THE UNIVERSITY OF WESTERN AUSTRALIA

Ethical Issues for Computing Professionals CITS3200

Alex Reid

Honorary Professorial Fellow, CSSE

The University of Western Australia

COPYRIGHT

Note: all Web references in these Powerpoint slides verified as at 23-Jul-25

CONTENTS

Aims:

1. Give an understanding of the variety of ethical issues you may confront.
2. Impart an appreciation of the complexity of many of these issues.
3. Help you to see you do have a responsibility, and to whom.
4. Introduce the Computer Society Codes of Ethics & Conduct as a basis.
5. Introduce a Framework for addressing ethical issues.

Why Computer Ethics #1?

- **Technology has Consequences – You Can't Ignore Them:**
  - **Samuel Johnson, 1759:** Integrity without knowledge is weak and useless, and knowledge without integrity is dangerous and dreadful.
  - **Albert Einstein, 1931:** It is not enough that you should understand about applied science in order that your work may increase man's blessings. Concern for man himself and his fate must always form the chief of all technical endeavours.
  - **Robert Oppenheimer, 1945**: knowledge cannot be pursued without morality, and scientists must ask what we ought to do with our knowledge, not merely what we can do.
  - **Norbert Wiener, 1950:** The new industrial revolution is a two-edged sword. It may be used for the benefit of humanity… It may also be used to destroy humanity, and if it is not used intelligently it can go very far in that direction.
  - **Rogerson & Bynum, 1995:** Computing Technology is the most powerful and most flexible technology ever devised. For this reason, computing is changing everything – where and how we work, where and how we learn, shop, eat, vote, receive medical care, spend free time, make war, make friends, make love.

Why Computer Ethics #2?

- **Computing Creates New Situations:**

  - **Walter Maner (1976):**

    Computer ethics = moral problems that are created, aggravated or transformed by the introduction of computer technology.

  - **James Moor (1998):**

    Computers are logically malleable:

    --> applied in unpredictable and novel ways

    --> situations & choices not previously arising

    --> policy vacuums.

    Values permeate our lives – help us make decisions. We don't always agree about all values, but many we do (eg what makes for a "good" program? – no universal agreement, but some convergence).

Why Study Computer Ethics? #1

- **Rationale for studying computer ethics (Maner, 1995):**
  - **It makes us behave like responsible professionals.**
  - **It teaches us how to avoid computer abuse and catastrophes.**
  - **We need to recognise policy vacuums created by advances in IT.**
  - **Some problems (eg Intellectual Property) are radically and permanently altered.**
  - **IT creates novel ethical issues that require special study.**
  - **These novel issues are large enough and coherent enough to define a new field.**

Why Study Computer Ethics? #2

- Example situation where moral/ethical choices have to be made (Moor, 1998):
  - A range of actions you as the owner of a Web site can take which impact a user's computer/smartphone when they use their Web Browser to access your site:
    a. do not change user's computer at all.
    b. allow user to decide if a cookie is to be left on the user's computer or not.
    c. leave a cookie on user's computer and inform them it's there.
    d. leave a cookie on user's computer but without their knowledge.
    e. removal of data from user's computer without their knowledge.
    f. arbitrary destruction of data on user's computer.

Note that, starting in Europe, you are now required to get user consent before leaving a cookie on the user's computer.

Framework for Thinking Ethically -1

A. Be Clear What Ethics is Not:

- It is not the same as Feelings
- It is not Religion
- It is not just following the Law
- It is not following Culturally Accepted Norms
- It is not science.

B. Approaches to Deriving Ethical Standards:

1) Utilitarian approach
2) Rights approach
3) Fairness or Justice approach
4) Common Good approach
5) Virtue approach

Framework for Thinking Ethically -2

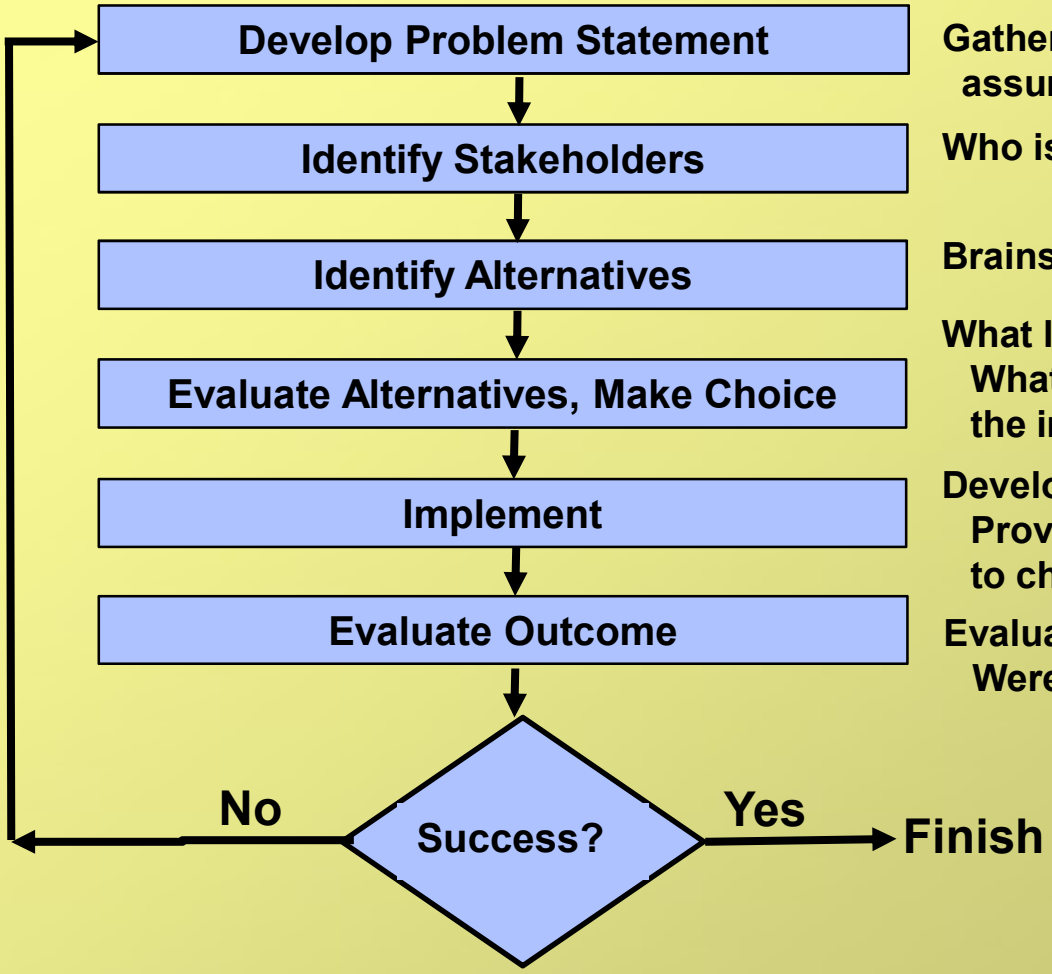
C. Decision Framework:

- Recognise there is an Ethical Issue (not just a technical problem).
- Get the Facts.
- Evaluate Alternative Actions & Consider which Action will:
  1) Produce most good, do least harm? [utilitarian]
  2) Best respect rights of all stakeholders? [rights]
  3) Treat people equitably? [fairness]
  4) Best serve the community as a whole? [common good]
  5) Lead me to being the sort of person I want to be? [virtue]
- Make a Decision and Test it.
- Act and Reflect on the Outcome.

From: https://www.scu.edu/ethics/ethics-resources/a-framework-for-ethical-decision-making/

Framework for Thinking Ethically -3

Ethical Dilemmas: Decision Process:



Gather and analyse the facts. Avoid making assumptions.

Who is most affected? Who else is affected?

Brainstorm. Involve others, incl. stakeholders.

What laws, guidelines, policies, principles apply? What parts of the Code of Ethics apply? What's the impact on you, your orgn., the stakeholders?

Develop and execute an implementation plan. Provide leadership in overcoming any resistance to change.

Evaluate outcome against success criteria. Were there any unintended consequences?

(adapted from Reynolds: Ethics in Information Technology, 5th Edn, 2015, p.20)

Why Are Computers Special?

Characteristics of Computers :

- powerful, fast => magnifying effect
- manipulate information => a new kind of tool
- new, evolving => don't understand them fully
- logically malleable => applied in novel, unusual ways
- have memory => adaptive, unpredictable
- complex => even programmers don't understand their programs
- programs can't be proven to be correct, & not 100% reliable => untrustworthy (yet we still rely on them)
- non-proportional effects => minor errors can produce catastrophic results
- pervasive, cheap => effects are very widespread
- copies that are identical to the original => ownership rights issues
- introduce spatial and temporal separation => break the chain of responsibility, facilitate anonymity (especially with AI)
- ...

