Web application for protein structure prediction

The breakthrough in protein structure prediction came in 2021 with Google DeepMind's AlphaFold, which streamlined lab research using deep learning software. However, accessibility to this tool is a challenge for scientists who are not familiar with the command line. This issue can be addressed by developing a user-friendly interface connected to high-performance hardware for easy computation submission. The input of this work is a familiar solution that no longer suits the current technical, security, and design requirements and the goal is to recreate a similar but improved software. Making this technology accessible contributes to science in different ways, such as drug discovery in medicine and the design of novel enzymes in biotechnology.

Author
Romana Ďuráčiová

BRNO FACULTY UNIVERSITY OF INFORMATION OF TECHNOLOGY TECHNOLOGY

Excel@FIT 2025

MUNI CERIT-SC

e-INFRA CZ

1 Motivation

Create a user-friendly web application for scientists that are not familiar with command line and running scripts.
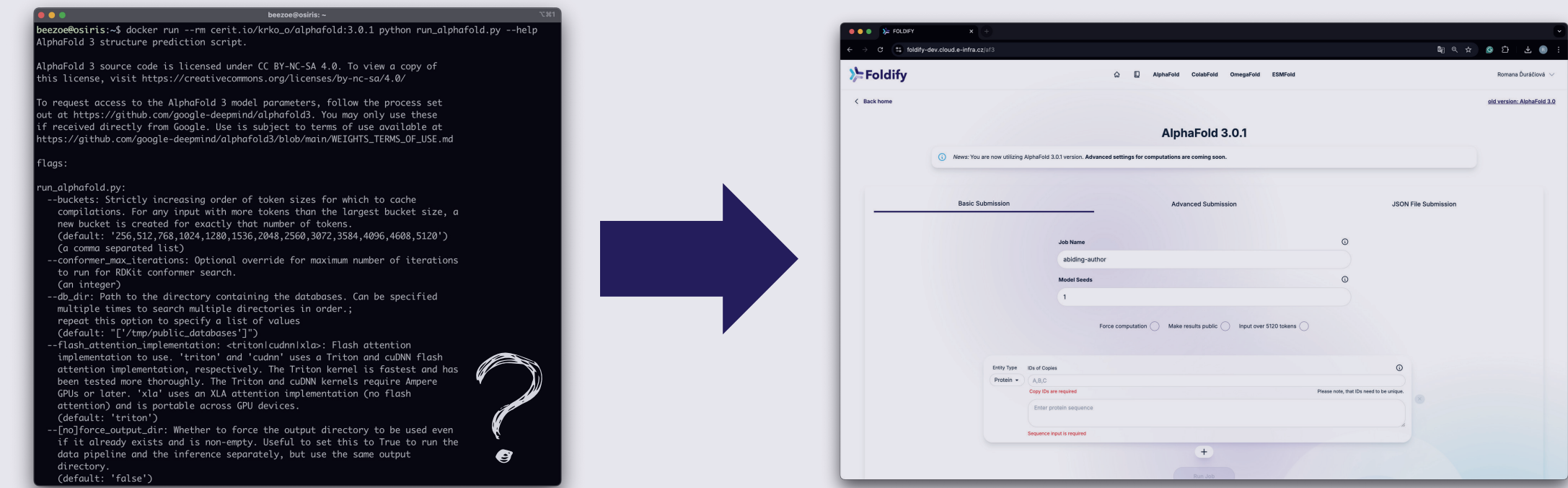

Figure 1. Motivation

2 From primary to quaternary structure

The aim of protein prediction is to derive the tertiary or quaternary structure from the primary sequence.

