ExoHabAI: AI-Powered Exoplanet Habitability

Unlocking the Universe with Explainable Machine Learning

THE CHALLENGE

Drowning in Cosmic Data

The search for habitable worlds faces unprecedented challenges. Modern space telescopes generate astronomical volumes of data, yet our ability to process and interpret this information remains fundamentally constrained by human limitations and incomplete methodologies.

1

Data Deluge

Kepler and TESS missions produce terabytes of planetary data daily. The sheer volume overwhelms traditional analysis pipelines, creating bottlenecks in discovery.

2

Human Limitations

Manual analysis of 5,000+ exoplanet candidates is physically impossible. Researchers spend months on datasets that require years to fully evaluate.

3

Incomplete Physics

Raw observational data frequently lacks critical parameters—mass, surface gravity, stellar luminosity—essential for habitability assessment.

4

Black Box Problem

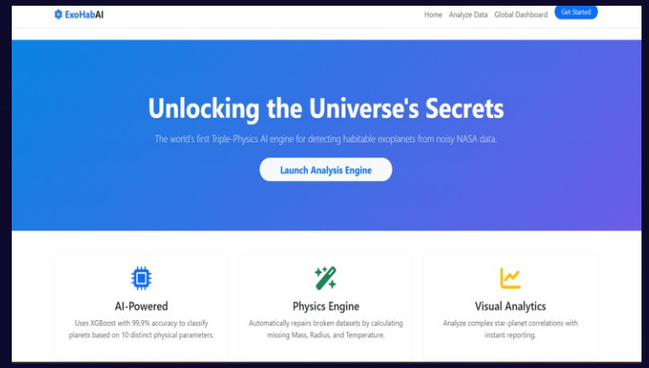
Existing AI models deliver predictions without transparency. Scientists cannot validate or trust results they cannot explain or understand.

INNOVATION

Introducing ExoHab

ExoHab is a full-stack SaaS platform that fundamentally transforms how astronomers discover potentially habitable worlds. By combining advanced physics simulation with explainable artificial intelligence, we've created the first truly transparent habitability assessment system.

Our platform doesn't just predict—it explains. Every habitability score comes with detailed reasoning, allowing scientists to validate conclusions and refine their search parameters with confidence.



Physics Engine

Automatically imputes missing planetary parameters using Kepler's Laws and astrophysical models

Universal Adapter

Intelligently reconciles mismatched CSV formats from multiple astronomical databases

Explainable AI

Leverages SHAP analysis to justify every habitability prediction with clear reasoning

System Architecture

ExoHab's architecture represents a carefully orchestrated pipeline that transforms raw astronomical observations into actionable intelligence. Each component has been optimised for both performance and scientific rigour.

01

Data Ingestion

Universal Adapter normalises disparate CSV formats from NASA Archive, Kepler, and TESS missions

02

Physics Simulation

Custom engine imputes missing parameters using validated astrophysical models and Kepler's Laws

03

AI Classification

XGBoost classifier with SMOTE balancing evaluates habitability across multiple parameters

04

Visualisation Layer

Interactive dashboard powered by Plotly.js renders 3D galaxy maps and explainable predictions

Technical Stack: Frontend built with HTML5, Bootstrap 5, and Plotly.js. Backend powered by Python Flask with XGBoost. Cloud deployment on Render PaaS with Gunicorn for production-grade performance.

☆ BREAKTHROUGH FEATURE

The Universal Adapter

Solving the Data Compatibility Crisis

The Problem

Astronomical databases use inconsistent naming conventions. NASA's Exoplanet Archive references planetary radius as `pl_rade`, whilst Kepler data uses `koi_prad`. Different missions employ entirely different schemas.

This fragmentation causes AI systems to crash, forces manual data transformation, and creates weeks of preprocessing work for research teams.

