

# Easy AlphaFold

## Package for fast and convenient summary of the AlphaFold data

Maciej Sikora, Małgorzata Sudół, Maria Bochenek, Kamil Pawlicki

Wednesday 18<sup>th</sup> May, 2022

# Project main goal

Easy-AlphaFold-Report is a package with the aim of processing data between databases easier and faster for the user.

Users can provide input in multiple formats - Pfam families, clans, Uniprot or PDB. Additionally, for even easier access for less tech-savvy users, the main script enables the use of a convenient GUI for a more click-oriented experience.

# Project main goal

As a result, users can expect an interactive and easy to read report with a summary of a provided dataset with a focus on AlphaFold database features. To help with the analysis, basic statistics are also calculated for a quick overview.

# Project main goal

Instead of clicking through multiple sites, and finding connections between databases, this package will do the work for you and if you still want to verify the results, or grab more data links on the report can send you to the respective websites.

Optionally during the summary process, the user can also choose to keep the data from AlphaFold in the .cif files (cif is a more descriptive alternative for PDB files).

# Project features

- Flexible input providing (direct input, files).
- Convenient usage via GUI.
- Various data type recognition and parsing (Pfam families, clans, PDB, Uniprot).
- Automatic database updates.
- ID verification.
- Data collection from multiple databases (PDB, PFAM, Uniprot, AlphaFold).
- Optional data download from AlphaFold (.cif files)
- Presenting results and statistics in the intuitive and interactive html.

For the basic version with GUI user needs to simply run the main python script:

- `python Easy_AlphaFold.py`

This should open an interactive window ready to use.

For programmatic access, terminal input is also available and uses the same script as a base - but with manual flags.

Description of the flags can be found in the Github README.md file.

## Input

In the interactive version, the user can provide ID manually via a dedicated box or use a file selector to provide a path to the file. The file should contain 1 ID per line.

