

Decoding Data Challenges in Diagnostics

Strand's Solutions

Preface

If you are a midsize to large diagnostics lab onboarding a new assay or trying to scale an existing assay, here are some common challenges you may face and our potential solutions.





Increased TAT Due to High Sample Volumes:



YOUR CHALLENGE:

Running only one flow cell at a time on DNAnexus leads to increased turnaround time (TAT) for samples and difficulty processing high sample volumes.



OUR SOLUTION:

We update the pipeline to execute multiple samples simultaneously, typically with a TAT of 5 and 9 days for WES and WGS samples, respectively, utilizing the full computational power of DNAnexus.

Data Storage Issues:



YOUR CHALLENGE:

Data storage on DNAnexus/ AWS s3/ ICA has already reached several petabytes, making it difficult for labs to scale in terms of space and cost with the increased sample volumes.



OUR SOLUTION:

We select a long-term storage solution and optimize the timing of data transfer. This involves determining which data types should be kept in an accessible storage and which should be moved to a deep archive storage. To ensure smooth workflows, we build detailed SOPs for data storage and archival. This strategy has aided us in storing NovaSeqX Plus data, saving 60% on storage costs for ~16 PB of data generated over the next decade. Find our recent poster here. Additionally, we explore alternative cloud storage solutions on AWS or Azure, which are more cost-effective.

Data Transfer Across Platforms:



YOUR CHALLENGE:

Moving data across various platforms (ICA to DNAnexus, DNAnexus to s3, etc.) is challenging due to the volume of data.



OUR SOLUTION:

We write custom scripts to automate the transfer of around ~10TBs of data per month. We also consider incorporating available data management platforms to improve how data is copied, moved, and managed.



Integration with LIMS:



YOUR CHALLENGE:

The current integration with LIMS and other bioinformatics platforms needs optimization.



OUR SOLUTION:

We tailor our solution to the specific needs of each laboratory. On a high level, we incorporate **HL7 APIs for smooth integration** between various systems.

Raw Data Access:



YOUR CHALLENGE:

Raw data needs to be made available to partners through different channels.



OUR SOLUTION:

We **centralize all raw data in one portal**, allowing users to download the required data to their preferred storage location.

Combining Multiple Data Types:



YOUR CHALLENGE:

Multiple data types need to be combined for unified reporting.



OUR SOLUTION:

We frequently combine various data types to facilitate comprehensive reporting. For instance, our tertiary analysis workflow in StrandOmics involves incorporating NGS results from StrandNGS along with the outcomes of other experiments, such as RNAseq and immunohistochemistry (IHC). This injection of multiple data types has facilitated accurate summarization and customized reporting for over 30,000 cases.

Launching and Tracking Pipelines:



YOUR CHALLENGE:

In addition to providing pipelines to be run, ICA offers an auto-launch feature that initiates, tracks, and reports the status of each pipeline. However, customized versions may be needed to obtain the required scale with this feature.



OUR SOLUTION:

We design an alternative **customized system for launching pipelines** and provide a **tailored dashboard** to track and monitor ICA run metrics. This system ensures a greater degree of surveillance and reliability in terms of pipeline failure reports and restarting runs.



We were very impressed with the quality of work and timeliness; you're definitely our go-to for bioinformatics consulting

- Director, Bioinformatics, Illumina



We were immensely impressed by Strand's ability to rapidly recruit a substantially sized clinical cohort of cancer patients, and to design and run a complex liquid biopsy panel on samples drawn from the cohort, all in roughly a year's time.

- Dr. Nishant Agarwal

Chief of Otolaryngology-Head and Neck surgery and director of Head and Neck Surgical Oncology, University of Chicago.



We have been using the StrandOmics pipeline to analyze and generate a report for our clinical cancer panels for over three years now. I would highly recommend using it to analyze data generated from clinical cancer NGS panels and the outputted clinical report provided after analysis.

- Senior Scientist/ Medical laboratory director for NY State, Prim Bio Research Institute



24+

YEARS OF EXPERIENCE

80,000+ Genetic Tests Reported

500+ Projects Executed for Genomics Majors Globally

Presence in 20+ Countries

strand

7th Floor, MSR North Tower, #144, Outer Ring Road, Nagavara, Bengaluru - 560045

+91 9980448044

hello@strandls.com

strandls.com

