CP5806

WEEK 5

JAMES COOK UNIVERSITY
AUSTRALIA

DUE DATES

- Assignment 2 due Sunday 9 August, 11:59 pm (Week 5)
- Assignment 3 due Wednesday 19 August, 11:59pm (Week 7)
  - Sisi will hold an extra session Saturday 15 August, 2:30-3:30pm (Week 6)

AUGUST							
	M	T	W	T	F	S	S
wk 4						1	2
wk 5	3	4	5	6	7	8	9
wk 6	10	11	12	13	14	15	16
wk 7	17	18	19	20	21	22	23
O Week	24	25	26	27	28	29	30
wk 1	31						

JAMES COOK UNIVERSITY AUSTRALIA

A2 STEP 8 ADDITIONAL EXAMPLE

STEP 8: DATA CUBE COMPUTATION (5%)

- Assume a base cuboid of 5 dimensions contains two base cells (a1, a2, a3, a4, a5) and (b1, b2, a3, a4, a5), measure is count;

a) How many nonempty aggregate (i.e., nonbase) cells will a full cube contain?

c) How many nonempty aggregate cells will an iceberg cube contain if the condition of the iceberg cube is "count $\geq$ 2"?

d) A cell, c , is a closed cell if there exists no cell, d, such that d is a specialisation of cell c (i.e., d is obtained by replacing a $*$ in c by a non-$*$ value) and d has the same measure value as c. A closed cube is a data cube consisting of only closed cells.

A2 STEP 8 ADDITIONAL EXAMPLE

STEP 8: DATA CUBE COMPUTATION (5%)

- Assume a base cuboid of 5 dimensions contains two base cells (a1, a2, a3, a4, a5) and (b1, b2, a3, a4, a5), measure is count;

Solutions:

a) For each base cell, 2 ^ 5 - 1 aggregate cells:

1: (*, *, *, *, *);

5: (a1, *, *, *, *),(*, a2, *, *, *),(*, *, a3, *, *),(*, *, *, a4, *),(*, *, *, *, a5);

10: (a1, a2, *,*,*), (a1, *, a3, *, *), (a1, *, *, a4, *), (a1, *, *, *, a5),(*, a2, a3, *, *), (*, a2, *, a4, *), (*, a2, *, *, a5), (*,*,a3,a4,*),(*,*,a3, *, a5), (*,*,*,a4,a5);

10: (a1,*,*,a4,a5), (a1,a2,*,*,a5), (a1,a2,a3,*,*), (a1,*,a3,*,a5), (a1,*,a3,a4,*)(a1,a2,*,a4,*), (*, a2,*,a4,a5), (*, a2,a3,*,a5), (*,a2,a3,a4,*),(*,*,a3,a4,a5);

5: (a1,a2,a3,a4,*), (a1,*,a3,a4,a5),(a1,a2,*,a4,a5), (a1,a2,a3,*,a5) ,(*,a2,a3,a4,a5);

Current total: 2 * (2 ^ 5 - 1)

There are 2 ^ 3 cells calculated twice:

(*,*, a3, a4, a5), (*,*,a3,a4,*), (*,*,a3,*,a5),(*,*,*,a4,a5), (*,*,a3,*,*), (*,*,*,a4,*), (*,*,*,*,a5), (*,*,*,*,*)

Final aggregate cells: 2 ^ 6 - 2 - 8 = 2 ^ 6 - 10

A2 STEP 8 ADDITIONAL EXAMPLE

STEP 8: DATA CUBE COMPUTATION (5%)

- Assume a base cuboid of 5 dimensions contains two base cells (a1, a2, a3, a4, a5) and (b1, b2, a3, a4, a5), measure is count;

Solutions:

c) Closed cell: no descendant cell d that has the same measurement value, e.g. cell (a1, a2, a3, a4, a5) is a descendant of cell (a1, *, *, a4, a5) or cell (*,*,*, a3,a4,a5) is a descendant of cell (*, *,*,*,a5) etc.

2 base cells are closed cells as they do not have descendant cells;

(*, *, a3, a4, a5) is a closed cell as all its descendant cells, i.e. (a1, *,a3,a4,a5) or (*, b2,a3,a4,a5) only have a count of 1;

A2 STEP 8 ADDITIONAL EXAMPLE

STEP 8: DATA CUBE COMPUTATION (5%)

1. When computing a cube of high dimensionality, we encounter the inherent curse of dimensionality problem: there exists a huge number of subsets of combinations of dimensions.

