



Faculty of  
Computer Science

Data Science and Business Analytics

Moscow  
2025

# Development of an Interactive Platform for Catalog Analysis and Visualization of Galaxy Clusters with AI Support

## **Presented By**

Pavlov Arkady,  
2<sup>nd</sup> year student,  
Faculty of Computer Science, HSE University

## **Project Supervisor**

Svetlana A. Voskresenskaia,  
Research Fellow,  
Faculty of Physics, HSE University



## About this project

Technical overview and resources

**Project Type:** Interactive astronomical research platform

### Core Technologies:

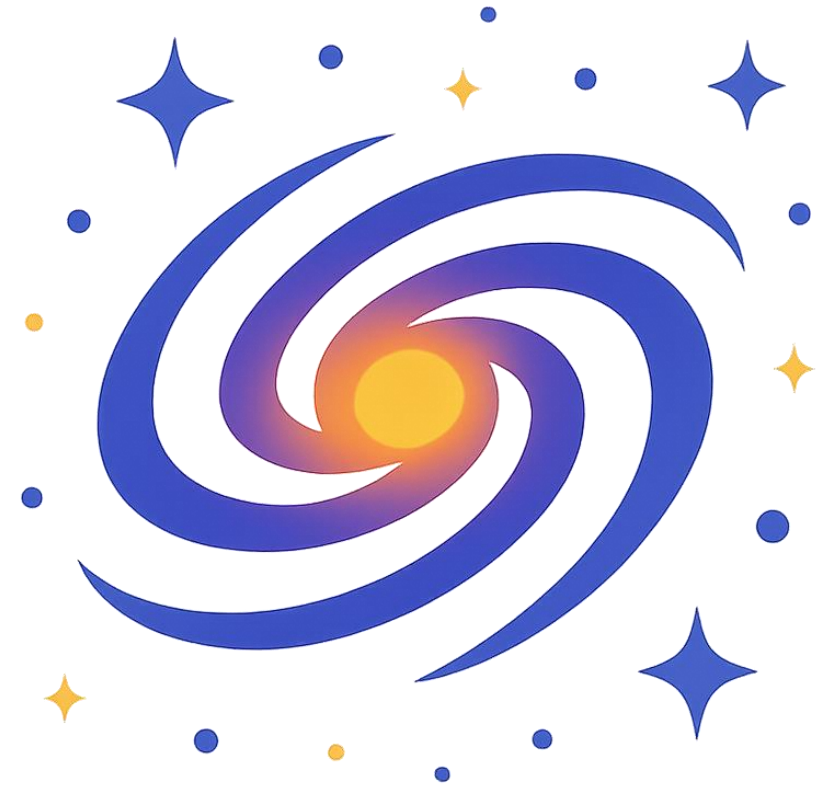
- Backend: Python, FastAPI, asyncio
- Frontend: JavaScript, TailwindCSS
- AI: Large Language Models (LLM)

**Deployment:** Docker, Nginx

**Data Sources:** SIMBAD, VizieR, NASA ADS, arXiv

**Repository:** <https://github.com/gcluster-tech>

**Development Period:** 2025 academic year



Our Project Logo



## Problem Statement

What's wrong with current workflow?

### Current Manual Workflow

- Search SIMBAD -> manually copy coordinates
- Search VizieR -> manually cross-reference object IDs
- Search ADS -> manually filter by object name
- Search arXiv -> manually validate relevance

### Our Automated Solution

- One search -> All databases simultaneously with intelligent cross-referencing, coordinate validation, and AI-powered literature filtering.



**astrophysics**  
data system



# Project Overview and System Architecture

What is this platform?

