Human-Computer Interaction & Emerging Technologies, Vol. 195, 2025, 1–14
https://doi.org/10.54941/ahfe1006218

Critical Foresight of Human-Computer Interaction: A Review on Methods to Assess Ethical Risks and Side-Effects of Emerging Technologies

Sarah Diefenbach[1], Johannes Stoll[1], and Daniel Ullrich[2]

[1]LMU Munich, Department of Psychology, 80802 Munich, Germany
[2]LMU Munich, Department of Computer Science, 80337 Munich, Germany

ABSTRACT

Innovative technologies often come with unforeseen side effects or even detrimental consequences. This may range from impacts on individual wellbeing, effects on the level of society and social dynamics, environmental impacts, or unwanted effects on an organizational level. The responsible design of human-computer interaction demands methods to systematically foresee and assess such potential negative consequences of technological innovations for humans and society, and ideally use such insights to sensibly design and adjust a product concept and its features. In a comprehensive literature review of the methods landscape, we identified a total of 40 future-oriented methods designed or adaptable to elicit negative ethical or societal consequences of technological innovations. This paper describes different clusters of methods, classifies the methods along different criteria, and lists examples for each cluster. Based on a cross-analysis of the methods and reported best practices, we discuss recommendations for a fruitful combination of structural elements of existing methods and sketch ideas for new approaches of a light-weight way to reflect on potential harmful consequences of emerging technologies.

Keywords: Human-computer interaction, Emerging technologies, Critical foresight, Method review

INTRODUCTION

Newly emerging technologies often have unforeseen side-effects or even detrimental consequences. These may range from impacts on individual wellbeing, effects on the level of society and social dynamics, environmental impacts, or unwanted effects on an organizational level. To give some examples of technologies in everyday life: Users may trust AI systems and large language models such as chatGPT beyond its true capabilities, and rely on information which may sound true but is actually wrong (Wach et al., 2023). Technologies such as Tik-Tok promote short attention spans and a quick switch to the next gratification opportunity, in turn, users may lose the ability to accept boredom or concentrate for longer periods on one topic (Asif & Kazi, 2024). Also, the regular use of beauty filters in social media among young people is often accompanied by increasing dissatisfaction with their real selves and an enhanced desire for cosmetic surgery, so-called snapchat dysmorphia (Ramphul & Mjias, 2018). Meanwhile, TikTok also announced to block teenagers from beauty filters, reacting to increasing concerns about rising anxiety and falling self-esteem (Booth, 2024). AI programs such as eternos.life allows people to interact with a digital replica of their loved ones – experts argue, that we do not understand the full range of consequences of such tools, and reveal numerous ethical risks and privacy issues (International Financial Magazine, 2024; Schönwandt, 2025).

The responsible design of human-computer interaction demands methods to systematically foresee and assess such potential negative consequences of technological innovations for humans and society, and ideally use such insights to sensibly design and adjust a product concept and its features. As formulated by Taneja (2019), "'Minimum viable products' must be replaced by 'minimum virtuous products'—new offerings that test for the effect on stakeholders and build in guards against potential harms." Especially in the realm of newly emerging technologies, this is not a trivial task. As already discussed in the 80ies, the attempt to influence or control the development of new technologies is even more challenging than for already existing systems, due to lacking experience and justifications of control (the so-called Collingridge dilemma, Collingridge, 1982). In the early stages, when interventions are still easy to accomplish, there is not enough knowledge about the technology's future consequences to justify controlling it. Besides lacking knowledge, we cannot rule out that there may be cases in which potential negative side effects are ignored for commercial interests. However, by the time negative social impacts become obvious, the technology and its devices have already become ingrained within society, and controlling it becomes much more difficult.

Being faced with such developments, a question at hand is: Could we have known better? Is there a way to foresee such side effects and negative consequences of technological innovations for humans and society – and ideally use such insights to sensibly design and adjust a product concept and its features? What can we do to prevent such effects? What tools are there to make it better?

Looking at the literature, various attempts have been made in order to systematically foresee the effects of innovative concepts of human-computer interaction (HCI) and newly emerging technologies. In a comprehensive literature review of the methods landscape, we identified a total of 40 future-oriented methods designed or adaptable to elicit negative ethical or societal consequences of technological innovations. In order to reveal the potential and specific characteristics of the different methods, we identified overarching clusters of similar methodological approaches (e.g., Technology Assessment, Speculative and Critical Design, Scenario Planning) and further classified the methods along different criteria (e.g., data handling, application domain, target audience, stakeholder involvement). In the following sections, we present our review and central insights in more detail. We describe the search process, identified clusters and sample methods, and classification criteria. Based on a cross-analysis of the methods and best practices, we discuss recommendations for a fruitful combination of structural elements of existing methods and briefly sketch new approaches of a light-weight way to reflect on potential harmful consequences of emerging technologies.