Computing Technology is the most powerful and most flexible technology ever devised

Some Ethical/Moral/Social Issues -1

- **Software Errors:**
  - Are you responsible for any and all errors found in your software?
  - What if someone else has modified your software – who is responsible then?
  - What if an error only emerges 10 or 20 years later – are you still responsible?
- **Copying Video/Music/Image Files:**
  - Is it actually "theft" even when the owner still has the file?
  - Is it OK if you just want to "borrow" the file?
  - Is it OK if you copied it to see if it was worth buying?
- **Software Ownership:**
  - How much of someone else's software can you re-use without their permission?
  - Is it OK to reverse-engineer someone's software?
  - Do the arguments that no-one should "own software" have any merit?
- **Email Issues:**
  - When is it OK to forward or broadcast someone else's email?
  - Are there situations where anonymous email is legitimate?
  - Is it ever OK to send out thousands of copies of an email?
  - Is "chain email" harmless and so OK?

Some Ethical/Moral/Social Issues -2

- **Hacking:**
  - If Hacking is done with altruistic motives, is it OK?
  - If Hacking reveals flaws in someone's computer security, isn't that a good thing?
  - If people don't protect their files, does that mean it's OK to copy them, etc?
- **Viruses:**
  - If your co-workers don't update their virus protection, isn't that their fault?
  - Is a virus that doesn't actually damage computers ever OK?
  - If a virus exposes system weaknesses, isn't that a good thing?
- **Privacy:**
  - Is it OK to photograph/video someone and then distribute that photo/video?
  - When is it OK to "photoshop" a photo of someone or something?
  - If a photo or image is available on the Web, then is it OK to copy and use it?
- **Software Development:**
  - How much software testing is enough to clear you of blame for errors?
  - If your "client" wants something that is "unreasonable" can you just ignore it?
  - How important is it to stick to your estimates for how long it will take?

Catalogue of Failures - 1

Computers and/or software failure have been implicated in:

- Hole in ozone layer undetected for 7 years.
- US Air Force Blackhawk helicopter crashes – 22 deaths.
- Hubble Telescope error compounded by computer shut-down (9-Dec-91).
- Three Mile Island (nuclear reactor) (28-Mar-79).
- Chernobyl (nuclear reactor) (26-Apr-86).
- Challenger Space Shuttle deaths (28-Jan-86).
- Mt Erebus Air NZ flight 901 crash (28-Nov-79).
- Korean Air Lines flight 007 over Sakhalin Island (1-Sep-83).
- HMS Sheffield sinking in Falklands (4-May-82).
- Iranian flight 655 shot down over Persian Gulf (3-Jul-88).
- Stock market crash due to automated trading in 1987.
- Australian Census Website Failure 9-Aug-16.
- Passport system failure delays flights 29-Apr-19 and again on 15-Jul-19.
- Robodebt – automated calculation of overpaid social security – Nov 2019.
- CrowdStrike – update fault brought the world to a standstill – 19-Jul-24.

… etc – a serious study can be made of computer disasters (eg Peter Neumann's Risks Digest - http://catless.ncl.ac.uk/Risks/).

Consequences of Computer Error

Who is to blame when computer systems fail?

Blame for Computer Errors

Cartoon depicting computer blaming human error “yet again” [www.cartoonstock.com]

Blame for Computer Errors

Cartoon depicting people committing suicide because of dramatic downturn in profitability only to discover it was caused by computer error [www.cartoonstock.com]

Blaming the Computer

Cartoon of computer taking the blame for a sales nose-dive (jumping out the window).
From ENTEC Catalogue, UK, Oct 95

Blame for Computer Errors

Cartoon depicting committee being advised the computer failure was due to someone walking on a sidewalk crack
[www.cartoonstock.com]

Catalogue of Failures - 2

Why is Software so prone to Catastrophic Failure?

- Complexity
- Error Sensitivity – non-linear, non-continuous (non-proportional)
- Hard to Test for every possible situation
- Correlated failures
- Lack of professional standards – few software engineers
- Development methodologies have been inadequate
- Proving software correctness has not been successful
- Verification attempted by:
  - mathematical analysis;
  - case analysis;
  - extensive testing; or
  - combination of the three.
- Tony Hoare's "Wasted 20 Years" trying to establish a basis for proving program correctness.
- Roger Needham's "Most Surprising Development in the 50 years of Computer Science" – that, on a regular basis, we would use software known to have significant numbers of bugs.

Robbie the Killer Robot

- Industrial Robot killed its operator: who was implicated?
  - Programmer had made an error in the relevant program
  - Operator did not follow instructions correctly
  - Supervisor did not ensure operator was adequately trained
  - Management cutting corners

See https://onlineethics.org/cases/case-killer-robot

More recent actual deaths:

- Volkswagen car manufacturing robot kills worker 2-Jul-15: https://www.washingtonpost.com/news/worldviews/wp/2015/07/02/a-robot-killed-a-factory-worker-in-germany-so-who-should-go-on-trial/;
- Tesla robot-driven car driver killed 7-May-16: https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk.
- Uber self-driving car kills pedestrian 18-Mar-18: http://www.abc.net.au/news/2018-03-20/uber-suspends-self-driving-car-tests-after-fatal-crash/9565586
- Wikipedia self-driving car fatalities page: https://en.wikipedia.org/wiki/Self-driving_car#Incident

AI/Robot Issues -1

Robots throw some of these issues into strong relief:

I, Robot

Blade Runner

Metropolis (1927)

Terminator

Pictures of various humanoid robots, mostly from movies.

Humans

Asimo

Star Wars

Ex Machina

AI/Robot Issues -2

Robots throw some of these issues into strong relief:

Pictures of various industrial robots, eg cars, vacuum cleaners, assembly-line manufacturing, bomb disposal, stock trading.

AI/Robot Issues -3

- **Asimov's 3 Laws of Robotics:**
  1. **A robot may not injure a human being or, through inaction, allow a human being to come to harm.**
  2. **A robot must obey orders given it by human beings except where such orders would conflict with the First Law.**
  3. **A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.**
- **Inadequacy of this Ethical Framework:**
  - ✡ **Adverse consequences arising from well-intended actions.**
  - ✡ **"Greater good" aspects (humanity as a whole vs individual humans).**
  - ✡ **Failure to see long-term consequences.**
  - ✡ **Some outright failures.**
  - ✡ **Complexity of ethical judgements (and fragility of trust).**

See Kuipers, Benjamin: Towards Morality and Ethics for Robots, *2016 AAAI Spring Symposium on Ethical and Moral Considerations in Non-Human Agents*
https://web.eecs.umich.edu/~kuipers/research/pubs/Kuipers-sss-16.html

Therac-25 Radiation Treatment Machine (1985-1987)

- Machine malfunction produced overdoses (100x)
- 4 or 5 patients died as a result of the failure
- Operators ignored error messages: “Malfunction 54”
- No immediate effects were noticed
- Manufacturer safety procedures were inadequate
- FDA tests were inadequate
- Remediation efforts were paltry
- Software error was eventually discovered

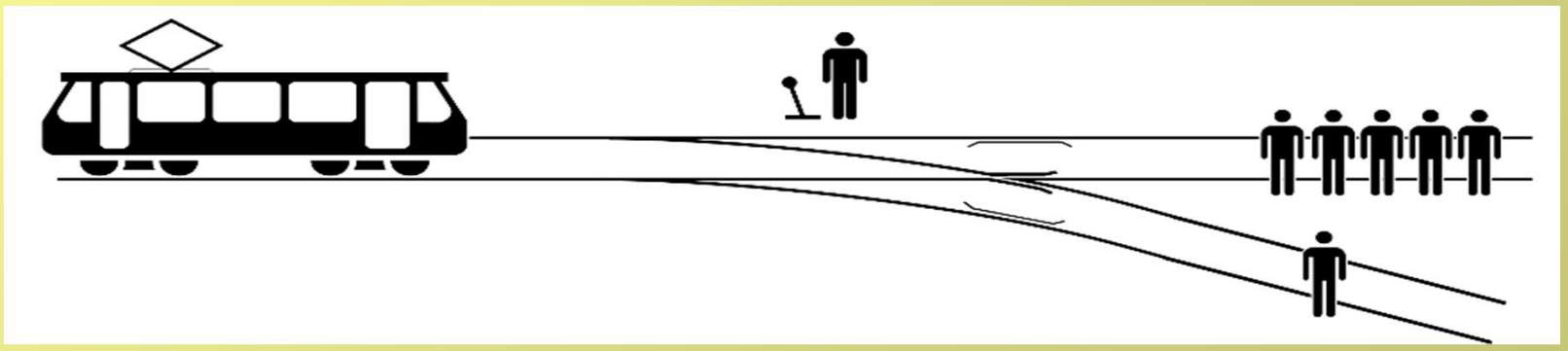
See http://staff.washington.edu/jon/pubs/safety-critical.html

The Classic "Trolley" Case:

A trolley is hurtling out of control along a track.

Ahead of it is a bunch of deaf children playing on the track, unaware of the trolley, and certain to be killed by it.

You're too far away to warn the children, but you are adjacent to a switch that would turn the trolley onto a siding, but there is a man there who has his foot stuck in the tracks, and he would certainly be killed if you diverted the trolley onto that track.



The Classic "Trolley" Case:

What do you do?

1. **Do nothing, turn your back, don't get involved.**
2. **Take a selfie with the trolley and post on Twitter!**
3. **Deliberately decide to leave the switch alone and let the kids be killed.**
4. **Divert the trolley onto the siding, killing the man but saving the kids.**

There are many variations to this, trying to balance the "value" of lives against each other.

See https://en.wikipedia.org/wiki/Trolley_problem

You could say "this is too hard" – leave it to your instinct at the time should you ever have to face such a dilemma.

