PuntoSicuro

ARTICOLO DI PUNTOSICURO

Anno 26 - numero 5536 di Lunedì 15 gennaio 2024

Big data analysis e intelligenza artificiale per migliorare la prevenzione

Un factsheet Inail su big data analysis e intelligenza artificiale presenta alcuni strumenti a supporto dei ricercatori nella gestione della moltitudine dei dati ? omici. Scienze omiche, nuove tecniche, BiTdata e intelligenza artificiale.

Roma, 15 Gen ? Nel 2022 il factsheet Inail " Applicazioni biotecnologiche. Gli aspetti normativi e i progetti Inail" aveva segnalato vari progetti di ricerca con riferimento ai temi connessi alla biotecnologia (utilizzo di agenti biologici, cellule o loro prodotti nei processi produttivi) e alla tutela della salute e sicurezza. Ed è in relazione a questi progetti che l'Inail ha prodotto una nuova scheda informativa (factsheet) che presenta le nuove **tecniche NGS** (*Next Generation Sequencing*) che consentono ai ricercatori di "studiare ed elaborare una grande quantità di dati depositati in repository pubblici che sono potenzialmente ricchi di informazioni su eventi cellulari".

Questi importanti dati possono essere utilizzati per "dare risposte a quesiti biologici non ancora studiati, consentendo ai ricercatori di fare nuove scoperte, estraendo e rianalizzando, con nuove domande biologiche, set di dati pubblici".

A fornire queste informazioni è la nuova scheda del Dipartimento innovazioni tecnologiche e sicurezza degli impianti, prodotti e insediamenti antropici (Dit) dell'Inail, intitolata "**Big data analysis e intelligenza artificiale: strumenti Inail a supporto dei ricercatori nella gestione della moltitudine dei dati - omici**" e a cura di E. Sturchio, P. Boccia, M. Zanellato, S. Gioiosa e L. Di Piramo.

La scheda si sofferma su alcuni strumenti Inail a supporto dei ricercatori nella gestione della moltitudine dei "dati ? omici", la banca dati molecolare BiTdata e l'applicazione di Intelligenza Artificiale (IA), di cui vengono schematizzate le potenzialità e alcune elaborazioni significative. Ricordiamo che con scienze "**omiche**" si intendono quelle discipline che utilizzano tecnologie di analisi per la produzione di informazioni/dati utili per la descrizione e l'interpretazione del sistema biologico studiato (ad esempio la genomica è la disciplina omica che studia le sequenze dell'intero genoma e le informazioni contenute al suo interno).

L'articolo di presentazione della scheda si sofferma sui seguenti argomenti:

- Le scienze omiche e il sequenziamento genetico di nuova generazione
- Big data analysis e banca dati molecolare
- Big data analysis e intelligenza artificiale

Pubblicità
<#? QUI-PUBBLICITA-MIM-[CSA139.D] ?#>

Le scienze omiche e il sequenziamento genetico di nuova generazione

La scheda sottolinea che lo sviluppo delle nuove tecnologie correlate allo studio delle scienze "omiche" ha portato ad una "**rivoluzione nel modo di fare ricerca**": si è passati "da un approccio basato su ipotesi ad un approccio basato sui dati, che a volte possono rispondere molto più velocemente a quesiti biologici ancora aperti".

Si indica, a questo proposito, che con il termine **bioinformatica** (Paulien Hogeweg, Ben Hesper, 1970) si intende descrivere '*lo studio dei processi informatici nei sistemi biotici*'. E la bioinformatica è "un campo interdisciplinare che sviluppa metodi e strumenti software per estrarre conoscenza dal dato biologico".

Si indica poi che per **sequenziamento genetico di nuova generazione** (**NGS**), noto anche come sequenziamento ad alto rendimento, "si intende l'insieme delle tecnologie di sequenziamento degli acidi nucleici che hanno in comune la capacità di sequenziare, in parallelo, milioni di frammenti di DNA. È quindi il termine generico utilizzato per descrivere una serie di diverse tecnologie, che consentono di sequenziare il DNA e l'RNA ovvero il sequenziamento del genoma, il risequenziamento del genoma, il profiling del trascrittoma (RNA-Seq), le interazioni DNA-proteina (sequenziamento del ChIP) e la caratterizzazione dell'epigenoma".