Primary Protein Structure   Secondary Protein Structure   Tertiary Protein Structure   Quaternary Protein Structure
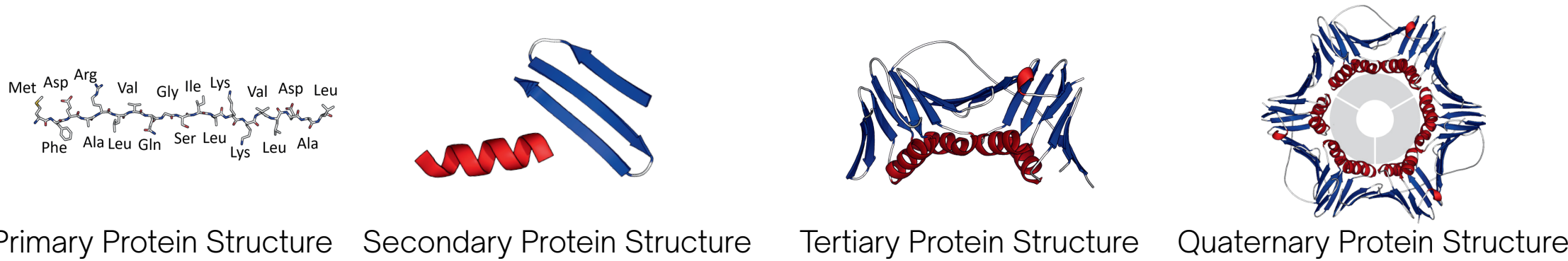
Figure 2. Protein Structure Levels

During the computation process, the user provides the primary structure in FASTA format, which is used by the AI to predict the 3D model. The details of the result are typically presented in either PDB or MMCIF format.
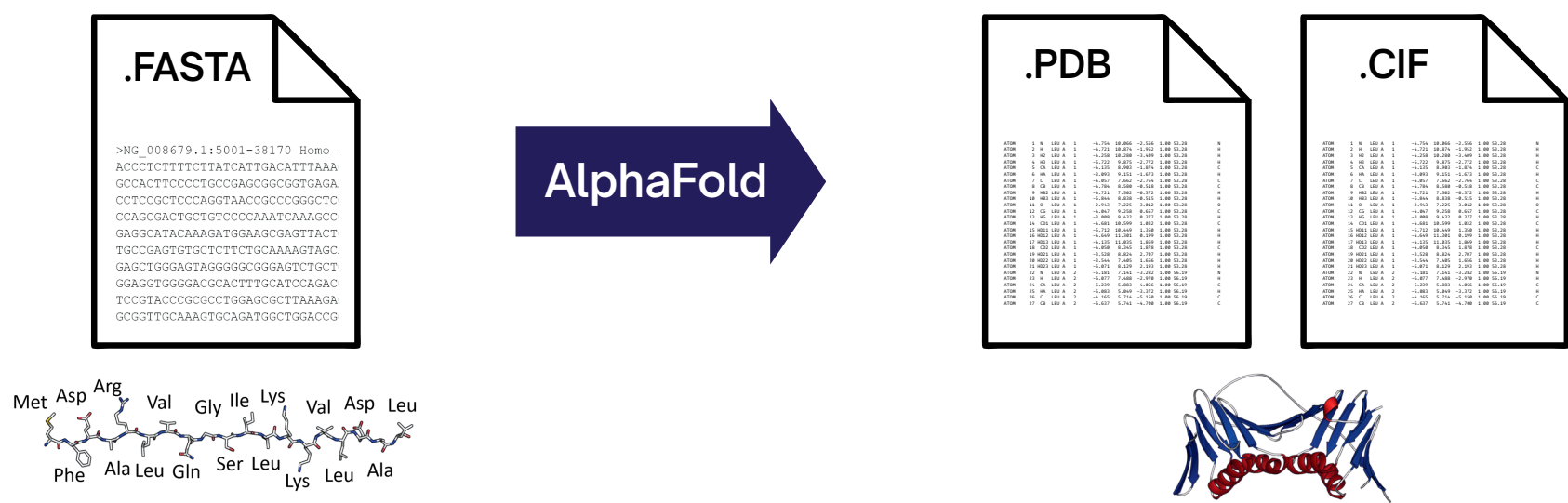

