Project Introduction

Artificial Intelligence (AI) algorithms, which are the heart of emerging aviation autonomous systems and autonomy technologies, are generally perceived as black boxes whose decisions are a result of complex rules learned on-the-fly. Unless the decisions are explained in a human understandable form, the human end-users are less likely to accept them, and in the case of aviation applications, certification personnel are less likely to clear systems with increasing levels of autonomy for field operation. Explainable AI (XAI) are AI algorithms whose actions can be easily understood by humans. This SBIR develops Explained Process and Logic of Artificial Intelligence Decisions (EXPLAIND), which is a prototype tool for verification and validation of AI-based aviation systems. The SBIR develops an innovative technique called Local Interpretable Model-Agnostic Explanation (LIME) for making the learning in AI algorithms more explainable to human users. LIME generates an explanation of an AI algorithm’s decisions by approximating the underlying model in the vicinity of a prediction by an interpretable one. We apply LIME to a NASA-developed aircraft trajectory anomaly detection AI algorithm (MKAD) to provide a proof-of-concept. EXPLAIND represents an important step towards user acceptance and certification of multiple AI based decision support tools (DSTs) and flight-deck capabilities planned to be developed under NASA's System Wide Safety and ATM-eXploration projects. EXPLAIND also benefits NASA's planned human-in-the-loop (HITL) simulations of machine learning (ML) algorithms using the SMARTNAS Testbed by providing techniques for making the algorithm's decisions more understandable to HITL participants. Moreover, with new European Union regulations soon requiring that any decision made by a machine be readily explainable, the EXPLAIND approach is also relevant to multiple non-aviation fields such as medical diagnosis, financial systems, computer law, and autonomous cars.

Anticipated Benefits

Applications include enhanced explainability AI/ML algorithms for (1) aviation anomaly detection and safety precursor identification for Real-time System-wide Safety Assurance, (2) ATD-3’s Traffic Aware Strategic Aircrew Requests (TASAR), (3) IDO traffic management, (4) UAM and UTM path planning, de-confliction, scheduling and sequencing, (5) AI explanation interfaces to support UAM and IDO HITLs using the SMARTNAS Testbed, (6) Science Mission Directorate’s distant planet discovery algorithms.

Primary application is for the FAA with goal for NASA to transfer a validated enhanced explainability AI concept (e.g., MKAD anomaly detection tool) to the FAA. Airline AI travel assistant tools are another application. With new European General Data Protection Regulation requiring that any decision made by a machine be readily explainable, EXPLAIND can be applied to non-aviation
fields like financial credit models, medical diagnosis, and self-driving car guidance systems.

**Primary U.S. Work Locations and Key Partners**

<table>
<thead>
<tr>
<th>Organizations Performing Work</th>
<th>Role</th>
<th>Type</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ames Research Center (ARC)</td>
<td>Lead Organization</td>
<td>NASA Center</td>
<td>Moffett Field, CA</td>
</tr>
<tr>
<td>ATAC</td>
<td>Supporting Organization</td>
<td>Industry</td>
<td>Santa Clara, CA</td>
</tr>
</tbody>
</table>

**Primary U.S. Work Locations**

California

**Organizational Responsibility**

**Responsible Mission Directorate:**
Space Technology Mission Directorate (STMD)

**Lead Center / Facility:**
Ames Research Center (ARC)

**Responsible Program:**
SBIR/STTR

**Project Management**

**Program Director:**
Jennifer L Gustetic

**Program Manager:**
Carlos Torrez

**Principal Investigator:**
Aditya Saraf

**Technology Maturity (TRL)**

Start: 1
Current: 3
Estimated End: 3

1 2 3 4 5 6 7 8 9

Applied Research Development Demo & Test

Closeout Documentation

Final Summary Chart
(https://techport.nasa.gov/file/37617)
SBIR/STTR

Explainable Artificial Intelligence based Verification & Validation for Increasingly Autonomous Aviation Systems, Phase I
Completed Technology Project (2018 - 2019)

Images

Project Image
(https://techport.nasa.gov/image/34838)

Technology Areas

Primary:
- TX10 Autonomous Systems
  - TX10.4 Engineering and Integrity
  - TX10.4.1 Verification and Validation of Autonomous Systems

For more information and an accessible alternative, please visit:
https://techport.nasa.gov/view/94579