REVIEW ON METHODS TO ASSESS ETHICAL RISKS AND SIDE-EFFECTS OF EMERGING TECHNOLOGIES

Search Process

Our review started with an initial "naïve" keyword search in different databases (e.g., ACM Digital Library, Web of Science, Springer Link, EBSCO, ScienceDirect, JSTOR) using combinations of relevant terms (e.g., technology assessment, predictive methods, future scenarios, responsible design, technology development, innovation, negative outcomes, ethical issues, unforeseen consequences). This initial scanning of the literature revealed related fields across multiple domains, like Technology Assessment (TA), Future Studies, Foresight Analysis, Speculative Design, Critical Design, and Responsible Research and Innovation. Subsequently, targeted research within those identified fields was conducted, using refined keyword strings with combinations (e.g., ethical foresight, anticipatory ethics). Backward and forward search was performed, to identify earlier and later work. This yielded a set of about 150 sources, which was reduced to a set of 40 methods after a more detailed screening with regards to our inclusion criteria (i.e., future-oriented plus focus on ethical and/or societal consequences of technological innovation).

Method Clusters

We organized the methods according to their commonalities and functions within foresight, assessment, or design processes, resulting in five broad clusters (n = number of methods subsumed in respective cluster).

(1) *Technology Assessment* (n = 9): Technology Assessment (TA) is a key methodology specifically designed for evaluating the social, cultural, political, and environmental impacts of new technological developments (Eijndhoven, 1997, for a review of TA methods see Tran & Daim, 2008). As formulated by Coates (2001), the aim of TA is to "better understand the consequences across society of the extension of the existing technology or the introduction of a new technology with emphasis on the effects that would normally be unplanned and unanticipated". Traditionally, TA is largely based on quantitative prediction-oriented knowledge, commonly using methods such as structural modelling. As argued earlier (Floridi & Strait, 2020), these quantitative approaches are not ideal to assess ethical impacts which are often complex and hardly transferable to measurable variables. However, there are also variants with a specific focus of stakeholder engagement or more interactive approaches such as Ethical Technology Assessment (Palm & Hansson, 2006), Ethical-Constructive Technology Assessment (Kiran et al., 2015), Designing-by-Debate (Ausloos et al., 2018), or Interactive Technology Assessment (Reuzel et al., 2001).

(2) *Speculative and Critical Design* (n = 19): In contrast to traditional, market-oriented design approaches, such methods emphasize critique and

speculation, to provoke thought and discussion about possible futures and societal developments (Dunne & Raby, 2013; Johannessen et al., 2019). Within this methods cluster, the sub cluster of *design fiction* methods involves creating narratives or prototypes of futures, exploring how technologies and artifacts might interact with society in unexpected ways such as value scenarios (Nathan et al., 2007), design fiction memos (Wong, 2021), timelines (Wong & Nguyen, 2021), pastiche scenarios (Blythe & Wright, 2006), or black mirror brainstorming (Jung et al., 2023). The sub cluster of *enactment* methods brings scenarios to life by creating experiential and tangible environments. Examples are the experiential futures ladder (Candy & Dunagan, 2017), immersive speculative enactments (Simeone et al., 2022), or speculative enactments (Elsden et al., 2017). Finally, the sub cluster of *unintended use* methods explores unexpected use cases and their societal impacts, may they be beneficial or detrimental. Examples are de-scription (Akrich, 1992), adversary personas (Miller & Williams, 2006), or security fictions (Merrill, 2020).

(3) *Scenario Planning* (n = 7): Methods of scenario planning explore possible futures so that decision-makers can identify potentially arising ethical risks and proactively develop counterstrategies. Methods of scenario planning show a great variety in the processes of designing scenarios and their ultimate characteristics (van Notten et al., 2003), whereby more recent approaches also rely on big data sources to improve the efficiency and the quality of scenario building. Examples are latent semantic analysis (Kwon et al., 2017), web mining (Kayser & Shala, 2020), or techno-ethical scenarios (Boenink et al., 2010).

(4) *Crowdsourced single prediction* methods (n = 2) focus on one specific outcome or prediction regarding future developments. For example, the approach of prediction markets makes use of the participants' "collective wisdom" and relies on the assumption that relevant information dispersed among individuals can be aggregated through market mechanisms (Williams et al., 2019). Participants buy and sell contracts based on their predictions about future events. Consequently, higher market prices indicate a greater perceived probability of the event occurring (Snowberg et al., 2008). Another example is the popular Delphi method, i.e., a structured, iterative communication technique that gathers insights from a panel usually composed of experts. In multiple rounds, experts provide judgments anonymously, after each round receiving a summary of the group's responses. Experts then revise their answers based on this feedback, until a more informed consensus is reached (Linstone & Turoff, 1975; Okoli & Pawlowski, 2004).