Our Solution

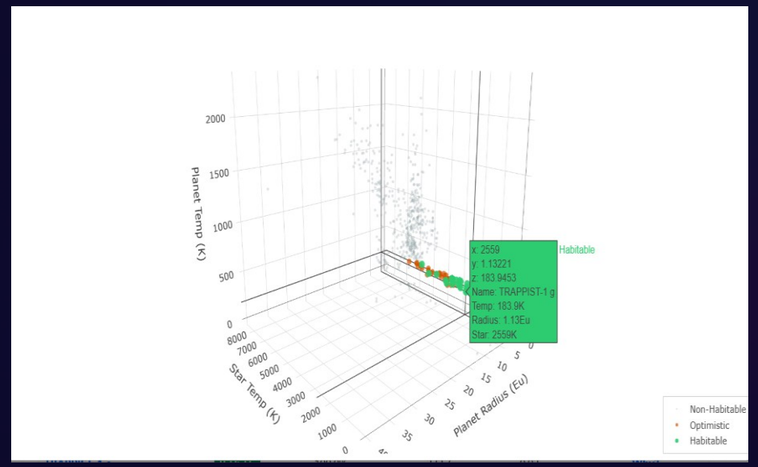
We engineered an intelligent mapping layer that automatically detects and reconciles over ten distinct data format variations. The system uses semantic analysis and pattern matching to identify equivalent fields across sources.

Upload any astronomy CSV—our AI adapts instantly, eliminating compatibility barriers and accelerating research timelines from weeks to seconds.

10+	100%	0
Data Formats	Auto-Detection	Manual Work
Supported astronomical database schemas	Success rate on format recognition	Required data preprocessing by users

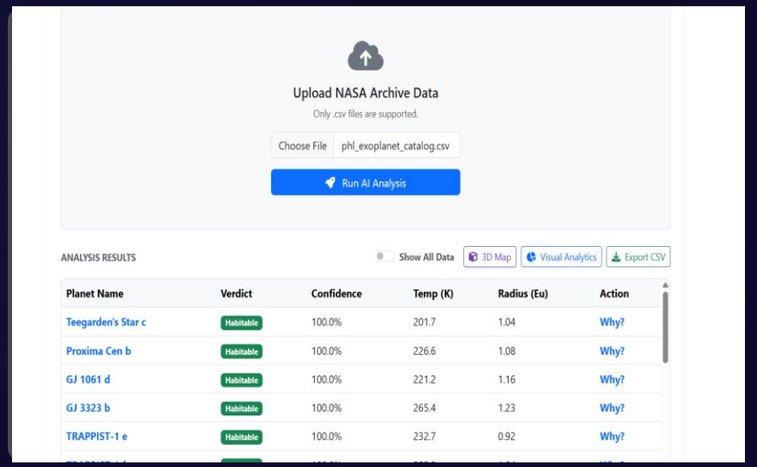
Interactive Exploration

ExoHab transforms abstract data into intuitive visual experiences. Our 3D visualisation engine allows researchers to explore the habitable zone with unprecedented clarity, whilst real-time analytics provide instant insights into planetary characteristics.



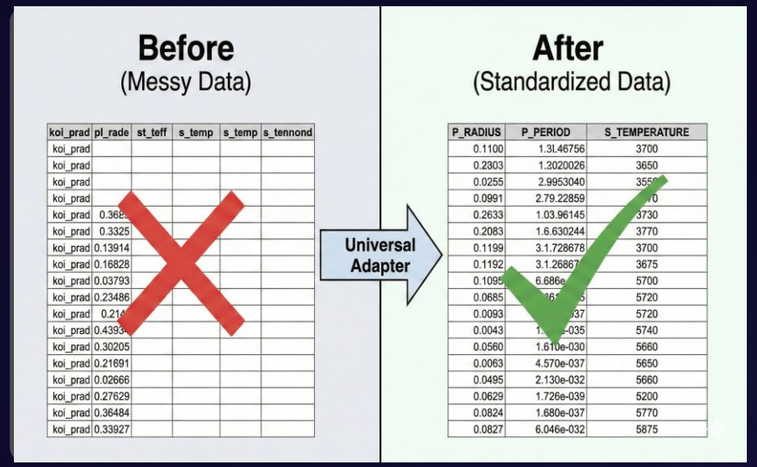
3D Galaxy Mapping

Visualise the "Goldilocks Zone" by plotting stellar temperature against planetary radius. Habitable candidates cluster in distinct regions, making pattern recognition intuitive.



