

IS  
20  
19

Zbornik 22. mednarodne multikonference

# INFORMACIJSKA DRUŽBA

Zvezek D

Proceedings of the 22nd International Multiconference

# INFORMATION SOCIETY

Volume D

Etika in stroka

Professional Ethics

Uredili / Edited by

Marjan Krisper, Franci Pivec, Matjaž Gams

<http://is.ijs.si>

9. oktober 2019 / 9 October 2019

Ljubljana, Slovenia

Zbornik 22. mednarodne multikonference  
**INFORMACIJSKA DRUŽBA – IS 2019**  
Zvezek D

Proceedings of the 22nd International Multiconference  
**INFORMATION SOCIETY – IS 2019**  
Volume D

**Etika in stroka**  
**Professional Ethics**

Uredili / Edited by

Marjan Krisper, Franci Pivec, Matjaž Gams

<http://is.ijs.si>

9. oktober 2019 / 9 October 2019  
Ljubljana, Slovenia


Uredniki:

Franci Pivec  
DOBA fakulteta, Maribor

Marjan Krisper  
Fakulteta za računalništvo in informatiko  
Univerza v Ljubljani

Matjaž Gams  
Odsek za inteligentne sisteme  
Institut »Jožef Stefan«, Ljubljana

Založnik: Institut »Jožef Stefan«, Ljubljana  
Priprava zbornika: Mitja Lasič, Vesna Lasič, Lana Zemljak  
Oblikovanje naslovnice: Vesna Lasič

Na naslovnici je uporabljena slika robota podjetja 

Dostop do e-publikacije:  
<http://library.ijs.si/Stacks/Proceedings/InformationSociety>

Ljubljana, oktober 2019

Informacijska družba  
ISSN 2630-371X

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni  
knjižnici v Ljubljani  
COBISS.SI-ID=302452992  
ISBN 978-961-264-161-0 (epub)  
ISBN 978-961-264-162-7 (pdf)

# PREDGOVOR MULTIKONFERENCI INFORMACIJSKA DRUŽBA 2019

Multikonferenca Informaci družba (<http://is.ijs.si>) je z dvaindvajseto zaporedno prireditvijo tradicionalni osrednji srednjeevropski dogodek na področju informacijske družbe, računalništva in informatike. Informacijska družba, znanje in umetna inteligenca so - in to čedalje bolj – nosilci razvoja človeške civilizacije. Se bo neverjetna rast nadaljevala in nas ponesla v novo civilizacijsko obdobje? Bosta IKT in zlasti umetna inteligenca omogočila nadaljnji razcvet civilizacije ali pa bodo demografske, družbene, medčloveške in okoljske težave povzročile zadušitev rasti? Čedalje več pokazateljev kaže v oba ekstrema – da prehajamo v naslednje civilizacijsko obdobje, hkrati pa so notranji in zunanji konflikti sodobne družbe čedalje težje obvladljivi.

Letos smo v multikonferenco povezali 12 odličnih neodvisnih konferenc. Zajema okoli 300 predstavitev, povzetkov in referatov v okviru samostojnih konferenc in delavnic in 500 obiskovalcev. Prireditvev bodo spremljale okrogle mize in razprave ter posebni dogodki, kot je svečana podelitev nagrad. Izbrani prispevki bodo izšli tudi v posebni številki revije Informatica (<http://www.informatica.si/>), ki se ponaša z 42-letno tradicijo odlične znanstvene revije.

Multikonferenco Informacijska družba 2019 sestavljajo naslednje samostojne konference:

- 6. študentska računalniška konferenca
- Etika in stroka
- Interakcija človek računalnik v informacijski družbi
- Izkopavanje znanja in podatkovna skladišča
- Kognitivna znanost
- Kognitonika
- Ljudje in okolje
- Mednarodna konferenca o prenosu tehnologij
- Robotika
- Slovenska konferenca o umetni inteligenci
- Srednje-evropska konferenca o uporabnih in teoretičnih računalniških znanostih
- Vzgoja in izobraževanje v informacijski družbi

Soorganizatorji in podporniki konference so različne raziskovalne institucije in združenja, med njimi tudi ACM Slovenija, SLAIS, DKZ in druga slovenska nacionalna akademija, Inženirska akademija Slovenije (IAS). V imenu organizatorjev konference se zahvaljujemo združenjem in institucijam, še posebej pa udeležencem za njihove dragocene prispevke in priložnost, da z nami delijo svoje izkušnje o informacijski družbi. Zahvaljujemo se tudi recenzentom za njihovo pomoč pri recenziranju.

V 2019 bomo sedmič podelili nagrado za življenjske dosežke v čast Donalda Michieja in Alana Turinga. Nagrado Michie-Turing za izjemen življenjski prispevek k razvoju in promociji informacijske družbe je prejel prof. dr. Marjan Mernik. Priznanje za dosežek leta pripada sodelavcem Odseka za inteligentne sisteme Instituta »Jožef Stefan«. Podeljujemo tudi nagradi »informacijska limona« in »informacijska jagoda« za najbolj (ne)uspešne poteze v zvezi z informacijsko družbo. Limono je dobil sistem »E-zdravje«, jagodo pa mobilna aplikacija »Veš, kaj ješ?!«. Čestitke nagrajencem!

Mojca Ciglarič, predsednica programskega odbora  
Matjaž Gams, predsednik organizacijskega odbora

# FOREWORD - INFORMATION SOCIETY 2019

The Information Society Multiconference (<http://is.ijs.si>) is the traditional Central European event in the field of information society, computer science and informatics for the twenty-second consecutive year. Information society, knowledge and artificial intelligence are - and increasingly so - the central pillars of human civilization. Will the incredible growth continue and take us into a new civilization period? Will ICT, and in particular artificial intelligence, allow civilization to flourish or will demographic, social, and environmental problems stifle growth? More and more indicators point to both extremes - that we are moving into the next civilization period, and at the same time the internal and external conflicts of modern society are becoming increasingly difficult to manage.

The Multiconference is running parallel sessions with 300 presentations of scientific papers at twelve conferences, many round tables, workshops and award ceremonies, and 500 attendees. Selected papers will be published in the *Informatica* journal with its 42-years tradition of excellent research publishing.

The Information Society 2019 Multiconference consists of the following conferences:

- 6. Student Computer Science Research Conference
- Professional Ethics
- Human – Computer Interaction in Information Society
- Data Mining and Data Warehouses
- Cognitive Science
- International Conference on Cognitonics
- People and Environment
- International Conference of Transfer of Technologies – ITTC
- Robotics
- Slovenian Conference on Artificial Intelligence
- Middle-European Conference on Applied Theoretical Computer Science
- Education in Information Society

The Multiconference is co-organized and supported by several major research institutions and societies, among them ACM Slovenia, i.e. the Slovenian chapter of the ACM, SLAIS, DKZ and the second national engineering academy, the Slovenian Engineering Academy. In the name of the conference organizers, we thank all the societies and institutions, and particularly all the participants for their valuable contribution and their interest in this event, and the reviewers for their thorough reviews.

For the fifteenth year, the award for life-long outstanding contributions will be presented in memory of Donald Michie and Alan Turing. The Michie-Turing award was given to Prof. Marjan Mernik for his life-long outstanding contribution to the development and promotion of information society in our country. In addition, a recognition for current achievements was awarded to members of Department of Intelligent Systems of Jožef Stefan Institute. The information lemon goes to the “E-Health” system, and the information strawberry to the mobile application “Veš, kaj ješ?!” (Do you know what you eat?!). Congratulations!

Mojca Ciglarič, Programme Committee Chair  
Matjaž Gams, Organizing Committee Chair

# KONFERENČNI ODBORI

## CONFERENCE COMMITTEES

### *International Programme Committee*

Vladimir Bajic, Južna Afrika  
Heiner Benking, Nemčija  
Se Woo Cheon, Južna Koreja  
Howie Firth, Škotska  
Olga Fomichova, Rusija  
Vladimir Fomichov, Rusija  
Vesna Hljuz Dobric, Hrvaška  
Alfred Inselberg, Izrael  
Jay Liebowitz, ZDA  
Huan Liu, Singapur  
Henz Martin, Nemčija  
Marcin Paprzycki, ZDA  
Claude Sammut, Avstralija  
Jiri Wiedermann, Češka  
Xindong Wu, ZDA  
Yiming Ye, ZDA  
Ning Zhong, ZDA  
Wray Buntine, Avstralija  
Bezalel Gavish, ZDA  
Gal A. Kaminka, Izrael  
Mike Bain, Avstralija  
Michela Milano, Italija  
Derong Liu, Chicago, ZDA  
Toby Walsh, Avstralija

### *Organizing Committee*

Matjaž Gams, chair  
Mitja Luštrek  
Lana Zemljak  
Vesna Koricki  
Marjetka Šprah  
Mitja Lasič  
Blaž Mahnič  
Jani Bizjak  
Tine Kolenik

### *Programme Committee*

Mojca Cigliarič, chair  
Bojan Orel, co-chair  
Franc Solina  
Viljan Mahnič  
Cene Bavec  
Tomaž Kalin  
Jozsef Györkös  
Tadej Bajd  
Jaroslav Berce  
Mojca Bernik  
Marko Bohanec  
Ivan Bratko  
Andrej Brodnik  
Dušan Caf  
Saša Divjak  
Tomaž Erjavec  
Bogdan Filipič

Andrej Gams  
Matjaž Gams  
Mitja Luštrek  
Marko Grobelnik  
Vladislav Rajkovič  
Grega Repovš  
Nikola Guid  
Marjan Heričko  
Borka Jerman Blažič Džonova  
Gorazd Kandus  
Urban Kordeš  
Marjan Krisper  
Andrej Kuščer  
Jadran Lenarčič  
Borut Likar  
Janez Malačič  
Olga Markič

Dunja Mladenič  
Franc Novak  
Ivan Rozman  
Niko Schlamberger  
Stanko Strmčnik  
Jurij Šilc  
Jurij Tasič  
Denis Trček  
Andrej Ule  
Tanja Urbančič  
Boštjan Vilfan  
Baldomir Zajc  
Blaž Zupan  
Boris Žemva  
Leon Žlajpah



## KAZALO / TABLE OF CONTENTS

<b><i>Etika in stroka / Professional Ethics</i></b> .....	<b>1</b>
PREDGOVOR / FOREWORD.....	3
Etika v raziskavah: razvoj vprašalnika za situacijsko analizo / Stres Špela .....	5
Etika v raziskavah, primer JRC "TTO Circle" organizacij / Stres Špela .....	9
Automation of Violence and the Disappearance of Moral Responsibility / Grušovnik Tomaž .....	12
Etični vidiki digitalizacije / Bavec Gene.....	15
Osnove računalniške etike / Gams Matjaž .....	19
Etična uporaba podatkov družbenih medijev / Štebe Janez, Sergeja Masten .....	23
Etični kodeks strokovnjakov računalništva in informatike / Krisper Marjan .....	26
Ethics First! / Pivec Franci .....	31
Etika naj postane del študija informatike / Pribac Igor.....	35
 <b><i>Indeks avtorjev / Author index</i></b> .....	 <b>39</b>





Zbornik 22. mednarodne multikonference  
**INFORMACIJSKA DRUŽBA – IS 2019**  
Zvezek D

Proceedings of the 22nd International Multiconference  
**INFORMATION SOCIETY – IS 2019**  
Volume D

**Etika in stroka**  
**Professional Ethics**

Uredili / Edited by

Marjan Krisper, Franci Pivec, Matjaž Gams

<http://is.ijs.si>

9. oktober 2019 / 9 October 2019  
Ljubljana, Slovenia



# PREDGOVOR

Konferenca »Etika in stroka« je namenjena etiki v času informacijske družbe. Po Rushworth Kidderju je etika "znanost o idealnem človeškem značaju" oziroma "znanost o moralni dolžnosti". Richard William Paul in Linda Elder definirata etiko kot "nabor konceptov in principov, ki nas vodijo glede primerne in neprimerne obnašanja". V času informacijske družbe pa vse kaže, da se stare norme podirajo, da je osnova uspešnega političnega delovanja čim bolj uspešno zavajanje množic preko socialnih omrežij in popolnoma pristranskih medijev. Kot da resnica, moralnost, etika ne bi več obstajali, oziroma so nepomembni.

Posebej se to kaže v odnosu do znanosti oziroma dejstev oz. resnice. Ali je cepljenje škodljivo ali koristno? Ali lahko nasprotniki cepljenja nekaznovano povzročajo škodo svojim otrokom in otrokom drugih? Ali je omejevanje resnice s strani tehnoloških gigantov kot Google ali Facebook omejevanje svobode, ali pa gre za omejevanje hujskanja množic in sovražnega govora?

Zanimive teme za prispevke so naslednje: Etični kodeksi v računalništvu/informatiki, občutljivost univerzalne (filozofske) etike za digitalni svet, etična odgovornost IT strokovnjakov, IT v službi zavajanja, etika v e-trgovanju, etika digitalnega sodišča, etičnost Googla in Facebooka, avtentičnost e-dokumentov v arhivih, etika in umetna inteligenca, etika v robotiki.

Poseben poudarek je na aktualni temi: Etične smernice za zaupanja vredno Umetno inteligenco, na podlagi dokumenta Evropske komisije Ethics Guidelines for Trustworthy Ai, ki zajema ključna etična vprašanja informacijske družbe.

Prvo leto konference smo zbrali osem zanimivih referatov, naslednje leto pa upamo na boljši odziv. Sočasno s konferenco bomo zbrali strokovna mnenja o etiki in predloge za izboljšanje. V skladu s pravili posveta v Državnem svetu bomo vprašanja oziroma pobude posredovali Vladi RS in predsedniku RS.

Franci Pivec, Marjan Krisper



# Etika v raziskavah: razvoj vprašalnika za situacijsko analizo

## Ethics in research issues: development of a situational analysis questionnaire

dr. Špela Stres  
Jožef Stefan Institute, Center for  
Innovation and Technology Transfer  
+386 1 477 3200  
[spela.stres@ijs.si](mailto:spela.stres@ijs.si)

### ABSTRACT

In this paper, we describe the development of a questionnaire to be used in situational analysis of ethics principles in research matters in the organizations.

### Keywords

Ethics, principles, survey

## 1. INTRODUCTION

As a part of the H2020 funded NewHorizon project on responsible research and innovation we were investigating the question of the importance of non-regulatory / non-conventional ethics and research integrity issues in European Public Research Organizations to look beyond standard ethics regulatory issues and processes.

In our everyday life of a Center for Innovation and Technology Transfer we meet many researchers. Whenever we ask them a question: "What do you think about ethics in research?", we get a specific answer. The answer is: "----"(silence and a perplexed look in their eyes).

Thus, we decided to develop a research tool to analyze the perception of ethical behavior in different situations within public research organizations from researcher's personal and institutional point of view.

## 2. PREPARATION OF THE SURVEY

### 2.1 The motivation

The ethics issues we were interested in, went beyond the conventional ethical issues (as integrity, responsibility, honesty, competence) [3 General research on ethics] or more philosophical ones (dignity, non-maleficence) [4 JRC handbook] and strived away from the IPR issues (privacy, confidentiality, justice) [5 WIPO Handbook].

Also, the issues we set to analyze, were broader, but due to addressing specific situations also more concrete than the ones included in the [6 Consensus statement]. Even though this document emphasizes that the responsibility for ethical research lies with everyone who is active in research, but especially with leaders in research performing organizations, it remains open that the researchers' morals alone cannot ensure research integrity; good conditions for exercising integrity must also be created at the level

of the organization and the research system. Also, in short-term, project-based positions, the role of the project leader in instilling ethical standards are crucial, as staff on shorter contracts are often not integrated in the organization to the same extent as permanent staff.

Having considered this, our analysis was based on a specific set of situations that were described as potentially ethically problematic by a group of scientists from three different countries, Sweden, UK and Slovenia. In the mentioned group Sweden represented the Nordic approach, UK the central European one and Slovenia the Balkan region approach to the ethical issues. The chosen group selected the initial set of situations to be surveyed and analyzed within the TTO Circle. The situations were modeled by JSI and the questions were developed at JSI.

## 3. RESULTS: THE QUESTIONNAIRE

### 3.1 The general motivation part

There were five main sets of questions in the questionnaire.

**The first set of questions** focused on the principles in the context of scientific research that would or would not require any of the solutions, focused on improving the ethics and ethical attitudes in those situations. An example of the questionnaire matrices is given below in Figures 2-5.

The general principles focused on topics as reliability in ensuring the quality of research, the honesty in developing and undertaking, reviewing, reporting research; the Respect for colleagues, participants, society; and the Accountability for the research from idea to publication, for its management and organization.

The main idea was to distinguish between different principles that should all be equally underlying in the research work.

There were four (4) principles and five (5) different categories offered with different levels of formalization, ranging from non-regulatory, through a code of conduct or a local policy, through legal framework in place and to on the spot enforcement. Also, the option of "no particular solution is required" was offered to enable those, who believe researchers need to be individualistic and of free

will with no external boundaries also when ethical behavior is considered, to share their opinion (see Fig 1).

1. In the context of scientific research, do you think the following principles require any of the following solutions? \*

	non-regulatory awareness raising via the listing of community expectations	a code of conduct or local policy	legal framework	enforcement	no particular solution is required
Reliability in ensuring the quality of research, reflected in the design, the methodology, the analysis and the use of resources;	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair, full and unbiased way;	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment;	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Accountability for the research from idea to publication, for its management and organization, for training, supervision and mentoring, and for its wider impacts;	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 1: The connection between research principles and required ethical monitoring solutions.

### 3.2 The personal involvement

The second set of questions focused on a set of specifically chosen situations, and questioned the necessity for each of these situations to be considered within an ethical assessment.

The focus within this set of questions was on a personal point of view, thus the described situations were described in a very addressing way. The purpose was to emotionally involve the respondent to answer from this or her personal point of view, neglecting the policies that could be in place in the environment he or she operates in.

The idea was to set the scene for later, organizational behavior-based questions. The inherent logic was that the questions and the answers should enable the topic of ethics in everyday research life to become internalized, to be considered as something that we are all connected to, before stating the position of the organization.

Thus, situations as “accepting invitations to panels that did not make a demonstrated effort in gender equality” or “asking a new student to do measurements and using these measurements in a paper without giving him credit” or “the process that senior authors decide the order of the author list” were introduced.

2. Do you consider this to be a SITUATION IN NEED OF AN ETHICAL ASSESSMENT? \*

	YES	NO
STEM cells research	<input type="checkbox"/>	<input type="checkbox"/>
research involving animals	<input type="checkbox"/>	<input type="checkbox"/>
research involving children	<input type="checkbox"/>	<input type="checkbox"/>
research involving adults	<input type="checkbox"/>	<input type="checkbox"/>
the process that senior authors decide who is included in author list of a publication	<input type="checkbox"/>	<input type="checkbox"/>
the process that senior authors decide the order of the author list	<input type="checkbox"/>	<input type="checkbox"/>
using an ICT tool for internal submission seen by everyone so anyone can request authorship	<input type="checkbox"/>	<input type="checkbox"/>
scheduling meetings outside 'core working hours'	<input type="checkbox"/>	<input type="checkbox"/>
organizing conferences that require travel at weekends	<input type="checkbox"/>	<input type="checkbox"/>
accepting invitations to panels that did not make a demonstrated effort in gender equality	<input type="checkbox"/>	<input type="checkbox"/>
formation of an interview board for hiring processes	<input type="checkbox"/>	<input type="checkbox"/>
asking a new student to do measurements and using these measurements in a paper without giving credit	<input type="checkbox"/>	<input type="checkbox"/>
researching new topics without a broad social agreement through a consultation on consequences	<input type="checkbox"/>	<input type="checkbox"/>
activities that result in personal financial benefit for the researcher	<input type="checkbox"/>	<input type="checkbox"/>

Figure 2: The situations listed to be assessed with respect to the necessity of ethical assessment from a personal point of view.

They were modeled in an inviting and personal way on purpose to incur personal responses.

### 3.3 The organizational view

The third part of the enquiry questioned about a similar, but less personally engaging set of situations, but from the point of view of the organization.

Does the organization itself perceive the situation to need an ethical assessment. Here we tried to see the difference between personal involvement in the ethical issues of personal situations that researchers need to live through and the ones recognized from the side of the institution. The difference to the previous set of questions is obvious. Here we address in a number of simple, straightforward situations, that can be easily understood in any research environment.

3. Does your organisation consider there to be ethical aspects and take them into account in the following settings? \*

	YES	NO	I am not sure
prizes/awards committees	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
performance reviews	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
recruitment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
letters of recommendation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
governing bodies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
steering or advisory boards	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
internal faculty funding	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
editorial boards	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
project invitations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
visibility	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
project coordination	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
CV preparation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
announcement of positions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
speaking events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
choosing research topics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figure 3: The situations listed to be assessed with respect to the necessity of ethical assessment from an organizational point of view.**

As such, they do not carry much emotional co-signature and it is thus more difficult to misinterpret them, but they are also not seen as situations in which the respondent is necessarily involved.

We wanted the respondents to enter the administration point of view, stating the views of the management of the research organization.

### 3.4 The possible remedies

The **fourth set of questions** focused on possible remedies or solutions that could be used by the organization to address the situations in need of ethical assessment and analyzed the use of such tools.

The listed solutions range from very light ones as for example awareness raising via listing of expectations from the community, which would majorly serve as an encouragement to think about such topics as ethics. Suggested solutions are: Through informal and formal guidelines and code of conduct or a local policy; to more strict and formal measures as the ethical review committee (and several of them in different fields of research) or the enforcement of some detailed procedural documents or policies; to the last and most prescribed solution in the form of a legal framework in terms of a national legislation.

