiver) utilizzatori finali dei servizi, saranno studiati opportuni meccanismi di profilazione e sicurezza per assicurare l'accesso da diverse piattaforme con garanzia di riservatezza e con la possibilità di gestire informazioni provenienti da dispositivi IoMT (Internet of Medical Things).
- Il modello di Data Lake "oncologico":
all'interno di questo stream, il progetto ha l'ambizione di costruire un sistema di raccolta, armonizzazione e normalizzazione nativa delle informazioni, superando l'organizzazione a silos delle applicazioni della rete oncologica, in modo da ampliare il volume di informazioni disponibili, abilitando l'uso efficace di algoritmi di AI anche con dati provenienti da fonti differenti e, infine, alimentando, con tali dati, le Identità Digitali Sanitarie (di seguito IDS).
- Gli algoritmi predittivi in ambito clinico:
per la parte di individuazione e personalizzazione dei trattamenti

e dei percorsi di cura, OncologIA intende approfondire i sistemi di IA e apprendimento automatico, specializzati nei casi d'uso medici citati in precedenza.

- Le Identità Digitali Sanitarie – IDS (Digital Twin):
per una generalizzazione del processo di simulazione, ma anche a supporto dell'individuazione di terapie potenzialmente più efficaci poiché personalizzate, OncologIA intende sperimentare un modello di "Identità Digitale Sanitaria" del paziente, creando di fatto un gemello digitale (Digital Twin) di quest'ultimo, al fine di fornire un quadro completo e integrato circa il suo stato di salute al personale clinico.
- Gli algoritmi predittivi a supporto della sanità territoriale:
OncologIA si propone di sviluppare un modello che tenga conto, nell'iter di cura, anche dei servizi offerti dalle strutture sanitarie territoriali, per una logistica più efficiente ed efficace, e delle caratteristiche ambientali, anch'esse potenzialmente impattanti sulla salute del paziente.

Di seguito sono approfonditi gli stream progettuali specifici di OncologIA al fine di descrivere gli avanzamenti tecnologici e le innovazioni d'ambito.

3. Le aree di innovazione di OncologIA

Il modello di erogazione in cloud

La trasformazione digitale in ambito sanitario passa anche attraverso la consapevolezza del ruolo del dato. Il volume dei dati digitali è in continua crescita e il loro valore è nell'estrarre informazioni utili e usabili. Il modello di OncologIA si basa sulla centralità del dato clinico e sulla capacità di ricostruire la storia clinica del singolo paziente, oltre che di aggregare dati di diversi pazienti in maniera tale da fornire strumenti per la cura e la ricerca.

Per supportare questa trasformazione basata sul dato - si parla di "data revolution" - sono necessarie adeguate infrastrutture tecnologiche e soluzioni analitiche finalizzate a costruire strumenti per il supporto decisionale, i controlli di appropriatezza, la predisposizione di scenari. Queste infrastrutture sono destinate a crescere in misura esponenziale, con le relative conseguenze anche in termini di sicurezza e tutela del dato sensibile. Un abilitatore fondamentale di trasformazione è il cloud, sia esso pubblico, privato o ibrido.

Le infrastrutture cloud offrono infatti facilità d'uso, flessibilità di configurazione e dimensionamento, ma possono introdurre nuovi rischi per la sicurezza dei dati.

Gli aspetti di sicurezza e privacy sono centrali soprattutto per la gestione del dato sanitario: in tale contesto è necessario che l'infrastruttura tecnologica permetta

di rispettare la normativa vigente senza inficiare la possibilità di implementare soluzioni analitiche con particolare riferimento a quelle basate sull'Intelligenza Artificiale.

Il potenziale innovativo di OncologIA in questo ambito risiede:

- nell'implementazione di un modello di gestione evoluta e certificata delle credenziali e dei profili di accesso che permetta di garantire riservatezza e rispetto delle normative;
- nella raccolta dei consensi privacy, tramite tecniche di Blockchain, per una gestione affidabile, trasparente e tracciata;
- nella creazione di un modello ibrido del Data Lake distribuito in cloud per la raccolta e analisi dei dati;
- nell'utilizzo di tecniche avanzate di crittografia standard e omomorfa, in meccanismi di differential privacy e nell'integrazione;
- nella coesistenza di questi aspetti all'interno di una piattaforma abilitante.