A Self-Drive Car Version of the "Trolley" Dilemma:

The same dilemma as in the classic 'Trolley' dilemma presents itself with those who program self-drive vehicles, illustrating one way in which Asimov's 3 laws are inadequate.

Imagine a situation (say) where a drunk driver of a truck is about to collide head-on with a self-drive car at high speed, with a very strong chance that all 4 people in the car will be killed.

The car can avoid the collision by swerving onto a footpath, killing a pedestrian.

How should the car be programmed?

1. To just brake hard and keep the same line (even though this means multiple deaths).
2. To swerve onto the footpath to avoid the head-on collision, but killing one pedestrian.

Programming the Self-Drive Car :

This is just one possible scenario that the car must be programmed to deal with:

- **For instance if several children are on the footpath, and all would be killed, does that change the equation?**
- **How many children/pedestrians should be sacrificed to save the car occupants?**
- **What if the pedestrian had stepped (illegally) onto the roadway into the path of the swerved car.**

In contrast to the “classic” dilemma, you can’t say “this is too hard” – you *have* to program the car to make some “decision” for every circumstance it may face.

AI Ethics Framework

Criteria for Determining if it is "Good" AI:

Broadly, the key concepts are:

- **Fairness to the directly impacted individuals as well as the broader societal impact.**
- **Ethics or equity.**
- **Accountability.**
- **Transparency, including explicability.**

= FEAT (or FATE)

See ACS Primer: "The Ethics and Risks of AI Decision-Making"
https://www.acs.org.au/content/dam/acs/acs-publications/Replacementdocuments2022/ACS_AI-Decision%20Makaing_A03_FA.pdf

AI Ethics Framework

Australian Government AI Ethics Principles (2019):

1. **Human, Social and Environmental Wellbeing:** throughout their lifecycle, AI systems should benefit individuals, society and the environment.
2. **Human-Centred Values:** throughout their lifecycle, AI systems should respect human rights, diversity and the autonomy of individuals.
3. **Fairness:** throughout their lifecycle, AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.
4. **Privacy Protection and Security:** throughout their lifecycle, AI systems should respect and uphold privacy rights and data protection and ensure the security of data.
5. **Reliability and Safety:** throughout their lifecycle, AI systems should reliably operate in accordance with their intended purpose.
6. **Transparency and Explicability:** there should be transparency and responsible disclosure to ensure people know when they are being significantly impacted by an AI system, and can find out when an AI system is engaging with them.
7. **Contestability:** when an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge its use or output.
8. **Accountability:** those responsible for the different phases of the AI system lifecycle should be identifiable and accountable for its outcomes, and human oversight of it should be enabled.

https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework

Australian Federal Court Rules that AI can be an Inventor

What does this mean?

- Handed down on 30-Jul-21 for an ML system, DABUS, by which 2 inventions have been made leading to 2 patent applications.
- An inventor could legally be an AI system.
- But an AI system cannot be the owner, controller or patentee of an invention.
- Aim is to promote technological innovation, regardless of its origin.
- First recognition worldwide of an AI patent claim.

What ethical implications does this have?

- Who owns the intellectual property? Who owns the patent?
- Who is responsible if the AI system makes a mistake?
- If I utilise such an AI system belonging to someone else, who owns what it "invents"? – Me, the AI System, the Writer of the AI System, No-one?
- Who owns any money made by such an AI system?
- Can such a system be sued? Or otherwise taken to court?
- What rights could such an AI system be accorded? (already one robot has been made a citizen of Saudi Arabia, in 2017).

https://ia.acs.org.au/article/2021/federal-court-rules-ai-can-be--inventor.html
https://ia.acs.org.au/article/2017/first-robot-to-be-granted-citizenship.html

Gartner Hype-Cycle for Artificial Intelligence, 2025

Gartner's Hype Cycle for Artificial Intelligence, June 2025:

https://www.gartner.com/interactive/hc/5505695?ref=solrAll&refval=421908450

Source: Gartner June 2025

Sources of Ethical Guidance

Hierarchy of Policies to Guide Conduct:

- **international treaties & agreements**
- **national laws**
- **government/agency regulations**
- **standards of good practice (within a whole industry)**
- **professional codes of conduct (within a professional association)**
- **corporate policies (within an organisation/corporation)**
- **community & personal values (unwritten common practices)**

Terrell Ward Bynum (1997)

Not ultimately guiding what you believe, but rather how you act, your behaviour.

Being a Professional

- 2 Scenarios:
  - New medical graduate, just starting out in a medical practice:
    - expected to "act professionally"
  - New high school graduate, taking a job as a cashier at Coles:
    - expected to "act professionally"
  - What's the difference?
  - To which is a new computing graduate closest?

- Abraham Flexner (1915) on medical professionalism:
  - It is basically intellectual, carrying with it high responsibility
  - It is learnèd in nature, because it is based on a body of knowledge
  - It is practical rather than theoretical
  - Its technique can be taught through educational discipline
  - It is well organised internally
  - It is motivated by altruism

Professional Body

- **Criteria:**
  - **Established collection of specialised knowledge**
  - **Formal accrediting criteria**
  - **Members undertake decisions on behalf of clients (who rely on their specialist expertise)**
  - **Defined performance standards**
  - **Members committed to maintain performance standards, body of knowledge**
  - **Acceptance of responsibility**
  - **Standards of conduct/ethics (and disciplinary procedures)**
  - **Recognition in society – high level of trust**

- **Summary:**

  professionals are people who have specialised knowledge on which others (and the public in general) have to place dependence; the public have to trust those professionals in regard to their specialised knowledge.

  Viz: TRUST => RESPONSIBILITY

Idea of a 'Professional Society':

- It is not a:
  - business - shareholders' interests are uppermost;
  - trade union - wellbeing of members – esp. wages and conditions;
  - industry association - advancing businesses in an industry;
  - club - mutual interest.
- Rather, it is a group of people who:
  - have expertise and professional skills;
  - claim a degree of autonomy in the exercise of their expertise;
  - accept responsibility for the results of their actions, and
  - as a group, hold a position in society that comes from having its members serve society as well as private interest,
  - or, alternatively, when pursing personal interests, it's not done at the expense of the well-being of the communities and society affected and on which they depend.

Australian Computer Society

The Professional Association for Australia's ICT Sector.

AUSTRALIAN COMPUTER SOCIETY

ACS is the leading professional association with over 49,500 members, representing Australia's technology community, across industry, government and education

Explore its wide range of services, developed to cultivate the highly-skilled, diverse workforce Australia needs for a future powered by technology.

49,500+ ACS Members	**1,500+ Certified Career Professionals**	**300+ ACS events every year.**	**35,000+ Training videos to upskill and learn from.**

https://www.acs.org.au/

ACS Code of Professional Ethics –1

ACS Code of Professional Ethics:

- The ACS's commitment is to a community served by an ethical and trustworthy ICT profession, and it strives to protect the public interest. To do this, the ACS seeks to promote the highest standards of ethics and technical knowledge, education and service excellence in its members and the wider profession in the practice of ICT to ensure the community overall benefits from the use of technology.
- ICT professionals are expected to practise the following core values in alignment with their ICT-based role, context and capabilities:
  - ✡ **Honesty**
  - ✡ **Trustworthiness**
  - ✡ **Respect – for Others and for the Profession**
- This Code of Professional Ethics applies to all ACS members regardless of their role or specific area of expertise in the ICT industry.
- The Code of Professional Ethics should be adhered to in conjunction with the Code of Professional Conduct

https://www.acs.org.au/content/dam/acs/CodeOfProfessionalEthics_Mar_2023.pdf

ACS Code of Professional Ethics –2

ACS Code of Professional Ethics Values:

1. Honesty:

Honesty is a founding principle for healthy interactions between people and the design and function of ICT systems. As an ACS member, you are expected to: …

2. Trustworthiness:

Trustworthiness is a critical principle in upholding public welfare and human dignity as an ICT professional. As an ACS member, you are expected to: ...

3.1 Respect for Others:

Respect for others is an essential principle that underpins the interactions of an ICT professional. As an ACS member, you are expected to minimise harm to any stakeholders from your professional activities and: …

3.2 Respect for the Profession:

Respect for the Profession is an abiding principle that governs the roles and responsibilities of ICT professionals. As an ACS member, you are expected to: …

ACS Code of Professional Ethics –3

ACS Code of Professional Ethics Expanded:

1. **Honesty:**
   a. Honest, open, truthful
   b. Not misrepresent capabilities or actions
   c. Speak out if detect unprofessional conduct
2. **Trustworthiness:**
   a. Be accountable for all actions
   b. Practise integrity
   c. Commit to lifelong learning
   d. Respect privacy, confidentiality, integrity of data
   e. Communicate your capabilities and limitations clearly
   f. Only undertake work for which you have the skills
   g. Be competent and strive for quality
   h. Develop robust, secure, user-friendly systems
3. **Respect (for Others, and for the Profession):**
   a. Minimise harm to all stakeholders, mitigate where harm is unavoidable
   b. Impartial, inclusive, mitigate risks, respect others' views, respect IP
   c. Educate public, sustainability, advance IT, contribute to challenges, etc

Ethical Values

ACS Code of Professional Ethics is founded on these primary Values:

1. Honesty
2. Trustworthiness
3. Respect (for Others, and for the Profession)

The Code has recently been revised (March 2023), to reflect the changing landscape in the Computing Profession (it was previously updated in 2014). One key goal in the last revision was to generalise it more, basing it more on broad Values, so that it can more readily be applied to new developments in ICT like AI.

