ENGAGING STUDENTS AND PROFESSIONALS IN ETHICAL REFLECTIONS ON NEW AND EMERGING INFORMATION AND COMMUNICATION TECHNOLOGIES

T. Børsen, boersen@plan.aau.dk
Aalborg University
Copenhagen, Denmark

P. Karadechev, petko@plan.aau.dk
Aalborg University
Copenhagen, Denmark

J. Contreras, jicc@ta4u.dk
Techno-Anthropology4U I/S
Copenhagen, Denmark

***Conference Key Areas**: Interdisciplinary engineering education, linking different disciplines both inside and outside engineering, linking with society; Sustainability and ethics, embedded and dedicated approaches.*
***Keywords**: Participatory Ethical Technology Assessment; Techno-Anthropology; Automated Decision-Making; Ethics in Engineering Education.*

ABSTRACT

Techno-Anthropology is an interdisciplinary Master's program offered by The Technical Faculty of IT and Design at Aalborg University that addresses human – technology interfaces and, hence, combines disciplinary elements of engineering and the humanities. In this concept paper, we present and analyze a three-day field trip to the annual TechFestival in Copenhagen that included a workshop for students, educators and IT professionals. The workshop was integrated into the Master's program of Techno-Anthropology as a means to teach the students technology ethics by direct interaction with external stakeholders. Hence, the purpose of the workshop was to design, implement and evaluate a participatory, ethical technology assessment teaching format, and consisted of a presentation, practical engagement activities and preparation of individual future road maps. This paper will focus on the context, content and results of the workshop activities. These were centered around a suggested model for collective ethical judgment, which includes the following stages:

1. Identifying ethical issues
2. Linking them to ethical values
3. Identifying ethical dilemmas and placing them in relevant human and non-human networks
4. Engaging participants in value-sensitive discussions aimed at addressing the identified ethical dilemmas provided by the workshop participants.

The paper will outline exemplary participatory design practices that will be of interest to educators and other professionals who work with new and emerging ICTs, and are interested in promoting ethical reflections. The presented analysis identifies the ethical issues, values, dilemmas, networks and solutions suggested by participants, and links them to different aspects of the socio-technical understanding of technology central to Techno-Anthropology. Future research will focus on how to translate the format of the presented experience into other contexts and technology domains that want to enhance their ethical reflection capabilities.

1 INTRODUCTION

1.1 What is Techno-Anthropology?

Techno-Anthropology is an interdisciplinary research area and a Master's program offered by the Technical Faculty of IT and Design at Aalborg University. The research and study program addresses human-technology interfaces. How do humans and technologies interact? This approach is labelled as the socio-technical understanding of technology. It focuses on technical artefacts and human dimensions such as culture and social relations. Techno-Anthropology combines disciplinary elements of both engineering and the humanities. The idea underpinning Techno-Anthropology is often visually illustrated as a triangle with technology users, technology artefacts and technology experts in its three corners (Figure 1).

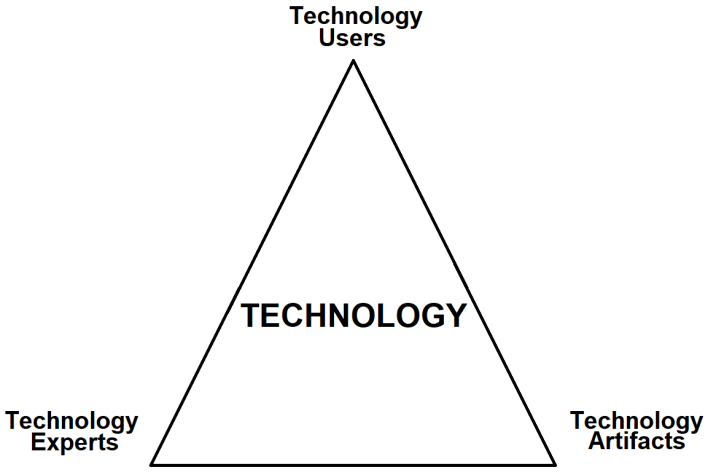

Figure 1. The Techno-Anthropological Field (based on [1]).