In particolare, la sperimentazione di tecniche di crittografia omomorfa risulta essere un aspetto importante all'interno di OncologIA. Gli algoritmi di Intelligenza Artificiale ricoprono un'importanza fondamentale e crescente, ma, come indicato in precedenza, il tipo di dato trattato in ambito sanitario

è estremamente sensibile. Un limite della crittografia tradizionale è che i dati necessari per le successive elaborazioni statistiche o per l'addestramento di modelli di Intelligenza Artificiale devono essere decifrati prima dell'uso, in contrasto con gli aspetti di privacy. In contrapposizione agli schemi di crittografia tradizionali, metodi di Crittografia Omomorfica permettono di operare direttamente sui dati crittografati.

In ultimo, al fine di elevare ulteriormente il livello di sicurezza e riservatezza delle informazioni trattate, verranno sperimentate specifiche tecniche di differential privacy – “aggiunta di rumore” ai dati - con l'introduzione di un certo grado di entropia nel dataset, tale per cui le informazioni contenute non siano più identificabili o correlabili in caso di eventuali “composition attack”.

In conclusione, attraverso il proprio modello di erogazione in cloud, il progetto OncologiaA ambisce a contribuire all'avanzamento tecnologico sulla sicurezza dei dati nell'ambito clinico sanitario, attraverso un modello integrato di tecniche innovative.

Il Data Lake oncologico

I dati digitali in ambito sanitario continuano a crescere in volumi, varietà e velocità, per cui le esigenze analitiche richiedono un'infrastruttura di data management che sia efficiente, flessibile, scalabile e distribuita. In questo scenario si concretizza la scelta di un Data Lake che rappresenta un repository centralizzato progettato per archiviare, elaborare e proteggere grandi quantità di dati di varia tipologia che possono essere strutturati, semistrutturati e non strutturati.

In tale contesto il progetto OncologiaA si pone come obiettivo la ricerca e lo sviluppo delle soluzioni per la realizzazione dell'infrastruttura elaborativa necessaria alla sua gestione. Il Data Lake rappresenta la base per le fasi di analisi e di intelligenza artificiale.

Si tratta, in termini più tecnici, dell'applicazione del cosiddetto schema on-read in cui lo schema dei dati viene fornito nel momento dell'analisi e non nel momento dell'archiviazione.

La gestione del dato all'interno di un Data Lake prevede:

- Data Ingestion e Storage, ossia la capacità di acquisire dati in real-time o in batch, e la capacità di conservare e accedere a dati strutturati, semi strutturati e non strutturati nel formato originario in cui sono prodotti, in cui è presente un sistema che permetta di definire i ruoli di accesso;
- Data Processing, ossia la capacità di a)



lavorare sui dati in precedenza acquisiti, al fine di renderli pronti per essere analizzati con procedure standard, b) implementare e ingegnerizzare le soluzioni di estrazione di valore dai dati, attraverso processi automatici e periodici, che sono il risultato delle operazioni di analisi;

- Data Analysis, ossia la capacità di creare modelli per l'estrazione sistematica di informazioni dai dati, che può avvenire in tempo reale o attraverso processi eseguiti su base periodica;
- Data Integration, ossia la capacità di agganciare alla piattaforma applicativi che consentano di interrogare ed estrarre dati in formati utilizzabili per scopi specifici.

L'infrastruttura elaborativa deve prevedere gli strumenti che permettano di gestire il dato secondo le modalità descritte, mantenendo centrali tutti gli aspetti di sicurezza e di dominio del dato di tipo sanitario.

OncologIA si pone come obiettivo all'interno del Data Lake di creare il profilo di ogni singolo paziente per tutto il ciclo della malattia, dalla fase preclinica per la definizione della patologia alla progettazione ed esecuzione della fase di terapia, fino alla valutazione dei risultati ottenuti e della sopportazione della terapia.

La tipologia dei dati, pur essendo in partenza suddivisa e differente per formato e contenuti, attraverso meccanismi di interoperabilità e un processo di

standardizzazione si uniforma, utilizzando il paziente come elemento di congiunzione e generando così il patrimonio informativo sanitario dello stesso, secondo una visione "paziente-centrica" del dato.

In OncologIA, l'adozione del modello Common Data Model (CDM) dell'Observational Medical Outcomes Partnership (OMOP) come riferimento per il design del Data Lake, con le opportune verticalizzazioni relative all'ambito oncologico, risulta di estremo valore, aprendo a soluzioni estendibili anche ad altri domini sanitari.

Gli algoritmi predittivi in ambito clinico

L'oncologia è una specializzazione medica che, dato il contesto in cui opera, necessita fortemente di dati storici e near-real time per poter offrire percorsi di cura sempre più efficaci e centrati sul singolo paziente. Lo sviluppo della tecnologia e dell'innovazione ha impattato questo ambito medico in diversi aspetti: quello della ricerca, degli strumenti di cura e diagnosi, della tipologia di dati e informazioni reperibili e analizzabili.

