

GenomSys contrassegna con successo il suo Codec Suite MPEG-G

La certificazione CE del Codec Suite MPEG-G permetterà all'industria genomica di beneficiare dei vantaggi dell'open standard MPEG-G per applicazioni clinico-diagnostiche.

Losanna, April 2021 – GenomSys, pioniere nello sviluppo dello standard aperto MPEG-G per la rappresentazione, l'archiviazione e l'elaborazione dei dati genomici, annuncia oggi l'ottenimento della marcatura CE per MPEG-G Codec Suite, muovendo un passo decisivo per portare la medicina personalizzata nella vita quotidiana.

GenomSys MPEG-G Codec Suite è una raccolta di strumenti software per l'elaborazione di dati genomici conformi allo standard ISO/IEC-23092 (MPEG-G). I tool permettono alle organizzazioni di implementare lo standard e di sfruttare i suoi vantaggi, quali ad esempio la gestione più efficiente dei dati, e analisi mirate grazie all'accesso selettivo possibile con MPEG-G, codificando e decodificando i dati genomici; includono anche esempi di codice sorgente, un manuale utente completo e un software aggiuntivo per transcodificare da/a formati legacy e per integrare le funzionalità in applicazioni e pipeline esistenti.

GenomSys Codec Suite è stato marcato CE come dispositivo medico diagnostico in vitro secondo la direttiva 98/79/CE, pertanto utilizzabile in ambito clinico.

MPEG-G (ISO/IEC-23092) offre diversi vantaggi per la genomica:

- ❖ **Standard internazionale non proprietario definito dall'ISO** che permette a tutti di scrivere il proprio codice interoperabile, indipendentemente da qualsiasi azienda
- ❖ **Un unico formato unificato** per tutti i formati di file connessi alla genomica (non allineati, allineati, di riferimento, rapporti) per assicurare la massima integrità dei dati e l'interoperabilità
- ❖ **Benefici significativi in termini di dimensioni** che consentono alle organizzazioni che gestiscono grandi volumi di dati genomici di ridurre i costi di archiviazione: ad esempio, file di Whole-exome sequencing di 2 GB invece di 7 GB (-70%) con formati precedenti, con benefici proporzionali alle dimensioni del sequenziamento
- ❖ **Tempo di accesso ai dati drasticamente ridotti**, eliminando tempi morti per i genetisti che eseguono l'analisi: ad esempio, il tempo di elaborazione per accedere a intervalli selezionati è di 2 secondi contro 450 secondi con i formati tradizionali (considerando anche l'external indexing richiesto in questo caso)
- ❖ **Elementi di sicurezza incorporati** che rafforzano la protezione della privacy dei dati genomici sensibili: possibilità di crittografare nativamente all'interno del file e concedere diverse autorizzazioni di accesso per singoli intervalli

Il sequenziamento del DNA è un tema molto sentito dalla comunità dei genetisti ma anche dal pubblico e dai pazienti perché ha cambiato l'approccio al percorso diagnostico delle malattie genetiche e sta entrando in molti altri settori della medicina ed in particolare nell'oncologia[1]. Anche se i fattori abilitanti sono sempre più presenti, come la drastica riduzione dei costi di sequenziamento e la crescente capacità di sequenziamento, la genomica personalizzata non è ancora riuscita ad affrancarsi del tutto [2,3]. Tra i punti aperti troviamo l'incertezza della protezione dei dati genetici e le sfide che riguardano la gestione e l'archiviazione nel lungo termine della mole di dati. MPEG-G Codec Suite, recentemente marcata CE, aiuta i laboratori a trasformare i loro file

formattati in MPEG-G e a sfruttare le efficienze che li accompagnano, unitamente e un elevato livello di sicurezza di questo standard di dati interoperabile.

La capacità di sequenziamento annuale prevista a livello mondiale di 500 milioni di gigabytes (500 Exabyte) entro il 2025 e i requisiti legali, almeno nei paesi europei, di conservare i dati genetici per almeno dieci anni richiedono soluzioni più efficienti in termini di costi nel campo della genetica[2].

"Affinché non solo si parli di medicina personalizzata ma la si faccia diventare realtà, è necessario stabilire un piano che permetta ai laboratori di implementare la diagnostica personalizzata in modo commercialmente ragionevole. Il sequenziamento è già molto efficiente da un punto di vista economico, mentre l'immagazzinamento successivo dei dati deve ancora essere rivisto in questo senso. Attraverso il formato MPEG-G, è possibile comprimere un file whole-genome a meno del 25% delle dimensioni dei formati legacy, con conseguente risparmio sui costi di archiviazione per le aziende", afferma il Chief Technology Officer & Co-Founder di GenomSys Claudio Alberti.