(5) *Combinatory Assessment* frameworks (n = 3) incorporate multiple foresight methods. Examples are the anticipatory technology ethics approach (Brey, 2012), analysing ethical impacts on the technology, artifact, and application level, or threatcasting (Vanatta & Johnson, 2018), which combines diverse inputs such as expert interviews, Delphi and others with a particular focus on cyber threats.

CLASSIFICATION CRITERIA AND METHOD MATRIX

In order to compare the different methods, reveal their specific potential, and identify the most suitable methods for a given application context, we classified the methods along different criteria, such as suggested application domain (e.g., technology, universal), target audience (e.g., policymakers, designers), type of stakeholder involvement (e.g., web data, market participants), bias mitigation mechanisms (e.g., diversity, negation, abstraction), or resource requirements. Table 1 lists a subset of the classification criteria and the categorizations of the methods. The method matrix with the full list of methods and classification criteria can be accessed online (https://osf.io/p2yqf/?view_only = 7238f986b10843faaae290f6fc5327e7).

Table 1: Method matrix (subset, for the full matrix and references see https://osf.io/p2yqf/?view_only=7238f986b10843faaae290f6fc5327e7). *Note: SCD = Speculative & Critical Design, qual = qualitative, quant = quantitative.*

Method	Cluster	Data	Target Audience	Time Horizon	Resource Requirements	Outcome Focus
Adversary Personas	Speculative & Critical Design	qual.	security professionals, developers, designers	short to medium-term	medium (detailed data on adversaries, expertise)	problem-related
Anachronistic Fiction	Speculative & Critical Design	qual.	designers, researchers	medium to long-term	medium (historical scenario, computing tools, design expertise)	problem-related
Anticipatory Action Learning	Scenario Planning	qual.	organizational leaders, educators, futurists	long-term	medium (workshop, documentation materials, knowledge)	problem-related
Anticipatory Technology Ethics	Combinatory Assessment	qual.	researchers, developers, policy makers, organizations	long-term	high (experts, time, funding)	problem-related
Black Mirror Brainstorming	Speculative & Critical Design	qual.	software developers, UX designers, design managers	medium to long-term	low (diverse group of participants, time, interview guide)	dystopian
Causal Layered Analysis	Scenario Planning	qual.	futurists, strategic planners, policymakers, researchers	no forecast	medium (documentation materials, knowledge)	problem-related
Delphi	Crowdsourced Predictions	qual., quant.	policymakers, academics	medium-term	high (experts, facilitator, questionnaires, time, software)	balanced
De-Scription	Speculative & Critical Design	qual.	designers, engineers, researchers	short to medium-term	low (documentation materials)	balanced
Design Fiction Memos	Speculative & Critical Design	qual.	designers, researchers, UX professionals	medium to long-term	medium (expertise, computing tools, field data, time)	problem-related

Continued

Table 1: Continued

Method	Cluster	Data	Target Audience	Time Horizon	Resource Requirements	Outcome Focus
Designing-by-Debate (DBD)	Technology Assessment	qual.	R&I consortia, organizations	medium-term (continuous)	high (ethical expertise, time, documentation materials)	problem-related
Dichotomy Mapping	Speculative & Critical Design	qual.	designers	short to medium-term	low (documentation materials)	balanced
DIODE	Technology Assessment	qual.	non-ethicists, technologists, policymakers, academics	medium-term (continuous)	high (flowcharts, documentation materials, time)	problem-related
Eth.-Const. Technology Assessment (ECTA)	Technology Assessment	qual.	technology developers, ethicists, policymakers	medium-term (continuous)	high (continuous micro-level involvement, time, documentation materials)	balanced
Ethical Impact Assessment	Technology Assessment	qual.	technologists, policymakers	medium-term (continuous)	high (time, delphi studies, ethical matrix)	problem-related
Ethical Matrix (Tool of EIA)	Technology Assessment	qual.	ethical decision-makers, policymakers	medium-term	low (ethical knowledge, documentation materials)	balanced
Ethical Risk Assessment	Technology Assessment	qual.	policymakers, academic organizations, nonprofit organizations	medium-term (continuous)	high (expertise, time, stakeholder knowledge)	problem-related
Ethical Technology Assessment	Technology Assessment	qual.	technology developers, ethicists, policymakers	medium-term (continuous)	high (time, expertise, documentation materials)	problem-related
ETICA	Combinatory Assessment	qual.	policymakers, industry professionals, researchers	long-term	high (bibliometric research, experts)	problem-related
Experiential Futures Ladder	Speculative & Critical Design	qual.	designers, futurists	medium to long-term	high (historical and cultural research, location, props and set design)	balanced
Futures Wheel	Scenario Planning	qual.	futurists, educators, corporate planners, policy advisors	long-term	low (pen, paper, creativity)	balanced
Method	Cluster	Data	Target Audience	Time Horizon	Resource Requirements	Outcome Focus
Immersive Speculative Enactments	Speculative & Critical Design	qual.	designers, researchers	medium to long-term	high (VR scenario, specialists, time)	balanced
Interactive Technology Assessment	Technology Assessment	qual.	policymakers, academics, ethicists	medium-term (continuous)	high (time, expertise, documentation materials)	problem-related
Inverted Behaviour Model	Speculative & Critical Design	qual.	designers	short to medium-term	low (documentation materials)	balanced
Latent Semantic Analysis	Scenario Planning	quant.	policymakers, tech companies	medium to long-term	medium (big data sources, data analysts, computing tools)	problem-related

