TECHETHOS

FUTURE ○ TECHNOLOGY ○ ETHICS

Enhancement of Ethical Frameworks and Outline of Detailed Ethics Framework

Deliverable 5.1

TechEthos receives funding from the EU H2020 research and innovation programme under Grant Agreement No 101006249. This deliverable and its contents reflect only the authors' view. The Research Executive Agency and the European Commission are not responsible for any use that may be made of the information contained herein.

WP5 (T5.1)			
Work Package	WP5		
Lead Partner	De Montfort University, DMU		
Author(s)	Nitika Bhalla (DMU), Sara Cannizzaro (DMU) Kathleen Richardson (DMU), Laurence Brooks (Sheffield/DMU)		
Contributor(s)	Laurynas Adomaitis (CEA), Bennet Francis (UT), Steven Umbrello (TUD), Michael J Bernstein (AIT), Pieter E Vermaas (TUD), Anaïs Resseguier (TRI), Gustavo Gonzalez (AIRI), Andrea Porcari (AIRI), Alexei Grinbaum (CEA)		
Due date	30th June 2023		
Submitted date	30th June 2023		
Version number	1	Status	final

Project Information	
Grant Agreement number	101006249
Start date	01/01/2021
Duration	36 months
Call identifier	H2020-SwafS-2020-1
Topic	SwafS-29-2020 - The ethics of technologies with high socio-economic impact
Instrument	CSA

The information, documentation and figures in this deliverable were produced by the TechEthos project consortium under EC grant agreement 101006249 and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

Dissemination Level	
PU: Public	☒

Quality Control	
Reviewed by:	Review date:
Julie Vinders [Trilateral Research]	07/06/2023
Mathijs Vleugel [ALLEA]	13/06/2023

Revision history

Version	Date	Description
1.0	19/05/23	Draft for review by ADIM board, references and other misc elements still to be tidied up.
1.0a	22/05/23	Removed section 4, as more relevant to D5.3
1.0b	05/06/23	Amendments/suggestions by ADIM board members
2.0	07/06/23 and 13/06/23	Revisions based on internal QA feedback

Keywords
Ethics, framework, technology, assessment, values, principles

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

How to cite

If you are using this document in your own writing, our preferred citation is:

Bhalla, N., Cannizzaro, S., Richardson, K., and Brooks, L., (2023) *TechEthos Deliverable D5.1: Enhancement of Ethical Frameworks and Outline of Detailed Ethics Framework.*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

The TechEthos Project

Short project summary

TechEthos is an EU-funded project that deals with the ethics of the new and emerging technologies anticipated to have high socio-economic impact. The project involves ten scientific partners and six science engagement organisations and runs from January 2021 to the end of 2023.

TechEthos aims to facilitate “ethics-by-design”, namely, to bring ethical and societal values into the design and development of new and emerging technologies from the very beginning of the process. Technologies covered are “climate engineering”, “digital extended reality” and “neuro-technologies”. The project will produce operational ethics guidelines for these technologies for users such as researchers, research ethics committees and policy makers. To reconcile the needs of research and innovation and the concerns of society, the project will explore the awareness, acceptance and aspirations of academia, industry and the general public alike and reflect them in the guidelines.

TechEthos receives funding from the EU H2020 research and innovation programme under Grant Agreement No 101006249. This deliverable and its contents reflect only the authors' view. The Research Executive Agency and the European Commission are not responsible for any use that may be made of the information contained herein.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No.101006249.

Table of contents

List of tables

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

List of figures

Definitions and abbreviations

Table 1: List of Definitions

Term	Explanation
Climate Engineering	Climate engineering is a family of technologies that enables the modification of natural processes and human activities looking to address and mitigate climate change locally and globally.
Digital Extended Reality	Extended Reality refers to AI-powered digital technologies (hardware and software) capable of perceiving and processing human sensorial outputs, e.g., voice, gestures, language, movement, emotions and other elements of human communication, as well as responding to these types of signals by creating an extended visual, audio, linguistic or haptic digital environment for users.
Neurotechnologies	Neurotechnologies are technologies that aim at affecting and emulating human-brain capabilities and functions through artificial replacements or add-ons in a two-way interaction between the brain and the external environment or systems.

Table 2: List of Abbreviations

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Term	Explanation
AI	Artificial Intelligence
ATE	Anticipatory Technology Ethics
DoA	Description of Action
eIA	Ethical Impact Assessment
ELSI	Ethical, Legal, and Social Implications
eTA	Ethical Technology Assessment
ETICA	Ethical Issues of Emerging ICT Applications
IEEE	Institute of Electrical and Electronics Engineers
IS	Information Systems
NLP	Natural Language Processing
OECD	Organisation for Economic Co-operation and Development
RRI	Responsible Research and Innovation
SRM	Solar Radiation Modification
TEAeM	TechEthos Anticipatory Ethics Matrix
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNFCCC	United Nations Framework Convention on Climate Change
VR	Virtual Reality
WP	Work Package

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Term	Explanation
XR	Extended Reality

Executive Summary

This deliverable explores ways to enhance existing ethical frameworks that can be useful and applicable to our three technology families, as well as other new and emerging technologies. The resulting outline of this deliverable is a detailed framework called 'TEAaM' (TechEthos Anticipatory Ethics Matrix) which supports the effective governance of new technologies in a broader sense, using a combination of existing frameworks such as ATE plus, Ethical Impact Assessment and a Future Studies approach. This composite framework has been developed using empirical insights from the TechEthos project, in particular WP2 and WP3 methodologies, in addition to consultation with expert stakeholders.

1. Introduction

1.1 Background

Why the need for an ethical framework? The ethics of emerging technology is the study of ethical issues at the Research and Development (R&D) and introduction stage of technology development through anticipation of possible future devices, applications, and social consequences [1]. Ethical considerations concerning the impacts of Research and Innovation are increasingly important, due to the fast pace of technological innovation and the ubiquitous use of the outcomes of RandI processes in society [2]. New and emerging technologies such as Neurotechnologies, Climate Engineering and Digital Extended Reality (TechEthos's technology families) have opened up the opportunity to stimulate questions and proposals to enhance existing ethical frameworks, which can help to mitigate some of these ethical challenges to technology and society.

An ethical framework is a set of principles that can provide a solid base for the development of applications that are consistent with the accepted social norms and moral principles and values in society. Agreeing on an ethical framework or a combination of frameworks will help to guide the developers and users of these technologies. It must be noted that such a framework will not eliminate *all* ethical risks due to the inherent nature of uncertainty when describing emerging technologies. Rather, the presence of an ethical framework could reduce the likelihood and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

potential negative impacts of ethical challenges, making researchers and policy makers aware of these possible implications.

The central problem for the ethics of emerging technologies is that we humans cannot predict the future, and therefore do not know which ethical issues will play out once the technology is fully developed and entrenched in society [1]. As the emerging technology is still evolving, many questions can arise about its nature, its future use, and its social consequences. However, if an ethical framework is to be useful in an area of emerging technology, it needs to be accepted by researchers/academics and policy makers *prior* to any activity that uses the technology or during the technology's development phase. Furthermore, the framework should be used in consultation at every stage of development and not just considered as an afterthought.

This deliverable first begins by analysing literature to identify current frameworks and methods that could be useful for developing an ethics framework for emerging technologies. Second, the results of the literature review were analysed and a critical evaluation of the selected frameworks with respect to the three technology families was carried out. Third, a detailed outline of a broad ethical framework was proposed using a combination of existing ethical frameworks that could be used for all three technology families as well as emerging technologies in general. This combination of frameworks was further 'enhanced' by incorporating methodologies from the TechEthos project, in particular WP2 and WP3, as well as expert review feedback.

2. Review of representative ethical frameworks

2.1 Introduction

A broad range of frameworks were reviewed, including: Anticipatory Technology Ethics (ATE), Ethical Matrix, IS Ethics Assessment Techniques (ETICA), Ethical Impact Assessment (EIA), SATORI CWA (Cen Workshop Agreement), Future Studies, Ethical Technology Assessment (eTA), Ethical Scenario Method and Anticipatory Technology Ethics (ATE) Plus. Section 2.1 provides a detailed overview of these frameworks. The literature was reviewed to assess the usefulness of each framework with respect to the three technology families.

The descriptions below describe the ethical frameworks, their main features and how they can be enhanced to incorporate emerging technologies.

2.2 Anticipatory Technology Ethics (ATE) [1]

This framework provides a conceptual understanding of emerging technologies with three levels to distinguish the ethical analysis.(see Figure 1) [3]

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

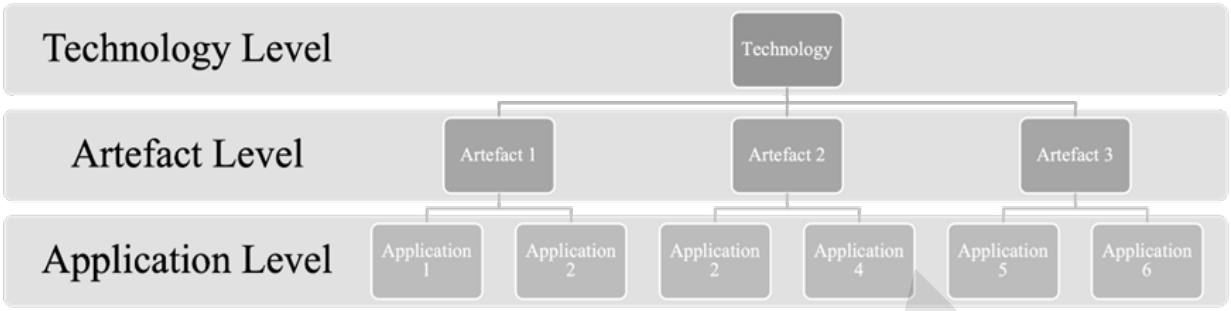

Figure 1: ATE uses three levels of ethical analysis [3]

(1) analysis of the technology (collection of techniques related to a common purpose or domain),

(2) analysis of the artefact (functional systems, artefacts and procedures based on a technology)

(3) analysis of the application level (the specific way in which artefacts are configured to be used).

