A prototype model for ethically aligned design in Construction and Construction A/IS

Philip McAleenan

Expert Ease International, Northern Ireland

Correspondence: expertease@confinedspaces.com

Abstract

Autonomous and intelligent systems (A/IS) (also known as artificial intelligence) are recent, developing and expanding technologies in use in construction and are the defining element of the 4th Industrial Revolution. New technologies bring promise to society but unless the wider social implications of their implementation are considered they also have the potential to negatively impact and change society in unintended and undesirable ways.

The use of A/IS raises moral and ethical questions relating to privacy, agency, data protection, etc., as well as the potential for unforeseen consequences in the hands of both benevolent and disingenuous actors. Ethically aligned design is a comprehensive set of recommendations developed by IEEE to guide designers and end users of A/IS towards considering the values and intentions of the organisation and the consequences for human agency and well-being beyond the technical objectives of using the system. Adapting and expanding these principles this model draws upon contemporary and classical thinking on moral responsibility and accountability and describes a methodology for establishing an agreed ethical foundation and moral objectives for projects in the context of socially responsible design and construction.

Keywords: designers knowledge, fourth industrial revolution, sustainability, UN Sustainable Development Goals.

Introduction

Humankind has so permanently influenced the major earth system processes, including atmospheric, geologic, hydrologic and biospheric that the current geological time period is defined as anthropogenic and accordingly has been named the *Anthropocene*. The global landscape that we see today is one that is in part, maybe a major part, that has been transformed by human activity (Berger 2010), and in that construction and agriculture are the most visible, though not sole vehicles for that transformation.

It is thus, with no pretensions to grandiosity that the engineering professions at the 2018 Global Engineering Congress (GEC) recognised that, "*[t]he infrastructure designed and built by our members to ensure human well-being is the critical interface with our planet. It dictates the patterns and flows through which we live our daily lives and affects our long-term prosperity,*" (ICE 2018a). In the context of the UN Sustainable Development Goals (UNSDG) (UN 2015), the GEC identified the problems that increasing infrastructural interconnectivity and resource demands make on sustainability (social, economic and environmental) and asked how would the engineering community remain consistent with environmental stewardship whilst maintaining their commitment to constructing for the well being of Mankind (ICE 2018a, ICE 2018b, ICE 2020).

In their work on the design and application of Artificial Intelligence (AI) systems (henceforth Autonomous and Intelligent Systems, (A/IS)), Sekiguchi et al. (2009) and Sekiguchi and Hori (2020) have advanced a model of a design thinking hierarchy that has a bearing in the wider field of construction engineering, particularly when the profession and the industry are being encouraged to adopt digital technologies (Barbosa et al. 2017) or are already beginning to embrace the use of such systems, albeit with some concern regarding unintended consequences, in areas of health and safety, (McAleenan et al. 2019).

When applied to construction project development it can be acknowledged that engineering design thinking begins with consideration of the user interface where-in the project to be constructed is viewed in the context of the environment in which it will sit (Dynn et al. 2006); what is to be built, its function and the constraints imposed by the environment in physical, economic, social and political terms. These are often, but not exclusively technical matters requiring technical solutions and engineers comfortably excel in this area. Sekiguchi and Hori (2020) have elaborated on the knowledge levels accessed in project design through advancing social and ethical levels of knowledge required to achieve a design that is ethically sound. When applied to construction, looking above the level of consideration of the interface between the construction artefact and its environment, there is a case for a deeper consideration of how the artefact's interaction with the environment creates change in both the physical and the social environment. It is therefore appropriate to evaluate this change in the context of social relationships and personal and societal values, (King 2000). This level of consideration effectively connects with the concerns raised at the GEC and engineers' commitment to environmental stewardship and human well being.

Sekiguchi and Hori (2020) have developed an "*organic and dynamic tool* [named *dfrome*, an abbreviation of "website for Design FROM the Ethics level"] *for use with a knowledge base on ethics to promote engineers' practice of ethical AI design to realise further social values"*. Their work has been substantially influenced by the Institute of Electrical and Electronics Engineers' Ethically Aligned Design (IEEE 2019), the Asilomar Principles developed by the Future of Life Institute (FLI) and the AI Network. Their tool is essentially a database of ethics knowledge designed to support engineers grasp of the ethics issues associated with their designs and in their use of it, recommends scenario paths that connect ethics with the technology.

