

Ethical Issues In Workflow Of Developing Affective Companion Technologies

Andrej Gogora¹ and Marek Debnár²

¹ Department of General and Applied Ethics, Faculty of Arts, Constantine the Philosopher University in Nitra, Nitra, Slovakia
agogora@ukf.sk

² Department of General and Applied Ethics, Faculty of Arts, Constantine the Philosopher University in Nitra, Nitra, Slovakia
mdebnar@ukf.sk

Abstract. In terms of professional ethics (as a part of applied ethics) we are aimed at the issue of professional performance of the computer system designers, software engineers, computer programmers (and related staff) who are developing affective companion technologies and emotion-aware technologies. From this point of view, we are not primarily interested in ethical concerns related to the consequences of such systems and technologies in potential everyday use, as well as not in some social reflections and expectations on affective companions (both an important questions). Instead, at this stage of research we are focusing on the ethical consideration of individual members of working group (stakeholder, project manager, software engineers, programmers) which play a crucial role in creating the overall design of computer systems intended to be used as affective companion technologies. We are interested to observe and analyse the respective ethical issues of working processes, moral professional dilemmas, way of thinking about ethical issues related to the development of such technologies, way of decision-making of morally ambiguous work assignment, as well as way of communicating on these ethical/moral issues in the work team and to the public.

Keywords: professional ethics, affective companion technologies, computer ethics, computing professionals, workflow design

1 Introduction

Ubiquity computing influences society on technological as well as on moral or ethical level. Proponents and opponents of these new technologies have opened a huge debate with strong emphasis on positive and negative ethical consequences of these facilities for well-being of users as well as on moral responsibility of computing professionals and stakeholders. The potential users, decision makers and whole society need to be sure that the ethical principles and moral values of new technologies related to AI/AS (artificial intelligence/autonomous systems) should be/are inevitable implemented in these technologies before they are introduced into everyday use. The intention of our study is to analyze how moral values and ethical principles are embodied in creating of technological artefacts and how applied ethicists can be a part of this process.

2 Field of Research and Aims

Affective companion technologies (ACT) are next generation of intelligent systems aimed at human-machine (emotional) interaction for purpose of improving the quality of life, which smartly adapt to individual requirements, needs and adjust user's emotional state or disposition. These features are realized by integrating technical functionality of hardware and software systems with a combination of cognitive processes and ethical principles as a basic part of these systems. It means that the system should be able to recognize and interpret user's intentions and react at user's behavioral strategies. Consequently, AI/AS systems need to be interpreted, reflected and created from the perspective of individual and social ethics, with aim to prioritize the maximum benefit to humanity and mitigate risks and negative impacts [Etzioni, A. - Etzioni, O., 2017].

In terms of professional ethics (as a part of applied ethics) we are aimed at the issues of professional performance of the computer system designers, software engineers, computer programmers (and related staff) who are designing and developing affective companion technologies and emotion-aware technologies (e.g. IEEE Code of Conduct, IEEE Code of Ethics, IEEE Ethically Aligned Design). From this point of view we are not primarily interested in ethical concerns related to the consequences of such systems and technologies in potential future everyday use, as well as not in some social moral reflections and expectations on affective companions (both an important questions) or in some political implications and policy making [Erdélyi - Goldsmith, 2018]. Instead, we are mainly focusing on the ethical

made-to-measure consideration of individual members of working group (stakeholder, project manager, software engineers, programmers) and their respective workloads which play a crucial role in creating the overall design of computer systems intended to be used as affective companion technologies. In short, we are not interested in “ethics by design” or “ethics in design” but in “ethics for design” [Dignum, 2018].

More precisely, from the point of view of professional ethics we are aimed at the ethical issues emerging in the whole process of developing of these systems, e.g. initial idea and first draft plan, creating of complex theoretical model (conceptual level), selection of co-workers and team settings, direction of working meetings and form of personal negotiations, decision-making processes, hierarchical and economical relationships (managerial and decision-making level), data mining and storage, code conducting, application process, test phase, corrective interventions (computing and applicative level), personal work habits and relationships, confessions, beliefs and personal backgrounds, gender proportion (personal and interpersonal level).

We are interested to detect the ethically critical points in this workflow, to observe and analyze the respective ethical issues and moral dilemmas present in the professional working process of building ACT, and on the basis of that to outline the ethical standards with respect for responsibility and transparency (in a specific workgroup environment). Moreover, we are concerned ourselves with the way of thinking of working group and its members about ethical issues related to the development of such technologies, also with the way of decision-making and performance of morally ambiguous work assignment, as well as with the way of communicating on these ethical and moral issues in the work team and to the public, what we consider to be one of the fundamental questions.

2.1 Localization and Interdisciplinarity

We are convinced that future development of AI/AS systems but also some type of another advanced technologies need broader interdisciplinary cooperation, which brings together scientists from humanities, social science and engineering disciplines. In short, applied ethics analysis and consideration should be a significant part of the workflow of developing affective companion technologies. According to the structure of examined issue, the scientific approach we are proposing is located at the intersection of various applied ethics such as mainly professional ethics, computer and information ethics [Gotterbarn - Miller - Rogerson, 1997; Gotterbarn, 1991], managerial

ethics and social ethics, as well as STS (science and technology studies) and some kind of social anthropology of work environment. The cooperation of information science, applied informatics (and disciplines related to the professional developing of affective companion technologies) with humanities discipline such as applied ethics should be beneficial to its technologically successful and morally unimpeachable implementation into everyday use.