The ATE postulates an identification stage at which ethical impacts are identified and descriptions of a technology (at the three levels mentioned above) are analysed by means of a list of ethical values and principles i.e. 'Brey's checklist'[1]. In addition, the framework proposes an evaluation stage, during which the relative importance of ethical impacts is assessed along with their likelihood of occurring.

As well as the three levels of analysis, ATE presents some suggestions for the future, since different forecasting methods are required for the technology, artefact and application levels. Some suggestions for future forecasting include the utilisation of existing studies in forecasting and Technology Assessment about the technology (to the extent that these are available). These provide ethicists with a first view of artefacts and applications that are likely to emerge in the future. Additionally, according to ATE, ethicists should initiate expert surveys and roundtable discussions with experts that yield expert predictions of possible or likely future artefacts and applications. Furthermore, other EU funded projects like SIENNA such as deliverable D6.1 (*General ethical assessment*) coupled the ATE framework with various foresight methodologies such as environmental scanning, science and technology roadmapping, multiple perspectives, and, optionally, future visioning. Foresight approaches where stakeholders beyond those of engineers or domain experts were considered in order to align the ATE approach with more general concerted efforts of Responsible Research and Innovation initiatives to include broader stakeholder communities. However, involving non-expert stakeholders remained only a contingent and optional step, rather than a requirement in the development of technology.

This framework is useful when considering the ethical analysis for new and emerging technologies that is comprehensive yet flexible enough to be used and tailored in different ways. Although the framework does suggest foresight activities, projections of the future are needed due to the level of uncertainty. Therefore, we recommend engagement with future studies.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

2.3 Anticipatory Technology Ethics (ATE) Plus [4]

This is a more comprehensive framework of the ATE approach that highlights further nuanced ways for distinguishing the levels and objects of analysis to better reflect the ontology of emerging technologies. The framework proposes a series of modifications to the levels and objects of ATE ethical analysis and the methods of foresight.

ATE Plus enhances the ATE framework to encompass the variety of human processes and material forms, functions, and applications that comprise the socio-technical systems in which these technologies are embedded, thus providing insights into the challenges of anticipating and responding to the potential impacts of emerging technologies. It does so by providing an analytical tool complementary to ethics-by-design approaches which consists of steps, namely, description, investigating philosophical ideas, identification of values and principles, narrative analysis, engagement with technology stakeholders and creation of a list of design questions. Further details of the ATE Plus approach can be found in Section 3 of this deliverable. The enhanced ethics assessment tool could potentially provide a more nuanced basis to develop ethics guidance in terms of informing ethics-by-design approaches where ATE is used early on in the design process to tease out important ethical issues.

2.4 Ethical Technology Assessment (eTA) [5]

The main purpose of this framework is 'to provide indicators of negative ethical implications at an early stage of technological development'[5]. The focus of eTA is on the whole life-cycle of technology development, from initial R&D to impacts on society. To attain an adequate understanding of future developments, eTA relies on studies in Technology Assessment (TA) and on close interactions with developers of technology. However, the goal of eTA is not to predict far into the future, but rather to continually assess current practices in technology development and provide feedback to designers and policy makers. The ethical analysis of an emerging technology takes place by confronting projected features of the technology or social consequences with ethical concepts and principles. The framework also proposes an ethical checklist of nine issues to identify the most common ethical issues in emerging technologies. This list contains issues like privacy, sustainability, issues of control, influence and power and issues of gender, minorities and justice.

Palm and Hansson's approach is one of the first ethical approaches explicitly targeted at emerging technologies. However, the approach has a few limitations. Firstly the methodology is vague, as it does not specify in detail what kind of knowledge needs to be acquired from technology developers and from the TA, and how it should be acquired. In addition, it does not describe in detail how ethical analysis can be performed on the basis of this knowledge. Furthermore, the ethical checklist of nine items seems somewhat limited, as many recognised moral values and principles are not found on the list, such as autonomy, human dignity, informed consent, distributive justice, etc. So it would seem one would need a much longer list to be able to do comprehensive ethical assessments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

of new technologies. To identify such issues, exploring moral intuitions of either stakeholders or the analyst may be in order [1].

2.5 Ethical Scenario Method [6]

This framework focuses on ethical assessments that are intended to help policy makers to anticipate ethical issues of emerging technologies. It relies on scenario analysis, which is a well-established approach within future studies. A unique feature of the approach is that it aims to anticipate the mutual interaction between technology and morality, and changes in morality that may result from this interaction. Such changes need to be taken into account when ethically assessing new technologies, so that new technologies are not evaluated from within a moral system that may not have the same validity by the time an emerging technology has become '*embedded*' in society. The ethical scenarios approach involves three steps. The first step, "sketching the moral landscape," aims to describe the new technology in question, as well as current moral beliefs, practices and regulations that are directly or indirectly relevant to the technology. The second step, "generating potential moral controversies, using NEST-ethics," aims to identify ethical issues and arguments regarding the new technology. This is done using the approach of NEST-ethics [7], which is an approach for identifying ethical issues and arguments in a new technology using a taxonomy of issues and arguments that have been used in past ethical controversies on technology. Finally, the third step of this approach is "constructing closure by judging plausibility of resolutions". In this step, the multitude of views and arguments from step 2 is reduced by imagining which resolution of the debate is the most plausible. The intention is to use steps 1 to 3 to develop a scenario of how the new technology will develop in the future, how this affects the moral landscape (i.e., moral beliefs, practices and regulations), and how moral closure is eventually reached. The scenario approach has some advantages such that it takes into account moral change over a larger time frame, However there are also some limitations to this framework. For example, it is a descriptive and predictive approach, rather than a normative and prescriptive one. It describes moral issues that are likely to emerge as the technology progresses, not ones that ought to emerge from an ethical point of view. In addition, the ethical issues are unlikely to draw much attention from the public that may be important. The ethical scenario approach may include moral controversies that are based on a false or misguided understanding of the technology or its social consequences. Such moral controversies do not present moral issues that ought to be considered in assessing emerging technologies, because they are based on false premises [1]. Furthermore, moral controversies that may emerge in public debate through stakeholder engagement may be different from moral issues that may result from thorough ethical assessments, even though there may be a large overlap in practice between the two. The TechEthos results from WP3 scenarios could be useful to understanding the ethical issues that arise from experts in the field of emerging technologies, as well as the linked third party workshops that would help to give insight into the public debate around ethics from a citizens perspective.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

2.6 Ethical Matrix [8]

The ethical matrix is a conceptual tool designed to help decision-makers reach sound judgements about the ethical acceptability of new and emerging technologies in the food and agriculture sector. The standard principles upon which the matrix is based are *respect for wellbeing, autonomy and fairness* (the columns in the matrix). The tool includes the affected parties that are relevant to these ethical issues consisting of different groups of people, such as consumers and food producers, but also non-humans, such as farm animals (the rows in the matrix). The ethical matrix is a starting point for ethical deliberation, and its essential features are a) the different perspectives of stakeholders or effected parties, b) different concerns according to which the ethical impact of a proposed technology may be analysed, c) *prima facie* application of ethical principles (where they are assigned a different weight according to the specific case they are applied to), and d) a theoretical, visual method for recording assessment.

Respect for:	**Wellbeing**	**Autonomy**	**Fairness**
Producers	Satisfactory income and working conditions	Managerial freedom	Fair trade laws
Consumers	Safety and acceptability	Choice	Affordability
Treated organisms	Welfare	Behavioural freedom	Intrinsic value
Biota	Conservation	Biodiversity	Sustainability

Figure 2: Generic ethical matrix [8]

Advantages of the ethical matrix include a clear definition of stakeholder groups into governmental advisory committees, ethics committees, non-governmental organisations, participants in exercises in public deliberation, commercial companies. Advantages also include flexibility of application, coupled with subjectivity and creativity in the evaluation process. This means that the matrix or parts of it could be adapted to frame the ethical acceptability of the TechEthos technology families and other technology families.