The IEEE's Ethically Aligned Design (1st edition (EAD1e)) is founded on the pillars of Universal Human Values, Political Self-determination and Data Agency, and Technical Dependability. These pillars connect to general principles that are designed to guide all manner of A/IS design. The principles are:

- Human Rights
- Well-being
- Data Agency
- Effectiveness
- Transparency
- Accountability
- Awareness of Misuse
- Competence

These principles are not exclusively the preserve of A/IS design and development; with minor adjustments they can also effectively contribute to the design thinking process in construction, as Sekiguchi and Hori's *dfrome* aims to do.

However these principles (as adjusted) coupled with the type of ethics knowledge base as envisaged by Sekiguchi and Hori (2020) would not be sufficient to address the concerns of engineers for a consistent system of construction that recognises the transformative nature of human interactions on the planet and facilitates a morally positive transformation. McAleenan (2020) outlined a range of difficulties associated with developing morally neutral A/IS, difficulties that EAD1e was designed to address, and all difficulties that stem from a human inability to agree on what ought to be done in any particular circumstance, or what constitutes morally good actions. For example McAleenan (2019) opines on the Camden Bench as something that is simple and seemingly innocuous, designed as an aesthetically pleasing point of rest for pedestrians, but;

> *"… its fundamental purpose it is a tool of social control. It is an icon of "aggressive architecture" that has been designed to prevent anti-social behaviour such as graffiti, drug dealing, congregating and homeless people sleeping, (Swain 2013). [When] put in the context of many other features of architectural furniture (stones under bridges, spikes, anti-stick walls, etc., and we see construction contributing to the control of users of city spaces".*

Outlined below is a prototype model that builds upon the principles detailed in EAD1e, the design thinking hierarchy outlined by Sekiguchi et al. (2007) and Sekiguchi and Hori (2020), and which draws on the work of Floridi (2016) on faultless responsibility in social networks. The model is praxic in nature, iteratively integrating an ethically aligned thinking process with the design activity and the construction process.

Moral Responsibility and Ethics Information Base

Distributed morality (DMA/DMR)

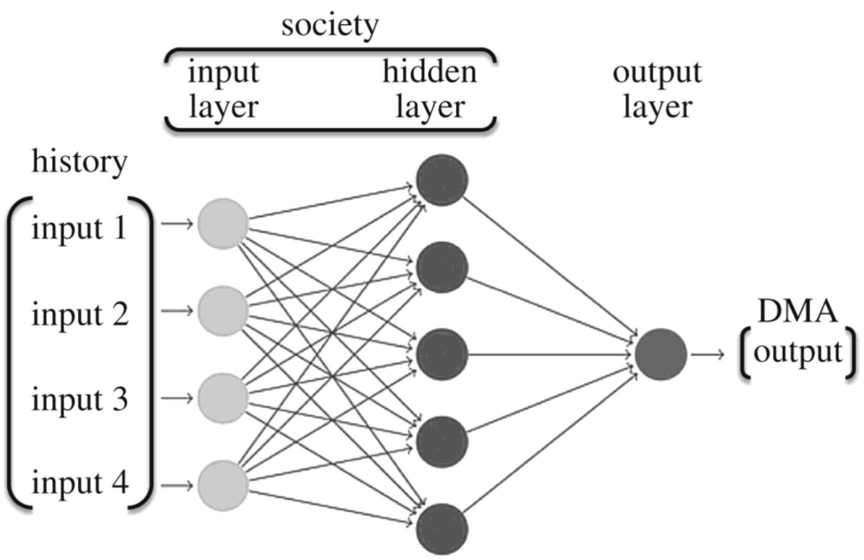
Ethically aligned design is grounded on an awareness of moral responsibility and the potential for actions by autonomous agents to have moral consequences, whether good or bad (morally loaded). When there is a direct and immediate connection between the agent carrying out an act (which includes decisions) and the recipient (person, property or environment) of that action then blame or praise can be clearly attributed. However in complex social networks where the acts of different agents are essentially neutral, it becomes more difficult to attribute blame or praise for the outcomes. This difficulty has contributed to a barrier to thinking on moral responsibility where there is a substantial distance between the initiating act and ultimate output with a complex intermediary web of agents propagating the act towards the output, (Floridi 2016).

