Initial Work Plan

1. Become familiar with the current research project

The first step will be to understand the drug repurposing research currently being conducted in the group.

I will provide:

- a **PowerPoint presentation** describing the project and the current pipeline;
- a **draft of the paper** I am currently writing, which provides a more detailed description of the methodology and results;
- examples of pipeline outputs for selected diseases.

The goal is to understand:

- the main research questions;
- how **Open Targets** disease–gene associations are used;
- how disease gene modules are constructed;
- how the **INDRA mechanistic knowledge graph** is used;
- how drug targets are connected to disease modules;
- how mechanistic paths are identified;
- how the random baseline and network-based drug ranking work;
- how different ranking methods are generated and evaluated;
- how known drugs can be used to validate the results.

2. Understand the INDRA and PPI-based pipelines

We currently have two complementary network-based approaches:

- **INDRA-based pipeline:** uses a directed mechanistic network with biological statements, relation types, belief/evidence information, mechanistic paths, and supporting publications. h

- **PPI-based pipeline:** a parallel version of the drug repurposing pipeline based on protein–protein interaction networks, developed by an ELTE student.

The goal is to understand both approaches and how their results can be compared and potentially integrated.

3. Drug Repurposing Explorer

As an initial development task, we would like to build a **web-based Drug Repurposing Explorer** for the project.

The system should eventually provide a common framework for exploring results from both the **INDRA and PPI pipelines**.

Developing the Explorer will also be part of the onboarding process, since working with the data and pipeline outputs will provide a deeper understanding of the overall project.

4. Detailed analysis of a specific disease

After becoming familiar with the project and the existing methods, the next research task will most likely be to select **one specific disease** and study it in greater depth using the algorithms and pipelines that have already been developed.

This may include:

- analyzing the disease gene module;
- comparing INDRA- and PPI-based results;
- examining highly ranked drug candidates;
- studying mechanistic paths and supporting evidence;
- evaluating known drugs and potential repurposing candidates;
- identifying strengths and limitations of the existing ranking approaches.

5. Longer-term direction

Once the existing pipeline and the selected disease have been studied in detail, we can identify a more specific research question based on the results.

This could later lead to the development or improvement of algorithms for areas such as:

- INDRA–PPI integration;
- mechanistic path analysis;
- evidence integration;
- drug ranking and re-ranking;
- or other methodological questions that emerge from the disease-specific analysis.

The immediate priority, however, is to **understand the existing work first, contribute to the Drug Repurposing Explorer, and then apply the existing methods to a specific disease in depth**.