Drawing from TechEthos results from Deliverable 2.2, the matrix can be adapted and specified to the specific technology family like NT, CE or XR, so as to provide a structured way of working through ethical concerns (the ethical matrix specifically says that it is flexible and adaptive to specific contexts). Specifically, the Ethical Matrix's proposed ethical principles of wellbeing (XR), dignity (which can be seen to equate to 'autonomy') (NT) and justice (which can be seen to equate to 'fairness') (CE) can apply respectively to the technology families. A future studies-oriented approach could also be used to enhance the matrix framework, in order to project the ethical concerns of the technology families across a 20 year life span at the very least.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

2.7 IS Ethics Assessment Techniques [9]

There are a number of different ethical and ethics related frameworks within the umbrella of IS ethics assessment techniques. One of the leading approaches seen within this is the Ethical Issues of Emerging ICT Applications (ETICA, www.etica-project.eu), which was an EU funded project focused on the societal impact of perceived high level technologies and their possible consequences. This project adopted a future perspective, as in what might be expected from ethics issues in emerging IS technologies (see Table 1). This approach aimed to prevent or minimise the possible future negative outcomes associated with a technology, through offering possible solutions.

Starting from a definition of emerging technologies as those likely to see significant development within 10-15 years (so not too far in the future), it used public and academic literature to identify 11 'key' probable emerging technologies (Affective Computing; Ambient Intelligence; Artificial Intelligence; Bioelectronics; Cloud Computing; Future Internet; Human-machine symbiosis; Neuroelectronics; Quantum Computing; Robotics and Virtual/Augmented Reality). Using a mix of bibliographic analysis and structured literature review, the main ethical concerns were uncovered, see Table 3 [10]. As is common to other findings (including TechEthos), some of these concerns are cross-cutting (i.e. Privacy, Autonomy, Digital Divide/Equity, Informed Consent) and applicable to a wide range of technologies, while others are much more technology specific. The ETICA project then ordered these ethical issues and identified how they could be addressed through governance arrangements.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

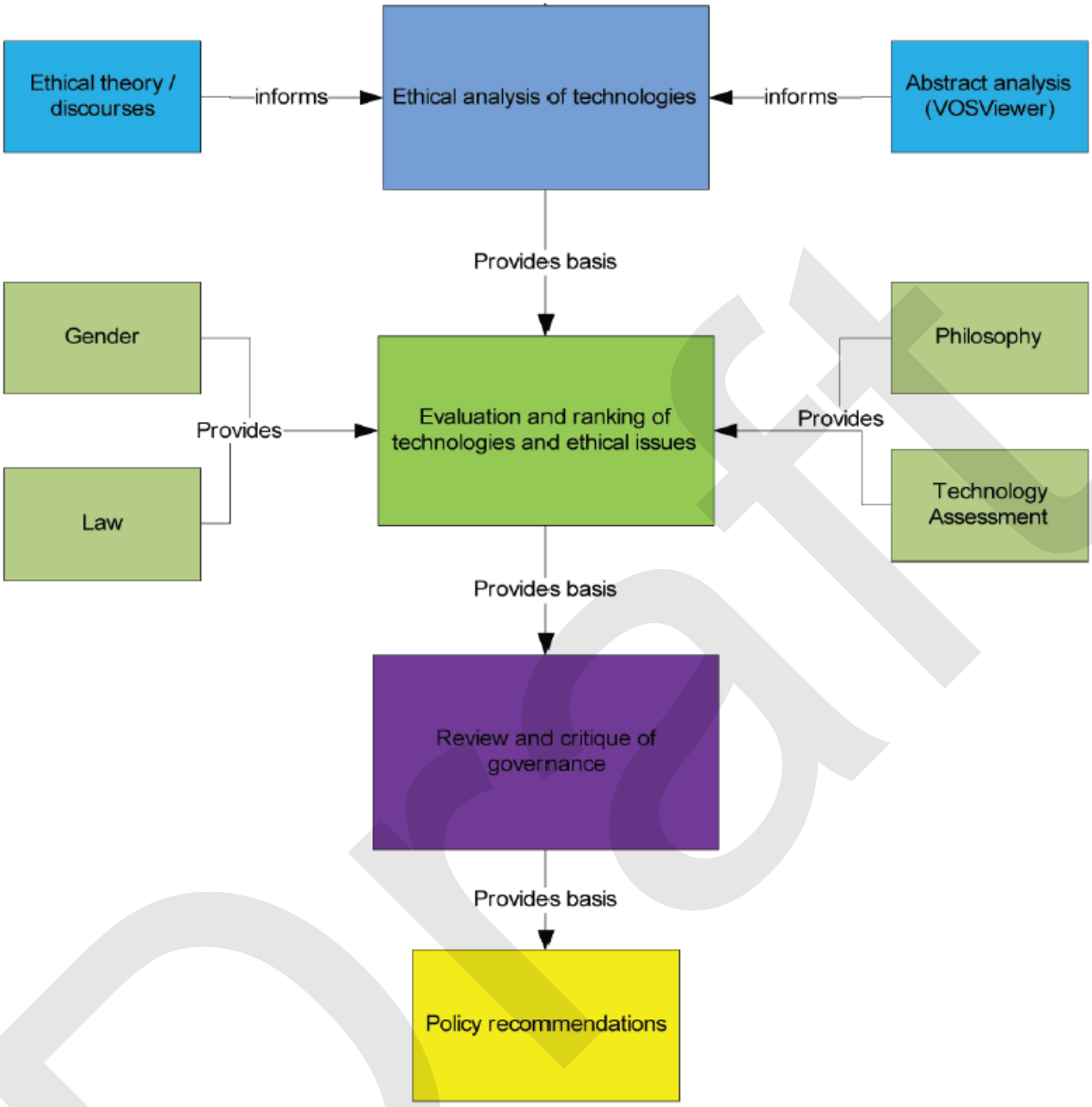

Figure 3: Representation of part of the ETICA process [9]

Advantages of the ETICA approach are that having identified the ethical issues in 11 emerging ICT applications (at that time), it then suggested possible recommendations and governance for addressing the ethics of these emerging ICTs. It also suggested another set of recommendations which were developed for industry, researchers and civil society organisations (CSOs), who could be seen as users of these technologies. However, ETICA also faced a number of problems. Rather like TechEthos, the concept of 'emerging ICT' was difficult to identify, especially what might constitute what is meant by both 'ICT' and 'emergence' in this context.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Given that in TechEthos the first stage of ETICA, i.e. the identification of the emerging technologies, has already been done through a specific horizon scanning process in WP1, it is only the second stage that might be applicable/useful (see Figure 3). This is the use of bibliographic analysis and structured literature review to identify the main ethical concerns for each of the three technology families, CE, NT, dXR. This has been carried out in Work Package 2, notably in Deliverable 2.1 and Deliverable 2.2.

Table 3: Types of emerging technologies and their ethical concerns [10]

Emerging technologies	**Brief description**	**Ethical concerns of the technology**
Affective computing	Also known as emotional computing which can be used to compute human emotion. This is done by increasing social-emotional intelligence in agents and robots. Anthropomorphism Privacy Cultural differences Responsibility Informed Consent - Affective Contractualism Disclosive ethics	• Conceptual Muddle and trust • Measurement and interpretation errors • Persuasion and coercion
Ambient Intelligence	Ambient Intelligence (AMI) technologies that are embedded, interconnected and unobtrusive in a user's environment. They adapt and anticipate as well as context aware of a user's environment.	• Privacy, Surveillance and Data Protection • Autonomy, • Freedom and Agency • Equity • Liability
Artificial Intelligence	Artificial intelligence (AI) is diverse and deals with machine intelligence. Can be used for data mining, industrial robotics, and speech recognition among others.	• Responsibility gap • Moral worth of machines • Human replacement • Privacy • Privacy – Artificial Intelligence • Surveillance • Digital Divide

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Emerging technologies	Brief description	Ethical concerns of the technology
Bioelectronics	Bioelectronics is a multidisciplinary research area and integrated into the research areas of cognitive sciences, nanotechnology or neurotechnologies. The possible application areas are infinite and numerous and can include wearable as well as monitoring technologies.	• Autonomy
Cloud Computing	Cloud computing is a recent trend in IT that moves computing and data away from desktop and portable PCs into large data centres. It basically means that software, different kinds of services and applications are all delivered as services over the Internet as well as to the actual cloud infrastructure [11], [12].	• Control and responsibility • Problem of many hands • Self-determination • Accountability • Ownership • Function creep • Monopoly and lockin • Precautionary • principle • Privacy in relation to the technology • Privacy across in relation to cultural borders • Cultural imperialism and dealing with diversity
Future Internet	One important aspect of FI is that the Internet will extend outside the traditional computer devices so that any objects in the environment can be connected to it. This is called the Internet of Things (IoT).	· Privacy and Security · Trust · Acceptance · Digital Divide - FI · Intellectual Property Rights Issues · Openness · Energy

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Emerging technologies	Brief description	Ethical concerns of the technology
Human-machine symbiosis	One of the major defining characteristics of human machine symbiosis is that of interaction. For humans and machines to mutually work together and be effective in for instance brain-computer interaction and/or in performing other highly challenging activities, interactivity is central to the technology.	· Therapy vs. Enhancement · Normality · Human Dignity · Risk and responsibility · Humanness · Immortality · Good life · Self-centeredness · Identity and personality · Autonomy - HMS · Freedom of choice/autonomy · Special groups · Privacy - HMS · Equality and fairness · Social disruptions and institutional problems
Neuroelectronics	Neuroelectronics, sometimes referred to as neurotechnology, is the discipline that deals with the interface between the human nervous system and electronic devices. Neuroelectronics is a highly complex and interdisciplinary field with contributions from computer science, cognitive science, neurosurgery and biomedical engineering.	· Brain Computer Interface and neural stimulation · Safety · Technology risks · Responsibility for harm · Social effects · Social pressure · Authenticity · Agency and Autonomy · Paradox of recovery · Sub-personal use of human beings