4. In what form does your organisation take ethical issues into account? \*

	YES	NO
Awareness raising via listing of the expectations from the community	<input type="radio"/>	<input type="radio"/>
Informal guidelines	<input type="radio"/>	<input type="radio"/>
Formal guidelines	<input type="radio"/>	<input type="radio"/>
Code of conduct or local policy	<input type="radio"/>	<input type="radio"/>
Ethical review committee	<input type="radio"/>	<input type="radio"/>
Field and context dependent ethical committees	<input type="radio"/>	<input type="radio"/>
Procedural documents (in terms of policy)	<input type="radio"/>	<input type="radio"/>
Legal framework (in terms of national legislation)	<input type="radio"/>	<input type="radio"/>
My organisation does not take ethical issues into account in the described settings	<input type="radio"/>	<input type="radio"/>

**Figure 4: The possible tools to address the situations in need of ethical assessment, from an organizational point of view.**

We also offered the option where the respondents could state that their organization does not take ethical issues into account in the described settings.

There were also some **content questions**, assessing the reasons about why and how does/ does not a specific research organization consider there to be ethical aspects and does/ does not take them into account to allow for personal explanations of the ethics in research as seen by the respondents. (in their own words).

In this part we dealt with issues of ethics in research beyond regulatory compliance from a personal point of view: single and outlying views were sought for here. We were interested in issues that the respondents see of interest beyond what is regulated in the sense of animals' rights and informed consent but could be related to management of science and technology and multilateral collaborations etc.

### 3.5 The tools available vs. the need

In the **final part** the questionnaire focused on matching the situations in need of ethical assessment with the tools available to deal with them – in particular it focused on the usage of the tools in practice.

Here we tried to understand the relation between the theoretical assessment of the need and the readiness to act on such needs institutionally.

This complex matrix was developed to obtain an insight to the actual type of remedy that the respondent should think would need to be available from the side of the institution, in order to fully address the ethical views on particular situations. The idea was to obtain actual strength of the remedy to be considered with respect to a particular situation. The starting point of the research was that some situations would require stronger remedies than other, but



perhaps not all of them would be allowed the same level of interference due to the requested autonomy of the researchers.

**ADDITIONAL 3: What type of INPUT do you think SHOULD be AVAILABLE from your Public Research Organization on ethical issues in the following settings (non optional extention of question 3)**

	1. Awareness raising	2. Informal guidelines	3. Formal guidelines	4. Code of conduct	5. Ethical review committee	6. Several field dependent ethical committees for different contexts, situations, settings	7. Several field dependent ethical committees for different fields of science
prizes/awards committees	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
performance reviews	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
recruitment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
letters of recommendation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
governing bodies	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
steering or advisory boards	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
internal faculty funding	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
editorial boards	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
project invitations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
visibility	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
project coordination	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CV preparation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
announcement of positions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
speaking events	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
choosing research topics	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Figure 5: The usage of the tools to address the situations in need of ethical assessment, from the organizational point of view.**

#### 4. CONCLUSIONS

The set up matrix of questions will allow further work and analysis about the use of ethical principles in public research institutions from a multiplicity of points of view: the general attitude, the more personalized point of view, the view of the organization and the possible remedies.

We plan to continue our work in terms of applying the questionnaire to a chosen set of relevant public research organizations in Europe, the JTC TTO Circle [1, 2].

#### 5. ACKNOWLEDGMENTS

Our thanks to Michael Bernstein for allowing us the time to develop the attitudes to the questionnaire and to supply us with some important documents in the field. Thanks to Johan Benesch (Chalmers University) and Huw Jones (Aberystwyth University)

for their inspiring support in the first phases of setting up the research questions.

#### 6. REFERENCES

[1] JRC TTO Circle Publication, 18.8.2019, <https://ec.europa.eu/jrc/communities/en/community/tto-circle-community/>.

[2] JRC TTO Circle Publication, 18.8.2019, <https://ec.europa.eu/jrc/communities/en/community/european-tto-circle/page/tto-circles-members>

[3] Eric Breit et al.. 2016. Promoting Integrity as an Integral Dimension of Excellence in Research. EU project No. 665926 report.

[4] Boucher Philip et al.. Ethics dialogues. Experiencing Ethics through ‘things’. 2014. ISBN 978-92-79-45049-5

[5] WIPO. 2017 WIPO. Successful Technology Licensing. Publication No. 903E 2015 Edition ISBN 978-92-805-2633-2

[6] Ellen-Marie Forsberg et al.. 2018. EUROPEAN CONSENSUS STATEMENT. Sci Eng Ethics, 24:1023–1034. Working with Research Integrity — Guidance for Research Performing Organizations: The Bonn PRINTEGER Statement. <https://doi.org/10.1007/s11948-018-0034-4>

# Etika v raziskavah, primer JRC "TTO Circle" organizacij; Ethics in research issues, an example of JRC "TTO Circle" organizations

dr. Špela Stres  
Jožef Stefan Institute, Center for  
Innovation and Technology Transfer  
+386 1 477 3200  
[spela.stres@ijs.si](mailto:spela.stres@ijs.si)

## ABSTRACT

In this paper, we describe the use of ethics principles in research matters in the organizations, which form part of the Joint Research Center's TTO Circle ([1]).

## Keywords

Ethics, principles, survey, TTO Circle, JRC

## 1. INTRODUCTION

As a part of the H2020 funded NewHorizon project on responsible research and innovation we were investigating the question of the importance of non-regulatory / non-conventional ethics and research integrity issues in European Public Research Organizations to look beyond standard ethics regulatory issues and processes.

This work is based on previously developed situational analysis questionnaires for analysis of ethics principles in research matters in the research organizations, presented in our article "Ethics in research issues: development of a situational analysis questionnaire".

We carried out a short survey about the perception of ethical behavior in different situations within public research organizations from researcher's personal and institutional point of view.

## 2. PREPARATION OF THE SURVEY

The ethics issues we were interested in, went beyond the conventional ethical issues (as integrity, responsibility, honesty, competence) [3 General research on ethics] or more philosophical ones (dignity, non-maleficence) [4 JRC handbook] and strived away from the IPR issues (privacy, confidentiality, justice) [5 WIPO Handbook]. Also, the issues we set to analyze, were broader, but due to addressing specific situations also more concrete than the ones included in the [6 Consensus statement].

The survey took place among the members of the TTO Circle, which currently are 31. This is a finite amount of data, but highly prominent at the same time, as high-level officials of Public Research Organizations are involved in the TTO Circle operations ([1]).

The TTO Circle stands for the European Technology Transfer Offices circle. This is a network of research institutions, established with the aim to bring together the major public research organizations in order to share best practices, knowledge and

expertise, perform joint activities and develop a common approach towards international standards for the professionalization of technology transfer.

The European TTO circle gathers the largest public research organizations across Europe. The network comprises currently 31 organizations (198349 scientific staff, 5243 softwares, 34338 patents and 4143 start-ups). The partners signed a Memorandum of Understanding formalizing their collaboration. They agreed to strengthen Europe's ability to create innovative products and services for the market. ([2]).

Jožef Stefan Institute has become a member of the TTO Circle in 2016. During the 2018 and 2019 a survey and an analysis of research ethics attitudes and behaviors has been carried out to understand the level of inclusivity for ethics and ethical assessments within the situations that arise in research operations.

## 3. EXECUTION OF THE SURVEY

A separate platform has been built to allow only single responses from specific institutions. It made available the basic information and rights of the respondent at any time and collected responses that could be reviewed but not changed after submission.

A developed situational analysis questionnaire has been uploaded to the platform and the link sent to the selected representatives of the member public research organizations of the TTO Circle.

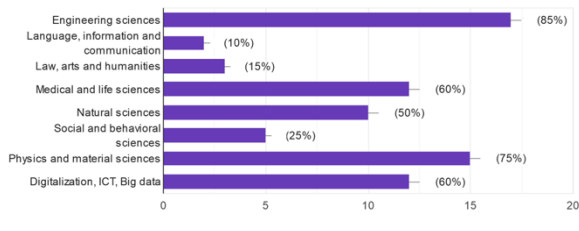
The questionnaires were answered by the selected representatives. We collected 22 filled out questionnaires out of the 31 institutions. The TTO Circle organizations are a representative sample of the highest level of the European research attitudes. An almost 71% response rate confirms a high involvement of these organizations with the ethical issues.

The research organizations interviewed engage in different fields of research work (Fig.1), but the majority of them is involved in research and development also in the IT field (85% engineering sciences, 10% language, information and communication, 60% digitalization, ICT, big data).

These organizations are closely related to the issues of ethics of IT specialists and other specialists working in the research field and is crucial for the relation of science or facts to the truth as accepted by the general society.

However, there is also a wider question to be answered, namely, is the scientific and research operation itself immune to ethical issues?

Main research fields of your organisation (tick all that apply)



**Figure 1: The distribution of research fields of the interviewed research organizations.**

And how are particular situations, principles, themes addressed when faced in the research community, either from a personal point of view of a certain researcher or from the point of view of the whole organization?

## 4. RESULTS

### 4.1 The personal attitudes

The **necessity to involve ethical assessments in any of the four principles (reliability, respect, honesty and accountability)** was analyzed within the first set of questions.

The analysis showed that “reliability in ensuring the quality of research reflected in the design, the methodology, the analysis and the use of resources”, is the most unifying principle among the four. More than 68% of respondents believe a code of conduct or a local policy should be set in place to assure reliability in research quality.

On the other hand, the only principle that evoked a response “no practical solution is required”, even if in only 9% of respondents, was “respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment”. This can be understood as a proof that the respect for all involved in the research actions is either already very high (and thus no particular solution is required) or the situation is in the minds of the researchers to be left unaddressed (best as it is and as left to the research autonomy). Mind that 45% of respondents think that non-regulatory and 45% of respondents think that a code of conduct or local policy would be needed to address the issues of mutual respect in the research society.

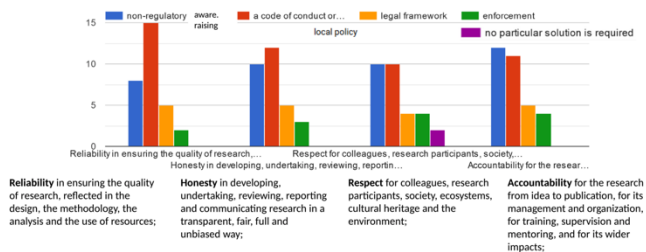
The latter (non-regulatory, awareness raising or a code of conduct and local policy) were also by far the most popular answers in all four principles in question (reliability, honesty, respect and accountability). As opposed to more than 77% of respondents asking for non-regulatory, awareness raising, code of conduct or local policy actions, the legal framework would be used by 21% of respondents on average and enforcement by only 14% of respondents (see Fig.2).

The **personal level of necessity to establish ethical assessment in some particular common research situations** was evaluated within the second set of questions. There were some situations that clearly showed preference of the respondents towards execution of ethical assessments: research involving animals 95%, research involving children 91%, STEM cells research 86%, activities that result in personal financial benefit for the researcher 86%, research

involving adults 82%, asking a new student to do measurements and using these measurements in a paper without credit 82%.

There were situations that clearly showed the opposite preference – that the ethical assessment would be not needed: scheduling meetings outside the core working hours (77% thought this is not an issue in need of ethical assessment); and 73% thought that organizing conferences that require travel at weekends is not problematic either. It would be important to observe the distribution of answers here in terms of female and male respondents, but we do not have this data available.

1. In the context of scientific research, do you think the following principles require any of the following solutions?



**Figure 2: The solutions needed to assure the four basic principles in the context of scientific research.**

There was a set of situations where a slight majority would require an ethical assessment: researching new topics without broad social agreement through a consultation on consequences 57%; the process that senior authors decide who is included in the author list of a publication and the process that the senior authors decide the order of the author list, both 55%; formation of an interview board for hiring processes 53%.

There was also a set of situations where a slight majority was against having the ethical assessments: using an ICT tool for internal submission seen by everyone so anyone can request authorship with 59% against and accepting invitations to panels that did not make a demonstrated effort in gender equality 55% against any ethical assessment needed.

The analysis showed a clear bias of the researchers towards already established procedures (which were named as necessary to be in ethical assessment), but less prevalent were topics that affect everyday life (work-life balance) and topics that are as of yet not part of the regulatory system (agreement with the cultural and social environment about the research topics).

Having understood these personal points of view we then also tried to analyze the organizational attitudes and set-ups.

### 4.2 The organizational attitudes

As expected, the organizational attitude has proven to be a more difficult issue than the personal attitude of the respondents, as on the average within the set of 16 questions almost 18% of respondents declared that they are not sure of what type the consideration of ethical issues within their organization is.

On the other hand there were some settings that clearly showed institutional orientation towards ethical assessment: for recruitment 83% of those who knew the institutional orientation regarded this setting as one that is covered by ethical assessment in their organization, 65% of institutions considered there to be necessity

for ethical assessment for announcement of positions and 70% of institutions considered there to be ethical aspects and take them into account in terms of members of governing bodies and their operations with 56% considered there to be ethical aspects and take them into account in their advisory boards.

There were also some settings that clearly showed reluctance to involve any ethical aspect consideration: letters of recommendation 78% of the ones who knew the organizational position claimed their organizations do not consider there to be ethical aspects and do not take them into account; 75% for project invitations and 70% for operations of editorial boards; CV preparation (73%), speaking events (73%), but also (impressively) 64% for internal faculty funding allocation and for memberships.

For some settings the balance was slightly positive (prizes/award committees with 58% in favor of ethical assessments) and for some it was slightly negative (performance reviews 64%, visibility 60%, choosing research topics 58% for rejecting the ethical assessments).

The distribution of institutional orientations shows that the institutions try to address first the situations, connected to ethical aspects, that are also connected to financial and research position benefits, whereas every day research and research work aspects are currently being less considered.

If we wanted to **consider the forms in which the organizations take ethical issues into account**, we found out that the majority of organizations relies on awareness raising via listing of the expectations from the community to address the ethical issues (72%) and on formal guidelines (76%).

Less formalized option of informal guidelines (66%) and a more formal option of the code of conduct or local policy (57%) together with an ethical review committee (53%) were also used. More than 62% of organizations also have a procedural document (in terms of policy) to address the ethical issues and take them into account. The least used tools are field and context dependent ethical committees (76% do not use them) and legal framework in terms of national legislation (53% do not have it).

Surprisingly, 83% of the respondents stated that it is not true that their organizations would not take ethical issues into account in the described settings, namely that only 17% of the respondents are employed in institutions that would not be aware of the importance of ethical issues and ethical assessments in the described settings.

The analysis of forms in which the organizations take ethical issues into account shows that the organizational level of ethical responsibility is higher than the personal one, which could lead to improvements in ethical assessments of particular settings in the future years.

We also investigated **the type of input that the respondents think should be available from their organizations** in particular settings where ethical assessment might be necessary.

In project invitations, visibility, prizes/awards and committees the prevailing requested input from the research organizations to the side of research community would be awareness raising. Whereas in performance reviews the community would require formal guidelines and in project coordination a code of conduct. (and also, awareness raising of the same level). Surprisingly, in CV preparation there is an equal request for awareness raising and for legal framework in terms of national legislation.

Governing bodies should have formal guidelines in ethical aspects of their operation and in recruitment, speaking events and internal faculty funding a strong majority requires procedural documents in

terms of a policy, closely followed by a request for formal guidelines.

Surprisingly, in many situations the research community requests their research organizations to impose a more transparent rule set in a more organized, more formal and more systematic way.

This shows that even though the organizations do take ethical issues into account and are leading the way to a more relaxed personal point of view of a particular researcher, there are improvements to be made in the way how organized and transparent are the processes of imposing these onto the research community.

## 5. CONCLUSIONS

The analysis showed a clear bias of the researchers towards already established procedures (those were named as necessary to be in ethical assessment), but less prevalent were topics that affect everyday life (work-life balance, agreement with the cultural and social environment about the research topics).

The distribution of institutional orientations showed that the institutions try to address first the situations, connected to ethical aspects, that are also connected to financial and research position benefits, whereas every day research and research work aspects are currently being less considered.

The analysis of forms in which the organizations take ethical issues into account showed that the organizational level of ethical responsibility is higher than the personal one, which could lead to improvements in ethical assessments of particular settings in the future years.

The type of input that the respondents think should be available from their organizations in many situations where ethical assessments might be necessary, shows that the research community requests their research organizations to impose a more transparent rule set in a more organized and systematic way. This showed that even though the organizations do take ethical issues into account, there are improvements to be made in the way how organized and transparent are the processes of imposing these onto the research community.

## 6. REFERENCES

- [1] JRC TTO Circle Publication, 18.8.2019, <https://ec.europa.eu/jrc/communities/en/community/tto-circle-community>.
- [2] JRC TTO Circle Publication, 18.8.2019, <https://ec.europa.eu/jrc/communities/en/community/european-tto-circle/page/tto-circles-members>
- [3] Eric Breit et al.. 2016. Promoting Integrity as an Integral Dimension of Excellence in Research. EU project No. 665926 report.
- [4] Boucher Philip et al.. Ethics dialogues. Experiencing Ethics through 'things'. 2014. ISBN 978-92-79-45049-5
- [5] WIPO. 2017. Successful Technology Licensing. Publication No. 903E 2015 Edition ISBN 978-92-805-2633-2
- [6] Ellen-Marie Forsberg et al. 2018. EUROPEAN CONSENSUS STATEMENT. Sci Eng Ethics, 24:1023–1034. Working with Research Integrity — Guidance for Research Performing Organizations: The Bonn PRINTEGER Statement.

# Automation of Violence and the Disappearance of Moral Responsibility

Assoc. Prof. dr. Tomaž Grušovnik  
University of Primorska, Faculty of Education,  
Cankarjeva 5,  
6000 Koper, Slovenia  
++386 (0)5 663 12 60  
[tomaz.grusovnik@pef.upr.si](mailto:tomaz.grusovnik@pef.upr.si)

## ABSTRACT

Committing violence often results in Perpetration-Induced Traumatic Stress, and thus humans go at length to avoid exposure to behaviors that involve physical force intended to hurt, damage or kill other human beings. Consequently, trends in twentieth-century warfare focused extensively on automation of killing by mechanical distancing of perpetrators from the victims and, more recently, by replacing human soldiers with UAVs (unmanned aerial vehicles) and other military robots. This substitution, of course, had problematic ethical consequences because it increased the propensity to use force. Today, however, we are witnessing an even harsher and more radical process of distancing, namely the introduction of automated decision-making which seemingly removes moral responsibility from human agents. Individual moral responsibility thus starts to disappear in an intricate web of military structures populated with automata, increasingly relegating the decision-making to information-processing algorithms. This relegation of moral responsibility could, in turn, contribute to heightened and irresponsible use of force, whereas the automated information-processing could help with the enforcement of unprecedented conformism.

## Keywords

Violence, perpetration-induced traumatic stress, UAV, moral responsibility, automated decision-making, automatic target recognition

## 1. INTRODUCTION

Contemporary wars are being increasingly waged with machines. Even though the concept of “drone,” for instance, is an old one, the technology became popular only during and after the Balkan war in 1990s and “massively expanded under President Barack Obama” [6]. The use of machines such as UAVs results in soldiers having less and less physical and personal contact with enemy troops on the ground. This “distancing” [5], however, is a key prerequisite for killing and violence. Physical distance between enemy troops as an important “killing-enabling factor” (KEF) can thus be translated into emotional distance, bracketing spontaneous empathy and aversion to killing and resulting in easier decision to commit violence.

If, however, decision-making in the battlefield is itself automated with the help of AI, moral responsibility for casualties could be relegated to non-human actors (such as information processing algorithms), thus further increasing the distance and, consequently, contributing to increased likelihood for conflict.

## 2. HUMAN SPONTANEOUS AVERSION TO VIOLENCE AND THE AUTOMATION OF KILLING

Humans have spontaneous aversion towards aggression and killing of other human beings as well as animals that are closely related to humans (e.g. mammals and even vertebrates such as birds and fish). indeed, in order to commit violence to humans or, for that matter, to animals one has to “psychically numb” (7) or emotionally withdraw” (4) oneself. Despite this withdrawal, Perpetration Induced Traumatic Stress (PITS) is a likely outcome for those who commit violence [2]. Therefore, there is a twofold tendency to a/ conceal violence (removal of cues by, for instance, removing slaughtering from public view [3]) and to b/ further distance the perpetrator from the victim.

### 2.1. Distancing the Perpetrator from the Victim

Distancing the perpetrator from the victim is a fairly old technique of decreasing the so-called “nonfiring rate.” Indeed, as Dave Grossman shows, this rate was very high in traditional battles, including WWII, encompassing from 80 to 85 % of all soldiers, since “only a small percentage of musketeers in a regimental firing line were actually attempting to shoot at the enemy while the rest stood bravely in line firing above the enemies’ heads or did not fire at all.” [5, p. 26] In order to lower it, military tacticians used psychological training and conditioning techniques as well as aforementioned emotional distancing. However, one of the most successful tactics in this case is still pure physical distancing from the ‘target’. Dave Grossman quotes Gwynne Dyer’s observation in order to illustrate this point:

... the intervention of distance and machinery between them [soldiers] and the enemy [is important]; they can pretend they are not killing human beings.

On the whole, however, distance is a sufficient buffer: gunners fire at grid references they cannot see; submarine crews fire torpedoes at “ships” (and not, somehow, at people in the ships); pilots launch their missiles at “targets” [5, p. 108]

One aspect of this distancing is so-called “mechanical distance,” characterized by use of different tools and increasingly UAVs and other military robots. Use of mechanical warfare interfaces in contemporary battlefield thus increases the likelihood of using force by decreasing the nonfiring rate in soldiers: “there has never been any difficulty in getting artillerymen, bomber crews, or naval personnel to kill. [5, p. 107]

Bearing this in mind, we can safely conclude that the future of warfare lies in still greater mechanical distancing of soldiers from enemies, in increased use of UAVs and military robots. The reason for this is not only safety of soldiers but, as shown above, successful lowering of nonfiring rates. Somewhat analogously to war industry one can expect that the slaughtering process in the industrial meat production complex will similarly become more and more automatized, thus lowering the stress and PITS symptoms in slaughterhouse workers while simultaneously lessening the likelihood of health and safety-related incidents at the workplace.

