

Table of contents

Introduzione	1
Qualità dei dati: di che si tratta?	3
Interoperabilità semantica ed ontologie	4
Perché raccogliere dati con modelli condivisi è fondamentale?	5
Modelli condivisi	7
Health Level Seven International (HL7)	8
DICOM	10
International Classification of Diseases (ICD)	11
SNOMED	12
LOINC	13
I Data model	14
Maturity Model: il ruolo di HIMSS	17
Stato dell'arte	18
L'adozione degli standard in Italia	19
Conclusioni	23
Bibliografia	25
Autori	30

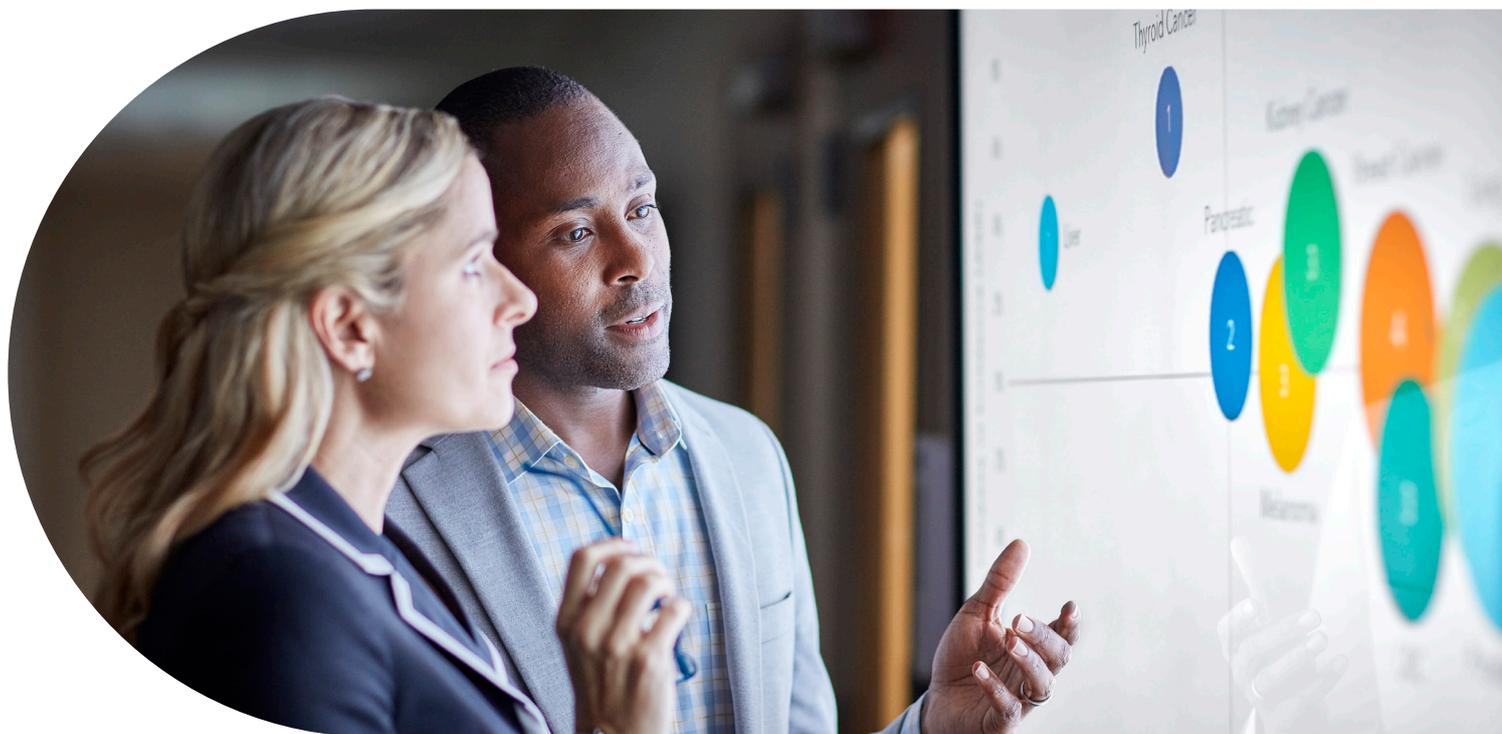
Introduzione

Nell'era della digitalizzazione e dei Big Data, la sanità moderna non è più un insieme disconnesso di servizi, bensì un ecosistema complesso in cui il flusso di dati tra diversi attori gioca un ruolo chiave. La gestione e la strutturazione dei dati stanno assumendo un ruolo sempre più centrale in tutti i campi della ricerca e della pratica clinica; in particolare, nell'ambito della ricerca medica e sanitaria, la capacità di raccogliere, organizzare e analizzare grandi volumi di dati è significativa per accelerare le scoperte, migliorare le terapie e ottimizzare i percorsi di cura. IQVIA, leader mondiale nell'uso dei dati per l'evoluzione dell'utilizzo delle informazioni e per lo sviluppo della sanità, enfatizza da sempre l'importanza di poter "contare sui dati"¹.

In questo contesto, "contare" assume un doppio significato: affidarsi alla qualità e all'affidabilità dei dati e, allo stesso tempo, utilizzare metodi statistici e analitici per estrarre il massimo valore da questi ultimi. La capacità di trasformare grandi volumi di dati non strutturati in informazioni azionabili infatti è fondamentale per guidare la ricerca e la pratica clinica verso nuovi orizzonti.

Come sottolineato anche dalla Scuola Normale Superiore di Pisa², la gestione dei dati della ricerca implica una serie di attività che vanno dalla pianificazione alla conservazione e alla condivisione degli stessi. Queste pratiche non sono solo essenziali per garantire l'integrità e la riproducibilità degli studi scientifici, ma sono anche obbligatorie per rispettare le normative nazionali e internazionali, come il Regolamento Europeo sulla Protezione dei Dati (GDPR). Garantire che questi dati siano facilmente accessibili e utilizzabili in maniera sicura ed efficace rappresenta una sfida significativa.

A supporto della condivisione e utilizzabilità dei dati, l'interoperabilità in campo sanitario — ovvero la capacità dei diversi sistemi elettronici di accedere, utilizzare, integrare e scambiare dati in modo efficace, sicuro e standardizzato — è fondamentale per rendere le comunicazioni delle informazioni cliniche fluide e sicure, migliorando così la qualità dell'assistenza e contribuendo alla creazione di una sanità digitale più efficiente e collaborativa.



¹<https://informatore.it/editoriali/lavoro/perche-possiamo-contare-dati-iqvia/>

²<https://www.sns.it/it/guida/la-gestione-dei-dati-della-ricerca>

Tra i principali vantaggi derivanti dall'interoperabilità dei dati sanitari si annoverano:

- **La riduzione degli errori medici:** l'accesso a una storia clinica completa del paziente può prevenire diagnosi errate e trattamenti inappropriati;
- **Il miglioramento del coordinamento delle cure:** possibilità per gli operatori sanitari di avere una visione completa e aggiornata delle condizioni del paziente per una migliore collaborazione;
- **L'efficienza operativa:** riduzione della duplicazione dei dati e delle inefficienze legate alla gestione di informazioni frammentate.

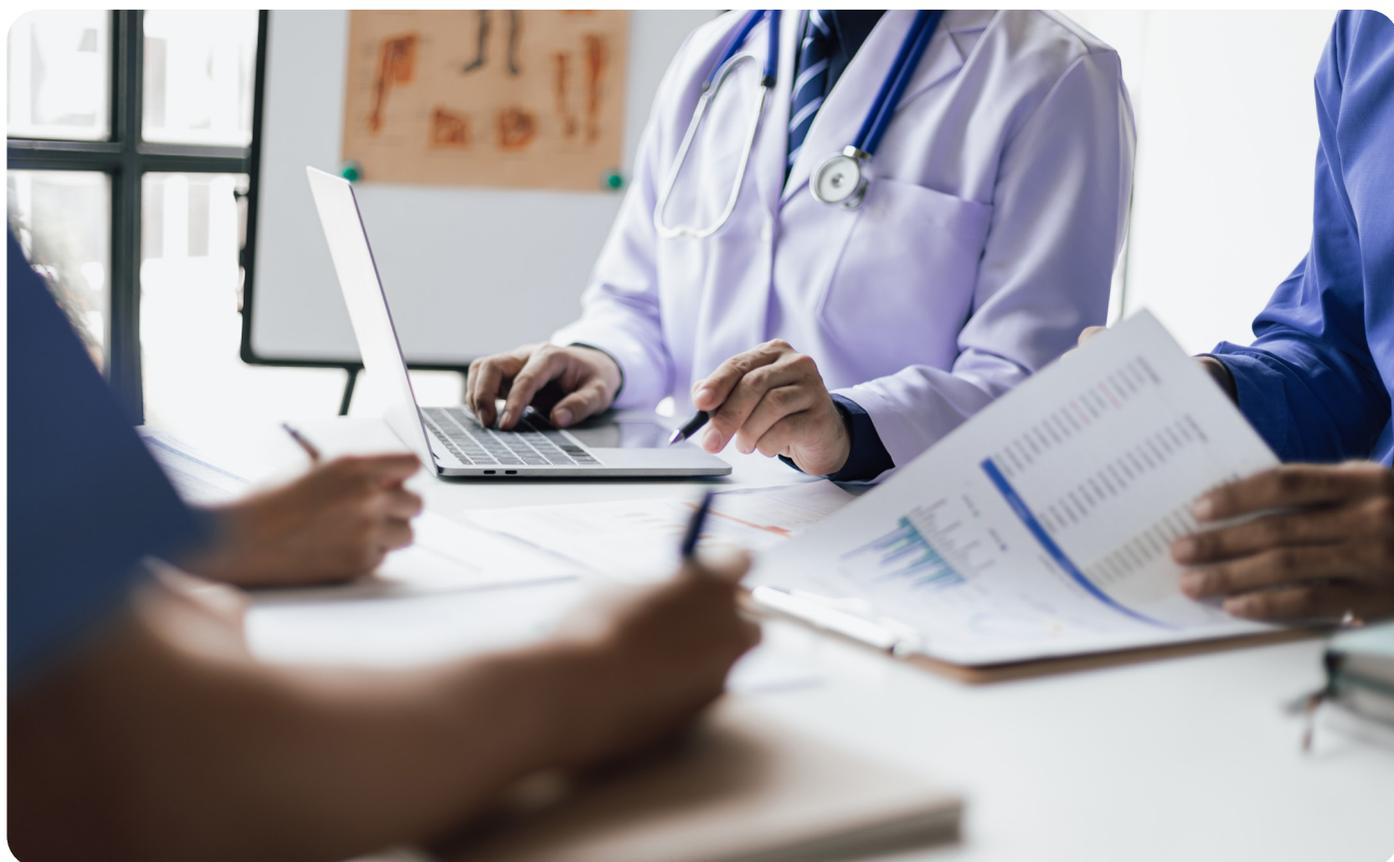
Nonostante i benefici, si riscontrano ancora molte sfide da affrontare, come la frammentazione dei sistemi e la mancanza di standard uniformi, ovvero di *Common Data Model* (CDM).

A tal proposito, è opportuno affermare che un modello di dati ben strutturato è la chiave per una gestione efficace delle informazioni, in quanto fornisce un *framework* logico che descrive la relazione tra differenti

tipi di dati e facilita l'interoperabilità tra sistemi e piattaforme diverse.

I *Common Data Model* sono fondamentali nel settore sanitario per diversi motivi:

- **Standardizzazione dei dati:** il CDM definisce una struttura comune per i dati, consentendo agli operatori sanitari di rappresentare informazioni in modo uniforme. Questo facilita la condivisione e l'analisi dei dati tra diverse organizzazioni;
- **Interoperabilità:** grazie al CDM, i sistemi sanitari possono comunicare tra loro senza problemi. Ad esempio, se un paziente cambia ospedale, il nuovo ospedale può facilmente accedere ai dati storici del paziente;
- **Analisi e ricerca:** i dati standardizzati consentono analisi più approfondite. I ricercatori possono esaminare grandi dataset per identificare tendenze, effetti collaterali dei farmaci e altre informazioni utili;
- **Evidenza basata su dati:** il CDM aiuta a creare evidenze basate su dati reali, migliorando la qualità delle decisioni cliniche e delle politiche sanitarie.



Sostanzialmente, i *Common Data Model* rappresentano uno strumento per migliorare la gestione dei dati sanitari e promuovere la salute pubblica.

Per raggiungere l'obiettivo dell'adozione di standard e modelli di dati, è essenziale utilizzare un approccio multidisciplinare che coinvolga informatici, statistici, clinici e ricercatori. La collaborazione tra queste figure professionali garantisce che i modelli di dati siano non solo tecnicamente validi, ma anche clinicamente rilevanti e in grado di rispondere alle domande di ricerca più impegnative.

Altro tema rilevante è la qualità del dato. Solo dati di elevata qualità possono costituire il presupposto per scoprire nuovi farmaci, per comprendere l'origine delle malattie, per riconoscere quale trattamento possa dirsi migliore; la qualità del dato altresì è importante per gestire i pazienti, trattare le loro malattie e per mettere in pratica delle strategie di prevenzione.

Ma la qualità del dato è anche presupposto necessario per l'attività regolatoria che deve essere finalizzata a costruire un clima di fiducia tra i vari *stakeholder*, nel contesto dell'ambiente sanitario e del rispetto dei pazienti e di tutti coloro che necessitano dei servizi sanitari.

La qualità del dato rappresenta poi un elemento imprescindibile nell'ambito delle best practice, affinché diverse pratiche siano confrontabili tra loro e per la creazione di un ecosistema dove ci sia condivisione e confronto tra i risultati della ricerca e della pratica clinica.

