SHARE SPACE .eu

# Ethics by Design: A Roadmap for Good Enough Ethics (GEE)

**Policy Brief**

Kathleen Richardson, Jessica Sutherland & Kathleen Bryson

Funded by the European Union

Grant Agreement No. 101092889

The development of technology - from AI to sustainable computing - can disrupt, harm or improve our way of life.

**In SHARESPACE we believe that considerations about human users, broader society, nonhuman animals and the environment should be at the heart of the design and development of extended reality technologies. Using an ethics by design approach within this project therefore ensures that the development of SHARESPACE technology is ethical (enough), equitable (enough) and environmentally friendly (enough).**

We have developed Good Enough Ethics (GEE) to offer guidance and actionable tools to shape the ethical design and development of technologies. The GEE approach is an interdisciplinary design methodology rooted in realistic challenges and development practices.

We propose six GEE principles and ways of implementing them throughout the design process. All through the SHARESPACE project, we will be implementing and testing these principles to evaluate their effectiveness.

This policy brief was developed in collaboration with the SHARESPACE (Embodied Social Experiences in Hybrid Shared Spaces) Consortium and their development of AR/VR/XR technologies. Further information can be found at sharespace.eu. This policy brief focuses on GEE as a methodology for technology development in general, not in the specific AR/VR/XR contexts.

# Contents



# Acronyms

| | |
|---|---|
| **DNSH** | Do No Significant Harm |
| **EbD** | Ethics by Design |
| **GDPR** | General Data Protection Regulation |
| **GEE** | Good Enough Ethics |
| **HCD** | Human-Centred Design |
| **UX** | User Experience |
| **UCD** | User-Centred Design |

# Introduction

Novel developments in technology bring with them ethical challenges and considerations. It can be tempting to see such advancements, such as artificial intelligence (AI), the internet of things (IoT) and extended reality (XR) in purely technical ways. However, as technology is developed and adopted, it falls on individuals and societies to navigate its impacts on humans, nonhuman animals and the natural world.

Technological progress can be an amazing thing for the world, but it comes with concomitant risks. The year 2024 is set to be the year of AI adoption, with worldwide IT spending predicted to increase 8% to approximately 5 trillion USD.[1]

**The global VR market size is predicted to exceed 22 billion USD by 2025.[2]**

**The global AI market is predicted worth 370.2 billion USD by 2025.[3]**

Ethics is a discipline concerned with the development of moral principles, virtues and values. It is typically associated with a branch of philosophy, but is also central to the way social animals organise their lives, even if they have never picked up a moral philosophy book.

## Moral Philosophy as an Ethical Deliberation Tool

Ethics as a philosophical discipline explores morality in terms of right and wrong, and good and bad. It also considers what constitutes a good life, what are our rights and what duties we have to each other.[4] The philosophical study of ethics can be broadly defined into three different fields: metaethics, normative ethics and applied ethics.[5] Metaethics asks the big questions and tries to understand the origins of moral principles and values. Normative ethics is generally concerned with what is right and wrong and developing ethical theories to guide our actions. Finally, applied ethics looks at how to put ethics into practice in the real world and attempts to give us guidance on how to act in specific situations.5

Figure 1 shows the way that the three fields of ethics interact with each other, with normative ethics roughly attempting to translate the nature of morality explored in metaethics into ethical principles, and applied ethics putting these principles into action.

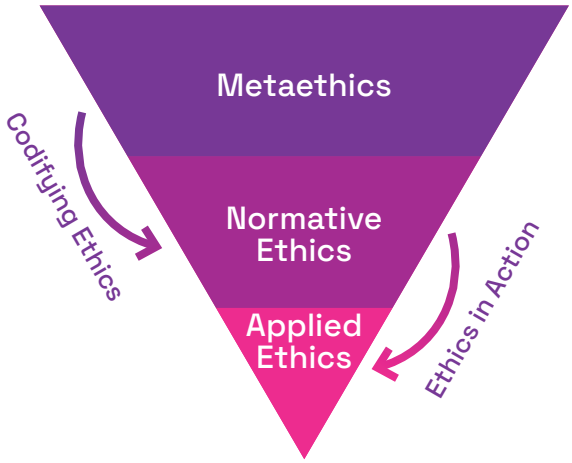


Fig. 1. The three fields of moral philosophy

This is not to say these fields are distinct: applied ethicists will often ask the big meta-ethical questions (e.g. what is the nature of reality?) and metaethicists will draw upon our ethics practices from an applied standpoint.

## Conflicts Between Ethics and Technology

It is often suggested that technology and ethics are incompatible, given that they arise from different objectives and paradigms. Thinking about ethics is therefore often seen to slow down technological progress and stifle innovation.[6] This perceived conflict has led us towards a development-first, consequences-second approach. For example, there is evidence of harmful effects of social media use on public discourse, and young people's mental health. Constant innovation may also harm the environment by requiring more and more resources, unsustainable materials and creating non-recyclable e-waste. Whether there is truth in the notion that ethics 'gets in the way' is up for debate; however, the sentiment remains - how can we mediate between technological advancement and ethical considerations?