The sides of the triangle identify sub-disciplines of Techno-Anthropology: Moving from the artefacts to the users, we see "Technology assessment" emerging [2]. Technology assessment identifies and assesses intentional and unintentional effects of technologies on users. If we go in the opposite direction, from the users towards the artefacts, we find "Anthropology-driven design" which is an endeavour that maps users' needs and translates them into design specifications [3]. We can also move from the users towards the technical experts. In this interface we find "Participatory design" [4] and "Action research" [5]. When we go from the technical experts towards the users, subdisciplines like "Transdisciplinarity" [6] and "Studies of expertise and experience" emerge [7]. When we go from Technical experts to the Technical artefact we are in the conventional domain of engineering such as "Software engineering". However, at Techno-Anthropology we also embrace "Responsible research and innovation" [8] and "Value-sensitive design" [9] as subdisciplines. Moving from the technical artefact to the technical expert, we see how technical artefacts form and influence research and entrepreneurship. An example of a Techno-Anthropological research project in this area is "How can professionals incorporate AI to make professional estimates". Another Techno-Anthropological research project in this area investigates how university faculty members use digital technologies in their teaching. We have not been able to identify sub-disciplines devoted to the study of how technology influences technology experts.

In this paper we work within one sub-discipline of Techno-Anthropology – between the technology experts and technology artefacts – as we investigate how to promote ethical reflections among technology Master's students, experts and educators in the domains of AI and machine learning. This is a topic relevant to include in engineering education [10]. Engineering ethics teaching objects usually takes the form of lectures, written assignments, case-studies and role plays [11]. Teaching ethics by direct engement with external stakeholders are rarely found and should be promoted [12]. In this paper

we present a technology ethics teaching object that centers around an equal meeting between Master's students of Techno-Anthropology and IT professionals.

1.2 Techfestival

In 2019, we offered a techno-anthropological workshop[1] for Master's students and private and public sector IT professionals at the annual Techfestival[2] in Copenhagen. The Techfestival attracts diverse audiences from sectors such as software engineering, public sector organizations, consultancy, etc., which are actively looking for technical and socio-technical solutions to various problems. This context offered us an interdisciplinary group of participants to engage with. The workshop was integrated into the Master's program of Techno-Anthropology as a means to teach technology ethics by direct interaction with external stakeholders. Hence, the purpose of the workshop was to design, implement and evaluate a participatory ethical technology assessment format. The format consists of a presentation, practical engagement activities and individual future road maps for the workshop participants. This paper will focus on the context, methodology and results of the workshop activities.

2 METHODOLOGY & RESEARCH DESIGN

In this short paper, we present and analyze a three-day field trip to the annual Techfestival in Copenhagen that included a workshop for students and IT professionals, interested in the ethical effects of A.I. The paper will outline exemplary participatory design practices that will be of interest to educators and other professionals who work with new and emerging ICTs, and are interested in promoting ethical reflections.

2.1 Quick and Proper Participatory Ethical Technology Assessment

In this paper we present the Quick and Proper Participatory Ethical Technology Assessment (PETA) model/format[3] that is developed by the authors. PETA promotes collective ethical judgement of an existing or emerging technology, and is suitable for use in ethics teaching of Master's students and professionals working with new and emerging technologies. The idea is to set up an extended peer-community, e.g. by organising a workshop where different stakeholders are invited. The idea is to collect differentiated experiences with the effects of the assessed technology. The model includes the following steps:

2.1.1 Identifying Ethical Issues

The first step in PETA is to identify possible effects of the technology under assessment. This is done by asking the assessed technology's extended peer-community[4] to brainstorm and discuss i) the intended good consequences of the technology under assessment, ii) how it can be misused, iii) what unintended consequences the technology can have, and iv) how it might affect human culture and societies in the long run.

2.1.2 Linking Ethical Issues to Ethical Values

An ethical value is understood as a normative criterion against which one can compare the wider consequences and circumstances of the use of a given technology. During the 2018 Techfestival an interdisciplinary group of festival participants produced the Copenhagen Catalogue[5], which is a list of 150 ethical "*principles for a new direction in technology*". The catalogue can guide the development and application of hardware and software. The values were formulated through participatory co-design.