Da ultimo le caratteristiche dei dati hanno rilevanza per i modelli di *governance*: possiamo immaginare la scambiabilità dei dati, l'accessibilità, l'interoperabilità solo se immaginiamo che questi dati siano affidabili e quindi che rispondano a requisiti di qualità.

Qualità dei dati: di che si tratta?

Ma qual è il concetto di qualità a cui facciamo riferimento? Quando all'interno del ciclo di vita di un dato dovrebbe essere implementata la sua qualità? Che caratteristiche dovrebbero avere i processi affinché possano garantire la qualità del dato? Chi dovrebbe creare questi quadri di riferimento sulla qualità del dato per la sua implementazione? In letteratura troviamo più di una definizione di qualità del dato.

La qualità del dato si riferisce alla misura in cui i dati soddisfano le esigenze per cui sono stati raccolti e utilizzati. Ecco alcune delle dimensioni principali della qualità dei dati:



Mantenere alta la qualità dei dati è essenziale per garantire che le analisi e le decisioni basate su di essi siano accurate e affidabili.

Un concetto di qualità più trasversale è quello dell' idoneità allo scopo. Idoneità in questo caso si riferisce alla capacità dei dati di soddisfare le esigenze specifiche per cui sono stati raccolti e utilizzati. Pertanto possiamo effettivamente definire se un dato è di qualità o meno solo se abbiamo chiaro l'ambito all'interno del quale quel dato viene utilizzato: non si può prescindere dall'interpretare la qualità del dato all'interno del contesto di riferimento, all'interno della realtà che quel dato riflette.

Ecco alcuni aspetti chiave:

- **Adeguatezza:** i dati devono essere sufficienti e appropriati per rispondere alle domande o agli obiettivi dell'analisi. Ad esempio, se si sta analizzando il comportamento dei clienti, i dati devono includere informazioni rilevanti come acquisti, preferenze e feedback.
- **Pertinenza:** i dati devono essere strettamente correlati allo scopo dell'analisi. Dati non pertinenti possono introdurre rumore e distorcere i risultati.
- **Limitazione:** secondo il GDPR, i dati devono essere limitati a quanto necessario rispetto alle finalità per le quali sono trattati. Questo significa che non dovrebbero essere raccolti dati superflui che non contribuiscono direttamente agli obiettivi dell'analisi. Mantenere alta la qualità dei dati rispetto allo scopo aiuta a migliorare l'efficacia delle analisi e a prendere decisioni più informate e precise.

Interoperabilità semantica ed ontologie

L'interoperabilità semantica e l'uso delle ontologie rappresentano due pilastri fondamentali per l'efficace utilizzo dei dati in un contesto scientifico e tecnologico.

Un presupposto essenziale per l'uso efficace dei dati è la definizione e condivisione del loro significato. Questo non si limita al valore o all'informazione singola, ma si estende al contesto più ampio in cui l'informazione è inserita, ovvero il "concetto". Questo implica l'uso di un linguaggio comune e l'attribuzione di un significato uniforme alle frasi, non solo alle parole singole.

Il primo passo verso questo obiettivo è garantire l'interoperabilità semantica.

L'adozione di standard o di un modello di dati comune può facilitare questo processo, ma non assicura automaticamente l'omogeneità semantica delle informazioni raccolte. È fondamentale che tutti coloro che condividono le informazioni utilizzino lo stesso linguaggio, non solo in termini di archiviazione e condivisione dei dati, ma anche in termini di interpretazione e comprensione. Questo è particolarmente rilevante quando le informazioni vengono utilizzate in sistemi avanzati di analisi dei dati.

Tuttavia, la condivisione di un "linguaggio comune" non è sufficiente. È altrettanto importante condividere un modo comune di rappresentare i concetti e le relazioni tra di essi³.

Qui entrano in gioco le ontologie, che svolgono un ruolo cruciale nell'ambito dell'Intelligenza Artificiale e della gestione avanzata della conoscenza. Le ontologie permettono di utilizzare modelli di dati strutturati, in cui ogni singola informazione è parte di un modello organizzato in modo gerarchico e relazionale.

Questo approccio facilita l'interoperabilità tra diversi sistemi eterogenei, un aspetto di crescente importanza nell'era dell'apprendimento automatico e dell'intelligenza artificiale: l'uso efficace dei dati per supportare l'apprendimento automatico richiede infatti un alto grado di interoperabilità, che può essere raggiunto solo attraverso l'adozione di un linguaggio e di un modello di dati comuni.

³Haendel MA, Chute CG, Robinson PN. Classification, Ontology, and Precision Medicine. N Engl J Med. 2018 Oct 11;379(15):1452-1462. doi: 10.1056/NEJMra1615014. PMID: 30304648; PMCID: PMC6503847.

Perché raccogliere dati con modelli condivisi è fondamentale?

La raccolta di dati tramite modelli condivisi è un'attività rilevante per garantire l'efficacia, l'accuratezza e la protezione dei dati stessi, specialmente nel contesto delle regolamentazioni come il GDPR e l'European Health Data Space (EHDS). La crescente complessità e il volume dei dati richiedono approcci strutturati e uniformi per gestire le informazioni in modo efficiente e sicuro, rispondendo così a esigenze sia di qualità che di protezione dei dati. È importante evidenziare alcuni concetti chiave:

1. La **Data Quality** è fondamentale per garantire che le decisioni basate sui dati siano affidabili e precise⁴. Raccogliere dati con modelli condivisi ne migliora la qualità assicurando coerenza e standardizzazione. Modelli uniformi permettono una raccolta sistematica delle informazioni, riducendo il rischio di errori e discrepanze. La standardizzazione facilita l'analisi e l'interpretazione dei dati, consentendo una visione più chiara e dettagliata delle informazioni raccolte⁵. Senza un modello condiviso, la qualità dei dati può variare notevolmente, rendendo difficile il confronto e l'aggregazione delle informazioni.
2. Raccogliere dati con **modelli condivisi** promuove l'efficienza e il coordinamento tra diverse entità e sistemi. In ambito sanitario, ad esempio, l'EHDS stabilisce linee guida comuni per l'uso dei dati, sia primario che secondario, al fine di migliorare la ricerca e l'innovazione⁶. Modelli condivisi permettono una gestione più fluida dei dati tra diversi attori, riducendo i duplicati e migliorando l'integrazione dei dati provenienti da fonti diverse. Questo approccio non solo semplifica il processo

di raccolta e analisi dei dati, ma ottimizza anche le risorse e riduce i costi operativi.

3. Il Regolamento EHDS e il GDPR⁷ stabiliscono rigide normative per la protezione dei dati, richiedendo che essi siano gestiti in modo sicuro e conforme alle leggi. L'EHDS, approvato nell'aprile 2024⁸, fornisce un quadro giuridico armonizzato per l'**uso secondario** dei dati sanitari, introducendo requisiti per la protezione dei dati, la qualità e l'accesso. Modelli condivisi aiutano a garantire la conformità con queste normative, semplificando il rispetto dei requisiti di protezione e sicurezza. Il GDPR, in particolare, stabilisce principi chiave come la **Data Protection by Design** e la **Data Protection by Default**, che richiedono che i dati siano raccolti e trattati in modo sicuro e rispettoso della *privacy* fin dalla progettazione dei sistemi⁹. Utilizzare modelli condivisi aiuta a implementare questi principi in modo uniforme, riducendo il rischio di violazioni e garantendo una gestione adeguata dei dati personali.
4. La protezione dei dati e la *privacy* sono priorità fondamentali nella gestione dei dati, specialmente quando si tratta di dati sensibili come quelli sanitari. L'EHDS enfatizza l'importanza della sicurezza dei dati, prevedendo che l'accesso avvenga in ambienti sicuri e che i dati siano **anonimizzati o pseudonimizzati**. Modelli condivisi facilitano l'applicazione di misure di sicurezza uniformi e tecniche di anonimizzazione, riducendo il rischio di esposizione non autorizzata e garantendo che i dati siano trattati in conformità con le normative. Inoltre, il GDPR conferisce alle persone il diritto di accedere, correggere e cancellare i propri dati, e richiede che le organizzazioni implementino misure di protezione adeguate. Modelli condivisi possono migliorare la gestione di questi diritti, assicurando che le richieste di accesso o cancellazione dei dati siano gestite in modo coerente e tempestivo.

⁴Redman, T. C. (2016). *Data Driven: Creating a Data Culture*. Harvard Business Review Press.

⁵English, L. P. (2015). *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*. Wiley.

⁶Kok, I., et al. (2024). The European Health Data Space: A Major Step Towards a Digital Health Ecosystem. *European Journal of Public Health*. doi:10.1093/eurpub/ckaa257.

⁷European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation).

⁸European Commission. (2024). Factsheet: European Health Data Space (EHDS).

⁹Wright, D., & De Hert, P. (2018). *Enforcing GDPR Compliance: The Role of the Supervisory Authorities*. Springer.

5. Il Regolamento EHDS¹⁰ e il Data Act¹¹ stabiliscono specifiche regole per l'uso secondario dei dati sanitari, che includono la ricerca scientifica e l'innovazione. Per garantire che i dati siano utilizzati in modo efficace e conforme, è essenziale avere **modelli condivisi che facilitino l'accesso e la gestione dei dati in conformità con le normative**. L'EHDS prevede che i dati siano utilizzati solo per finalità legittime, come la ricerca scientifica, e richiede che le richieste di accesso siano autorizzate da organismi competenti. I modelli condivisi aiutano a standardizzare le pratiche di richiesta e autorizzazione, assicurando che l'uso dei dati secondari avvenga in modo trasparente e regolamentato. Questo approccio garantisce che i dati siano utilizzati in modo etico e che i diritti dei soggetti interessati siano rispettati.
6. Modelli condivisi facilitano l'innovazione e la collaborazione tra diverse organizzazioni e settori. In ambito sanitario, ad esempio, una raccolta

di dati coerente e standardizzata consente a ricercatori e sviluppatori di lavorare insieme per sviluppare nuove soluzioni e trattamenti. Questo approccio collaborativo è fondamentale per accelerare il progresso scientifico e migliorare le pratiche sanitarie.

Raccogliere dati con modelli condivisi è essenziale per garantire la qualità, l'efficienza e la protezione dei dati. Questi modelli non solo migliorano la coerenza e la precisione dei dati, ma facilitano anche la conformità con normative come il GDPR e l'EHDS. La standardizzazione nella raccolta e gestione dei dati permette di sfruttare appieno le potenzialità dei dati, promuovendo al contempo la sicurezza e la *privacy*. In un contesto in cui i dati sono una risorsa preziosa e le normative sono sempre più rigorose, adottare modelli condivisi è fondamentale per garantire un utilizzo efficace e responsabile delle informazioni.



¹⁰European Commission. (2024). European Health Data Space (EHDS) - Proposal for a Regulation.

¹¹European Commission. (2021). Data Governance Act (DGA) - Proposal for a Regulation.

Modelli condivisi

Nel vasto panorama dell'informatica sanitaria, l'adozione di standard oggi è basilare per garantire l'interoperabilità, la sicurezza e l'efficacia delle soluzioni informatiche impiegate nella pratica e nella ricerca clinica. Questi standard forniscono linee guida comuni per la raccolta, la gestione e lo scambio di informazioni sanitarie, promuovendo una migliore qualità dell'assistenza, la condivisione di dati tra sistemi eterogenei e la facilitazione della ricerca scientifica.

Tuttavia, i risultati oggi ottenuti sono frutto di diversi anni di studi e ricerche che, più o meno in parallelo con l'evoluzione tecnologica, hanno cercato il modo migliore per rappresentare e normalizzare le informazioni disponibili, al fine di gestirle e sfruttarle nel miglior modo possibile.

I modelli di dati (*Data Model*) sono **rappresentazioni astratte della struttura dei dati e delle loro relazioni all'interno di un sistema** informatico¹². Nel contesto sanitario e della ricerca clinica, i modelli di dati hanno avuto un ruolo cruciale nel supportare l'organizzazione, l'archiviazione e l'analisi delle informazioni mediche e scientifiche. Dal punto di vista storico, i primi modelli di dati furono sviluppati negli anni '60 e '70, in parallelo con l'evoluzione dei database. Il modello gerarchico e il modello reticolare furono tra i primi modelli utilizzati per rappresentare le relazioni tra i dati. Nel 1970, Edgar F. Codd propose il modello relazionale¹³, che divenne rapidamente dominante grazie alla sua semplicità e flessibilità. Questo modello si basava su tabelle (relazioni) e rappresentava i dati sotto forma di righe e colonne, permettendo operazioni come selezione, proiezione e join.

A partire dagli anni '80, i database relazionali divennero lo standard de facto per la gestione dei dati in molti settori, compreso quello sanitario. I sistemi di gestione di database relazionali (RDBMS) come Oracle, IBM

DB2, e Microsoft SQL Server divennero ampiamente adottati. È proprio durante questo periodo che, nel campo sanitario, iniziò a emergere la necessità di standardizzare i dati per migliorare la condivisione delle informazioni e l'interoperabilità tra sistemi diversi.

Sul finire degli anni '90, con la crescita di Internet e la necessità di scambiare dati tra sistemi eterogenei, XML (*eXtensible Markup Language*) diventò popolare come formato di interscambio dati. XML permetteva una rappresentazione flessibile dei dati e fu utilizzato per sviluppare modelli di dati specifici per la sanità, come HL7 (Health Level Seven), che vedremo in seguito.

L'introduzione e la diffusione delle cartelle cliniche elettroniche (EHR) ha reso essenziale l'utilizzo di modelli di dati standardizzati. Questi sistemi devono gestire enormi quantità di dati strutturati e non strutturati, inclusi diagnosi, prescrizioni, referti di laboratorio, immagini mediche e molto altro.