## 2.2. Distancing with the Help of Automated Decision-Making

If distancing from the victim or, for that matter, the enemy (or in some contexts the animal) is a trend in modern warfare and if twentieth and the beginning of twenty-first century witnessed this distancing in the form of mechanical interfaces, then the future development of distancing seems to belong to AI and automated decision-making and profiling. Indeed, UAVs are being increasingly used in different applications, including military domain. [1] Moreover, autonomous UAVs have or will soon have advanced ATR (automatic target recognition) algorithms and will, for instance, be able to execute complex maneuvers in air confrontation. [8] Autonomous UAVs with ATR will thus further unburden the pilots of moral responsibility in combats, thereby making it still easier to use and misuse the lethal technology.

Indeed, it seems that the relegation of responsibility of killing the enemy soldiers to wholly impersonal algorithms represents the last and final step in distancing the perpetrator from the victim. In fact, one could claim that what emerges with this process is a completely new concept of “distance” between the executor and victim, since conceptually the perpetrator and the victim become categorically different entities: the first being an impersonal algorithm and the second a “traditional” human being, a moral agent and patient. The process of distancing the perpetrator from the victim thus finally results in a categorical gap between the executor and the victim.

## 3. MORAL PROBLEMS RELATED TO AUTOMATED DECISION-MAKING

As noted above, categorical distance between the perpetrator and the victim seems to be the last step in the distancing of soldiers from enemies. If first distancing – the distancing with the help of mechanical interface that separated the killer from the victim – was morally problematic, the second distancing – the categorical

gap between the executioner and the victim – seems to be even more so. The reason for this is threefold, as demonstrated below.

### 3.1 The Problem of Relegating Responsibility to Decision-Making Algorithms

Firstly, with the automated decision-making in combat situations the moral responsibility seems to start to dissolve and disintegrate to the point of its disappearance. The notion of moral responsibility is namely bound up with the concept of “moral agent”, i.e. a person who can act morally and is thus also obliged to follow certain moral norms. Algorithms, i.e. decision-making programs, are ontologically speaking not entities in such a way that they could be regarded as “moral agents”. Thus, if an autonomous UAV kills a civilian in a war zone by mistake it cannot, by definition, be morally blamed. Thus the question “Who is responsible for the death of the civilian casualty?” necessarily arises. But the answer to this question is not as straightforward as in more traditional combat situations – are we to blame the programmers from faulty algorithm, or sensor engineers, or the personnel that authorized the use of UAVs in the mission? It seems that moral responsibility in such cases becomes indefinitely relegated to non-specifiable others, thus making it difficult to ensure accountability.

### 3.2 Encouraging the Use of Force

Secondly, the disappearance of moral responsibility and accountability can encourage the use of force in the similar way that more traditional mechanical distancing decreased the nonfiring rate. That is, if the personnel is aware that moral responsibility can be relegated to unspecified or hardly specifiable moral agents, then the staff could be inclined to use the technology more often and in more precarious ways. Indeed, if the moral responsibility becomes lost in complex web of ontologically and categorically different actors (machines, algorithms, humans), then the temptation to abuse the technology with impunity seems to become tangible.

### 3.3 Encouraging Conformism and Uniformity

Thirdly, the use of automated decision-making could induce human subjects to unconsciously conform to the criteria by which algorithms are supposed to derive at decisions. This process might inadvertently significantly limit the freedom of expression of individuals. Similarly to “social credit systems,” ATR algorithms could induce people to act and behave in such a way that makes them less likely to become targets, thereby infringing upon the scope of human expression.

## 4. CONCLUSION

The distancing of perpetrators from the victims as a decisive KEF with the help of mechanical interfaces (long-range artillery, bombers, submarines and more recently UAVs) was predominant in the twentieth century warfare that saw unprecedented number of war casualties. However, new and categorically different form of distancing started to take place in the twenty-first century with the advent of AI and ATR. Increasingly, actors in modern battlefield are algorithms that are ontologically different from human beings and thus cannot be regarded as typical “moral agents”, i.e. subject that have the ability and obligation to act morally. This results in the dissolution of moral responsibility which starts to disappear in a complex web of human and non-

human entities that populate modern warzones. Consequently, the temptation to use and abuse such technology emerges. The most dire scenario that can be imagined for the future warfare is a world full of autonomous military and policing robots that will enforce interests of no particular or traceable origin. In order to avoid that scenario, new and perhaps unprecedented regulations will have to come in power, limiting the potential abuse of military and policing technology. The form and shape of these regulations could be an interesting platform for further research.

## 5. REFERENCES

- [1] Becerra, V. M. 2019. Autonomous Control of Unmanned Aerial Vehicles. *Electronics* 2019, 8, 452. DOI – doi:10.3390/electronics8040452
- [2] Dillard, J. 2008. A Slaughterhouse Nightmare: Psychological Harm Suffered by Slaughterhouse Employees and the Possibility of Redress through Legal Reform. *Georgetown Journal on Poverty Law & Policy*, 15, 2, 391-408.
- [3] Fitzgerald, A. 2010. A Social History of Slaughterhouses: From Inception to Contemporary Implications.” *Human Ecology Review* 17, 1, 58-69
- [4] Fromm, E. 1973. *The Anatomy of Human Destructiveness*. Holt, Rinehart and Winston, New York.
- [5] Grossman, D. 2009. *On Killing: Psychological Cost of Learning to Kill in War and Society*. Little, Brown & Company, New York.
- [6] History of Drone Warfare. The Bureau of Investigative Journalism. Retrieved from: <https://www.thebureauinvestigates.com/explainers/history-of-drone-warfare>
- [7] Joy, M. 2010. *Why We Love Dogs, Eat Pigs and Wear Cows? – An Introduction to Carnism*. Conari Press, San Francisco.
- [8] Zhang, X., Liu, G., Yang, C., Wu, J. 2018. Research on Air Confrontation Maneuver Decision-Making Based on Reinforcement Learning. *Electronics* 7 (11), October 2018, 278. DOI – <https://doi.org/10.3390/electronics7110279>

# Etični vidiki digitalizacije

## Ethical aspects of digitalization

Cene Bavec

Univerza na Primorskem

Titov trg 4, 6000 Koper

cene.bavec@guest.arnes.si

### POVZETEK

Digitalna preobrazba pogloblja že obstoječe etične dileme podedovane še iz industrijske dobe, hkrati prinaša tudi popolnoma nove etične izzive, na katere smo le malo pripravljeni. Z zornega kota normativne in uporabne etike smo osvetlili nekatere izzive za razvijalce in ponudnike, uporabnike ter regulatorje digitalnih storitev. Čeprav imajo vse omenjene skupine resne probleme povezane z etiko in vprašanji, kaj je dobro in kaj slabo za posameznike in družbo, je osrednje sporočilo namenjeno predvsem strokovnjakom s področja digitalizacije. Ob vsem strokovnem in razvojnem delu se morajo zavedati tudi svoje etične odgovornosti, ki je povezana predvsem z zasebnostjo in varnostjo podatkov. V bližnji bodočnosti pa bodo prišla v ospredje popolnoma nove etične zadrege okoli uporabe umetne inteligence in robotike.

### Ključne besede

Digitalizacija, etični izzivi, normativna etika, uporabna etika, zasebnost, digitalna globalizacija

### ABSTRACT

Paper title: Ethical aspects of digitalization

Digital transformation deepens already existing ethical dilemmas inherited from the industrial age. At the same time, it brings also entirely new ethical challenges that we are not well prepared for. From the perspective of normative and applied ethics, we have highlighted some of the challenges for developers and service providers, users and regulators of digital services. Although all of these groups have serious problems related to ethics, and the issues of what is good and what is bad for individuals and society, the central message is primarily intended for experts in the field of digitalization. In addition to their professional and development work, they must also be aware of their ethical responsibility, which is primarily related to privacy and data security. However, ethical concerns about the use of artificial intelligence and robotics are the coming new issues.

### Keywords

Digitization, ethical challenges, normative ethics, applied ethics, privacy, digital globalization

## 1. UVOD

Vsaka človekova dejavnost, ki vpliva na druge posameznike ali družbo, nosi v sebi tudi etične probleme in dileme. Nobenega dvoma ni, da digitalna preobrazba družbe dramatično spreminja vlogo posameznika in samo strukturo družbe ter odnose v njej.

Zato je razmislek o etičnih vidikih digitalizacije pomemben del razmišljanj o naši bodočnosti.

Informatiki in ostali strokovnjaki na področju digitalizacije živijo in delajo v okoljih, kjer so izzivi pretežno tehnične in kognitivne narave, zato pogosto o etičnih izzivih in posledicah svojega dela niti ne razmišljajo. Vendar se stanje izjemno hitro spreminja, saj smo v zadnjih dveh ali treh letih priča resnim razpravam o etičnosti nekaterih informacijskih storitev, posledicah totalnih monopolov tehnoloških multinacionalk, informacijskim zlorabam v političnem okolju in širjenju lažnih novic, da ne omenjamo še drugih tem.

Po drugi strani pa tudi filozofi in teoretiki s področja etike precej počasi dojemajo, kaj vse prinaša digitalizacija in sobivanje realnega in virtualnega sveta (Bavec in Manzin, 2011). Seveda ni vse tako črno, saj je vedno več sociologov in filozofov, ki so se posvetili tem temam. Mnogi povezujejo digitalno preobrazbo z družbenimi tveganji (Beck, 2009, Bauman, 2016), saj je digitalna bodočnost le delno predvidljiva, drugi se naslanjajo na področje uporabne etike (ang. Applied Ethics), ki pa daje precej poenostavljeno sliko celotne problematike (Frey in Wellman, 2008). Mimogrede, uporabna etika je študijsko področje tudi na znanih tehnološko usmerjenih univerzitetnih programih v Evropi in ZDA (npr. University of Oxford, Utrecht University, Norwegian University of Science and Technology ali University of California, Los Angeles - UCLA).

V vmesnem svetu pa živijo managerji in drugi navadni smrtniki, ki v večini primerov pragmatično sprejemajo, kar jim tehnologija ponuja in nimajo niti časa niti motiva razmišljati o etičnosti svojih digitalnih dejanj, pa naj bodo aktivna ali pasivna (Tarafdar, D'Arcy, Turel in Gupta, 2015). Tudi managerji, ki narekujejo digitalizacijo, praviloma enačijo etičnost z zakonitostjo in spodbujajo informacijske storitve, ne glede na morebitno škodo posameznikom (na primer osebni podatki) in družbi kot celoti. Od neetičnosti do nezakonitosti je namreč zelo dolga pot in zakonitost ne more biti alibi za neetičnost.

## 2. ŠIRŠI KONTEKST DIGITALNE ETIKE

Etika je filozofski nauk ali pogled na to, kaj je dobro in kaj slabo za posameznika ali družbo. Kot velja za vsak filozofski nauk, lahko tudi etiko obravnavamo z različnih zornih kotov in različnih predpostavk. Da vsaj malo omejimo področje, ki ga obravnavamo v tem prispevku, vzemimo kot izhodišče Sokrata in njegov pristop k etiki, ki je preprosto povedano naslednji. Ne dopoveduj ljudem, kako naj konkretno živijo, ampak jih spodbujaj in uči, kako naj vrednotijo dobra in slaba dejanja. Da se izognemo nepotrebnim začetnim nesporezumom, pa še dodajmo ugotovitev, da etika ne daje vedno pravih odgovorov na moralna vprašanja.



Z zornega kota digitalne preobrazbe družb, ko smo šele na začetku poti in niti ne vemo točno, kam gremo, je to kar primeren metodološki pristop. Glede na to, da šele spoznavamo družbene posledice digitalizacije, smo očitno še v fazi, ko se srečujemo z novimi etičnimi zadregami, na katere niti nismo pripravljeni (Russo, 2018). Zato so pretežno filozofska vprašanja več kot na mestu.

Sodobne teorije delijo etiko na meta etiko, normativno etiko in uporabno etiko:

1. **Meta-etika** je v bistvu skepticizem in je osredotočena na vprašanja, kako sploh razumemo, kaj je prav in kaj ne (na primer, ali bomo na področju digitalizacije in ožje umetne inteligence sploh lahko zagotovo vedeli, kaj je etično ali neetično). Del meta-etike je tudi kognitivizem, ki predpostavlja, da etični stavki izražajo predloge, ki so torej lahko pravilni ali nepravilni. Anti-kognitivizem pa seveda to srdito zanika.
2. **Normativna etika** se osredotoča na vprašanja, kaj je etično ali moralno, ko se odločamo o določenem dejanju. V bistvu presoja standarde, kaj je prav in kaj ne v konkretnem dejanju. Vprašanja povezana z digitalizacijo se na to področje zelo počasi prebijajo.
3. **Uporabna etika** pa poskuša uporabiti etična načela v realnem življenju in zajema vrsto specifičnih področij, kot so inženirska etika, bio-etika, poslovna etika in seveda *digitalna etika*. Zato bomo v nadaljevanju izhajali predvsem iz stališč uporabne etike.

Velika večina vseh razprav povezanih z etiko in digitalizacijo spada pod uporabno etiko, saj se največ etičnih vprašanj skriva v uporabi informacijskih tehnologij. Vseeno kaže vsaj omeniti tudi veliko širša etična vprašanja, ki spadajo pod meta-etiko in normativno etiko, kot so, na primer:

- Ali veljajo v digitaliziranem in virtualnem svetu enaka etična načela, kot veljajo v realnem svetu? Kaj sploh pomeni virtualni svet z etičnega zornega kota?
- Umetna inteligenca postaja človeku partner pri odločanju, ki pa ima vedno tudi etične vidike. Kje je točka singularnosti (trenutek, ko se bodo stroji lahko s pomočjo umetne inteligence sami učili in razvijali neodvisno od ljudi), ko bo človek prepustil strojem tudi etično presojo (Prinz, 2012)? Kako bomo takrat opredelili etičnost?
- Kaj z etičnega zornega kota pomeni izguba zasebnosti, kot jo poznamo v realnem svetu? Kaj pomeni nezmožnost medmrežja, da pozablja? Ali je to sploh etična dilema?
- Kaj je etičnega in kaj neetičnega v globalizaciji, ki jo spodbuja digitalizacija? Ali je potrebno upoštevati različna etična načela različnih kultur in okolij (van der Velden, 2008)? Ali sploh lahko obstaja globalna digitalna etika (Bao in Xiang, 2006)?
- Ali je programska koda govor? Nenavadno vprašanje, ki ga je sprožil Apple in s katerim se je že ukvarjalo ameriško sodstvo (Henshall, 2018), saj je povezano z ameriško ustavo, ki opredeljuje svobodo govora.

Podobnih načelnih in filozofskih vprašanj je seveda še veliko več, kar daje razmišljanjem o etiki in digitalizaciji poseben raziskovalni čar. Vendar se bomo v nadaljevanju omejili le na nekaj pomembnejših dilem, ki spadajo na področje *uporabne etike*.

### 3. UPORABNA DIGITALNA ETIKA

Digitalna preobrazba družbe ima svoje korenine v industrijski družbi, zato je večino etičnih dilem podedovala in jih ni sama povzročila. Marsikatera od teh podedovanih dilem pa je okrepila. Vendar je ta preobrazba tako radikalna, da je morala prinesiti tudi radikalno nove etične izzive. Nekateri so že vidni, drugi pa nas še čakajo, za večino prihodnih izzivov pa skoraj zagotovo sploh ne vemo.

Etične izzive digitalizacije lahko v grobem razdelimo v tri skupine glede na njihove nosilce, ki se v določenih primerih lahko delno prekrivajo, vendar se v svojem bistvu razlikujejo:

1. Razvijalci in ponudniki digitalnih storitev
2. Uporabniki digitalnih storitev
3. Regulatorji digitalnih storitev

Vsaka od teh skupin ima svojo specifično vlogo v digitalizaciji in posledično drugačne etične dileme in vprašanja. Pomembno pa je, da so med seboj usklajene, saj predstavljajo verigo, katere trdnost določa najšibkejši člen. Neetičnost katere koli od teh skupin se neposredno prenaša na drugi dve in s tem ogroža etičnost celotnega sistema digitalizacije (Floridi, 2019).

Omenimo samo trenutno največji etični izziv digitalne preobrazbe, to je ogrožanje zasebnosti. Del neetičnosti izhaja iz samih ponudnikov storitev, kot sta Facebook ali Google, ki razpolagata z osebnimi podatki in z njimi upravljata kot s tržnim blagom. Nedorazumljivo niso niti vsi uporabniki teh storitev, saj so agresivno trolanje, napadi na zasebnost drugih ali razširjanje lažnih in pogosto tudi nevarnih informacij vsakodneven pojav. Podobno velja tudi za države, ki igrajo vlogo regulatorjev in še zdaleč ne opravljajo svoje funkcije, kot bi si jo želeli. Po drugi strani pa so države tudi »interni ponudniki« digitalnih storitev in imajo resne etične probleme z obdelavo osebnih podatkov.

#### 3.1 Razvijalci in ponudniki digitalnih storitev

Razvijalci in ponudniki globalnih digitalnih storitev imajo veliko več etičnih problemov, kot so pripravljeni priznati. Internet je omogočil velikim multinacionalkam, da so postale glavni nosilci in promotorji digitalne preobrazbe na globalni ravni, kar jim seveda prinaša ogromne dobičke. Zato niso posebej občutljive na etična vprašanja, ki bi te dobičke tako ali drugače omejevala (Bavec, 2012).

Tipičen primer je poslovni model Facebooka, ki bi nas moral resno zaskrbeti. Ni pa edini. Problem je predvsem v tem, da so se večinoma vsi spravili na osebne podatke, saj je milijarda uporabnikov neizčrpen vir podatkov in seveda denarja. Če k temu dodamo še različne zakonske pristope k varovanju podatkov v ZDA in EU, ali solo akcije Kitajske, potem te multinacionalke nimajo posebnih razlogov za dvigovanje svojih etičnih načel, saj jim večina uporabnikov te grehe odpušča. Trenutno večino naporov usmerjajo v lobiranje za svoje interese.

Včasih pa gredo tudi te multinacionalke preko etičnih načel, ko širša javnost tega enostavno ne more sprejeti. Tak primer so bila Googlova očala, ki bi v bistvu ubila koncept zasebnosti. Zgražanje je čutil tudi Facebook, ki je individualne osebne podatke dajal okoli 120 organizacijam in ne samo Cambridge Analitiki. Takih primerov je bilo kar nekaj.

Posebno, vendar izjemno pomembno področje digitaliziranih storitev, ki je polno neetičnih praks, so digitalni mediji, vendar je to tako specifična problematika, da zahteva posebno obravnavo. Še pomembnejše je področje digitalnih storitev v zdravstvu, ki se

sicer naslanja na Hipokratovo etiko, vendar so se razmere v zadnjih 2400 letih spremenile, saj sta vsaj telemedicina in umetna inteligenca v medicinski diagnostiki popolnoma nova izziva tudi za zdravstveni sistem.

Če vse nekoliko poenostavimo, potem je trenutno največji etični izziv razvijalcev in ponudnikov digitalnih storitev, ne samo na globalni ampak tudi lokalni ravni, vezan na dilemo: spoštovati etična načela pri obdelavi osebnih podatkov in s tem bistveno zmanjšati dobiček, ali pa obratno. Trenutno so njihove odločitve skrajno vprašljive z etičnega zornega kota. Najbrž tudi ne bo šlo brez zakonske prisile, ki pa mora biti globalna zaradi globalnega značaja teh storitev.

V kategorijo ponudnikov digitalnih storitev spadajo tudi državne uprave in javni sektor, vključno z zdravstvom. Na primer, v Sloveniji se lahko upravičeno sprašujemo, ali je etično digitalizirati zdravstveni sistem, ne da bi zagotovil res visoko stopnjo varovanja osebnih podatkov. Slaba tolažba je, da je tak pristop za davkoplačevalce bistveno cenejši, čeprav je etično vprašljiv. Podobno etično problematično področje je videonadzor, ki ni več samo v funkciji varovanja določenih objektov, ampak je v funkciji nadzora nad gibanjem posameznikov na širšem območju.

Čeprav je uporaba podatkov v službah državne uprave vsaj v EU bistveno boljše in strožje regulirana, kot je v prej omenjenih multinacionalkah, se bomo morali o etiki digitalizacije državnih uprav še veliko pogovarjati.

### 3.2 Uporabniki digitalnih storitev

Kot smo že omenili, pa tudi vsi uporabniki niso povsem imuni na neetično uporabo digitalnih storitev. Agresivno trolanje in žalitve na raznih forumih so v popolnem nasprotju z etiko, pa naj jo gledamo s katerega koli zornega kota hočemo. Seveda lahko takoj podvomimo tudi v etična načela organizacij (običajno spletnih časopisov ali blogov), ki take komentarje objavljajo na svojih uradnih spletnih straneh. Vsaj načeloma obstajajo rešitve za ta problem, vendar se jih izogibamo, ker obstaja nevarnost, da gremo v drugo neetično in tudi politično skrajnost, to je omejevanje svobode govora.

Z etične plati je pomembno tudi področje računalniških iger, čeprav o tem skoraj nikoli ne govorimo. Računalniške igre so izredno obsežna industrija s stotinami milijoni rednih uporabnikov in brez dvoma spada v okvir digitalizacije. Imajo pa dva etična problema. Po eni strani so neetična dejanja lahko vključena v samo logiko iger, kar posledično lahko spreminja etična načela uporabnikov. Po drugi strani pa je etično vprašljiva izdelava nečesa, kar zasvaja ljudi in jim v bistvu škodi. Čeprav je primerjava nekoliko groba, je vse skupaj podobno tobačni industriji ali pa proizvodnji in prodaji neomejenih količin opioidnih analgetikov. Seveda pa vse igre ne spadajo v to kategorijo, saj je veliko iger izobraževalne narave in imajo lahko izrazito pozitiven učinek, posebej na mlajšo populacijo.

