Research ethics guidelines for the computer and information sciences

Philip Brey

University of Twente
The Netherlands

Situation in computer science until recently

Variety of codes of ethics (professional codes for computer scientists)

- ACM Code of Ethics and Professional Conduct
- IEEE Code of Ethics
- Specialized codes for subfields

No published research ethics guidelines !!!

Until about 2015, few RECs for computer science

SATORI and CEN CWA standards for research ethics

CEN CWA 17145-1 (2017)

Ethics assessment for research and innovation - Part 1: Ethics committee

With procedural standards for research ethics committees and with research ethics guidelines for various fields

Including ethics guidelines for the

computer and information sciences

cen

EUROPEAN COMMITTEE
FOR STANDARDIZATION

SATORI ethical principles for all fields

General ethical principles for all fields of research and innovation

1. Research integrity
2. Social responsibility
3. Protection of human research participants
4. Protection of animals used in research
5. Protection and management of data and dissemination of research results
6. Protection of researchers and the research environment
7. Avoidance of and openness about potential conflicts of interest

Differences in ethical principles between fields

Medical sciences: Focus on the *relationship between medical researcher and human subject*. Medical ethical principles such as autonomy, informed consent, beneficence, nonmaleficense, human dignity, and justice.

Life sciences: Focus on *relationship of researchers to living biological systems, ecosystems and the environment*. Ethical principles include animal welfare, ecosystems integrity, sustainability, health and environmental risks, naturalness and playing God.

Natural sciences: Focus on *truth and accuracy.* Principles: scientific integrity, data integrity, freedom from bias, and honesty

Differences in ethical principles between fields

Engineering sciences: Focuses on the *technological intervention into society.* Ethical principles include social responsibility, well-being, impacts on rights, the precautionary principle, sustainability, and the good of society.

Social sciences: Focuses on the *relation between the researcher and human beings and social structures.* Principles: informed consent, equality, anonymity, confidentiality, privacy, fairness, nondiscrimination, human rights, avoidance of cultural and social bias, and respect. Also: data integrity, research integrity, objectivity

Humanities: Focuses on *the study of human culture and the human condition*. Ethical issues include concern interpretation and construction of narratives, the role art in society, and our responsibilities in the preservation of cultural heritage. Also: principles and issues of social sciences.

Differences in ethical principles between fields

Computer and information sciences: Focus on the *processing, storage and dissemination of information.*

Ethical principles: information privacy, security, intellectual property, censorship, freedom of information, and fairness & non-discrimination

Research ethics guidelines for computer science

- *Seven sets of general guidelines*
  Most importantly: guidelines for protection of human participants and for social responsibility

- *Nine sets of specific guidelines*

1. Protection of privacy and personal data

- (Summary) General ethical guidelines for the use of personal information in research (applies to all fields)

- (Summary) Ethical guidelines specifically for computer science research, including guidelines for data preparation, data mining, and the development of techniques for surveillance and human subject monitoring.

2. Avoidance of security risks

Six guidelines. Selection:

- Ensure that new research concepts and innovations offer reasonable protection against any potential unauthorized disclosure, manipulation or deletion of information and against potential denial of service attacks, e.g. protection against hacking, cracking, cyber vandalism, software piracy, computer fraud, ransom attacks, disruption of service

- Do not engage in research that involves attempts to make unauthorized access to telephone systems, computer networks, databases or other forms of ICT; such research is illegal and unethical, regardless of motivation

- Treat with extreme caution the dissemination of research involving the identification of undiscovered security weaknesses in existing systems

3. Respect for freedom of expression

- Ensure that new research concepts and innovations do not pose unjustified inherent risks to the freedom of individuals to express themselves through the publication and dissemination of information, or to their freedom of access to information

- If research or innovation involves the use of censorship methods, strike an appropriate balance between the need for content control and the right of individuals to express themselves freely.

4. Respect for intellectual property

- Ensure that new research concepts and innovations do not pose unjustified inherent risks to the intellectual property rights of individuals or organisations.

- Avoid research that could generate copyright issues, such as research involving peer-to-peer networking or file sharing and distribution.

5. Respect for other individual rights and liberties

- (Summary) Ensure that new research concepts and innovations do not pose inherent risks to autonomy, authenticity or identity and do not take away control of users.

- (Summary) Ensure that decisions made by information systems take into account the rights, values and interests of stakeholders, and make them transparent and hence accountable.

- Take into account how responsibilities and liabilities are assigned between humans and machines when information systems are involved in decision-making.