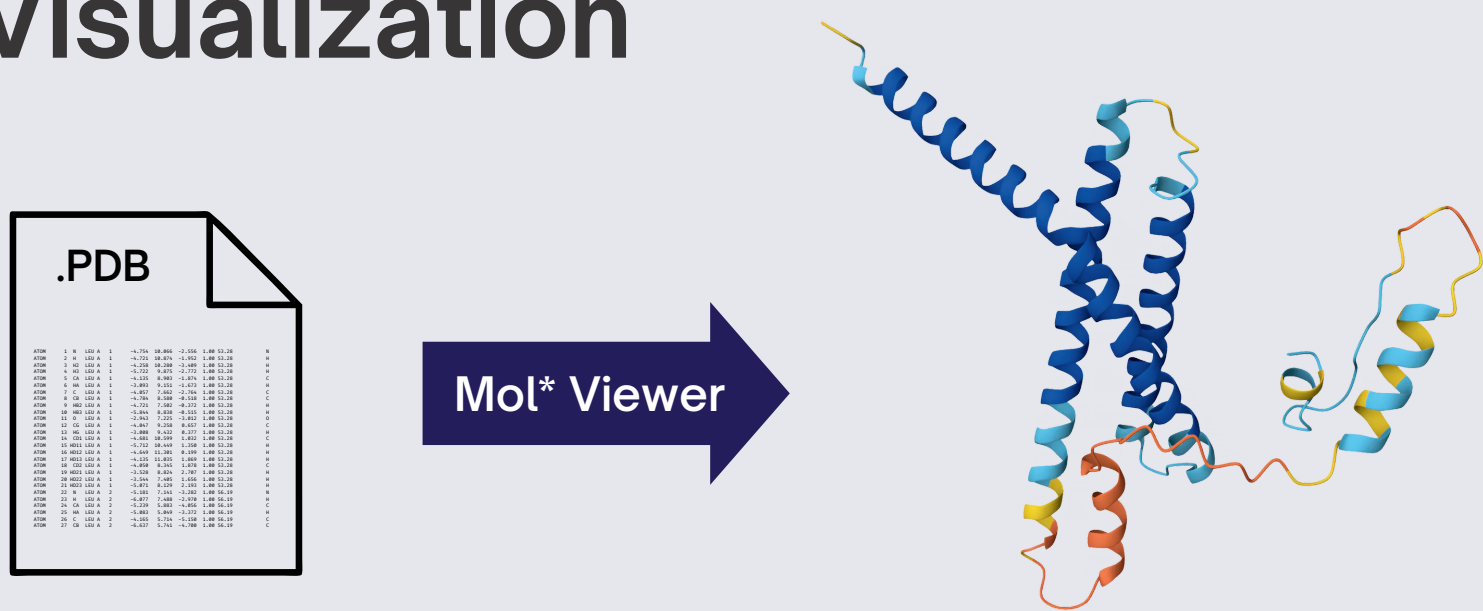**Figure 3.** Input and Output Data Formats

3 Data Visualization

Figure 4. Protein Visualization

4 Behind the Scenes

Server
Back-End
- handling user requests
- managing job execution
- retrieving and serving data

