Empowering your AI Missions through Independent Validation & Testing

An Introduction to CalypsoAI
CalypsoAI was established to solve the big challenges facing AI deployment today

AI Experience coming from the US National Security community, and key investors who invest across AI categories

Customers and Strategic Relationships:

Founded: 2018
Locations: San Mateo, CA and Virginia
Stage: Series A; Series B exp. late 2022
Investors: Paladin Capital, Lightspeed Ventures, 8VC, and Lockheed Martin Ventures.
Contract Vehicles: JAIC T&E BPA, NASA SEWP, GSA, Army CHESS ITES-SW2, other contracts available upon request
Deployment Options: Built on Kubernetes, we support containerized deployment and are hardware/platform agnostic. Available via AWS GovCloud and on-prem

Named Gartner®
Cool Vendor
AI Core
Technologies 2022
These issues are exacerbated when:

- Models are trained on limited data that does not represent the deployment environment
- Models are not developed to withstand rapid change in environments or real-world conditions
- Models are at risk of strategic adversarial noise injection
- Model test and evaluation is not automated
- There is a lack of model version control
- There is a lack of model performance standardization
The solution that builds Trust in your AI adoption by independently validating and testing your AI/ML models.

Confidently enable and accelerate your Mission!
Our Product: VESPR Validate

Ensuring your AI/ML can achieve organizational goals, securely, in real-world conditions.

**Stress Test Real-World Performance**

Utilizing 3D maps, gaming engines, physics-based simulations, and quantified noise distributions mirroring real-world data gaps we test models in adverse environments to provide confidence of accurate performance in operational environments. These include:

- Weather Conditions
- Blur
- Brightness
- Defocus

**Inversion / Privacy Testing**

Performing rapid systematic attacks on the model to inference sensitive training data, we determine if this data is secure.

**Adversarial Security**

We use cutting edge adversarial attacks on the model to trigger model failure utilizing the Minimal Attack Surface to test model vulnerability to adversarial image attacks.
Case Study
Automated Target Recognition (ATR) from MQ-9 Full Motion Video

Data: Full Motion Video (FMV)
Source: MQ-9 Reaper
Data Type: Infrared
Target: Tank
Model Type: Pytorch Classifier
Number of Classes: 11

Vendor Model Performance Metrics
F1 Score .58
Global Accuracy .59
Precision .61
Recall .59
**Case Study**

Sail Drone, Automated Identification of Iranian State Actors, Faris Island

**Location:** 27.9900° N, 50.1700° E  
Near: Farsi Island, W, NW

**Data:** Full Motion Video (FMV)  
**Source:** Unmanned Sail Drone  
**Data Type:** Optical, Infrared  
**Target:** Iranian Patrols, Small Boats  
**Model Type:** Pytorch ResNet V1.5  
**Number of Object Classes:** 11

**Vendor:** Model Performance Metrics  
**F1 Score:** .87  
**Global Accuracy:** .86  
**Precision:** .9  
**Recall:** .89
VESPR Validate works across the MLOps pipeline and can easily integrate with MLOps tools such as but not limited to Azure Machine Learning, Scalabel, DataRobot, Dataiku, Arize AI, and many more.

VESPR Validate offers critical automated Test, Evaluation, Validation & Verification (TEVV) components to enable organizations to create a robust MLOps platform that ensures models function correctly in operational environments characterized by rapid change, adversarial activity, and varying mission profiles.