### 3.3 Regulatorji digitalnih storitev

V Evropski uniji polagamo velike upe v zakonske ukrepe in drugo regulativo, s katero bi omejili nezakonito delovanje ponudnikov in tudi uporabnikov digitalnih storitev (Steviss-Gridneff, 2019). Vendar ti ukrepi še zdaleč ne bodo rešili problema etičnosti, saj zakoni niso primerno orožje za boj z neetičnostjo. Spoštovanje etike je zavestna odločitev posameznikov in organizacij vpletenih v digitalno preobrazbo. Zato problema etičnosti v digitalizaciji

zlepa ne bomo dokončno rešili. Lahko bomo veseli vsake »etapne« zmage.

Z zornega kota regulatorjev obstaja še težji problem. Digitalna preobrazba je globalen proces, zato ga lahko na lokalni ali državni ravni rešujemo le v zelo omejenem obsegu. Potrebni so globalni dogovori in konkretni normativni ukrepi (Zhao, 2018). Ampak tudi v tem primeru bodo etični problemi ostali, med drugim tudi zato, ker imajo tudi različne kulture različne poglede na etičnost. In, kot smo že omenili, ali sploh lahko govorimo o sprejemljivi globalni etiki na področju digitalizacije.

## 4. ŠE KRAJŠI RAZMISLEK

V tem prispevku večkrat omenjamo etičnost organizacij, vendar v teh primerih govorimo o deklarirani ali formalni etičnosti, ki jo organizacija pripisuje sama sebi. V praksi pa se večina problemov pokaže na ravni posameznikov, saj odločitve ne sprejema imaginarna organizacija, ampak jo v njenem imenu sprejemajo posamezniki (Yardley, 2018). Na področju digitalizacije so to strokovnjaki s področja informacijskih tehnologij in kognitivnih znanosti ter managerji, ki opredeljujejo cilje organizacije in digitalizacije.

Zato bi se morali tudi strokovnjaki in razvijalci digitalnih storitev, ki lahko vključujejo preproste informacijske sisteme ali kompleksno uporabo umetne inteligence, zavedati pomena etike in tega, kaj je prav in kaj ne. Kakor koli obrnemo, oni izdelujejo orodja za morebitno neetično uporabo. Reakcija enega od bivših razvijalcev pri Facebooku, ki se je na nekem intervjuju prijel za glavo z izjavo: »O bog, kaj smo storili«, je vredna globokega razmisleka. Tudi običajno opravičilo, da je orodje le orodje in da le uporabnik odloča, kako ga uporablja, je precej vprašljivo (Pinterič, 2015, Lomas, 2018). Informacijska tehnologija, za razliko od klasične tehnologije, še zdaleč ni nevtralno orodje. Tudi podatki niso nevtralni.

Trenutno se večina etičnih zadreg vrti okoli osebnih podatkov, vendar hitro prihajajo popolnoma novi izzivi, ki jih spodbuja razvoj umetne inteligence in robotike (Hagerty in Rubinov, 2019), na operativni ravni pa razvoj industrije 4.0. O etičnih izzivih umetne inteligence in raznih algoritmov na tem mestu ne bomo posebej govorili, saj posegajo daleč na filozofsko področje. Kljub trenutno zelo razgretim razpravam, nam še zdaleč ni jasno, kaj pomeni etično sobivanje inteligentnih robotov in ljudi. Še vedno razpravljamo o tem, ali bi roboti potrebovali uradno zabeleženo identiteto, ali bi morali plačevati davke namesto zaposlenih, ki so jih nadomestili, kdo bo odgovoren za morebitne napake ali zlorabe in podobno. Čeprav se nam zdi, da je na akademski ravni zadeva več ali manj jasna, v praksi še zdaleč ni. Paziti namreč moramo, da ne zaidemo v etične pasti ljudizma, ki so zaznamovale začetne odpore proti industrializaciji.

Za konec zapustimo Sokrata in njegovo misel, da ne pridigaj, ampak spodbujaj k razmišljanju in kljub vsemu podajmo nekaj aplikativnih etičnih nasvetov, posebej razvijalcem digitalnih storitev.

Z etičnega zornega kota so tudi tehnološko usmerjeni razvijalci dolžni v svojih rešitvah upoštevati etična načela povezana z zasebnostjo, varnostjo in tudi splošno poštenostjo. Vendar to ni tako preprosto, kot se zdi na prvi pogled. Pogosto je neka aplikacija razvita v popolnoma drugačnem kontekstu, kot je kasneje uporabljena (na primer, video tehnologije v vojaških dronih in podobno). Zato ni slučajno, da tudi tehnološko usmerjeni programi na tujih univerzah poučujejo osnove uporabne

etike v različnih okoljih, vključno z digitalno etiko. Študenti s področja informacijskih tehnologij pri nas takih znanj nimajo in jih na fakultetah o tem tudi ne poučujemo.

ACM (Association for Computing Machinery) ima svoj kodeks etičnega in profesionalnega delovanja (ACM Code of Ethics and Professional Conduct), ki se začneja s stavkom: »Aktivnosti računalniških strokovnjakov spreminjajo svet. Da bi delovali odgovorno, se morajo ozirati na širše učinke svojega dela ...«. Kodeks je izjemno poučen in nazoren, zato bi ga morali poznati tudi naši informatiki. Zanimiva je reakcija enega od bralcev tega kodeksa, ki je izjavil, »da sedaj ve, kaj mora povedati svojemu šefu, če bo od njega zahteval kaj neprimernega« (Pancace, 2018).

## 5. ZAKLJUČEK

V prispevku smo preleteli nekatere etične dileme digitalne preobrazbe. Teh je očitno veliko več in so tudi veliko težje, kot se nam zdi na prvi pogled. Vendar o etiki razpravljamo že tisočletja, pa še vedno nimamo vseh odgovorov in verjetno jih tudi nikoli ne bomo imeli. Skoraj zagotovo to velja tudi za etiko na področju digitalizacije. Kljub temu pa se moramo tudi tehnološko usmerjeni strokovnjaki, ki delamo na področju digitalizacije, zavedati etičnih načel in v skladu z njimi tudi delovati.

Naš osnovni problem je, da z izjemo neodvisnih raziskovalcev na univerzah in inštitutih, delamo po željah naročnikov in jim skoraj podzavestno prepuščamo tudi odločitve o etičnosti njihovih zahtev. Tako, kot velja za zdravnike ali raziskovalce, ki delajo z živimi bitji, bi se morali tudi mi držati etičnih načel, saj tudi mi neposredno vplivamo na človeška življenja in delovanje družbe. Vendar se zgodba šele začneja in zelo verjetno bomo v bodoče morali prilagoditi tudi nekatera že zakoreninjena etična načela novim razmeram, ki jih povzroča digitalizacija. Najbolj očitno bo ponoven razmislek o tem, kaj je zasebnost.

In še osebno priznanje. Ko sedaj razmišljam o etiki in digitalizaciji, včasih smo temu rekli informatizacija, spoznavam, da bi kar nekajkrat moral reči svojim šefom: »Ne, tega pa ne bom naredil na tak način!«

## 6. LITERATURA

- [1] Bao, Z. in Xiang, K. 2006. Digitalization and global ethics. *Ethics and Information Technology*, 81, 41-47.
- [2] Bauman, Z. 2016. *Postmoderna etika*. Filozofska fakulteta, Ljubljana, 2016.
- [3] Bavec, C. 2012. Social Responsibility and Professional Ethics in Management: Some Empirical Evidences at Country Levels. *Managing Global Transitions: International Research Journal*.
- [4] Bavec, C. in Manzin, M. 2011. *Management virtualnih organizacij*. Univerza na Primorskem, Fakulteta za Management.
- [5] Beck, U. 2009. World risk society and manufactured uncertainties. *Iris* Vol. 1, No. 2, pp. 291-299.
- [6] Floridi, L. 2019. Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. *Philosophy & Technology*, 1-9.
- [7] Frey, R. G. in Wellman, C. H. Eds. 2008. *A companion to applied ethics*. John Wiley & Sons.
- [8] Hagerty, A. in Rubinov, I. 2019. Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence.
- [9] Henshall, A., 2018. *What is Digital Ethics? 10 Key Issues Which Will Shape Our Future*. Process.st. <https://www.process.st/digital-ethics/>
- [10] Hongladarom, S. Ed. 2006. *Information Technology Ethics: Cultural Perspectives: Cultural Perspectives*. IGI Global.
- [11] Lomas, N. 2018. *Gartner picks digital ethics and privacy as a strategic trend for 2019*. TC. <https://techcrunch.com/2018/10/16/gartner-picks-digital-ethics-and-privacy-as-a-strategic-trend-for-2019/>
- [12] MIT. 2019. *Asia's AI agenda, The ethics of AIMIT*. Technology Review Insights. 2019.
- [13] Pancake, C.M. 2018. *Programmers need ethics when designing the technologies that influence people's lives*. The Conversation. <http://theconversation.com/programmers-need-ethics-when-designing-the-technologies-that-influence-peoples-lives-100802>
- [14] Pinterič, U. 2015. *Spregledane pasti informacijske družbe*. Fakulteta za organizacijske študije, Novo mesto. 2015
- [15] Prijatelj, V., Hudernik Preskar, A., & Krstov, L. 2010. Pravna in etična vprašanja ob uporabi zdravstvenih storitev na daljavo. *Informatika Medica Slovenica*. p. 26.
- [16] Prinz, J. 2012. Singularity and inevitable doom. *Journal of Consciousness Studies*, 197-8, 77-86.
- [17] Russo, F. 2018. Digital Technologies, Ethical Questions, and the Need of an Informational Framework. *Philosophy & Technology*.
- [18] Stevis-Gridneff, M. 2019. *E.U.'s New Digital Czar: 'Most Powerful Regulator of Big Tech on the Planet'*. The New York Times. <https://www.nytimes.com/2019/09/10/world/europe/margreth-e-vestager-european-union-tech-regulation.html>
- [19] Tarafdar, M., D'Arcy, J., Turel, O. in Gupta, A. 2015. The dark side of information technology. *MIT Sloan Management Review*, 562, 61.
- [20] van der Velden, M. 2008. Invisibility and the Ethics of Digitalization: Designing so as not to Hurt Others. *Information security and ethics: Concepts, methodologies, tools, and applications* pp. 859-871. IGI Global.
- [21] Yardley, D. 2018. *The top five ethical / moral principles for digital transformation*. Consultancy.uk. <https://www.consultancy.uk/news/16602/the-top-five-ethical-moral-principles-for-digital-transformation>
- [22] Zhao, H. 2018. *New regulatory frontiers - opportunities and challenges*. International Telecommunication Union.

# Osnove računalniške etike

## Computer science ethics

Matjaž Gams  
Institut "Jožef Stefan"  
Jamova 39  
1000 Ljubljana  
Slovenija  
[Matjaz.gams@ijs.si](mailto:Matjaz.gams@ijs.si)

### ABSTRACT

In this paper, several basic ethical issues are addressed in particular in relation to computer science and artificial intelligence.

### Keywords

Etika, računalništvo in informatika

### 1. UVOD

Vprašanje etike se zadnje čase upravičeno postavlja čedalje bolj eksplicitno. V uvodu predstavimo dva primera: lažne novice in neetično poročanje o Melaniji Trump.

V zadnjih letih imamo izrazit pojav lažnih novic. Razmere so prišle tako daleč, da se v javnosti pojavljajo trditve, da večina znanstvenih prispevkov ni točnih oz. v ekstremnem primeru, da je bil članek o lažnih novicah v resnici lažna novica (COSMOS, Science of everything, Fake news journal paper revealed as fake news).

V resnici znanost še naprej na večini področij dosega odlične rezultate, čeprav je tudi tu očiten prodor ideologije zlasti v raznovrstne družboslovne vede. V naravoslovju in inženirstvu pa je na osnovi anonimnega recenziranja vsaj v kvalitetnih revijah in na kvalitetnih konferencah praktično nemogoče objaviti ideološko obarvane polresnice ali neresnice. Zato je toliko pomembneje iz prave znanosti in izmed pravih znanstvenikov izločati tiste, ki dajejo ideologiji prednost pred stroko.

Glede medijske trditve, da je bil prispevek v resni reviji o lažnih novicah v resnici lažna novica, pa tole: često citiran prispevek v kvalitetni reviji Nature Human Behavior na temo lažnih novic je bil umaknjen na zahtevo avtorjev [1], ker so našli napako. Torej ne gre za lažno novico, ampak za normalen postopek objav, preverjanja in popravljanja. Prispevek v medijih je torej napačen ali zavajajoč. Zdi se, da javnost ne razume, da raziskovanje temelji na preiskovanju, postavljanju in preverjanju hipotez in da je v resnici tudi »resnica« dinamična, teži pa k ugotavljanju več in več dejstev in zakonitosti.

Zadnjih nekaj let se razkorak med znanjem, spoznanji in javnostjo oz. mediji poslabšuje in to opažamo vsak dan v slovenskih in svetovnih medijih. Preprost primer so prispevki o prvi ameriški dami Melaniji, kako ima narobe oblečen čevlji, kako se je narobe prestopila itd. Pri tem se najprej pojavi prispevek v »main-stream« ameriških medijih, nato se kmalu pojavi še slovenska inačica.

To morda ni najbolj čista oblika lažnih novic, ker gre le za blatenje najbolj znane Slovenke, ki se s politiko sploh ne ukvarja, ampak se ukvarja z vprašanji, kako izboljšati življenje otrok in podobno. Ker pa je poročena s politikom, ki razburja, upravičeno ali ne, množico posameznikov in skupin, je tudi ona stalno kritizirana. Mediji so – tako levi kot desni – žal postali glasila političnih strank in ideologij (politokracija), posledica pa je, da se znanstveno dokazani argumenti čedalje manj cenijo, pa naj gre za cepljenje ali globalno segrevanje. Gotovo ni etično napadati zakonskega partnerja nekoga,

ki ga politično ne mara, je pa poročanje o Melaniji eklatanten primer, kako so se mediji etično spustili na bistveno nižji nivo kot pred leti.

Osnovno poslanstvo avtorja tega prispevka je prinesiti čim več stroke v odločanje in čim manj politike v stroko, predvsem pa povečati pomen raziskovanja, stroke, razvoja in seveda podporo v smislu upoštevanja in financiranja znanosti [2,3]. Slovenija še vedno ostaja na repu Evrope, tretja najslabša po financiranju akademske znanosti. Razkorak med stroko in ideologijami, politiko, lažnimi novicami, se povečuje v škodo stroke, zato je toliko pomembneje, da opozorimo na to stranpot.

### 2. ETIKA

Etika je po slovenski Wikipediji: »filozofski nauk o nravnosti, o dobrem in zlu. Razumemo jo lahko kot filozofsko disciplino, ki raziskuje temeljne kriterije moralnega vrednotenja, pa tudi splošno utemeljitev in izvor morale, je skupek moralnih principov. Po nekaterih filozofih pa je etika filozofska disciplina, ki se ukvarja s tematiko človeškega hotenja in ravnanja z vidika dobrega in zlega, moralnega in nemoralnega.«

Etika je podobno kot socialni odnosi (prijatelji, nasprotniki, nadrejeni ...) opazna pri vseh socialno živeti skupinah živali in ljudi. Npr. če se kakšen osebek iz skupine počuti odrinjenega, zapostavljenega, protestira, ker se početje očitno ne sklada s predstavo o tem, kaj je »fer« oz. »pošteno«. Tako tudi za etiko lahko rečemo, da je sestavljena iz dveh komponent: ena je splošna za vsa živa bitja in druga je specifična, povezana z okoljem, kontekstom, kulturo določene skupnosti.

V tem prispevku etiko obravnavamo skozi koncepte, kot so poštenost, objektivnost, korektnost, koristnost za slovensko, evropsko in človeško civilizacijo [4]. V nadaljevanju sledi analiza vpliva tehnoloških gigantov (kdo vpliva na mnenja ljudi), povezano z vprašanji etike in digitalizacije [5], asilomarskih AI principov, v zaključku pa so povzeta osnovna etična stališča.

### 3. PROBLEM TEHNOLOŠKIH GIGANTOV

V tem prispevku se ukvarjamo z vprašanjem, ali tehnološki giganti vplivajo na ljudske množice, ali pa delujejo objektivno in tehnološko – kar bolj ali manj sami trdijo.

IT tehnološki giganti s praktično monopolnim položajem so recimo Facebook, Google, Apple, Twitter, in Uber, oziroma drugi navajajo: Facebook, Amazon, Microsoft, Google in Apple (FAMGA).

Posamezne afere kot ona z »The Cambridge Analytica«, kjer so bili zbrani podatki za vplivanje na volitve brez privolitve avtorjev informacij, so le vrh ledene gore. Pogosto se govori o problemih le

skozi prizmo enega ali drugega političnega ali ideološkega stališča. V resnici pa praktično vse vpletene ideološke in politične strani nelegalno izkoriščajo informacije in skušajo vplivati, da bi drugo stran čim bolj mogoče izločili iz javnega vpliva, tako da jih obtožujejo sovražnega govora, lažnih novic in podobnega. Pri tem postaja čedalje bolj očitno, da imajo ti tehnični giganti enostavno preveč koncentrirane moči, da nad njimi ni nobenega neodvisnega nadzora in da imajo izredno velik vpliv na ljudske mase. Pojavljajo se masovne histerije, čarovniški pogromi nad drugače mislečimi, grožnje s smrtjo, rasistični in vse mogoči izpadi. Programi AI, ki so začeli recenzirati »objektivno« in izločati npr. posnetke z nasiljem nad živalmi, so sicer naredili velik korak naprej kljub občasnim problemom, recimo izločanjem robotskih bitk ali npr. posnetka, ko bogomolka zagradi in poje kobilico.

Kljub pozitivnemu vplivu AI pri izločanju neprimernih besedil je na osnovi več poročanj v javnosti postalo jasno, da ali vrh ali vodilni zaposleni v IT gigantih tako ali drugače v cenzuro vpeljujejo svoje osebne in politične poglede. Torej smo iz časov npr. komunizma, kjer je komunistična partija nadzirala in usmerjala delovanje in miselni svet ljudi, kot je še danes recimo primer v Severni Koreji, prišli v družbo, kjer možgane perejo vodilni v tehnoloških gigantih. Njihov vpliv je daleč večji in močnejši kot so bili mehanizmi vpliva na možgane v komunističnih režimih stare Jugoslavije ali SSSR, zato imajo predvsem mladi močno oprane možgane. Za preprosto primerjavo se spomnimo, koliko političnih ali proti-komunističnih šal je bilo v stari Jugoslaviji (veliko in so si jih ljudje pripovedovali v ožjih krogih) in koliko je sedaj političnih in proti-ideoloških šal (malo). S stališča šal je sedanje stanje bolj podobno razmeram v komunistični SSSR ali Severni Koreji. Zdi se, da so predvsem mladi izrazito pod vplivom vodilnih ideologij.

Avtor prispevka je poskusno spraševal razne skupine ljudi naslednja vprašanja:

1. Kolikšna je razlika med pokojnino moškega in ženske v Sloveniji?
2. Koliko je razlika med beneficijami ene ure moškega in ženske?
3. Ali so lahko objektivni programi rasistični?
4. Kakšen IQ ima Donald Trump?
5. ... in vprašanja tipa »Ali je cepljenje povezano z avtizmom?«
6. .... in vprašanja tipa »Ali je globalno segrevanje?«

Odgovori na prva štiri vprašanja so izstopali, medtem ko je bila večina odgovorov na vprašanja tipa 5 in 6 bolj ali manj pričakovanih. Večinski odgovori so bili: Ženske imajo občutno manjše pokojnine in plače. Programi so seveda lahko rasistični. Donald Trump ima pod- ali povprečen IQ. Na vprašanja kot 5 so bili odgovori v glavnem skladni in na vprašanja tipa 6 so bili praktično vedno skladni s poznanimi trditvami v literaturi. Za prva 4 vprašanja navajamo pravilne odgovore:

1. Povprečna ženska ima večjo pokojnino kot moški v Sloveniji. Razlika je minimalna.
2. Povprečna letna plača moškega v Sloveniji je nekaj večja kot je povprečna plača ženske, vendar je glede na uro dela praktično enaka in glede na večjo prispevno stopnjo in glede na daljšo življenjsko dobo v resnici ženske dobijo precej več izplena iz ure dela. No, iz te razlike ne bi veljajo delati slona, raje bodimo ponosni, da smo ena najbolj spolno izenačenih združb na svetu. Ko pa pogledamo indeks enakosti, ki je za Slovenijo okoli 0.7, je jasno, da so taki indeksi nerealen odraz stanja.
3. Ali so lahko objektivni programi rasistični?

Program je seveda lahko rasističen, če npr. naredi odločitev na osnovi barve kože. Ko pa program naredi objektivno analizo, npr. statistiko ali strojno učenje, pa je govorjenje o rasističnih programih napad na znanost zaradi ponorele ideologije. V skrajnem primeru je potem tudi  $1 + 1 = 2$  tudi rasistično, saj objektivni programi v resnici le izračunavajo neko, resda bolj zapleteno enačbo kot zgornjo. Seveda je možno tudi objektivne programe spremeniti v subjektivne oz. rasistične preko spreminjanja učnih primerov. Tako lahko neko specifično rasno skupino diskriminiramo preko namerno generiranih rasističnih učnih primerov, česar ne more odkriti praktično nobena inšpekcija. Na podoben način so v Googlu izločali neželene pomembne posameznike – namesto »Janezek Starček« so generirali veliko negativnih učnih primerov tako, da so npr. spustili eno črko - »Jnezek Starček« in program se je naučil obravnavati »Janezek Starček« kot nezaželenega ter bil zato izločen iz socialnih omrežij. Program se je korektno naučil, da je »Janezek Starček« nekaj slabega, vendar so bili učni primeri diskriminatorno zbrani.