I modelli di dati utilizzati nelle EHR sono progettati per garantire che i dati possano essere facilmente recuperati, aggiornati e condivisi tra diverse istituzioni sanitarie. Proprio l'accorgersi della necessità di "far lavorare insieme" diversi sistemi ha portato alla ribalta l'interoperabilità, diventata così nel tempo una delle principali sfide nell'informatica sanitaria. Standard come HL7 FHIR (Fast Healthcare Interoperability Resources) sono stati sviluppati per affrontare questa sfida, fornendo una base comune per rappresentare e scambiare dati clinici in modo uniforme.

Con l'avvento dei Big Data, i modelli di dati si sono evoluti per gestire grandi volumi di informazioni provenienti da fonti diverse come dispositivi indossabili, genomica, imaging medico e trial clinici¹⁴. La ricerca clinica ha beneficiato di modelli di dati più complessi che integrano dati strutturati e non strutturati, consentendo l'applicazione di tecniche di machine learning per scoprire nuovi *pattern* e supportare la medicina personalizzata.

¹²Cfr. Elmasri, R., & Navathe, S. B. (2010). *Fundamentals of Database Systems* (6th Edition). Addison-Wesley.

¹³Codd, E. F. (1970). A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6), 377-387.

¹⁴Raghupathi, W., & Raghupathi, V. (2014). Big Data Analytics in Healthcare: Promise and Potential. *Health Information Science and Systems*, 2(1), 1-10.

Organizzazioni internazionali, come l'ISO (*International Organization for Standardization*¹⁵) e il CDISC (*Clinical Data Interchange Standards Consortium*), hanno sviluppato standard per i modelli di dati utilizzati nella ricerca clinica. Standard che aiutano a garantire che i dati raccolti durante i trial clinici siano accurati, coerenti e utilizzabili per l'analisi statistica. Ne è un esempio il modello CDASH (*Clinical Data Acquisition Standards Harmonization*), che fornisce linee guida per la raccolta dei dati clinici, garantendo la qualità e la conformità con i requisiti normativi.

In questa panoramica, esamineremo alcuni dei principali standard utilizzati in ambito sanitario.

Health Level Seven International (HL7)

Oggi uno degli standard più rilevanti e citati è l'HL7 (Health Level Seven International)¹⁶, un insieme di protocolli e formati di messaggi che mira a definire come i dati clinici debbano essere scambiati tra i sistemi informativi sanitari.

Si tratta di uno degli standard più utilizzati nell'informatica sanitaria per lo scambio di dati clinici e amministrativi tra software e sistemi ospedalieri. Definisce una serie di regole che permettono di inviare messaggi contenenti informazioni sanitarie — referti di laboratorio, dati delle cartelle cliniche, programmi di appuntamenti, e altre comunicazioni — tra le varie applicazioni software utilizzate dalle strutture sanitarie¹⁷.

HL7 si propone di standardizzare il modo in cui le informazioni sanitarie vengono scambiate e gestite tra diversi sistemi informatici. Le finalità principali di HL7 includono:

- **Interoperabilità dei sistemi:** HL7 fornisce un *framework* per l'integrazione e la comunicazione tra diversi sistemi sanitari, inclusi i sistemi di gestione delle cartelle cliniche elettroniche (EHR), i sistemi di laboratorio e i sistemi di *imaging*;

- **Standardizzazione della trasmissione dei dati:** consente la trasmissione coerente e comprensibile delle informazioni sanitarie, facilitando la condivisione di dati tra diversi fornitori e piattaforme;
- **Supporto alla qualità delle cure:** migliora la qualità delle cure attraverso una comunicazione più efficace e una condivisione più fluida delle informazioni cliniche tra professionisti della salute;
- **Efficienza e riduzione degli errori:** automatizza e standardizza i processi di scambio di informazioni, riducendo la possibilità di errori e migliorando l'efficienza operativa.

HL7 è stato fondato nel 1987 come un'organizzazione no-profit per sviluppare standard per la comunicazione elettronica in ambito sanitario¹⁸. Il nome "Health Level Seven" riflette il focus dell'organizzazione sul **settimo livello del modello OSI** (*Open Systems Interconnection*), che riguarda l'applicazione e l'interazione tra applicazioni.

Il primo standard HL7, HL7 Version 2.x, è stato pubblicato nel 1989 e ha introdotto un linguaggio comune per l'interscambio di messaggi tra sistemi sanitari. Questo standard ha avuto un impatto significativo, venendo adottato da molti ospedali e cliniche per gestire i dati clinici e amministrativi.

Successivamente, HL7 ha sviluppato HL7 Version 3.x, che introduce un *framework* basato su XML e una maggiore precisione nella modellazione dei dati clinici. La versione 3 ha cercato di risolvere le limitazioni del modello 2.x, ma ha incontrato alcune sfide nell'adozione. Nel 2009, HL7 ha pubblicato HL7 FHIR (Fast Healthcare Interoperability Resources), un nuovo standard progettato per migliorare l'interoperabilità attraverso API web moderne e un'architettura basata su risorse. FHIR ha guadagnato un'ampia accettazione grazie alla sua flessibilità e alla compatibilità con le tecnologie web moderne.

¹⁵ISO/TC 215 (2017). ISO 13606: Health Informatics—Electronic Health Record Communication. International Organization for Standardization.

¹⁶Health Level Seven International (2020). HL7 Standards. HL7 International.

¹⁷Dolin, R. H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F. M., Biron, P. V., & Shabo, A. (2006). HL7 Clinical Document Architecture, Release 2. *Journal of the American Medical Informatics Association*, 13(1), 30-39.

¹⁸Hersh, W., & Hickam, D. (2008). Health Level Seven (HL7) Standards: A Brief History and Overview. *Journal of the American Medical Informatics Association*, 15(3), 317-325.

HL7 utilizza una varietà di standard e specifiche per gestire le informazioni sanitarie:

- HL7 Version 2.x: utilizza un formato basato su messaggi con segmenti e campi predefiniti per la trasmissione delle informazioni.
- HL7 Version 3.x: adotta un approccio basato su XML e utilizza modelli di informazioni per una rappresentazione più strutturata e coerente dei dati clinici¹⁹.
- HL7 FHIR: si basa su una serie di risorse standardizzate che possono essere facilmente integrate tramite API web, consentendo una comunicazione più agile e flessibile.

HL7 è ampiamente adottato in tutto il mondo e ha un impatto significativo sull'interoperabilità dei sistemi

informativi sanitari. FHIR, in particolare, ha facilitato l'integrazione di nuove tecnologie e migliorato la comunicazione tra sistemi diversi. HL7 continua a evolversi, con aggiornamenti regolari e nuovi standard per rispondere alle esigenze emergenti della sanità digitale. FHIR è una specifica più recente che utilizza i moderni standard web, inclusi XML e JSON, per trasmettere dati sanitari. È progettato per essere più facile da implementare rispetto alle versioni precedenti di HL7 e supporta una varietà di modelli di interoperabilità, compresi RESTful APIs e SOAP. FHIR definisce risorse di base che possono essere utilizzate da sole o combinate per gestire complesse necessità cliniche e di assistenza sanitaria, rendendo lo scambio di dati più intuitivo e flessibile²⁰.



¹⁹Grannis, S. J., & Overhage, J. M. (2006). HL7 Version 3: An Overview and Future Directions. *Journal of Biomedical Informatics*, 39(2), 190-195.

²⁰Murphy, S. N., & Mendis, M. (2012). FHIR: An Agile Standards Development Methodology for Healthcare Interoperability. *AMIA Annual Symposium Proceedings*, 2012, 139-143.

DICOM

DICOM (*Digital Imaging and Communications in Medicine*) è uno standard globale per la gestione, l'archiviazione, la trasmissione e la visualizzazione delle immagini mediche. Creato per facilitare l'interoperabilità tra dispositivi di imaging e sistemi informatici in ambito sanitario, DICOM è essenziale per la pratica della radiologia e altre specialità che utilizzano immagini diagnostiche. Originariamente, è stato sviluppato per affrontare la crescente complessità e varietà delle immagini mediche e per assicurare che tali immagini e i dati associati potessero essere condivisi e utilizzati in modo efficace tra diverse macchine e software.

Le principali finalità di DICOM includono:

- Interoperabilità delle immagini mediche: DICOM consente la comunicazione e lo scambio di immagini mediche tra apparecchiature di imaging (come TAC, RMN, ecografi) e sistemi informativi, indipendentemente dal produttore o dal fornitore del software²¹;
- Standardizzazione dei formati di immagine: assicura che le immagini prodotte da diversi dispositivi abbiano un formato standard, facilitando l'archiviazione, il recupero e l'analisi delle immagini;
- Integrazione dei sistemi informativi sanitari: DICOM permette l'integrazione delle immagini nei sistemi di gestione delle cartelle cliniche elettroniche (EHR) e nei sistemi PACS (*Picture Archiving and Communication System*), migliorando l'accessibilità e la gestione dei dati;
- Supporto alla diagnostica e al trattamento: fornendo un accesso uniforme e standardizzato alle immagini mediche, DICOM supporta il processo decisionale clinico e la pianificazione dei trattamenti.

DICOM è stato sviluppato a partire dal 1983 dall'American College of Radiology (ACR) e dalla National Electrical Manufacturers Association (NEMA), con

l'obiettivo di creare uno standard universale per la trasmissione e l'archiviazione delle immagini mediche. La prima versione è stata pubblicata nel 1985, ma presentava alcune limitazioni che ne ostacolavano l'adozione su larga scala: queste difficoltà hanno portato allo sviluppo della versione successiva, DICOM 3.0, rilasciata nel 1993²². Questa versione ha introdotto un significativo miglioramento nella flessibilità e nelle capacità di integrazione del sistema, rendendolo uno standard globale per le immagini mediche. Oggi, DICOM supporta non solo le immagini bidimensionali, ma anche immagini tridimensionali, video, e altri dati multimediali associati all'imaging medico.

DICOM definisce sia il formato dei file di immagini mediche che i protocolli di comunicazione utilizzati per lo scambio di tali immagini. Gli elementi chiave della struttura DICOM includono²³:

- File header: ogni file DICOM inizia con un header che contiene informazioni sul paziente, sull'esame e sulle impostazioni tecniche dell'immagine;
- Data set: il corpo del file contiene i dati effettivi dell'immagine, codificati in un formato standardizzato;
- Tags DICOM: ogni elemento di dati è identificato da un tag DICOM, che specifica la natura dell'informazione (es. nome del paziente, data dell'esame, parametri tecnici dell'immagine).

DICOM è ora utilizzato in tutto il mondo come standard *de facto* per l'imaging medico. Viene supportato da tutti i principali produttori di apparecchiature mediche e software di imaging, e la sua adozione è obbligatoria in molti sistemi sanitari nazionali (tra cui quello italiano). La sua continua evoluzione garantisce che rimanga rilevante e adattabile alle nuove tecnologie emergenti nel campo della sanità digitale.

²¹Huang, H. K. (2010). *PACS and Imaging Informatics: Basic Principles and Applications* (2nd Edition). Wiley-Blackwell.

²²Bryan, S., Weatherburn, G. C., Watkins, J. R., & Buxton, M. J. (1999). The Digital Imaging and Communications in Medicine (DICOM) Standard: A Brief Overview and Its Impact on Radiological Practice. *Clinical Radiology*, 54(9), 640-647.

²³Mildenberger, P., Eichelberg, M., & Martin, E. (2002). Introduction to the DICOM Standard. *European Radiology*, 12(4), 920-927.

International Classification of Diseases (ICD)

Per quanto riguarda la gestione dei dati relativi ai pazienti e alle caratteristiche degli eventi clinici come ricoveri e accessi ambulatoriali, l'adozione di standard per l'identificazione univoca di diagnosi ed interventi è essenziale.

ICD (*International Classification of Diseases*) è un sistema standardizzato di classificazione²⁴, sviluppato dall'Organizzazione Mondiale della Sanità (OMS), per codificare diagnosi, malattie, lesioni e altre condizioni di salute ed è estremamente diffuso per la raccolta e l'analisi dei dati sanitari a livello globale, regionale e nazionale.

L'ICD è stato creato per fornire un linguaggio comune che permetta ai professionisti sanitari di documentare e comunicare in modo omogeneo le condizioni di salute dei pazienti. Le principali finalità di ICD includono:

- Monitoraggio della salute pubblica: ICD è ampiamente utilizzato per monitorare l'incidenza e la prevalenza delle malattie a livello locale, nazionale e globale, facilitando la pianificazione sanitaria e la risposta alle emergenze;
- Codifica delle diagnosi: fornisce un sistema di codifica per le diagnosi cliniche, che viene utilizzato nei registri medici, nei sistemi di fatturazione sanitaria e per scopi statistici;
- Ricerca clinica ed epidemiologica: consente ai ricercatori di aggregare e confrontare i dati sulla salute provenienti da diverse fonti, supportando studi epidemiologici e clinici su larga scala;
- Politiche sanitarie e finanziamento: le codifiche ICD vengono utilizzate dai governi e dalle compagnie assicurative per determinare il finanziamento delle cure sanitarie e per definire le politiche di rimborso.

L'ICD utilizza una struttura gerarchica per classificare le malattie e le condizioni di salute. Ogni condizione è identificata da un codice alfanumerico che corrisponde a una specifica categoria.

²⁴Cfr. World Health Organization (2019). ICD-11 for Mortality and Morbidity Statistics (ICD-11 MMS). World Health Organization.