---

[1] Invitation to the workshop: https://techfestival.co/event/effects-ai-decision-making/
[2] https://techfestival.co
[3] The model side of PETA is focused on defining concepts (technical, socio-technical, ethical, methodological, etc.). The format side of PETA deals with how these concepts are applied in a real-world setting.
[4] Extended peer-communities covers both members of a technology's traditional peer-community (e.g. software engineers, UX designers, etc.), and members of non-obvious peer-communities, who are facing ethical issues posed by the assessed technology. Our workshop facilitates brainstorming in discussion for communities that may not traditionally interact with each other.
[5] https://www.copenhagencatalog.org/

Hence, the second step of PETA is to liaise the ethical issues identified during step one to relevant ethical values. We provided the participants with a card deck including all 150 values, divided the participants in groups and asked them to identify the values they found most relevant for A.I. and machine learning in their own and their fellow group member's professional fields.

2.1.3 Identifying Ethical Dilemmas

PETA's third step operates with a distinction between an unethical situation and an ethical dilemma. We have an unethical situation if the technology under assessment violates ethical standards without being justified by any other ethical value. An ethical dilemma refers to a situation where ethical values are in collision.

The purpose of the third step of the PETA model is to decide whether a technology generates an unethical situation and hence is ethically wrong and should not be developed or used, is not associated with ethical problems and can be applied without ethical concerns, or creates an ethical dilemma, where the situation is ambivalent. Latter cases call for ethical judgment, decision-making and value-sensitive design.

At Techno-Anthropology the first, second and third steps have been applied to a number of cases: risk reducing surgery [13], an ethical pig stable [14], the use of DDT [15], psychotropics and other enhancing technologies [16], and automated decision-making [17]. These texts can be read in preparation to attending in a PETA workshop.

2.1.4 Engaging Participants in Value-Sensitive Discussions

Value-sensitive design is an ethical approach to overcome ethical dilemmas by design. The idea is to maintain already inscribed ethical values while changing the design so that new ethical values, those who are violated in the assessed socio-technical configuration, are inscribed by changing the design.

The socio-technical configuration needs to be changed in order to address the ethical dilemma by involving interdisciplinary competencies. This means to 1) identify actors and actants that can influence and drive a change in the technology under assessment, and 2) arrange a shared theoretical and practical ground for actionable discussions to occur. This includes identifying interests of human actors and intentionality of nonhuman actants, which should be recognizable and actionable by either or both sides [18].

Addressing ethical dilemmas using value-sensitive design is not a value-neutral act; as such the fourth PETA step demands high levels of reflexivity and awareness regarding participation (who/what is an actor/actant), issue visibility (what can each side perceive as an ethical issue), and commitment (specific steps that relevant human and non-human networks can be positioned around to address the identified ethical dilemmas).

The learning objectives associated with the PETA teaching approach is both to generate knowledge of a technology's ethical issues, link them with ethical criteria and to make ethical judgments in a transdisciplinary context of different and complex perspectives.

2.2 The Workshop

In September 2019, the authors organized an interactive workshop titled "The Effects of A.I. and Automated Decision-Making" (ADM) at Copenhagen's Techfestival. We used the term ADM instead of Artificial Intelligence or Machine Learning to include automated, reflexive, unconscious decision-making processes in humans as well as in machines, where A.I. and M.L. apply. The workshop enacted the PETA model by engaging an audience of students and technical experts from different industries. There were four main goals that the workshop addressed: 1) creating a common ground for concepts and

definitions; 2) creating an engaging technological narrative to serve as a boundary object facilitating a shared understanding of ADM and the ethical assessment methodology; 3) creating a participatory environment; 4) setting up an ethics teaching object based on interaction between students and external stakeholders.

The workshop was attended by ~90-100 participants, which were arranged in small groups of ~6-7 people. The groups produced 3 deliverables as they went through 3 rounds of exercises. The deliverables include 2 lists (intentional and unintentional effects of ADM and ethical values relevant to the participant's own line of work) and 1 personal, actionable roadmap to solve the identified issues.