In tale contesto, un ruolo sempre più determinante è giocato dall'Intelligenza Artificiale. Molte e diverse sono state le esperienze e le applicazioni nella letteratura scientifica e nella pratica clinica, soprattutto orientate a progettare strumenti che aiutassero a valutare la compatibilità di e tra i trattamenti, e apportassero valore aggiunto nella parte di imaging, sia in fase di diagnosi

o screening sia di follow up.

Tuttavia, una delle aree per cui è sempre più necessaria una sperimentazione in ambito clinico sanitario è quella relativa agli strumenti predittivi e prognostici all'interno del percorso del paziente, utilizzando dati, diversi per tipologia e provenienti dalle fonti cliniche più eterogenee (RIS-PACS, LIS, EHR etc.). I modelli di machine learning e deep learning hanno il potenziale di fornire tali strumenti attraverso la loro capacità elaborativa.

Il progetto OncologIA si inserisce all'interno di tale filone di ricerca, approfondendo e sperimentando algoritmi che possano attingere a dati provenienti da diverse fonti, intese anche come step all'interno dell'iter di cura: da quelli caratterizzanti la prima richiesta di approfondimento diagnostico e stadiazione, fino a quelli di tutto il percorso di follow-up, rappresentando uno strumento di supporto prognostico nelle decisioni puntuali per il clinico, grazie alla complessiva e olistica valutazione e considerazione dei dati.

In particolare, OncologIA intende sviluppare e sperimentare tali algoritmi nei diversi snodi del percorso terapeutico, sia per finalità di monitoraggio, supporto alle decisioni e predizione, sia per manipolare e trasformare i dati grezzi provenienti da più fonti, come ad esempio referti e cartelle cliniche elettroniche, ovvero documenti contenenti informazioni in forma non strutturata o descrittiva, in dati fruibili, attraverso specifiche tecniche di Data Augmentation,

Data Enrichment e Natural Language Processing (NLP). Il risultato è l'estensione, l'arricchimento e la migliore qualità dei dataset a disposizione. Inoltre, la possibilità di elaborare informazioni, caratterizzate da una vasta quantità di variabili, permette di costruire indicatori sintetici specifici per il paziente, in modo da aumentare l'accuratezza clinica – ossia la qualità della valutazione del medico - sullo stato specifico del singolo paziente, in un'ottica di medicina di precisione.

In sostanza, l'impiego delle tecniche di intelligenza artificiale ha lo scopo di supportare il tumor board in una definizione più accurata dello stato di salute del paziente, in determinati snodi del percorso terapeutico, attraverso la costruzione di ulteriori indicatori sintetici di tipo diagnostico e prognostico, in accordo con i Percorsi Diagnostici Terapeutici Assistenziali (PDTA) di riferimento.

Sempre nell'ottica di fornire strumenti al personale medico, per una medicina che garantisca centralità dell'assistito e che benefici di una più varia tipologia di dati fruibili, ulteriori aspetti considerati sono il patient engagement e la valutazione della sua Quality-of-Life (QoL).

OncologIA pone le basi per lo studio e la prototipazione di strumenti di supporto al medico per automatizzare la comunicazione con il paziente, al fine di svolgere il follow-up domiciliare senza gravare sulle attività del personale sanitario e aumentando la "vicinanza" all'assistito

in seguito al contatto con la struttura sanitaria. L'impiego dell'Intelligenza Artificiale, in questo caso, sarà il fattore abilitante per estendere la raccolta di informazioni direttamente dal paziente, automatizzando la somministrazione di Patient Reported Experience Measures (PROMs) e Patient Reported Experience Measures (PREMs) attraverso il linguaggio naturale e traducendo il parlato acquisito in informazioni strutturate. Su tali dati potranno essere realizzati algoritmi e strumenti predittivi sulla QoL e la relativa visione paziente-centrica. L'impiego di tecniche di NLP è alla base dell'utilizzo del canale vocale, con il vantaggio di non richiedere device dedicati o altri strumenti che, soprattutto per la popolazione anziana, possono essere di difficile fruizione. Collezionare in modo strutturato i risultati delle survey offrirà al clinico una visione d'insieme sull'andamento della somministrazione dei questionari e sul trend di partecipazione, segnalando eventuali criticità per i singoli pazienti.

In conclusione, OncologIA ambisce a innovare lo scenario degli algoritmi predittivi in ambito clinico: a) sperimentando la potenziale integrazione tra fonti di dati eterogenee, b) integrando le informazioni sui potenziali *needs* e feedback del paziente, a supporto del personale sanitario e c) realizzando gli strumenti di analisi, monitoraggio e predizione attraverso una modalità circolare e di continuo confronto con gli attori coinvolti, secondo un approccio di co-design per l'innovazione

tecnologia in ambito sanitario.