Continued

Table 1: Continued

Method	Cluster	Data	Target Audience	Time Horizon	Resource Requirements	Outcome Focus
MANOA	Scenario Planning	qual.	policymakers, business people	long-term	medium (flipchart, mapping tools, workshop)	balanced
Motivation Matrix	Speculative & Critical Design)	qual.	designers	short to medium-term	low (documentation materials)	balanced
Pastiche Scenarios	Speculative & Critical Design	qual.	designers, researchers	medium to long-term	medium (design expertise, fictional and literary works, writing skills)	dystopian
Poetics of Design Fiction	Speculative & Critical Design	qual.	designers, researchers, students in HCI	medium to long-term	medium (time, writing skills, collaboration with literary practitioners)	dystopian
Prediction Markets	Crowdsourced Predictions	quant.	policymakers, academics	medium-term	high (market platform, computing tools, regulation)	balanced
Real-Time Technology Assessment	Technology Assessment	qual., quant.	policymakers, researchers, engineers, social scientists, public	medium-term (continuous)	high (expertise, time, data collection tools, documentation materials)	balanced
Science Fiction Prototyping	Speculative & Critical Design	qual.	engineers, designers, scenario planners	medium to long-term	high (time, literature analysis, experts, workshops)	balanced
Security Cards	Speculative & Critical Design	qual.	software developers, designers	short to medium-term	low (game materials, cards)	problem-related
Security Fictions	Speculative & Critical Design	qual.	software developers, designers	short to medium-term	low (game materials, physical space)	problem-related
Speculative Enactments	Speculative & Critical Design	qual.	researchers, designers in HCI	medium to long-term	high (design artifacts, users, actors, enactment spaces, time)	balanced
Tarot Cards of Tech	Speculative & Critical Design	qual.	designers	short to medium-term	low (game materials, cards)	balanced
Techno-Ethical Scenarios	Scenario Planning	qual.	policymakers, scholars	long-term	medium (time, literature analysis, experts, workshops)	problem-related
Threatcasting	Combinatory Assessment	qual.	military strategists, researchers	long-term	high (expert based Delphi, workshops)	dystopian
Timelines	Speculative & Critical Design	qual.	technology practitioners, research teams, educators	medium to long-term	low (white board, sticky notes, index cards, drawing materials)	balanced
Value Scenarios	Speculative & Critical Design	qual.	designers, policy makers	medium to long-term	medium (design social sciences expertise, stakeholder data)	problem-related
Web Mining	Scenario Planning	quant.	researchers, strategists	medium to long-term	medium (big data sources, data analysts, computing tools)	problem-related