By way example, the New York High Line project that transformed a derelict rail infrastructure into a raised linear public park through Manhattan had the effect of gentrifying areas of the city, raising rents and thereby forcing out long established residents and local small business who could no longer afford to remain in their districts, (McAleenan & McAleenan, 2017). The transformative nature the project was such that whilst the initial intent was beneficial the unintended consequences created other transformations that were morally loaded vis the residents of the area, yet conversely beneficial to the newer influx of wealthier residents. The client, design-team and contractors could in truth not be held legally liable for these negative consequences of the project, but morally could be considered responsible for them.

Floridi (2013) uses the term distributed morality to refer "*only to cases of moral actions that are the result of otherwise morally-neutral or at least morally negligible interactions among agents constituting a multi-agent system, which might be human, artificial* [A/IS]*, or hybrid"*. Notwithstanding the many ethical nuances outlined by Floridi the concept may be simplified as a moral action that ultimately causes harm to a recipient even though that was not the intent of the agent immediately causing the initiating action nor indeed of any other agent in the system. For example, an accident occurs that causes harm to a person, property or the environment though the action causing it was not intended (by definition). That causal action is rarely the initiating action; it is often the final causal action in a series of causal events and circumstances involving a network agents none of whom may have any direct connection to or intention regarding the recipient of the final action. Such an action propagating forward through a networked system is a distributed moral action (DMA).

Distributed moral responsibility (DMR) is not concerned with guilt, but rather with faultless responsibility (Floridi 2016), namely when propagating back from the DMA the initiating action(s) without which the DMA could not have occurred is ascertained. In this sense agents in the network may be deemed responsible though guiltless for the end results of their decisions or actions where those actions or decisions are necessary precursors to the DMA and the agents themselves are autonomous, able to interact with each other and capable of learning from and changing their interactions. Figure 1 is Floridi's (2016) concept of the network.

Figure 1: A multi-agent system as a multi-layered neural network, (Floridi 2016)



In the context of design and construction the input layer can be taken to represent the design decisions and the agents involved in making them. The hidden layer is the network of agents and their actions and decisions in constructing and use the project designed. At the output layer the recipient(s) of the DMA are, e.g., construction workers, end users, society and the environment.

In a simple Prevention through Design example, the decisions of the client, architect, engineer, financial backer (inputs) to opt for a glass enclosed atrium at the fourth floor of a ten floor structure impacts ultimately on a worker who task is to clean the external glass and any negative DMA (output) is traceable back to the initiating decisions and agents responsible for them (DMR).

However in a small multi-agent network and a short propagation path some or all of the initiating agents may be liable, criminally and civilly, e.g.,the Express Park, Bridgwater fatality where the contractor and the architect were both held liable, (Croner-i 2010). But in more complex networks with longer propagation durations, some of decades or even centuries the picture changes. The social need for mobility and the manufacturing industry's production of faster, safer and more affordable transport vehicles cannot be held liable for the vast global roads infrastructure with it concomitant habitat fragmentation and atmospheric pollution, but they are responsible for it and thus morally responsible its consequences.

Information base

Sekiguchi and Hori (2020) have observed that, in relation to A/IS ethics studies have not sufficiently been incorporated into the research and development of A/IS, that A/IS ethics it too broad and the information base very large, engineers have little time to understand and construct a relationship between A/IS ethics and their own research, and that engineers find A/IS ethics too abstract. These observations relate very much to the nature of ethics as a field of study and the difficulties of non-specialists, whether A/IS engineers or construction professionals, in getting a sufficient grasp of the subject matter so as to inform their work of its ethical implications and moral consequences, (McAleenan 2020). Their project is designed to fill these gaps through the provision of a knowledge base of A/IS ethics that can deal with the complex relationships between them and provide engineers with scenario paths that connects the ethics to their research. It is interactive and has scope for addressing varying social values and personal ethics.