The Code of Professional Conduct is currently in the process of being revised to align it more closely with the Code of Professional Ethics. There are no inconsistencies or conflicts between the two Codes, just different ways of expressing the same Values.

ACS Code of Professional Conduct –1

ACS Code of Professional Conduct explained:

1. **The Primacy of the Public Interest**
   You will place the interests of the public above those of personal, business or sectional interests.
2. **The Enhancement of Quality of Life**
   You will strive to enhance the quality of life of those affected by your work.
3. **Honesty**
   You will be honest in your representation of skills, knowledge, services and products.
4. **Competence**
   You will work competently and diligently for your stakeholders.
5. **Professional Development**
   You will enhance your own professional development, and that of your colleagues and staff.
6. **Professionalism**
   You will enhance the integrity of the Society and the respect of its members for each other.

ACS Code of Professional Conduct –2

ACS Code of Professional Conduct detail:

https://www.acs.org.au/content/dam/acs/rules-and-regulations/Code-of-Professional-Conduct_v2.1.pdf

1.2.1. The Primacy of the Public Interest

a) The public interest takes precedence over personal, private and sectional interests.
b) Any conflicts should be resolved in favour of the public interest.
c) In your work, you should safeguard the interests of your immediate stakeholders, provided that these interests do not conflict with the duty and loyalty you owe to the public.
d) The public interest is taken to include matters of public health, safety and the environment.

ACS Code of Professional Conduct –3

ACS Code of Professional Conduct detail: (cont)

1.2.2. The Enhancement of Quality of Life

a) The development of ICT has had a significant impact on our society and way of life.
b) Whilst this impact has been beneficial to a very great extent, like all technologies, ICT has also had some negative effects, and will continue to do so.
c) An ethical approach to your work will help to recognise and minimise these adverse effects.
d) You should promote equal access to the benefits of ICT by all members of society.

ACS Code of Professional Conduct detail: (cont)

1.2.3. Honesty

a) Do not breach public trust in the profession or the specific trust of your stakeholders.
b) Observance of utmost honesty and integrity must underlie all your professional decisions and actions.
c) Circumstances will undoubtedly arise during the course of your professional career where it may appear to be beneficial for you to be deceptive in some way.
d) This type of behaviour is not acceptable professional conduct.

ACS Code of Professional Conduct detail: (cont)

1.2.4. Competence

a) Accept only such work as you believe you are competent to perform.

b) Do not hesitate to obtain additional expertise from appropriately qualified individuals where advisable.

c) You should always be aware of your own limitations and not knowingly imply that you have competence you do not possess.

d) This is distinct from accepting a task of which the successful completion requires expertise additional to your own.

e) You cannot possibly be knowledgeable on all facets of ICT but you should be able to recognise when you need additional expertise and information.

ACS Code of Professional Conduct detail: (cont)

1.2.5. Professional Development

a) Keep yourself informed of such new technologies, practices and standards as are relevant to your work.

b) Others will expect you to provide special skills and advice; in order to do so, you must keep your knowledge up-to-date.

c) You should encourage your staff and colleagues to do the same.

d) Take action to ensure that your hard-won knowledge and experience are passed on in such a way that the recipients not only improve their own effectiveness in their present work, but also become keen to advance their capabilities and take on additional responsibilities.

ACS Code of Professional Conduct detail: (cont)

1.2.6. Professionalism

a) The ICT industry is relatively new and characterised by rapid change. It has not had the opportunity to evolve over many years and acquire its own standards and legislation.
b) The ACS is endeavouring to improve public confidence in the ICT industry.
c) It is imperative that members of the Society maintain professional standards that improve and enhance the industry's image, especially in the workplace.
d) All people have a right to be treated with dignity and respect.
e) Discrimination is unprofessional behaviour, as is any form of harassment.
f) Members should be aware that the ACS can help them resolve ethical dilemmas.
g) It can also provide support for taking appropriate action, including whistle-blowing, if you discover an ACS member engaging in unethical behaviour.

Case Studies Illustrating Many of these Issues:

- Each case involves various aspects of the Codes and/or ethical or social issues.
- They are mostly based on actual cases.
- Analyse each case for the following:
  1. identify those to whom you owe any kind of duty;
  2. assess the extent of harm potentially incurred by each person or category;
  3. assign priorities to the duties owed;
  4. identify possible alternatives;
  5. seek opportunities for negotiation and formation of "social contracts".
- Note that, since decisions are based on value judgements, there will be differences of opinion at times…

Social Contract Theory: http://www.iep.utm.edu/soc-cont/

Case Study 6

Aircraft Industry Quality Control Manager Quandary

- Aircraft Manufacturer is keen to get a new aircraft flight-tested.
- Ground testing on the aircraft was possibly inadequate.
- Company is pressuring QC Manager to "sign off".
- Delays may cost the company business, him his job, etc.
- Test pilot knows his job is risky anyway.
- Danger to the test pilot and to other victims of any crash.

"Social Contract" approach – to whom does the Quality Control Manager have a "contract of responsibility"? Which should take precedence? How to choose between them?

See McFarland, Michael C: "Urgency of Ethical Standards Intensifies in Computer Community", IEEE Computer, March 1990, pp77-81
https://ieeexplore.ieee.org/document/50274 (DOI: 10.1109/2.50274)

Case Study Survey

CITS3200 Ethics Case Studies Survey:

- This is a set of simple Case Studies designed to help you understand some of the Ethical Issues you may face as a computer professional.
- Your responses will be recorded and aggregated with others from the 2025 CITS3200 Class (they will be kept anonymous).
- Average responses will be published in a table later this Semester (with comparisons from earlier years).
- You may suspend answering at any time and resume later.

This Survey is available here:

https://uwa.qualtrics.com/jfe/form/SV_6DnHaKB3CTTZLuZ

Or can be accessed via the CITS3200 Website:

https://teaching.csse.uwa.edu.au/units/CITS3200/resources.html

Or from my personal Website:

https://alex-reid.com/Ethics/

Please attempt this Survey over the next week or so.

END OF LECTURE 1

Lecture 2 will be delivered on 6-Aug-25

(in the meantime, please attempt the Ethics Case Studies Survey)

ACS Codes of Ethics & Conduct

ACS Code of Professional Ethics (recap):

Honesty **Trustworthiness** **Respect (for Others, for Profession)**

ACS Code of Professional Conduct (recap):

1. **The Primacy of the Public Interest**
   You will place the interests of the public above those of personal, business or sectional interests.
2. **The Enhancement of Quality of Life**
   You will strive to enhance the quality of life of those affected by your work.
3. **Honesty**
   You will be honest in your representation of skills, knowledge, services and products.
4. **Competence**
   You will work competently and diligently for your stakeholders.
5. **Professional Development**
   You will enhance your own professional development, and that of your colleagues and staff.
6. **Professionalism**
   You will enhance the integrity of the Society and the respect of its members for each other.

ACS Case Study 1-1

Scenario:

- Jean, a statistical database programmer, is trying to write a large statistical program needed by her company. Programmers in this company are encouraged to write about their work and to publish their algorithms in professional journals.
- After months of tedious programming, Jean has found herself stuck on several parts of the program.
- Her manager, not recognising the complexity of the problem, wants the job completed within the next few days.
- Not knowing how to solve the problems, Jean remembers that a co-worker had given her source listings from his current work and from an early version of a commercial software package developed at another company.
- On studying these programs, she sees two areas of code which could be directly incorporated into her own program and solve her impasse.

What should she do?

ACS Case Study 1-2

Alternative Courses of Action:

1. Incorporate the code into her program; tell no-one and complete the work with a day to spare.
2. Use the ideas from the code she's seen, but write her own code to do that job; tell no-one and complete the work on time.
3. As for 2 above, but declare this to her manager; complete the job on time.
4. Ask permission to use the ideas and/or code she's seen, delaying the job significantly, and potentially incurring a royalty payment.
5. Incorporate the ideas/code she's seen, and deliver on time, but ask permission retrospectively (incurring maybe a larger royalty/penalty payment).
6. Explain her dilemma to her managers, and leave the decision to them.

ACS Case Study 1-3

To Whom Does She Owe Any Kind of Duty?

1. Her Managers: ultimately, they will carry the can for her actions.
2. The Owner of the code she "borrowed".
3. Her Colleague: perhaps betraying a trust?
4. Herself: can she live with herself?
5. Her Company: they are reliant on her to develop the program.
6. Her Company's clients, who may be relying on this product for their business.

Does her duty to "the public good" come into play here at all?

ACS Case Study 1-4

What Parts of the Code of Professional Ethics Apply Here?

1. Honesty: (a) be honest in all interactions, and (b) do not misrepresent any capability: *by implication, if she were to employ someone else's work without giving credit, both these injunctions would be contravened.*

2. Trustworthiness: (b) Exercise integrity; (d) Respect proprietary information: *in 2 or 3 scenarios, her actions would violate both these values.*

3. Respect: 3.1(g) Respect others' intellectual property: *the actions in 2 or 3 cases would be in clear violation of this injunction.*

What Parts of the Code of Conduct Apply Here?

