

## **Abstract**

This thesis explores the differences and data flow among eleven protein-protein interaction (PPI) databases. It addresses the lack of user awareness regarding the independence of databases and data processing methods. The study analyzes data collection methods, types of information provided, organism coverage, and the mutual dependencies between databases. Significant variability was found in the number of proteins and interactions, validation approaches, and species coverage. The thesis provides a graphical map of data flow and species overlap, emphasizing the need to consult multiple databases. The work aims to help readers navigate PPI databases and select appropriate sources for their needs.