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Emerging technologies	Brief description	Ethical concerns of the technology
Quantum Computing	Quantum computation is strongly seen to efficiently solve some of the most difficult problems in computational science and in a way change dramatically the development and implementation of information and communication systems of the future (e.g. integer factorisation, discrete logarithms, and quantum simulation and modelling that are intractable on any present or future conventional computer).	• Errors and misunderstanding • Encryption • Natural ethics • Control on research
Robotics	Robots are machines with motor functions that are able to perceive their environment and operate autonomously so that they can replace human effort. Below are a number of features that define robots in the military, households and healthcare.	· Sensory Perception, privacy and · surveillance · Tele-presence · Robot Autonomy · Robot Responsibility · Moral Obligation · Privacy - Robotics · Overtaking Humankind · Robot Rights · Man–machine Interaction
Virtual/Augmented Reality	VR originally referred only to a completely immersive virtual reality or virtual environment. Today the term virtual reality is also used to describe non-immersive or partial immersive applications, although the boundaries are becoming obscure (Beier 1999).	· Escapism · Personal Harm caused by virtual reality · Blurring of Real and Virtual · Violent Content · Denial of Virtual Harm · Alternative Rules · Access resulting in Digital Divide · Autonomy related to virtual reality · Privacy related to virtual reality · Addiction · Designers Responsibility

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

2.8 Ethical Impact Assessment (EIA) [13]

The framework identifies key social values and ethical issues, provides some brief explanatory contextual information which is then followed by a set of questions aimed at the technology developer or policy-maker. The aim of this framework is to facilitate consideration of ethical issues, in consultation with stakeholders, which may arise in their undertaking. In addition to consultation with stakeholders, the framework includes a set of ethical tools and procedural practices which can be employed as part of the ethical impact assessment. The ethical tools help the technology developer to get a better idea of how the technology is perceived ethically by stakeholders. Furthermore, the framework provides a diagrammatic pathway which is useful to technologists in order to check and review potential ethical challenges and to mitigate some of the risks. The EIA framework consists of the following steps: 1) conducting an EIA threshold analysis, 2) preparing an EIA plan, 3) identifying ethical impacts 4) evaluating the ethical impacts (step 3 and 4 are to be carried out in consultation with stakeholders), 5) formulating and implementing remedial actions, 6) reviewing and auditing the EIA.

The EIA framework does not account for emerging technologies in the future, but investigates continuously the ethical implications of what is known about the technology under development. However, as there are often inherent privacy issues such as equality and human dignity etc. in new and emerging technology, research has also been carried out to integrate privacy impact assessment into EIA [14].

2.9 SATORI CWA 17145-2 [15]

The SATORI (Stakeholders Acting Together On the ethical impact assessment of Research and Innovation) research project, funded by the European Commission, developed a framework for common, basic ethical principles and joint approaches and practices for improving ethics assessment practices of Research and Innovation. The Cen Workshop Agreement (CWA) consists of two parts, the first part provides recommendations for ethics committees on practices and procedures; the second part provides researchers and organisations with guidance on ethical impact assessment.

The SATORI CEN Workshop Agreement has the following unique features: it is the first international standard document for ethics committees and ethical impact assessment of research and innovation; it is based on an extensive study of hundreds of existing ethics documents; it is a comprehensive standard covering all fields; it includes ethics assessment guidelines; it addresses procedures for the establishment and composition of ethics committees; it addresses quality assurance in ethics assessment; it contains a clear methodology for assessing ethical impacts; it can

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

be adapted to different value systems and cultural contexts; it is applicable to different organisational contexts.

The CWA is a comprehensive approach for ethically assessing the actual and potential mid- and long-term impacts of research and innovation on society. This particular approach does not seem to have a future element and it would benefit from integration of a Future studies approach, and also with scenario-building exercises as embedded in WP3 in TechEthos.

2.10 Future Studies [16]

Future studies is an eclectic approach to forecasting which considers the sociological, anthropological, technological and scientific approaches to the future that 'is still being made: it is what people can shape and design through their own actions' [16], [17] acknowledging that recognition of a problem, for example global warming, does not necessarily provoke actions.

The passage of time is experienced as ontogenetic (level of being) and phylogenetic (level of the species). Known by other nomenclatures such as 'futures research', 'futuristics' or 'prognostics', their theories aim to 'discover or invent, propose, examine, and evaluable possible, probable, and preferable futures.' be [16], [18]. There are practical dimensions to the approach, constructing visions to achieve desirable futures. Since the genre has its origins in fiction, Henry More's classic *Utopia* [16] is a case in point, otherness is inherent to it. There is some *other future* that is possible to construct. Hence future studies explore the social constructedness of possible futures. The French philosopher of the Enlightenment and French Revolution Marquis de Condorcet [1743-1794] took the role for the future from God, into the hands of human beings (women would have largely been excluded from shaping this conversation until the twentieth century). Contemporary approaches draw on an eclectic interdisciplinary heritage. Van Lente and Peter's [17] work draws on the imagination, art and the aesthetic experience. In a counter move to the original Enlightenment goal of planning, they instead point to the work of John Dewey's *Art as Experience* and regard art as a "mode of prediction not found in charts and statistics" (p. 1).

Early advances in Future Studies, such as the work of W.F Ogburn's informed technology assessment [16], which later informed the 'ethics technology assessment' (aTE) developed by Brey (2012). Future studies theorists analyse state level interventions including in health, banking, education or politics, but often outcomes diverge from intentions. But, the approach can be traced to Norman Henchey's 1978 paper *Making Sense of Future Studies*. He identified three factors which led to interest in the field: Disillusionment, Anxiety and Expectation.

Technological progress is strongly connected to growth and economic expansion with many different publications shaping the political goal of the relentless growth of capitalism. The Limits to Growth (1972) was regarded as a landmark publication in this respect. Technologically organised economic growth is resource intensive, activities continue as unabated, but with different

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

configurations. A post-industrial Europe and North America has given way to dynamic East Asian economies, with China accounting for 28 per cent of global manufacturing in 2018 (WEFORUM).

Future-focused methodologies have innovated techniques such as *backcasting*, working from the premise of an alternative future, then working backwards to map the stages of achieving it [19]. In the last decade Future Studies approaches to technologies has led to work on global warming [17], the internet of things [20] and imaginaries of digitally mediated governance (Mager and Katzenbach 2021). Moreover, some of its approaches take an extreme social constructivist approach, implying that the external world is shaped by perception of it, and that the way one sees the external objective world shapes it, and the possible futures [21]. The approach's usefulness comes from recognising the dynamic interplay of heterogeneous actors, with cause and effect contingent, but taken to its logical conclusion there is no objective world outside the actors that create the future. That said, attention is paid to the sex, race, class and ability/disability hierarchies that structure future relations, as well as conflating science and myth, allowing a broader spectrum of ideas to be valued and methodologically included in planning, forecasting, backcasting, and theorising. Inspired by predictive social sciences, critical philosophy and poststructuralism, Future Studies narratives places' ethics within larger socio-political processes.

Due to its theoretical soundness and its intent to problematise the very concept of 'future' itself, future studies provides a much needed critical lens with which to approach the study of the ethical impact of emerging technologies. For this reason it has been incorporated in several stages of the TechEthos project, including digital ethnographies and ethical analysis in WP2, and scenario building in WP3.

3. Methodology for ethical framework development

The methodology commenced with a review of current ethical frameworks in the literature (see section 2), to see how these frameworks can be complemented by methodologies from WP2 (D2.2 *Identification and specification of potential ethical issues and impacts and detailed analysis of ethical issues*) and WP3 (D3.1 *Evolution of advanced Techethos scenarios*). The aim of this task is to propose recommendations to enhance the existing ethical frameworks to include emerging technologies in general, as well as the three specific technology families.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

4. Selecting ethical frameworks to enhance for emerging technologies

Having reviewed several approaches to ethical frameworks that exist in the literature and have previously been applied in a range of technology contexts. While this is not a comprehensive list, it aims to identify the key criteria in each approach in order to assess the usefulness with respect to emerging technologies. This can be illustrated in the Appendix I of this report.

Based on this critical evaluation of frameworks, the key criteria that were identified were:

- What are the advantages of the framework i.e. does it have demonstrable benefits? Or has it been implemented and used by organisations.
- What are the disadvantages (e.g. no evidence of the framework being used)
- Can the framework be applied to emerging technologies and how can we measure its effectiveness?
- Does the framework have a futures element?
- Can the framework be enhanced or refined with respect to the methodologies of TechEthos (in particular, WP2 and WP3)

Based on these criteria, the three approaches selected for review are:

1. Anticipatory Technology Ethics plus (ATE+) (which is an enhancement of the current Brey ATE framework),
2. Ethical Impact Assessment (EIA)
3. The Future Studies approach

4.1 ATE plus (Enhancement of Brey's ATE framework)

A strong foundation for evaluating potential issues/challenges with novel or developing technology is provided by the original ATE framework. However, the TechEthos project identified some gaps that needed to be addressed before the ATE approach could be implemented, including the necessity to bring values and principles into an *a priori* conversation with technology. In this spirit, we began the project with the original ATE formulation. However, the project did bring to light certain requirements that needed to be addressed for a more relevant technology assessment, which is essential for ethics-by-design.