#### 4. Kakšen IQ ima Donald Trump?

Praktično vsi ameriški predsedniki so izjemni posamezniki, pametni, sposobni, bogati, pa naj gre za Trumpa ali Obama ali Busha ali Clintona .... Praktično vsi imajo IQ med 120 in 150, kar jih po IQ uvršča med raziskovalce in profesorje. Postati predsednik je pač izredno težko in le izredni posamezniki v tem tudi uspejo. Če imamo ideološko oprane možgane preko IT gigantov in masovnih medijev ter določene posameznike vidimo kot svetnike in druge kot peklenščke, to ne pomeni, da niso vsi po vrsti pametni. Seveda se lahko osebno strinjamo ali ne z določenimi pogledi in aktivnostmi, ampak zato ne smemo zanikati objektivnih števil.

Kako je torej mogoče, da celo na Institutu »Jožef Stefan« večina, zlasti mlajših, ne pozna objektivnih dejstev oz. informacij, ampak ponavlja ideološko obarvane trditve? Ali je odgovor iskati v smeri IT podjetij?

Kako so IT giganti prišli do izjemnega vpliva in moči, je preprosto pojasniti. Vsi po vrsti prinašajo nove storitve, ki jih z veseljem uporabljamo. Pravzaprav smo že postali odvisni od njih, morda celo preveč. Brez njih nam ni živeti. V resnici prinašajo mnogo pozitivnih novosti v naša življenja. Kdor je kdaj poskušal iskati zanimive novice po spletu in si ogledoval npr. YouTube, je pridobil izredno zanimive posnetke in informacije z vseh virov točno s področij, ki zanimajo individualnega gledalca, medtem ko so poročila v medijih veliko bolj nefleksibilna in za širše mase.

Hkrati pa neodvisnemu opazovalcu hitro postane jasno, kako YouTube in podobni perejo možgane. Če recimo začnete gledati posnetke o kolesarstvu, vam bo sistem predlagal čedalje več zanimivih kolesarskih posnetkov. Če začnete gledati levo usmerjene posnetke, vam sistem sam od sebe predlaga čedalje več tovrstnih in kmalu ste mnenja, da imajo vsi levi vse prav, vsi desni pa vse narobe. Ko pa začnete gledati desno usmerjene posnetke, vas zanese desno. Ko gledate CNN, začnete sovražiti desne. Ko gledate FOX News, začnete sovražiti leve. Kako se otresti tega potenciranja? Potrebno je le pazljivo nekaj časa gledati CNN in nekaj časa FOX, da se informacije uravnotežijo in je lažje videti skozi ideološko retoriko in propagandno navijanje. Vsaj tak je recept oz. način ogledovanja informacij avtorja prispevka.

Ni jasno, v koliki meri je pranje možganov in polariziranje množic stranski produkt, vsaj avtorju tega prispevka ne, in v koliki meri gre za namerno forsiranje določenih pogledov in miselnosti. Je pa

nesporno, da čedalje več strokovnjakov opozarja na omenjeni pojav.

Še iz enega razloga se pogled na IT gigante počasi spreminja. Nekaj vodij omenjenih ogromnih podjetij je prišlo v ameriški parlament pojasnjevati delovanje njihovih podjetij. Po koncu njihovih nastopov je bilo celo njihovim zagovornikom jasno, da imajo v rokah izredno močne mehanizme za manipuliranje javnega mnenja, vplivanja na volitve, hote ali nehoti vnašanja svojih osebnih pogledov v cenzuro, hkrati pa ni nobenega neodvisnega organa, ki bi bil sposoben nadzirati te gigantske sisteme. Tako, malo po malem, so omenjeni giganti zapravili zaupanje ne samo intelektualcev in politikov, ampak tudi množic.

En vtis je jasen: kljub občasnemu leporečju o etiki, je delovanje IT gigantov daleč od etično korektnega in še bolj pomembno – bolj ali manj se na etiko požvižgajo kot na nekaj nebodigatreba. Denimo Frank Pasquale, avtor »The Black Box Society: The Secret Algorithms That Control Money and Information« [6] piše, da je postalo jasno, da se ti sistemi ne bodo samo-regulirali, da jih politiki v ZDA ali Evropi ne bodo mogli obvladati, ampak da bodo še naprej manipulirali z uporabniki, jih izrabljali in zavajali.

Med drugim Pasquale trdi: "I don't think tech companies can have these discussions until a regulatory framework forces them to do so. They were warned about the perils of lax application of their own guidelines, and they have ignored or marginalized their critics".

Vtis je, da so – tako kot večina despotov oz. politikov na oblasti – IT giganti opijanjani od svoje moči in pomembnosti, prepričani, da so edinstveni nosilci napredka na svetu (kar po svoje tudi so v tehničnem smislu), da so tudi edini pozvani (od koga?), da usmerjajo ljudske množice v pravo smer, in da so zaradi pomembnosti in sposobnosti kupovanja podpore sposobni na svojo stran pridobiti vsaj politike, če že uporabnikov ne. Ampak njihova »PRAVA SMER« je njihov ideološki pogled in smisel vrednot, ki so daleč od optimalnih, za ljudi in človeško civilizacijo najbolj koristnih. Direktorji in lastniki so večinoma usmerjeni v povečevanje svojega bogastva in moči in imajo prej dokaj čudne poglede kot normalne. Oblast in bogastvo kvarita ljudska srca!

Večina vodstev IT velikanov vidi delovanje demokracije, ljudstev in politike, prav tako tudi etike kot nekaj neumnega, primitivnega, škodljivega. Sebe vidijo kot edine prave svetovne voditelje, kot rešitelje, za katere je etika nepotrebna ovira pri snovanju novih IT rešitev.

Po eni strani so rešitve na dlani: potrebno je vgraditi mehanizme kot so etika, objektivnost, neodvisnost od ideoloških in političnih pritiskov. Potrebno je izločiti vpliv omenjenih posameznikov na delovanje sistema in zagotoviti neodvisen nadzor.

Neodvisni strokovnjaki so glede omenjenega dali vrsto dobrih smernic, taki so recimo »asilomarski principi«, opisani v tem prispevku. V knjigi: Woodrow Hartzog, »Privacy's Blueprint: The Battle to Control the Design of New Technologies« [7], so eksplicitna opozorila, da je preveliko oblast praktično nemogoče nadzorovati in usmerjati. Da je vsaka aplikacija v bistvu tako ali drugače povezana z etiko in vplivom na ljudi. Hartzog pravi: "every aspect of the design of Facebook is bent towards its mission to get you to never stop sharing and to feel good about it in the process". Posledično namenoma oz. posredno nezavedno silijo uporabnike v odvisnost od sistema do nezdravih razmer.

Ena od rešitev je učenje etike [8] med rednim študijem, kot to predlaga ACM [9]. Za umetno inteligenco pa so posebej relevantni t.i. asilomarski principi.

## 4. ASILOMARSKI AI PRINCIPI

Umetna inteligenca je že proizvedla koristna orodja, ki jih vsak dan uporabljajo ljudje po vsem svetu. AI vsak dan naredi 300 trilijonov odločitev. Nadaljnji razvoj AI, primerno etično voden, bo ponujal neverjetne priložnosti za pomoč in krepitev položaja ljudi v desetletjih in stoletjih, ki so pred nami. Ker je AI med najbolj udarnimi tehnološkimi področji, so se tu postavila prva vprašanja, kako usmeriti raziskave in razvoj AI v pravo smer.

Januarja 2017 je bila konferenca BAI 2017, kjer so strokovnjaki umetne inteligence definirali 23 asilomarskih principov o etični uporabi umetne inteligence (AI). Mnenje skupnosti je bilo precej skladno tedaj in je sedaj (op. prevod je prirejen): "Prihaja do velikih sprememb v neznanih časovnih okvirih, vendar v vseh delih družbe, zato imajo odločujoči veliko odgovornost in priložnost, da ga oblikujejo v dobro ljudi in civilizacije."

Prva naloga organizatorjev je bila sestaviti seznam skupnih mnenj o tem, kaj mora družba storiti, da bi v AI v desetletjih najbolj koristila ljudem. S tega seznama so organizatorji, kolikor so lahko, zbrali mnenja v temeljni sklop principov z veliko stopnjo soglasja. Osnova je bila skupno dogovarjanje in ne preglasovanje. O podrobnih besedilih so nato sodelovale manjše skupine, jih natančno opredelile in komentirale. Ta postopek je ustvaril izboljšane različice osnovnih principov. Na koncu so pregledali celoten nabor principov, da so določili raven podpore za vsako različico in končno izbrali najbolj sprejemljive verzije.

Po končanem postopku se obdržali le principe z visokim soglasjem – pogoj je bil, da se je strinjalo vsaj 90% udeležencev. 23 principov je bilo razvrščenih v raziskovalne strategije, pravice do podatkov in prihodnje zadeve, vključno s potencialno superinteligenco, ki so jo podpisali tisti, ki želijo svoje ime povezati s seznamom. Smoter načel je, da bodo podala nekaj smernic, kako lahko moč AI uporabimo za izboljšanje življenja vsakogar v naslednjih letih.

Na spletnih straneh Future of Life Institute [10] lahko dobite izvorno besedilo.

### 4.1 Raziskovalna vprašanja

1) Cilj raziskovanja: Cilj raziskav AI bi moral biti ustvarjanje ne kakršne koli inteligence, ampak koristne inteligence.

2) Financiranje raziskav: Naložbe v AI bi morale spremljati financiranje raziskav o zagotavljanju koristne uporabe, vključno z etičnimi vprašanji iz računalništva, ekonomije, prava in socialnih študij, kot so:

- Kako lahko naredimo prihodnje AI sisteme robustne, da bodo delali, kar hočemo, ne da bi pri tem delali nepravilno ali se zlomili?

- Kako lahko z avtomatizacijo izboljšujemo svojo blaginjo, hkrati pa ohranjamo človekove vire in namen?

- Kako lahko posodobimo naše pravne sisteme, da bi bili bolj pravični in učinkoviti, da bi bili v koraku z AI in obvladovali tveganja, povezana z AI?

- S katerimi vrednotami naj se AI uskladi in s kakšnim pravnim in etičnim statusom?

3) Povezava med znanostjo in politiko: obstajati mora konstruktivna in zdrava izmenjava med raziskovalci AI in oblikovalci politike.

4) Raziskovalna kultura: Med raziskovalci in razvijalci AI je treba spodbujati kulturo sodelovanja, zaupanja in preglednosti.

5) Izogibanje tekmovanja za vsako ceno: Skupine, ki razvijajo sisteme AI, bi morale aktivno sodelovati, ne pa v dirki za prevlado kršiti varnostne standarde.

## 4.2 Etika in vrednote

6) Varnost: AI sistemi morajo biti varni in zaščiteni v celotni življenjski dobi in preverljivi, kadar je to primerno in izvedljivo.

7) Preglednost neuspeha: Če sistem AI povzroči škodo, bi bilo treba ugotoviti, zakaj.

8) Preglednost - pravna: Vsaka udeležba avtonomnega sistema pri odločanju sodišča mora zagotoviti zadovoljivo razlago, ki jo mora pregledati pristojni človeški organ.

9) Odgovornost: Oblikovalci in snovalci naprednih sistemov AI morajo biti odgovorni za moralne posledice njihove uporabe, zlorabe in dejanj, z odgovornostjo in priložnostjo za oblikovanje teh posledic.

10) Prireditev vrednosti: Visoko avtonomni sistemi AI morajo biti zasnovani tako, da se lahko njihovi cilji in vedenja uskladijo s človeškimi vrednotami med njihovim delovanjem.

11) Človeške vrednote: AI bi morali biti zasnovani in upravljani tako, da so združljivi z ideali človekovega dostojanstva, pravic, svoboščin in kulturne raznolikosti.

12) Osebna zasebnost: Ljudje bi morali imeti pravico do dostopa do podatkov, ki jih ustvarijo in nadzirajo AI sistemi, da lahko te podatke analizirajo in uporabljajo.

13) Svoboda in zasebnost: Uporaba AI na osebnih podatkih ne sme neupravičeno omejevati resnične ali zaznane svobode ljudi.

14) Skupne koristi: AI tehnologije bi morale koristiti in opolnomočiti čim več ljudi.

15) Skupna blaginja: Gospodarsko blaginjo, ki jo ustvari AI, je treba širše deliti, da bi koristili celotnemu človeštvu.

16) Človeški nadzor: Ljudje bi morali izbrati, kako in ali bodo delegirali odločitve na sisteme AI, da bi dosegli cilje, ki jih je izbral človek.

17) Ne-subverzija: Moč, ki jo podeljuje nadzor nad zelo naprednimi sistemi AI, bi morala spoštovati in izboljšati družbene in državljske procese, od katerih je odvisno zdravje družbe.

18) Avtonomna orožja AI: Izogibati se je treba tekmi orožja s smrtonosnim avtonomnim orožjem.

## 4.3 Dolgoročne dileme

19) Previdnost glede zmogljivosti: Če ni soglasja, bi se morali izogibati trdnim predpostavkam o zgornjih mejah prihodnjih zmogljivosti AI.

20) Pomembnost: Napredni AI sistemi bi lahko pomenili preveliko spremembo v zgodovini življenja na Zemlji, zato bi jih bilo treba načrtovati in z njimi sorazmerno upravljati.

21) Tveganja: Tveganja, ki jih predstavljajo sistemi AI, zlasti katastrofalna ali eksistenčna tveganja, morajo biti predmet načrtovanja in ukrepov za ublažitev, sorazmernih s pričakovanim učinkom.

22) Rekurzivna samopopolnitev: Za sisteme AI, ki so zasnovani za rekurzivno samopopolnjevanje ali samo ponovitev na način, ki bi lahko privedli do hitrega povečanja kakovosti ali količine, morajo veljati strogi varnostni in nadzorni ukrepi.

23) Skupno dobro: Superinteligenco je treba razvijati samo v službi široko razširjenih etičnih idealov in v korist celotnega človeštva, ne pa ene države ali organizacije.

## 5. DISKUSIJA

V današnjih časih, ko postaja očitno, da so ljudje čedalje bolj zasvojeni in izgubljeni, konfliktni in z opranimi možgani, je potrebno uveljaviti varstvo uporabnikov predvsem s strani IT gigantov in medijev. Pri tem imamo na voljo dobre usmeritve, kot so npr. asilomarski principi.

Pomembno je, da stroka in neodvisni intelektualci opozarjajo na sedanje problematično stanje, ki počasi ogroža celo razvoj zahodne civilizacije. Prav tako je pomembno, da politika sliši glas stroke.

Druga pomembna rešitev je ureditev IT gigantov, kjer ima Evropa škarje in platno v svojih rokah glede varstva podatkov ali zasebnosti, pa vendar se zdi, da ji ne uspeva zaščititi uporabnike. Poglejmo samo izogibanje davkov. Verjetno je precej bolj možna regulacija v ZDA, recimo razbitje podjetij zaradi preveč monopolnega položaja. Seveda pa je take rešitve potrebno previdno uvesti, da ne bomo škodili razvoju najbolj tehnološko naprednih podjetij na svetu, ki so med najbolj zaslužnimi za razvoj človeške civilizacije.

## 6. ZAHVALA

Zahvalil bi se sodelavcem Odseka za inteligentne sisteme na Institutu »Jožef Stefan« za mnogo zanimivih pripomb.

## 7. LITERATURA

- [1] Qiu, X., Diego F., Oliveira, D.F.M., Shirazi, A.S., Flammini, A., Menczer, F. (2019). Retraction Note: Limited individual attention and online virality of low-quality information Nature Human Behaviour, volume 3, page102. Retraction of: Nature Human Behaviour <https://doi.org/10.1038/s41562-017-0132>, published online 26 June 2017.
- [2] Državni svet, posveti, <http://www.ds-rs.si/?q=posveti-konference-predavanja>.
- [3] Gams, M. Državni svet, blog, <https://dis.ijs.si/ds/>
- [4] Nastran, J., Šircelj, B., Bokal, D. Gams, M. (2018). Sensitivity analysis of computational models that dissolve the Fermi paradox. V: LUŠTREK, Mitja (ur.), PILTAVER, Rok (ur.), GAMS, Matjaž (ur.). Slovenska konferenca o umetni inteligenci : zbornik 21. mednarodne multikonference Informacijska družba - IS 2018, 8.-12. oktober 2018, Ljubljana, Slovenija, ISSN 2630-371X.
- [5] Bavec, C. (2012). Social Responsibility and Professional Ethics in Management: Some Empirical Evidences at Country Levels. *Managing Global Transitions: International Research Journal*.
- [6] Pasquale, F. (2016). *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard College.
- [7] Hartzog, W. (2018). *Privacy's Blueprint: The Battle to Control the Design of New Technologies* Hardcover, Harvard College.
- [8] ACM Code of Ethics and Professional Conduct, <https://www.acm.org/code-of-ethics>
- [9] Embedded EthiCS, Integrating Ethics Broadly Across CS Education. (2019), CACM, vol. 26, No. 8.
- [10] Future of Life Institute; Asilomar Principles; <https://futureoflife.org>

# Etična uporaba podatkov družbenih medijev

## Ethical Social Media Data Use

Janez Štebe  
ADP - Arhiv Družboslovnih Podatkov,  
Univerza v Ljubljani, FDV  
Kardeljeva pl. 5  
1000 Ljubljana  
+386 1 5805 292  
Janez.Stebe@FDV.Uni-Lj.Si

Sergeja Masten  
ADP - Arhiv Družboslovnih Podatkov,  
Univerza v Ljubljani, FDV  
Kardeljeva pl. 5  
1000 Ljubljana  
+386 1 5805 277  
Sergeja.Masten@FDV.Uni-Lj.Si

### POVZETEK

Podatki družbenih medijev so zanimivi za raziskovanje na različnih področjih. S tem pa so povezane možnosti zlorab in različne dileme, tako glede varovanja zasebnosti, percepcije javnosti informacij, ter možnosti objavljanja in deljenja podatkov za druge. Priporočila projekta SERISS ponujajo pregled tovrstnih vprašanj in razmisleke o načinih reševanja.

### Ključne besede

Podatki, etika, družbeni mediji, deljenje podatkov, privolitve.

### ABSTRACT

Social media data are gaining attention in various fields of research. The dilemmas about privacy protection, public or private perception of information and options for data sharing and publishing are among the threats. Guidelines from SERISS project gives an overview of questions and offers some considerations about how to find solutions to the dilemmas.

### Keywords

Data, ethics, social media, data sharing, consent.

## 1. UVOD

Podatki družbenih medijev so zelo zanimivi za raziskovanje in lahko nadomeščajo ali pa dopolnjujejo druge vire. Nastajajo sproti in tako vsebujejo neprimerljivo podrobno časovno in zaradi velikih količin tudi prostorsko granularnost, iz njih se da izluščiti tako psihološke osebnostne lastnosti kot politične in druge preference, pa tudi neposredno opazovati obnašanje in navade posameznikov ter skupin [1]. S podatki so povezane tudi različne omejitve, kot so vprašanja samoselekcije in nepokritosti populacije, pristranega prikazovanja sebe, tehničnih omejitev pridobivanja in nestrukturiranosti informacij, zaradi česar je težko oceniti kakovost podatkov [2, 3].

Posebnost podatkov družbenih medijev (DM) je, da so to vsebine, ki jih uporabnik DM dajejo sami na lastno pobudo in neodvisno od morebitne naknadne raziskovalne rabe. Raziskovalec te vrste 'organske' podatke, če gre za velike količine, 'najde' in ne 'generira' oz. ustvarja s pomočjo interakcije z opazovanim subjektom kot na primer v anketah [4]. Zaradi tega se za nazaj v kontekstu raziskovalne rabe 'najdenih' podatkov odprejo pravne in etične dileme glede tega, ali smemo te podatke uporabljati in s kakšnimi omejitvami [5].

Pri tem z rabo podatkov razumemo zbiranje, analizo, spravilo, objavljanje rezultatov in spravilo podatkov ter njihovo deljenje za drugo rabo, se pravi celotni življenjski krog podatkov [6]. Primeri

spornih rab podatkov družbenih medijev kažejo, kako aktualna je ta tema in kako potrebno je jasno postaviti meje dovoljenega. Ogorčenja je bil deležen Facebookov projekt manipulacije z vsebinami za proučevanje privzemanja čustev brez presoje projekta s strani etične komisije [7] in nazadnje razvpiti primer analize milijonov podatkov Facebookovih uporabnikov brez obvestila in privolitve za sodelovanje s strani Cambridg Analytica za ciljano prikrito oglaševanje.

## 2. PRIPOROČILA PROJEKTA SERISS

Znotraj EU H2020 projekta št. 654221 Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS) je bil delovni sklop 6 posvečen novim oblikam podatkov in z njimi povzanimi pravnimi, etičnimi in problemi kakovosti.

Rezultati dela na nalogi 6.1 glede pravnih in etičnih izzivov podatkov DM so poročilo [8], vodič [9] in hitra priporočila [10]. Pri tej nalogi sta kot partnerja projekta sodelovala ESS (European Social Survey) in CESSDA (Consortium of European Social Science Data Archives), kot s slednjo povezana enota pa tudi slovenski ADP (Arhiv družboslovnih podatkov) pri Univerzi v Ljubljani in drugi (NSD – Norwegian Centre for Research Data, UK Data Service in ČSDA – Czech Social Science Data Archive).

V prispevku bomo predstavili nekaj izvlečkov na obravnavano tematiko vezanih priporočil in zaključkov projekta SERISS s povabilom, da se slovenska javnost z njimi podrobneje seznanja.

