COM6655 Professional Issues
Autumn 2021-22

Professional Ethics for Information and Computer Scientists (Part 2)

Dr Maria-Cruz Villa-Uriol
Department of Computer Science, University of Sheffield
m.villa-uriol@sheffield.ac.uk

Content

Usefulness of the code

- Neither the BCS nor the other codes have succeeded in becoming a central point of reference for computer professionals.
- **Q. Why not?**

Limitation to a specific group of professionals

- In general these codes make clear that they apply only to their members.
- So, either:
  - Only members of the these organisations are considered to be computing professionals; or else
  - Other professionals can practice with a different ethical code.

Lack of attention to the rights of professionals

- These professional codes do not place any emphasis on the rights of its members.
- For example, what if a member disagrees with a term in the ethical code?

Justification

- These professional codes fail to indicate what the basic source of justification is.
- For example, all but one term in the BCS code of conduct is of the form ‘You shall do something’.
- **Q. Is this a problem?**

Ambiguity

- Some computing professionals find the codes unhelpful.
- One reason for this is the ambiguity of some of the terms. For example one item of the code is
  - … act with integrity and respect in your professional relationships with all members of BCS and with members of other professions with whom you work in a professional capacity...
- Is this useful?
- Much of these existing professional codes are open to personal interpretation.

Prioritising

- Neither code establishes priorities among its different obligations. For a particular ethical problem, more than one principle of the code is likely to apply.
- This could lead to contradictory advice when applying the codes to a concrete problem.
- For example, it is not hard to imagine the following items of the BCS code being placed in competition.
  - **NOT** disclose or authorise to be disclosed, or use for personal gain or to benefit a third party, confidential information except with the permission of your Relevant Authority, or as required by Legislation
  - **NOT** misrepresent or withhold information on the performance of products, systems or services (unless lawfully bound by a duty of confidentiality not to disclose such information), or take advantage of the lack of relevant knowledge or inexperience of others.

For and Against Ethical Codes

- **Advantages:**
  - Define a common standard of behaviour, so clients know what to expect;
  - Codes can support a refusal to behave unethically, so even 'obvious' rules have their uses.
- **Disadvantages:**
  - Codes can give contradictory advice;
  - May give rise to complacency - practitioners think that so long as they are obeying the code, they need not concern themselves with ethical issues.
  - Could draw attention away from major ethical issues (e.g., how technology should be introduced and controlled) towards smaller immediate issues (the behaviour of individuals).
  - May suggest a dichotomy between ethical decision making in personal life and in professional life.

Computer ethics and value sensitive design

Computer ethics and value sensitive design

- Helen Nissenbaum (2001) - How computer systems embody values
  - Information technology can change society – *but technology is also developed on the basis of certain values*.
  - Complex interplay between system and "those who built it, what they had in mind, its conditions of use, and the natural, cultural, social, and political context in which it is embedded—all these factors may feature in an account of the values embodied in it."

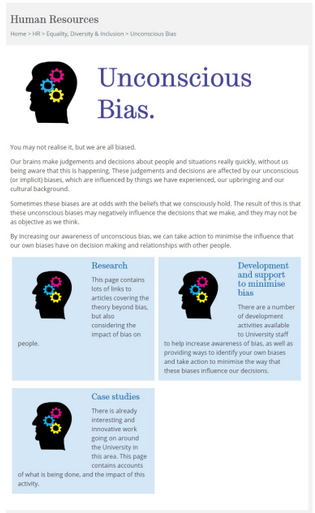
Embodied values

The Guardian, 16/Nov/2017.
https://www.theguardian.com/technology/2017/nov/15/tim-berners-lee-world-wide-web-net-neutrality

Bias in computer systems

- Friedman and Nissenbaum (1996) *Bias in computer systems*.
- Bias can take different forms
  - Pre-existing bias
  - Technical bias
  - Emergent bias