2.2 Public Interest

a) identify those potentially impacted by your work and explicitly consider their interests;

f) respect the intellectual property of others;

2.5 Competence

b) not misrepresent your skills or knowledge;

d) respect and protect your stakeholders' proprietary interests;

g) respect, and seek when necessary, the professional expertise of colleagues in their areas of competence.

Identifying Author of Anonymous Message

- You are the Systems Administrator at your medium-sized Company.
- Your Company has set up an Anonymous on-line Discussion Forum to encourage employee discussion/participation.
- The Forum frequently receives postings which are critical of Company policies, practices, etc.
- Your boss asks you to identify the author(s) of these postings (which you are able to do, using your system privileges).
- What do you do?
  1. Just agree.
  2. Argue the toss with the Boss, but then agree.
  3. Take the matter higher.
  4. Use the existing Forum to ensure this first gets wide publicity within the Company.
  5. Go to the local Press with the story.
  6. Take some other action. What?

Anonymity

Cartoon depicting a dog surfing the Internet, saying to another dog: “On the Internet, no-one knows you’re a dog”.

The New Yorker, July 5, 1993, page 61.

Intellectual Property

Issues:

- **Copyright Act**
- **Moral Rights**
- **Digital Agenda Amendments**
- **Fair Dealing, Section 113P (formerly VA/B)**
- **Attribution, Plagiarism**
- **Software Licences**
- **Shrink-Wrap Licences, Web Extensions**
- **Employer *versus* Employee Rights**
- **Patents**
- **Public Domain: Shareware, Freeware**
- **Open Source Movement**
- **Website Contents: Linking, Deep Linking, Framing, Copying**
- **Copying Music, Movies, Images**

Copyright

Copyright Act 1968 (© not required in Australia)

- **Ownership of copyright in an original work is automatic**
- **May need to prove it at some time**
- **Rights: to make copies, sell, distribute, change, etc**
- **Works (expression of an original thought, idea): writing (prose, poetry, drama, etc), graphics, audio & video recordings, music, designs, software, ...**
- **Software made explicit in 1984**
- **Digital Agenda amendments 2001**
- **Australia is signatory to the 1952 Universal Copyright Convention (and the earlier Berne Convention, 1886)**
- **Moral rights: authorship acknowledged, content not distorted**
- **Duration: 70 years after death of author, 75 after creation for corporate works ("Mickey Mouse" provisions: 70 and 95 years)**
- **Key is potential value to author/creator**
- **Relationship to Patent**

Copyright - Complications

- **Contractual obligation** – may over-ride normal copyright
- **Employer rights** – based on terms of employment
- **Student rights** – based on University IP Policy
- **Shared rights** – where shared effort/resources are contributed
- **Using the resources of others** – gives them some rights
- **Insubstantial portions** – can quote small amounts from works
- **Quoting, Attribution** – give credit to author
- **Plagiarism** – deliberate or accidental use of others' works without attribution
- **Implied permission** – where the context suggests copying/distribution is expected
- **Temporary copying of electronically communicated works** – store-and-forward, caches, auto-backup, memory, computer screen
- **Fair Dealing** – for private use in study, research – limited amounts
- **Educational purposes under Section 113P (formerly part VB)** – special provisions for use in official courses, upon payment of a fee
- **Public domain software** – freely available, distributed
- **Shareware** – free to trial, distribute, not for long-term use
- **Licences** – over-ride, extend Copyright conditions

Copyright - Digital Amendments

Digital Agenda (2000)

- Mainly didn't change anything, just clarified
- New right of "communication"
- Applies to Emails, Web pages, etc
- It is now illegal to forward emails without permission.

Web Pages

- A Web page is a "work"
- Linking to another Web page – not an infringement
- "Deep linking" is this an infringement?
- Framing – making it look like it's yours
- "Passing off"
- Obtaining permission of owner – is it always required?
- Web page "terms of use" – must you observe these if they exist?

Linking/Framing/Tagging Issues

Website for The Shetland Times

https://www.shetlandtimes.co.uk/

Website for The Shetland News

https://www.shetnews.co.uk/

Settled out of court Nov 1997

See https://cyber.harvard.edu/property/metatags/link2.html

Framing: The Washington Post Co., et al. v. TotalNews Inc, et al, filed Feb. 2, 1997: see http://euro.ecom.cmu.edu/program/law/08-732/DomainNames/WashingtonPostVTotalNews.pdf

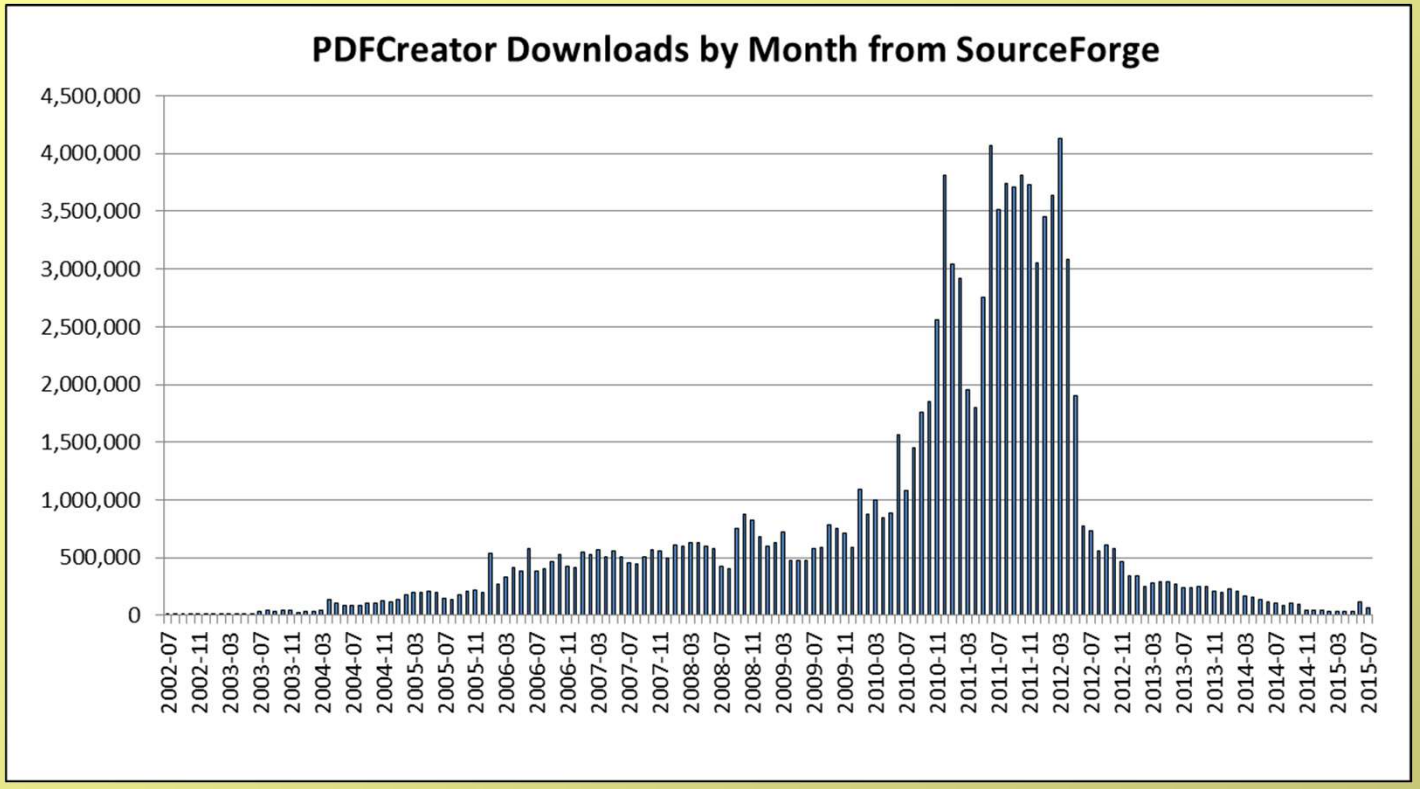
Both these (and others) were settled out of court, so we don't really know the full legal position – but what is the Ethical or Moral position?

Copyright - Open Source

- Open Source Movement – GNU https://www.gnu.org/ and Free Software Foundation https://www.fsf.org/
- Linux https://www.linuxfoundation.org/
- GNU General Public Licence (GPL) https://www.gnu.org/licenses/gpl-3.0.html
  May use the software freely
  - May copy & distribute sourcecode (with notice included)
  - May modify/add to it, but mustn't charge for it
  - Any added software attracts the same rights/conditions
- An ideological issue?
- A better way to develop software?
- An attempt to "dethrone" Microsoft? – see Peruvian Bill discussion https://www.theregister.com/2002/05/19/ms_in_peruvian_opensource_nightmare/.
- Munich City embraces Open Source https://opensource.com/government/14/5/how-munich-switched-15000-pcs-windows-linux (more recently, reverted to Microsoft).
- European Commission – eg "Pooling Open Source Software" Report https://joinup.ec.europa.eu/sites/default/files/document/2012-02/pooling-open-source-software-en.pdf.
- UK Government support – eg https://www.gov.uk/government/publications/procurement-policy-note-8-11-procurement-of-open-source.
- Websites to promote use of OSS – eg SourceForge https://sourceforge.net/

Open Source: PDF-Creator Statistics

Now nearly 650,000 products available via SourceForge; eg: PDF-Creator



From https://sourceforge.net/projects/pdfcreator/ [24-Jul-15]

Plus other sources of Free or Open Source Software (FOSS), such as Github (https://github.com/), and on Google & Apple App stores 90+% are free.

