Assessing Trustworthiness of Autonomous Systems

Greg Chance[1], Kerstin Eder[1]

Trustworthy Systems Lab, University of Bristol, Bristol, UK

***Abstract*–As autonomous systems (AS) become more ubiquitous in society, more responsible for our safety and our interaction with them more frequent, it is essential that they are trustworthy. Assessing the trustworthiness of AS is a major challenge for the verification and development community (practitioners and researchers). Assessing trustworthiness must extend beyond conventional verification and validation (V&V) and safety-critical systems assurance, and now consider the the manner in which people will interface and interact with AS across the broad range of current and future artificial intelligence (AI) applications. The meta-expression 'trustworthiness' is examined in the context of AS, capturing and condensing the current understanding in the literature. A list of challenges are presented in the form of a process that can be used as a trustworthiness assessment framework for AS.**

1 Introduction

The use of AI-enabled and autonomous systems is pervasive in current society and is set to become even more so with current technological trends. Systems with embedded AI and machine learning (ML) algorithms can be found in numerous applications from mobile phones and vacuum cleaners to medical diagnostics [1].

Verification and validation are important to earning trust and gaining confidence in the safety of AS.

In the following, related work is reviewed in Section

2 Related Work

3 Conclusion

References

[1] I. Kononenko. "Machine learning for medical diagnosis: history, state of the art and perspective". In: *Artificial Intelligence in medicine* 23.1 (2001), pp. 89–109.

Statements about authorship contribution. Greg Chance (e-mail: greg.chance@bristol.ac.uk), and Kerstin Eder (e-mail: kerstin.eder@bristol.ac.uk) are with the Trustworthy Systems Lab, Department of Computer Science, University of Bristol, Merchant Ventures Building, Woodland Road, Bristol, BS8 1UQ, United Kingdom.

`https://medium.com/gobeyond-ai/artificial-intelligence-ai-in-mobile-phones-is-it-a-good-thing-fe044f20ea6c`

`https://blog.tensorflow.org/2020/01/ecovacs-robotics-ai-robotic-vacuum.html`