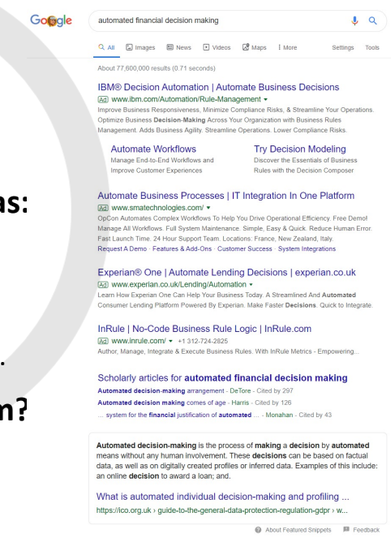
ACM Transactions on Information Systems, Volume 14 Issue 3, July 1996 Pages 330-347.



https://www.sheffield.ac.uk/hr/equality/ub

Example: pre-existing bias

- An important day-to-day application of software is automated decision making.
- A system could implement a **pre-existing bias**:
  - An automated loan advisor might negatively weight applicants who live in “undesirable” locations, such as low-income or high-crime neighbourhoods, as indicated by their postcode.
- **Q. Is this bias sensible, or is there a problem?**



Example: technical bias

- Some technical decisions can influence matters significantly for some users, e.g.
  - When listing flights, which are listed first? Direct or indirect?
  - The system may have a bias towards placing certain flights first – for instance those with all segments on a single carrier – which could favour certain airlines
  - Might miss better alternatives for someone flying from (say) Phoenix, USA to New York and then London.

Example: emergent bias

- This is bias which emerges over time as the system’s use changes, e.g.,
  - In USA, the computerised National Resident Medical Match program is used to place medical students in their first jobs.
    - Developed in the ‘70s it was assumed that only one member of the family needed placing.
    - But increasing numbers of women led to more marriages between residents, and it was biased against couples.

Algorithmic bias

- Why not use an objective algorithm, trained on data, rather than risk human prejudice?
- Then we could rely on algorithms to recruit people for a job, give them a loan, identify them as a suspect in a crime, send them to prison or grant them parole…
- BUT
  - Increasingly clear that such algorithms often reflect pre-existing bias
  - Machine learning algorithms opaque and unregulated
  - If data used for training is biased, bias persists

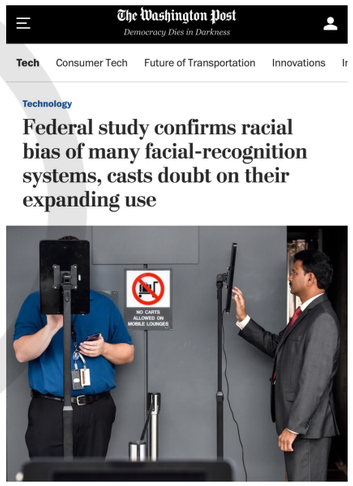
Automatic justice

- Criminal justice system has been shown to have systematic racial biases in USA. Black people are:
  - arrested more often than whites, even when they commit crimes at the same rate.
  - sentenced more harshly and more likely to be searched or arrested during a traffic stop
- COMPAS: widely used algorithm that assesses whether defendants and convicts are likely to reoffend in the future. Seems to work: white and black defendants given higher risk scores tend to reoffend at the same rate.
- BUT(*)
  - black defendants twice as likely to be mislabelled as high risk than white defendants
  - white defendants twice as likely to be mislabelled as low risk
  - black defendants may have higher previous offence counts due to existing racial bias, so the system's inputs potentially have bias built into them no matter how good the algorithm is

(*) The system's developers dispute the situation. See https://fivethirtyeight.com/features/technology-is-biased-too-how-do-we-fix-it/amp/ for a fuller discussion of technological bias.

Example – face recognition

- Facial-recognition systems in the US misidentified people of colour more often than white people.
- Asian and African American people were up to 100 times more likely to be misidentified than white men.
- Women were more likely to be falsely identified than men.
- Middle-aged white men generally benefited from the highest accuracy rates.
- **Q. What's going on?**



Officials program iPads loaded with new facial-recognition scanners last year at Dulles International Airport. (Bill O'Leary/The Washington Post)

Can we remove bias?