Client
Front-End

HTTP request
HTTP response

Initates Computation
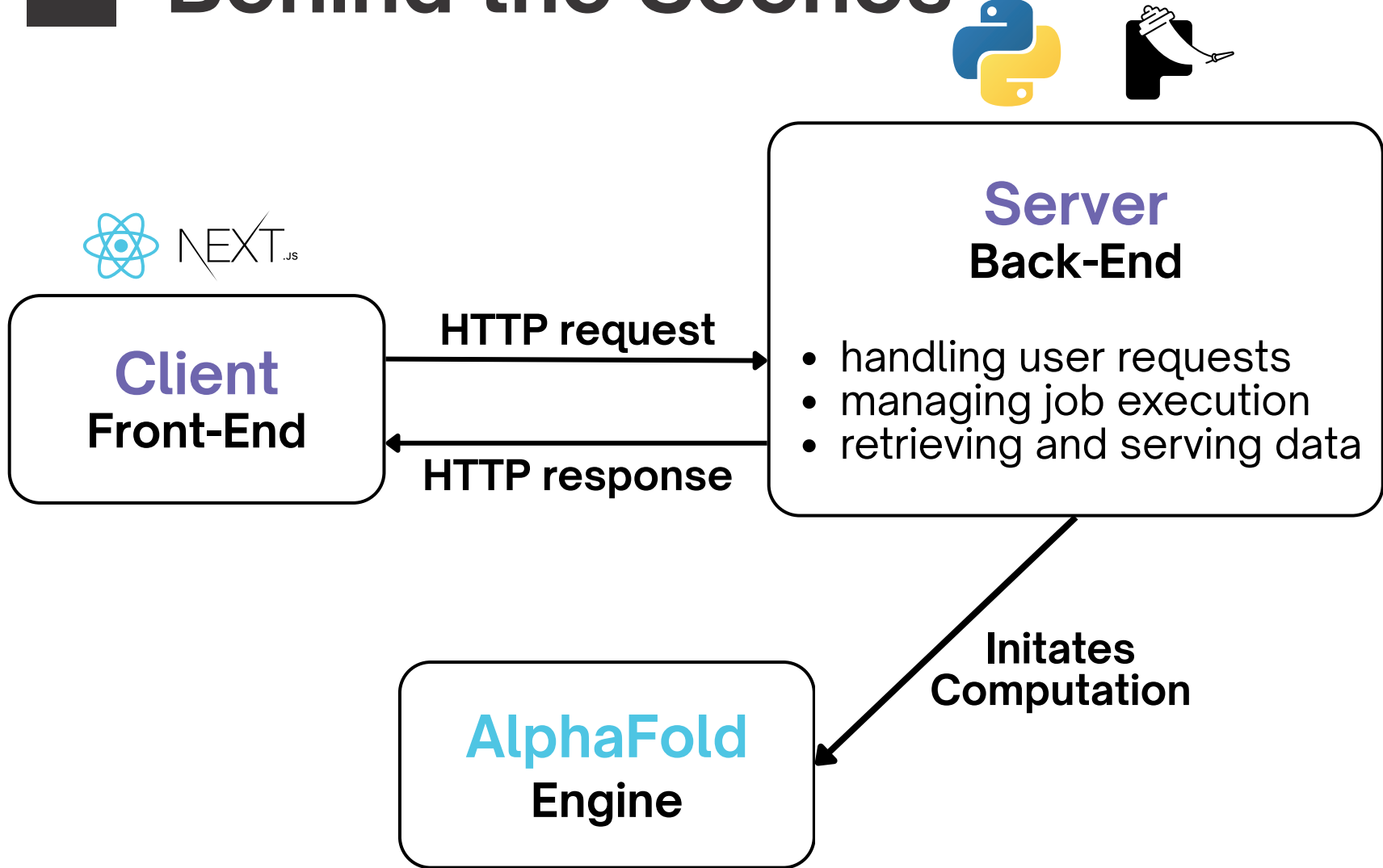
AlphaFold
Engine

Figure 5. Application Architecture

5 Result

Foldify