²⁵Butler, M., & McCreedy, E. (2018). Use of ICD-10 and ICD-11 Classifications in Clinical and Health Service Delivery Applications. *Health Information Management Journal*, 47(1), 29-33.

La struttura principale include:

- Capitoli: raggruppamenti di condizioni correlate, ad esempio malattie infettive, neoplasie, malattie del sistema circolatorio, ecc;
- Blocchi: sottogruppi all'interno di ogni capitolo che organizzano le malattie in categorie più specifiche;
- Categorie: le singole condizioni, ognuna delle quali ha un codice unico, spesso con sottocategorie per distinguere variazioni della stessa condizione.

La codifica ICD è utilizzata in una varietà di contesti sanitari, inclusa la documentazione delle cartelle cliniche, la gestione delle assicurazioni sanitarie, la valorizzazione economica dei ricoveri, la pianificazione delle risorse sanitarie e il monitoraggio delle epidemie e delle pandemie.

Attualmente, sebbene in Italia si faccia ancora largo uso della versione 9, è in uso internazionalmente la decima revisione del sistema, nota come ICD-10, che ha introdotto un livello di dettaglio molto maggiore rispetto alle versioni precedenti. L'ICD-10 è composto da migliaia di codici alfanumerici che forniscono una descrizione dettagliata delle diagnosi mediche e dei processi patologici. Ogni codice è strutturato in modo da rappresentare una specifica condizione, consentendo una classificazione precisa.

Una nuova revisione, l'ICD-11²⁵, è stata adottata nel 2019 e si prevede che entrerà in vigore a partire dal 2022. L'ICD-11 promette di essere più aggiornata rispetto ai progressi scientifici e alle esigenze di salute pubblica. È stata progettata per essere più accessibile e più facile da implementare, grazie all'uso della tecnologia digitale e alla possibilità di integrarsi meglio con altri standard e sistemi di classificazione. L'ICD-11 introduce anche nuovi capitoli e concetti, come la medicina tradizionale e i determinanti sociali della salute, ampliando così la sua applicabilità e rilevanza in diversi contesti culturali e sanitari.

L'ICD ha radici nel XIX secolo, con la prima versione pubblicata nel 1893 come "International List of Causes of Death". Inizialmente sviluppato dall'Office International d'Hygiène Publique (precursore dell'OMS), fu adottato da diversi paesi per standardizzare la registrazione delle cause di morte. Con il tempo, l'ICD è stato ampliato per includere non solo le cause di morte, ma anche una vasta gamma di malattie e condizioni cliniche. L'OMS ha assunto la gestione dell'ICD nel 1948 e ha pubblicato la 6ª revisione (ICD-6), che fu la prima a includere sia le malattie che le cause di morte²⁶. Le successive revisioni hanno ampliato ulteriormente l'ambito e la granularità dell'ICD.

SNOMED

SNOMED CT (*Systematized Nomenclature of Medicine – Clinical Terms*) è un sistema di classificazione e terminologia clinica internazionale progettato per rappresentare in modo dettagliato e coerente le condizioni di salute, le osservazioni cliniche e le procedure mediche. SNOMED CT è utilizzato per garantire che i dati clinici siano descritti in modo uniforme e preciso, facilitando la condivisione e l'analisi delle informazioni sanitarie.

SNOMED CT ha diverse finalità principali che includono:

- Standardizzazione della terminologia clinica: fornisce un vocabolario comune e standardizzato per descrivere le condizioni mediche, le osservazioni e le procedure, migliorando la coerenza nella documentazione clinica;
- Facilitazione dell'interoperabilità: consente l'integrazione e l'interoperabilità dei dati clinici tra diversi sistemi e applicazioni, supportando la condivisione delle informazioni tra strutture sanitarie diverse;
- Supporto alla qualità e alla ricerca: migliora la qualità della documentazione clinica e facilita la raccolta di dati per la ricerca e l'analisi epidemiologica.

- Assistenza alla decisione clinica: fornisce una base dettagliata per i sistemi di supporto alle decisioni cliniche, aiutando i professionisti della salute a prendere decisioni più informate.

SNOMED CT è organizzato in un sistema gerarchico che include:

- Concetti: le unità fondamentali che rappresentano condizioni mediche, sintomi, procedure e altri elementi clinici;
- Descrittori: sinonimi e definizioni per ogni concetto, che aiutano a chiarire e standardizzare i termini utilizzati;
- Relazioni: connessioni tra concetti che descrivono le loro interazioni e attributi, come le relazioni di "causa", "sito" e "stato";
- Categorizzazioni: strutture che organizzano i concetti in categorie e sotto-categorie per facilitare la navigazione e l'utilizzo.

SNOMED CT è ampiamente adottato in tutto il mondo come uno standard di terminologia clinica. La sua continua evoluzione e aggiornamento garantiscono che rimanga rilevante e utile nel contesto della sanità moderna e digitale. SNOMED CT è un componente rilevante²⁷, per quanto non ancora pienamente fruito in Italia, per l'integrazione dei dati clinici e il supporto alla medicina basata su prove.

SNOMED è stato sviluppato originariamente negli anni '60 come una terminologia medica progettata per migliorare la qualità della documentazione clinica. Il sistema è stato inizialmente conosciuto come SNOMED RT (*Reference Terminology*) e ha evoluto nel tempo per includere una gamma più ampia di concetti clinici e relazioni.

Nel 2002, SNOMED è stato ampliato e riorganizzato come SNOMED CT, una terminologia clinica globale che incorpora una vasta gamma di concetti e relazioni cliniche. SNOMED CT è stato sviluppato dal *College of*

²⁶Moriyama, I. M., Loy, R. M., & Robb-Smith, A. H. T. (2011). History of the Statistical Classification of Diseases and Causes of Death. World Health Organization.

²⁷Elkin, P. L., & Brown, S. H. (2008). Historical Perspective on SNOMED and Its Impact on Clinical Data Standards. *Studies in Health Technology and Informatics*, 136, 19-23.



American Pathologists (CAP) e successivamente gestito e mantenuto dalla *International Health Terminology Standards Development Organisation* (IHTSDO), ora conosciuta come SNOMED International.

Il sistema SNOMED CT è stato adottato a livello globale e integrato in numerosi sistemi di gestione delle informazioni sanitarie, migliorando la qualità e la coerenza dei dati clinici a livello internazionale.

LOINC

LOINC (*Logical Observation Identifiers Names and Codes*) è uno standard globale per l'identificazione di test di laboratorio, misurazioni cliniche, osservazioni e documenti sanitari.

Creato per facilitare la condivisione e l'integrazione dei dati tra diversi sistemi informativi sanitari, LOINC fornisce un vocabolario standardizzato che consente la codifica univoca delle informazioni cliniche, migliorando l'interoperabilità e la qualità delle cure.

Prima della sua introduzione, i laboratori e i sistemi sanitari utilizzavano spesso codici e terminologie proprietarie, rendendo difficile lo scambio accurato di informazioni tra diverse istituzioni e piattaforme software. Questo ostacolava la continuità delle cure, la ricerca clinica e la gestione della salute pubblica.

Le finalità principali di LOINC includono:

- Interoperabilità dei dati clinici: LOINC permette la condivisione precisa e coerente dei dati clinici tra diverse organizzazioni sanitarie, facilitando l'integrazione delle informazioni provenienti da fonti diverse, come laboratori, ospedali e cliniche;
- Standardizzazione delle osservazioni cliniche: fornisce un sistema comune di codifica per test di laboratorio, misurazioni fisiologiche, e altre osservazioni cliniche, standardizzando il modo in cui queste informazioni vengono registrate e trasmesse;

- Supporto alla ricerca e alla sanità pubblica: utilizzando LOINC, i ricercatori e gli enti di sanità pubblica possono aggregare e analizzare i dati clinici provenienti da fonti diverse, consentendo studi epidemiologici e monitoraggio delle malattie su larga scala;
- Facilitazione della medicina basata sui dati: LOINC aiuta a integrare i dati clinici nei sistemi di supporto alle decisioni mediche, migliorando l'accuratezza diagnostica e il trattamento dei pazienti.

LOINC utilizza una struttura a sei dimensioni per codificare ogni test o osservazione clinica. Queste dimensioni includono²⁹:

- Component: descrive ciò che viene misurato, osservato o valutato;
- Property: indica la caratteristica della misura (es. concentrazione, presenza/assenza).
- Timing: specifica il momento o l'intervallo di tempo della misurazione (es. puntuale, 24 ore);
- System: rappresenta il sistema o la parte del corpo a cui si riferisce il test (es. sangue, urina);
- Scale: definisce la scala di misura (es. quantitativo, qualitativo);
- Method: indica la tecnica utilizzata per effettuare la misurazione (es. metodo enzimatico).

²⁸Spackman, K. A., & Campbell, K. E. (2004). The Evolution of SNOMED: From SNOMED RT to SNOMED CT. *International Journal of Medical Informatics*, 73(4), 277-287.

²⁹Huff, S. M., Rocha, R. A., McDonald, C. J., De Moor, G. J., Fiers, T., Bidgood, W. D., ... & Dolin, R. H. (1998). Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary. *Journal of the American Medical Informatics Association*, 5(3), 276-292.

LOINC è stato sviluppato nel 1994 dal Regenstrief Institute, un'organizzazione di ricerca sanitaria senza scopo di lucro con sede a Indianapolis (Usa)³⁰ in risposta alla crescente complessità dei sistemi sanitari e alla necessità di migliorare la comunicazione e lo scambio di dati tra laboratori e fornitori di servizi sanitari³¹. In particolare, la mancanza di standardizzazione nelle modalità con cui i risultati dei test di laboratorio venivano registrati e trasmessi rappresentava un ostacolo significativo per la condivisione dei dati e la collaborazione tra diversi enti sanitari.

Nel corso degli anni, LOINC è stato continuamente aggiornato e ampliato per coprire nuove aree della medicina e per adattarsi alle esigenze emergenti della sanità globale. LOINC è ora integrato in molti sistemi di gestione delle informazioni sanitarie, inclusi software di EHR (*Electronic Health Records*), sistemi di laboratorio e reti di sanità pubblica. Oggi, il suo utilizzo è promosso da organizzazioni come l'Organizzazione Mondiale della Sanità e gli istituti nazionali di sanità pubblica in diversi Paesi, contribuendo a una maggiore coerenza e qualità nell'assistenza sanitaria globale.

I Data model

Affinché le informazioni sanitarie possano essere utilizzate in modo efficace, è essenziale che siano organizzate secondo modelli di dati strutturati che permettano di **interpretare, condividere e analizzare i dati in modo standardizzato**, esportabile in ambito transnazionale, con risultati ripetibili e verificabili. La standardizzazione dei dati clinici secondo i Data model ha diversi vantaggi, come ad esempio l'accelerare il time-to-market dei nuovi farmaci e trattamenti, migliorando la comunicazione tra tutti gli *stakeholder* coinvolti nel processo di sviluppo terapeutico.

OMOP (*Observational Medical Outcomes Partnership*) è un'iniziativa di ricerca collaborativa che si concentra sulla creazione di una base di dati e di standardizzazione per l'analisi dei dati sanitari provenienti da studi

osservazionali³². L'obiettivo principale di OMOP è migliorare la qualità e l'affidabilità degli studi osservazionali attraverso l'adozione di standard comuni e metodi analitici avanzati³³.

Le principali finalità di OMOP includono:

- Standardizzazione dei dati: OMOP promuove l'uso di standard comuni per la raccolta, la gestione e l'analisi dei dati sanitari, al fine di garantire la coerenza e la comparabilità tra diversi studi e database;
- Promozione della ricerca: fornisce strumenti e risorse per facilitare la ricerca su effetti avversi dei farmaci e altre questioni di salute pubblica utilizzando dati reali provenienti da studi osservazionali;
- Miglioramento della qualità degli studi: attraverso metodologie standardizzate, OMOP mira a migliorare la qualità e l'affidabilità degli studi basati su dati osservazionali;
- Facilitazione della collaborazione: facilita la collaborazione tra ricercatori, istituzioni e aziende farmaceutiche per migliorare la ricerca e la condivisione delle conoscenze.

La struttura di OMOP è basata su:

- *Common Data Model* (CDM): fornisce una struttura uniforme per la rappresentazione dei dati clinici, che include tabelle e campi standardizzati per le informazioni sui pazienti, le diagnosi, i trattamenti e gli eventi avversi;
- Analytical tools: include strumenti e metodologie per l'analisi dei dati, come il Achilles e il ATLAS, che aiutano a esplorare e visualizzare i dati all'interno del CDM;
- Standardized vocabularies: utilizza vocabolari standardizzati per garantire la coerenza nella rappresentazione dei dati clinici e facilitare l'integrazione di informazioni provenienti da fonti diverse;

³⁰McDonald, C. J. (1996). The LOINC Project: A Universal Code System for Tests, Observations, and Measurements. *Clinical Chemistry*, 42(5), 839-848.

³¹McDonald, C. J., Huff, S. M., Suico, J. G., Hill, G., Leavelle, D., Aller, R., ... & Maloney, P. C. (2003). LOINC, a Universal Standard for Identifying Laboratory Observations: A 5-Year Update. *Clinical Chemistry*, 49(4), 624-633.

³²Hripcsak, G., & Albers, D. J. (2013). Next-generation phenotyping of electronic health records. *Journal of Biomedical Informatics*, 46(5), 792-799.

³³Stang, P. E., & Ryan, P. B. (2013). Common Data Model for Observational Research: OMOP's Contributions to Standardizing Data. *Health Information Management Journal*, 42(3), 35-42.

- Collaborative research network: una rete globale di ricercatori e istituzioni che collaborano per migliorare la qualità della ricerca basata su dati osservazionali e promuovere l'uso del CDM.

