

GenomSys successfully CE marks its MPEG-G Codec Suite

The CE mark allows the genomic industry to fully exploit the MPEG-G open standard's benefits for clinical-grade diagnostic applications.

Lausanne, April 2021 – GenomSys, a pioneer developing the ISO-defined open standard for genomic data storage and processing, announces today the successful CE marking of its MPEG-G Codec Suite, taking the next step in delivering an operational framework to establish personalized medicine for everyone.

GenomSys MPEG-G Codec Suite is a collection of software tools to process genomic data compliant with ISO/IEC-23092 genomic data standard (MPEG-G). The tools enable organizations to implement the standard and leverage its benefits by encoding and decoding genomic data; they also include source code examples, a comprehensive user manual, and additional software to transcode from/to legacy formats (to preserve past investments) and to integrate functionalities into existing applications and pipelines.

Our Codec Suite is CE marked as an in-Vitro Diagnostic Medical Device according to 98/79/CE directive and hence approved for clinical-grade diagnostic purposes.

MPEG-G format offers several advantages for genetics:

- ❖ **Non-proprietary ISO-defined international standard** enabling everyone to write its own interoperable code and independence from any company's strategy or continuing support
- ❖ **Single unified format** for all genomic-connected file formats (unaligned, aligned, reference, reports) to assure the highest data integrity and interoperability
- ❖ **Significant size benefits** enabling cost savings for organizations handling large volumes of genomic data: e.g., Whole-exome sequencing file of 2GB instead of 7GB (-70%) with legacy formats, with benefits growing with the size of the sequencing
- ❖ **Dramatically faster data access time** reducing latency for geneticists running the analysis: e.g., processing time to access selected intervals only 2secs vs. 450secs with legacy formats
- ❖ **Built-in security elements** strengthening privacy protection of sensitive individual genomic data: possibility to encrypt natively within the file and grant different access permissions by intervals

Genetic testing is becoming an increasingly important tool within healthcare and especially in disease prevention[1]. Although enablers are increasingly in place, such as dramatic reduction in sequencing costs and growing sequencing capacity, personalized genomics has not yet been able to take off fully[2,3]. The bottlenecks lie with the uncertainty of data privacy and the challenges dealing with significant amount of data. The newly CE marked MPEG-G Codec Suite helps laboratories transform their legacy formatted files into MPEG-G and leverage the accompanying efficiencies and high-security level from this widely interoperable data standard.

The expected worldwide annual sequencing capacity of 500 million gigabytes (500 Exabytes) by 2025 and the legal requirements, within European countries at least, to store genetic data for ten years require more efficient solutions in terms of storage cost-savings in the field of genetics[2].

"In order for us not only to talk about personalized medicine but to make it a reality, a framework must be established that enables laboratories to implement personalized diagnostic in a commercially reasonable way. Sequencing is already very efficient from an economic point of view, whereas subsequent data storage has yet to be considered more from this perspective. Through the MPEG-G format, it is possible to compress a whole-genome file to less than 25% of the size of legacy

formats, resulting in cost savings for companies.”, explains GenomSys’ Chief Technology Officer & Co-Founder Claudio Alberti.

This remarkable amount of stored data, maintaining the most intimate personal information, will need further clarification regarding data protection. In this context, transnational improvements in guidelines for handling genetic data and - as called for by experts - the use of technology to control data flow could increase data privacy for each citizen’s genomic data[4].

In addition to the benefits in data protection and storage, MPEG-G has considerable advantages with regard to the processing of genomic data. Currently, when turn-around times for sequencing are steadily decreasing - whole-genome sequencing can now be done in a matter of days compared to multiple months as five years ago - data processing time is still lagging behind. MPEG-G can substantially shorten processing time since it does not require any preprocessing steps during the analysis, consequently cutting time by more than 89% to access each sample. The paramount decrease in time lies inside MPEG-G’s indexing structure, allowing rapid access to the desired offset without any time spent on sorting and indexing - a actual selective access.

“Time is money,” says Claudio. “The genomic community was able to massively cut down sequencing time continuously since the Human Genome project in 2003, with the challenges lying now outside the wet-lab and in the analysis. The open standard MPEG-G delivers a substantial benefit in the processing time of a genomic dataset. For example, the access of multiple regions of a gene within a whole-genome sequencing dataset takes with MPEG-G only 2 seconds compared to minutes for legacy formats.”

The MPEG-G Codec Suite’s CE marking demonstrates GenomSys’ ambition to provide a clinical-grade solution for the genetic community to leverage the standard’s benefits and an essential step towards real personalized medicine. The CE mark was successfully issued on March 31st.

-
- [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 is a Swiss health technology software company that develops secure, clinical-grade solutions bringing personalized medicine to everyday life.

GenomSys first took genomics to the digital era by addressing the issues of legacy genomic text-based data formats. The company developed and led the ISO international standardization of MPEG-G (ISO/IEC-23092), a digital data format for genomic data representation that delivers substantial benefits in terms of efficient compression, interoperability, native protection, and selective access.

Now GenomSys is taking the next step, enabling laboratories and end-users to take advantage of the new standard: MPEG-G native analytical pipelines for more efficient storage and focused analyses via selective access, and the ability for people to store and analyze their own genomic data in their smartphones, bringing genomics into the mobile world.

In doing so, GenomSys defines a new paradigm with a highly differentiated solution: a consistent value proposition emphasizing privacy and convenience to truly deliver personalized medicine to individuals by leveraging the uniqueness of the new digital genomic standard and the pervasiveness of mobile, directly or through business partners.

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

Website: www.genomsys.com