Le Identità Digitali Sanitarie (IDS)

Un'Identità Digitale Sanitaria (IDS) è una rappresentazione virtuale che utilizza la tecnologia per monitorare e gestire l'intero ciclo di vita dei pazienti, dai dati sui loro sintomi e condizioni di salute, alle diagnosi e ai trattamenti.

L'IDS si inquadra all'interno della famiglia dei gemelli digitali (Digital Twin) che permettono di replicare nel virtuale le caratteristiche, potenziali ed effettive, di persone, luoghi, processi, oggetti, infrastrutture, sistemi e dispositivi. In generale, un Digital Twin può essere progettato utilizzando diverse tecnologie, come l'intelligenza artificiale, la realtà aumentata e la realtà virtuale.

In campo medico, il suo potenziale è legato alla possibilità di creare modelli estremamente dettagliati e precisi dei pazienti - ma anche delle patologie stesse - che possono aiutare a migliorare la qualità dell'assistenza sanitaria e a ridurre i costi associati alla cura. Ulteriori potenzialità offerte sono: simulazione e test di diverse strategie di trattamento, identificazione precoce dei problemi e previsione delle necessità dei pazienti. Inoltre, può essere utilizzato per la formazione del personale sanitario, migliorando la capacità di cura.

I casi d'uso implementabili attraverso la creazione di un gemello digitale in ambito oncologico sono vari, integrabili, e tutti rispondenti ad esigenze dei pazienti e del

personale sanitario. Ad esempio, nell'ambito del monitoraggio dei dati clinici a supporto alle decisioni, le IDS possono consentire al personale medico di rilevare pattern di evoluzione patologica e inviare alert al case manager in caso di anomalie, per poi fornire supporto nell'applicazione di eventuali adattamenti alla terapia in tempo reale. In termini diagnostici, le IDS, unite alla capacità di calcolo degli algoritmi di IA, possono restituire analisi di dati provenienti da diverse fonti, come i dati clinici del paziente, le informazioni sulle terapie e le analisi dei tessuti tumorali, al fine di identificare le caratteristiche specifiche del tumore e quindi potenziare la personalizzazione della terapia.

In sostanza, il potenziale dello sviluppo dell'idea del gemello digitale in ambito sanitario, con la creazione delle Identità Digitali Sanitarie, risulta essere di grande impatto attraverso la restituzione strutturata e integrata di informazioni provenienti da più fonti, al fine di fornire al personale medico un'immagine chiara e sinottica del paziente e della sua condizione, senza perdere il potenziale dell'analisi di grandi moli di dati per la valutazione dei fenomeni generali.

Oncologia intende approfondire e sperimentare il potenziale rappresentato da tali modelli digitali, sia per le finalità di monitoraggio del paziente che di supporto alle decisioni del medico. In particolare, attraverso le attività di ricerca industriale e sviluppo sperimentale, si intende creare le basi per la costruzione di una soluzione complessa che possa sviluppare i seguenti casi d'uso:

- **Monitoraggio:** definire una soluzione che raccolga sia informazioni sulle terapie prescritte - come la durata del trattamento e la posologia dei farmaci - che, attraverso l'utilizzo di sensori e/o dispositivi indossabili, quelle relative al benessere del paziente, provenienti dal monitoraggio dell'attività fisica, del sonno e di altre attività da identificare con il personale clinico. Attraverso tali dati e le funzionalità che verranno implementate, l'applicazione fornirà al paziente la possibilità di segnalare eventuali sintomi o effetti collaterali dei farmaci, al fine di monitorare la loro gravità e supportare il medico a valutare l'opportunità di modificare la terapia. Inoltre, l'applicazione integrerà i dati degli esami di laboratorio e delle visite di follow-up, al fine di monitorare l'evoluzione della malattia e l'efficacia della terapia nel tempo. In questo modo, saranno garantite una maggiore precisione nella gestione della terapia post-intervento e una maggiore tempestività nella gestione dei problemi di salute del paziente.
- **Supporto decisionale sulle scelte terapeutiche:** definire una soluzione che, attraverso l'accesso ai dati clinici del paziente e l'elaborazione di tali dati attraverso algoritmi di intelligenza artificiale, fornirà informazioni sulle opzioni terapeutiche disponibili, raccomandazioni personalizzate e basate anche sui dati individuali, per aiutare i medici nelle scelte terapeutiche. Avrà un'interfaccia utente intuitiva, e sarà aggiornata regolarmente

per garantire la sicurezza e la privacy dei dati del paziente oltre a suggerimenti allo stato dell'arte in tema clinico.