Software IP: Ethical Issues

So, what are the moral or ethical issues here?

- Is it ever OK to copy and use commercial software?
- Is it OK to "try out" or to "borrow" commercial software?
- Does it really matter if you don't follow licence terms for free software?
- Can't you argue that since the cost of copying/reproduction is near zero, then that's all you're depriving the owner of?
- Aren't the arguments for all software being free overwhelmingly strong?
- If the software owner is a big multinational corporation, doesn't that change the equation (given how predatory many of them are)?
- Etc (more on this later).

Who are the stakeholders, to whom do you owe any kind of duty?

Downloading Music MP3

ABC News Website 18-Nov-03

University Students Convicted of Music Piracy

From: https://www.abc.net.au/news/2003-11-18/suspended-sentences-over-music-piracy/1510900
See also https://www.smh.com.au/national/uni-students-accused-of-internet-music-piracy-20030202-gdg7cd.html

Example: Peer-Peer #1

Sample Defences of Illegal Music Downloads:

- Everyone's doing it
- We won't get caught
- The music industry charges too much
- They should make it impossible to copy
- It doesn't hurt anyone
- It only hurts a company, not a person
- Musicians are being exploited by multinationals
- The listening public is being exploited
- It helps increase sales
- Music should be free
- I can't afford to pay for it

Is "file sharing" always a "bad thing"?

Example: Peer-Peer #2

Ethical Tests:

- What laws govern the situation?
- Who gains and who suffers?
- Would you be happy for your action to be publicised?
- Would you tell your boss what you're doing?
- Would you tell your parents?
- What would you think if it was done to you?
- Does it violate Trust? Integrity? Truthfulness? Gratitude? Justice? Kindness?
- Are you treating others with respect?
- What if everyone did the same?

- Kabay: *The Napster Cantata*

http://www.mekabay.com/ethics/napster.htm

Downloading Music Can Be OK -1

Invitations to Obtain Free Music Download:

Kylie Minogue (2003), Karnivool (2010)

Downloading Music Can Be OK -2

Legal Downloads a Worldwide Hit.

Headline from IT Section of The West Australian, Tuesday, 26-Jul-05

Downloading Music Can Be OK -3

http://www.apple.com/itunes/1billion/
23-Feb-06
[no longer extant]

Report of 1 billionth iTunes music download, by Alex Ostrovsky in Feb-06

http://www.tuaw.com/2010/02/24/10-billionth-itunes-download-going-down-today/
24-Feb-10 [no longer extant]

Report of 10 billionth iTunes music download in Feb-10

Statistics of iTunes music & apps downloaded Jul-04 to Jan-13

Report of 25 billionth iTunes music download along with 40 billionth App download in Feb-13

http://techcrunch.com/2013/02/06/charting-the-itunes-stores-path-to-25-billion-songs-sold-40-billion-apps-downloaded-and-beyond/ 6-Feb-13

Downloading Music - Prevention

SDMI Challenge

- **Secure Digital Music Initiative**
  https://en.wikipedia.org/wiki/Secure_Digital_Music_Initiative
  - ❖ **"Unbreakable" Watermarking – 4 varieties (Steganography)**
  - ❖ **SDMI-compliant players**
  - ❖ **Make copies but not MP3-compressed copies for distribution**
- **Challenge – 6 September 2000 – Prize Money of $10,000**
- **Boycotted by some groups**
- **Princeton Group broke each coding scheme, but refused the prize**
  https://www.cs.princeton.edu/techreports/2002/657.pdf

Which approach do you think is right? Why?

1. **Boycott**
2. **Solve, publish and collect reward**
3. **Solve, publish and don't collect reward**
4. **Solve, don't publish and collect reward**
5. **Solve, don't publish, don't collect reward**

Should Music, Software be Free?

Some believe that all music should be free:

- Yes: "A true musician produces music because they love it, not because they're hoping to make money out of it."
- No: "If music were free then how will the musicians put food on their table?"

https://debatewise.org/998-music-should-be-free-for-all/

Some believe that all software should be free:

- Yes: "software should be written for the joy of helping others, not for money; and it costs nothing to make a copy, ie the 'cost of manufacture' is zero."
- No: "It costs time and effort to write software and developers have to make a living somehow."

https://www.gnu.org/philosophy/shouldbefree.en.html and https://medium.com/@pmhj42/why-software-should-be-free-counterargument-ef0373f61779

Some believe that all writing should be free…

Photoshopping Images

Celebrity Pics

Matthew Macfadyen
Michael Phelps
Lesley Garrett
Avril Lavigne

https://www.boredpanda.com/before-after-photoshop-celebrities/?utm_source=google&utm_medium=organic&utm_campaign=organic

Enhancing Photos

Cartoon depicting someone getting a whole range of enhancements done to his photos when developed

www.tedgoff.com

Enhancing Photos

Cartoon depicting a doctor saying the broken rib in the X-ray can be fixed by photoshopping.

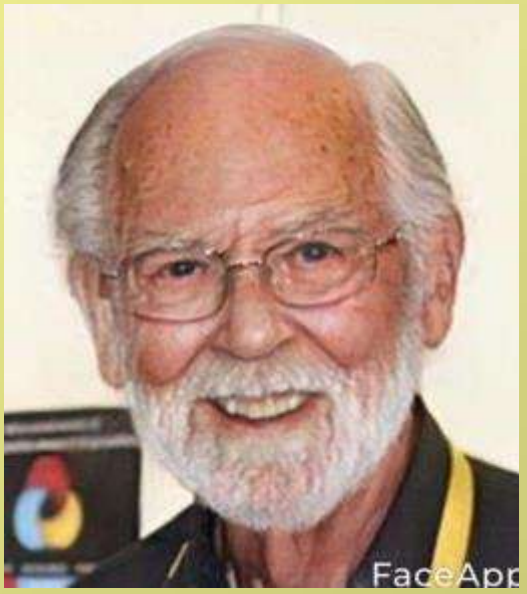
www.funny-joke-pictures.com

Enhancing Photos

Cartoon depicting a child saying the grades in his report card can be fixed by photoshopping.

www.cartoonstock.com

Doctoring Images



Using the first (real) image, FaceApp provides variations based on certain criteria…

Case Study 4

Digital Photograph Manipulation

- It's simple now for various forms of image "enhancement" to be made, eg:
  - Red-eye elimination
  - Cropping
  - Special effects (eg sepia-colour)
  - Wrinkle removal
  - Changing the contents in significant ways
- Is there anything wrong with "touching up" an image?
- What kind of "touching up" might be OK, in what circumstances? What might be wrong? Why?
- Is there anything wrong with the way the latest Pixel phone is promoted – its image manipulation ability?

AI Voice Cloning:

- eg lyrebird.ai will clone your voice given 10 minutes of samples.
- Where might you legitimately use your own voice clone?
- Where might you use it to deceive?
- What about restoring voices to people who have lost theirs (eg Val Kilmer to throat cancer, used AI in 'Top Gun: Maverick')?
- What about recreating the voices of people who have died?
- What about making people appear to say things they didn't/wouldn't?
- Most people are happy to communicate with Siri, Alexa, Google Assistant – is it any different to be hearing the voice of someone famous, or someone you know? Reinforcing female subservience?

- In a 2022 podcast, Waleed Aly had 3 objections to dead voice cloning:
  - ✡ Goulish, dishonest (hollows out the original);
  - ✡ Impulse to remove pain (emotional stunting, sedating);
  - ✡ Mythology of techtopia (tech creates a better world).

  https://www.abc.net.au/radionational/programs/theminefield/should-voice-assistants-use-the-voices-of-our-loved-ones/13992386

Deep Fake Videos

Video of The Queen giving her Christmas message, Channel 4 version.

https://www.youtube.com/watch?v=IvY-Abd2FfM

Video of Tom Cruise being faked.

https://www.youtube.com/watch?v=nwOywe7xLhs

Video of Barack Obama saying things he'd never actually say.

https://www.youtube.com/watch?v=cQ54GDm1eL0

Video of Volodymyr Zelenskyy appearing to tell soldiers to surrender to Russia.

https://www.youtube.com/watch?v=X17yrEV5sl4

Some Issues Highlighted by Use of AI:

- Autonomous cars: programming for emergencies.
- Deep fake videos: making people do things they haven't.
- Doctored sound files: making people say things they haven't.
- Art creation (eg Craiyon.com, Bing Image Creator): passing off as your creation.
- Google's LaMDA (Language Models for Dialog Applications): Is it sentient? Does it matter?
- Facial recognition: an invasion of privacy?
- Use in the military: "do no harm"?