- How might we remove bias?
  - Greater transparency?
  - Require human review?
  - Spread awareness that algorithms can be biased?

What about changing the way we design things?

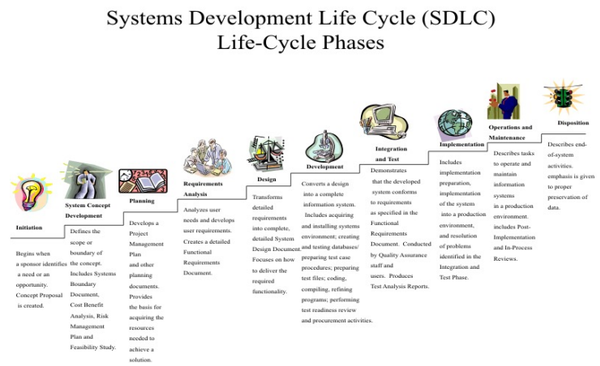


Image: By US Department of Justice - INFORMATION RESOURCES MANAGEMENT, Public Domain, https://commons.wikimedia.org/w/index.php?curid=5061812

Value sensitive design

- Friedman, Kahn and Borning (2008)
  Chapter 4 of *The Handbook of Information and Computer Ethics*: "Value Sensitive Design and Information Systems"
- Value sensitive design is "*a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process*"
- Extend the traditional criteria used to evaluate a system (e.g. reliability and correctness), and also consider its relationship to core human values.

- See https://en.wikipedia.org/wiki/Value_sensitive_design for more.
- See: Himma & Tavani (2008) The Handbook of Information and Computer Ethics. Wiley. http://www.cems.uwe.ac.uk/~pchatter/2011/pepi/The_Handbook_of_Information_and_Computer_Ethics.pdf

Example (Friedman and Kahn, 2003)

- **The relationship between values and technological design**
  - Inuit people were given snowmobiles:
    - This moved transportation methods away from dog sleds, but also introduced a symbol of social status and a dependence on money economy.

A better approach

- Value sensitive design is concerned with shared human values such as
  - Well-being
  - Human dignity
  - Justice
  - Human rights
- e.g., the right to privacy

Image by Gerd Altmann from Pixabay

How to apply it

1. Start with value, technology or context of use
2. Identify direct and indirect stakeholders (indirect – those who don't use the system but could be affected by its use)
3. Identify potential harms and benefits of system for stakeholders
4. Map harms and benefits onto corresponding values
5. Conceptually investigate values (philosophical work can be useful here)
6. Identify value conflicts (e.g. Privacy vs security, trust vs security, accountability vs privacy)
7. Explore effects of different technical designs on value conflicts, and on different groups of stakeholders
8. Design for flexibility, and subsequent modification based on feedback.

What it involves

- Value sensitive design involves an interactive process that examines conceptual, empirical and technical issues.
- **Conceptual** – how are values supported or diminished by designs
- **Empirical** – researching understandings and experiences of those affected by application
- **Technical** – identifying values and developing technical mechanisms that support them

Example: cookie technology

applying conceptual, empirical and technical investigations

- **Conceptual issues**
  - What values are implicated? Trade-offs? Weighting (e.g., between moral values such as privacy, and non-moral values e.g. aesthetic preferences)
  - Value of informed consent is involved, protection of other human values such as privacy, autonomy, and trust.
  - Do people understand what they are agreeing to?
- **Empirical issues**
  - Observing people's use of and understanding of cookies on websites
- **Technical issues**
  - Designing systems to support the values that were identified as relevant
  - Looking at how the technological properties support or hinder human values.