Simple interface full of features deployed within a Kubernetes cluster. Successfully predicted over **700 proteins** after five months of operation.
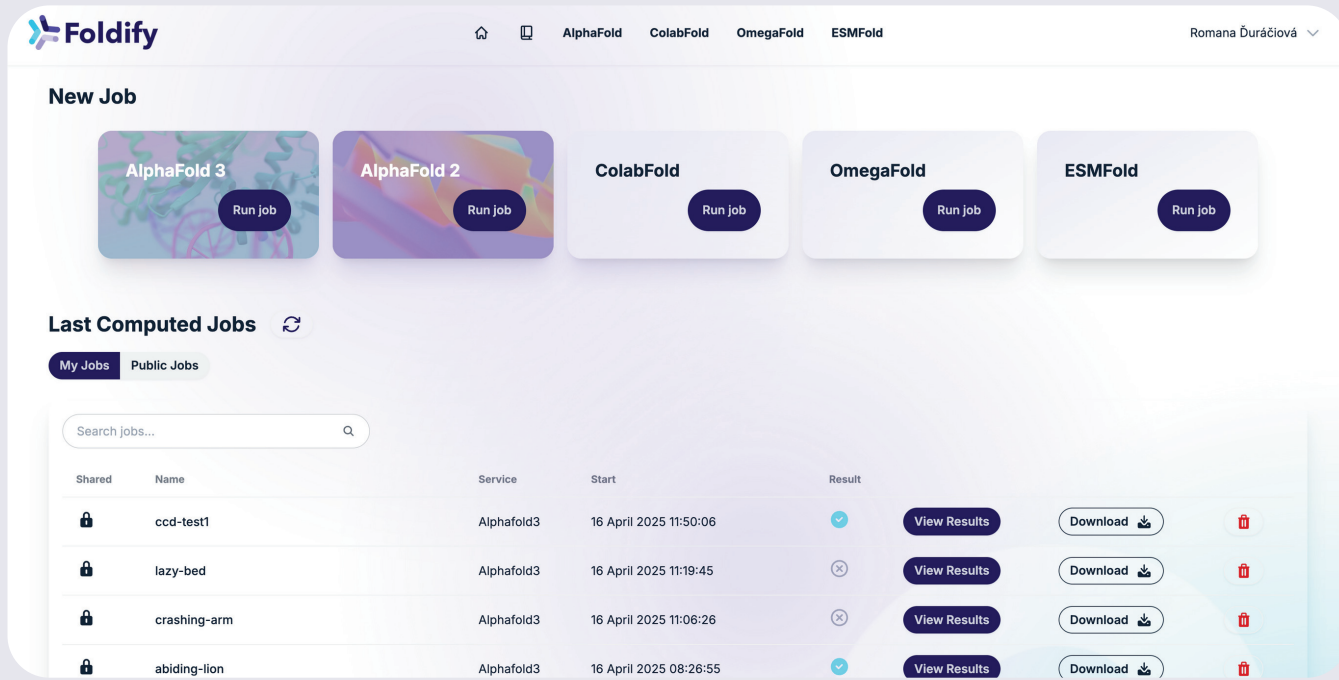