As the table shows, most of the methods (37 of 40) focus on qualitative data, only five (e.g., prediction markets, latent semantic analysis) include quantitative data (partly combined with qualitative data). Regarding the application domain, 30 of the 40 methods are specifically referring to technology and 10 methods are described as universal methods. Referring to the target audience, i.e., the people who are imagined to apply the method, the main categories of mentions are designers (18), policymakers (15), and researchers (13). Even if not reported explicitly, all methods include some type of mechanism to prevent bias and one-sided perspectives – at least in that the methods suggests including data from participants with diverse backgrounds (e.g., experts, non-experts). In addition to diversity as bias mitigation mechanisms, the most frequent mechanisms are negation (29), i.e., deliberately considering opposing perspectives, and abstraction (28), i.e., strategies which encourage thinking beyond established mental models and recognizing higher-level concepts and principles. Regarding the resource requirements (e.g., materials, special expertise, time resources, relevant data, documentations), we broadly categorized the requirements into low (e.g., only pen & paper needed, low time requirements), medium (e.g., design and social sciences expertise, data on stakeholders required) and high (VR scenario, specialists, high time requirements). Here, ten methods were categorized as low, twelve as medium, and sixteen as high requirements methods. Finally, another interesting aspect is the outcome focus of the method and the types of consequences they aim to foresee: Seventeen methods take a balanced anticipatory perspective, considering positive as well as negative outcomes. Twenty-three methods take a problem-related angle towards future occurrences, and rather focus on negative or problematic outcomes. Four of these (i.e., black mirror brainstorming, pastiche scenarios, poetics of design fiction, threatcasting) deliberately take their problem-related perspective to extremes and explicitly encourage the exploration of dystopian or catastrophic scenarios.

CROSS-ANALYSIS AND IMPLICATIONS FOR FUTURE WORK

Based on a cross-analysis of the methods and reported best practices, we identified several elements that turned out as particularly valuable. Such elements or combinations of existing approaches might be considered when sketching new methods.

First, while it is beneficial to adopt the general openness of speculative and critical design approaches during processes of brainstorming of possible future artifacts or ethical consequences, it is also recommended to provide some systematic to make sure to consider effects for different individuals or groups on different levels. For example, this could pick up the STEEP (Social, Technological, Environmental, Economic, and Political) categories, following the example of MANOA (Schultz, 2015). Also, one could pick up the concept from Anticipatory Technology Ethics and ETICA, differentiating distinct technological levels such as technology, artifact, and application (Brey, 2012; Stahl & Flick, 2011). Distinguishing these levels, makes it easier to systematically analyse the wide range of ethical issues associated with emerging technologies, from their fundamental principles to their specific applications in society.

Second, a subsequent evaluation of relevance seems advisable, where ratings of plausibility/probability/severity provide guidance which effects are most important to consider in (re)designing the technology. This phase may incorporate principles of the Delphi method, such as anonymous, iterative polling among stakeholders, or aspects of Prediction Markets, offering incentivized anonymous betting on the most probable outcome. Another way to explore plausibility is historical analysis, evaluating how contemporary or future moral issues have been influencing societies during the past.

Third, there should be some visual presentation of potential future outcomes which is quickly comprehensible and also allows "to play" with the different elements. For example, the Futures Cone (Wong & Nguyen, 2021) is a timeline format that includes dimensions of time and probability, which could further be enhanced by color-coding events according to their perceived negativity. Lastly, there are various examples how a methods' structure and procedure may involve bias mitigation mechanisms, such as diversity by incorporating a diverse stakeholder base, negation through consideration of negative perspectives, and abstraction by enabling participants to think outside their usual mental frames.

Forth, placing attention to bias mitigation seems an important issue. While the here discussed methods already involve at least one bias mitigation mechanism, combining different mechanisms of diversity, negation, or abstraction seems advisable to further reduce the possibility of biases.

Fifth, when aiming at prevention, a deliberate focus on negative or dystopian consequences may be more suited than a neutral perspective, that in comparison might lack emotional impact. This could help in mitigating optimism bias, create a sense of urgency, and drive proactive measures, by illustrating potential worst-case scenarios if negative trends continue. Stakeholders may feel motivated to take immediate and decisive action to avert such outcomes (Claisse & Delvenne, 2015). Future methods could borrow from Pastiche Scenarios or Anachronistic Fiction to frame foresight exercises within societal circumstances of fictional works such as Nineteen Eighty-Four or actual historical episodes of oppression such as the European witch hunts. This way, technologies, artifacts, applications, and resulting ethical issues can be explored within extreme societal circumstances, while at the same time remaining within reasonable narrative boundaries to avoid over-speculation. An additional benefit of this approach is that it provides a safe space for participating stakeholders to express opposing opinions freely, negating potential concerns of being perceived as overly pessimistic.

Finally, as an overarching principle and to enhance the actual impact of a method, it seems important to consider the balance between rigor and resources. On the one hand, foreseeing all possible risks requires to cover a multitude of perspectives, at best based on reliable data and prediction models. On the other hand, it seems advisable to reduce the entry hurdle for companies, policy makers and all other stakeholders to apply the method. If a method requires too much resources or comes with other difficulties of acceptance, it won't be used. In this sense, a light-weight way of risk-assessment may be better than no risk-assessment at all.