OMOP, attraverso OHDSI, ha avuto un impatto significativo sulla ricerca basata su dati osservazionali, promuovendo standardizzazione e collaborazione globale. Oggi, uno degli obiettivi principali dell'OMOP è quello di consentire ai ricercatori di condurre analisi trasversali su diverse popolazioni e sistemi sanitari, migliorando la capacità di generare evidenze cliniche reali (Real-World Evidence, RWE). L'adozione del CDM e dei suoi strumenti è in crescita, con numerosi studi e progetti che utilizzano il modello per migliorare la qualità e l'affidabilità della ricerca sanitaria, riducendo la complessità e i costi associati alla trasformazione dei dati da fonti eterogenee, garantendo che le analisi siano riproducibili e affidabili su scala internazionale.

OMOP è stato avviato nel 2008 con il supporto dell'FDA (*Food and Drug Administration*) e del *Pharmaceutical Research and Manufacturers of America* (PhRMA). L'iniziativa è nata in risposta alla crescente necessità di standardizzare e armonizzare i dati sanitari provenienti da studi osservazionali, che sono fondamentali per l'analisi dell'efficacia e della sicurezza dei farmaci: di fatto, il modello OMOP è stato inizialmente progettato per facilitare studi di farmacovigilanza, valutazione comparativa dell'efficacia e studi osservazionali su larga scala. Nel 2011, OMOP ha lanciato il *Common Data Model* (CDM), che ha fornito una struttura standardizzata per la rappresentazione e l'analisi dei dati clinici. Il CDM ha facilitato la comparazione e l'integrazione dei dati da diverse fonti e ha migliorato l'analisi dei risultati della ricerca. Nel 2013, OMOP è stato integrato nel *Observational Health Data Sciences and Informatics* (OHDSI), una comunità globale di ricerca che continua a espandere e sviluppare il modello e le risorse di OMOP³⁴.

Continuando l'exkursus sui Data Model, citiamo anche **CDISC** (*Clinical Data Interchange Standards Consortium*), un'organizzazione internazionale che sviluppa e promuove standard per la raccolta, la gestione e l'analisi dei dati clinici. CDISC si concentra sulla creazione di standard che facilitino l'interoperabilità e l'efficienza nella ricerca clinica e nella gestione dei dati³⁵.

Le principali finalità di CDISC includono:

- Standardizzazione dei dati clinici: fornisce standard uniformi per la raccolta e la gestione dei dati clinici, migliorando la qualità e la coerenza dei dati attraverso studi clinici e ricerche;
- Facilitazione della conformità regolatoria: aiuta le aziende farmaceutiche e le istituzioni di ricerca a rispettare le normative e i requisiti regolatori attraverso l'adozione di standard comuni;
- Promozione dell'interoperabilità: migliora la capacità di integrare e condividere dati tra diversi sistemi e applicazioni, facilitando la comunicazione e la collaborazione tra ricercatori e organizzazioni;
- Ottimizzazione dei processi di ricerca: semplifica e accelera il processo di raccolta e analisi dei dati clinici, migliorando l'efficienza e riducendo i costi associati alla ricerca.

La struttura di CDISC è organizzata intorno a:

- Standard di dati: include modelli e linee guida per la raccolta, la tabulazione e l'analisi dei dati clinici, come SDTM e ADaM³⁶;
- Rilevanza regolatoria: garantisce che gli standard siano allineati con i requisiti regolatori globali, facilitando l'invio di dati alle autorità come la FDA e l'EMA;
- Infrastruttura di supporto: fornisce risorse e strumenti per supportare l'implementazione degli standard, inclusi manuali, linee guida e strumenti di convalida;

³⁴Ryan, P. B., & Schuemie, M. J. (2014). The Observational Health Data Sciences and Informatics (OHDSI) initiative: Collaborative research to improve the quality of observational data. *Journal of Biomedical Informatics*, 52, 105-116.

³⁵Wang, X., & Reddy, S. (2016). Clinical Data Interchange Standards Consortium (CDISC): A Historical and Technical Review. *Journal of Biomedical Informatics*, 59, 319-327.

³⁶Devine, J., & Daugherty, J. (2010). The Development and Implementation of CDISC Standards: A Historical Perspective. *Drug Information Journal*, 44(1), 15-21.

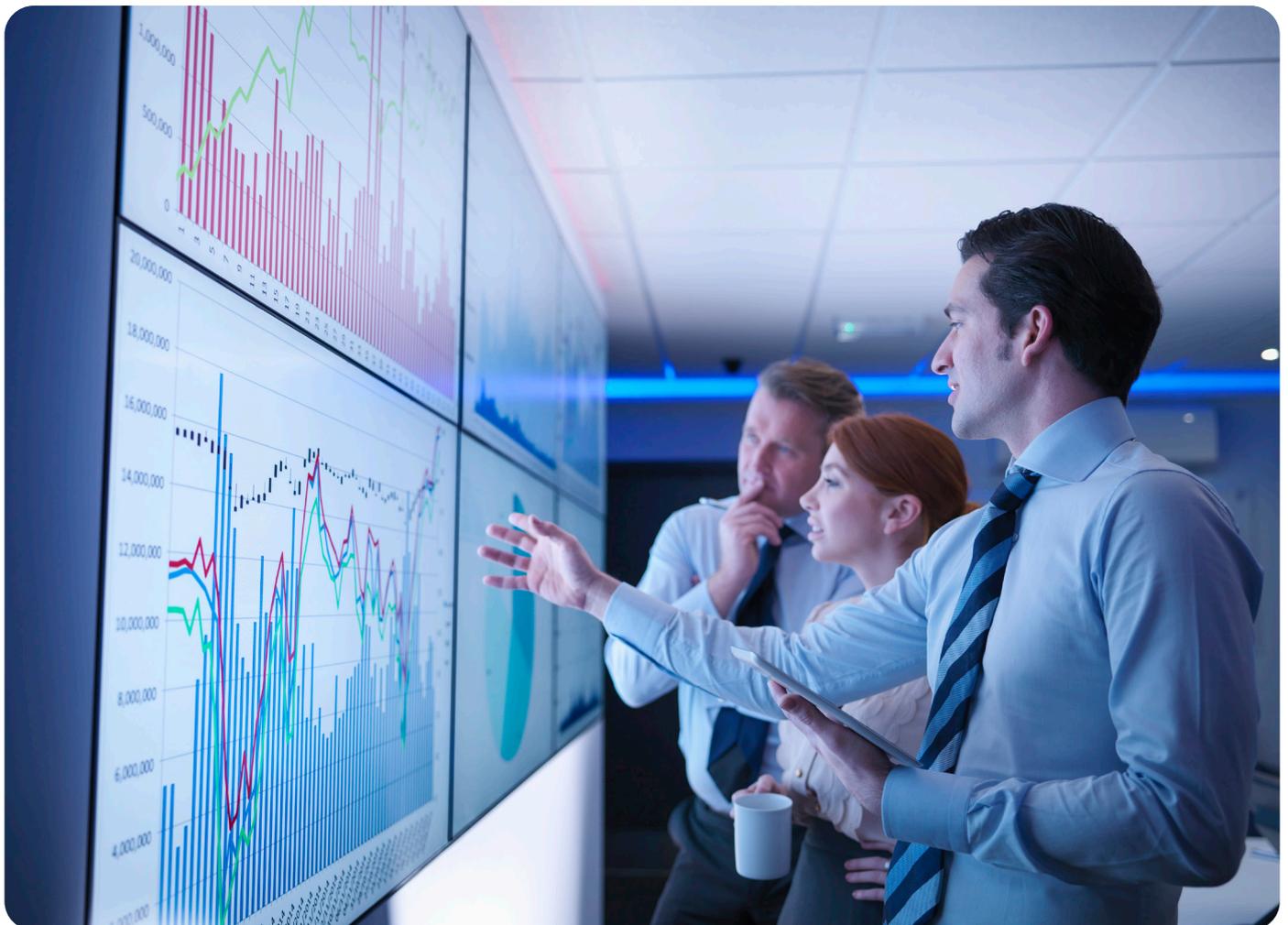
- Collaborazione e aggiornamento: CDISC collabora con ricercatori, industrie e autorità regolatorie per aggiornare e migliorare continuamente gli standard in base alle esigenze emergenti e alle evoluzioni del settore.

CDISC ha avuto un impatto notevole sulla standardizzazione dei dati clinici e continua a essere un leader nella promozione dell'interoperabilità e della qualità nella ricerca clinica. Questi standard sono riconosciuti a livello globale e sono raccomandati dalle agenzie regolatorie, come la FDA negli Stati Uniti e l'EMA in Europa, per la presentazione dei dati nei processi di approvazione dei nuovi farmaci e dispositivi medici.

CDISC è stato fondato nel 1997 come una organizzazione no-profit con l'obiettivo di migliorare la qualità e l'efficienza della ricerca clinica attraverso

la standardizzazione dei dati³⁷. L'organizzazione ha sviluppato una serie di standard fondamentali, tra cui:

- CDASH (Clinical Data Acquisition Standards Harmonization): un insieme di linee guida per la raccolta e la gestione dei dati clinici.
- SDTM (Study Data Tabulation Model): un modello per la tabulazione dei dati clinici che facilita la loro presentazione e l'invio alle autorità regolatorie.
- ADaM (Analysis Data Model): fornisce un modello standard per la preparazione dei dati analitici necessari per la revisione dei risultati degli studi clinici.
- SEND (Standard for the Exchange of Nonclinical Data): un standard per la rappresentazione dei dati non clinici, utilizzato principalmente nella ricerca preclinica.



Maturity Model: il ruolo di HIMSS

Il maturity model, o modello di maturità, rappresenta uno strumento essenziale nel contesto della sanità digitale per misurare e guidare il progresso tecnologico e organizzativo delle istituzioni sanitarie. L'organizzazione globale HIMSS (*Healthcare Information and Management Systems Society*) è un ente di riferimento nel settore, dedicato a promuovere l'adozione e l'implementazione delle tecnologie dell'informazione in ambito sanitario per realizzare il pieno potenziale della salute di ogni persona nel mondo. HIMSS non solo si distingue per la sua pervasività globale, ma anche per il suo approccio pratico e collaborativo nel promuovere best practice e innovazione. La missione di HIMSS è chiara: migliorare la salute globale attraverso l'uso efficace delle tecnologie dell'informazione. Questa organizzazione svolge un ruolo fondamentale non solo nella promozione di soluzioni IT avanzate, ma anche nel fornire una piattaforma per lo scambio di conoscenze e esperienze tra professionisti della salute e organizzazioni sanitarie di tutto il mondo, utile per affrontare le sfide quotidiane e per permettere a diverse organizzazioni di apprendere dalle soluzioni adottate altrove.

Tra le molteplici attività di HIMSS, i modelli di maturità sono forse uno degli strumenti più noti e utilizzati. Questi modelli offrono un *framework* per valutare il livello di avanzamento digitale di un'organizzazione sanitaria e per tracciare un percorso di crescita strutturato e basato su standard globali. I modelli di maturità di HIMSS non sono solo metriche di misurazione, ma fungono anche da roadmap per guidare le organizzazioni attraverso il complesso processo di digitalizzazione. Uno dei modelli più riconosciuti è l'EMRAM (Electronic Medical Record Adoption Model), che misura l'adozione dei registri medici elettronici all'interno delle strutture sanitarie. Questo modello si articola su otto livelli, da 0 a 7, dove il livello 7 rappresenta l'eccellenza nell'uso dei sistemi elettronici per la gestione dei dati clinici. Nel contesto attuale, il modello EMRAM è stato aggiornato per

includere elementi di continuità delle cure, riflettendo l'importanza della gestione integrata dei dati del paziente al di là del semplice ricovero ospedaliero. Questo aggiornamento è rilevante poiché riconosce la necessità di considerare l'intera storia clinica del paziente, inclusa la prevenzione e la continuità delle cure, come componenti essenziali per un sistema sanitario efficiente e centrato sul paziente. Come evidenziato da esperti, uno dei maggiori ostacoli alla realizzazione di un sistema sanitario digitale uniforme in Europa sono la disomogeneità³⁸ e il tema del coinvolgimento dei clinici. Le diverse realtà nazionali e regionali spesso si trovano a stadi differenti di maturità digitale, il che rende difficile l'implementazione di soluzioni tecnologiche standardizzate e interoperabili su scala continentale.

HIMSS si propone di colmare questo *gap*, fornendo strumenti pratici e metodologie testate che possono essere adattati alle specifiche esigenze di ogni organizzazione sanitaria. Il valore di questi modelli risiede nella loro capacità di offrire una *roadmap* chiara e basata sull'esperienza globale, permettendo alle organizzazioni di progredire verso livelli più alti di maturità digitale in modo strutturato e consapevole. Un altro aspetto cruciale è la gestione del dato. Nei livelli inferiori del modello EMRAM, viene data grande importanza all'organizzazione e alla qualità dei dati attraverso l'inclusione del *Clinical Data Repository* (CDR) già al secondo livello di maturità. Questo riflette la consapevolezza che la qualità del dato è alla base di qualsiasi processo decisionale efficace e che l'architettura di gestione dei dati deve essere solida e ben strutturata fin dalle prime fasi di digitalizzazione.

Senza una gestione dei dati adeguata, i benefici derivanti dall'adozione di tecnologie digitali rischiano di essere limitati. HIMSS, pertanto, enfatizza l'importanza di una solida architettura dei dati come prerequisito per il successo di qualsiasi iniziativa di digitalizzazione sanitaria. Questo è particolarmente vero per le organizzazioni che aspirano a raggiungere il livello 7 del modello EMRAM, un traguardo che richiede una riorganizzazione profonda del modo in cui i dati sono raccolti, gestiti e utilizzati.

