

D5.6 - Embedding of ethics and cultural concerns into AML based world models - Initial

This document is deliverable (D) 5.6 of the ALMA project, *Embedding of ethics and cultural concerns into AML-based world models – Initial*.

The research reported here addresses an everyday example of interactions between what should be done ideally (normative ethics) and what is done in practice (behavioral ethics). In particular, between the normative ethics of road traffic regulations and the behavioral ethics of vehicle drivers. Road traffic regulations are an example of everyday culturally-bound social choices for the implementation of normative ethics. This is done through, for example, traffic lights that indicate where and when drivers must stop their vehicles to reduce the risk of road accidents. For roads in sparsely populated areas, there may be few traffic lights. By contrast, for roads in densely populated areas, there can be many traffic lights. This increase in the number of traffic lights in relation to increasing population density is a manifestation of the culturally-bound social choice to reduce risks to human life from road vehicles by regulating the behaviour of drivers.

However, despite one hundred years of increasing traffic regulations, there continue to be road traffic accidents. Predicting the violation of traffic regulations, such as driving through red traffic lights, is an established topic in machine learning. In this deliverable, we report our research into the potential for AML to improve upon existing machine learning applications. We applied AML to an existing traffic violation dataset. Our findings indicate that further development of AML is needed to reach the performance of current state of the art machine learning methods. This document comprises six sections. First, there is an introduction section. Second, the world models construct is explained. Third, machine learning world models are discussed and related to AML. Fourth, the embedding of ethics and cultural concerns in AML-based world models is discussed in relation to road traffic regulation. Fifth, application of AML to predict violation of road traffic regulations is reported. In conclusion principal findings are stated and directions for further research are proposed in the final section.

After this deliverable, concluding research in Work Package 5 will be reported in *D5.7 Embedding of ethics and cultural concerns into AML based world models - Final*. In that future deliverable, we shall report our further research into the potential of Algebraic Machine Learning (AML) to improve interactive human-centric machine learning systems in complex settings.