## 3. ETIČNE DILEME UPORABE PODATKOV DRUŽBENIH MEDIJEV

### 3.1 Privolitev

Seznanjena privolitev je osnovni način, da smemo zbirati osebne podatke. Z nastopom Splošne uredbe o varstvu osebnih podatkov (GDPR) je izpolnjevanje zahtev glede privolitve še strožje določeno kot je bilo do sedaj, saj naj bi bila pridobljena brez pogojevanja z ustreznimi informacijami, podrobna in dokazljiva.

Pri podatkih DM je najprej vprašanje, ali so to sploh osebni podatki. Če gre za javne profile institucij ali organizacij, seveda niso. Tudi objave javnih oseb, kadar nastopajo v svoji javni vlogi, so izvzete iz obveznosti varovanja osebnih podatkov. Drugače pa so podatki, tudi če ne vsebujejo imen ali naslovov, skoraj vedno osebni, saj objavljene slike, psevdonimi ipd. še vedno posredno skozi povezane informacije lahko vodijo do prepoznavnega posameznika. V tem primeru je torej privolitev obvezna. Iz tega sledi tudi, da je podatke SM težko anonimizirati. In nadalje, kadar povezujemo podatke SM s podatki istega posameznika iz anket, se upošteva možnost prepoznavanja prvih tudi slednji šteti kot



osebni podatki, četudi 'pseudoanonimizirani' [11]. Se pravi, da so oboji podvrženi tako jasno izraženi določni privolitvi, kot posebnemu režimu varovanja pred možnostjo zlorabe.

Izjeme, predvidene v GDPR in implementirane v slovenski osnutek Zakona o varstvu osebnih podatkov, so kadar je zaradi velikih količin podatkov ali podobnih razlogov pridobivanje privolitve praktično neizvedljivo. V teh primerih so še vedno potrebni ukrepi, ki npr. med drugim vključujejo etično presojo s strani pooblaščenih organov kot so etične komisije [510] pomena raziskovalnih ciljev projekta, posebnosti in občutljivosti vsebin (politična in druga prepričanja, spolnost, zdravstveno stanje ipd.) in tveganj glede udeležene populacije (manjšine, bolniki, otroci) [10].

Ločujemo tudi pravne in etične vidike privolitve. Tudi če je osnova za uporabo podatkov druga kot pa privolitev pravnega značaja, je npr. če podatki niso občutljivi, niti ne gre za izpostavljeno populacijo, primerna oblika t.i. izločitvena ('opt-out') privolitev etične narave, ko posameznike kontaktiramo z obvestilom o zbiranju podatkov in podatke nato uporabljamo v kolikor nismo dobili odklonitve, kar sicer po GDPR ne bi bilo sprejemljivo [12].

### 3.2 Javna ali zasebna informacija

Poenostavljeno si včasih predstavljamo, da so objave na DM že tako ali tako javne, in jih lahko uporabljamo brez vsakih omejitev [13]. Naravo zasebnosti objav moramo presojati glede na okoliščine, npr. če gre za zaprte forume ali skupine, ali pa občutljivo vsebino [10]. Večina uporabnikov ne pričakuje, da bi lahko njihove objave uporabljali za različne druge analitične namene, pri tem pa so tudi razlike glede uporabe za javne raziskovalne ali pa privatne tržne ipd. namene [14].

### 3.3 Deljenje podatkov

Priprava podatkov za raziskovalni projekt ali pa kot tematsko zbirko zahteva določena znanja, čas in sredstva [15]. Enkrat pripravljene podatki bi bili lahko zanimivi še za koga. Omogočanje dostopa do podatkov in programov za njihovo obdelavo je tudi etična obveznost raziskovalcev, ki na ta način dokazujejo integriteto na podatkih temelječih objav v člankih. Pri tem pa je vprašanje tudi, kje in na kakšen način deliti podatke, po možnosti pri usposobljenih podatkovnih centrih [16].

Poleg izzivov varovanja zasebnosti in vprašanja privolitve za arhiviranje in deljenje podatkov je najbolj aktualno vprašanje, kakšni so pogoji in kaj dovolijo zasebna SM podjetja, ki posedujejo podatke. Rešitev, ki se je poslužujejo npr. v primeru zagotavljanja izpolnjevanja omejitev podjetja Twitter, je deljenje ID-jev posameznih objav skupaj s kodo, ki omogoča ponovno generiranje vsebine [17]. Rešitev ima med drugim to pomanjkljivost, da uporabniki DM posamezne objave umaknejo in tako ni mogoče pridobiti podatkovnega seta istovetnega prvotnemu. Obenem pa je seveda na ta način zagotovljeno spoštovanje volje posameznika pri umiku vsebine.

## 4. ZAKLJUČEK

Vidimo, da so pravne omejitve npr. glede intelektualne lastnine in etična načela raziskovanja lahko v nasprotju. Raziskovalci podatkov družbenih medijev so iz različnih področij, tako tehničnih kot iz družboslovja, humanistike, medicine idr. Z različno usposobljenostjo in priučeno občutljivostjo za etičnost

raziskovanja. Zato je toliko pomembneje vzbuditi pozornost do nekaterih dilem in problemov etične rabe tovrstnih podatkov.

Člane etičnih komisij bi bilo potrebno usposobiti za presojo etičnih dilem pri novih vrstah podatkov. Raziskovalci in uporabniki podatkov pa bi morali z lastno presojo in v primeru dilem, z aktivnim posvetovanjem pri pristojnih organih, razreševati dileme na način, da bi zbužali zaupanje javnosti.

## 5. ZAHVALE

Zahvaljujemo se sodelavcem pri projektu SERISS za prispevke pri posameznih vsebinah, ki smo jih izkoriščali pri pisanju in Evropski komisiji za podporo projektu.

## 6. LITERATURA

- [1] Social Media Research Group. 2016. Using social media for social research: An introduction. Government social research. <https://dera.ioe.ac.uk/26600/>.
- [2] Kleiner, B., A. Stam and N. Pekari 2015. Big data for the social sciences. FORS Working Papers. [https://forscenter.ch/wp-content/uploads/2018/07/fors\\_wps\\_2015-02\\_kleiner.pdf](https://forscenter.ch/wp-content/uploads/2018/07/fors_wps_2015-02_kleiner.pdf)
- [3] Callegaro M., Yang Y. 2018. The Role of Surveys in the Era of "Big Data". In: Vannette D., Krosnick J. (eds) The Palgrave Handbook of Survey Research. Palgrave Macmillan, Cham. [https://doi.org/10.1007/978-3-319-54395-6\\_23](https://doi.org/10.1007/978-3-319-54395-6_23)
- [4] Groves, R.M. 2011. Three Eras of Survey Research, *Public Opinion Quarterly*, Volume 75, Issue 5, Special Issue 2011, Pages 861–871. DOI= <https://doi.org/10.1093/poq/nfr057>.
- [5] OECD. 2016. Research ethics and new forms of data for social and economic research. OECD Science, Technology and Industry Policy Papers, No. 34, OECD Publishing, Paris. <http://dx.doi.org/10.1787/5j1n7vnpxs32-en>.
- [6] CESSDA Training Working Group. 2017 - 2018. *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. Pridobljeno iz <https://www.cessda.eu/DMGuide>.
- [7] Bishop EL. and Gray D. 2017. Ethical challenges of publishing and sharing social media research data. In Woodfield K (ed) *The Ethics of Online Research*. Bingley, UK: Emerald Publishing, pp.159–187
- [8] Hagen, S. Straume, Ø., Bishop, E., Vavra, M., Koščik, M., Štebe, J., Ryan, L., Rød, L.M., Payne, E., Somby, A., L'Hours, H., Emery, T., Cizek, T., Krejci J., Høgetveit Myhren, M., Jessop, C. 2019. Report on legal and ethical framework and strategies related to access, use, re-use, dissemination and preservation of social media data. Deliverable 6.3 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. V pripravi na: <https://seriss.eu/resources/deliverables/>
- [9] –. 2019. Guidelines on the use of social media data in survey research. Deliverable 6.2 of the SERISS project funded under the European Union's Horizon 2020 research and innovation programme GA No: 654221. Dostopno na: <https://seriss.eu/resources/deliverables/>
- [10] –. 2019. Social Media and Research. 10 Legal and Ethical issues to consider. SERISS project funded under the European Union's Horizon 2020 research and innovation

- programme GA No: 654221. [https://seriss.eu/wp-content/uploads/2019/07/Social-media-guidelines\\_quick-ref.pdf](https://seriss.eu/wp-content/uploads/2019/07/Social-media-guidelines_quick-ref.pdf)
- [11] Sloan, L., Jessop, C., Al Baghal, T., & Williams, M. 2019. Linking Survey and Twitter Data: Informed Consent, Disclosure, Security, and Archiving. *Journal of Empirical Research on Human Research Ethics*. DOI: <https://doi.org/10.1177/1556264619853447>
- [12] Social Data Science Lab. 2016. Flowchart for Publishing Twitter Posts. <http://socialdatalab.net/wp-content/uploads/2016/08/EthicsSM-SRA-Workshop.pdf>
- [13] Zimmer, M (2010) But the data is already public: on the ethics of research in Facebook. *Ethics and Information Technology* 12(4): 313–325. DOI: <http://dx.doi.org/10.1007/s10676-010-9227-5>
- [14] Williams ML, Burnap P and Sloan L (2017) Towards an Ethical Framework for Publishing Twitter Data in Social Research: Taking into Account Users' Views, Online Context and Algorithmic Estimation. *Sociology* 51(6): 1149–1168. DOI: <https://doi.org/10.1177/0038038517708140>
- [15] Ljubešić, N., Erjavec, T. and Fišer, D. 2017. Twitter corpus Janes-Tweet 1.0, Slovenian language resource repository CLARIN.SI, <http://hdl.handle.net/11356/1142>
- [16] Weller, K., Kinder-Kurlanda, K. 2015. “Uncovering the Challenges in Collection, Sharing and Documentation: The Hidden Data of Social Media Research?”, <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10657>
- [17] Kinder-Kurlanda, K., Weller, K., Zenk-Möltgen, W., Pfeffer, J., and Morstatter, F. 2017. “Archiving information from geotagged tweets to promote reproducibility and comparability in social media research”, <https://doi.org/10.1177/2053951717736336>

# Etični kodeks strokovnjakov računalništva in informatike

## Code of Ethics for IT Professionals

Marjan Krisper  
Fakulteta za računalništvo in  
informatiko  
Večna pot 113  
1000 Ljubljana  
Slovenija  
marjan.krisper@fri.uni-lj.si

### POVZETEK

Dejanja strokovnjakov za računalništvo in informatiko spreminjajo svet. Da bi ravnali odgovorno, bi morali razmišljati o širših vplivih svojega dela in dosledno podpirati javno dobro. Kodeksi etike in poklicnega ravnanja naj izražajo vest stroke. V tem prispevku bomo obravnavali razvoj kodeksov skozi čas in njihove glavne značilnosti. Hiter razvoj novih tehnologij in njihov prodor v vsa področja človekovega delovanja in infrastrukturne sisteme kot so npr. zdravstvo, promet, energija, itd. zahteva posebno pozornost strokovnjakov, njihovih združenj in vedno bolj tudi uporabnikov teh sistemov.

### ABSTRACT

The actions of computer and information science professionals are changing the world. In order to act responsibly, we should think about the broader impacts of our work and consistently support the public good. Codes of ethics and professional conduct should express the conscience of the profession. In this article we will discuss the evolution of codes over time and their main features. Rapid development of new technologies and their penetration into all areas of human activity and infrastructure systems such as healthcare, transport, energy, etc. requires the special attention of professionals, their associations and, increasingly, users of these systems.

### 1. UVOD

Po drugi svetovni vojni je Norbert Wiener (Wiener 1948) napisal knjigo z naslovom Kibernetika, v kateri je opisal novo vejo uporabne znanosti in opredelil tudi nekatere družbene in etične posledice uporabe računalnikov. Vprašanja, ki jih je opredelil v tej in še dveh naslednjih dveh knjigah so vključevala teme, ki so danes še kako pomembne: računalniki in varnost, računalniki in brezposelnost, odgovornost računalniških strokovnjakov, računalniki za invalide, informacijska omrežja in globalizacija, navidezne skupnosti, delo na daljavo, človeških telesa in stroji, etika robotov, umetna inteligenca, računalniki in religija ter številne druge teme. (Bynum 2000).

Leta 1976, skoraj tri desetletja po izdaji Wienerjeve knjige o kibernetiki, je Walter Maner opazil, da se etična vprašanja in težave, obravnavane v njegovem tečaju Medicinska etika na univerzi Old Dominion, pogosto zapletejo ali bistveno spremenijo, ko vključijo računalnike. Včasih je dodajanje računalnikov, kot se zdi Manerju, dejansko ustvarilo povsem nove težave z etiko, ki ne bi obstajale, če računalnikov ne bi izumili. Zaključil je, da mora obstajati nova veja uporabne etike, podobna že obstoječim področjem, kot sta medicinska etika in poslovna etika. Potem ko je preučil ime "informacijska etika", se je odločil, da bo predlagal novo področje in ga poimenoval "računalniška etika". (V tistem času Maner ni vedel za dela Norberta Wienerja o računalniški etiki). Razvil je eksperimentalni tečaj računalniške etike, namenjen predvsem študentom univerzitetnih programov računalništva. Leta

1978 je napisal "Starter Kit" o poučevanju računalniške etike, ki ga je pripravil za udeležence delavnic, ki jih je vodil, in govorov, ki jih je imel na različnih znanstvenih konferencah v ZDA. Bil je tudi polaga za učne načrte na univerzah. Rezultat tega je, da so se številni znanstveniki, zlasti filozofi in računalničarji, seznanili z računalniško etiko. (Bynum 2000)

## 2. ETIČNI KODEKSI STROKOVNJAKOV RAČUNALNIŠTVA IN INFORMATIKE

V devetdesetih letih so različna mednarodna in nacionalna strokovna združenja kot so napr. ACM (Association for Computing Machinery) in IEEE (Institute of Electrical and Electronics Engineers) sprejela etične kodekse oziroma načela strokovnega ravnanja računalničarjev. ACM-ov etični kodeks je bil sprejet 1992 in v letu 2018 (ACM 2018) dopolnjen z novimi poudarki zaradi tehnoloških sprememb in njihovega prevladujočega vpliva na ljudi, organizacije in družbo.

Slovensko društvo informatika je prvi kodeks, ki je bil narejen na podlagi ACM-ovega, sprejelo leta 1998 in ga dopolnilo v letu 2010 (SDI 2010).

IFIP (International Federation for Information Processing) je ob koncu devetdesetih analiziral kodeks tridesetih organizacij, svojih članic in prišel do splošne ugotovitve da glede na svojo mednarodno in globalno delovanje ne more sprejeti univerzalnega kodeksa, veljavnega za vse članice. Je pa v analizi predstavil nekatere skupne značilnosti tridesetih analiziranih kodeksov (IFIP 1996):

- Spoštovanje

To vključuje spoštovanje interesov ali pravic vpletenih, spoštovanje prestiža poklica, spoštovanje interesov ali pravic javnosti ter spoštovanje blaginje in zdravja javnosti ter kakovosti življenja.

- Osebnostne (ali institucionalne) lastnosti

Sem spadajo vestnost, poštenost, pozitiven odnos, usposobljenost in učinkovitost. V praksi se izraza "vestnost" in "poštenost" pogosto srečujeta pod izrazoma

"sprejemanje odgovornosti" in "integriteta. Druge teme, ki se nanašajo na "vestnost" in "poštenost", so: "strokovnost", "zasluge za delo drugih", "dobra vera ali dobra volja", "skrb za doseganje splošnih ciljev" in podobno. Glede "usposobljenosti" in "učinkovitosti" sta pogosta še dva pojma: "poklicni razvoj in usposabljanje" ali "samooomejitev glede na lastne kompetence

- Zasebnost informacij in celovitost podatkov

"Zaupnost" zahtevajo skoraj vsi splošni kodeksi združenj IFIP. "Zasebnost" in "spoštovanje intelektualne lastnine" se pogosto pojavljata. Druge teme, kot so "računalniški kriminal", "piratstvo ali zloraba informacij", "integriteta podatkov", so manj pogoste.

- Obdelava in dostop do informacij

Večina kodeksov zahteva dostop do informacij vpletenih strank ali ljudi in "obveščanje javnosti". Polovica jih zahteva "celovite informacije." Pri tem velja, da lahko "zasebnost informacij" in "prosti pretok informacij" postaneta nasprotujoča. Oba koncepta je treba uravnovežiti.

- Odnos do regulative

"Regulativa" se v kodeksih ne pojavlja kot glavna tema. Manj od polovice kodeksov zahteva "spoštovanje kodeksa", "spoštovanje zakonov" in "spoštovanje IT in poklicnih standardov."

Še nekatera pomembna etična vprašanja pri uporabi informacijske tehnologije:

- Nepooblaščen nosilci obdelave in hrambe informacij
- Vprašanje lastništva programske opreme
- Preučitev videnja računalnikov kot razmišljajočih strojev in absolutnih ter nezmotljivih ustvarjalcev resnic
- računalniški kriminal in varnost sistemov

- piratstvo ter pravice intelektualne lastnine
- hekanje in virusi
- pomanjkanje zanesljivosti informacijskih sistemov in težave s kakovostjo
- shranjevanje podatkov in zasebnost
- umetna inteligenca in ekspertni sistemi
- informatizacija delovnega mesta.
- Itd.

## 2.1 ACM etični kodeks

Za razumevanje vloge in evolucije etičnih kodeksov je zgovoren primer omenjenega prenovljenega etičnega kodeksa ACM. V preambuli kodeksa je napisano:

Kodeks je zasnovan tako, da navdihuje in vodi etično ravnanje vseh računalniških strokovnjakov, vključno s sedanji in novimi ambicioznimi izvajalci, inštruktorji, študenti, vplivneži in vsemi, ki računalniško tehnologijo uporabljajo. Poleg tega Kodeks služi kot osnova za sanacijske ukrepe v primeru kršitev. Kodeks vključuje načela, oblikovana kot izjave o odgovornosti, ki temeljijo na razumevanju, da je javnemu dobru potrebno posvetiti glavno pozornost. Vsako načelo je dopolnjeno s smernicami, ki ponujajo razlago za pomoč računalniškemu strokovnjaku pri njegovem razumevanju in uporabi. Struktura kodeksa:

### 1. SPLOŠNA ETIČNA NAČELA

- 1.1 Prispevajte k družbi in blaginji ljudi, ob priznanju, da so vsi ljudje udeleženi pri uporabi računalnikov.
- 2 Ne škodujte.
- 1.3 Bodite pošteni in vredni zaupanja.
- 1.4 Bodite pošteni in ukrepajte proti diskriminaciji.
- 1.5 Spoštujte delo, potrebno za ustvarjanje novih idej, izumov in ustvarjalnih del in računalniških sistemov.
- 1.6 Spoštujte zasebnost.

1.7 Spoštujte zaupnost.

### 2. STROKOVNE ODGOVORNOSTI.

- 2.1 Prizadevajte si doseči visoko kakovost tako v procesih kot tudi pri izdelkih strokovnega dela.
- 2.2 Vzdržujte visoke standarde strokovne usposobljenosti, ravnanja in etične prakse.
- 2.3 Poznati in spoštovati obstoječa pravila v zvezi s poklicnim delom.
- 2.4 Sprejmite in zagotovite ustrezen strokovni pregled.
- 2.5 Podajte celovite in temeljite ocene računalniških sistemov in njihovih vplivov, vključno z analizo možnih tveganj.
- 2.6 Delo opravljajte samo na področjih, ki so v vaši pristojnosti.
- 2.7 Spodbujajte ozaveščenost javnosti in razumevanje računalništva, povezanih tehnologij in njihovih posledic.
- 2.8 Do računalniških in komunikacijskih virov dostopajte le, če se to ujema z javnim dobrom.
- 2.9 Oblikujte in izvedite sisteme, ki so zanesljivi in uporabni.

### 3. NAČELA STROKOVNEGA VODENJA.

- 3.1 Zagotovite, da je javno dobro v vseh strokovnih dejavnostih glavna skrb.
- 3.2 Izrazite in spodbujajte, sprejemanje in ocenjevanje izpolnjevanja družbenih odgovornosti članov organizacije ali skupine.
- 3.3 Upravljajte osebje in vire za izboljšanje kakovosti dela.
- 3.4 Izpostavljajte, uporabljajte in podpirajte takšne politike in procese, ki odražajo načela kodeksa.
- 3.5 Ustvarjajte priložnosti, za strokovno rast organizacije.
- 3.6 Bodite previdni pri spreminjanju ali zamenjavi sistemov.

3.7 Prepoznajte in posvetite posebno skrb sistemom, ki se integrirajo v družbeno infrastrukturo .

#### 4. IZPOLNJEVANJE KODEKSA.

4.1 Podpirajte in spodbujajte spoštovanje načel kodeksa.

## 2.2 Analiza sprememb ACM kodeksa

Posodobitve so najbolj vidne v treh dopolnitvah: pomen novih nastajajočih sistemov, diskriminacija in zasebnost podatkov (techrepublic 2018):

- Na prvem mestu so novi nastajajoči sistemi kot posledica penetracije računalniške tehnologije na vsa področja delovanja družbe. V oddelku 3.7 Kodeksa članice ACM poziva, naj posvetijo posebno skrb sistemom, ki se vključijo v družbeno infrastrukturo. Ko nove tehnologije prodirajo na področja prometa, zdravstva, energetike itd, bi morali ustvarjalci teh sistemov delovati tako, da zagotovijo pravičen dostop do sistemov in raven integracije upoštevaje potrebe javnosti.
- Vprašanje diskriminacije je posebej obravnavano. Posebej je izpostavljena odgovornost strokovnjakov, da tehnološke sisteme naredijo čim bolj vključujoče in dostopne vsem ljudem. Diskriminacija na podlagi starosti, barve kože, invalidnosti, narodnosti, družinskega statusa, spolne identitete, članstva v sindikatu, vojaškega statusa, državljanstva, rase, vere ali prepričanja, spola, spolne usmerjenosti ali katerega koli drugega neprimernega dejavnika predstavlja izrecno kršitev kodeksa, Nadlegovanje, vključno s spolnim nadlegovanjem, nasiljem in drugimi zlorabami oblasti in avtoritete, je oblika diskriminacije, ki med drugim

omejuje dostop do virtualnih in fizičnih prostorov, kjer se to dogaja.