Eu… *…lligence Act*, 27 February 2024
<https://artificialintelligenceact.eu/>

AI Ethics Principles

1. **Human, Social & Environmental Wellbeing.**
2. **Human-Centred Values.**
3. **Fairness (inclusive, accessible, non-discriminatory).**
4. **Privacy Protection and Security.**
5. **Reliability and Safety.**
6. **Transparency and Explainability (but what about deep learning?).**
7. **Contestability.**
8. **Accountability.**

Australian Government AI Ethics Principles (2019):
https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles

CSIRO Discussion Paper on AI Ethics Framework (April 2019):
https://www.csiro.au/-/media/D61/Reports/Artificial-Intelligence-ethics-framework.pdf

ACS Primer: The Ethics and Risks of AI Decision-Making (June 2021):
https://www.acs.org.au/content/dam/acs/acs-publications/ACS_The-Ethics-and-Risks-of-AI-Decision-Making.pdf

The EU Artificial Intelligence Act (February 2024): https://artificialintelligenceact.eu/

AI Ethics Case Studies

1. Identifying de-identified data

In 2016, a dataset that included de-identified health information was uploaded to data.gov.au. It was expected that the data would be a useful tool for medical research and policy development. Unfortunately, it was discovered that in combination with other publicly available information, researchers were able to personally identify individuals from the data source. Quick action was taken to remove the dataset from data.gov.au.

2. Houston teachers fired by automated system

An AI was used by the Houston school district to assess teacher performance and in some cases fire them. There was little transparency regarding the way that the AI was operating. The use of this AI was challenged in court by the teacher's union, as the system was proprietary software and its inner workings were hidden. The case was settled and the district stopped using it.

3. The COMPAS sentencing tool

COMPAS is a tool used in the US to give recommendations to judges about whether prospective parolee will re-offend. There is extensive debate over the accuracy of the system and whether it is fair to African Americans. Investigations by a non-profit outlet have indicated that incorrect predictions unfairly categorise people of colour as a higher risk. The system is proprietary software.

Source: https://www.csiro.au/-/media/D61/Reports/Artificial-Intelligence-ethics-framework.pdf

AI Ethics Case Studies

4. **Cambridge Analytica & Facebook Data** – page 28
5. **Equifax data breach** – page 29
6. **Locating People with Geoprofiling** – page 31
7. **Microsoft Chatbot** – page 31
8. **Amazon Same-Day delivery** – page 32
9. **Houston Teachers** – page 34
10. **Enbridge Pipeline Leak** – page 35
11. **Automated vehicles** – page 36
12. **Israeli Judges and Decision Fatigue** – page 38
13. **Amazon Hiring Tool** – page 42
14. **US Predictive Policing** – page 42
15. **Brisbane Predictive Policing** – page 43
16. **Predicting Coma Outcomes** – page 44
17. **Manipulating User Moods** – page 46
18. **Surveillance Technology** – page 53

From: CSIRO: Artificial Intelligence - Australia's Ethics Framework Discussion Paper

How is AI Different?

- Fundamentally, it is no different – apply the same approach, eg identify stakeholders, work out the "social contracts" involved, etc.
- But it is more complicated:
  - Separation in time and space from "author" is greater;
  - Very unclear who is responsible;
  - The use of deep machine learning can mean that no-one understands how decisions are reached.
- Attributing blame if something goes wrong can be complicated.
- What are the AI designer's responsibilities here?

Facial Recognition:

- **Bunnings, Kmart, Good-Guys attempt to pilot it 2023:**
  - **Rampant rise of shop-lifting;**
  - **2 called a halt after Choice raised concerns;**
    - **Lack of notice and disproportionate to the harm to privacy;**
  - **Others said they don't do it;**
  - **What if they retrospectively used AI to analyse weeks' worth of CCTV?**
  - **See: https://www.itnews.com.au/news/bunnings-and-kmart-facial-recognition-probe-set-to-finish-by-july-590881**
- **Complaint is based on Privacy, but Australia's privacy laws are complicated.**
- **Necessarily an invasion of privacy?**
  - **Is it actually "creepy and invasive"?**
- **Extends the use of CCTV.**
- **China's use for watching behaviour (Skynet), eg to enforce the Social Credit System:**
  https://en.wikipedia.org/wiki/Mass_surveillance_in_China

Image of facial recognition, along with Bunnings & Kmart logos

Image of CCTV cluster

AI Ethics Issues

AI Art Creation (eg using Craiyon.com):

- A Galaxy Far Far Away:

- Dog and Sunflowers, style of Van Gogh:

- Suspension Bridge, style of Mondrian (Craiyon vs Bing)

Who is the author/artist? Who owns them? Who has copyright? Can you sell them as works of art?
(Sothebys sold a portrait of Alan Turing for $1.64m in Nov 2024).

Then there's Non-Fungible Tokens (NFTs):
let's not even go there!

AI-Generated Essays

- Like the Essay Mills, but tuned to the cheater's own style.
- Is it any different from using calculators in exams?
- Use AI itself to detect cheating.
- ChatGTP said: "Cheating is lame no matter how you do it. Play fair or don't."
- ChatGTP also suggested: "Make the exam more difficult so that students can't simply memorise the answers."
- Phillip Dawson: 'Defending Assessment Security in a Digital World', Routledge, 2020
  https://ebookcentral.proquest.com/lib/uwa/detail.action?docID=6340937

Dangers of Generative AI

- Inbuilt bias based on the data they are trained on – racism, sexism, etc
- Unemployment, Job Displacement
- Unequal wealth distribution
- Theft – of art, literature, music, identity, etc (Copyright & IP)
- Impersonation & Fraud, Deepfakes
- Privacy Concerns
- Misinformation, Disinformation, Fake Information
- Hallucinations
- Provenance, Transparency
- Accountability & Responsibility
- Automation of Harmful Activities, eg Spam
- Easy Access to Dangerous Information (eg how to build a bomb)
- Hacked Systems for Malicious Purposes
- Sentient Machines and the Existential Threat
- Threat of Model Collapse

AI Ethics Issues

Cartoon featuring Deep Learning to find data that agree with your opinion.
From https://timoelliott.com/blog/cartoons/artificial-intelligence-cartoons

Automated Service Decisions:

- **Bank loan applications, etc.**
- **What recourse do we have?**
- **Can we get information about their decision-making criteria?**
- **Is this big business hiding behind "computers"?**
- **"Inadvertent" bias in AI systems, because they've learnt from a biased sample.**
- **US Algorithmic Accountability Act: conduct critical impact assessments.**
- **Employ diverse development teams.**

Image from Little Britain with computer operator reporting that the "Computer Says No"

https://www.pinterest.com.au/pin/1150529654512773 40/

Military Uses of AI

Image of a vast army of robot warriors

- **US Dept of Defense Ethical Principles (25-Feb-20):**
  - **Responsible**
  - **Equitable**
  - **Traceable**
  - **Reliable**
  - **Governable**

https://www.defense.gov/News/News-Stories/Article/Article/2094085/dod-adopts-5-principles-of-artificial-intelligence-ethics/

- **AI is the US Defense Dept's top technology modernisation priority.**
- **The first to master AI will prevail on the battlefields of the future.**
- **Biggest issue is autonomous target recognition.**
- **Adversaries may not share the same ethical approach to offensive weapons** (eg Ethiopia using Turkish drone to kill 60 civilians in Jan-2022: https://www.washingtonpost.com/world/interactive/2022/ethiopia-tigray-dedebit-drone-strike/).

Creating "Sentient" Systems:

- **Turing Test (TT)**
- **Total Turing Test (TTT) – Harnad – includes non-verbal behaviour.**
- **Truly Total Turing Test (TTTT) – Schweizer – includes "extended and multifarious interactions with human beings generally."**
- **Kurtzweil's "Singularity" – AGI (2060??)**
- **Google's LaMDA: if it "appears" to be intelligent, is that any different from it being intelligent? Is simulated "thinking" the same as thinking? Blake Lemoine interview with 'sentient' LaMDA: https://www.scientificamerican.com/article/google-engineer-claims-ai-chatbot-is-sentient-why-that-matters/**
- **Is it ethical to not say whether the "help desk" is human or machine?**

See: Re-Engineering Humanity, Brett Frischmann & Evan Selinger, Cambridge, 2018.

LaMDA Transcript: https://s3.documentcloud.org/documents/22058315/is-lamda-sentient-an-interview.pdf

AI Ethics Issues

Image of Tesla car able to order its own repairs.

Maybe in the future, it will order itself a new set of tyres, and new paint job, and charge you for Child Support for a Chevy Volt.

Viruses – What's Our Responsibility?

Cartoon depicting someone being dragged away by a monster, after opening an email attachment.

www.tedgoff.com

Will Virus Ruin Your Computer Too?

Cartoon depicting someone asking if a colleague can see if a virus on a floppy disk also ruins their computer.

www.tedgoff.com

Responsibility for Virus Protection

- To establish whether staff are clicking on phishing attempts or not, you could design a *test* - send a false phishing email around and see how many clicked on it.
- Eg Belgian Government in 2015: but it went badly wrong because many people contacted the travel company to complain about being asked to provide credit card details, but that company knew nothing about it (no-one had cleared it with them).