In the initial ATE formulation, there were also some potential issues that were identified. Firstly, the original ATE framework did not give sufficient consideration to foresight concepts and activities. These concepts and activities should be made core, rather than optional in an enhanced framework; secondly, a wider spectrum of stakeholders and publics would also provide opportunity for an enhanced framework, since ATE exclusively focuses on 'expert' stakeholders. Instead, the wider

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

stakeholder engagement could include, for example, consultation with under-represented groups as carried out in WP3 of TechEthos; an enhanced framework would also need to consider the different temporal horizons affecting the impacts of emerging technologies. For example the scenario workshops in WP3 took into consideration new and emerging technology in a 20 year time span. Lastly, ATE ethical analyses were heavily focused on potential negative effects, and neglected to include any significant potential positive effects (such as the counterfactual argument which highlights the potential negative effects of not using a specific technology, under extreme conditions). The consequence is that ATE would leave out conflicts about the subjectivity of assigning positive and negative 'effects' to a technology - in other words, would rule out the opportunity for interpreting the overall impact of an emerging technology. Judgements about fairness concerns over how technology's obligations and benefits are distributed.

Therefore, based on this review of shortcomings and on the TechEthos project requirements, we have been able to highlight some lacunae in the original ATE approach, in Table 4 below,

Table 4: Lacunae in ATE [4]

Lacunae
Meaningful consideration of foresight concept/activities
Thoughtful inclusion of non-expert stakeholders and publics
Clear explanation of time horizons to consider when adopting the approach
Considering impacts beyond those that are negative

A number of changes to the levels and objects of ethical analysis have been proposed for ATE, including the methods of foresight, as well as the methods of ethical analysis, in light of the benefits of ATE and the lacunae identified. The following phases are the suggested improvements to the process of ethical analysis in ATE, which cut through the many layers, based on our experience with TechEthos (as addressed in WP2 and published in Adomaitis, Grinbaum, and Lenzi, 2022). In order to provide an enhanced opportunity for anticipatory technology ethical analysis, ATE could:

1. Describe relevant objects, procedures, techniques, approaches, applications, and use cases (for instance, natural language processing, virtual reality, or the usage of digital twins in training or healthcare);

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

2. Investigate fundamental philosophical ideas and conundrums that provide conceptual support for the moral challenges (e.g., is there a natural preference for real life over virtual reality?);
3. Identify values and principles (e.g., transparency, dignity) and return to step 2 for clarification if necessary;
4. Utilise narrative analysis to distinguish between morally clear ethical concerns and morally murky presuppositions in technological judgement of the values and principles described in step 3 (such as "Be careful what you wish for" and "The rich get richer, the poor get poorer");
5. Engage key technology stakeholders ethnographically through narratives as opposed to the inclusion of open-ended questions.
6. Create a list of operationalised design questions that can be asked about the use of techniques (or applications and use cases), such as: Does the XR system account for prospective changes in its users' behaviour? Who benefits from the behavioural changes, and how are the changes sparked?

Given that TechEthos' main goals are to provide assistance for ethically creating technologies, the project has made changes to the ATE framework to inform the creation of ethics-by-design guidelines. Therefore, through TechEthos, some of the lacunae in ATE have been matched to their potential modes for enhancement as described below in Table 5.

Table 5: Lacunae in ATE matched to modes for enhancement [4]

Lacunae	**Potential Tools/Variables for Enhancement**
Meaningful consideration of foresight concept/activities	Narratives approaches, including lay narratives, cultural narratives
Thoughtful inclusion of non-expert stakeholders and publics	Uncertainty (mapping procedures of how to characterise uncertainty)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Clear explanation of time horizons to consider when adopting the approach	The inextricability of some opaque element even when transparency is promoted.
Consideration of potential impacts beyond negative ones	Including socially beneficial impacts

4.1.1 Description of potential Tools/Variables for Enhancement

Narratives approaches

In the TechEthos project, we focused on examining the construction of narratives during technology development. We created contrasting future scenarios (WP3) to elicit a range of perspectives on social and ethical issues. To ensure that the scenarios were effective, we made sure that they were plausible and maintained internal consistency within the social, technical, economic, environmental, political, and value dimensions of the scenarios [22]. Instead of identifying the most probable scenarios, this approach centred on exploring multiple plausible futures, each highlighting different ethical aspects. Therefore, in the operationalisation of ATE Plus in TechEthos, we proposed shifting the focus from "likely futures" to questions about "plausible futures", so as to encourage reflection on the social, ethical, environmental, economic, and other potential impacts.

Uncertainty

Another concern with the original ATE framework is that it does not explicitly address uncertainty or ambiguity in foresight activities related to future considerations. Research on decision-making under conditions of uncertainty (i.e. related to the use but also non-use) demonstrates that context can impact perceptions of uncertainty [17]. The emotional aspect of certainty evaluations [24] and the tendency to replace the concept of delay with that of risk [25] are also important factors to consider. By neglecting uncertainty in favour of likelihood, the original ATE framework failed to fully explore the governance of science and technology that is intrinsically linked to uncertainty [26]. Furthermore, grappling with uncertainty head-on can surface important ethical dilemmas, such as those related to constructing uncertainty as a political device [27], [28]. Thus, addressing uncertainty in ethical analyses is essential for exploring the governance and ethical dilemmas inherent in science and technology. In TechEthos we addressed uncertainty by investigating imaginations of the future created by businesses involved in the development of commercial technological applications of the technology families through digital ethnographies (WP2).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Inextricability of time horizons

In the original ATE framework, the analytical considerations extend to both present and future states of technologies, artefacts, and applications. TechEthos employs Future Studies methods such as forecasting, technology assessment, expert surveys and ethnographies (WP2), workshops with under-represented groups and scenarios to address future-oriented concerns (WP3).

Probability of consequences

The ATE framework concentrates on the probability of potential unintended consequences of technology development or artefact application, as assessed by experts, within their respective contexts. However, the emphasis on "considered likelihood" in the original ATE raises an important question about human perception. That is, whose perception is being considered likely, what evidence is it based on, and what motivations are influencing these perceptions? Consultation with non-expert stakeholders, brings to light insights from under-represented groups and the public.

4.2 Future Studies Approach

Future Studies is not a coherent body of literature, but its approaches centre around contingency, subjectivity, and conscious planning. The view of these approaches is that the future does not merely come into being but is socially constructed by its citizens. The stratification of citizens according to power dynamics - sex, race, class, ability - subsequently shapes what futures are permissible, considered or occur - thus Future Studies incorporate these aspects into its epistemological concerns. The future is not a neutral point in a clearly defined temporal point, but a series of concatenating heterogeneous factors. State and business have extensive resources to shape the narrative of the future, but citizens provide alternative perspectives. Future Studies approaches examine the multiple layers of political engagement, recognising the social, economic, political, and legal shaping of the future. In other ways, Future Studies can be regarded as an extreme example of social constructivism, and thus words and subjectivity can be overplayed and valued beyond material resources or barriers. While non-state and non-corporate actors contribute significantly to influencing narratives, the resources at their disposal are severely limited compared to state and corporations. Ultimately, Future Studies approaches regard ‘the Future’ as a fiction that can be indefinitely created by actors that inform it, while attempting, but not always succeeding, to balance out conflicting interests and priorities mediated through power, resources, wealth, influence, physical, social or legal trajectories. In its favour it offers an approach to the future that aims to synthesise contingency, agency and imagination of actors and offers backcasting as a way to read the past as a guide to the time that has not yet arrived.

Therefore Future Studies offer a problematisation of the future that can form a useful theoretical framework to underpin any study that claims to be looking at time - so any research project that looks at time will benefit from a more thorough conceptualisation of what is fundamentally a

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No.101006249.

philosophical concept. As such although not strictly framed as ethical, Future Studies can enhance any ethical frameworks for analysis that set out to investigate the long term impact of new emerging technologies.

4.3 Ethical Impact Assessment (EIA)

As the ATE Plus focuses on ethical values and principles, and Future Studies focusses on problematising the future, the EIA framework emphasises stakeholder engagement. The EIA framework is specifically about people and public dialogue and therefore offers a different and complementary perspective. This framework raises questions aimed at the technology developer or policy maker to facilitate consideration of ethics, in consultation with a variety of stakeholders. Although this framework does not specifically account for future emerging technologies, it investigates continuously ethical implications of what is known about the technology under development. Essentially, the framework is supported by ethical tools that aim to help the developer to get a better idea of how the technology is perceived ethically by stakeholders and what measures could be adopted to ensure that the technology is ethically acceptable or what alternatives might be available. This particular approach can be enhanced by methodologies of the TechEthos project, for example the key ethical values and principles from WP2 (deliverable D2.2) as well as ethical issues raised in engagement with expert stakeholders through scenarios and public consultation including under-represented groups (WP3). The involvement and emphasis of diverse stakeholders supports the systematic reflection of ethical issues in decision-making through independent evaluation, and supports the explicit communication about values [13] .