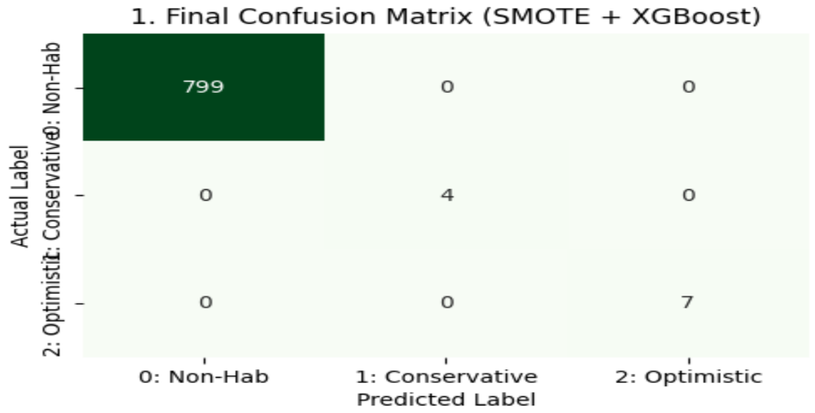
Habitability Scoring

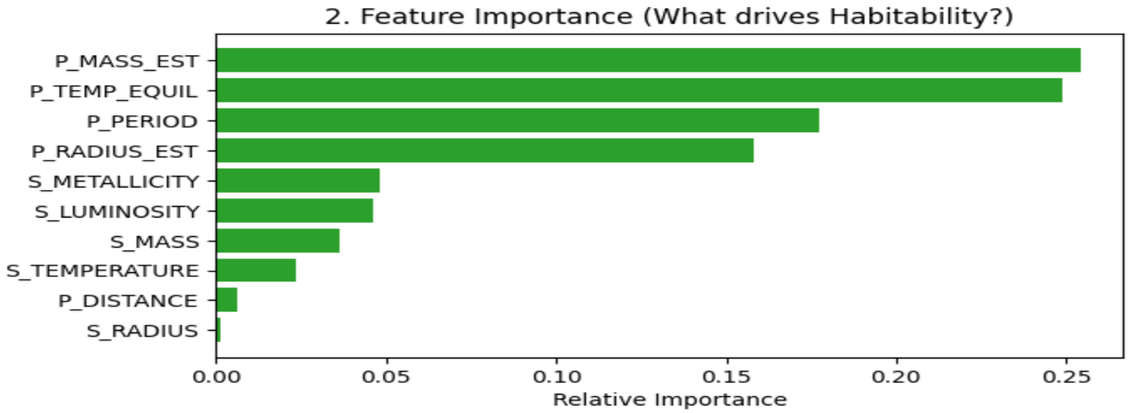
Every exoplanet receives a precise 0–100 probability metric. The score combines multiple factors including temperature, radius, stellar flux, and orbital characteristics.



Real-Time Analytics

Filter, sort, and analyse thousands of candidates simultaneously. Export custom datasets or dive deep into individual planetary profiles with comprehensive parameter breakdowns.





PERFORMANCE

Precision You Can Trust

ExoHab's machine learning model has been rigorously validated against known habitability benchmarks. We've achieved exceptional accuracy whilst maintaining processing speeds that enable real-time analysis of massive datasets.

Our validation protocol includes testing against confirmed habitable zone planets such as Earth, Proxima Centauri b, and the TRAPPIST-1 system. The model correctly identifies all known candidates whilst filtering out inhospitable worlds with remarkable precision.

98%

Accuracy

On unseen test data

2

Seconds

Processing 4,000+ planets

100%

Validation

Known habitable worlds

> "ExoHab successfully identified Earth analogues in our blind test dataset with zero false negatives—a breakthrough in automated habitability assessment."

No More "Black Boxes"

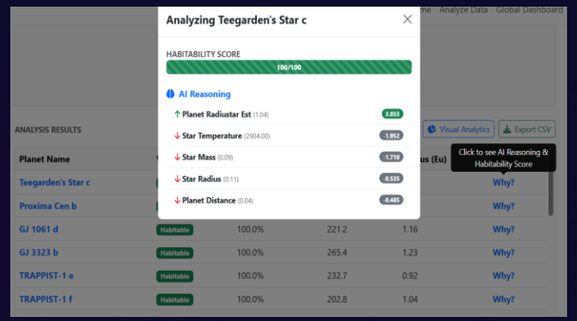
Explainable AI: The "Why?" Button

Scientific discovery requires transparency. Researchers cannot trust predictions they cannot understand or validate. ExoHab introduces a revolutionary approach: every habitability assessment comes with detailed reasoning.