Applying these observations to the field of construction design Floridi's schematic (Figure 1) can be modified to include types of information and knowledge informing the input layer, (Figure 2)

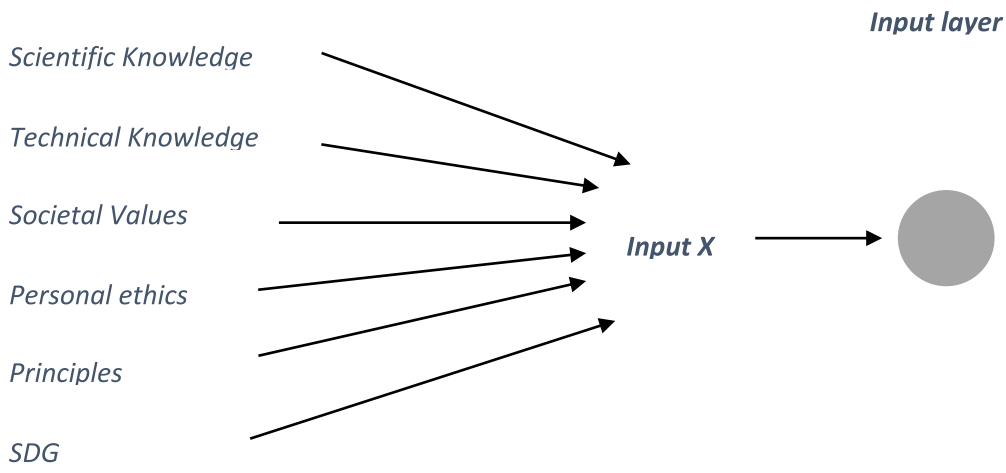


Figure 2: Knowledge base informing inputs to Floridi's multi-agent system

To exercise ethically aligned design thinking the knowledge base must grow substantially beyond the technical knowledge necessary to meet the physical requirements to construct the project and the ethics associated with professional codes of conduct.

The anthropogenic influence on the planet has not been wholly positive and it is important for designers to have a deeper awareness of how construction impacts on the environment if they are to effectively mitigate the harm it causes. Current environment considerations are often inadequate, primarily because the knowledge used in informing environmental design is limited and therefore inadequate. By way of example, the effects of construction, e.g., roads or even something as simple as a cabin in the wilderness, on trophic systems can be detected many kilometres into an unfragmented habitat (Soulé 2010). The nature of those impacts and their harmfulness varies on the basis of the construction project and the changes it manifests and continues to manifest on the flora and fauna in the immediate area of the project. Changes to mega biomes are observable to the average person, micro biomes in the soils and water sources need specialist investigation to understand what changes have occurred. Inadequate knowledge of the science means that environmental policies of engineering and contracting companies can be ineffective, or even more harmful (e.g., laying nutrient rich soils to seed wildflowers on road verges) when it comes to remediating environmental impacts.

If construction is to benefit humanity (ICE 2018b) a knowledge of the social sciences is also important if designers are to understand the impacts of projects on society. Disciplines such as sociology, psychology and economics are pertinent to that understanding. Mega sporting projects such as the Olympics and the World Cup are held up as opportunities for urban improvement and economic growth through increased trade, jobs and income. But often the reality is very different, with host cities ending up in massive debt and poor populations reallocated when gentrification occurs on the sports campus, (Kumar 2012, McBride 2016).

Table 1 maps the General Principles of ethically aligned design (EAD1e) against the UN sustainable development goals. Principle 3 has been broadened from “data agency” to “agency" and Principle 7 “awareness of misuse” amended to “awareness of consequences”. There is insufficient scope in this paper discuss the interpretations of how the principles relate to the goals but it is sufficient to illustrate that two major societal initiatives interrelate. These are to be considered an adjunct to the societal values necessary to effective ethical thinking.

UN Sustainable Development Goals / EAD1e General Principles	1. Human Rights	2. Well-being	3. Agency	4. Effectiveness	5. Transparency	6. Accountability	7. Awareness of consequences	8. Competence
1. No Poverty	✔	✔		✔				
2. Zero Hunger	✔	✔		✔				
3. Good health & well-being	✔	✔		✔				
4. Quality Education	✔	✔	✔	✔				
5. Gender equality	✔	✔		✔				
6. Clean water & sanitation	✔	✔		✔				
7. Affordable & clean energy	✔	✔		✔				
8. Decent work & economic growth	✔	✔		✔				
9. Industry, innovation & infrastructure	✔		✔	✔	✔	✔	✔	✔
10. Reduced inequalities	✔	✔	✔	✔	✔	✔		
11. Sustainable cities and communities	✔	✔		✔				
12. Responsible consumption and production	✔	✔	✔	✔	✔	✔	✔	✔
13. Climate Action	✔	✔		✔				
14. Life below water	✔			✔				
15. Life on land	✔			✔				
16. Peace, justice & strong institutions	✔	✔	✔	✔	✔	✔	✔	✔
17. Partnerships for the goals	✔	✔	✔	✔	✔	✔	✔	✔