CASSANDRA

As one possible approach, combining such insights into a new, light-weight method, we sketched the so-called Cassandra method (also see Diefenbach & Ullrich, 2024). The method borrows its name from the Greek myth of Cassandra. Because she was extraordinarily beautiful, the god Apollo gave her the gift of prophecy. However, after she constantly rejected his attempts to win her over, he cursed her so that no one would ever believe her prophecies. Cassandra could foresee future negative developments but her prophecies were not considered and the people ran into misfortune (like the tragedy of Troy). Probably, many of us know Cassandra moments from daily life – where we see bad developments coming but no one is interested in the risks we see. Other days, we might take the other role and refuse to listen what others or our "inner Cassandra" have to say. Our method aims to utilize the power of Cassandra and peoples' ability of role taking, to systematically look at the unnoticed, unformulated, or carelessly ignored side effects of technological innovations. But importantly in a lightweight and playful way, which balances out the over-criticism that many tech managers fear. Cassandra is not meant to be a test to be passed (or tricked) but it aims at really taking a different perspective – a true change in mindset which could lead to reject or modify an initial idea, and making a product better, more sustainable and hopefully even more accepted and successful. In line with the above insights, when sketching the Cassandra method, we decided for a deliberate focus on dystopian visions, structural elements to counteract biases such as groupthink or over-optimism, and some kind of visual representation of negative effects and their relevance, which allow to reflect on it in a playful light way. With the help of different tools and visualizations (e.g. card set, colour coded diagram), the method collects negative effects on different levels, assesses its severity and probability, and ideally draws implications how to redesign and improve a product concept. A more detailed description of the method and related tools can be accessed online (https://osf.io/p2yqf/?view_only=7238f986b10843faaae290f6fc5327e7).

CONCLUSION

As revealed in the present review, we are not the first to discuss the issue of negative side effects of emerging technologies. Given the multitude of methods, one could assume that the problem is basically well recognized and there should be tools to properly address it. Of course, another question is whether the problem is recognized by the right persons in the relevant positions, and to what degree commercial or other personal interests are in conflict with acknowledging the risks of the taken direction. Still, we should also acknowledge that per se, thinking about the risks related to an idea that one believes in, is a challenging task – from a psychological perspective and existing design culture.

On the one side, when thinking about new product concepts, progress and innovation critically depends on positive views and people who believe in great visions without too many reservations. Design research even names an optimistic mindset, "believing that designers can make the world a better place", as one of the fundamental principles of responsible design (Boehnert et al., 2010). On the other side, responsible design requires critical reflection and thinking about the many ways how a product may be misused or cause harm besides the envisioned positive use-case, and consider the unfolding of a design's impact in a system (e.g., Hernandez & Goni, 2020; Salamanca et al., 2019). Also, even with good intentions, different psychological mechanisms can add to an over-optimistic vision of how new technologies will change for the better and not seeing the risks. Psychologically seen, self-criticism or looking for negative aspects of an idea or project one sees as somehow belonging to oneself is contradictory to the natural tendency to protect one's self-worth, attitudes, and beliefs. Accordingly, information search and perception are often biased, only confirming what we already belief and neglecting contradictory information, the so-called confirmation bias (for an overview see Klayman, 1995). Fueled by a great vision and fascinated by technological opportunities, development teams may instead focus on what they wish to see, and mutually confirm their (biased) opinions, so-called groupthink (for an overview see Park, 1990). In consequence, it is not too surprising that experts in the field show over-optimism with regard to technology assessment and foresight (Tichy, 2004). All the more, we believe that methods such as the here presented can be a step to make this task a bit easier and form an important building block for the systematic inclusion of responsible design decisions in the field of HCI and emerging technologies.

REFERENCES

Akrich, M. (1992). The De-scription of Technical Objects. Shaping Technology-Building Society: Studies in Sociotechnical Change, 205–224. https://hal.archives-ouvertes.fr/halshs-00081744

Asif, M & Kazi, A. (2024). Examining the Influence of Short Videos on Attention Span and its Relationship with Academic Performance. International Journal of Science and Research, 13, 4, 1877–1883. https://dx.doi.org/10.21275/SR24428105200

Ausloos, J., Heyman, R., Bertels, N., Pierson, J. & Valcke, P. (2018). Designing-by-Debate: A Blueprint for Responsible Data-Driven Research & Innovation. In SpringerBriefs in research and innovation governance (S. 47–63). https://doi.org/10.1007/978-3-319-73207-7_8

Blythe, M., & Wright, P. C. (2006). Pastiche scenarios: Fiction as a resource for user centred design. Interact. Comput., 18, 1139–1164. https://doi.org/10.1016/j.intcom.2006.02.001

Boehnert, J., Sinclair, M. & Dewberry, E. (2022). Sustainable and Responsible Design Education: Tensions in Transitions. Sustainability, 14(11), 6397. https://doi.org/10.3390/su14116397