**Unified astronomical research platform** combining multiple data sources

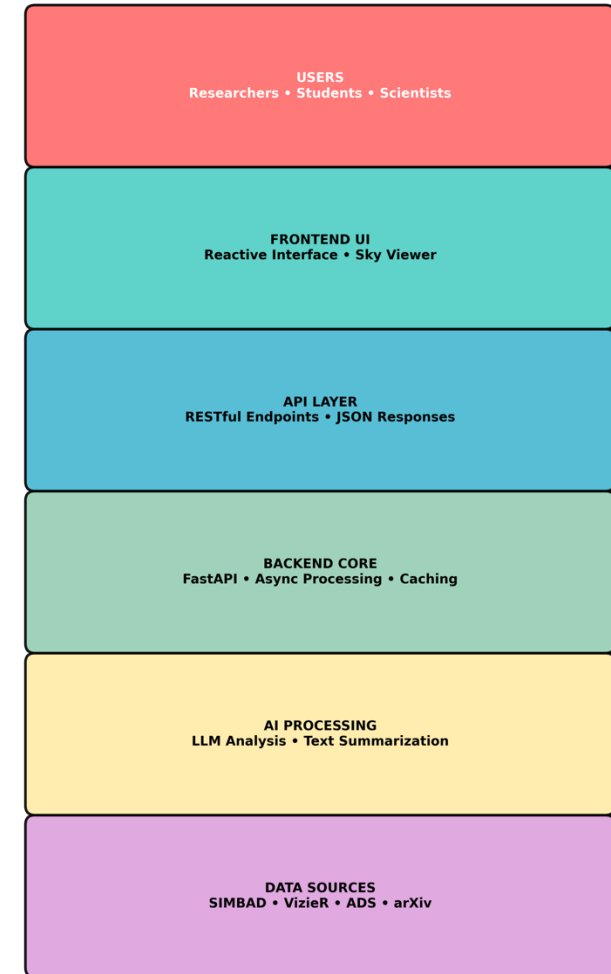
**FastAPI backend** with async/await architecture, database integrations and LLM processing for literature

**Web-based solution** with modern UI/UX design based on TailwindCSS with multi-language support and Aladin Lite sky viewer

**Real-time integration** with leading astronomical databases:

- SIMBAD - object database with identifiers & measurements
- VizieR - various astronomical catalogs
- NASA ADS - astrophysics literature search
- arXiv - preprint server for astronomy

**AI-powered** literature summarization using LLM



Tower of technologies



# Main Interface

Landing page (mobile and desktop versions)

GC

RU

## Что такое скопления галактик?

Скопления галактик — крупнейшие гравитационно связанные структуры во Вселенной, включающие сотни и тысячи галактик, горячий газ и тёмную материю.

### Методология

Узнайте о методах исследования скоплений.

### Aladin + Поиск по координатам

Интерактивный астрономический поиск с интеграцией звездного атласа.

### Статьи по координатам и типу объекта

Найдите статьи по координатам и типу объекта.

### Проекты команды и описания каталогов

Изучите проекты команды и информацию о каталогах.

Galaxy Clusters

EN RU

## What are galaxy clusters?

Galaxy clusters are the largest gravitationally bound structures in the universe, containing hundreds to thousands of galaxies, hot gas, and dark matter.

### Method

Learn about the methodology for cluster research.

### Aladin + Coordinate Search

Interactive astronomical search with sky atlas integration.

### Articles by coordinate and object type

Find articles by coordinates and object type.

### Team projects & catalog descriptions

Explore team projects and catalog info.

### Contact us

Get in touch with the team.

© 2025 Galaxy Clusters Research. All rights reserved.



# Core Functionality

What makes us different?

## 1. Intelligent Cross-Database Querying

- Automatic coordinate conversion between databases
- Smart object ID resolution across catalogs
- **vs manual copy-paste between sites**

## 2. Unified Results Correlation

- Automatic merging of duplicate entries
- Distance-based object matching
- **vs separate results from each database**

## 3. AI-Enhanced Literature Analysis

- Context-aware paper filtering by object relevance
- Automated abstract summarization with key findings
- **vs manual reading of hundreds of abstracts**

[< Back](#)

## Galaxy Clusters

### Research Methodology & Chat Bot Overview

#### Astronomical Object Search Bot

Our chat bot is a CLI-based tool designed for astronomers to quickly retrieve information about astronomical objects. It accepts either object names or coordinates and returns comprehensive data from multiple astronomical databases.