**This policy brief outlines a new approach to technological development that is an outcome of a collaboration between technologists and ethicists. GEE centres the impacts that new technology can have on humans, nonhumans and the environment. Born out of an ethics by design approach, the GEE methodology aims to guide decision-making in each step of the research and development of new technologies towards an ethical (enough) technology.**



[1]Statista, 2024. 'Tech Trends 2024 - Statistics and Facts.' Available at: https://www.statista.com/topics/9025/tech-trends/#topicOverview Last accessed: 1st May 2024.
[2]Statista, 2024. 'Virtual Reality (VR) - Statistics and Facts.' Available at: https://www.statista.com/topics/2532/virtual-reality-vr/#topicOverview Last accessed: 1st May 2024.
[3]Statista, 2024. 'Artifical Intelligence - Worldwide.' Available at: https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide Last accessed:1st May 2024.
[4]Hendricks, C & Matthews, G, 2020. Introduction to Philosophy: Ethics. United States: Rebus Foundation, p.1.
[5]LaFollette, H & Persson, I, 2013. 'Introduction' In The Blackwell Guide to Ethical Theory, edited by Hugh LaFollette, and Ingmar Persson. Newark: John Wiley & Sons, p.2.
[6]Fontrodona, J, 2013. "The relation between ethics and innovation." In Social Innovation: Solutions for a Sustainable Future, pp. 23-33. Berlin, Heidelberg: Springer.

# Good Enough Ethics as Ethics by Design

Thinking about ethics is not, and should not be, restricted to ethics professionals. To create ethically robust technology, ethicists and technologists must work together to bridge the gap between what is technically possible and what is ethically desirable.

The GEE methodology is influenced by the concept of 'good enough parenting',[7] where parental perfection is not emphasised, but rather a parent raises their child with sufficient autonomy, values and skills to eventually tackle the world by themselves. It was developed by Donald Winnicott, an English paediatrician and psychoanalyst in the 1950s.[7] He was working at the time of another prominent theorist John Bowlby, founder of attachment theory. These thinkers prioritised family relationships as their underpinning philosophy. This is also what we do in the GEE methodology.

Moreover, GEE builds on this intuitive, dialogical and co-creative model to develop robustly ethical technology from the beginning. It aims to recognise that the development of perfectly ethical technology is simply unrealistic, and perhaps even impossible. The 'good enough' in GEE may seem controversial to some readers as it implies an ill-considered approach. On the contrary, ethicists may prescribe multiple ethical values, but these can be divorced from the practicalities of the real world and the complexities of human relationships. Moreover, ethical values are not separate from political motivations and often conceal power structures (e.g. ethics washing).

As good enough parenting focuses on familial relationships as the basis to all community and world relationships, some may view GEE as characterising technological artefacts as 'children' that respond to nurture. This is not the case. GEE is about acknowledging the power of human relationships and their creative, transformative potential, while at the same time accepting an imperfect world. GEE works with technologists to draw on their experiences and intuitions. It is not a top-down approach, but one developed dialogically in collaboration.

**We have identified 6 core ethical values, organised through a recursive, relational typology of *family*, *community* and the *world*. For family, the key values are attachment and trust; for community, autonomy and privacy; and for the world, justice and responsibility.**

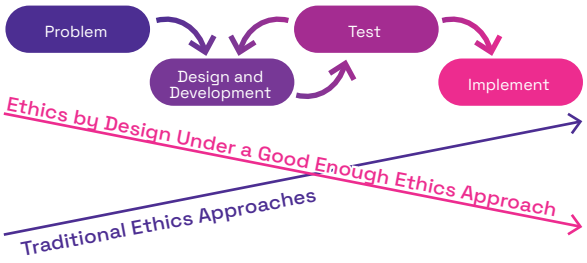


Fig. 2. The level of involvement of ethics by design compared to other, more traditional ethics approaches

Figure 2 illustrates the spectrum of ethical technologies, starting from simple legal compliance towards a perfectly ethical technology. What GEE attempts to do is to find a path between these two extremes to ensure that the technology developed is safe and contributes to the wellbeing of humans, nonhuman animals and the environment, without stifling beneficial technological development.

GEE builds on an ethics by design approach, which guides our decision-making in the design and development of new technologies using ethical principles. Ethics is sometimes thought to be a set of rules or standards that people should follow to do right and to be good. However, this is too simplistic a picture of what ethics by design can be. Ethics by design aims to proactively anticipate ethical issues, rather than simply react to them when they arise. This is in contrast to more traditional ethical approaches which may identify some ethical issues when conceptualising the problem to overcome, yet with most of the ethical deliberation taking place after the development of the technology. However, this is not always possible, especially in the emerging technologies space because of unforeseen problems arising. The GEE approach draws on the resourcefulness of stakeholders to deal with these unforeseen ethical issues and accidents in order to minimise future issues.