Furthermore, this accepted framework has been implemented in the SATORI CEN CWA 17145-2 ethics assessment pre-standard for Research and Innovation. This CEN Workshop Agreement (CWA) sets requirements and provides guidelines for ethics assessment of research and innovation. It is a policy-oriented guide for researchers and ethics assessors on the different stages of the ethical impact assessment (EIA) process. This reinforces the effectiveness of the EIA framework and illustrates how it can be further enhanced when used in combination with ATE plus and Futures studies, to make this useful for new and emerging technologies.

5. Outline of the framework that supports the ethical governance of new technologies

Various ethical frameworks from literature have been analysed to assess their usefulness in anticipating potential impacts with respect to emerging technologies. Although many of the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

existing approaches share similarities, they each have their own strengths and limitations. Therefore, in order to achieve an outcome capable of being applied across a range of emerging technologies, we have chosen to take one approach and add to it from the empirical studies carried out as part of the method used within TechEthos.

The enhanced ATE Plus, which builds on the original ATE approach [3], aims to assess technological innovations by providing an analytical tool complementary to ethics-by-design approaches to engineering novel technologies. Furthermore, in combination with Ethical Impact Assessment (EIA) which emphasises stakeholder engagement, and Future Studies the element of uncertainty will be considered through forecasting and scenario development. Therefore, this report proposes an enhanced ethical framework to support the effective governance of new technology which has been further enhanced by results of the TechEthos project, namely the methodologies of WP2 and WP3. The 'TechEthos Anticipatory ethics Matrix' (TEAeM) will be useful for researchers, academics and policy-makers wanting to assess the ethical issues of emerging technologies and to mitigate these risks (Figure 4). Note that TEAeM is intended to be used in such a way that the ordering of the various matrix elements can be done in a range of ways, depending on the specific emerging technology under scrutiny.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

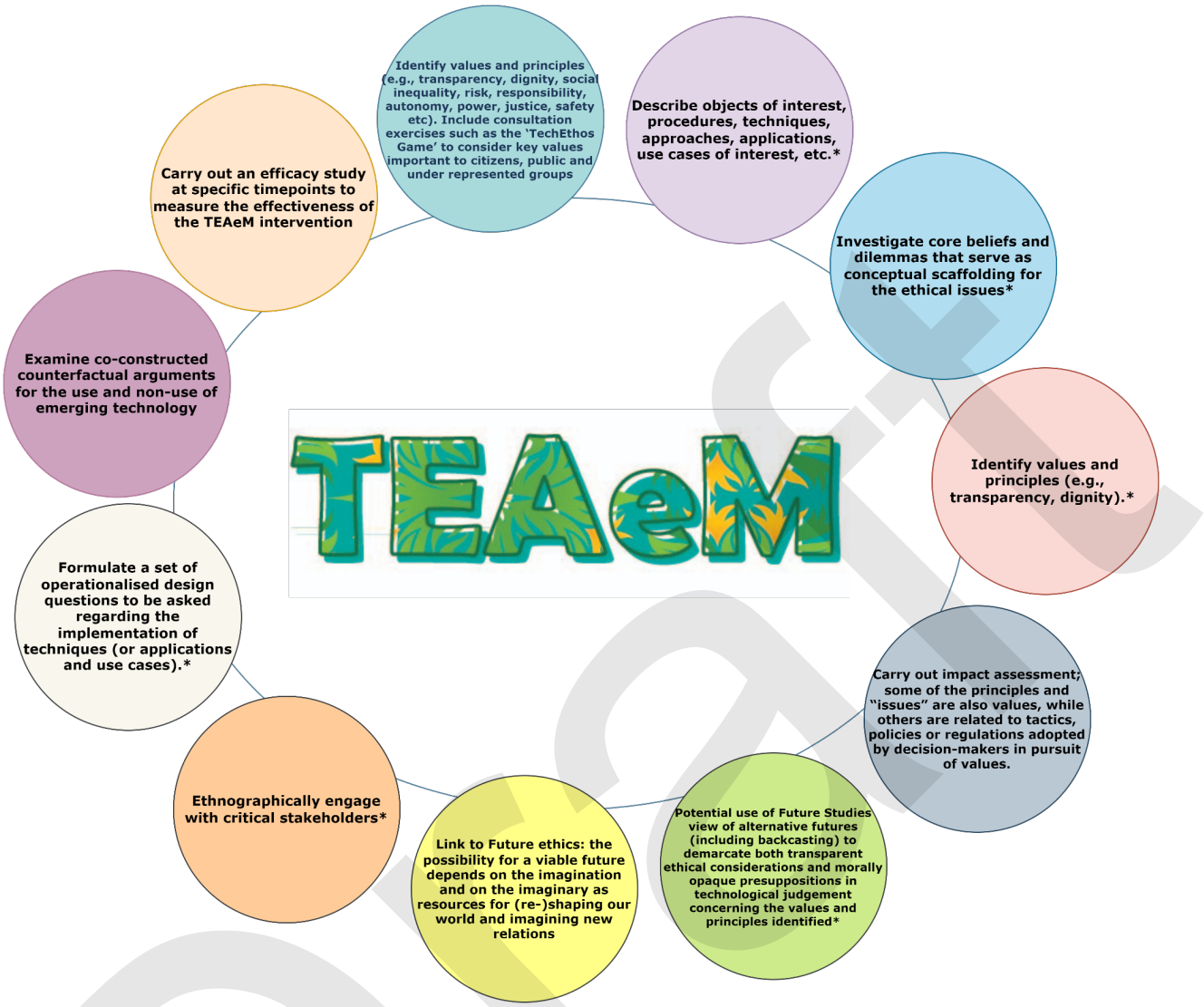

Figure 4: *The TechEthos Anticipatory ethics Matrix (TEAeM)*

Further explanation of the various elements that comprise TEAeM are presented in Table 6, in order to provide the starting points or relevant questions that could be asked in each of the matrix elements. As has been noted above, although TEAeM is presented in a tabular format in Table 6, and could be carried out in this way, it is intended that this is a flexible approach, that is responsive to the needs for specific emerging technology and so steps can be returned to or the order adjusted as and when needed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Table 6: Explanation of TEAeM elements

TEAeM	Explanation of the TEAeM elements
Describe objects of interest, procedures, techniques, approaches, applications, use cases of interest, etc.	What are the main goals or features of the technology, application, us e case, etc.?
Investigate core beliefs and dilemmas that serve as conceptual scaffolding for the ethical issues	Starting from societal, cultural, religious and legal issues in location of development
Identify values and principles (e.g., transparency, dignity, social inequality, risk, responsibility, autonomy, power, justice, safety etc). Include consultation exercises such as the 'TechEthos Game' to consider key values important to citizens, public and under-represented groups	Relevance to each technology family (if appropriate include cross cutting ethical issues too), eg. TechEthos deliverable D2.2 of key values and principles; eg. TechEthos deliverable D3.1 on the outcomes of using TechEthos game with underrepresented communities
Carry out impact assessment. Some of the principles and "issues" are also values, while other issues are related to tactics, policies or regulations adopted by decision-makers in pursuit of values (like data protection).	Use one of a range of impact assessment tools (accepted I.A or company specific) to identify what are the potential impacts of the technology, as it currently stands. Use of academic and grey literature, as well as potentially relevant policy documents, to establish the set of values that have been linked to technology or application in question
Potential use of Future Studies view of alternative futures (including backcasting) to demarcate both transparent ethical considerations and morally opaque presuppositions in technological judgement concerning the values and principles identified.	For example, creation of scenarios and other stakeholder engagement activities around various emerging technologies in the near and middle future contexts to help developers, users and others to think about the range of issues, both transparent and opaque.
Ethnographically engage with critical stakeholders	Use LinkedIn to search for companies working in the particular technology area and then review websites/videos, etc., using a direct or digital ethnography approach.
Link to Future ethics: the possibility for a viable future depends on the imagination and on the imaginary as resources for (re-	Use of future oriented analysis in the direct or digital ethnography, to establish what kind of future is being envisioned by developers and

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

TEAeM	Explanation of the TEAeM elements
)shaping our world and imagining new relations.	application experts and organisations. Embed contingency into the analysis.
Link to empirical data: aim to stay in contact with technology developers during the whole developmental process and discuss different approaches to problems that arise. … Continuous dialogue and repeated assessments are preferable to one single large-scale assessment.	Engage with developers and users in ongoing dialogue with them about problems that arise in the development and application processes. Use of databases, such as Cordis, to identify research projects in the appropriate field and contact them to establish a set of experts that can also be consulted with
Formulate a set of operationalised design questions to be asked regarding the implementation of techniques (or applications and use cases).	Use the results from the various analysis carried out in the previous stages to create the set of relevant design questions, using an ethics-by-design approach
Carry out an efficacy study at specific timepoints to measure the effectiveness of the TEAeM intervention	Review and reflect on the TEAeM process, with measures to identify any changes seen, eg. whether developers incorporated any of these changes into their practices
Examine co-constructed counterfactual arguments for the use and non-use of an emerging technology	Reflect on the ethical conundrum of risks of omission or inappropriate prevention (non-use of a technology with desired outcomes), which stand in tension with risks of commission (e.g. undesired consequences from technology use), eg. CDR non-use results in greater harm to humans and environments compared to the world with CDR use. SRM non-use results in a possibility of more severe harm than with SRM use.