In conclusione, l'Identità Digitale Sanitaria (IDS) che si intende sviluppare con Oncologia rappresenta un elemento di profonda innovazione del processo assistenziale. Grazie ad essa, il concetto di "paziente al centro" trova una sua concretizzazione inedita. L'IDS può essere immaginato come un "angelo custode digitale" che va ad affiancarsi al sistema organizzativo e specialistico che ha in cura il paziente, potenziandolo. Esso mantiene costantemente "in primo piano" l'unicità rappresentata dal paziente. Al modello addestrato su ampi volumi di dati viene anche affidato il compito di effettuare una verifica continua della rispondenza del modello rispetto ai casi clinici sui quali è applicato e arricchito.



Gli algoritmi predittivi per la sanità territoriale

Oltre agli impatti sul merito della cura, la tecnologia digitale in ambito sanitario può contribuire in maniera decisiva anche nel metodo, e quindi nell'aspetto organizzativo della cura. In sanità, gli aspetti di processo impattano sugli outcome terapeutici in modo altrettanto importante rispetto a quelli della cura.

Il Progetto Oncologia intende integrare gli strumenti di monitoraggio e predizione della cura con quelli di pianificazione del processo e del piano terapeutico, in un'ottica di ottimizzazione di risorse, di logistica e spostamenti del paziente, sulla base di servizi e specializzazioni territoriali. Tale opportunità di innovazione è in totale continuità con il già esistente modello organizzativo delle reti oncologiche in hub&spoke, e lo rafforza attraverso la possibilità di avere in maniera istantanea informazioni elaborate attraverso algoritmi di AI.

Come già evidenziato, le tecniche di intelligenza artificiale sono in grado di elaborare e sintetizzare in un unico algoritmo le diverse tipologie di dato, sia quelle che riguardano il dato territoriale, sia quelle relative al dato clinico, per fornire un indicatore finale che possa suggerire il miglior percorso terapeutico sul territorio regionale.

In particolare, Oncologia intende sperimentare un modello che permetta l'ottimizzazione del percorso di cura del

paziente, considerando gli aspetti logistici - quindi il suo domicilio e la posizione geografica delle strutture sanitarie - gli aspetti clinici - quindi lo stato di salute, le patologie da cui è affetto e l'iter terapeutico da seguire - e gli aspetti prestazionali, ovvero i servizi socioassistenziali forniti dalle strutture stesse. L'ottimizzazione del percorso di cura ha lo scopo di alleggerire l'intero iter terapeutico del paziente, indirizzandolo agli ospedali e alle strutture che risultano più vicini al proprio domicilio e che forniscono le giuste prestazioni, necessarie per completare il percorso di cura.

L'integrazione degli algoritmi di intelligenza artificiale all'interno del Data Lake Oncologico con i dati territoriali, i dati clinico-sanitari relativi allo specifico paziente e i dati prestazionali degli ospedali e delle altre realtà del sistema sanitario, alimenterà un sistema di raccomandazione a supporto delle strutture sanitarie e socioassistenziali. In particolare, attraverso l'elaborazione dei dati territoriali e quelli clinici del paziente, il case manager oncologico sarà supportato nell'individuazione delle strutture sanitarie più adatte alle necessità contestuali e specifiche del singolo paziente. A fini esplicativi, di seguito si riportano varie tipologie di dati che potranno essere considerati:

- Strutture sanitarie sul territorio e loro capacità di accoglienza
- Strutture socioassistenziali
- Medici specialisti sul territorio

- Prestazioni erogate per tipologia di specializzazione
- Tecniche utilizzate dalla singola struttura per la gestione delle singole prestazioni
- Condizioni ambientali dell'area territoriale di riferimento
- Percorso clinico del paziente individuato
- Dati clinici del paziente (anamnesi, patologie, interventi, terapie)
- Prestazioni e cure di cui necessita il paziente.

