

# Hybrid information systems: who is in control?

Abstract submission to IS4SI 2021 DIGHUM Workshop

**Daniel Boyd**

**Independent researcher**

The main motivation behind Digital Humanism is the concern that, unless carefully controlled, digital technologies could develop in ways that are detrimental to individual humans and to humanity as a whole. Preventing such scenarios requires proactive and effective intervention. The threat is often framed as that of digital technologies (or their owners) exploiting the populace to further their own interests. In this adversarial scenario, what is required is a shift of control of technological development to instead represent the interests and rights of the people. However, as the level of interaction between technologies and their users intensifies the demarcation between the two blurs. It is then no longer appropriate to consider the situation as one involving two separate and competing entities. Instead, they together form a single integrated system whose behaviour is the emergent product of the properties of the component parts and the interactions between them. Neither component is individually in control. Steering the outcome in an ethically desirable direction is consequently not possible using unilateral measures directed at just one of the participating components.

Influencing such hybrid systems necessitates a deeper understanding of the way they function. This is hampered by the very different systemic nature of the two components. On the one hand we have designed digital technologies; on the other self-organising human societies. Traditionally, these two realms of knowledge employ very different methods and theoretical models and are therefore difficult to unify. To this we can add as a third category Artificial Intelligence systems whose function is not the product of direct programming but of various forms of autonomous and directed training, evolution and development. Influencing the behaviour of AI systems requires a different approach to that used for traditional programmed computer software.

This article considers such hybrid systems from the perspective of Emergent Information Theory, which provides an explanatory framework for the origins and functions of both designed technological and evolved biological information based systems. This commonality permits extension to the hierarchical levels of organisation above the level of the physically separate components. In the simplest case, this can be applied to the system created by interaction between a single human and the digital technologies this person uses. However, most of the concerns of Digital Humanism apply to societal or global systems. In these situations both the digital and human sides of the equation are themselves emergent systems composed of many parts: a society is more than a collection of independent individuals; large scale digital systems are composed of many different functional elements. This intermediate level of organisation therefore also needs to be considered. On the basis of the unified perspective offered by Emergent Information Theory the systemic challenges of controlling systems that combine designed and self-organising components are discussed. The conclusions drawn provide theoretical foundations to support the realisation of the goals of the Vienna Manifesto.