As noted already, predicting the future, especially one where technology is involved and could spiral into many different directions, is almost impossible. Equally, to create an ethics framework that works for one specific technology would have been easier and perhaps more obviously applicable. However, in developing TEAeM for researchers/academics and policy makers, the aim was to enable future and emerging technologies to be able to be developed in a more ethically informed way (i.e. ethics-by-design) and as such we cannot yet know what those technologies might look like. Hence, we created an ethical framework that could support the ethical governance of the broadest range of future technologies and in doing so support a more ethical society.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

6. References

[1] P. A. E. Brey, 'Anticipatory Ethics for Emerging Technologies', *NanoEthics*, vol. 6, no. 1, pp. 1–13, Apr. 2012, doi: 10.1007/s11569-012-0141-7.

[2] W. Reijers, 'Methods for Practising Ethics in Research and Innovation: A Literature Review, Critical Analysis and Recommendations'', *Sci. Eng. Ethics*, vol. 24, no. 5, pp. 1437–1481, 2018, doi: 10.1007/s11948-017-9961-8.

[3] P. A. E. Brey, 'Anticipating ethical issues in emerging IT', *Ethics Inf. Technol.*, vol. 14, no. 4, pp. 305–317, Dec. 2012, doi: 10.1007/s10676-012-9293-y.

[4] S. Umbrello, 'From speculation to reality: enhancing anticipatory ethics for emerging technologies (ATE) in practice'.

[5] E. Palm and S. O. Hansson, 'The case for ethical technology assessment (eTA)', *Technol. Forecast. Soc. Change*, vol. 73, no. 5, pp. 543–558, Jun. 2006, doi: 10.1016/j.techfore.2005.06.002.

[6] M. Boenink, T. Swierstra, and D. Stemerding, 'Anticipating the interaction between technology and morality: A techno-ethical scenario study of experimenting with humans in bionanotechnology', *Stud. Ethics Law Technol.*, vol. 4, no. 2, pp. 1–38, 2010.

[7] T. Swierstra and A. Rip, 'Nano-ethics as NEST-ethics: Patterns of Moral Argumentation About New and Emerging Science and Technology', *NanoEthics*, vol. 1, no. 1, pp. 3–20, Mar. 2007, doi: 10.1007/s11569-007-0005-8.

[8] B. Mepham, M. Kaiser, E. Thorstensen, S. Tomkins, and K. Millar, 'Ethical Matrix Manual', *Agric. For. Meteorol. - AGR For. METEOROL*, Jan. 2006.

[9] B. C. Stahl *et al.*, 'Identifying the Ethics of Emerging Information and Communication Technologies: An Essay on Issues, Concepts and Method', *Int. J. Technoethics*, vol. 1, no. 4, pp. 20–38, Oct. 2010, doi: 10.4018/jte.2010100102.

[10] K. J. Wakunuma and B. C. Stahl, 'Tomorrow's ethics and today's response: An investigation into the ways information systems professionals perceive and address emerging ethical issues', *Inf. Syst. Front.*, vol. 16, no. 3, pp. 383–397, Jul. 2014, doi: 10.1007/s10796-014-9490-9.

[11] M. Dikaiakos and G. Pallis, 'Cloud Computing:Distributed Internet Computing for IT and Scientific Research', *IEEE Internet Comput.*, pp. 10–13, Jan. 2009.

[12] 'Gartner's Hype Cycle Special Report for 2009', *Gartner*. https://www.gartner.com/en/documents/1108412 (accessed Jun. 05, 2023).

[13] D. Wright, 'A framework for the ethical impact assessment of information technology', *Ethics Inf. Technol.*, vol. 13, no. 3, pp. 199–226, 2011.

[14] D. Wright and M. Friedewald, 'Integrating privacy and ethical impact assessments', *Sci. Public Policy*, vol. 40, no. 6, pp. 755–766, Dec. 2013, doi: 10.1093/scipol/sct083.

[15] SATORI CEN Workshop Agreement, 'Ethics assessment for research and innovation - Part 2: Ethical impact assessment framework', CEN CWA 17145-2 2017 Workshop Agreement, 2017. [Online]. Available: http://satoriproject.eu/media/CWA17145-23d2017.pdf

[16] W. Bell, 'The Sociology of the Future and the Future of Sociology', *Sociol. Perspect.*, vol. 39, no. 1, pp. 39–57, Mar. 1996, doi: 10.2307/1389342.

[17] H. van Lente and P. Peters, 'The future as aesthetic experience: imagination and engagement in future studies', *Eur. J. Futur. Res.*, vol. 10, no. 1, p. 16, Dec. 2022, doi: 10.1186/s40309-022-00204-8.

[18] N. Henchey, 'Making Sense of Future Studies', 2023.

[19] S. E. Bibri, 'A methodological framework for futures studies: integrating normative backcasting approaches and descriptive case study design for strategic data-driven smart sustainable city planning', *Energy Inform.*, vol. 3, no. 1, p. 31, Dec. 2020, doi: 10.1186/s42162-020-00133-5.

[20] S. Kumar, P. Tiwari, and M. Zymbler, 'Internet of Things is a revolutionary approach for future technology enhancement: a review', *J. Big Data*, vol. 6, no. 1, p. 111, Dec. 2019, doi: 10.1186/s40537-019-0268-2.

[21] S. Inayatullah, 'Futures Studies: Theories and Methods'.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

[22] S. Urueña, 'Understanding "plausibility": A relational approach to the anticipatory heuristics of future scenarios', *Futures*, vol. 111, pp. 15–25, Aug. 2019, doi: 10.1016/j.futures.2019.05.002.
[23] P. D. Windschitl and E. U. Weber, 'The interpretation of "likely" depends on the context, but "70%" is 70%—right? The influence of associative processes on perceived certainty', *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 25, pp. 1514–1533, 1999, doi: 10.1037/0278-7393.25.6.1514.
[24] L. Z. Tiedens and S. Linton, 'Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing', *J. Pers. Soc. Psychol.*, vol. 81, pp. 973–988, 2001, doi: 10.1037/0022-3514.81.6.973.
[25] B. J. Weber and G. B. Chapman, 'The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect?', *Organ. Behav. Hum. Decis. Process.*, vol. 96, no. 2, pp. 104–118, Mar. 2005, doi: 10.1016/j.obhdp.2005.01.001.
[26] O. Renn, A. Klinke, and M. van Asselt, 'Coping with Complexity, Uncertainty and Ambiguity in Risk Governance: A Synthesis', *AMBIO*, vol. 40, no. 2, pp. 231–246, Mar. 2011, doi: 10.1007/s13280-010-0134-0.
[27] Oreskes, N., 'Merchants of Doubt', *Bloomsbury*. https://www.bloomsbury.com/us/merchants-of-doubt-9781608193943/ (accessed Jun. 02, 2023).
[28] J. Stilgoe, 'Stilgoe et al - Developing a framework for responsible innovation | Elsevier Enhanced Reader'. https://reader.elsevier.com/reader/sd/pii/S0048733313000930?token=50A038F72AFB67DFA0B9A30A83FB9348D2E57B69803BB129D3EE4B5EA576EA6E8FA07C4966A33ACC356AB04D3F0A6E86&originRegion=eu-west-1&originCreation=20211029142753 (accessed Oct. 29, 2021).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

Appendix I

	Ethical Framework for Task 5.1		
	ATE	**ATE plus (manuscript)**	**Ethical matrix**
Description	ATE employs three levels of ethical analysis, the technology, artifact and application level, which each contain various objects of analysis. Knowledge of them is acquired through forecasting, including the use of existing forecasting studies, expert panels and surveys, and self performed futures studies. Ethical analysis, finally, is performed at two initial stages, the identification and evaluation stage. At the identification stage, moral values and principles are operationalized and cross-referenced with technology descriptions resulting from the forecasting stage. The values and issues are derived from an ethical checklist as well as from the technology ethics literature and bottom-up analyses. At the evaluation stage, the potential importance of identified ethical issues is evaluated and these issues are elaborated. Evaluations may subsequently be used for improving technology development, for better governance of technology, or for other purposes. ATE may be applied to particular emerging information technologies	Enhancement of ATE framework: 1. Describe objects of interest, procedures, techniques, approaches, applications, use cases of interest, etc. (e.g., natural language processing; virtual reality; digital twins in training or health); 2. Investigate core philosophical notions and dilemmas that serve as conceptual scaffolding for the ethical issues (e.g., Is there an inherent preference for material reality over virtual reality?); 3. Identify values and principles (e.g., transparency, dignity) and return to step 2 for clarification if necessary; 4. Use narrative analysis to demarcate both transparent ethical considerations and morally opaque presuppositions in technological judgment concerning the values and principles identified in step 3 (e.g., "Be careful what you wish for", "The rich get richer, the poor get poorer"); 5. Ethnographically engage with critical stakeholders associated with technologies based on narratives instead of an addition to open-ended questions. 6. Formulate a set of operationalised design questions to be asked regarding the implementation of techniques (or applications and use cases) (e.g., does the XR system take stock of the potential changes of behaviour in its users? Who profits from the changes in behaviour and how are the changes incited?).	The ethical matrix starts with generally-accepted ethical principles and interprets these ethical concerns according to all affected parties' situations involved in agricultural biotechnology issues: The three principles, care for wellbeing, respect for dignity, and justice, are suggested as they seem to cover most of the ethical concerns in the field and correspond to major theoretical approaches in ethics. The aim of the ethical matrix is to help users identify ethical issues raised by the use food & agriculture technologies and to arrive at intellectually defensible decisions
Advantages - demonstrable benefits (eg, usability (claim thats its been done in industry), implementation/cost effectiveness of framework)	Advantages of ATE is that it is capable of detailed and comprehensive ethical analysis of emerging technologies	enhances the ATE framework to encompass the variety of human processes and material forms, functions, and applications that comprise the socio-technical systems in which these technologies are embedded.	Used in biotechnology in stem cell research https://www.eurostemcell.org/stem-cells-ethical-matrix and also used in food & agriculture
Disadvantages (no evidence of framework used)	Limitation - uncertainity of future	proposed framework, manuscript in-press	Overall matrix applies to bio/agriculture based technologies, does not take into account emerging technologies. however the ethical principals such as wellbeing (XR), dignity(NT) and justice (CE) can apply to 3 tech families.
Applicable to emerging tech?	yes	yes	not specifically for emerging tech. Ethical Matrix offers a structured way of working through equine welfare concerns in light of competing interests and outcomes,
Futures element?	The forecasting approach has as an advantage over the generic approach that it is able to consider more ethical issues, by including not only those that are generic to the technology but also those that are specific to projected future devices and their uses.	focus on 'plausible futures' through techethos scenarios WP3	no
Can be enhanced/refined? WRT tech families and techethos empirical findings?	yes	yes	no - but use of matrix can be adapted as it provides a structured way of working through ethical concerns in light of competing interests and outcomes,
specific to a tech family	all 3 tech families	all 3 tech families	no
more general to emerging tech (crossing boundaries & broad insights)	yes	yes	Overall matrix applies to bio/agriculture based technologies, does not take into account emerging technologies. however the ethical principals such as wellbeing (XR), dignity(NT) and justice (CE) can apply to 3 tech families.
How effective? can it be measured	NA	NA - manuscript in press	used in specific technologies in agriculture and bio