L'impatto di una tale soluzione può essere enorme: i pazienti oncologici sono sottoposti a terapie periodiche e di forte intensità da effettuare nelle specifiche strutture. Gli spostamenti e l'organizzazione della vita del paziente sono spesso a carico dei familiari, dei caregiver e del personale assistenziale. L'ottimizzazione degli aspetti logistici, aumentando la vicinanza geografica tra paziente e strutture, permetterebbe una gestione più semplice dei trasferimenti giornalieri dei pazienti e ne migliorerebbe complessivamente la qualità della vita. Inoltre, una maggiore attenzione nel gestire l'accesso ai servizi ospedalieri e ambulatoriali e un più efficiente bilanciamento del carico di pazienti su tutto il territorio regionale rafforzerebbe l'erogazione delle prestazioni clinico-sanitarie, ottimizzando in generale tutte le risorse messe a disposizione dal servizio sanitario. Sotto l'aspetto del percorso di cura, il miglioramento degli aspetti logistici e clinico-sanitari all'interno del percorso del paziente oncologico

si traduce in una maggior aderenza alle cure con una conseguente miglior applicabilità del protocollo terapeutico. Infine, l'ottimizzazione della distanza e della quantità di spostamenti tra domicilio e struttura sanitaria permetterebbe una drastica diminuzione delle emissioni di CO₂ nell'atmosfera, con conseguenti vantaggi energetici e ambientali che mirano a migliorare l'intera sostenibilità regionale.

In sintesi, il progetto Oncologia fornirà alle strutture ospedaliere e di ricerca clinica strumenti innovativi che, integrati con quello di monitoraggio e predizione dello stato di salute del paziente, permetteranno al personale clinico di poter contare su

un sistema di supporto alle decisioni logistico/organizzative nel piano di cura. Tale possibilità non solo rafforzerà l'organizzazione territoriale in hub&spoke, ma porterà il modello a un'evoluzione delle reti di cura, in cui tutte le strutture sanitarie presenti su un territorio sono valorizzate per le proprie caratteristiche specialistiche, geografiche, etc., al fine di offrire un approccio sempre più integrato nella cura del paziente e un'erogazione dei servizi sanitari sempre più efficiente ed efficace.



4. Conclusioni

La rapida e progressiva digitalizzazione dei diversi ambiti della società sta portando alla produzione e condivisione di dati e informazioni che rivoluzionano il nostro approccio verso la conoscenza. Tale contesto ci offre la possibilità di capitalizzare questo patrimonio di informazioni per ampliare le modalità di ricerca di soluzioni per migliorare la qualità della vita delle persone.

Attraverso la disponibilità dei dati e la capacità delle macchine di elaborarli, è possibile investigare i fenomeni in maniera multidimensionale, ossia finalizzata a indagare un fenomeno nella sua complessità multidisciplinare e, quindi, nella relazione con un'eterogeneità di variabili che lo caratterizzano, secondo un approccio olistico alla costruzione della conoscenza.

Il mondo della Sanità, e in particolare quello della specializzazione oncologica, è sempre più caratterizzato dalla multidisciplinarietà dell'equipe medica. In ambito oncologico, infatti si parla di tumor board (o cancer board), in cui le squadre mediche sono multi-specialistiche al fine di permettere un confronto ampio, una valutazione di tutti gli aspetti clinici caratterizzanti i pazienti e individuare le diverse opzioni terapeutiche percorribili. Un board in cui il paziente, nelle sue molteplici specificità, è realmente al centro.

Come evidenziato in precedenza, l'utilizzo delle tecnologie digitali ha un potenziale

enorme nel supportare i gruppi medici e di ricerca nelle rispettive analisi e decisioni. Durante il periodo pandemico, questa opportunità di innovazione è diventata una necessità che di fatto ha aumentato l'attenzione verso le soluzioni di medicina digitale.

Almaviva Digitaltec e Almax si sono inserite in questo contesto, con le proprie competenze in ambito IT e nella gestione di progetti complessi, con l'ambizione di creare uno strumento che recepisca le necessità di tipo medico - relative alla multidisciplinarietà dei team (e quindi della conoscenza) e la centralità del paziente - e le supporti con l'uso e l'elaborazione di ampi volumi di dati attraverso la tecnologia, con particolare attenzione alla sicurezza e alla privacy.

OncologIA nasce da questa ambizione. Partendo dal progetto di ricerca industriale e sviluppo sperimentale, le due società del Gruppo Almaviva intendono costruire una soluzione in campo medico-oncologico - e non solo - che integri le seguenti caratteristiche:

- a) essere in grado di accogliere e strutturare grandi quantità di dati sanitari provenienti dalle fonti più eterogenee (ospedali, centri di ricerca, strutture sanitarie private, sociosanitarie, etc.) e che siano eterogenei nella forma in cui si presentano (testo, immagini, dataset, etc.), attraverso la costruzione

- di un Data Lake specializzato e orientato alla ricerca clinica
- b) accesso ed elaborazione di tali dati compliant con le norme vigenti in termini di sicurezza e privacy, attraverso l'utilizzo di tecniche di crittografia omomorfa e tecnologia blockchain
 - c) essere in grado di elaborare tali moli e tipologie di dati (da quelli clinici a quelli territoriali) per fini analitici e predittivi sia nella cura del paziente che nell'organizzazione della cura sul territorio, attraverso algoritmi di AI/ML, in modo da fornire letture multidimensionali e integrate, nonché sistemi di supporto alle decisioni
 - d) garantire una centralità sullo stato di salute del paziente attraverso la costruzione di un suo gemello digitale (Identità Digitale Sanitaria) che ne monitori puntualmente le condizioni e in cui i dati si integrino con quelli ad ampio spettro provenienti da altre fonti, secondo una visione di medicina di precisione unita a studi di tipo statistico.