Table 1: UNSDG mapped against EAD1e General Principles

Prototype Model for Ethically-Aligned Design Thinking

A design thinking methodology

Awareness of the nature of faultless responsibility and the designer's role as an initiating causal agent in a multi-agent network combined with access to an appropriate information base (via personal knowledge or an A/IS system such as dfrome (Sekiguchi and Hori 2020)), is the platform on which design thinking can take place. Given that ethics and moral behaviour is a complex field of study fraught with disagreements about what ought to be done and contractions between and varying interpretations of competing theories (McAleenan 2020) such a platform cannot produce ready made solutions to the ethical questions posed by the design project. Indeed it can be as difficult to frame the correct ethical question as it is to interpret a solution. On top of this, not all moral agents act morally, including agents who input to the decision making process.

To align what can be done with what ought to be done requires an appreciation of consequentialism (even a deontic or a virtue ethic must be carried out in an appreciation of the consequences) and the ability to decide between a range of consequences in order achieve an outcome that promotes the wellbeing of mankind. This includes decisions on what is to be accepted as "wellbeing".

Figure 3a outlines a basic praxic model for design thinking. It is a methodology wherein theory (design thought) and practice (design output and construction action) are integrated and weighed against each other in a continuing iterative process of critical evaluation of design and output in a moral matrix.

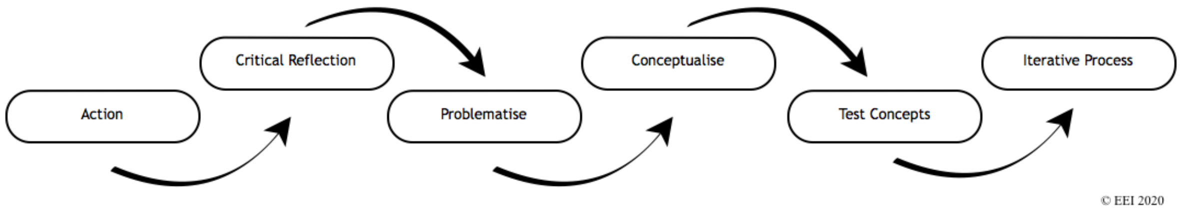


Figure 3a: Prototype praxic model for Ethically Aligned Design

Several steps in the process are highlighted in Figure 3b. The foundational platform described above forms the starting point for a critical evaluation of the project being undertaken, an aspect of which is consideration of the project as an object in space and time rather than as something static on the landscape. From its inception it is recognised that it has come with a history of matters and events that preceded it and is being developed in anticipation of what will come after. In other words there are reasons for its development, and for its function and purpose going forward.

A new high-rise office block in a city centre comes into being on the basis of a complex multi-disciplinary history of the development of the city and its centre. It further exists in a dynamic environment that will continue to change into the future; changes that may be as a result of the office block being built or changes that in combination with the office block will impact the local climate/weather and natural, social, political and economic environments, etc.

But the overt reasons for a project do not explain the project completely. It will impact on the environment immediately and over time, it will impact and alter social, economic and political relations through processes such as gentrification, economic corridors and urban expansion. Critical evaluation reflects on these causes and effects in order to understand them and to achieve a controlled transformation.