After establishing shared concepts and definitions, **the first round** aimed to identify, reflect and share thoughts about selecting relevant ethical values / principles from the 2018 Techfestival's Copenhagen Catalogue, and how they are relevant for each group member. The **second round** focused on the ambivalence of ADM, where participants became a "human algorithm" and were instructed to sort the 150 values / principles in the Copenhagen Catalogue by very specific criteria (size, shape, etc.). Ambivalence arose when discussing why each group had a different answer to the exact same set of criteria. The **third round** inspired participants to be professionally and socially engaged via ethical discussions and ethical dilemmas in the design of ADM. They were to translate their group reflections to actionable plans for their specific technical, professional or business area. All participants gathered in plenary with concrete examples from their professional experiences. They were identifying new values that could support the detection of ethical tensions in their own field with the help of fellow participants. Two volunteers debated their individual approaches in front of the entire audience.

Figure 2. Personal roadmap to support individual ethical reflections on automated decision-making

As a final task, participants used a roadmap instrument (see Figure 2, above) to facilitate prospective thinking in the future, and as a tool to share the workshop experience with their colleagues at work.

Techfestival 2019 marks the identification of ethical values (among those 150 included in the Copenhagen Catalogue) during the workshop as a starting point (red point). Participants were encouraged to repeat the exercise identifying ethical values relevant to additional specific cases and dilemmas in at least two additional milestones (yellow & blue points). A follow-up meeting was proposed for Techfestival 2020 where participants can share the experience and results of implementing the workshop's exercises.

3 RESULTS

3.1 Ethical Assessment

The presented ethical analysis identifies only the ethical issues suggested by the workshop participants. How the ethical issues liaise with ethical values, networks and solutions was not agreed upon. The results of the shared ethical assessment from the workshop are four ethical issues regarding automated decision-making, and are summarized here as follows:

3.1.1 Black Boxed Information Selection

A central point in the ADM domain is that decisions made by computers are rarely based on shared (socio-)technical principles, and similar technologies can have wildly different results. Practically, this makes it impossible for humans to know *how* decisions are made in an ADM context, thus resulting in a classical black-boxed situation.

3.1.2 Wicked Algorithms

If the input is flawed, so is the output. This goes under the heading "junk in, junk out". It is also possible to manipulate the data input to prevent an A.I. system from serving a beneficial purpose. A traditional understanding of algorithms as impartial tools is complicated into *wicked* algorithms that are malleable in unforeseen ways.

3.1.3 Unequal Distribution of the Fruits of Digitalization

A digital transformation will affect different target groups in different ways. Digitalization can both benefit workers by improving their job conditions, but also result in loss of jobs all together. Stakeholders with knowledge and power advantages are unlikely or unable to share the benefits of digitalization where they are most needed.

3.1.4 Political Consequences of the Internet of Things

The more information is generated by internet-connected IoT devices, the more politicians can promote decision-making to an "input-output formulae" as is seen in China, where some public behaviour is regulated through registering citizens' digital traces and making them a subject for peer assessment. Such wide-ranging effects of internet-connected devices, many of which operate under ADM processes, have deep political consequences, which require attention.

3.2 Processual Lessons from the Workshop

We present four main lessons for organizers of workshops with focus on ethical assessment of ADM processes.

3.2.1 Consider Participants Individually and Engage Personal Experience

Relevant stakeholders and interested parties may instinctively be approached as different types of groups: different professions, disciplines, sectors, minorities, and so on. It is critical to engage individual perspectives from different points of view, which hold ADM expertise in a broad sense (including user perspectives). For example, a workshop participant who uses machine learning models to evaluate university student retention and participation[6] was engaged by another participant who provided consulting services to African governments as part of his position in the Danish Foreign Ministry. These different perspectives, engaged individually, were able to better explain and reflect upon their role in the functioning of their respective institutions, as well as to receive relevant professional feedback they would not have looked for otherwise.