Participatory design

- This involves stakeholders, designers, researchers and end users in design process to help ensure that the end product meets the needs of the intended user base
- Value-sensitive design is linked to participatory design, but also involves the inclusion of marginalised perspectives in the design process.
- E.g.: maybe none of the current participants has a disability, but accessibility still needs to be considered

Research integrity and ethics at The University of Sheffield

https://www.sheffield.ac.uk/rs/ethicsandintegrity

Good Research & Innovation Practices Policy

THE UNIVERSITY OF SHEFFIELD

POLICY ON GOOD RESEARCH AND INNOVATION PRACTICES

This Policy is in three sections:

(1) Good Research and Innovation Principles
This explains the principles governing all the University's research and innovation (R&I) activities, the purpose of the Policy, its value and whom it applies to;

(2) Good Research and Innovation Practices
This clarifies the University's expectations concerning good practices in research (R) and/or in innovation (I) activities;

(3) Annex
This contains information on what the University means by unacceptable R&I practices (sometimes referred to as research misconduct, questionable practices, or the use of unfair means), provides additional detailed information on good R&I practices, lists the University's policies that this Policy complements, and lists the sources which have been consulted in developing the Policy.

All three sections have the authority of University of Sheffield policy.

For the sake of brevity the *Policy on Good Research and Innovation Practices* is referred to as the *GRIP Policy*, the University of Sheffield as *the University* and research and innovation as *R&I*.

Good Research & Innovation Practices Policy

The University of Sheffield's Good Research & Innovation Practices Policy

CONTENTS

Good Research & Innovation Practices Policy

4.3 RESEARCH ETHICS

Two important dimensions of research integrity are how to ethically undertake research with human participants, and research with animals. Research ethics with respect to humans is defined by the University's *Ethics Policy Governing Research Involving Human Participants, Personal Data and Human Tissue*. The GRIP Policy provides the governing framework for the *Ethics Policy*. Research ethics with respect to animals is defined by the University's *Ethical Policy on the Use of Animals*.

4.4 PROFESSIONAL ETHICS

Many researchers are members of professional bodies that may have their own policies and rules. As professionals, researchers have responsibilities in addition to those of the general population, not least because of the standing of the research profession in society. Drawing upon their disciplinary expertise, as professionals researchers are capable of making informed observations, publications, presentations or statements that the general public cannot.

In the research context remaining professional does not always mean divorcing oneself from, or remaining independent of, cultural, political or religious perspectives (for example, being a member of a pressure group or political party may be relevant to, and provide a motivation for, research). However, remaining professional means to remain objective and, where relevant, researchers should be open and transparent about their personal views, perspectives or beliefs. As professionals, and as members of society, researchers should endeavour not to cause harm. If an R&I project has the potential to cause harm it should be designed in a way that minimises the potential to harm, and this potential for harm should be explicit.

4.5 UNACCEPTABLE R&I PRACTICES

What is meant by unacceptable R&I practices (sometimes referred to as corrupt practices, questionable practices or the use of unfair means) is clarified in the Annex. The causes of unacceptable practices vary, but a common cause is a lack of awareness and understanding of good R&I principles and practices. Relatively few instances of unacceptable practice are caused by deliberate, dangerous, reckless or negligent deviations from good R&I practices. Therefore, depending on the type of unacceptable R&I practice, the remedy may range from advice, guidance, mentoring or formal training through to an investigation of potential research misconduct in accordance with the University's *Policy on Investigating and Responding to Allegations of Research Misconduct*. The latter may lead to disciplinary action. If and when an unacceptable R&I practice occurs, and the appropriate remedy is educational, then the affected researcher(s) and their

Summary

Summary

- **Codes of Conduct** are a means by which professional organisations can control their membership and avoid external regulation.
  - They also act as a guide to professional decision making.
  - Ethical problems often arise because of the special nature of computer technology.
- The BCS , IEEE and ACM have established professional codes of conduct. They are flawed, but have some merit.

Summary (continued)

- **Undesirable:** Computing professionals are bound by a fixed set of rules that distance them from their own ethical standards and those of society in general.
- **Desirable:** Professional ethical codes act as guidelines for ethical decision making in the context of professional employment, not in life in general. The guidelines place an emphasis on the personal ethics of the individual and their place within the ethics of society as a whole.

Summary (continued)

- Value sensitive design emphasises the idea of values in design
  - **Conceptual** – how are values supported or diminished by designs
  - **Empirical** – researching understandings and experiences of those affected by application
  - **Technical** – identifying values and developing technical mechanisms that support them

Take-home message

It is important to think about the effect of your work on human society