Of course, compliance with legal and ethical rules is important for guiding action and design. It is the baseline of ethical action. But frequently, ethical acts are not adequately captured in existing laws, rules and regulations. Sometimes our laws, rules and regulations can promote unethical actions (e.g. slavery, subordination of women, child exploitation). GEE as an ethics by design approach acknowledges this tension and aims to work towards compliance with relevant rules and regulations (for example, GDPR[8] and the EU AI Act[9]), whilst promoting ethical values and principles that go beyond this.

[7]Winnicott DW. Transitional objects and transitional phenomena; a study of the first not-me possession. Int J Psychoanal. 1953;34(2):89–97.
[8]European Parliament, 2016. 'Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)'.
[9]Council of the EU, 2023. 'Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world'. Available at: https://www.consilium.europa.eu/enpress/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/
Last accessed: 15th December 2023.
[10]Floridi, L, 2018. Soft ethics, the governance of the digital and the General Data Protection Regulation. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133).

# 6 Principles of Good Enough Ethics

The GEE approach is an intuitive, relational and dialogical process. While there are many ethical principles, we had to think carefully about which ethical values should be prioritised. The below six principles aim to guide design behaviours towards the development of ethical (enough) technologies.

## Family

The basis of all life - without it there would be no basis for relationships in community or in the world.

### Attachment

Balancing the benefits of technology-mediated connections with the preservation of authentic human attachment is a critical consideration in the development of such technologies.

### Trust

Trust is a value for social cohesiveness and cooperation. Developing trust between users, technology developers and the technology developed is important. However, this trust cannot be simply unearned or automatic.

## Community

The basis of our lives with others, including our co-animals and other living beings. Supported by our experiences within the family.

### Autonomy

The notion of autonomy encompasses human freedom, including freedom of thought, expression and association. Valuing autonomy means considering the ways in which technologies may manipulate, nudge or alter human action and societies.

### Privacy

Privacy can be understood both in terms of the collection and processing of personal data in general, but is crucially context-dependent.

## World

Materialist as well as socially constructed domains that may encompass the physical, political and/or spiritual.

### Justice

Justice manifests differently across diverse practical contexts, encompassing retributive, distributive, intergenerational and social justice. Emerging technologies have brought justice concerns to the forefront e.g. algorithmic justice.

### Responsibility

Related to the principle of trust, responsibility asks us to consider accountability of both users and technologists. Importantly for the development of systems using AI is the notion of a 'human in the loop'.[11]



[11]Boddington, P, 2023. AI Ethics: A Textbook. United States: Springer, p.43.

# 6 Principles of Good Enough Ethics

## Identifying Ethical Risks and Solutions

Figure 4 outlines the five questions to consider when identifying ethical risks and proposing solutions. This tool asks us to reflect on the project objectives, and then map the relevant GEE principles onto these objectives. In doing so, we need to recognise who will be affected, and in which ways. Finally, we can then begin to consider potential solutions under four themes: new approach, adjustments, mitigations and explanations.

**What is the objective?**

- Fulfilling a user need or preference
- Improving user experience
- Accessibility goals
- Improving efficiency, reliability or availability

**Which GEE principles are relevant?**

- Attachment
- Autonomy
- Justice
- Trust
- Privacy
- Responsibility

**Who is affected?**

- Individuals - users, non-users, stakeholders
- Society - identities, social groups
- Nonhuman animals
- Environment

**How are they affected?**

- Physical
- Financial
- Psychological
- Emotional
- Moral
- Exploitation

**How can this be solved?**

- New approach needed
- Adjust the method e.g. different dataset
- Mitigate the effects e.g. safety protocols
- Explain the problem and how to avoid it e.g. user manual

Fig. 4. Ethical risks questionnaire

## Mitigating Risks: Designing Guardrails

Various themes may come to mind when considering these nine principles. Below are a selection of these themes, with examples of how these may bear on technological development.

| |
|---|
| Environmental concerns: Production, use and end life of technology under DNSH12 |
| Health and safety: Wearable technology, eye strain, neck strain, cybersickness |
| Privacy of users: Collection and storage of data, realistic representation of users |
| Consent: Understandibility of technology, informed consent in proof of principles and use cases |
| Futurecasting: Potential (mis)use of technologies |
| Objectification: Representation of users |
| Influence of the built environment: Nature of the research environment influencing innovation |
| Manipulation: Nudging techniques changing human behaviour |
| Blurring of reality: VR/AR/XR as both distinct from, and as part of, the real world, online personas |

[12]European Commission, 2021. 'Do No Significant Harm.' Available at: https://knowledge4policy.ec.europa.eu/glossary-item/do-no-significant-harm_en Last accessed: 18th December 2023.

Funded by
the European Union

For more information about SHARESPACE, our Good Enough Ethics approach and published research, visit sharespace.eu

SHARE SPACE .eu