³⁸Gordon, P. (2004). Clinical Data Interchange Standards Consortium (CDISC): Overview and Benefits. *Clinical Trials*, 1(3), 207-213.

Stato dell'arte

La gestione dei big data in ambito sanitario rappresenta una delle sfide principali dell'informatica biomedica. L'efficace centralizzazione e integrazione delle informazioni sono fondamentali per la scoperta di conoscenze e per il progresso nella ricerca sanitaria. Tuttavia, le soluzioni ottimizzate per fornire risorse informative di alta qualità e facili da usare non sono ancora completamente esplorate. Le lacune tra le attuali pratiche di gestione dei dati e le esigenze emergenti evidenziano la necessità di avanzamenti significativi nel campo della gestione dei big data.

Proviamo ora a riepilogare quanto descritto sinora. Le metodologie per l'uso dei dati reali (*Real-World Data*, RWD) variano notevolmente tra le istituzioni sanitarie, il che rende cruciale la standardizzazione per un utilizzo clinico efficace. L'adozione di standard come quelli del *Clinical Data Interchange Standards Consortium* (CDISC) possono facilitare la condivisione e l'analisi dei dati tra diverse istituzioni e paesi. Un approccio pratico è convertire i dati dai *Data Warehouse* (DWH) nel formato *Fast Healthcare Interoperability Resources* (FHIR), che permette di mappare gli standard nazionali a quelli internazionali. Questo approccio non solo migliora l'interoperabilità, ma prepara anche il terreno per una piattaforma di dati che aderisca alle normative nazionali e internazionali, permettendo una condivisione più ampia dei dati.

Un'altra dimensione importante è la qualità dei dati, che deve essere garantita attraverso modelli di maturità e standardizzazione. La qualità dei dati è cruciale non solo per l'uso primario, ma anche per l'uso secondario, che può includere ricerche e decisioni regolatorie. Il progetto Sequoia negli Stati Uniti, per esempio, ha dimostrato che la qualità dei dati raccolti deve essere continuamente testata e verificata per l'uso secondario. Questo processo di verifica continua è essenziale per garantire che i dati siano affidabili e utili.

In Europa, il progetto del *European Health Data Space* (EHDS) rappresenta una risposta a queste sfide.

L'EHDS mira a centralizzare e rendere interoperabili i dati sanitari attraverso l'adozione di standard comuni e tecnologie avanzate. Tuttavia, l'implementazione di questi standard presenta sfide significative, tra cui le difficoltà tecniche e le preoccupazioni legate alla *privacy*. Per affrontare queste sfide, è essenziale un impegno verso la standardizzazione e l'adozione di modelli di maturità che valutino la qualità e l'interoperabilità dei dati, anche alla luce del fatto che le performance degli specialisti possono essere influenzate in maniera significativa grazie alla condivisione di informazione clinica nei network³⁹.

L'*Institute for Healthcare Improvement* (IHE) gioca un ruolo cruciale in questo contesto. IHE sviluppa e promuove standard e profili di interoperabilità che aiutano le istituzioni sanitarie a integrare e scambiare dati in modo più efficace. Attraverso l'adozione di questi standard, le istituzioni possono superare le barriere all'interoperabilità e garantire che i dati siano utili e di alta qualità.

Il passaggio dal dato documentale al dato atomico è una trasformazione fondamentale per migliorare l'efficacia della gestione dei dati. Questa transizione è particolarmente evidente in Italia, che si distingue per la sua ambizione di avanzare nel Piano Nazionale di Ripresa e Resilienza (PNRR) con obiettivi ambiziosi per l'integrazione e l'uso dei dati sanitari. Tuttavia, il progresso deve affrontare le sfide tecniche di interoperabilità e le preoccupazioni per la *privacy*, che possono rallentare l'implementazione dei progetti.

A livello globale, la pandemia di COVID-19 ha rivelato la necessità di migliorare l'interoperabilità dei dati. In molti casi, le risposte alla pandemia hanno mostrato lacune significative nei sistemi di gestione dei dati, con conseguenti difficoltà nella condivisione e nell'analisi dei dati. Questa esperienza ha sottolineato l'importanza di sviluppare sistemi che siano non solo

³⁹Saccavini, Claudio & Troncin, Cristiana & Arkhipova, Daria & Lee, Gwanhoo & Vaia, Giovanni. (2018). The Effect of Network Centrality of Medical Specialists on Their Performance: Evidence from an Italian Health Information Exchange Platform.

interoperabili ma anche orientati alla gestione dei dati in modo globale e integrato.

La gestione dei big data nella salute richiede un'attenzione continua alla qualità e alla standardizzazione dei dati. La standardizzazione, attraverso l'adozione di modelli e standard internazionali come IHE e Data Model, è essenziale per garantire che i dati siano utilizzabili e condivisibili. L'implementazione di standard e modelli di maturità, insieme a un impegno per l'interoperabilità, è fondamentale per affrontare le sfide attuali e per costruire una base solida per la scoperta di conoscenze e per la ricerca sanitaria avanzata.

L'adozione degli standard in Italia

In Italia, la penetrazione e l'adozione degli standard sanitari internazionali come HL7, FHIR e CDA hanno acquisito un'importanza strategica negli ultimi anni, specialmente in relazione allo sviluppo del Fascicolo Sanitario Elettronico (FSE) e delle cartelle cliniche elettroniche regionali.

Il Fascicolo Sanitario Elettronico 2.0 (FSE 2.0), in particolare, rappresenta un importante punto di svolta, volto a migliorare l'interoperabilità e la condivisione dei dati sanitari a livello nazionale e tra le varie Regioni, promuovendo al contempo l'adozione di standard tecnici avanzati. Nonostante l'impegno per l'adozione di questi standard, le Regioni italiane devono affrontare una serie di sfide tecniche e organizzative. Una delle principali difficoltà è l'integrazione dei sistemi legacy con i nuovi standard interoperabili come FHIR e OMOP. Molte infrastrutture sanitarie regionali sono ancora basate su sistemi informatici sviluppati su misura, spesso non facilmente compatibili con i nuovi standard.

L'Agenzia per l'Italia Digitale (AgID) ha dato un forte impulso all'adozione di HL7 e FHIR per garantire un'interoperabilità più efficace tra i sistemi sanitari regionali e nazionali. Il Fascicolo Sanitario Elettronico (FSE) è stato pensato come piattaforma centrale per la raccolta e la condivisione di dati sanitari e clinici tra i diversi attori del sistema sanitario italiano, e uno dei suoi obiettivi principali è l'adozione di standard per la comunicazione tra i vari sistemi informativi regionali. A

tal fine, FHIR è stato indicato come uno degli standard principali per lo scambio di informazioni sanitarie, vista la sua capacità di favorire l'interoperabilità a livello di API e di semplificare l'accesso ai dati.

Nelle Regioni, ci sono state diverse iniziative mirate all'adozione di standard. La Regione Veneto, ad esempio, ha implementato la Cartella Clinica Elettronica (CCE), inserendo FHIR come standard di riferimento per consentire una migliore comunicazione tra i diversi attori sanitari. La Lombardia sta seguendo un percorso simile, con un forte impegno nella digitalizzazione del sistema sanitario attraverso l'uso del Fascicolo Sanitario e della Cartella Clinica Elettronica.

Il ruolo del Fascicolo Sanitario Elettronico 2.0 (FSE 2.0)

L'iniziativa del Fascicolo Sanitario Elettronico 2.0 mira a standardizzare l'infrastruttura tecnologica del sistema sanitario italiano, adottando il profilo HL7 per i dati clinici strutturati e FHIR per favorire l'interoperabilità tra i sistemi regionali. L'obiettivo del FSE 2.0 è non solo migliorare l'accesso ai dati sanitari da parte dei cittadini, ma anche consentire agli operatori sanitari di accedere a informazioni cliniche in tempo reale, facilitando decisioni basate su dati concreti e aggiornati. Il successo del FSE dipende fortemente dalla sua capacità di standardizzare e condividere informazioni attraverso l'intero ecosistema sanitario.

Progetti internazionali e l'adozione di OMOP

Oltre agli standard già consolidati, stiamo assistendo anche a un crescente interesse per l'adozione del modello OMOP per la standardizzazione dei dati clinici e sanitari a fini di ricerca. OMOP è stato sviluppato per supportare la mappatura dei dati sanitari in un *Common Data Model* (CDM), favorendo l'armonizzazione dei dati provenienti da fonti diverse e il loro utilizzo in studi clinici osservazionali e ricerche basate sui big data.

In Italia, diversi centri di ricerca e università hanno aderito all'iniziativa EHDEN (*European Health Data & Evidence Network*), che promuove l'uso del modello OMOP per la condivisione e l'analisi dei dati sanitari in Europa. Partecipare a reti internazionali come EHDEN permette ai centri italiani di migliorare la loro capacità

di analizzare grandi quantità di dati e di contribuire alla ricerca europea su tematiche cliniche e sanitarie di rilevanza strategica. I centri italiani che partecipano a EHDEN stanno lavorando per mappare i propri dati clinici nel formato OMOP, promuovendo così l'integrazione di dati sanitari a livello europeo e aprendo la strada a una ricerca clinica più efficace e basata sull'evidenza.

Uno sguardo all'adozione degli standard nel mondo

A livello mondiale, l'adozione degli standard sanitari come HL7, FHIR, OMOP e CDISC è divenuta un tema centrale per lo sviluppo di sistemi sanitari interoperabili e per la gestione efficiente dei dati clinici. Gli Stati Uniti, il Regno Unito e molte organizzazioni internazionali hanno promosso fortemente l'adozione di questi standard, con lo scopo di migliorare la qualità dei dati, l'integrazione tra sistemi e l'efficacia delle cure.

In termini di legislazione, negli Stati Uniti un momento significativo per l'adozione di standard come HL7 e FHIR è stato il varo del "21st Century Cures Act" nel 2016⁴⁰. Questo atto ha spinto le istituzioni sanitarie a implementare tecnologie in grado di garantire l'interoperabilità dei dati, incoraggiando l'utilizzo di FHIR per l'accesso e lo scambio di dati sanitari tramite API standardizzate. Il 2020 ha rappresentato un punto di svolta con l'introduzione delle regole di interoperabilità da parte dell'*Office of the National Coordinator for Health Information Technology* (ONC), che hanno reso obbligatorio l'uso di FHIR per promuovere il libero scambio di informazioni tra fornitori, pazienti e sistemi tecnologici.

Nel Regno Unito, il *National Health Service* (NHS) ha adottato FHIR per migliorare la comunicazione e l'interoperabilità dei dati tra le varie strutture del sistema sanitario. Un esempio significativo è la spinta verso la digitalizzazione e l'interoperabilità del sistema sanitario britannico con l'introduzione del "*NHS Long Term Plan*"⁴¹

nel 2019, che ha sottolineato l'importanza di adottare standard aperti come FHIR per supportare la continuità delle cure.

In Europa, l'Agenzia Europea per i Medicinali (EMA) ha sostenuto l'uso di standard come CDISC per facilitare la raccolta e l'analisi dei dati clinici nei trial regolatori, migliorando così la coerenza e la qualità dei dati trasmessi a livello europeo.

A livello globale, l'OMS sta promuovendo l'adozione di standard come HL7 per migliorare la gestione delle informazioni sanitarie nelle nazioni in via di sviluppo e potenziare la capacità di risposta a crisi sanitarie, come dimostrato durante la pandemia di COVID-19. Nello specifico, l'OMS e HL7 hanno avviato una collaborazione per promuovere l'adozione globale di standard aperti di interoperabilità nel settore della salute digitale. Questa iniziativa mira a supportare la Strategia Globale sulla Salute Digitale dell'OMS 2020-2025, facilitando lo scambio di dati sanitari tra sistemi. L'accordo prevede lo sviluppo di specifiche globali di interoperabilità, l'accesso ai protocolli dell'OMS e l'uso delle classificazioni OMS all'interno della comunità HL7 FHIR, con risorse tecniche per aiutare i Paesi a implementare questi standard⁴².

Anche la Food and Drug Administration ha adottato standard come CDISC per la raccolta e gestione dei dati degli studi clinici, rafforzando la capacità di analisi e validazione delle nuove terapie. In particolare, FDA sta collaborando con HL7 per integrare lo standard FHIR nel settore farmaceutico, in particolare per migliorare la qualità dei processi chimici, di produzione e di controllo (PQ/CMC). L'obiettivo è modernizzare la gestione dei dati regolatori attraverso l'interoperabilità, semplificando la condivisione di informazioni tra aziende farmaceutiche e agenzie regolatorie. L'adozione del FHIR consentirà di velocizzare le approvazioni, garantire una maggiore trasparenza e migliorare l'efficienza dei processi di regolamentazione⁴³.

⁴⁰<https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act>

⁴¹<https://www.longtermplan.nhs.uk/>

⁴²<https://www.who.int/news/item/03-07-2023-who-and-hl7-collaborate-to-support-adoption-of-open-interoperability-standards>

⁴³<https://www.fda.gov/industry/pharmaceutical-quality-chemistry-manufacturing-controls-pqcmc/pqcmc-and-hl7-fhir>

Progetti Comunità Europea (es: Darwin EU)

EHDEN (*European Health Data & Evidence Network*)⁴⁴ e DARWIN EU (*Data Analysis and Real-World Interrogation Network*) sono due notevoli iniziative volte alla condivisione dei dati sanitari e la ricerca clinica in Europa, promosse con il coinvolgimento di vari *stakeholder*, tra cui l'Agenzia Europea dei Medicinali, istituzioni sanitarie, e partner industriali.