2.2 Theoretical and Methodological Assumptions

The specific behavior of AI/AS technologies must be safe and predictable for humans in many domains, with many consequences, including the ethical problems the engineers never envisioned. In general, all criteria that we apply to humans performing social functions are criteria that must be considered in technology's algorithm to replace human judgement. The process of utilizing multiple ethical approaches to probably aligned end user values must provide a key competitive differentiator in the algorithm. Technologies which are aligned to humans should be instructed in terms of our moral values and ethical principles. Therefore ethical cognition itself must be taken as a subject matter also in engineering. Professional ethics and ethical reflection need to be a core subject for engineers beginning at university level and for all advance degrees. Analyses are impossible to be done behind the borders of research at other institution via armchair, but the professionals in ethics participating at interdisciplinary cooperation need to be present from the beginning of the AI/AS research.

On this basis we are primarily not interested in traditional ethical analysis that is dealing with very abstract, detached and boresome moral concepts and principles. We are not proposing to strictly attribute these concepts and principles to new technology and to convincingly conduct the cultural criticism of its contemporary effects of surveillance and de-personalization. Instead, we declare our commitment to observe and analyze ethical and moral issues directly in the processes of professional designing and creating of affective companion technologies and related computational systems (in a made-to-measure way). This approach is theoretically and methodologically based on the actor-network theory [Latour - Woolgar, 1986; Latour, 2005; Law - Hassard, 1999; Law, 1999] that strictly insist upon the tracking of actors (any subject of the investigation) directly in its own environment (and trying not to assign them extrinsic feature by means of analysis).

Thus, if we assume that the way ACT (and other AI technologies) react to humans is clearly determined by the way their algorithms are designed and

programmed by software developer, we are obliged to ethically analyze the workflow and the whole process of professional creating of these technologies. These professional processes are generating (intentionally or unintentionally) the specific ethical issues and moral threats of new technologies such as problem of humanity, tech addiction, directing of human actions, AI security, robot rights, singularity of human, artificial stupidity and other unforeseen ethical consequences. The lifecycle of the respective workflow is the right place to capture these ethical and moral issues by means of thick description, professional consultation, ethical expertise, working team discussion and highly situationally responsive ethical analysis.

2.3 Outputs and Purposes

The supposed benefits of this approach should be the reflection of specialized ethical problems and dilemmas directly in professional environment, interdisciplinary sharing of methodology and research experience between software developers (and related stuff) and applied ethicists, as well as various form of their cooperation such as consulting expertise, ethical coaching and ethical committees decision-making. We are convinced that the successful fulfillment of the aims should be useful for the balanced human-machine relationship.

3. Summary

In conclusion, we are convinced that focusing ourselves as applied ethicists on professional ethics and made-to-measure ethical analysis of the specific workgroup developing ACT and other AI technologies (and respective workloads) should be more beneficial to better understanding of ethical consequences of these technologies for individual users and society (and their dynamics) than the traditional armchair ethical consideration of these issues.

References

1. Etzioni, A. & Etzioni, O.: *The Journal of Ethics* (2017) 21: pp. 403-418.
<https://doi.org/10.1007/s10892-017-9252-2>
2. Erdélyi, O. J., Goldsmith, J.: *Regulating Artificial Intelligence Proposal for a Global Solution*.
http://www.aies-conference.com/wp-content/papers/main/AIES_2018_paper_13.pdf
3. Dignum, V.: *Ethics and Information Technology* (2018) 20: pp. 1-3.
<https://doi.org/10.1007/s10676-018-9450-z>
4. Gotterbarn, D., Miller, K., Rogerson, S.: *Software Engineering Code of Ethics*, In: *Information Society* (1997) 40(11): pp. 110–118.

5. Gotterbarn, D.: Computer Ethics: Responsibility Regained, In: *National Forum: The PhiBeta Kappa Journal*,71 (1991): pp. 26–31.
6. Latour, B., Woolgar S.: *Laboratory life: the construction of scientific facts*. Princeton, New Jersey: Princeton University Press (1986).
7. Latour, B.: *Reassembling the social: an introduction to actor-network-theory*. Oxford New York: Oxford University Press (2005).
8. Law, J, Hassard J.: *Actor network theory and after*. Oxford England Malden, Massachusetts: Blackwell/Sociological Review (1999).
9. Law, J.: *Organizing modernity*. Oxford, UK Cambridge, Massachusetts, USA: Blackwell (1994).
10. Ramsey, W., Keith F. (eds.): *Cambridge Handbook of Artificial Intelligence*.Cambridge University Press (2011).
11. Biundo, S., Wendemuth A. (eds.): *Companion Technology: A Paradigm Shift in Human-Technology Interaction (Cognitive Technologies)* 1st ed. (2017).
12. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems (A/IS)*, version 2 (2018).
http://standards.ieee.org/develop/indconn/ec/ead_v2.pdf