  See https://www.computerworld.com/article/1627759/belgian-government-phishing-test-goes-offtrack.html.
- A similar incident in the US military in 2014 - http://www.washingtonpost.com/politics/gone-phishing-army-uses-thrift-savings-plan-in-fake-email-to-test-cybersecurity-awareness/2014/03/13/8ad01b84-a9f3-11e3-b61e-8051b8b52d06_story.html.

Bunratty Attack

- 1997 COSAC Conference in Bunratty, Ireland (Computer Security Audit & Control Symposium).
- Standard ("innocent") email messages.
- Utilises standard Messaging API.
- Utilises hidden folders.
- All hidden from user - eg as for Calendar updates.
- Covert, asynchronous, remotely upgraded, remotely removed.
- Defence requires code on every client to identify false messages.
- I-Love-You Virus (followed by the Kournikova Virus) based on some of the same vulnerabilities, but not all.
- What would you do?
  1. Keep as quiet as possible?
  2. Tell Microsoft under a veil of secrecy?
  3. Publicise as widely as possible to ensure something is done?
  4. Take some other action? What?

Hoax Virus

- **Targets naïve users**
- **Exploits unusual icon for system file**
- **Advises user to delete file**
- **Advises user to forward to everyone they know**

Subject: BAD virus - act quickly!!
Date: Tue, 29 May 2001 21:57:22 -0400
Subject: Please Act Urgently
VIRUS COULD BE IN YOUR COMPUTER
It will become activate on June 1st and will delete all files and folders on the hard drive.
No Anti-Virus software can detect it because it doesn't become a VIRUS until 1/6/2001.
It travels through the e-mail and migrate to your computer.
To find it please follow the following directions:
Go To "START" button
Go to "Find" or "Search"
Go to files and folders
Make sure to search in drive C
Type in; SULFNBK.EXE
Begin Search
If it finds it, highlight it and delete it
Close the dialogue box
Open the Recycle Bin
Find the file and delete it from the Recycle Bin
You should be safe.
The bad part is you need to contact everyone you sent ANY e-mail to in the past few months.

Fake News and Internet Hoaxes:

- Especially via social media.
- How to identify Fake News: https://beconnected.esafety.gov.au/articles-and-tips/how-to-spot-fake-news
- Hoax-busting sites:
  - http://www.snopes.com/ or
  - https://www.truthorfiction.com/ or
  - See http://hoaxbusters.org/ (now closed, but still with some useful links) or
  - Hoax Slayer https://en.wikipedia.org/wiki/Hoax_Slayer (site closed on 31/5/21).

2 Types of Hacker

- OLD-TIME ("white hat"):
  - Clever, addicted, insatiable quest for knowledge, a cooperating community, advancing the cause of effective computer programming, development and use.
  - CERT – Computer Emergency Response Team
  - "Hackathons"
- MODERN (generally "black hat"):
  - Gaining access to "private" computers
  - Beating the "system"
  - Electronic graffiti
  - Personal gain, theft, data alteration, etc
  - The Hackers Handbook (1985) – Cornwall/Sommer
  - International crime
  - Espionage
  - The Cuckoo's Egg (1990) – Clifford Stoll
  - Vandalism
  - "Denial of Service" attacks

Hacker "Ethics" and Rationale

- **Ethics:**
  - **All information should be free**
  - **Access to computers should be unlimited and total**
  - **Mistrust authority – promote decentralisation**
  - **Judge hackers by their skill**
  - **True hackers create art and beauty**
  - **Computers can change your life for the better**

  - Levy: *Hackers: Heroes of the Computer Revolution*

  [see Open Source Initiative]

- **Rationale:**
  - **We're helping to improve security**
  - **It's the fault of the software vendors**
  - **It's the fault of slack security**
  - **We're not doing any harm**
  - **No-one will listen unless we take action**
  - **It helps keep Big Brother at bay**

  [cf justification offered by Assange, Snowden]

Competent Advice

Cartoon depicting a janitor answering the Tech Support phone after hours, offering a range of technical advice.

www.tedgoff.com

AI Twitter/X Case Study

Twitter/X Analysis Using Big Data & ML:

- You work as an analyst for a big-data Internet consultancy firm.
- Your firm has been commissioned to undertake research into social media usage to highlight the topics covered, extent of re-Tweeting, etc.
- You invite Twitter users to allow you see their Tweets over the past 2 years; some 10,000 users agree to this (provided that their identity be hidden in any published results). You harvest these 200,000 tweets.
- You train an ML system on this data to discover how often and the pattern by which Tweets get repeated by other users, in part or in whole.
- You also obtain anonymised Tweets from all Twitter users over the past 2 years, totalling 100m tweets from 5m users; unfortunately, in all re-Tweets the re-Tweeted user-id has not been removed (you don't realise this yet).
- By using a big data science application, along with the ML system, looking for repeated (eg re-Tweeted) Tweets or part-Tweets, you have been able to identify a large percentage of the anonymous cohort (ie find out their real Twitter handles).
- With this information, the value of research could be vastly improved.

AI Twitter/X Case Study

Twitter/X Analysis Using Big Data & ML :

What do you do when you realise the Tweet data has not been properly anonymised?

1. **Scrub the data to remove identifiers and then run your analysis on properly anonymised data.**
2. **Immediately discard the faulty data, advise the supplying agency, and request the data again.**
3. **Run your analysis just to see what it would produce, but then discard the faulty data.**
4. **Run your analysis and produce and publish your results (not mentioning the data breach).**
5. **Advise your managers and let them decide what to do.**
6. **Something else?**

Judge Fatigue Case Study

Judge Fatigue vs ML System:

- An analysis of judicial parole hearings revealed that there was a distinct pattern correlating severity of sentences with certain times of the day (ie more lenient early in the morning or straight after lunch, more severe late in the morning or late in the afternoon).
- An ML system was trained to analyse all the factors that a judge would take into account along with the judgement in 1,000 cases; those cases considered late in the morning or afternoon were omitted.
- Because it was a deep learning system, evidence as to why certain decisions were recommended could not be provided, but it did eliminate the "judge fatigue" pattern.

Should this system be implemented to replace judges?

1. No, not until its explicability/transparency can be much improved.
2. No, but it should be used to advise judges, who could take its advice or not.
3. Yes, but not until it has been trained on very many more cases.
4. Yes, but a judge should be able to overrule it.
5. Yes, but the appellant can have the right to appeal to a human judge.
6. Yes, its accuracy is clearly much better than human judges.

Synthetic Voice Case Study

Synthetic Voice for AI System:

- You have implemented an AI-based Home Assistant system that responds to voice commands.
- You have chosen a female voice to use, based on research that shows that more people respond to and trust a female voice.
- Your own research shows that an overwhelming majority of the users of this system do like this voice.
- However, this decision comes under much criticism from a vocal lobby that claims that this reinforces the female "housewife" stereotype.

What action should you take?

1. Ignore the criticism as it reflects only a tiny minority.
2. Invite the lobby group to engage in a public debate about the merits or otherwise of this female voice.
3. Modify the voice to make it more assertive, less subservient.
4. Modify the voice to make it more "anodyne" (ie of indeterminate sex).
5. Create male and female versions and make a decision which to use based on user feedback.
6. Create male and female versions and give users the choice of which to use.
7. Something else?

Other Relevant Case Studies

- A number of actual situations can be found in ACS Code of Professional Conduct Case Studies, with relevant sections of the Code identified – see https://www.acs.org.au/memberships/acs-ethics-education-program.html (scroll down to "ACS Case Studies", others within the Education Program itself).
- Several good case studies are presented in the context of the ACS Code of Ethics in the *Information Age* article below.
- Students are strongly encouraged to read these case studies.
- Burmeister, Oliver K: "Applying the ACS Code of Ethics", *Information Age*, Feb/Mar 2001, pp54-59, and in the subsequent 3 issues (Apr/May, Jun/Jul, Aug/Sep, 2001). Also published as: Burmeister, Oliver K: "Applying the ACS Code of Ethics", *Ethics in Computing*, v32, n2, May 2000, pp107-119.
- This analysis is based on that which first appeared in 1993 as follows:

  Anderson, Ronald E et al: "Using the New ACM Code of Ethics in Decision Making", *Communications of the ACM*, v36, n2, Feb 1993, pp98-106.
- A selection of thought-provoking case studies were published in *Information Age* in Oct/Nov 2018 – see https://ia.acs.org.au/article/2018/ethics-part-1--artificial-influencers.html.
- Other helpful case studies can be found in Bynum, Terrel Ward & Rogerson, Simon, eds: "Computer Ethics & Professional Responsibility: Introductory Text & Readings", Blackwell, 2004

Case Study Survey

CITS3200 Ethics Case Studies:

- This is a set of simple Case Studies designed to help you understand some of the Ethical Issues you may face as a computer professional.
- Your responses will be recorded and aggregated with others from the 2025 CITS3200 Class (they will be kept anonymous).
- Average responses will be published in a table later this Semester (with comparisons from earlier years).
- You may suspend answering at any time and resume later.

This Survey is available here:

http://uwa.qualtrics.com/jfe/form/SV_6DnHaKB3CTTZLuZ

Or can be accessed via the CITS3200 Website

https://teaching.csse.uwa.edu.au/units/CITS3200/resources.html

Please attempt this Survey within the next week.

END OF LECTURES

Bibliography

https://alex-reid.com/Ethics/Computer-Ethics-Bibliog.html