In particolare si ricorda che le più importanti **applicazioni NGS** includono:

1. lo studio della regolazione dell'espressione genica, comprese modifiche epigenetiche, interazioni tra proteine e sequenze regolatorie, predizione di varianti di splicing dell'RNA messaggero (mRNA);
2. l'analisi del genoma e la ricerca sugli SNP, nelle regioni codificanti e non codificanti del genoma, la previsione della struttura delle proteine;
3. la diagnosi precoce di patologie attraverso la ricerca e monitoraggio dei biomarcatori".

Big data analysis e banca dati molecolare

Per individuare **set di dati rilevanti a fini prevenzionistici**, il Dipartimento DIT, in collaborazione con Inail-Dcod e l'Università degli Studi "Sapienza" di Roma, ha realizzato una "**Banca dati molecolare INAIL**, denominata **BiTdata**, in assonanza ai Big data".

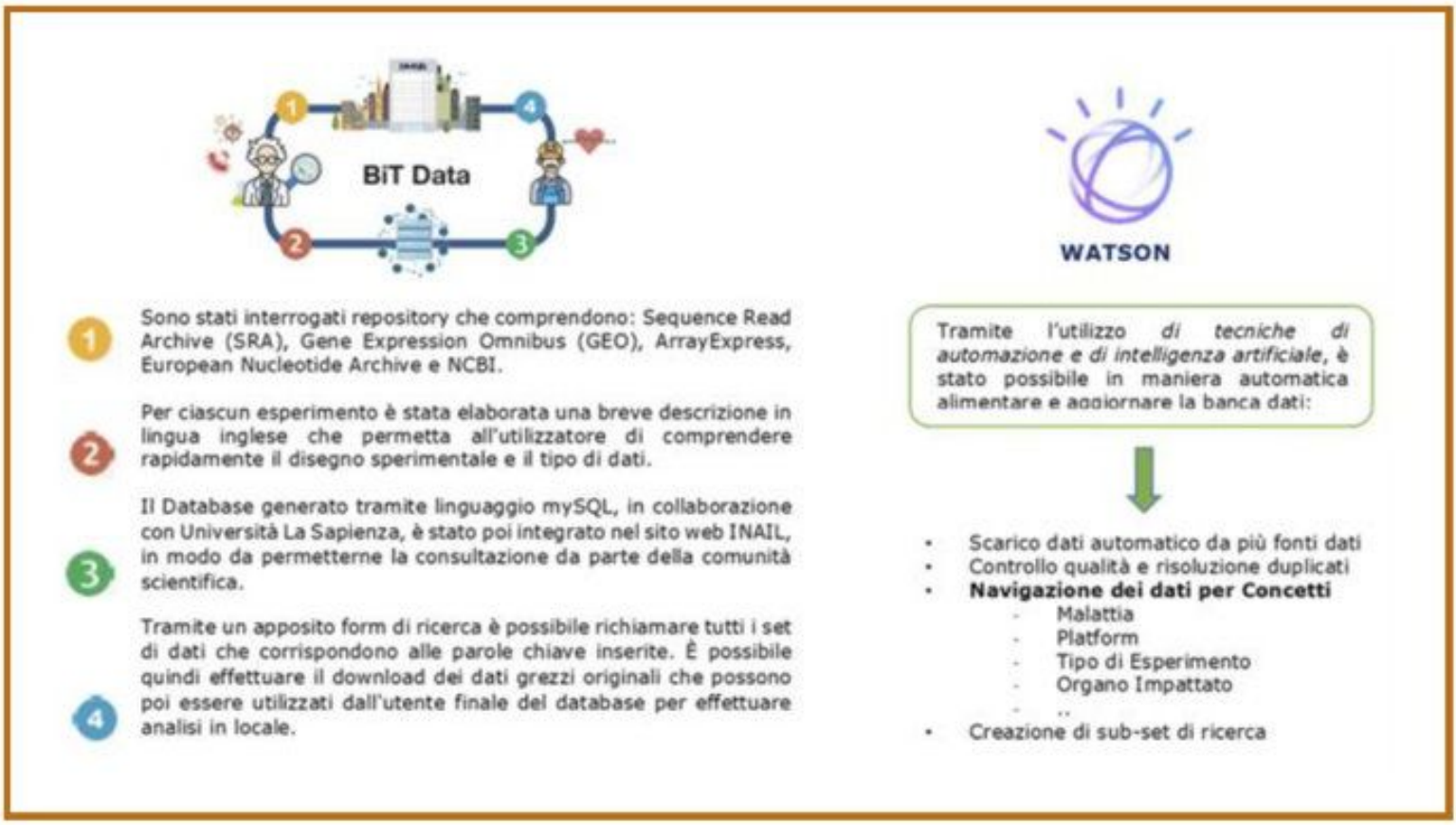
Si tratta ? continua la scheda - di "dati molecolari BioTecnologici" che fungono quasi da '**Biglietto Tecnologico**' dell'**esposizione occupazionale**".