Boenink, M., Swierstra, T., & Stemerding, D. (2010). Anticipating the Interaction between Technology and Morality: A Scenario Study of Experimenting with Humans in Bionanotechnology. Studies in Ethics, Law, and Technology, 4, 1–38. https://doi.org/10.2202/1941-6008.1098

Booth, R. (2024). TikTok to block teenagers from beauty filters over mental health concerns. https://www.theguardian.com/technology/2024/nov/26/tiktok-to-block-teenagers-from-beauty-filters-over-mental-health-concerns
Brey, P. A. E. (2012). Anticipatory Ethics for Emerging Technologies. NanoEthics, 6(1), 1–13. https://doi.org/10.1007/s11569-012-0141-7
Candy, S., & Dunagan, J. (2017). Designing an experiential scenario: The People Who Vanished. Futures, 86, 136–153. https://doi.org/10.1016/J. FUTURES.2016.05.006
Claisse, F., & Delvenne, P. (2015). Building on anticipation: Dystopia as empowerment. Current Sociology, 63, 155–169. https://doi.org/10.1177/0011392114556579
Coates, J. F. (2001). A 21st century agenda for technology assessment. Technological Forecasting and Social Change, 67, 303–308.
Collingridge, D. (1982). The Social Control of Technology. Journal of the Operational Research Society, 33(8), 771. https://doi.org/10.2307/2634327
Diefenbach, S. & Ullrich, D. (2024). The Cassandra Method – Dystopian Visions as a Basis For Responsible Design. Engineering Proceedings, 68(1), 2. doi: 10.3390/engproc2024068002
Dunne, A. & Raby, F. (2013). Speculative everything: Design, fiction, and social dreaming. Choice Reviews Online, 51(10), 51–5390. https://doi.org/10.5860/choice.51-5390
Eijndhoven, J. V. (1997). Technology assessment: Product or process? Technological Forecasting and Social Change, 54, 269–286. https://doi.org/10.1016/S0040-1625(96)00210-7
Elsden, C., Chatting, D., Durrant, A. C., Garbett, A., Nissen, B., Vines, J., & Kirk, D. S. (2017). On Speculative Enactments. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 5386–5399. https://doi.org/10.1145/3025453.3025503
Floridi, L., & Strait, A. (2020). Ethical Foresight Analysis: What it is and Why it is Needed? Minds and Machines, 30(1), 77–97. https://doi.org/10.1007/s11023-020-09521-y
Hernandez, R. J. & Goni, J. (2020). Responsible Design for Sustainable Innovation: Towards an Extended Design Process. Processes, 8(12), 1574. https://doi.org/10.3390/pr8121574
International Finance Magazine (2024). AI resurrections: Reviving the dead or exploiting grief? https://internationalfinance.com/magazine/technology-magazine/ai-resurrections-reviving-the-dead-or-exploiting-grief/
Johannessen, L. K., Keitsch, M., & Pettersen, I. N. (2019). Speculative and Critical Design—Features, Methods, and Practices. Proceedings of the Design Society: International Conference on Engineering Design. https://doi.org/10.1017/DSI.2019.168
Jung, K. S., Makam, N., Jung, S. D., Kim, S. & Kim, S. (2023). Black Mirror: A novel application of speculative design to facilitate context-aware design thinking. In Proceedings of The Design Society, 3. Cambridge University Press, 787–796. https://doi.org/10.1017/pds.2023.79
Kayser, V. & Shala, E. (2020). Scenario development using web mining for outlining technology futures. Technological Forecasting and Social Change, 156, 120086. https://doi.org/10.1016/j.techfore.2020.120086
Kiran, A. H., Oudshoorn, N. & Verbeek, P. (2015). Beyond checklists: toward an ethical-constructive technology assessment. Journal of Responsible Innovation, 2(1), 5–19. https://doi.org/10.1080/23299460.2014.992769