##### Input Formats

- Object names (e.g., M31, NGC 4472)
- Coordinates in degrees (10.6847 41.269)
- HMS format (00:42:44.3 +41:16:09)

##### Output Data

- Object parameters & coordinates
- Scientific articles with abstracts
- Catalog cross-references

#### External Data Sources

Our system integrates with the world's leading astronomical databases to provide comprehensive and up-to-date information:

##### SIMBAD

Astronomical object database with identifiers, basic measurements, and bibliography

##### VizieR

Access to astronomical catalogs including Abell, Planck, ROSAT, and other cluster catalogs

##### NASA ADS

Astrophysics literature search and citation database

##### arXiv

Preprint server for astronomy and astrophysics papers

## Core Functionality

What makes us different?

### Traditional Approach

#### Manual Steps (up to 15 min):

1. Open SIMBAD -> search "Vega"
2. Copy coordinates manually
3. Open VizieR -> paste coordinates
4. Check for catalog matches
5. Open ADS -> search by name
6. Filter papers manually
7. Read abstracts one by one
8. Cross-check object IDs

**Error-prone, time-consuming**



### Our Platform

#### Automated Workflow (up to 2 min):

1. Type "Vega" -> press Enter
2. **AI automatically:**
  - Queries all 4 databases
  - Validates coordinates
  - Merges duplicate results
  - Filters relevant papers
  - Summarizes key findings
3. Review unified results

**Fast, accurate, comprehensive**



# Search Result Interface

Comprehensive data display (mobile and desktop versions)

## Search Results Features:

- **Interactive tables** with sorting
- **Literature links** to papers
- **Article abstracts**

## Real Example

Search for "10 22" returns  
cleared data from NASA ADS

< Назад GC < Back

## Research Articles

### Search Articles

10 22

20 articles

Search

### Search Results

Found 107 articles (showing 20)

Sorted by publication year (newest first)

Nº1 2024  
**i-PI 3.0: A flexible and efficient framework for advanced atomistic simulations**  
**Authors:** Litman, Yair, Kapil, Venkat, Feldman, Yotam M. Y. et al.  
**Citations:** 16 **Reads:** 194  
**Journal:** Journal of Chemical Physics  
**▼ Abstract**  
Atomic-scale simulations have

## Galaxy Clusters

## Research Articles

### Search Articles

Sky background effects

15 articles

Search

### Search Results

Found 7 articles (showing 7)

Sorted by publication year (newest first)

Nº1 2025  
**Deriving physical parameters of unresolved star clusters. IX. Sky background effects in the aperture photometry**  
**Authors:** Daugevičius, Karolis, Stonkutė, Rima, Kriščiūnas, Eimantas et al.  
**Citations:** 0 **Reads:** 0 **Journal:** arXiv e-prints  
**▼ Abstract**  
The aperture photometry method is a powerful tool that enables us to study large star cluster systems efficiently. However, its accuracy depends on various factors, including the stochasticity of the stellar initial mass function and variations in the sky background. Previously, in the eighth paper of this series, we established the best achievable limits of the aperture photometry method for star cluster studies in the local universe. The aim of this study is to determine how the sky background affects the limits and applicability of the aperture photometry method in star cluster analysis. We used a large sample of star cluster models spanning the parameter space of M 31 clusters. To determine how the background affects star cluster photometry, we placed images of simulated clusters into five background fields of different stellar density from the Panchromatic *Hubble* Andromeda Treasury (PHAT) survey and measured them using aperture photometry. We determined age and mass limits for the M 31 disc star clusters at which photometric uncertainties are low enough to enable the determination of cluster parameters using the aperture photometry method. We demonstrated that for typical-size clusters, optimal aperture diameters are of ~3 half-light radii. We assessed cluster detection completeness in relation to varying sky background densities, based on the M 31 PHAT survey data. Our results suggest that a significant selection bias towards more compact clusters may exist in the PHAT survey. We derived new mass limits of the cluster mass function (CMF) in the PHAT survey reaching down to masses of 500 M<sub>☉</sub> for sizes