[6] This participant wrote a blog post on her experiences at the workshop: https://www.version2.dk/blog/forudsigelse-eller-prognose-1089370

3.2.2 *Create Shared Experiences and Understandings of what Algorithms are Doing*

One way to visualize nuanced ethical issues with regards to automated decision-making is to transform workshop participants into metaphorical logic gates that enact criteria set by a supposedly objective and clear algorithms. When asked to perform "objective" criteria on a set of cards, our workshop participants immediately identified the flawed aspect of an "objective" algorithm, regardless of their professional or educational background. This simple exercise (see Figure 3 below) illustrates how each individual can highlight ethically problematic aspects of ADM quickly by comparing their professional experience to other expertise.

Figure 3. Workshop participants order Techfestival Principles by a simple and "objective" criteria (e.g. font size, etc.) in the *Becoming the algorithm* exercise

3.2.3 *Recognition of Ethical Issues in Personal Work Cases*

It was possible for the workshop participants to reach agreement on the identified ethical issues. No consensus on which ethical values were linked to which ethical issues was reached, however. Neither did the group reach unanimity with regard to solutions and networks. We do not know if consensus can be achieved in these areas. We need more research to draw such conclusions.

3.2.4 *Learning by Opening Up*

One can explain participation as a balanced interaction between power and learning. One of the assumptions for participation is that those who gain power relying on technical expertise are required to open up for learning. Users' contributions were sharing knowledge with technical experts, and at the same time they gained power. We as techno-anthropologists assumed the role of facilitators in the interplay of power exchange. The Copenhagen Catalogue was useful to focus on ethical discussions and made participants relate to them. Interdisciplinary participation became visible when participants exposed a myriad of interpretations of the same values / principles.

Figure 4. Example of the Techfestival Principles

4 CONCLUSION

As techno-anthropologists we use methodologically and theoretically flexible ways to address and analyse the entangled relationships between humans and technologies. What separates techno-anthropological approaches from other approaches is a formal structure to interdisciplinary mediation. Blurring boundaries between disciplines diminish the impact of participation and co-creation. Techno-Anthropology acts as translator and facilitator among subjects; thus, uncertainty and confusion in the borders can decrease. In this regard, and as stated in the techno-anthropological triangle (Figure 1), the workshop was a relevant avenue to assess ADM as a technical artifact. The workshop systematically illustrated the tension between users and technical experts with regards to ethical issues and values related to technologies.

It is noteworthy that the exercise allowed attendees to alternate between roles. The participants were acting as technology experts when identifying what elements are relevant for strengthening their understanding of ADM technologies while listening to users (Learning). They were anticipating and developing, as technology users, concrete actions, and new values required to build confidence and minimise undesirable effects of the technology, thus raising empowerment (Power). Beyond the ethical reflections, another consequence of the workshop demonstrated the different aspects such as networks, values, ethical issues, and possible solutions to understanding the socio-technical configuration of technology central to Techno-Anthropology. Finally, having presented the PETA model and the road map to help attendees to implement the workshop in their fields, it was our contribution to pinpoint the weight of responsible research and innovation as a subdiscipline that requires more attention in novel technologies like ADM.

The attending students and other participants learned both about AMD technologies, their ethical issues (black boxed information selection, wicked alghoritms, unequal distribution of risks and benefits of digitalisation and political issues of the internet of things), as well as of the advantages of involving different stakeholders in ethical technology assessment, exchanging power and learning within an extended peer-community, creating shared understanding and translating all this into personal action. Hence, the presented PETA ethics teaching object is relevant to include in engineering ethics education because it aligns well with what Jeroen van den Hoven calls for when he states that:

> Multidisciplinary and interdisciplinary collaboration is very important. This is a much needed development in academia. The solution to the big and urgent problems in the world will not be found in one discipline, in one journal or in one book. Moreover, adequate solutions will always be systems' solutions, and they will most likely deal with technology and human behavior with values and norms. [19]

REFERENCES

[1] Børsen, T (2013) Identifying Interdisciplinary Core Competencies in Techno-Anthropology: Interactional expertise, social responsibility competence, and skills in anthropology-driven design, What is Techno-Anthropology? Børsen, T, Botin, L (editors) Aalborg University Press, Aalborg, p. 35-66.