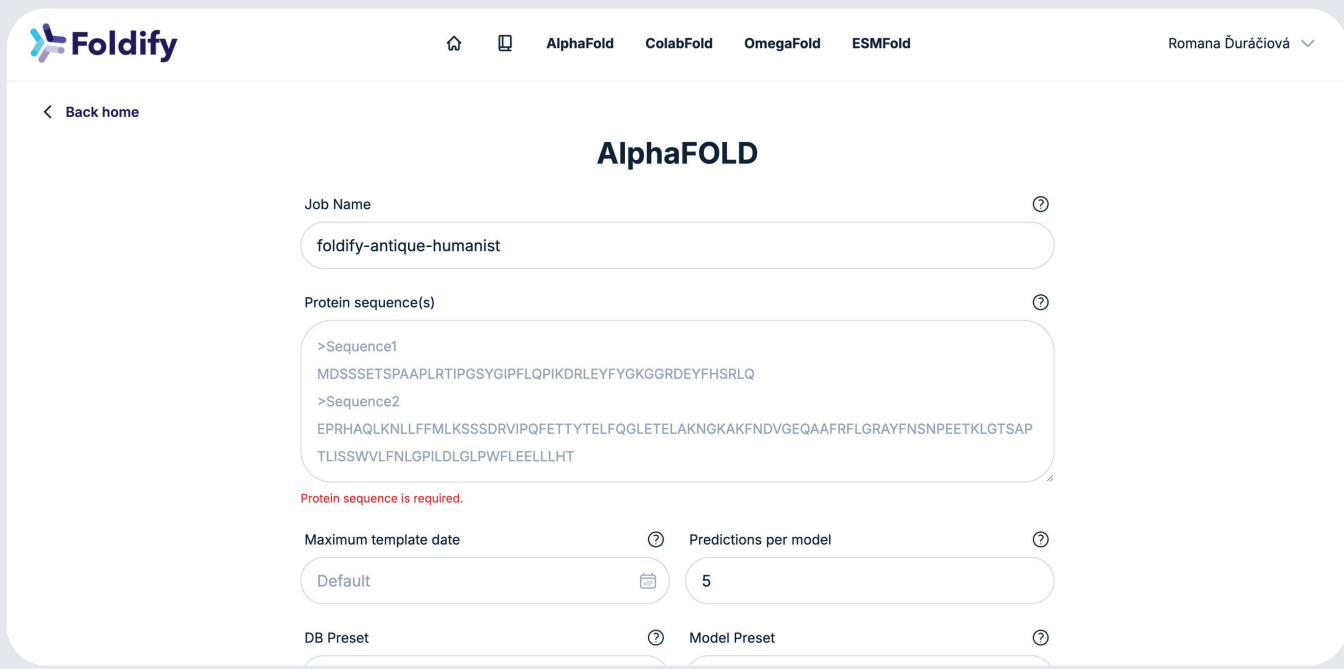**Figure 6.** Dashboard Screenshot
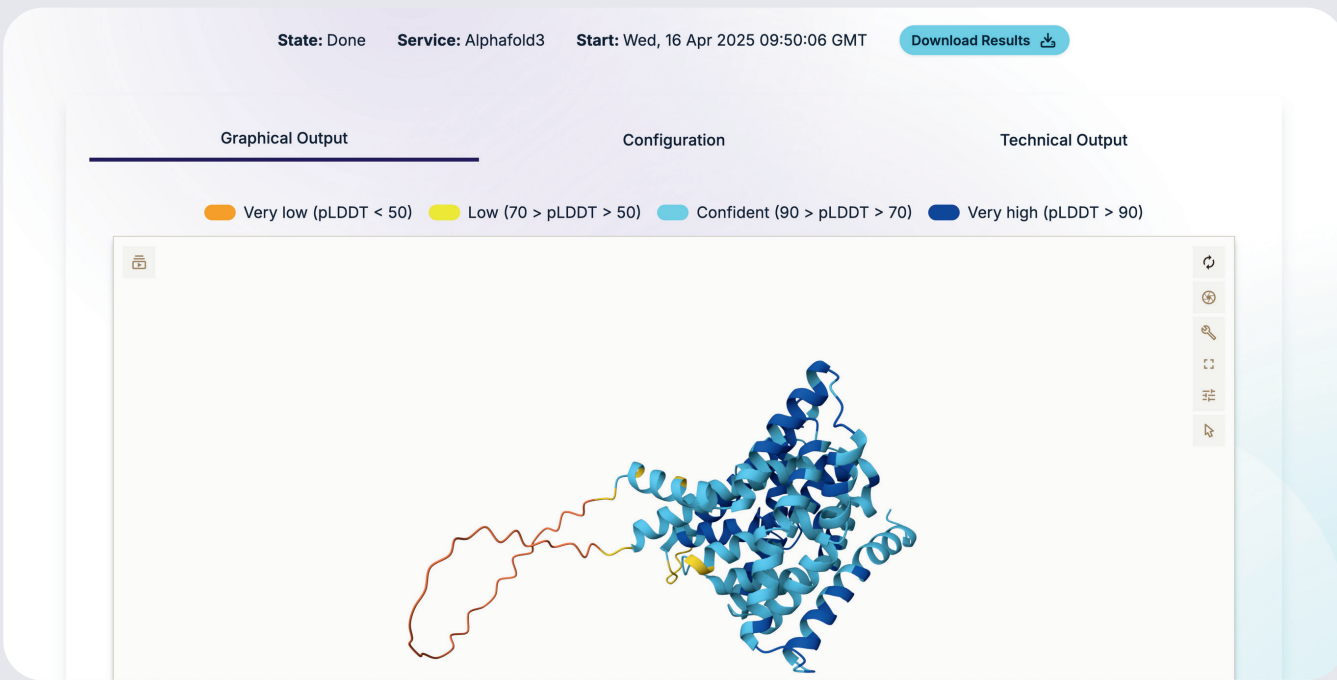

Figure 7. Submission Form Screenshot

Figure 8. Result Screenshot