EHDEN mira a creare un ecosistema europeo di dati sanitari armonizzati, facilitando l'interoperabilità attraverso l'adozione dello standard OMOP CDM. Questo progetto è stato lanciato nel contesto dell'iniziativa IMI (*Innovative Medicines Initiative*) per affrontare la frammentazione dei dati sanitari e creare un'infrastruttura che consenta l'analisi multi-paese e collaborativa. Il consorzio EHDEN lavora con ospedali, università, aziende farmaceutiche e organizzazioni sanitarie, riducendo i costi e il tempo di preparazione dei dati per la ricerca clinica. Ad oggi, sono stati finanziati numerosi progetti per standardizzare i dati sanitari in oltre 100 istituzioni in tutta Europa, promuovendo la generazione di prove di alta qualità per migliorare la cura del paziente.

DARWIN EU è un'iniziativa guidata dall'EMA e nata per fornire prove basate su dati reali (RWD) al fine di supportare decisioni regolatorie rapide e basate sull'evidenza. Lanciato nel 2022, DARWIN EU⁴⁵ si concentra su analisi in tempo reale dei dati clinici e sanitari provenienti da vari sistemi sanitari europei, creando un'infrastruttura che collega fonti di dati sanitari per fornire informazioni utili su sicurezza, efficacia e qualità dei farmaci. Questa rete è parte integrante della strategia dell'EMA per l'utilizzo dei RWD nella regolamentazione, aumentando la velocità e l'affidabilità della ricerca osservazionale e supportando decisioni informate sull'approvazione e il monitoraggio di nuovi trattamenti.

Entrambe le iniziative si basano su una rete di partner scientifici e industriali, con il contributo di varie agenzie governative e organismi di regolamentazione, come le autorità nazionali per la sanità pubblica.

L'obiettivo comune è superare le barriere legate alla frammentazione dei dati e consentire alle organizzazioni sanitarie e regolatorie di accedere a dati armonizzati e di alta qualità. I primi risultati di EHDEN e DARWIN EU indicano una maggiore efficienza nella raccolta e analisi dei dati sanitari, con effetti positivi nella ricerca clinica e nella regolamentazione dei medicinali.

EMA, che ha un ruolo centrale nel coordinamento di DARWIN EU, sfrutta questa infrastruttura per ottenere dati reali a supporto delle sue decisioni e migliorare i processi di valutazione dei farmaci, specialmente nelle fasi post-marketing. Collaborando con i partner di EHDEN, l'EMA è in grado di utilizzare dati armonizzati per condurre studi osservazionali su larga scala, migliorando la sicurezza dei pazienti e l'efficacia delle terapie.

European Health Data Space

Come visto da quanto descritto in precedenza, la digitalizzazione è essenziale per il futuro dell'assistenza sanitaria ai cittadini e per sostenere l'innovazione nell'industria medica dell'UE. Dovrebbe ormai esser chiaro come i dati costituiscano una componente indispensabile del mondo di oggi — se utilizzati responsabilmente e nel pieno rispetto dei diritti fondamentali — e come possano apportare enormi benefici a tutti gli aspetti della vita quotidiana, compresa la salute.

I sistemi sanitari degli Stati Membri generano, trattano e conservano già un'ingente quantità di dati sanitari che forniscono ai servizi sanitari e ai ricercatori informazioni potenzialmente preziose.

Si stima che il riutilizzo dei dati sanitari valga circa 25-30 miliardi di euro all'anno. Secondo le previsioni, tale cifra dovrebbe raggiungere circa 50 miliardi di euro nei prossimi 10 anni. Tuttavia è ancora spesso difficile per i cittadini accedere per via elettronica ai propri dati sanitari e per i ricercatori utilizzare tali dati al fine di migliorare le diagnosi e le terapie. Tale difficoltà nell'accesso crea pertanto ostacoli alla prestazione di assistenza sanitaria e all'innovazione in questo settore.

⁴⁴<https://www.ehden.eu/>

⁴⁵<https://www.darwin-eu.org/>

Il 24 aprile 2024, dopo l'accordo politico raggiunto con il Consiglio dell'Unione europea, il Parlamento europeo ha approvato con 445 voti favorevoli, 142 contrari e 39 astensioni la creazione di uno Spazio europeo dei dati sanitari, l'iniziativa dell'UE finalizzata alla costruzione di un ecosistema digitale per la condivisione sicura dei dati sanitari, che costituisce una parte integrante della transizione digitale dell'assistenza sanitaria nell'UE e della strategia europea in materia di dati, nonché un pilastro fondamentale della costruzione di un'Unione europea della salute. L'obiettivo primario dell'EHDS è quello di superare le sfide legate alla *privacy*, alla sicurezza e all'interoperabilità per creare uno spazio comune in cui i dati sanitari possano fluire liberamente e in modo sicuro.

Lo *European Health Data Space* (EHDS) rappresenta una delle più ambiziose iniziative dell'Unione europea per trasformare la gestione e l'uso dei dati sanitari attraverso un'integrazione armonizzata e centralizzata a livello continentale. Questa proposta di regolamento, attualmente in fase di negoziazione presso il Consiglio dell'UE, è progettata per facilitare un accesso più fluido e sicuro ai dati sanitari e per promuovere una serie di benefici significativi sia per i cittadini che per i sistemi sanitari e di ricerca in tutta Europa.

L'EHDS mira a disciplinare sia l'uso primario che secondario dei dati sanitari, cercando di armonizzare e migliorare la loro gestione all'interno dell'UE. Il regolamento affronta una serie di tematiche cruciali per il successo dell'iniziativa, inclusi la compatibilità con le normative esistenti, come il Regolamento Generale sulla Protezione dei Dati (GDPR), il *Data Act* e il *Data Governance Act*, e aspetti fondamentali come la qualità dei dati, l'interoperabilità, la *governance* e la cybersicurezza.

L'uso primario dei dati, trattato nel Capo II del regolamento, si concentra sulla facilitazione dell'accesso ai dati sanitari per i pazienti e per i professionisti della salute, anche al di fuori del proprio Paese. Questo aspetto rappresenta un'evoluzione rispetto alla direttiva esistente sull'assistenza sanitaria transfrontaliera, migliorando l'accesso e il controllo dei dati da parte degli individui e garantendo che i dati possano essere utilizzati per esigenze di cura in tutta l'Unione Europea.

L'uso secondario dei dati, descritto nel Capo IV, si riferisce invece all'impiego dei dati sanitari per finalità diverse dalla cura diretta dei pazienti, come la ricerca scientifica, la politica sanitaria e la salute pubblica. Questo uso secondario ha il potenziale per generare benefici significativi attraverso la scoperta di nuove malattie, lo sviluppo di trattamenti terapeutici, e il miglioramento dei sistemi sanitari. Inoltre, può sostenere il mercato dell'*health-tech* e facilitare l'addestramento di sistemi di intelligenza artificiale, contribuendo così alla crescita economica e alla competitività dell'Unione Europea.

Il regolamento EHDS offre opportunità senza precedenti per la medicina di precisione, predittiva e preventiva. Ad esempio, l'analisi dei dati genomici e ambientali può migliorare la comprensione delle malattie e delle loro interazioni con l'ambiente. Ciò consente studi approfonditi sulla familiarità o ereditarietà di certe condizioni patologiche, migliorando la capacità di misurare la variazione fenotipica e comprendere meglio le manifestazioni delle malattie in diverse popolazioni.

Per la medicina predittiva, l'EHDS permetterà di realizzare studi di associazione su ampi set di dati, facilitando la scoperta di varianti genetiche e fattori di rischio associati a malattie, anche sconosciuti fino ad ora. Questo potrà supportare la creazione di sistemi di sorveglianza in tempo reale per prevedere e gestire emergenze sanitarie, come focolai pandemici, e realizzare piani di intervento rapidi ed efficaci.

In termini di medicina preventiva, l'EHDS potrà potenziare gli studi sulla salute della popolazione, permettendo di valutare l'efficacia degli interventi di salute pubblica e di ridurre il carico delle malattie attraverso strategie *data-driven*. La capacità di monitorare e analizzare i dati sanitari su larga scala offrirà opportunità per migliorare la qualità dell'assistenza sanitaria, ridurre errori medici e gestire problemi complessi come le infezioni correlate all'assistenza e la resistenza antimicrobica.

Un aspetto rilevante dell'EHDS è la standardizzazione e interoperabilità dei dati. La creazione di uno spazio

europeo per i dati sanitari richiede che i dati siano compatibili e facilmente scambiabili tra diversi sistemi e Paesi. Per raggiungere questo obiettivo, il regolamento prevede l'adozione di standard comuni e infrastrutture digitali condivise che facilitano la *governance* dei dati e garantiscono un livello elevato di cybersicurezza.

Nonostante i numerosi benefici potenziali, la creazione e l'implementazione dell'EHDS presentano delle sfide significative. La protezione della *privacy* dei dati e la loro sicurezza rimangono priorità assolute. È fondamentale che i sistemi implementati rispettino le normative vigenti e siano progettati per prevenire violazioni e abusi.

Inoltre, garantire la qualità dei dati è essenziale per massimizzare i benefici dell'EHDS. I dati devono essere accurati, completi e aggiornati per essere utili in contesti di ricerca e cura. Le tecnologie e le pratiche per assicurare la qualità dei dati devono essere integrate nei processi di raccolta e gestione dei dati.

Per quanto riguarda i prossimi passi, il testo del regolamento per la creazione dell'EHDS dovrà essere formalmente approvato dal Consiglio europeo ed entrerà in vigore venti giorni dopo la pubblicazione nella Gazzetta Ufficiale dell'UE. Il regolamento sarà applicabile solo a partire dai due anni successivi alla pubblicazione, quindi non prima dell'estate 2026. Sono previste tuttavia alcune eccezioni, tra cui quelle relative all'uso primario e secondario dei dati, che si applicheranno da quattro a sei anni dopo la pubblicazione, per cui occorrerà attendere rispettivamente il 2028 e il 2030. Per ovviare a tale differimento di efficacia, gli Stati Membri potranno esigere che i dati sanitari siano resi disponibili in formato elettronico prima dell'applicazione del regolamento.

Vi è inoltre da considerare che gli Stati Membri mostrano diversi livelli di maturità digitale nel settore sanitario; l'EHDS mira a standardizzare le pratiche di sanità digitale in tutta l'UE entro il 2028, con ulteriori categorie di dati da raggiungere entro il 2030. Questo lasso di tempo consente una preparazione adeguata per gli Stati membri e i fornitori di assistenza sanitaria.

Conclusioni

L'adozione di standard per l'organizzazione dei dati sanitari è di grande importanza per superare le sfide poste dalla crescente complessità e dalla vasta quantità di informazioni generate nel contesto della ricerca e della pratica clinica. Questi modelli consentono ai clinici e ai ricercatori di aggregare e analizzare dati provenienti da fonti disparate, mantenendo la consistenza e l'integrità necessarie per trarre conclusioni affidabili.

Nella pratica clinica, l'utilizzo di modelli di dati potrà aiutare gli operatori sanitari a monitorare gli esiti dei pazienti, a migliorare la qualità dell'assistenza e a prendere decisioni terapeutiche informate. Ad esempio, i dati strutturati possono essere utilizzati per identificare *pattern* di malattie, per valutare l'efficacia di interventi sanitari o per monitorare la sicurezza dei farmaci in uso clinico.

L'interoperabilità dei dati è un altro aspetto fondamentale che viene potenziato dall'adozione di standard: la capacità di scambiare dati tra sistemi diversi senza perdita di informazioni o significato è essenziale per la continuità delle cure, per la ricerca multicentrica e per le collaborazioni internazionali. Inoltre, l'interoperabilità supporta lo sviluppo di sistemi di supporto decisionale clinico che si basano su algoritmi in grado di processare grandi volumi di dati strutturati per fornire raccomandazioni basate sull'evidenza ai clinici.

Un altro vantaggio significativo dell'uso di Data Model è legato alla *governance* dei dati. Con la crescente attenzione alla *privacy* dei pazienti e alla sicurezza dei dati, avere un sistema di dati ben definito e conforme agli standard internazionali è indispensabile per garantire la protezione delle informazioni sensibili. OHDSI e CDISC forniscono un *framework* che aiuta le organizzazioni sanitarie a soddisfare i requisiti normativi, come il *General Data Protection Regulation* (GDPR) in Europa, e a instaurare fiducia nei pazienti riguardo al modo in cui i loro dati vengono gestiti e utilizzati.

L'integrazione di modelli di dati strutturati richiede un investimento significativo in termini di formazione, tecnologia e cambiamento organizzativo. Tuttavia, i benefici a lungo termine superano di gran lunga i costi iniziali. La formazione del personale sanitario e dei ricercatori sull'uso degli standard di dati è essenziale per massimizzare l'utilizzo di tali risorse e per assicurare che i dati siano utilizzati in modo etico e responsabile.

Dal punto di vista comunitario, l'EHDS rappresenta una svolta significativa nella gestione dei dati sanitari in Europa. Offrendo un ambiente regolamentato e armonizzato per l'accesso e l'uso dei dati, l'Unione Europea può promuovere l'innovazione, migliorare

la salute pubblica e sostenere lo sviluppo di nuove tecnologie. Tuttavia, per realizzare appieno il potenziale dell'EHDS, è essenziale affrontare le sfide legate alla protezione dei dati, alla qualità e all'interoperabilità. La creazione di uno spazio europeo dei dati sanitari non solo migliorerà la qualità della cura e la ricerca scientifica, ma potrà anche posizionare l'Unione Europea come leader globale nell'uso dei dati per la salute. Per raggiungere questi obiettivi, è fondamentale una collaborazione continua tra Stati Membri, istituzioni europee e *stakeholders* del settore sanitario e tecnologico.