[2] Botin, L, Børsen, T (editors) (Forthcoming) Techno-Anthropological Contributions to Technology Assessment. Aalborg University Press, Aalborg.

[3] Christensen, LR (2013). Techno-Anthropology for Design, What is Techno-Anthropology? Børsen, T, Botin, L (editors) Aalborg: Aalborg University Press; p. 385-404.

[4] Kanstrup, AM, Bertelsen, P. Participatory Reflections: Power and Learning in User Participation, What is Techno-Anthropology? Børsen, T, Botin, L (editors) Aalborg University Press, Aalborg; p. 405-430.

[5] Thorsen N, Børsen T (2018) Aktionsforskning på kandidatuddannelsen i Tekno-Antropologi, Den Ufærdige Fremtid, Bilfeldt, A, Jørgensen MS, Andersen, J, Perry, KA (editors). Aalborg University Press, Aalborg; p. 179-205.

[6] Brown, VA, Harris, J, Russell, J (editors) (2010) Tackling wicked problems: through the transdisciplinary imagination. Earthscan, London and Washington, DC.

[7] Collins, H, Evans, R (2008) Rethinking expertise, University of Chicago Press, Chicago and London.

[8] Von Schomberg, R, Hankins, J (2019) International Handbook on Responsible Innovation. Edward Elgar Publishing, Cheltenham and Northampton, MA.

[9] Van den Hoven, J (2013) Value sensitive design and responsible innovation, Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society, Owen, R, Bessant, J, Heintz, M (editors), John Wiley & Sons, Chichester, Oxford, & Malden MA, pp. 75-83.

[10] Polmear M, Bielefeldt AR, Knight D, Canney N, Swan C. Analysis of macroethics teaching practices and perceptions in engineering: a cultural comparison. European Journal of Engineering Education, Vol. 44, No. 6, pp. 866-81.

[11] Zandvoort, H, Børsen, T, Deneke, M, Bird, SJ (2013) Editors' overview perspectives on teaching social responsibility to students in science and engineering. International Journal of Science and Engineering Ethics. Vol. 19, No. 4, pp. 1413-1438.

[12] Børsen, T, Serreau, Y, Reifschneider, K, Baier, A, Pinkelman, R, Smetanina, T, et al. (2020) Initiatives, experiences and best practices for teaching social and ecological responsibility in ethics education for science and engineering students, European Journal of Engineering Education, E-publication before print.

[13] Børsen, T, Danborg, PB (2013) Techno-Anthropological Ethics and Health Information Systems Technologies, Techno-Anthropology in Health Informatics: Methodologies for Improving Human-Technology Relations, Botin, L, Bertelsen, P, Nøhr, C (editors) IOS Press, Amsterdam, Berlin, Tokyo, & Washington DC, pp. 83-94.

[14] Børsen, T, Baunerhøj, AC (2016) En etisk vurdering af fremtidens svineproduktion, Fødevareetik, Coff, C (editor) Munksgaard, Frederiksberg, pp. 271-92.

[15] Børsen, T, Nielsen, SN (2017) Applying an Ethical Judgment Model to the Case of DDT, HYLE–International Journal for Philosophy of Chemistry, Vol. 23, No. 1, pp. 5-27.

[16] Birkholm, K (2013) Human Enhancement as Techno-Anthropology par Excellence, What is Techno-Anthropology? Børsen, T, Botin, L, (editors), Aalborg University Press, Aalborg, pp. 91-116.

[17] Birkholm, K, Frølich, N (editors) (2018) De skjulte algoritmer: Teknoantropologiske perspektiver, DJØF Forlag, Copenhagen.

[18] Blok, A, Jensen, TE (2011). Bruno Latour: Hybrid thoughts in a hybrid world. Routledge, London & New York.

[19] Maedche, A (2017) Interview with Prof. Jeroen van den Hoven on why do ethics and values matter in business and information systems engineering? Business & Information Systems Engineering, Vol. 59, No. 4, pp. 297-300.