

# 1. Introduction

We now provide an overview of a problem domain and a discussion of the intended users.

## 1.1. PROBLEM STATEMENT

Astrophysicists studying binary stars use a program called POSYDON (POpulation SYnthesis with Detailed binary-evolution simulatiONs) to simulate binary star interactions. Studying stars directly can be time-consuming and costly, so astrophysicists often conduct primary research via simulations before verifying their results with real astronomical observations. In particular, POSYDON studies systems where two stars interact with each other.

Given the number of simulations it performs and the criteria it collects, POSYDON generates a large amount of data and stores this across numerous spreadsheet files in Comma Separated Values (CSV) format. Due to the dispersed nature of the data, there is currently no efficient way to search and analyze the information. To address this, our project will provide a software tool to import this data into a common location. Following data processing, users will be able to search collected binary star data using standard English language requests, by utilizing pre-generated queries, or by writing custom queries using Structured Query Language (SQL).

Without a tool such as this, users would have difficulty formulating their requests for obtaining particular subsets of the simulated data that satisfies their desired properties. Specifically, for every desired query, a complex program will need to be constructed - as opposed to having a compact, high-level language providing a declarative ("what") without the need for procedural ("how") details. By simplifying data access and supporting custom data, this project aims to make binary star data querying more accessible to students, educators, and researchers, allowing for greater collaboration. In addition, similar queries can be reused almost verbatim by changing the values in their parameters.

## 1.2. INTENDED USERS

Our intended users are divided into three primary groups: astrophysicists, educators, and students. Each of these groups has different needs when utilizing our tool.

Astrophysicists primarily utilize the data produced by POSYDON to compare data points between the stars. At the same time, however, they are often very busy, and the current product does not support efficient access to this dataset. Many of these researchers may not be familiar with querying languages such as SQL. As a result, they will need a tool that can efficiently retrieve the data using natural language queries instead of requiring complex SQL commands. In turn, our product will provide opportunities for increased productivity for researchers by reducing the need to look through multiple files. Instead, they will be able to write a short request and receive the data they need in a timely manner.

Educators would primarily use the tool to help facilitate an engaging and interactive classroom learning environment. Since they may not be experts in binary star systems, educators would need this tool to be intuitive and easy to use so that creating exercises and classroom activities is simple. This would assist educators in showing students the complex relations between these astrophysical events and how they connect to larger cosmic anomalies.

Students will primarily use this tool for their own study of binary star systems to assist with coursework and personal research. Since students may not be familiar with SQL, students need a tool that provides an intuitive, user-friendly interface that simplifies the querying process. This simplified tool will help the student focus on understanding and analyzing the data rather than struggling with the technical challenges.