# AI-Powered Research Assistant

Beyond simple database queries

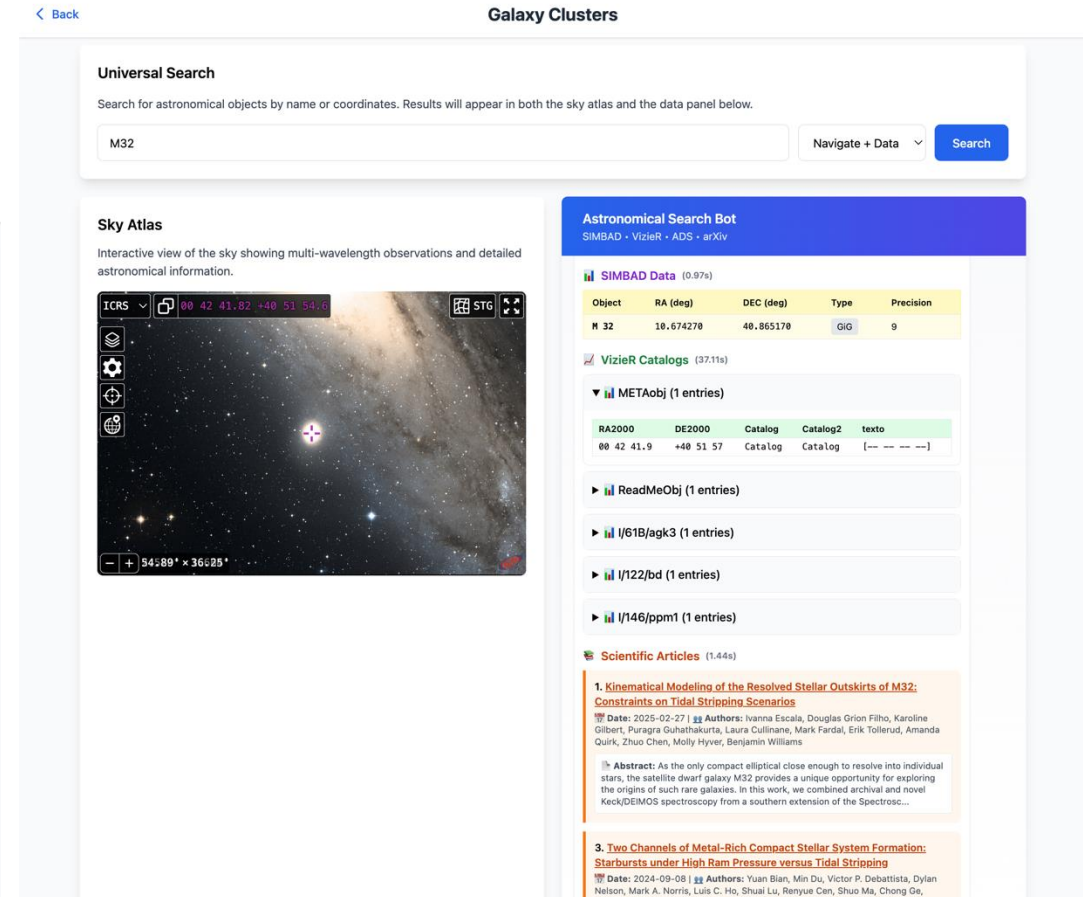
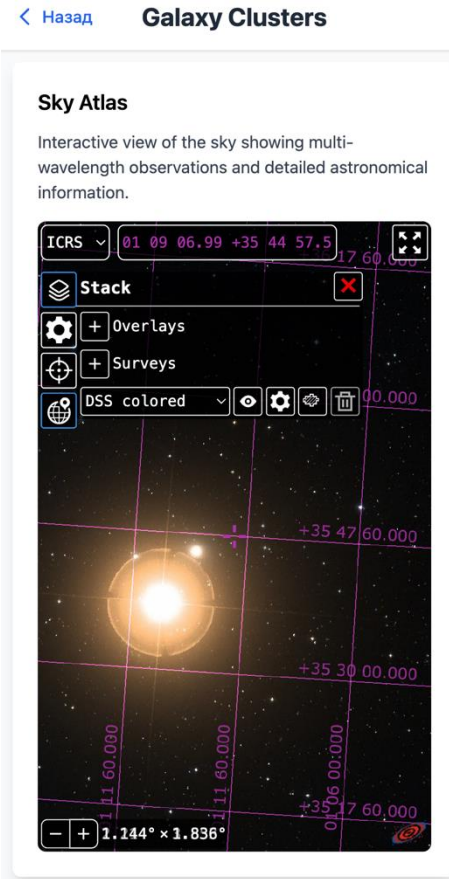
## Unique AI Capabilities:

