



I big data a supporto della ricerca e della cura

Gemelli Generator: il progetto della Fondazione Policlinico Universitario A. Gemelli IRCCS che trasforma i dati in valore

Andrea Damiani, Responsabile Gemelli Generator RWD: “Nel settore health, l’innovazione funziona solo quando i benefici tornano al paziente. Quello è il fine. Nulla deve rimanere puramente accademico o commerciale, ma ogni progetto deve invece avere un impatto sulla comunità”.

L’azienda

La Fondazione Policlinico Universitario Agostino Gemelli IRCCS è un policlinico universitario di Roma e il più grande ospedale d’Italia, con oltre 1500 posti letto. Inaugurato il 10 luglio 1964, dal 2015 ha assunto la forma giuridica di fondazione. La sua mission è di offrire ai pazienti umanità, eccellenza e alta specializzazione nelle cure, favorendo l’innovazione della medicina e formando i professionisti del futuro. Nel 2020 ha avviato il progetto Gemelli Generator, con l’obiettivo di dotarsi di strumenti in grado di raccogliere, conservare e analizzare l’ingente quantità di dati generati ogni giorno all’interno del Policlinico, rendendoli comprensibili e fruibili, nel completo rispetto della privacy, dell’integrità dei dati e della proprietà intellettuale, a beneficio di tutti. Gemelli Generator si compone di quattro facilities.

L’idea in breve

La facility Real World Data (RWD) offre servizi di intelligenza artificiale e big data processing per l’integrazione dei dati clinici e lo sviluppo di modelli predittivi a supporto di diagnosi, terapie e processi clinici. L’obiettivo è la valorizzazione clinica e scientifica dei dati della Fondazione Policlinico Universitario A. Gemelli IRCCS, attraverso servizi avanzati di estrazione, raccolta e analisi dati.

L’innovazione con Sas

Gemelli Generator offre uno spazio virtuale dove il committente interessato può fare uso, tramite



algoritmi concordati, di tabelle estratte da un Data Warehouse senza tuttavia vedere tali dati, a tutela della privacy e della proprietà intellettuale. Ciò che il committente riceve è l'output del modello, certificato per data e contenuto prima di essere reso disponibile per il download. Gemelli Generator usa la tecnologia sandbox per dare accesso a terze parti, in un ambiente controllato e sicuro, a dati che non potranno mai essere visti. L'idea nasce da un'intuizione del prof. Vincenzo Valentini, Direttore del Dipartimento di Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia del Policlinico Gemelli, di poter estrarre conoscenza a partire dai dati reali (Real World Data) raccolti durante la pratica clinica all'interno del Policlinico, a supporto della medicina Knowledge-Based ed Evidence-Based. Alle tradizionali modalità di raccolta dati per l'effettuazione di studi clinici, si aggiunge una nuova, grande opportunità: l'uso dei dati già a disposizione dell'ospedale perché raccolti durante l'ordinaria pratica clinica. Gemelli Generator ha sviluppato una versione di studio con strumenti open source, insieme ai dottorandi di ricerca dell'Università Cattolica del Sacro Cuore, e un prototipo, funzionale e di facile utilizzo, realizzato con la società di software Sas, già partner tecnologico del Policlinico Gemelli. Il team sta ora lavorando all'interfaccia utente e sta individuando terze parti per effettuare test sul prototipo.

Data products nel rispetto della privacy e della proprietà intellettuale

Emersa l'esigenza di valorizzare i dati sia per i ricercatori interni sia per terze parti esterne (come ricercatori, case farmaceutiche, assicurazioni, committenti istituzionali), bisognava però garantire la protezione dei dati e la privacy dei pazienti, in conformità con il GDPR, e proteggere la proprietà intellettuale dei dati (che derivano da ciò che è stato elaborato all'interno del Policlinico, come le refertazioni). La soluzione scelta risponde a entrambi i requisiti: all'utente non si offrono dati ma Data Products. Grazie all'ambiente protetto sandbox, il committente può ottenere statistiche aggregate e tabelle di dati (che non vedrà mai), può accedere a una collezione di algoritmi e scegliere quale lanciare su quelle tabelle. Tutto in base ad accordi precedenti con il Policlinico. Ogni documento è certificato per data e contenuto. L'altro aspetto importante che sarà oggetto di approfondimento per prossima implementazione è la certificazione dell'origine del dato, con tracciamento puntuale fino alla creazione del modello in output.

Benefici

Grazie a ricerca e tecnologia, la Fondazione Policlinico Universitario Agostino Gemelli IRCCS è riuscita a dotarsi di strumenti che permettono di raccogliere, conservare, elaborare e utilizzare un'ingente mole di informazioni, per rispondere alle necessità di cura della popolazione e trovare soluzioni cliniche adeguate. Gemelli Generator innesca una catena del valore, con il fine ultimo di portare benefici alla comunità dei pazienti. Valorizza l'enorme quantità di dati raccolti dal Policlinico. Permette di creare modelli utili per la ricerca e per la clinica, rendendo fruibili dai professionisti numerose informazioni altrimenti non disponibili. Risponde anche alla necessità del settore farmaceutico di velocizzare i processi. Sulla base dei risultati medico-scientifici prodotti, i



Milano, 11 – 12 ottobre 2022



pazienti potranno beneficiare di un'assistenza sanitaria più efficiente e di cure sempre più personalizzate.

Nome azienda: Fondazione Policlinico Universitario Agostino Gemelli IRCCS
Settore: Sanità
Sito web: www.policlinicogemelli.it