Il percorso clinico del paziente, e quindi le informazioni e dati che lo contraddistinguono, risulta essere l'asset fondamentale da valorizzare, in modo da permettere al personale medico di prendere le corrette decisioni, per contenuto e per tempistica. Le tecnologie di intelligenza artificiale sono un alleato e un'innovazione importante per il mondo sanitario. Il motivo è legato sia alla loro capacità di analizzare, elaborare, valutare grandi volumi di dati -

diversi per tipologia, forma e provenienza, dando una sintesi personalizzata, preventiva e predittiva - sia come strumento di valorizzazione a livello informativo e di contenuto di quelle tipologie di dato che non nascono espressamente numeriche o strutturate.

In questo scenario, OncologIA contribuisce a dare risposte alle necessità di automazione di alcuni processi sanitari in ambito oncologico e di gestione del paziente, che attualmente non esistono o che sono realizzate tramite gestione umana. Le sperimentazioni tecnologiche e applicative di OncologIA permetteranno di ottimizzare l'intera gestione dell'assistito e della cura, finalizzata a una riduzione dei tempi e una maggiore precisione dei risultati, andando ad integrare più informazioni, quindi aprendo a migliori servizi per i pazienti.

In conclusione, OncologIA si inquadra nell'ampio trend evolutivo che vede sempre più importante l'impiego di tecnologie ML/AI in ambito clinico. Attraverso le attività di ricerca industriale e sviluppo sperimentale di OncologIA, si intende apportare innovazione in termini di:

- nuove opportunità di superamento della logica di costituzione in-house delle infrastrutture necessarie all'erogazione di questa tipologia di servizi, andando verso un nuovo modello di erogazione in cloud e distribuito, che assicuri aderenza alle normative di sicurezza e privacy

- creazione di nuovi modelli di cooperazione dei dati, in grado di aggregare informazioni provenienti da fonti eterogenee (diagnostiche per immagini, terapie farmacologiche, risposte e dati forniti dai pazienti, laboratori di analisi, dati territoriali, etc.) in modo sicuro, garantendone la coerenza rispetto all'elaborazione richiesta pur nel rispetto di riservatezza e proprietà del dato
- modellazione e industrializzazione dell'approccio adottato in Oncologia, per la realizzazione di un framework adattabile a molteplici domini clinici, creando i presupposti per avvicinare le soluzioni sviluppate a un numero elevato di contesti, con l'obiettivo di facilitare sempre più persone e più professionisti a beneficiare di questa importante forma di innovazione del processo assistenziale.

Tali elementi rappresentano una sfida importante, il cui successo consentirà di arricchire i modelli assistenziali e tecnologici delle aziende sanitarie per quanto attiene ai processi clinici legati alla diagnostica, alla gestione dei percorsi di cura e all'organizzazione dei servizi territoriali.



Gli attori di Oncologia

Il progetto Oncologia vede la collaborazione tra una serie di realtà pubbliche e private nella sua implementazione. Le due organizzazioni principali del progetto sono Almaviva Digitaltec S.r.l (azienda proponente) e Almawave S.p.A. (azienda aderente), entrambe appartenente al Gruppo Almaviva. Almaviva Digitaltec (ADT) nasce a fine 2017 per sviluppare attività basate su edge technologies per i mercati della Pubblica Amministrazione Centrale, Locale e industry. L'azienda rappresenta l'anima digital di tutto il Gruppo Almaviva e si caratterizza per le sue competenze nelle aree mobile & portals, IoT, GIS, API economy, microservices, containers, big data & analytics.

ADT è molto attiva, inoltre, nella costruzione di una rete nazionale e locale in ottica Open Innovation: con docenti dell'Università di Napoli "Federico II" e Almawave ha costituito lo spin-off accademico "Data Jam", specializzato in servizi di Machine Learning e Intelligenza Artificiale e relativo ambito di ricerca e innovazione. È presente sul territorio nella ricerca di talenti e soluzioni innovative partecipando all'organizzazione di Hackathon con la Pubblica Amministrazione, Academy IT e progetti di ricerca nazionali e internazionali.

