ACAI 2021
The advanced course on AI on Human Centered AI

Tutorial Outline

Operationalising AI Ethics: Conducting Socio-Technical Assessment
Andreas Theodorou, Umeå University & VeRAI AB; Virginia Dignum, Umeå University & VeRAI AB

Timetable:

09:00 - 10:45	Introduction to ESLEC issues
11:00 - 13:00	Protostrategos serious game
14:00 - 14:30	Protostrategos serious game wrapup
14:30 - 17:30	Context-specific AI Assessment
17:30 - 18:00	Wrapup discussion / reflections

Reading list:

Protostrategos case **(must be read per instructions on the document).** The document will be made available by the end of September at: https://umeauniversity-my.sharepoint.com/:f:/g/personal/anth2768_ad_umu_se/EnuhwD8Fqj5KtWjr4y2ec3wB6YTkbzEoy8Y0tgrN7f6gSg?e=q5EGfQ

Policy documents:

- European Commission's High-Level Expert Group on AI "Ethics Guidelines for Trustworthy AI" Available at: https://wayback.archive-it.org/12090/20210728013426/https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
- UNESCO's "Recommendations on AI" Available at: https://en.unesco.org/artificial-intelligence/ethics

Required readings:

- Winfield A.F.T., Booth S., Dennis L.A., Egawa T., Hastie H., Jacobs N., Muttram R.I., Olszewska J.I., Rajabiyazdi F., Theodorou A., Underwood M.A., Wortham R.H. and Watson E. (2021) IEEE P7001: A Proposed Standard on Transparency. Front. Robot. AI 8:665729. https://doi.org/10.3389/frobt.2021.665729
- Aler Tubella, A., Theodorou, A., Dignum, F., & Dignum, V. (2019). Governance by Glass-Box: Implementing Transparent Moral Bounds for AI Behaviour. Proceedings of the 28th International Joint Conference on Artificial Intelligence. Presented at the 28th International Joint Conference on Artificial Intelligence (IJCAI-19), Macao, China, August 10-16, 2019. https://doi.org/10.24963/ijcai.2019/802

Optional readings:

- van de Poel I. (2013) Translating Values into Design Requirements. Available at: http://ndl.ethernet.edu.et/bitstream/123456789/74675/1/14.pdf#page=261

- Aler Tubella, A., Theodorou, A., Dignum, V., & Michael, L. (2020). Contestable Black Boxes. Rules and Reasoning. RuleML+RR 2020., 159–167. https://doi.org/10.1007/978-3-030-57977-7_12
- Winfield, A. F., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. Philosophical Transactions A: Mathematical, Physical and Engineering Sciences, 376(2133). https://doi.org/10.1098/rsta.2018.0085
- Bryson, J., & Winfield, A. (2017). Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. Computer, 50, 116-119. https://doi.org/10.1109/MC.2017.154