Klayman, J. (1995). Varieties of confirmation bias. Psychology of Learning and Motivation 32, 385–418. https://doi.org/10.1016/S0079-7421(08)60315-1
Kwon, H., Kim, J. & Park, Y. (2017). Applying LSA text mining technique in envisioning social impacts of emerging technologies: The case of drone technology. Technovation, 60–61, 15–28. https://doi.org/10.1016/j.technovation.2017.01.001
Linstone, H., & Turoff, M. (1975). The Delphi Method: Techniques and Applications. In Technometrics (Bd. 18). https://doi.org/10.2307/3150755
Merrill, N. (2020). Security Fictions: Bridging Speculative Design and Computer Security. DIS '20: Proceedings of the 2020 ACM Designing Interactive Systems Conference, 1727–1735. https://doi.org/10.1145/3357236.3395451
Miller, R., & Williams, L. A. (2006). Personas: moving beyond role-based requirements engineering. https://techrep.csc.ncsu.edu/2006/TR-2006-24.pdf
Nathan, L. P., Klasnja, P. V., & Friedman, B. (2007). Value scenarios. CHI '07 Extended Abstracts on Human Factors in Computing Systems. 2585–2590. https://doi.org/10.1145/1240866.1241046
Okoli, C. & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. Information & Management, 42(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002
Palm, E. & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). Technological Forecasting and Social Change, 73(5), 543–558. https://doi.org/10.1016/j.techfore.2005.06.002
Park, W.-W. (1990). A review of research on groupthink. Journal of behavioral decision making, 3(4), 229–245. https://doi.org/10.1002/bdm.3960030402
Ramphul, K. & Mejias, S. (2018). Is "Snapchat Dysmorphia" a real issue? Cureus 10, 3 (2018). https://doi.org/10.7759/cureus.2263
Reuzel, R., Van der Wilt, G., Have, H. T. & De Vries Robb, P. (2001). Interactive Technology Assessment and Wide Reflective Equilibrium. The Journal of Medicine and Philosophy a Forum for Bioethics and Philosophy of Medicine, 26(3), 245–261. https://doi.org/10.1076/jmep.26.3.245.3015
Schönwandt, F. (2025). Mein Mann lebt als KI weiter - Lieben und Sterben mit Künstlicher Intelligenz. Documentary film. https://www.ardmediathek.de/video/Y3JpZDovL21kci5kZS9zZW5kdW5nLzI4MTA2MC8yMDI1MDEyMDIzMDUvcmVwb3J0YWdlLWRva3UtaW0tZXJzdGVuLTM5NA
Schultz, W. (2015). Manoa: The future is not binary. APF Compass, 4, 4–8.
Simeone, A. L., Cools, R., Depuydt, S., Gomes, J. M., Goris, P., Grocott, J., Esteves, A. & Gerling, K. (2022). Immersive Speculative Enactments: Bringing Future Scenarios and Technology to Life Using Virtual Reality. CHI Conference on Human Factors in Computing Systems. 17, 1–20. https://doi.org/10.1145/3491102.3517492
Snowberg, E., Wolfers, J. & Zitzewitz, E. (2008). Prediction Markets: From Politics to Business (and Back). In Elsevier eBooks (S. 385–402). https://doi.org/10.1016/b978-044450744-0.50021-4
Stahl, B. C. & Flick, C. (2011). ETICA Workshop on Computer Ethics: Exploring Normative Issues. In IFIP advances in information and communication technology (S. 64–77). https://doi.org/10.1007/978-3-642-20769-3_6
Taneja, H. (2019). The Era of "Move Fast and Break things" is Over. https://hbr.org/2019/01/the-era-of-move-fast-and-break-things-is-over
Tichy, G. (2004). The over-optimism among experts in assessment and foresight. Technological Forecasting and Social Change 71, 341–363.

Tran, T. A. & Daim, T. (2008). A taxonomic review of methods and tools applied in technology assessment. Technological Forecasting And Social Change, 75(9), 1396–1405. https://doi.org/10.1016/j.techfore.2008.04.004

Vanatta, N. & Johnson, B. D. (2018). Threatcasting: A framework and process to model future operating environments. The Journal of Defense Modeling And Simulation Applications Methodology Technology, 16(1), 79–88. https://doi.org/10.1177/1548512918806385

Van Notten, P. W., Rotmans, J., Van Asselt, M. B. & Rothman, D. S. (2003). An updated scenario typology. Futures, 35(5), 423–443. https://doi.org/10.1016/s0016-3287(02)00090-3

Wach, K., Duong, C. D., Ejdys, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G.,... & Ziemba, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. Entrepreneurial Business and Economics Review, 11(2), 7–30.

Williams, L. V., Sung, M. & Johnson, J. (2019). Prediction markets: Theory, evidence and applications. International Journal of Forecasting, 35(1), 266–270. https://doi.org/10.1016/j.ijforecast.2018.11.001

Wong, R. Y. (2021). Using Design Fiction Memos to Analyze UX Professionals' Values Work Practices: A Case Study Bridging Ethnographic and Design Futuring Methods. CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 93, 1–18. https://doi.org/10.1145/3411764.3445709

Wong, R. Y., & Nguyen, T. (2021). Timelines: A World-Building Activity for Values Advocacy. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 616, 1–15. https://doi.org/10.1145/3411764.3445447