L'azienda è in forte crescita e dalla sede di Napoli si è passati all'apertura delle sedi di Bari, Cagliari, Venezia e Palermo. Da fine 2017 - anno di fondazione - a inizio 2023 ha raggiunto 358 dipendenti e ha un forecast

assunzioni per l'anno 2023 per un totale di 500 dipendenti con l'apertura, nel III Q 2023 delle sedi di Roma e Milano.

Almawave è l'azienda leader nell'intelligenza artificiale e nell'analisi del linguaggio naturale. Propone tecnologia proprietaria all'avanguardia e servizi applicati per concretizzare il potenziale dell'AI nell'evoluzione digitale di aziende e pubbliche amministrazioni. Ha una presenza internazionale in ben 29 Paesi. Conta 6 laboratori tecnologici e oltre 400 professionisti con forti competenze su tecnologie abilitanti e principali framework - Big Data, Data Science, Machine Learning, Architetture AI e Integration - oltre a una profonda conoscenza dei processi di business. La mission di Almawave è quella di rendere concreta la trasformazione digitale nella vita di tutti i giorni, attraverso un modello di esperienza naturale nell'interazione uomo-macchina, grazie all'uso evoluto dell'intelligenza artificiale. Asset tecnologici proprietari consentono di interpretare testo e voce in modalità multilingua, interagire in chiave multicanale, analizzare il dato e l'informazione in ottica di valorizzazione della conoscenza e di automazione. Almawave dispone di un'approfondita conoscenza tecnologica di data science, di processi di business e dei contesti operativi e ha realizzato Iride®, un'avanzata piattaforma basata su artificial intelligence che sfrutta l'omnicanalità per migliorare la customer experience, secondo una strategia citizen centric, abilitando

al contempo l'analisi cognitiva di tutti i contenuti e l'interazione in linguaggio naturale con gli utenti in oltre 40 lingue (voce e testo). Al perimetro di Gruppo si sono aggiunte due nuove società: The Data Appeal Company, operativa con un prodotto AI-based nel mondo della location intelligence per i settori Turismo, Fintech e Retail, e SisTer - Sistemi Territoriali, che offre soluzioni di decision support system, Location Intelligence e competenze Big Data nei settori Utilities e Government.

Le soluzioni Almawave fanno leva sui prodotti e sugli sviluppi realizzati nei propri laboratori di innovazione tecnologica:

- Digital Architecture for AI & Cognitive - Progettazione di architetture Enterprise che integrano una visione avanzata di tecnologie di mercato con gli ecosistemi cliente, i modelli cloud-oriented e le tecnologie proprietarie Almawave
- Data-driven Transformation & Augmented Analytics - Progettazione di soluzioni per l'adozione di nuovi modelli di gestione delle informazioni che, sfruttando le più evolute tecniche di modellazione della conoscenza, di AI e le metodologie di tipo statistico attuariale, consentono di integrare dati strutturati e non strutturati
- Open Data & Semantic Web - Applicazione dell'AI nella modellazione della conoscenza e delle informazioni con tecniche di rappresentazione simboliche o di apprendimento e progettazione di modelli di dati, anche virtuali, a supporto della creazione di

nuovi paradigmi di analisi in ambito Open Data, linked data e semantic web.

- Social Media & Web Data Science - Soluzioni di data science su fonti web e social per discovery e analisi di temi, contenuti e dinamiche di interesse, monitoraggio della digital reputation, grazie alle potenzialità delle soluzioni cognitive proprietarie per massimizzare il valore delle informazioni
- Conversational AI & Engagement - Supporto al cliente nell'utilizzo di assistenti virtuali multicanale e multimodali, attraverso la declinazione della piattaforma conversazionale proprietaria per semplificare l'ingaggio e l'accesso in linguaggio naturale a dati e processi per la creazione della migliore customer experience
- Cognitive Solutions & Smart Process Automation - Declinazione delle tecnologie cognitive, proprietarie per il Natural Language Processing, di mercato per la Computer Vision, per l'ottimizzazione e l'automazione dei processi aziendali, il miglioramento della customer experience e l'analisi della voice of the customer.

L'azienda collabora inoltre con università e partner scientifici (Università di Trento, Fondazione Bruno Kessler, Fondazione Ugo Bordoni, Università La Sapienza, Politecnico di Milano, etc.) e tecnologici leader di mercato, sulle migliori tecnologie, metodologie e best practice per fornire nuovi e migliori servizi ai cittadini e alle imprese.



Almaviva Digitaltec e Almawave
sono aziende del Gruppo Almaviva

www.almaviva.it