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

	Health technology assessment techniques (HTA)	ETICA approach	Ethical Impact assessment (EIA)
Description	A framework for systematic identification of ethical aspects of health technologies. The framework consists of twelve items with sub-questions, short explanations, and a concluding overall summary. The items are organized into four different themes: the effects of the intervention on health, its compatibility with ethical norms, structural factors with ethical implications, and long term ethical consequences of using the intervention.	How foresight and futures research can identify and address ethical issues in the field of Information Systems (IS). Starting from the premise that such IS are part of socio-technical systems, the interaction between technology and human actors raise ethical concerns. Early recognition of these concerns can address ethical issues and improve the use of the technology for a range of social and organisational goals. The aim of the ETICA approach is to provide comprehensive overviews of ethical issues for emerging technologies that are likely to play out in the medium-term future. The ETICA approach makes use of projections of the future which it derives from futures research. It aims to arrive at a foresight analysis, which is a forecasting analysis that considers multiple possible futures, out of which one is chosen as most desirable or important to consider. The ETICA approach relies on multiple futures methods and studies, under the assumption that while individual studies will contain biases and shortcomings, their aggregate use will tend to yield more reliable results.	The framework identifies key social values and ethical issues, provides some brief explanatory contextual information which is then followed by a set of questions aimed at the technology developer or policymaker to facilitate consideration of ethical issues, in consultation with stake holders. the framework includes a set of ethical tools and procedural practices which can be employed as part of the ethical impact assessment
Advantages - demonstrable benefits (eg, usability (claim thats its been done in industry), implementation/cost effectiveness of framework)	A framework for identifying ethical aspects of health technologies could be useful for NT but not CE or XR	useful for all emerging tech families (future studies)	The ethical tools will help the technology developer to get a better idea of how the technology is perceived ethically by stakeholders.
Disadvantages (no evidence of framework used)	may not be useful to CE and XR	ethical analysis undertaken in the ETICA project appear to refer to generic properties of the technologies that are studied, details at the 'artifact' and 'application' level are limited	does not account for emerging tech in the future, but investigates continuously the ethical implications of what is known about the technology under development.
Applicable to emerging tech?	more so for NT	yes	yes
Futures element?	nothing in literature to suggest	Future studies approach has some weaknesses, The main sources of the ETICA approach for locating ethical issues are government and political texts, scientific texts, and non-academic texts. Many of the nonacademic, scentific and government texts may not be based on scientific methods of futures research	not really - but needs to have a continuous exploration of ethical issues of what is known about the technology under devloepment
Can be enhanced/refined? WRT tech families and techethos empirical findings?	The Socratic approach is the result of a joint effort of experts in the field of ethics and HTA. Consensus is reached in the expert panel on an approach that is considered to be more clear, comprehensive, and applicable for addressing ethical issues in HTA.	yes	EIA uses stakeholder holder engagement as a key to identifying ethical issues which techethos has also used
specific to a tech family	NT	applicable to emerging tech	applicable to emerging tech
more general to emerging tech (crossing boundaries & broad insights)	socratic approach is useful and forms the basis of ADIM board meetings with experts - elements of this approach can be applied to all 3 tech families for example discussions with ADIM board/experts	yes can be applied to emerging tech	yes can be applied to emerging tech
How effective? can it be measured	NA	NA	EIA framework used in SATORI CWA

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

	SATORI CWA	Ethical technology assessment	Ethical Scenario method	Future studies
Description	CWA 17145-2 provides guidance when conducting an ethical impact assessment of a research or innovation project. From the initial idea of the project to the finalisation, ethics will have a key role in the project. Depending on the expected ethical impact of the innovation, measures are taken to prevent or mitigate negative impacts.This CEN Workshop Agreement (CWA) sets requirements and provides guidelines for ethics assessment of research and innovation. It consists of 2 parts, part 1 Ethics committee; Part 1 provides recommendations for the ethics committees on practices and procedures, part 2 Ethical impact assessment framework. This part provides a practical, policy-oriented guide for researchers and ethics assessors on the different stages of the ethical impact assessment (EIA) process	eTA can be conducted on the basis of a check-list that refers to nine crucial ethical aspects of technology: (1) Dissemination and use of information, (2) Control, influence and power, (3) Impact on social contact patterns, (4) Privacy, (5) Sustainability, (6) Human reproduction, (7) Gender, minorities and justice, (8) International relations, and (9) Impact on human values	Describes various possible (but not necessary likely futures, usually in narrative form. 1. Compose initial scenarios. Think what kind of ethical issues might be relevant related to your designs. 2. Evaluate your initial scenarios with different stakeholders. 3. Based on evaluations you can write first ethical guidelines or questions for your designs. 4. Build and realize your ideas and designs further on. Follow and modify your guidelines accordingly. 5. Involve different stakeholders to the evaluations of concept and prototypes. 6. Implement designs to the real world, take a few steps back to refine your design or start the developmental cycle again. It is preferable to test your design in real life settings as soon as possible.	explores alternative futures in order to assist people in choosing and creating their most desirable future.....
Advantages - demonstrable benefits (eg, usability (claim thats its been done in industry), implementation/cost effectiveness of framework)	The CWA aims to improve the quality of ethics assessment and harmonize ethics assessment practices. 1st international ethics assessment pre-standard for R&I - The SATORI CEN Workshop Agreement (2017) validity for 3 years - now seems inactive	Enables deliberation on desirability of new technologies. Focus on moral controversy and conflict	Supports imagination of future, Focus on technology in society Long term perspective, Applicable to emerging technologies	Also looks backwards as well as forwards to inform the futures
Disadvantages (no evidence of framework used)	does not seem to have a futures element	Applicable to technologies in development only. Morality conceived as a stable phenomenon	Morality conceived as a stable phenomenon	
Applicable to emerging tech	yes	yes, but dynamic interaction with technology and morality is neglected	yes	yes
Futures element?	does not seem to have a futures element	Limited time horizon	yes - but dynamic interaction with technology and morality is neglected. Scenarios used in techethos WP3	yes
Can be enhanced/refined? WRT tech families and techethos empirical findings?	European CWA is the result of extensive investigations and broad consultation and mutual learning processes with hundreds of stakeholders.	yes	yes	
specific to a tech family	applicable to emerging tech	applicable to emerging tech	applicable to emerging tech	NA
more general to emerging tech (crossing boundaries & broad insights)	yes	yes	yes	yes
How effective? can it be measured	Standard was valid for 3 years now seems inactive	NA	NA	used by planners (planning & design)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

TECHETHOS

FUTURE ○ TECHNOLOGY ○ ETHICS

Coordinated by

AIT AUSTRIAN INSTITUTE OF TECHNOLOGY

Partners

Airi ASSOCIAZIONE ITALIANA PER LA RICERCA INDUSTRIALE

allea All European Academies

cea

dmu.ac.uk DE MONTFORT UNIVERSITY LEICESTER

ecsite EUROPEAN NETWORK SCIENCE CENTRES & MUSEUMS

eurec

TRILATERAL RESEARCH

TU Delft

UNIVERSITY OF TWENTE.

Linked Third Parties

BUCHAREST SCIENCE FESTIVAL

CENTER FOR THE PROMOTION OF SCIENCE

iQ LANDIA

Consorcio Parque de las Ciencias

ScienceCenter NETZWERK

Vetenskap & Allmänhet VA – PUBLIC & SCIENCE