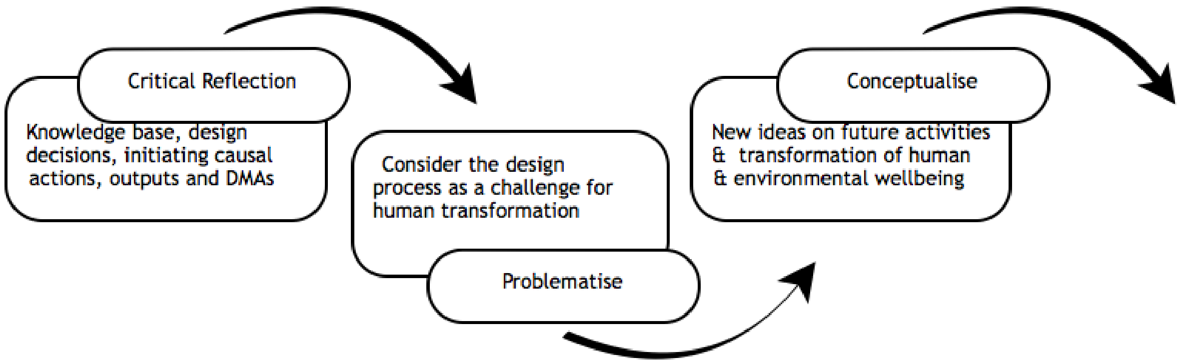


Figure 3b: Prototype praxic model for Ethically Aligned Design, detail

Problematisation is a key element of praxis in which the designer considers the project as a challenge with the purpose of transforming human and environmental wellbeing. As an element of critical thinking it is a method that rejects the taking of common knowledge for granted and opens the way for reflection, new viewpoints and ultimately brings about positive change.

However well or poorly done problematisation develops new ideas, solutions and activities that are aimed at achieving a positive transformation. The ideas are then tested in practice and are in turn are critically evaluated in a continual cycle of reflective action and active reflection. The iterative nature the process ensures continual advancement, continual improvement in the design process and implementation, not just technically but transformatively. This means that the engineer, the client and society through its government agents must envisage the changes that is both desirable and necessary and then design and build for that change.

The advantage of this model is that it overcomes fixed thinking and therefore fixed solutions. Praxic thought and action on ethically aligned design recognises that the built environment impacts of Man and nature and aims to ensure that that impact if beneficial in the immediate and long-term future.

Conclusion

The use of autonomous and intelligent systems is growing exponentially worldwide and though the construction industry is lagging behind other industries in its use it is being encouraged to adopt A/IS to exploit the potential for increased productivity (Barbosa et al. 2017). A major concern is that the pace of technical development of these systems far outpaces the development of ethical codes and practices that will regulate and control what is developed and how it is used.

At the same time species depletions, environmental degradation and climate change are immediate critical issues that humanity must address, not least because much of these are anthropogenic effects. Construction, extraction, agriculture, and manufacturing, etc., will of necessity continue and therefore humanity's impact on society and the environment will continue. A defining quality of humankind is each persons awareness of themselves as an ethical and moral being. Developments in moral theory, ethically aligned design and methodologies for critical thinking provide the tools for engineers to more effectively design for humanity and the welfare of the planet.

In construction the agents for change, clients, engineers, contractors and government, must begin by recognising and accepting their role as morally responsible agents for the moral consequences of their decisions and actions in an extended multi-agent network. In doing so they must act on that responsibility by quantifying the moral benefit or moral loadedness of any project they undertake. A/IS systems such as *dfrome* are useful tools but are no substitute for decision making by human agents who have made sufficient inquiry and continually reflects upon their decisions through to completion and onto the next project.

The model described draws upon social, ethical, technical and political initiatives, merging them into a workable process of ethically aligned design that prioritises societal well-being and environmental sustainability a step principal outputs of all construction projects.

References

Barbosa, F., Woetzel, J., Mischke, J., et al. .2017. Reinventing Construction: a Route to Higher Productivity. McKinsey Global Institute, New York, NY, USA. [online] Available at : https://www.mckinsey.com/%7E/media/McKinsey/Industries/Capital%20Projects%20and%20Infrastructure/Our%20Insights/Reinventing%20construction%20through%20a%20productivity%20revolution/MGI-Reinventing-Construction-Executive-summary.ashx [accessed 16 July 2020]

Berger, J. 2010. Fear-mediated food webs, in J, Terborgh and A, Estes , eds. 2010. Trophic Cascades. Island Press, Washington DC. pp. 241-253

Croner-i, 2010. Express Park and Oxford Architects v Health and Safety Executive. [online] Available at https://app.croneri.co.uk/law-and-guidance/case-reports/express-park-and-oxford-architects-v-health-and-safety-executive?product=133 [Accessed 22 August 2020]

Dynn, C.L., Agogino, A.M., Eris, O., Frey,D.D. and Leifer, L.J., 2006. Engineering design thinking, teaching, and learning. IEEE Engineering Management Review, [e-journal] 34(1) pp. 65-65, First Quarter 2006, doi: 10.1109/EMR.2006.1679078

Floridi, L. 2013. Distributed morality in an information society. Sci Eng Ethics, [e-journal] 19(3), pp.727-743. doi:10.1007/s11948-012-9413-4 [Accessed 16 July 2020].