La banca dati BiTdata - consultabile nella sezione dedicata alle attività di Ricerca e Innovazione tecnologica del sito Inail ? "prende in esame Piattaforme Informatiche Internazionali, che mettono a disposizione dataset completi dei principali cambiamenti genomici in seguito ad esposizione occupazionale ad agenti fisici chimici e biologici. È progettata, quindi, allo scopo di individuare set di dati rilevanti a fini della prevenzione, rendendo fruibili e accessibili i dati depositati nei 'repository' pubblici; consentirebbe di effettuare 'meta-analisi', ovvero analisi di campioni con caratteristiche biologiche comparabili, che rappresenta una delle più importanti sfide della bioinformatica".

Inoltre gli studi del **trascrittoma** (la totalità degli RNA trascritti a partire da un genoma), combinati con tecniche di data mining (estrazione dei dati), possono "fornire nuove informazioni sulla patogenesi di numerose patologie e possono contribuire all'identificazione di nuovi biomarcatori candidati con potenziale valore clinico".

Si segnala poi che l'attività di raccolta dei dati della BiTdata "è stata automatizzata, attraverso tecnologie di intelligenza artificiale, e viene rinnovata con cadenza regolare al fine di mantenere costantemente aggiornato il database".

Riprendiamo dalla scheda un'immagine che schematizza le funzionalità di Bitdata della piattaforma di IA:



Big data analysis e intelligenza artificiale

Si segnala poi che la piattaforma sviluppata è in grado di interrogare automaticamente i repository genomici scelti, che "mettono a disposizione in modalità 'open-source' i dati e a seconda di una lista di **sostanze** d'interesse da monitorare, preleva le informazioni rilevanti di ogni espe−rimento". E la lista "può essere aggiornata e rivista da un'in−terfaccia, in questo modo è stato possibile recentemente aggiungere anche SARS-CoV-2 fra le richieste di interrogazione".

Per ogni esperimento "l'applicazione salva non solo l'ab−stract, ma una serie di dati ed informazioni che il re−pository mette a disposizione, come il titolo dell'esperi−mento, l'anno, la tecnica di sequenziamento e vari altri". E successivamente degli algoritmi di intelligenza artificiale, appositamente sviluppati, "agiscono nella descrizione dell'esperimen−to per andare ad estrarre altre informazioni rilevanti, come la presenza di termini derivanti dal dizionario MeSH (Medical Subject Headings),

biomarcatori speci−fici, il tipo di esperimento se in vitro o in vivo".

Questa una lista completa delle **informazioni recuperate** direttamente dai repository o rilevate dall'utilizzo di tecniche di intelligenza artificiale:

- "Agente;
- Fonte dati;
- Tipo di piattaforma;
- Data di pubblicazione;
- Tipo di esperimento;
- Biomarcatore;
- Terminologia MeSH;
- Time Course;
- Numero di campioni;
- Tipo di soggetto;
- Tipo di studio".

Rimandiamo, in conclusione, alla lettura integrale del factsheet, che riporta molti altri dettagli, ad esempio in riferimento alla elaborazione di dati attraverso specifico tool della dashboard della piattaforma di IA.

RTM

Scarica il documento da cui è tratto l'articolo:

Inail, Dipartimento innovazioni tecnologiche e sicurezza degli impianti, prodotti e insediamenti antropici, " Big data analysis e intelligenza artificiale: strumenti Inail a supporto dei ricercatori nella gestione della moltitudine dei dati - omici", a cura di (Inail Dit), S. Gioiosa (Cineca), L. Di Piramo (IBM Italia), Factsheet, edizione 2023 (formato PDF, 838 kB).

Vai all'area riservata agli abbonati dedicata a " Big data analysis e intelligenza artificiale per la gestione dei dati e la sicurezza".

Licenza Creative Commons