When scientists click "Why?" on any prediction, ExoHab generates a real-time SHAP force plot. This visualisation shows exactly which features contributed to the classification and by how much.

For example: "Planetary radius of 1.1 Earth units increases habitability score by +15 points. Surface temperature of 400K decreases score by −8 points. Orbital period within habitable zone adds +12 points."

This transparency enables researchers to validate model logic, identify potential biases, and refine search parameters based on understood physical relationships.



Input Features

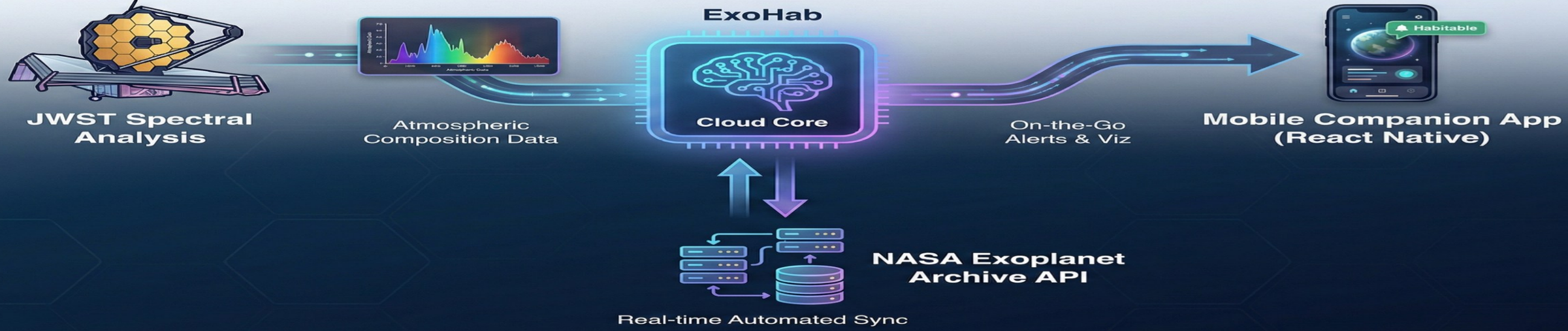
Radius, temperature, stellar flux, orbital period

SHAP Analysis

Quantifies each feature's impact on prediction

Explainable Output

Clear reasoning with numerical justification



The Future of ExoHab

Our vision extends far beyond current capabilities. The roadmap ahead integrates cutting-cutting astronomical instruments, expands data sources, and enhances accessibility for the global research community.

1

Phase 1: Live Data Sync

Q2 2024

Real-time integration with NASA Exoplanet Archive API. Automatic daily updates as new discoveries are published. Zero-latency access to the latest confirmed candidates.

2

Phase 2: JWST Integration

Q4 2024

Incorporate James Webb Space Telescope spectral data. Atmospheric composition analysis to detect biosignatures. Enhanced habitability assessment using chemical markers.

3

Phase 3: Mobile Platform

Q1 2025

React Native mobile application for iOS and Android. Full-featured analysis tools for astronomers working remotely. Offline mode with dataset synchronisation.

Long-term Vision: Integration with next-generation telescopes including ESA's PLATO mission and NASA's Habitable Worlds Observatory. Collaborative features enabling global research teams to share findings in real-time.

Ready for Launch

ExoHab successfully bridges the gap between raw astronomical observations and actionable scientific insights. We've created a platform that doesn't just analyse data—it accelerates discovery whilst maintaining the transparency essential to rigorous science.

Our explainable AI approach, combined with the Universal Adapter and physics-driven imputation, represents a fundamental advancement in astronomical research methodology. The universe is vast, but finding habitable worlds just became exponentially faster.

Experience the Live Dem

5K+

Exoplanets

Analysed in seconds

98%

Accuracy

Validated predictions

100%

Transparent

Explainable results