6. Avoidance of harms to justice and equality

- Consider how new research concepts and innovations could widen or narrow social inequalities in terms of the distribution of opportunities, powers and capabilities, civil and political rights, economic resources, income, risks or hazards

- Consider how new research concepts and innovations could harbour or counter unjust bias in terms of age, gender, sexual orientation, social class, race, ethnicity, religion or disability

- Consider how new research concepts and innovations could harm or promote the interests of vulnerable, disadvantaged, or underrepresented groups and communities in society, including those in low income and lower-middle income countries.

7. Promotion of well-being and the common good

- Consider how the research or innovation activity could harm or promote the general well-being of individuals and groups in society (e.g. effects on the quality of work or quality of life)

- Consider how the research or innovation activity could harm or promote the social skills and behaviour of individuals, and how it could harm or promote the learning or exercising of important virtues, such as patience and empathy

- Consider whether and how the research or innovation activity could harm or promote important social institutions and structures, democracy, and important aspects of culture and cultural diversity.

8. Promotion of environmental sustainability

- Optimize technologies for effective and cost-efficient resource use (including raw materials and energy), for resource recovery (recycling), and for lowering the production of environmentally harmful wastes and environmental pollution.

9. Dual use of computer science R&I

- Consider whether new research concepts and innovations could have military applications

- Consider whether new research concepts and innovations could contribute to the proliferation of weapons of mass destruction

- Consult proper authorities before publishing and adhere to relevant national and supra-national regulations if a technology has significant military applications or if it contributes significantly to the proliferation of weapons of mass destruction. Even if publication is allowed, find a proper balance between security and freedom of publication.

Recent ethics guidelines for AI

- **High-Level Expert Group on AI (HLEG)** of the European Commission: The Ethics Guidelines for Trustworthy Artificial Intelligence (AI)

- **IEEE** (Institute of Electrical and Electronics Engineers): Ethically Aligned Design

Recent ethics guidelines for AI

- **OECD** (Organisation for Economic Cooperation and Development): OECD Principles on AI

- **UNESCO**: First draft of the Recommendation on the Ethics of Artificial Intelligence

EU Ethical AI: Seven requirements

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and data governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Environmental and societal well-being
7. Accountability
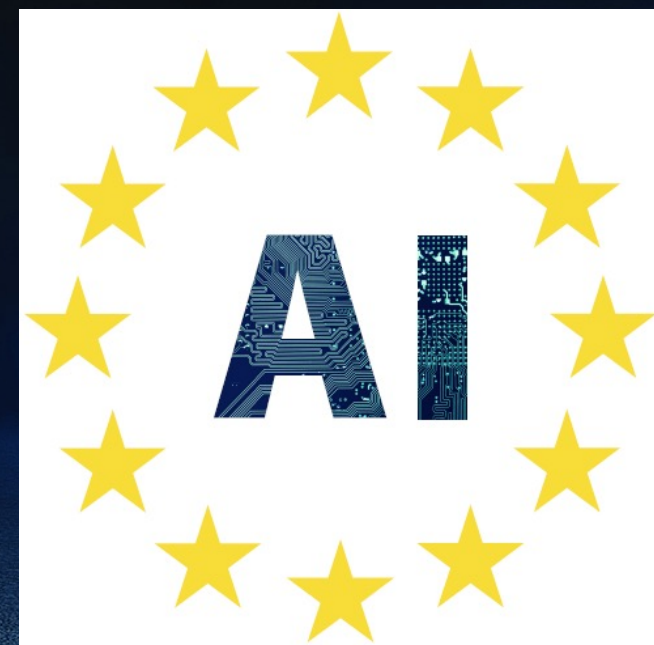
Relation between computer ethics guidelines and AI guidelines

- Significant overlap in principles and guidelines

- Computer ethics guidelines are broader and more complete

- AI guidelines are focused on AI-specific issues such as control, transparency, delegation of responsibility and decisionmaking

Recommendations

- The computer science ethics guidelines we developed in SATORI and in a CEN CWA in 2017 are still current

- AI ethics guidelines should not be used for computer science as a whole

- AI ethics guidelines can be used as stand-alone guidelines or incorporated into broader computer ethics guidelines

- satoriproject.eu (deliverable D4.7)

- Brey, P. (2022). Research Ethics Guidelines for the Engineering Sciences and Computer and Information Sciences. In Laas, Davis and Hildt (eds.), *Codes of Ethics and Ethical Guidelines*, Springer.

THANK YOU !