



Public report

Release of GitHub CAVAA repository and community engagement (D6.4)

Action number: [101071178](#)

Action Acronym: [CAVAA](#)

Action title: [Counterfactual Assessment and Valuation for Awareness Architecture](#)

Author(s): [Oscar Guerrero Rosado \(Radboud University\)](#)

Submission date: [01. September 2024](#)

The project document is published on www.cavaa.eu.



Funded by the
European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EISMEA – European Innovation Council and SMEs Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

European
Innovation
Council



Table of content

1. Introduction	1
2. Consortium.....	2
3. Description of work performed.....	2
3.1. Partners involved.....	2
4. Description of work performed.....	2
4.1. Repository overview.....	3
4.2. Computational models.....	3
5. Results.....	3
6. Conclusion, next steps.....	3

1. Introduction

The Counterfactual Assessment and Valuation for Awareness Architecture (CAVAA) project addresses the challenge of developing AI systems capable of awareness, a concept essential for survival in complex environments where many factors remain hidden. CAVAA aims to create an integrated computational architecture that combines perception, memory, and the simulation of potential future states to form a virtual representation of the world. This architecture is not only meant to explain awareness in biological systems but also to engineer it in technological systems, such as robots and artificial agents.

The release of the CAVAA GitHub repository marks a significant milestone in this effort. It provides access to the initial computational models developed under the project, which form the foundation for further research and development. These models will be used to explore use cases such as robot foraging, social robotics, and human-AI interactions, ultimately contributing to AI systems that are more explainable, adaptable, and trustworthy.

The repository is publicly accessible at the following link: <https://github.com/SPECS-Lab/CAVAA>.

This deliverable reports on the progress made in establishing the repository and the contributions from the Donders Institute for Brain, Cognition and Behaviour (Radboud University), which has led the initial phase of development. As the project progresses, contributions from other technical partners will be integrated to expand the repository's capabilities.

2. Consortium

Radboud University	Radboud University – Coordinator (NL)
CERTH CENTRE FOR RESEARCH & TECHNOLOGY HELLAS	CERTH - The Centre for Research & Technology, Hellas (GR)
UNIVERSITY OF TECHNOLOGY CHEMNITZ	University of Technology Chemnitz (DE)
SORBONNE UNIVERSITÉ	Sorbonne University (FR)
eodyne	Eodyne Systems (ES)
Robotnik	Robotnik (ES)
UPPSALA UNIVERSITET	Uppsala University (SE)
tp21	tp21 GmbH (DE)
UNIVERSITY OF OXFORD	University of Oxford (UK)
University of Sheffield	University of Sheffield (UK)

3. Description of work performed

The CAVAA repository has been established as a central hub for sharing the project's computational models and resources. In this initial phase, the Donders Institute for Brain, Cognition and Behaviour has been the primary contributor, developing key models that underpin the reactive and adaptive layers of the CAVAA architecture. These layers are critical for simulating the awareness processes that the project aims to model.

3.1. Partners involved

STICHTING RABDOUD UNIVERSITEIT - SRU (Lead Beneficiary):

- Responsible for the initial development and implementation of computational models in the repository.
- Focused on creating foundational models related to WP1 including allostatic control, perception, memory formation, and initial virtualization strategies.

Other partners, including **CERTH**, **University of Technology Chemnitz**, **Sorbonne University**, **Robotnik**, and **University of Sheffield**, are expected to contribute in the next stages, enhancing the repository with additional models and functionalities.

4. Description of work performed

The CAVAA GitHub repository has been structured to facilitate easy access and collaboration. It is designed to be modular, allowing for the seamless integration of new models and updates as the project evolves.

4.1. Repository overview

The repository contains a main branch where the core computational models are stored. Each model is accompanied by comprehensive documentation, including setup instructions, usage examples, and explanations of the underlying algorithms. The repository is organized to support contributions from multiple partners, with clear guidelines for adding new models and improving existing ones.

4.2. Computational models

- **Hypothalamus Model:** Contributing to the reactive layer, this model simulates the hypothalamus's role in regulating internal states based on external stimuli. It includes both a multiattractor version, which allows for the simulation of multiple internal needs, and a simplified version, which focuses on two primary needs. These models are crucial for enabling the agent to respond dynamically to changes in its environment, ensuring survival and goal achievement.
- **Hippocampal Model:** This model is a key component of the adaptive layer, simulating the role of the hippocampus in spatial memory and navigation. It uses a sparse convolutional autoencoder to create high-dimensional embeddings that capture spatial information, mimicking the formation of place cells. The model is designed to help the agent navigate its environment by forming and recalling spatial maps.
- **Sequential Episodic Control (SEC):** SEC leverages episodic memory for decision-making, integrating short-term and long-term memory processes to select actions based on past experiences. This model is particularly important for tasks that require the agent to recall and apply learned experiences to new situations, enhancing its ability to adapt to complex and changing environments.

5. Results

The release of the CAVAA GitHub repository has successfully established a foundation for ongoing research and development in the project. The initial computational models developed by the Donders Institute have been integrated into the repository, providing the necessary tools for simulating awareness in AI systems. These models have been thoroughly documented to ensure that they are accessible and usable by the wider research community.