Questa considerevole quantità di dati, che contengono le informazioni più preziose, costituiscono una categoria particolare di "dato personale", ma nonostante le precauzioni normative esistenti molti aspetti rimangono critici. In questo contesto, il miglioramento delle linee guida internazionali per il trattamento dei dati genetici e - come richiesto dagli esperti - l'uso della tecnologia per controllare il flusso dei dati potrebbe aumentare la privacy dei dati genomici di ogni cittadino[4].

Oltre ai vantaggi nella protezione e nell'immagazzinamento dei dati, MPEG-G ha notevoli vantaggi per quanto riguarda il trattamento dei dati genomici. Attualmente, mentre i tempi per il sequenziamento stanno diminuendo costantemente - il sequenziamento dell'intero genoma può ora essere fatto in pochi giorni rispetto a diversi mesi solo cinque anni fa - il tempo di elaborazione dei dati è ancora considerevole. MPEG-G può ridurre sostanzialmente il tempo di elaborazione poiché non richiede alcuna fase di pre-elaborazione durante l'analisi, riducendo conseguentemente il tempo di accesso ad ogni campione di oltre l'89%. I ridotti tempi di elaborazione risiedono nella struttura di indicizzazione di MPEG-G, che permette un rapido accesso all'offset desiderato senza tempi morti per l'ordinamento e l'indicizzazione - un vero accesso selettivo.

"Il tempo è denaro", spiega Claudio *"La tecnologia è stata in grado di ridurre massicciamente i tempi di sequenziamento in modo continuo, dal progetto Genoma Umano dal 2003, agli odierni flussi di lavoro dei laboratori di analisi. Lo standard aperto MPEG-G offre un vantaggio sostanziale nel tempo di elaborazione di un set di dati genomici. Per esempio, l'accesso a più regioni di un gene all'interno di un set di dati di sequenziamento dell'intero genoma richiede con MPEG-G solo 2 secondi rispetto ai minuti dei formati precedenti"*

Il marchio CE del MPEG-G Codec Suite, rilasciato con successo il 31 marzo 2021, dimostra l'ambizione di GenomSys di fornire una soluzione clinica che consenta alla comunità dei genetisti di sfruttare i benefici dello standard, muovendo un passo concreto verso una effettiva medicina personalizzata.

[1] Sandler S, Alfino L, Saleem M. The importance of preventative medicine in conjunction with modern day genetic studies. Genes Dis. 2018 Apr 12;5(2):107-111. doi: 10.1016/j.gendis.2018.04.002. PMID: 30258938; PMCID: PMC6146230.

[2] Stephens et al. (2015). Big Data: Astronomical or Genomical?. PLoS biology. 13. e1002195. 10.1371/journal.pbio.1002195.

[3] National Human Genome Research Institute Link: <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost>

[4] Bonomi, L., Huang, Y. & Ohno-Machado, L. Privacy challenges and research opportunities for genomic data sharing. Nat Genet 52, 646–654 (2020). <https://doi.org/10.1038/s41588-020-0651-0>

About GenomSys

GenomSys SA è una società svizzera di software e tecnologie nel campo della genomica che sviluppa soluzioni sicure e di livello clinico in grado di portare la medicina personalizzata nella vita quotidiana.

Dapprima GenomSys ha portato la genomica nell'era digitale affrontando i problemi dei formati testuali tradizionalmente utilizzati per i dati genomici. La società ha infatti sviluppato e guidato la standardizzazione internazionale ISO di MPEG-G (ISO/IEC-23092), un formato digitale per la rappresentazione dei dati genomici che offre vantaggi sostanziali in termini di compressione, interoperabilità, protezione dei dati e accesso selettivo.

Oggi GenomSys sta compiendo il passo successivo, consentendo a laboratori ed individui di beneficiare dei vantaggi del nuovo standard: software di analisi nativi per una gestione più efficiente dei dati genomici e analisi mirati grazie all'accesso selettivo possibile con MPEG-G, e la possibilità per gli individui di memorizzare e far analizzare i propri dati genomici direttamente dal proprio smartphone, portando quindi la genomica nel mondo mobile.

In questo modo, GenomSys sta definendo un nuovo paradigma con una soluzione di altamente differenziata: una value proposition coerente, che mette al centro privacy e convenienza, per fornire davvero una medicina di precisione all'individuo sfruttando l'unicità dello standard MPEG-G, in maniera diretta o attraverso partner commerciali.

Contatto: Lucas Laner, Marketing Manager: lucas.laner@genomsys.com

Website: www.genomsys.com