Floridi, L. 2016. Faultless responsibility: on the nature and allocation of moral responsibility for distributed moral actions. , Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, [e-journal] 374(2083), https://doi.org/10.1098/rsta.2016.0112 [Accessed 16 July 2020] .

IEEE, (Institute of Electrical and Electronics Engineers) 2019. Ethically Aligned Design: a Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, 1st edn. IEEE New York, NY, USA. [online] Available at : seehttps://standards.ieee.org/industry-connections/ec/ead1e-infographic.html [Accessed 16 July 2020].

Institution of Civil Engineers [ICE], 2018a. UK, US and Canadian civil engineering bodies commit to joint action on sustainable development. [online] Available at: https://www.ice.org.uk/news-and-insight/ice-press-centre/uk,-us-and-canadian-civil-engineering-bodies-commi [Accessed 16 July 2020] .

Institution of Civil Engineers [ICE], 2018b. The Institution of Civil Engineers Royal Charter and By-

Laws. [pdf] Institution of Civil Engineers. Available at: https://www.ice.org.uk/ICEDevelopmentWebPortal/media/Documents/About%20Us/ICE-Royal-Charter-and-By-laws-2018.pdf [Accessed 16 July 2020].

Institution of Civil Engineers [ICE], 2020. Sustainability Route Map. [online] Available at: https://www.ice.org.uk/knowledge-and-resources/sustainability-route-map [Accessed 16 July 2020].

King, R.J.H., 2000. Environmental Ethics and the Built Environment. Environmental ethics, June 2000, 22(2):115-131, DOI: 10.5840/enviroethics200022230.

Kumar, A., 2012. Want to cleanse your city of its poor? Host the Olympics. Ceasefire Magazine, April 2012. [online] Available at : https://ceasefiremagazine.co.uk/olympics-opportunity-cleanse-city/ [Accessed 17 July 2020].

McAleenan, P. and McAleenan, C., 2017. Re-envisaging construction in the context of humanisation. Proceedings Joint CIB W099 & TG59, International Safety, Health, and People in Construction Conference. Cape Town, 11-13 June 2017.

McAleenan, P., 2019. Contributory thoughts on social engineering, [pdf] An Spréach, Expert Ease International, February 2019. Available at: http://expertease.ie/data/documents/an-Spreach-1-Feb-2019.pdf [Accessed 22 August 2020].

McAleenan, P., McAleenan, C., Ayers, G., Behm, M. and Beachem, Z. 2019. The ethics deficit in occupational safety and health monitoring technologies. Proceedings of the Institution of Civil Engineers – Management, Procurement and Law 172(3), pp.93–100, https://doi.org/10.1680/jmapl.18.00027 [Accessed 16 July 2020].

McAleenan, P. 2020. Moral responsibility and action in the use of artificial intelligence in construction. Proceedings of the Institution of Civil Engineers – Management, Procurement and Law, https://doi.org/10.1680/jmapl.19.00056 [Accessed 16 July 2020].

McBride, J. 2016. The Economics of Hosting the Olympic Games. Council on Foreign Relations, July 2016. [online] Available at: https://www.cfr.org/backgrounder/economics-hosting-olympic-games [Accessed 17 July 2020].

Sekiguchi, K., Tanaka, K. and Hori, K. 2009. "Design with discourse" to design from the "ethics level". In: Information Modelling and Knowledge Bases XXI, Frontiers. In : Artificial Intelligence and Applications. T.W. Družovec, H. Jaakkola, Y. Kiyoki, T. Tokuda and N. Yoshida eds. IOS Press, Amsterdam, the Netherlands, vol. 206, pp. 307–314.

Sekiguchi, K. and Hori, K ., 2020. Organic and dynamic tool for use with knowledge base of AI ethics for promoting engineers' practice of ethical. AI & Society, 35(1), pp.51–71, https://doi.org/10.1007/s00146-018-0867-z [Accessed 16 July 2020].

Soulé, M. 2010. Conservation Relevance of Ecological Cascades. In J. Terborgh and A. Estes, eds. 2010, Trophic Cascades. Island Press: Washington DC. pp. 337-352.

United Nations [UN] 2015. 17 Goals to Transform Our World. [online] Available at: https://www.un.org/sustainabledevelopment/ [Accessed 16 July 2020].