- Končno kodeks, v razdelku 1.6., obravnava pojem zasebnosti in varstva podatkov. Tehnologija omogoča zbiranje množice podatkov. Strokovnjaki bi morali biti dobro seznanjeni z vsemi problemi v zvezi z varstvom zasebnosti. Strokovnjaki za računalništvo bi morali osebne podatke uporabljati samo za zakonite namene in brez kršenja pravic posameznikov in skupin. V kodeksu je zapisano, da je treba v sistemih hraniti le najmanjšo še potrebno količino podatkov in da morajo članice ACM delovati, tako da omogočijo preglednost svojih podatkovnih praks.

## 3. ZAKLJUČNA BESEDA

Ob analizi etičnih kodeksov v tem prispevku, lahko občudujemo dalekovidnost Norberta Wienerja, ki je že koncem 40 let prejšnjega stoletja predvidel vpliv rabe računalnikov na ljudi in družbo in celo taksativno naštel področja, ki so z vidika etike aktualna še danes. Lahko pričakujemo spremembe kodeksov tudi v bodoče. Nove tehnologije, v zadnjem času najbolj poudarjene umetne inteligence, prav tako množični podatki, množična robotizacija, družbena omrežja, internet stvari, itd., ob neetični rabi lahko predstavljajo realno grožnjo globalnih razsežnosti. Seveda kodeksi ne bodo dovolj. Potrebna bo tudi ustrezna regulativa, ki je veliko močnejše orožje v boju proti problemom, ki se bodo pojavljali v bodoče, ki pa vedno zamuja.

## 4. LITERATURA IN VIRI

- [1] Bynum, T. W. 2000, A Very Short History of Computer Ethics, Newsletter on Philosophy and Computing, Wiener, N., , Cybernetics: or Control and Communication in the Animal and the Machine, Technology Press. (1948)
- [2] Maner, W., 1996, "Unique Ethical Problems in Information Technology," in Terrell Ward Bynum and Simon Rogerson, eds., Global Information Ethics, Opragen Publications, 1996, pp. 137 – 52, (the April 1996 issue of Science and Engineering Ethics).

- [3] ACM Code of Ethics and Professional Conduct, 2018, <https://www.acm.org/code-of-ethics>
- [4] ACM updates computing ethics code for first time in 26 years, here's what's new, 2018, <https://www.techrepublic.com/article/acm-updates-computing-ethics-code-for-first-time-in-16-years-heres-whats-new/>
- [5] Etični kodeks slovenskega društva informatika, 2010, [www.drustvo-informatika.si > fileadmin > dokumenti > eticni\\_kodeks](http://www.drustvo-informatika.si/fileadmin/dokumenti/eticni_kodeks)
- [6] Codes of Ethics Within IFIP and Other Computer Societies1, 1996, [https://link.springer.com/content/pdf/10.1007/978-0-387-35663-1\\_1.pdf](https://link.springer.com/content/pdf/10.1007/978-0-387-35663-1_1.pdf)
- [7] Pivec, F., 2009, Pogledi na etično obravnavo informacijsko-komunikacijske tehnologije, Uporabna informatika, št. 3.

# Ethics First!

## Ethics guidelines for trustworthy AI

mag. Franci Pivec  
IZUM Maribor (ext)  
++386 (0) 51 320 137  
franci.pivec@ext.izum.si

### ABSTRACT

In the article we recapitulate the content of Ethic Guidelines for Trustworthy AI, the AI HLEG (2019) prepared for the European Commission offered for discussion last spring. In the Commentary we give a selection of views on the role of AI, on ethical dimensions of ICT and especially on the importance of AI ethics. We present critical responds on the document, particularly those from the Special interest group for computer ethics of the IFIP. In the conclusion we claim that the actual anecdotal approach toward the evidence of ethical problems and the offered “check-list” for problem solving needs to be complemented with urgent changes in the common ethics, since it hasn't been able to come up forward with the ethics of ICT / AI. For that reason intellectual strengths of technic and humanistic need to be joined, which has been missed. We suggest trying at home by connecting the Slovenian philosophical society and the Slovenian society Informatika.

/V prispevku smo povzeli vsebino Ethics Guidelines for Trustworthy AI, ki jih je letošnje pomlad ponudila v razpravo Evropska Komisija oz. v njenem imenu AI HLEG (2019). V Komentarju smo navedli izbor pogledov na vlogo UI, na etične razsežnosti IKT in še posebej na utemeljenost etike UI. Predstavili smo kritične odzive na document, posebej iz posebnointeresne skupine za računalniško etiko IFIP. V Sklepu smo ugotovili, da je treba sedanji anekdotični pristop beleženja etičnih problemov in ponujanje ‘ček-liste’ rešitev poglobiti s presojo nujnosti sprememb v obči etiki, ki ne dohaja vplivov IKT oz. UI. Za to pa je potrebno združiti intelektualne moči tehnike in humanistike, kar je doslej manjkalo. Lahko bi poizkusili doma, če bi se povezali Slovensko filozofsko društvo in Slovensko društvi ‘Informatika’./

### Keywords

Ethics Guidelines for Trustworthy AI, computer/information ethics, AI ethics.

### 1. UVOD

Predlog dokumenta je na osnovi številnih predhodnih posvetov z okoli 500 aktivnimi udeleženci pripravila 50-članska High Level Expert Group on Artificial Intelligence (AI HLEG) in ga objavila 8. aprila 2019, z načrtom, da ga bo Evropska Komisija obravnavala in sprejela v prvi polovici leta 2020. Skupina se je oprla na rezultate dela European Group on Ethics in Science and New Technologies (EGE), ki je stalna svetovalna skupina Evropske Komisije. Izhodišče je, da se kritična izguba zaupanja v letalstvo, nuklearno tehnologijo ali varno hrano tiče tudi umetne inteligence in pri tem ne gre zgolj za kakovost tehnologij, ampak za domišljenost kompleksnejših socio-tehničnih sistemov. Zato

morajo o tem razmišljati in ukrepati podjetja, organizacije, raziskovalci, javne službe, vladne agencije, institucije, civilna družba, posamezniki, delavci, potrošniki in drugi.

### 2. POVZETEK “ETIČNIH SMERNIC ZA ZAUPANJA VREDNO UMETNO INTELIGENCO” V EVROPI

Cilj Smernic je promoviranje zaupanja vredne umetne inteligence (ZVUI), ki zajema tri glavne sestavine: (1) biti mora zakonita, t.j. skladna z veljavno zakonodajo in pravili; (2) biti mora etična – privržena etičnim načelom i vrednotam; (3) biti mora zanesljiva, tako v tehničnem kot v socialnem smislu, da ne bi – čeprav z dobrimi nameni – povzročala poškodb. Vsaka od teh sestavin je nepogrešljiva sama zase, vendar nezadostna za doseganje ZVUI. Potrebno je, da so prisotne vse tri v medsebojni harmoničnosti ter se prekrivajo s svojim delovanjem. Če se pojavijo med njimi razhajanja, si mora skupnost prizadevati za njihovo čim večjo usklajenost.

#### 2.1 Etična načela

1. Razvijajte, upravljajte in uporabljajte sisteme umetne inteligence (SUI) na način, ki je skladen z etičnimi načeli spoštovanja človekove avtonomnosti, varovanja pred poškodbami, poštenosti in nedvoumnosti. prepoznavajte in naslavljajte obstoječa in potencialna razhajanja med temi načeli.
2. Bodite posebej pozorni pri vključevanju ranljivejših skupin kot so otroci, osebe s primanjkljaji in težavami ter drugih, ki so zgodovinsko prikrajšani ali so na robu vključenosti, ali pa živijo razmerah, za katere je značilna asimetrija moči ali informiranosti, npr. odnosi med delodajalci in delojemalci, med trgovci in potrošniki ipd.
3. Upoštevajte, da ob velikih koristih za posameznike in družbo, SUI prinašajo tudi določena tveganja in lahko povzročajo negativne vplive, ki jih je težko predvideti, identificirati ali meriti – npr. za demokracijo, vladavino prava, pravično razdelitev blagostanja ali tudi za način človekovega razmišljanja. Sprejmite ustrezne ukrepe za ublažitev takih tveganj, če je to potrebno glede na razsežnost tveganja.

#### 2.2 Priporočila

1. Zagotovite, da bodo razvoj, upravljanje in uporaba SUI skladni z naslednjimi priporočili za ZVUI: (1) človekovo obvladovanje in nadzor; (2) tehnična zanesljivost in varnost; (3) zasebnost in upravljanje podatkov; (4) transparentnost; (5) različnost,



nediskriminatornost in poštenost; (6) okoljsko in societalno blagostanje; (7) odgovornost.

2. Upošteвайте tehnične in ne-tehnične metode za zagotovitev in implementacijo navedenih priporočil.

3. Povečajte število raziskav ter inovacij, ki podpirajo dostopnost SUI in pospešujejo izpolnjevanje priporočil, spodbujajte odpiranje vprašanj v širšem družbenem okolju ter sistematično usposabljanje nove generacije strokovnjakov za etiko umetne inteligence.

4. Komunicirajte na jasen in proaktiven način z deležniki o zmožnostih in omejitvah SUI, nakazujte realistična pričakovanja ter načine implementacije; bodite transparentni glede dejstev, povezanih s SUI.

5. Podpirajte sledljivost in preglednost SUI, še posebej v kritičnih kontekstih in situacijah.

6. Vključite deležnike skozi celoten življenjski cikel SUI; krepite usposabljanje in izobraževanje, da se bodo vsi deležniki zavedli nujnosti ZVUI.

7. Bodite pozorni, da morebiti obstajajo temeljna razhajanja med različnimi načeli in priporočili, zato stalno ugotavljajte, vrednotite, dokumentirajte ter komunicirajte take primere in njihove rešitve.

## 2.3 Ocenjevanje

1. Sprejmite ocenjevalno listo za ZVUI v vseh primerih, ko razvijate, upravljate ali uporabljate SUI ter jo prilagodite specifičnim razmeram uporabe teh sistemov.

2. Upošteвайте, da takšna ocenjevalna lista ni nikoli dokončna, zato je treba zagotoviti kontinuirano identificiranje ter implementiranje priporočil, evalvacijskih postopkov in vključevanje njihovih rezultatov v življenjski cikel umetne inteligence, ob obveznem vključevanju deležnikov v ta proces

## 3. KOMENTAR

### 3.1 UI za ljudi

Japonska vlada je napovedla "Japonsko družbo 5.0", ki bo maksimalno izrabila zmogljivosti IKT za uresničitev "super pametne" kiber-fizične družbe, osredotočene na človeka, kar bodo omogočili socialni roboti, pametni agenti in entitete UI (Gladen, 2019). Danska vlada se je zavezala, da bo državo pripeljala med prvakinje odgovornega razvoja in uporabe UI (Danish Government, 2019). Podobne strateške dokumente sprejemajo tudi druge države, ob tem pa nastajajo tudi civilnodružbeni organizacije, ker si javnost želi jasne slike o razsežnostih in vplivu UI. Tako poslanstvo ima npr. AI4People, ki hoče usmeriti to izjemno silo v dobrobit družbe, slehernika v njej in okolij, ki so nam vzajemna (Floridi in drugi, 2018). Z njeno pomočjo bomo lažje odkrili, kaj lahko postanemo, kaj lahko naredimo, kaj lahko dosežemo in kako lahko sodelujemo med seboj na celi zemeljski obli. Presenetljivo, večina razmišljanja o UI so dejansko razmišljanja o človeku! "UI nas mora opolnomočiti za globlji razmislek o tem, kdo smo in kaj pomeni biti človek" (Blanco-Perez in drugi, 2019). Sploh ni presenetljivo, da se v tej zvezi čuti pozvano tudi katoliška cerkev (Wichmann, 2019).

### 3.2 Ko gre za ljudi, gre za etiko

Vsa znanost se zaklinja na človeka, a je vse pogostejše v zadregi, ko mora opravičevati moralne zdrse: izmišljanje, prikrajanje,

potvarjanje in izkrivljanje podatkov, plagiatorstvo in parazitstvo, zamolčani konflikti interesov, ipd.. (Klampfer, 2018). Ko gre za kibernetiko, jo je njen oče Norbert Wiener že pri njenem poimenovanju neločljivo povezal z etiko in še več: "Veliko pred Nagasakijem in javnimi svarili o atomski bombi, mi je bilo jasno, da imamo v rokah še drug potencial nezaslišanega pomena za dobro ali zlo" (Taddeo, Miller, 2016, 4). Bolj ali manj splošno je znano, da IKT posega v zdravljenje ljudi, v zasebnost, varnost, avtonomijo, identiteto, v zaupanje med ljudmi itd.. Zaradi nje se spreminjajo: narava družbe, kultura, nadzorovanje, odgovornost, legalnost, tveganja, enakost, lastnina, politična moč, ekologija, razmerje med spoloma in vsakodnevno zaznavamo nove in nove posledice (Stahl, 2011). Problem je, ker etične vidike večinoma zaznavamo posteriorno, nujno pa je proaktivno delovanje, česar pa nam stara deontologija v primeru novih tehnologij ne omogoča. Floridi in Sanders (2003) ugotavljata, da je egopoietica posameznikovih kreposti, nastala v Platonovi Republiki, nezadostna in da globalna omreženost zahteva sociopoietico oz. informacijsko etiko, ki pa jo zavira moralni eskapizem.

### 3.3 Etika UI

Alan Turing si je leta 1950 zastavil izrecno vprašanje, ali stroj lahko misli, in presodil, da nanj ni mogoče odgovoriti, ker ni jasno, kaj pomeni "misliti", niti kaj je "stroj". Lahko imitira človeka in na tem temelji t.i. Turingov test, ki je eno od izhodišč za etiko UI. Morda lahko na ta način pojasnimo tudi dogotrajno "zimo UI", ko se od 1960 do 2000 ni kaj dosti dogajalo, saj za to ni bilo potrebe, ker so že začetne rešitve zadoščale za računalniško imitacijo neštetihih nezahtevnih opravil v družbi? Definicij UI je sicer mnogo, ker je tudi v tem primeru nejasno, kaj je "umetnost" in kaj je "inteligenca", se je pa mogoče zadovoljiti s kartko kompilacijo: "UI je raziskovalno področje, katerega cilj je kreiranje artifaktov kot so računalniški programi ali stroji. Smoter večine aplikacij je reševanje problemov s pomočjo algoritmov – to so imitacije ali nadgradnje človeške inteligence. Poleg tega pa je osredotočena še na učenje iz rešitev kompleksnih problemov ter nadgrajevanje funkcionalnosti obstoječih aplikacij" (Richter in drugi, 2019, 146)

UI od samih začetkov sproža resne skrbi glede svoje etičnosti (Marley in drugi, 2019):

- nepopolnost zapopadenja problema (algoritmsko sklepanje je verjetno, a ne izključuje napak);
- nenatančnost vpogledov (algotitemske sisteme je težko kontrolirati, pregledovati, popravljati in so za uporabnike black-box);
- zavajajoči vpogledi (sklepanja so točna v isti meri kot so točni vhodni podatki);
- nepošteni izvidi (sprožajo diskriminatorna ukrepanja, če se odredotočajo le na nekatere družbene skupine);
- transformativni učinki (splošno razširjeni algoritmi profiliranja ogrožajo avtonomijo osenosti in informacijsko zasebnost);
- sledljivost (škodo, ki jo povzročajo algoritmi je težko izslediti, zato je moralna odgovornost zamegljena – kdo odgovarja za nesreče avtonomnih vozil?).

Tudi v primeru UI se zdijo najpriročnejši etični kodeksi ravnanja, ki vsebujejo bolj ali manj konkretne odgovore na zgoraj navedene moralne skrbi. Kašnih 70 jih je na razpolago in med njimi so tudi

tisti, ki jih ponuja industrija: Google, IBM, Microsoft, Intel; nadalje akademske institucije, strokovna združenja (IEEE, ACM.), vlade in končno tudi EU. Ob njih je vse bolj jasno, da vedeti kaj je etična UI, še ne pomeni tudi vedeti, kako narediti etično UI – prepad med načeli in prakso to stalno kaže.

### 3.4 Pripombe k Etičnim smernicam

Na Smernice EU se je takoj odzvala posebna interesna skupina za etiko, ki pri IFIP deluje že pol stoletja, od izida programskega dela Jacquesa Berleurja in Klause Brunnsteina *Etika računalništva* (Berleur, Brunnstein, 1966). Skupno vodi Dona Goterbarn, ki si posebej prizadeva za interdisciplinarni pristop k računalniški etiki, njegov prvi komentar v korespondenci znotraj skupine (e-mail 20/4 2019) pa je bil: “Videti je, da so končno tudi drugi odkrili pomembnost računalniške etike, a na nesrečo je tudi videti, da bodo prezrli vse dosedanje delo in izumili “novo etiko” ter ponovili tudi vse napake iz preteklosti.” Prva napaka je v tem, da spet potrpežljivo nakladajo problem in ponujajo odgovore nanje, kar zožuje etiko na navadno “ček-listo”. Dejansko pa se moramo odločati med mnogimi alternativami in izbira mora biti odvisna od argumentiranih premislekov o vseh danostih in okoliščinah. “Ček-lista” absolutno ne zadošča. ampak je celo nevarna, ker ne nudi okvira za globljo etično analizo.

Julie Cameron, članica skupine, posebej opozarja na nedorečenost “poštenosti” v dokumentu, saj je neenakopravnost različnih skupin državljanov očitna in če se o UI oglašajo vlade, naj predvsem popravijo, kar je njihova napaka. Treba je vzpostaviti pritožbeni sistem, kamor se lahko prizadeti obrnejo zaradi zlorab UI, saj so v sedanjih sodnih mlinih brez upa zmage proti multinacionalkam. Spet se je treba vrniti na zožko opredelitve med “opt-in” in “opt-out”, ko gre za naše osebne podatke. Celotna vlada nas veselo silijo v uporabo internet, tudi ko so druge alternative razumnejše, kar kaže na njihovo podpiranje nepoštenih ravnanj z osebnimi podatki, vključno s problematičnim nadziranjem ljudi. Da naj bi na etične dileme odzivali z DA ali NE, je korak nazaj od že dosežene osveščenosti uporabnikov. In še to, EU bi se morala zavzeti za globalne rešitve in zastaviti soj angažman za opolnomočenje OZN na tem področju.

Luciano Floridi, ki ima eno ključnih vlog pri nastajanju Etičnih smernic, ne zanika odprtih vprašanj o tem, kaj naj bila etično dobra UI v Evropi. Najbolj potrebno se mu zdi razčiščevanje naslednjih moralnih klastrov: etična vprašanja digitalne trgovine, etični problem zavajanja ljudi, etični vidik lobiranja, etičnost dumpinga, etični problem izmikanja odgovornosti. Povsod je prisotna tudi UI in njeni ponudniki se ne morejo sklicevati na to, da so zgolj tehniki-programerji in se jih ne tiče nobena etična grdobija, ki jo (menda nehote) podpirajo.

### 4. SKLEP

Ne navadno je, da se EU najprej loteva etičnosti UI, preden se je opredelila do temeljne ravni etične problematike računalništva oz. informatike. Prevladuje pogled, da klasična etika ponuja povsem zadovoljive paradigme za sedanji proctor in čas, ki pa se je pod vplivom IKT močno spremenil. Če je že treba, naj se pač opredeli obrtniška etika za računalničarje, kakršno so v Aristotelovih časih imeli tesarji ipd.. Vendar to zagotovo ni perspektivna pot, ker UI vpliva v vse smeri in v globino. Anekdotični pristop naštevanja etičnih napak in sestavljanje “ček-liste” njihovih poprav je veliko premalo za odpravo nezaupanja, ki se širi glede uporabe UI. To

bo resna ovira prihodnjemu razvoju in če se EU tega zaveda, naj gre pri iskanju prave rešitve do etičnih korenin. 70 obstoječih kodeksov za UI predstavlja nekonsistentno in konfuzno rastišče vseh mogočih prepričanja in navideznih rešitev, k čemur dodajamo še enega s strani pomembnega predlagatelja, od katerega bi pričakovali, da se bo znal poglobiti v temeljna vprašanja obče etike za današnji svet. Za tak korak je potrebno interdisciplinarno sodelovanje tehnike in humanistike, kar je dolej umanjalo.

Zakaj ne bi česa takega poizkusili doma in za skupno mizo povabili Slovensko filozofsko društvo in Slovensko društvo Informatika?