Key outcomes include:

- **Foundation for Further Development:** The repository now contains the essential models needed to begin exploring the CAVAA architecture's capabilities.
- **Documentation and Usability:** Each model is accompanied by detailed documentation, making it easier for other researchers to use and build upon these models.
- **Community Engagement:** The repository is open for contributions, with guidelines in place to facilitate collaboration among partners and the broader research community.

6. Conclusion, next steps

The first phase of the CAVAA repository development has laid a solid foundation for the project's objectives. The next steps involve integrating contributions from the other technical partners — CERTH, University of Technology Chemnitz, Sorbonne University, Robotnik, and University of Sheffield — who will add new models and enhance the existing ones. These additions will focus on more

advanced aspects of the CAVAA architecture, such as emulating alternative realities and virtualized representations, which are critical for achieving the project's goals.

The ongoing development and refinement of the CAVAA repository will be essential for advancing the project's aim of creating AI systems capable of awareness. As the project progresses, the repository will continue to be a central resource for sharing advancements and fostering collaboration within the research community.

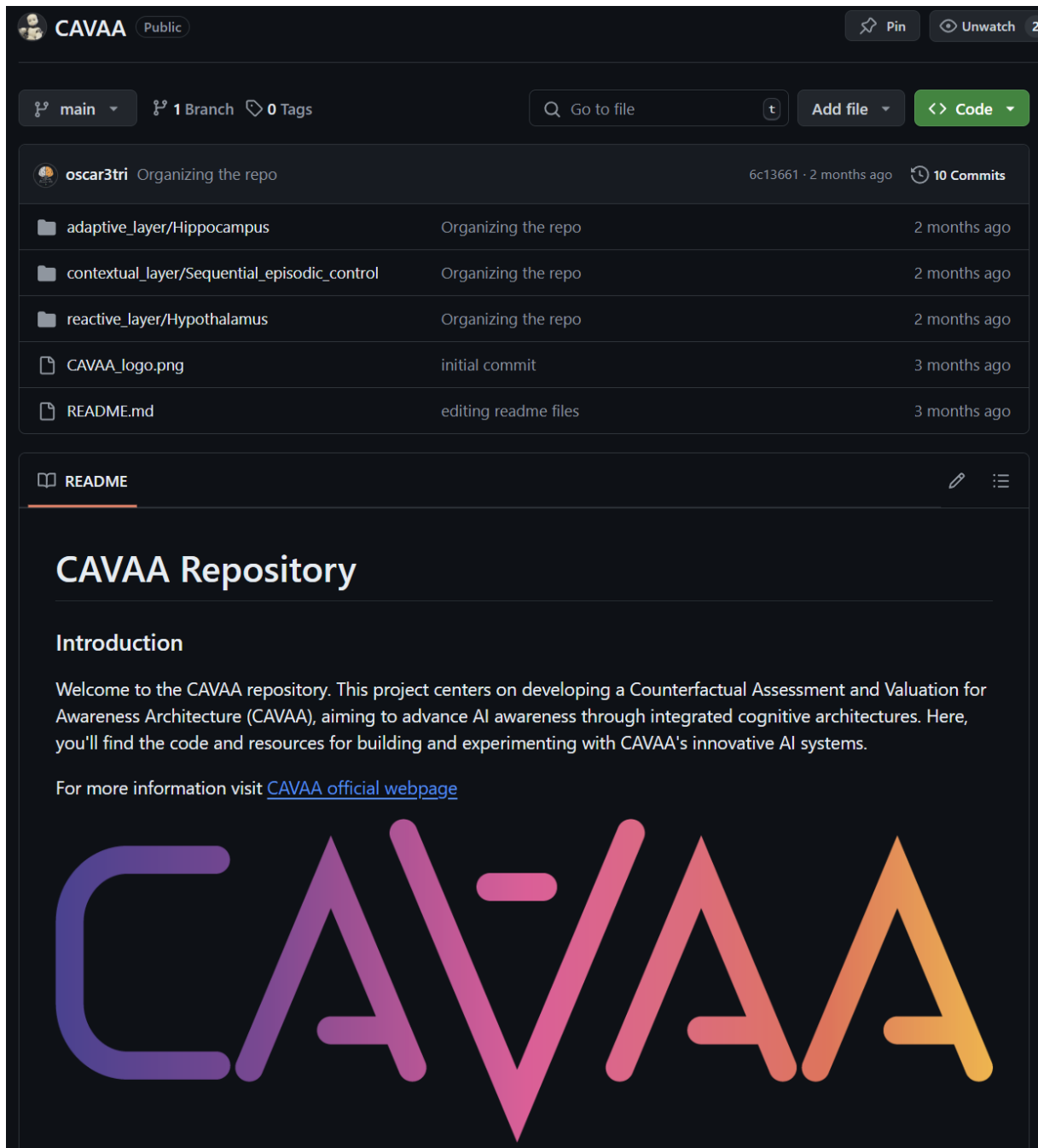


Figure 1: CAVAA GitHub repository.