GRAKN.AI

lifebit

Lifebit uses Grakn to improve genomic analyses over the cloud

Industry
Biotechnology

Use Case
Cloud Resource Optimisation

Challenge
To find the most efficient resources to run specific analyses in the cloud

Solution
To use real-time information of previous genomic analyses, store this in Grakn, and learn from them using Machine Learning

Result
Lifebit Deploit provides its users with optimised suggestions about which instances are the most efficient for their genomic analyses

Lifebit is a biotechnology business that offers a platform-as-a-service, called Deploit, which guides the user from data to insight and fully integrates with the way scientists already work. It advances their work through resource-management handling, cost optimisation and learning from previous analyses —all of this is wrapped around an intuitive UX/UI.

To provide the users of Deploit with recommendations on how to run their genomic analyses over the cloud, Lifebit's team chose to use Grakn. One of their Bioinformatician, Luisa Santus, explains: "Grakn helps us to efficiently store our data and easily use it to train a Machine Learning model".

Challenge

In the last few decades, sequencing technologies have evolved incredibly fast leading to drops in magnitutde in the cost of sequencing a genome. This fast evolution is leading into an unprecedented production of genomic data. So when bioinformaticians try to make sense of genomic data, they usually follow a path to insights which is very long and troublesome.

Because of this, Lifebit created a platform-as-a-service, called Deploit, which guides the user from data to insight and fully integrates with the way scientists already work. It automates their multi-omics and other big data HPC/cloud deployments, enabling them to leverage AI for breakthrough insights generation.

Lifebit recognised the challenge it faced in finding the best resource to run specific analyses in the cloud. They needed a technology which would help them to both efficiently store their data and easily use it to train a ML model. To address this challenge, Luisa's team decided to use Grakn.

"At Lifebit, we monitor real-time information of previous analyses, store it in Grakn and learn from them in order to provide suggestions to users on which instances are the most suitable for their analyses."

Luisa Santus
Bioinformatician
Lifebit

Why Grakn

At Lifebit, Luisa's team chose Grakn because of the fast onboarding process, the ease to quickly master Grakn's most complex functionalities, and the technology's evolvability. As Luisa mentions: "we need it to not only model our data today, but also be able to do so with an ever more complex and large dataset".

Once Luisa's team chose to use Grakn, they defined their schema accordingly, and what kind of data should be stored, in order to find the best resources to run a genomic analysis.

Benefits

Lifebit is now using Grakn to provide suggestions on resource optimisation to run genomics analyses over the cloud while minimising costs and runtimes. Luisa explains: "At Lifebit, we monitor real-time information of previous analyses, store it in Grakn and learn from them in order to provide suggestions to users on which instances are the most suitable for their analyses (depending on the nature of their workflows, data and needs)."

For their schema, Luisa defined a model in Grakn which represents a specific pipeline, on a specific dataset, on a specific machine. With this information, they can then learn from it to improve their recommendations. This is done by training a Machine Learning model in Python which gives them the answers they want.

Luisa concluded: "In this case, Grakn was again very helpful since it provides a Python client".

Grakn is the knowledge graph to organise complex networks of data and making it queryable, by performing knowledge engineering. Rooted in Knowledge Representation and Automated Reasoning, Grakn provides the knowledge foundation for cognitive and AI systems, by providing an intelligent language for modelling, transactions and analytics. Being a distributed database, Grakn is designed to scale over a network of computers through partitioning and replication.

Get in touch with us:
enterprise@grakn.ai
grakn.ai/community