### 5. REFERENCES

- (1) ACM (2018) *ACM Code of Ethics and Professional Conduct*. New York: ACM
- (2) AI HLEG (2019) *Ethics Guidelines for Trustworthy AI*. Brussels: EC <https://www.politico.eu/wp-content/uploads/2019/04/POLITICO-AI-ethics-guidelines-HLEG-final-April-4.pdf>
- (3) Berleur, J., Brunnstein, K. (1966) *Etics of Computing: Codes, Spaces for Discussion and Law*. London: Chapman&Hall
- (4) Blanco-Perez, C., Perez-Casares, A., Rodriganez-Riesco, R. (2019) Educating for Future: Empowering the Human and Redefining Values and Citizenship in the Age of Technological Disruption. *Cadmus*, 3 (6) 129-147.
- (5) Bughin, J. in drugi (2019) *Notes from the frontier tackling Europe's gap in digital and AI*. McKinsey Global Institute <http://www.mckinsey.com/mgi/publications/multimedia/>
- (6) Dainow, B., O. (2017) Threats to Autonomy from Emerging ICT's. *Australian Journal of Information Systems*, 21, 1-15
- (7) Danish Government (2019) *National Strategy for Artificial Intelligence*. Kobenhavn: Ministry of Finance, Ministry of Industry, Business and Financial Affairs. <http://www.regeringen.dk>
- (8) Floridi, L., Sanders, J. W. (2003) *Internet Ethics: The Constructionist Values of HomoPoieticus*. Research Raport 25. 02. 03, ed, by G. M. Greco, Oxford University <http://web.comlab.ox.uk/oucl/research/areas/ieg>
- (9) Floridi, L. in drugi (2018) AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risk, Principles, and Recommendations. *Minds and Machines*, 26. nov. 2018.
- (10) Floridi, L. (2019a) What the Near Future of Artificial Intelligence Ciul Be. *Philosophy&Technology*, 32, 1-15
- (11) Floridi, L. (2019b) Establishing the rules for building trustworthy AI. *Nature Machine Intelligence*, 7. maj 2019.
- (12) Floridi, L. (2019c) Translating Principles into Practices of Digital Ethics: Five Risk of Being Unethical. *Philosophy&Technology*, 23. maj 2019.
- (13) Gladden, M. (2019) Who Will Be the Members of Society 5.0? Towards an Anthropology of Technologically Posthumanized Future Societies. *Social Science*, 8 (148) 3-39

- (14) Stahl, B., C. (2011) IT for a better future: how to integrate ethics, politics and innovation. *Journal of Information, Communication&Ethics in Society*, 9 (3) 140-156
- (15) Klampfer, F. (2018) Etične zadrege sodobne znanosti. *Dialogi*,10/2018, 30-61
- (16) Morley, J., Floridi, L., Kinsey, L., Elhalal, A. (2019) From What to How: An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices. <http://aexiv.org/ftp/arxiv/papers/1905/1905.06876.pdf>
- (17) Richter, A., Gačić, T., Kölmel, B., Waidelich, L., Glaser, P. (2019) A Review of Fundamentals and Influential Factors of Artificial Intelligence. *International Journal of Computer and Information Technology*,8 (5) 142-156
- (18) Stahl, B., C. in drugi (2014) From computer ethics to responsible research and innovation in ICT. *Information&Management*, 30, 1-9  
<http://dx.doi.org/10.1016/j.im.2014.01.001>
- (19) Tadeo, M., Miller, K. (2016) *Ethics and Information Technologies: History and Themes of Research Field*. V: M. Taddeo and K. Miller (ur.) *The Ethics of Information Technologies*. London: Routledge.
- (20) Wichmann, M. (2019) Kirche digital. *Die Tagespost*, 2. maj 2019.

# Etika naj postane del študija informatike

## Ethics should Become Part of the Informatics Curriculum

Igor Pribac

Filozofska Fakulteta

Aškerčeva 2

1000 Ljubljana

Slovenija

[Igor.Pribac@ff.uni-lj.si](mailto:Igor.Pribac@ff.uni-lj.si)

### POVZETEK

Medicinska izobrazba tipično vključuje seznanjanje študentov z etiko in medicinskimi etičnimi dilemami. Tehniška izobrazba tipično ne vsebuje etičnih vsebin. Trdim, da za takšno razlikovanje ni dobrih razlogov. Etiški premisleki so potrebni, kjerkoli kdo deluje in spreminja dobrobit drugih bitij z etično vrednostjo. Mogočnejše so tehnologije, s katerimi upravljamo, močnejši so razlogi, da se ovemo etičnih implikacij njene uporabe. Informatika je ena najmočnejših tehnologij sodobnega časa, vključitev informacijske etike v študij informatike je zato moralno upravičena zahteva.

**Ključne besede:** inženirska etika, biomedicinska etika, človekove pravice.

### ABSTRACT

Medical education typically require that students of medicine get some knowledge of ethics and ethical challenges specific to medicine. On the other hand, technical education typically does not involve any of this. I argue that there are no good reasons for making such a difference. Ethical reasoning is needed wherever someone acts in a way that change the well-being of others than him bearing ethical value. The mightier the technology we use, the stronger are the reasons to become aware of the ethical implications of its usage. Informatics is arguably one of the mightiest technologies of our time; the inclusion of information ethics in the study of informatics is a morally legitimate request.

**Keywords:** engineering ethics, biomedical ethics, human rights.

#### 1. I.

V enem od treh letošnjih božičnih predavanj Kraljevega inštituta, ki veljajo za najprestižnejša znanstvena predavanja v Veliki Britaniji namenjena širši javnosti, bo letošnja izbranka, da jih izvede, matematičarka Hannah Fry, zagovarjala tezo, da je v študij uporabne matematike in informatike potrebno od njunega začetka vgraditi etiko in da ti poklicni profili potrebujejo dokument, ki bo imel podobno vlogo, kakršno ima v medicini Hipokratova prisega. Prepričana je, da so to področja, na katerih znanstveni in tehnološki razvoj drvi, ne da bi ga spremljalo ustrezno etično premišljevanje znanstvenih dosežkov in njihovih uporab (Sample).

Strinjati se je mogoče, da je tehnološki razvoj čedalje silnejši in da je prav informatizacija in razvoj informacijske družbe eno od njegovih najpomembnejših žarišč, verjetno celo njegovo najpomembnejše (Bynum; Castells, 1996, 1998). Biotehnologije, ki so najresnejši konkurent za ta status, so prav tako v silnem razmahu, a se v genskih posegih močno opirajo na računske sposobnosti računalnikov, medtem ko obratno ne velja (Rifkin, 2001). Mogoče

je zagovarjati tezo, da z razvojem interneta stvari, robotike in širjenjem pametnih strojev informatika postaja to, kar je bila za razvoj prve industrijske revolucije težka industrija – sektor, katerega proizvodi se vgrajujejo v proizvodne postopke v vseh drugih industrijskih sektorjih in v rastoči delež potrošniških blag, zlasti tistih z daljšo življenjsko dobo. Informatika in uporabna matematika ustvarjata algoritme, ki so jedro dodane vrednosti tako rekoč vseh drugih tehničnih naprav, proizvodnih in potrošniških. Na življenje vseh nas čedalje bolj vplivajo algoritmi, ki so jedro pametnih naprav in določajo relevantne okvire naših izbir.

Pisci teh algoritmov, ki torej kreirajo okvire naših izbir, določajo privzete opcije in razvrščajo odstopanja od njih, so pogosto spodbudno nagrajeni mladi inženirji informatike, ki dobijo dokaj določno opredeljene cilje in naloge. V tržni ekonomiji so ti cilji in naloge določeni z logikami, ki nas navsezadnje privedejo do dobičkonosnosti kot poslednjega razloga nastanka celotnega pogona. Pogosto akterjem v proizvodni verigi funkcionalnost in hierarhija ciljev njihovega dela ni znana in inženirji zato ne morejo predvideti in moralno oceniti posledic svojega dela (Castells 1996, 365 in nasl.).

Ta razmah računalniških programov je doslej potekal v normativnem okolju, ki ga je zaznamovala precejšnja odsotnost ali ohlapnost moralnih in pravnih norm in pravil. Zakonodaja je nepresenetljivo zaostajala za realnim tehnološkim razvojem, vendar je svojo mrežo temeljnih vrednot in človekovih pravic, ki se jih je zavezala spoštovati, kljub neprilagojenim področnim zakonodajam vseeno ponujala podlago za razsodbe sodišč. Te so, zbrane, lahko pomemben orientir pri določanju področnih pravil ravnanja in delovanja tako na pravni kakor na etični ravni. Drugi pomemben vir opore so nastajajoči področni in stanovski etički kodeksi računalniške in inženirske etike. Uveljavlja se misel, da so Splošna deklaracija o pravicah človeka in drugi podobni dokumenti skupaj z nacionalnimi ustavami normativni vir, ki določa meje sprejemljivega ravnanja zlasti vsem akterjem javnega sektorja, in da so ti viri referenčni tudi za zasebne akterje. (Pribac, 2012)

Področje, na katerem je vse jasneje čutiti manko, je odsotnost širše, kontinuirane in več nivojsko zastavljene javne razprave o moralnih, pravnih in razvojnih dilemah, ki jih ta skokovit tehnološki preboj IKT sproža. Takšna razprava je v družbi pomembna predvsem zato, ker se skozi lahko izrišejo kritična področja, na katerih bi kazalo poseči in jih družbeno ustrezno pravno in moralno regulirati, in ker se ta področja v takšni razpravi, v katero se pomembno vključujejo tudi področni strokovnjaki, lahko pokažejo jasneje in prej, kakor če bi jih skušali detektirati in preučiti z ad hoc raziskavami. Razprava, ki bi tematizirala moralno in pravno želene in neželene vidike novih tehnologij, je po moji oceni pri nas pičila. Število družboslovcev in humanistov, ki se ukvarjajo s temi vprašanji je skromno in svoje refleksije novih tehnologij običajno prispevajo s precejšnjim časovnim zamikom glede na njihovo vpljavo. Pogosto tudi zato,

ker sami niso posebno večji in zahtevni uporabniki naprednih IKT tehnologij in potrebujejo več časa, da pridejo z njimi v stik in jih doumejo. Moralni premisleki rabe novih tehnologij težko shajajo brez tehnološke ekspertnosti, ki je družboslovci in humanisti ne premorejo. A prav tako res je, da inženirji v splošnem ne premorejo sposobnosti, da bi premišljali moralne in socialne implikacije in konsekvence IKT proizvodov, ki jih načrtujejo. V svojem izobraževanju jih večinoma niso pridobili, ker jih v učnih načrtih večinoma ni. Je tako prav? V nadaljevanju bom razvil splošen argument, da bi seznanitev z osnovami etike in osnovnimi dilemami posamezne področne etike morala vstopiti v učne programe vseh inženirskih poklicev podobno, kot je splošno privzeto, da je področna etika del študijskih programov zdravniških poklicev in tudi izobraževanj za zdravstvene tehnike, sestre in babice. Del tega argumenta je tudi teza, da je to tem bolj res, kolikor moralno pomembnejše so posledice delovanja posameznega poklicnega profila – kar informatike in uporabne matematike postavlja v sam vrh inženirskih strok, ki bi nujno potrebovale etiške vsebine v svojih programih (prim. Noorman, Harris in dr., 6-9; Bynum).

K etični presoji svojega strokovnega početja tehniki doslej niso bili spodbujeni. Krepitev IKT tehnologij z utopističnimi obeti, ki jih je ta tehnologija zbuvala, na pragmatični ravni teh etičskih premislekov dolgo ni potrebovala. Razvoj osebnega računalništva in svetovnega spleta je v osemdesetih in devetdesetih oživel iluzijo zlivanja tehnološkega napredka z moralnim. To iluzijo so spihnila dognanja o kopičenju, analiziranju, profiliranju in posredovanju tretjim osebam osebnih podatkov, najdenih na računih članov in obiskovalcev omrežja. Delodajalci, zavarovalniški agenti in – kot smo lahko spoznali zlasti ob zadnjih volitvah v ZDA – politični strategji se lahko okoristijo z njimi. Vsesplošna dosegljivost, prenosljivost in izsledljivost podatkov je trčila ob varstvo temeljnih svoboščin in pravic, na katerih temelji liberalno demokratični svet, katerega del smo. Naša demokracija je liberalna prav zato, ker kot nezastarljiv namen vsakega političnega delovanja postavlja varstvo temeljnih osebnih svoboščin in človekovih pravic.

## 2. II.

Splošno sprejeto je, da je del medicinske izobrazbe tudi poznavanje etike. Medicinske fakultete v liberalnih demokracijah v svoje študijske programe zato praviloma vključujejo tudi predmet ali modul, ki se ji posveča. Vsaj del statusa, ki ga uživa etika v medicinski stroki, gre pripisati statusu, ki ga uživa medicinska stroka v družbi. Medicina velja za etični poklic, tj. poklic, ki ima etično poslanstvo vdelano v svoje jedro, saj 'pomaga ljudem'. Sama medicina je ta svoj status utrdila, ko je še pred našim štetjem izdelala svoj stanovski etični kodeks (ti. Hipokratova prisega), si torej zadala moralna načela, ki so se uveljavila onkraj posameznih političnih meja v prostoru in času kot nekakšna moralna ustava, ki so jo zdravniki spoštovali ne glede na politično oblast, pod katero so opravljali svoj poklic. Velika večina drugih poklicev si je takšne kodekse omislila šele nedavno. Poleg tega so prav na področju medicine v drugi polovici 20. stoletja začela nastajati nacionalna in nadnacionalna posvetovalna telesa. Ta so skušala osvetljevati etične in družbene dileme, ki spremljajo razvoj tehnologij in praks na medicinskem področju.

Toda ta tradicionalni in še vedno prevladujoči pogled na poseben status medicine med znanostmi, se zamaje, če medicino preučimo primerjalno in se vprašamo, kaj jo razlikuje od drugih znanosti, zlasti od tehniških. Izhajajoč iz razdelitve vseh znanosti in ved na 4 poglobitve skupine (naravoslovje, tehnika, družboslovje, humanistika), je za nekaj znanosti in ved lahko diskutabilno, v

katero skupino jih kaže uvrstiti. A medicine ni med njimi. Razmeroma nedvomno jo lahko uvrstimo med tehniške znanosti, čeprav tradicija takšne uvrstitve ne podpira. Opora za takšno razvrstitev je razlikovanje med naravoslovnimi in tehniškimi vedami. Osnovno razlikovanje med obema skupinama je, da prva vsebuje deskriptivni diskurz o tem, kaj nekaj je, medtem ko so znanosti druge opisni diskurzi o nečem, česar še ni. In dalje: prve načeloma ne posegajo ali čim manj posegajo v to, kar preučujejo (saj bi v nasprotnem primeru njihova lastna dejavnost vstopala v predmet preučevanja, kot vemo od Heisenbergovega načela nedoločljivosti naprej), medtem ko so prizadevanja tehniških ved od samega začetka usmerjena v poseganje v realen svet: tehniška naravnost je naravnost v oblikovanje predlogov za transformativne posege v realni svet. Skupno vsem tem predlogom je, da zase trdijo, da bo svet, če bodo predlogi uresničeni, v nekem pomembnem pogledu boljši, kot bi bil, če posega ne uresničimo, in tudi, da bo boljši tudi v primerjavi z drugimi možnimi posegi.

V soju tako opredeljenega razlikovanja med naravoslovjem in tehniko se jasno pokaže, da ima medicina svoj naravoslovni sestavni del (diagnostika) in svoj tehniški sestavni del (terapijstva in preventiva). Oba sta nepogrešljiva in med seboj ne recipročno povezana: brez točne diagnoze stanja posameznika ali populacije ne more biti izbrana uspešna terapija ali preventiva; sama diagnoza brez terapije ali preventive pa je, četudi dobra, jalova. Cilj medicinskega prizadevanja, ki je maksimiranje zdravja, lahko torej doseže le terapevtski poseg ali preventivni ukrep. In oba sta v svojem formalnem jedru inženirska projekta, podobna projektom, ki jih na drugih področjih oblikujejo domiselni strokovnjaki drugačnih znanstvenih kompetenc.

Izboljšava, ki jo ti predlogi ukrepov ali posegov ponujajo, je lahko različna. Doseženi so lahko novi cilji, že doseženi cilji so lahko s predlaganim ukrepom doseženi lažje, ceneje, hitreje, udobneje itn. Vsi cilji tehniškega prizadevanja so v določenem pogledu vsaj za neko skupino ljudi neka korist oz. prednost. Ta prednost ali korist nekaterih pa je lahko družbeno sporna, saj lahko negativno vpliva na dobrobit druge skupine ljudi, ki lahko zato temu ukrepu ali posegu nasprotuje. Previdnostni premislek, ki cilja le na maksimiranje koristi akterja, ne zagotavlja moralne upravičenosti ravnanja, ki ga vodi.

Konflikt dveh vrednotenj nekega tehničnega posega nas privede do nujnosti moralne (v tem besedilu ne razlikujem med moralo in etiko) presoje ukrepov in posegov. Moralna presoja nečesa je v nečem podobna tehnični, saj presoja glede na neko dobro. Tako kot tehnika tudi etika posega in spreminja realnost. Vendar je dobro, na katerega se sklicuje etika drugačno od tistega, na katerega se sklicuje presoja v tehniki, kolikor je zgolj tehnika. Slednja se bo zadovoljila z ugotovitvijo, da predlagan poseg prinaša določeno korist nekaterim. Morala letvico postavi višje in zahteva, da je načrtovana korist nekaterih ne pomeni neupravičenega poslabšanja stanja drugih. Moralna presoja se razlikuje od presoje koristi akterjev ali posameznih skupin, univerzalna je in varuje interese vseh tistih in vsega tistega, kar ugotovi, da ima moralno vrednost (Singer, 36-38). Ta zahtevnejša presoja je seveda nujno omejujoča in selektivna glede na presojo tehnično izvedljivega in koristnega za nekatere.

Ker je vsak odgovoren za posledice svojih dejanj, nosijo tudi inženirji del odgovornosti za moralne vsebine naprav in posegov, ki jih načrtujejo ali z njimi upravljajo (Harris, 91-110). Poznavanje teh vsebin je nujni pogoj sposobnosti moralne presoje teh tehnologij in njihove rabe, njihovo nepoznavanje pa je odločen korak k opustitvi vsakega etičnega pomisleka povezanega z njimi. Poznavanje osnov etičnih teorij, praktičnega silogizma in moralne

argumentacije v kontekstu priznavanja človekovih pravic kot moralnega temelja sodobnih družb je zato skupaj z razvojem sposobnosti prepoznavanja moralno relevantnih lastnosti načrtovanih ali obstoječih posegov in naprav (ti. *disclosive ethics*) nujni pogoj dejavnega prevzemanja moralne odgovornosti inženirjev na njihovih področjih delovanja. S tem bi tudi ti izobrazbeni profili lahko obveljali kot zaznamovani z moralnostjo, nič manj kakor medicina. Primarni cilj vpeljevanja etičkih vsebin v tehniške študije tako ni dogmatičen, ampak je glede odgovorov odprt. Prizadeva si pri študentih okrepiti sposobnost zaznavanja moralnih vprašanj in jih opremiti za njihovo reševanje. Začne se z ugotovitvijo, da etika ni del nobene tehniške stroke, da je iz znanosti neizpeljiva in nanjo nezvedljiva in da mora zato svoja oporišča oblikovati sama (Singer, 21-34; Johnson, 609).

Poleg tega moralnega argumenta za uvedbo etike v poklicno izobraževanje informatikov in drugih inženirjev, je mogoče oblikovati še previdnostnega, ki se ozira le po koristih za akterja samega (Karoof). Mogoče je trditi, da v svetu, ki ga oblikuje čedalje gostejša in vplivnejša mreža pravnih in moralnih, formalnih in neformalnih pravil delovanja, v katerem je beseda tudi o etični zasnovi izdelkov (*ethical design*), sposobnost ravnanja v skladu z njihovim duhom pomeni primerjalno prednost za akterja. In obratno: njihovo nepoznavanje lahko velja za primerjalni zaostanek. Poznavanje področne etike in sposobnost razreševanja etičnih izzivov ima poleg intrinzične torej tudi instrumentalno vrednost.

### 3. LITERATURA

- [1] Bynum, Terrell, "Computer and Information Ethics", *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2018/entries/ethics-computer/>
- [2] Castells, Manuel (1996): *The Rise of the Network Society: Economy, Society and Culture 1.* knjiga, Blackwell.
- [3] Castells, Manuel (1998): *End of Millenium, The Information Age: Economy Society and Culture 3.* knjiga, Blackwell.
- [4] Harris, E. Charles; Pritchard, Michael S.; James, Ray; Englehardt, Elaine (2015): *The Engineering Ethics: Concepts and Cases*, Wadsworth.
- [5] Johnson, Deborah G. (2003): »Computer Ethics«, v: *A Companion to Applied Ethics* (ur. Frey, G. R. in Christopher Heath Wellman), Blackwell, str. 608-619.
- [6] Karoof, Paul (2019): »Embedding Ethics in Computer Science Curriculum«, *Harvard Gazette* (25. 1. 2019)
- [7] <https://news.harvard.edu/gazette/story/2019/01/harvard-works-to-embed-ethics-in-computer-science-curriculum/>
- [8] Noorman, Merel, "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/spr2018/entries/computing-responsibility/>
- [9] Pribac, Igor (2012): »Javna etika«, v: *Javna etika in integriteta: odgovornost za skupne vrednote: integriteta, odgovornost, pravna država* (ur. Kečanovič, Bećir), Komisija za preprečevanje korupcije, str. 141-150.
- [10] Rifkin, Jeremy (2001): *Stoletje biotehnologije*, Založba Krtina.
- [11] Sample, Ian (2019): »Maths and tech specialists need Hippocratic Oath, says Academic« (Guardian, 16. 8. 2019)
- [12] <https://www.theguardian.com/science/2019/aug/16/mathematicians-need-doctor-style-hippocratic-oath-says-academic-hannah-fry>
- [13] Singer, Peter (2008): *Praktična etika*, Založba Krtina.



## Indeks avtorjev / Author index

Bavec Cene.....	15
Gams Matjaž .....	19
Grušovnik Tomaž.....	12
Krisper Marjan .....	26
Pivec Franci.....	31
Pribac Igor.....	35
Sergeja Masten.....	23
Štebe Janez.....	23
Stres Špela.....	5, 9





IS  
20  
19

Konferenca / Conference

Uredili / Edited by

**Etika in stroka /  
Professional Ethics**

Marjan Krisper, Franci Pivec, Matjaž Gams