Bibliografia

ACRONIMO	DEFINIZIONE
CDA	Clinical Document Architecture — standard HL7 per la struttura dei documenti clinici elettronici
CCE	Cartella Clinica Elettronica — sistema di gestione elettronica dei dati clinici regionali in Italia
CDISC	Clinical Data Interchange Standards Consortium — standard per lo scambio di dati clinici
CDM	Common Data Model — Modello comune di dati per armonizzare i dati sanitari
CTMS	Clinical Trial Management System — sistema per la gestione degli studi clinici
DARWIN EU	Data Analysis and Real-World Interrogation Network — rete europea per l'analisi di dati sanitari real-world
DICOM	Digital Imaging and Communications in Medicine — standard per la gestione di immagini mediche
EHDEN	European Health Data & Evidence Network — progetto per armonizzare i dati sanitari in Europa
EHDS	European Health Data Space — spazio europeo per la condivisione e l'accesso ai dati sanitari
EMA	European Medicines Agency — Agenzia Europea dei Medicinali
EMRAM	Electronic Medical Record Adoption Model — modello di adozione della cartella clinica elettronica
EMR	Electronic Medical Record — cartella clinica elettronica
EUDRAMED	European Databank on Medical Devices — banca dati europea sui dispositivi medici
ePRO	Electronic Patient Reported Outcomes — strumenti digitali per la raccolta di dati direttamente dai pazienti
EU/UE	European Union/Unione Europea
FDA	Food and Drug Administration — Agenzia per la regolamentazione di cibo, farmaci e dispositivi medici negli Stati Uniti
FHIR	Fast Healthcare Interoperability Resources — standard per l'interoperabilità dei dati sanitari
FSE 2.0	Fascicolo Sanitario Elettronico 2.0 — iniziativa italiana per la digitalizzazione dei dati sanitari regionali
GDPR	General Data Protection Regulation — Regolamento generale sulla protezione dei dati
HDRP	Health Data Research Platform — piattaforma per la gestione e l'analisi di dati sanitari
HIMSS	Healthcare Information and Management Systems Society — organizzazione globale che promuove la sanità digitale
HL7	Health Level 7 — standard internazionale per lo scambio di dati sanitari elettronici

Bibliografia

ACRONIMO	DEFINIZIONE
IVD	In Vitro Diagnostic — diagnostico in vitro, riferito ai dispositivi medici regolamentati
LOINC	Logical Observation Identifiers Names and Codes — standard per la codifica di osservazioni mediche
MDR	Medical Device Regulation — regolamento UE per la sicurezza e le prestazioni dei dispositivi medici
NHS	National Health Service — servizio sanitario nazionale del Regno Unito
OMOP	Observational Medical Outcomes Partnership — modello comune di dati per la standardizzazione dei dati osservazionali
OMS	Organizzazione Mondiale della Sanità — agenzia dell'ONU per la salute pubblica
PQ/CMC	Pharmaceutical Quality/Chemistry Manufacturing and Controls — gestione della qualità farmaceutica legata alla produzione

Bibliografia

1. Bryan, S., Weatherburn, G. C., Watkins, J. R., & Buxton, M. J. (1999). The Digital Imaging and Communications in Medicine (DICOM) Standard: A Brief Overview and Its Impact on Radiological Practice. *Clinical Radiology*, 54(9), 640-647.
2. Butler, M., & McCreedy, E. (2018). Use of ICD-10 and ICD-11 Classifications in Clinical and Health Service Delivery Applications. *Health Information Management Journal*, 47(1), 29-33.
3. Codd, E. F. (1970). A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, 13(6), 377-387.
4. Devine, J., & Daugherty, J. (2010). The Development and Implementation of CDISC Standards: A Historical Perspective. *Drug Information Journal*, 44(1), 15-21.
5. Dolin, R. H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F. M., Biron, P. V., & Shabo, A. (2006). HL7 Clinical Document Architecture, Release 2. *Journal of the American Medical Informatics Association*, 13(1), 30-39.
6. Elkin, P. L., & Brown, S. H. (2008). Historical Perspective on SNOMED and Its Impact on Clinical Data Standards. *Studies in Health Technology and Informatics*, 136, 19-23.
7. Elmasri, R., & Navathe, S. B. (2010). *Fundamentals of Database Systems (6th Edition)*. Addison-Wesley.
8. English, L. P. (2015). *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*. Wiley.
9. European Commission. (2021). *Data Governance Act (DGA) - Proposal for a Regulation*.
10. European Commission. (2024). *European Health Data Space (EHDS) - Proposal for a Regulation*.
11. European Commission. (2024). *Factsheet: European Health Data Space (EHDS)*.
12. European Parliament and Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)*.
13. Gordon, P. (2004). Clinical Data Interchange Standards Consortium (CDISC): Overview and Benefits. *Clinical Trials*, 1(3), 207-213.
14. Grannis, S. J., & Overhage, J. M. (2006). HL7 Version 3: An Overview and Future Directions. *Journal of Biomedical Informatics*, 39(2), 190-195.
15. Haendel MA, Chute CG, Robinson PN. Classification, Ontology, and Precision Medicine. *N Engl J Med*. 2018 Oct 11;379(15):1452-1462. doi: 10.1056/NEJMra1615014. PMID: 30304648; PMCID: PMC6503847.
16. Health Level Seven International (2020). *HL7 Standards*. HL7 International.
17. Hersh, W., & Hickam, D. (2008). Health Level Seven (HL7) Standards: A Brief History and Overview. *Journal of the American Medical Informatics Association*, 15(3), 317-325

Bibliografia

18. Hripcsak, G., & Albers, D. J. (2013). Next-generation phenotyping of electronic health records. *Journal of Biomedical Informatics*, 46(5), 792-799.
19. Huang, H. K. (2010). *PACS and Imaging Informatics: Basic Principles and Applications* (2nd Edition). Wiley-Blackwell.
20. Huff, S. M., Rocha, R. A., McDonald, C. J., De Moor, G. J., Fiers, T., Bidgood, W. D., ... & Dolin, R. H. (1998). Development of the Logical Observation Identifier Names and Codes (LOINC) Vocabulary. *Journal of the American Medical Informatics Association*, 5(3), 276-292.
21. ISO/TC 215 (2017). *ISO 13606: Health Informatics—Electronic Health Record Communication*. International Organization for Standardization.
22. Kok, I., et al. (2024). The European Health Data Space: A Major Step Towards a Digital Health Ecosystem. *European Journal of Public Health*. doi:10.1093/eurpub/ckaa257
23. La gestione dei dati della ricerca | ScuolaNormaleSuperiore. (n.d.). <https://www.sns.it/it/guida/la-gestione-dei-dati-della-ricerca>
24. McDonald, C. J. (1996). The LOINC Project: A Universal Code System for Tests, Observations, and Measurements. *Clinical Chemistry*, 42(5), 839-848.
25. McDonald, C. J., Huff, S. M., Suico, J. G., Hill, G., Leavelle, D., Aller, R., ... & Maloney, P. C. (2003). LOINC, a Universal Standard for Identifying Laboratory Observations: A 5-Year Update. *Clinical Chemistry*, 49(4), 624-633.
26. Mildenerger, P., Eichelberg, M., & Martin, E. (2002). Introduction to the DICOM Standard. *European Radiology*, 12(4), 920-927.
27. Moriyama, I. M., Loy, R. M., & Robb-Smith, A. H. T. (2011). *History of the Statistical Classification of Diseases and Causes of Death*. World Health Organization.
28. Murphy, S. N., & Mendis, M. (2012). FHIR: An Agile Standards Development Methodology for Healthcare Interoperability. *AMIA Annual Symposium Proceedings*, 2012, 139-143.
29. Papagni, A. (2022, September 7). Perché possiamo contare sui (con) i dati IQVIA. *Informatore*. <https://informatore.it/editoriali/lavoro/perche-possiamo-contare-dati-iqvia/>
30. Raghupathi, W., & Raghupathi, V. (2014). Big Data Analytics in Healthcare: Promise and Potential. *Health Information Science and Systems*, 2(1), 1-10.
31. Redman, T. C. (2016). *Data Driven: Creating a Data Culture*. Harvard Business Review Press.
32. Restifo, Nicola & Locatelli, Paolo & Montefusco, Vittorio & Sini, Elena & Facchini, Roberta & Torresani, Michele. (2013). *The Role of ICT and Mobile Health to Improve Clinical Process Management. An Overview on the Therapy Management Process and a Real Case*.

Bibliografia

33. Saccavini, Claudio & Troncin, Cristiana & Arkhipova, Daria & Lee, Gwanhoo & Vaia, Giovanni. (2018). The Effect of Network Centrality of Medical Specialists on Their Performance: Evidence from an Italian Health Information Exchange Platform.
34. Spackman, K. A., & Campbell, K. E. (2004). The Evolution of SNOMED: From SNOMED RT to SNOMED CT. *International Journal of Medical Informatics*, 73(4), 277-287.
35. Stang, P. E., & Ryan, P. B. (2013). Common Data Model for Observational Research: OMOP's Contributions to Standardizing Data. *Health Information Management Journal*, 42(3), 35-42.
36. Ryan, P. B., & Schuemie, M. J. (2014). The Observational Health Data Sciences and Informatics (OHDSI) initiative: Collaborative research to improve the quality of observational data. *Journal of Biomedical Informatics*, 52, 105-11
37. Wang, X., & Reddy, S. (2016). Clinical Data Interchange Standards Consortium (CDISC): A Historical and Technical Review. *Journal of Biomedical Informatics*, 59, 319-327.
38. World Health Organization (2019). ICD-11 for Mortality and Morbidity Statistics (ICD-11 MMS). World Health Organization.
39. Wright, D., & De Hert, P. (2018). *Enforcing GDPR Compliance: The Role of the Supervisory Authorities*. Springer.

Autori



MARCO BECHINI
Healthcare Integration
Director, Consulting,
IQVIA

Marco Bechini è entrato in IQVIA nel 2022 con il ruolo di Healthcare Integration Director nell'unità RW. Dal 2019 al 2021 è stato responsabile della Business Unit Interoperability and DataPlatform Solutions di InterSystems Italia.

Nei precedenti nove anni di permanenza in Dedalus S.p.A ha avuto la responsabilità del business della Interoperabilità e del reparto R&D. E' stato CEO di Codices S.r.l. (Soluzioni per Interoperabilità), di Dedalus Clinic S.r.l. (IT clinical solutions), di HIM.IT S.r.l. (Soluzioni per la telemedicina) e di Radiosity S.r.l. (clinical imaging).

Dal 1997 al 2006 ha lavorato in Siemens S.p.A. Medical Solutions in ambito delle soluzioni di Imaging Digitale e dal 2006 to 2009 è stato Business Unit Manager (Healthcare Informatics) in FUJIFILM Italia Spa E' stato membro di diverse iniziative di standardizzazione come IHE (Integrated Healthcare Enterprise) e HL7.

Autore di diverse pubblicazione scientifiche, è stato coinvolto in progetti di ricerca della comunità europea, sempre nell'ambito dell'e-health.

Ha conseguito la laurea in Giurisprudenza ed il Master ManSan (Management Aziende Sanitarie — Università di Pisa). E' inoltre DPO, Esperto Qualificato di Radioprotezione e Tecnico di Radiologia.



SARA CAZZANIGA
Institutional Engagement and
Scientific Partnership Director,
IQVIA

Sara Cazzaniga è una professionista con oltre 20 anni di esperienza nel settore farmaceutico.

Ha conseguito la Laurea in Biologia presso l'Università

degli Studi di Milano e la specializzazione in patologia Clinica presso la facoltà di medicina.

Attualmente ricopre il ruolo di Direttore delle relazioni istituzionali e delle partnership scientifiche presso IQVIA. In precedenza, ha lavorato nella ricerca clinica e nel Medical Affairs presso Janssen Cilag.

Sara è membro di diversi gruppi di lavoro presso società scientifiche e autrice di pubblicazioni scientifiche.



CHRISTIAN FEDELI
Healthcare Solution
Architect Manager,
IQVIA

Christian Fedeli è entrato in IQVIA nel 2022 con il ruolo di Solution Architect nel team di Healthcare & Implementation Science.

Dal 2020 al 2022 è stato responsabile del Gruppo Product and Professional Services delle Soluzioni per la Medicina Territoriale di Dedalus Italia S.p.A. In precedenza, ha ricoperto il ruolo di Product Manager per cinque anni del Primary Care System per Dedalus Group S.p.A. Per il medesimo Gruppo, ha ricoperto nei precedenti sei anni I ruoli di Project Manager in progetti di entità nazionale, regionale e locale, volti all'implementazione di soluzioni e tecnologie per la digitalizzazione dei processi in ambito ospedaliero, diagnostico e territoriale; è stato inoltre responsabile del team di Operation di Sicilia Sistemi Tecnologie s.r.l. (Gruppo Dedalus), e responsabile dei Servizi di Presales dell'Area Sud.

Ha inoltre accumulato esperienza sul campo, lavorando come Tecnico di Presidio per la digitalizzazione dei processi ospedalieri per conto di TechSystem s.r.l. dal 2007 al 2009. Da anni partecipa in qualità di docente o relatore ad eventi e corsi ECM.

Laureato in Scienze della Comunicazione Istituzionale e di Impresa, ha inoltre certificazioni nell'ambito del Project Management (AgilePM e Prince2) e del Service Management (ITIL).

CONTACT US

IQVIA Solutions Italy

Via Fabio Filzi 29

20124 Milano

CentroStudi_Italia@iqvia.com

iqvia.com