Záměr týmového projektu

**Studijní program:** Softwarové a datové inženýrství

**Typ projektu:** Softwarový projekt (NPRG069)

**Studenti:** Katarína Bucková, Richard Fedák, Samuel Karaš, Milan Truchan

**Vedoucí:** doc. RNDr. David Hoksza, Ph.D.

**Konzultanti:** Mgr. Marian Novotný, Ph.D., Bc. Lukáš Polák, Mgr. Petr Škoda, Ph.D.

**Název a téma projektu:** Rozšírenie projektu Prankweb

**Prehľad**

Cieľom práce je rozšírenie existujúceho projektu Prankweb, ktorý umožňuje, okrem iného, na základe štruktúr proteínov predikovať a vizualizovať potenciálne väzobné miesta ligandov. Rozšírenie spočíva v pridaní nového spôsobu nachádzania väzobných miest pomocou

1. pLM (protein Language Model), ktorý bude predikovať väzobné miesta zo sekvencie proteínu.
2. AHoJ-DB obsahujúcej predpočítané štruktúry pre danú sekvenciu. Ku každej štruktúre databáza poskytuje 2 typy väzobných miest pre ligandy:
   1. APO (ligand free) - miesto s nenaviazaným ligandom
   2. HOLO (ligand bound) - miesto s naviazaným ligandom
3. MMSeq2 obsahujúcej homologické sekvencie iných živočíšnych druhov.

Projekt je rozdelený na 4 časti: Frontend, Backend, Dátová časť, Integrácia pLM.

**Frontend**

Cieľom je vizualizovať štruktúry proteínov a ich (predikované) väzobné miesta (TODO aj s ligandami?). Taktiež je cieľom rozšíriť systém o možnosť vizualizovať viacero štruktúr s predikovanými väzobnými miestami (TODO ligandami?).

**Backend**

TODO

**Dátová časť**

TODO

**Integrácia pLM**

TODO

**Približný priebeh**

Odhadovaná dĺžka projektu je štandardných 9 mesiacov. Časti projektu sú nasledujúce:

1. Oboznámenie sa s doménou a štruktúrou projektu
2. Definícia funkčných požiadaviek
3. Tvorba špecifikácie projektu
4. Implementácia rozšírenia
5. Návrh a vývoj prípadných dodatočných funkcionalít
6. Testovanie
7. Tvorba dokumentácie