- **Contextual Understanding** – Knows when "M31" and "Andromeda" refer to the same object
- **Smart Paper Filtering** – Distinguishes relevant papers from noise
- **Research Gap Detection** – Identifies understudied aspects of objects

## Other Features:

- **Touch-optimized** interface
- **Responsive tables** with horizontal scroll
- **Simplified navigation** for small screens
- **Fast loading** on mobile networks
- **Gesture support** for sky viewer

**Fully cross-platform** – Works seamlessly on desktop, tablet, and mobile devices with adaptive UI.





## Deployment & Scalability

Production-ready solution

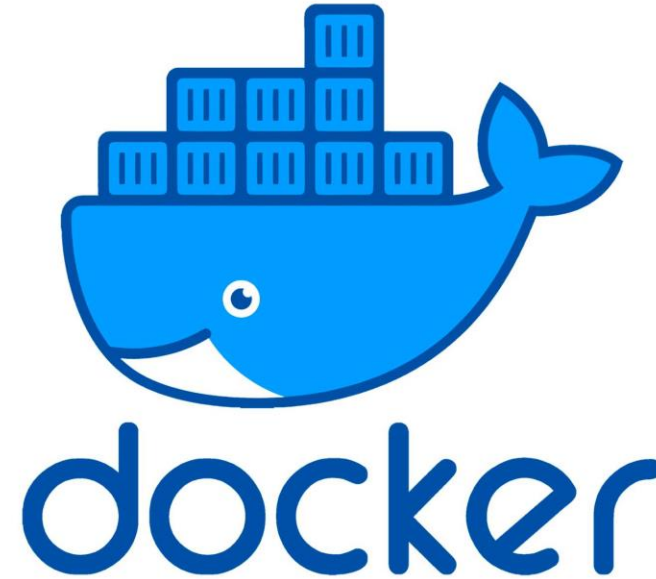
### Infrastructure

- **Docker containerization** for consistent deployment
- **Nginx reverse proxy** for load balancing
- **SSL/TLS encryption** for secure connection
- **Health monitoring** with automated checks
- **Log aggregation** for debugging & analytics

### Scalability Features:

**Horizontal scaling ready:** Load balancer + multiple backend instances

**Database optimization** by applying caching layer that reduces API usage





## Conclusion

### Key Achievements

#### Theorem (Project Success)

*Successfully created a **unified platform** that integrates major astronomical databases with AI-powered analysis, reducing research time and improving data accessibility.*

#### Technical Metrics:

- **Over 10,000 objects** accessible via unified search
- **AI-powered** literature summarization in seconds
- **Real-time cross-database** queries (SIMBAD, VizieR, ADS, arXiv)
- **Fully responsive** interface for desktop and mobile



**Live Demo:** [gclusters.tech](https://gclusters.tech)



## Questions and Discussion

Thank you for your attention!

### Contact Information:

**Student:** Arkady Pavlov

**Email:** [aspavlov@hse.ru](mailto:aspavlov@hse.ru)

**Personal GitHub:** <https://github.com/Tokollo5/>

**Supervisor:** Svetlana A. Voskresenskaia

**Email:** [savoskresenskaya@edu.hse.ru](mailto:savoskresenskaya@edu.hse.ru)



**Live Demo:** [gclusters.tech](https://gclusters.tech)