   (a) **(2 val.)** Suppose that there are only two base cells, namely $(a_1, a_2, a_3, a_4, a_5, a_6, \ldots, a_{100})$ and $(a_1, a_2, a_3, a_4, a_5, b_6, \ldots, b_{100})$, in a 100-dimensional base cuboid, each with a cell count of 100. Compute the number of nonempty *aggregate* cells (note that, for example, $(a_1, a_2, a_3, \ldots, a_{100})$ is not considered an aggregate cell).

   Solution:

   Each base cell generates $2^{100} - 1$ aggregate cells. We subtract 1 because $(a_1, a_2, a_3, \ldots, a_{100})$, for example, is not an aggregate cell. Thus, the two base cells generate $2 \times (2^{100} - 1) = 2^{101} - 2$ aggregate cells.

   However, several of these cells are counted twice. In particular, any cell that aggregates over the dimensions 6 to 100 will be counted twice. There is a total of $2^5 = 32$ such cells. Therefore, the total number of cells generated is $2^{101} - 34$.

   (b) **(1 val.)** Suppose we are to compute an iceberg cube from the above. If the minimum support count in the iceberg condition is 20, how many aggregate cells will there be in the iceberg cube? Indicate five such cells.

   Solution:

   All cells in the previous question have a cell count of at least 100. Therefore, the total number of aggregate cells in the iceberg cube will be $2^{101} - 34$. Some such cells: $\{(a_1, a_2, a_3, a_4, a_5, * \ldots, *)$, $(a_1, *, \ldots, *)$, $(*, a_2, *, \ldots, *)$, $\ldots$, $(*, *, *, *, a_5, *, \ldots, *)$, $(*, *, \ldots, *)\}$.

A2 STEP 8 ADDITIONAL EXAMPLE

Problem 1. (**23** points total)
Suppose the base cuboid of a data cube contains two cells:
$(a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}) : 1, (\underline{a_1}, b_2, \underline{a_3}, b_4, b_5, b_6, b_7, b_8, \underline{a_9}, b_{10}) : 1$
where $a_i \neq b_i$ for any i.

(a) (3 points) How many cuboids are there in this data cube?

Answer: 2^{10}. Since we have 10 dimensions with no concept hierarchy, there are 2^{10} cuboids and all of them should not be empty.

(b) (5 points) How many (nonempty) closed cells are there in this data cube?

Answer: 3. Two base cells and $(a_1, *, a_3, *, *, *, *, *, a_9, *)$.

(c) (5 points) How many (nonempty) aggregate cells are there in this data cube?

Answer: 2038. For each base cell, there are $2^{10} - 1$ aggregated cells. However, there are $8 = 2^3$ cells that are counted twice since there are 3 common dimensions. Therefore, the total number of nonempty aggregate cells is $2 \times (2^{10} - 1) - 2^3 = 2038$.

(d) (5 points) How many (nonempty) aggregate closed cells are there in this data cube?

Answer: 1. $(a_1, *, a_3, *, *, *, *, *, a_9, *)$.

(e) (5 points) If we set minimum support = 2, how many (nonempty) aggregate cells are there in the corresponding iceberg cube?

Answer: 8. These two base cells have common value in 3 dimensions; therefore, there are $2^3 = 8$ nonempty cells with support = 2 and all of them are aggregate cells.

WEEK 5 LEARNING OUTCOMES

- Give examples of breaches of ethics in IT environments
- Use codes of ethics to make ethical decisions
- Review scenarios and identify ethical issues associated with security
- Classify different security crime types
- Apply privacy preserving techniques
- Summarise privacy issues related to freedom of expression and intellectual property.

JAMES COOK UNIVERSITY AUSTRALIA

SOME NOTES ON ASSESSMENT 1

- Some people did not properly adhere to APA style:
  - See: https://libguides.jcu.edu.au/apa6th/home
  - Slightly nit-picky but in the reference list:
    - *Titles should be written in sentence case*
    - *NOT In Title Case*
      - Except for journals, e.g. *Journal of Title Case Studies*
- Avoid statements like: Smith, writing in *Some Book of Facts*, contends *XYZ* (Smith, 1999).
  - Instead: Smith (1999) contends *XYZ*.

SOME NOTES ON ASSESSMENT 1

- See https://libguides.jcu.edu.au/writing/writing4c for a guide on academic language
  - Academic writing shouldn’t use contractions or slang.
- Make sure to engage with the question asked, and all parts of it, and explicitly take a stance when asked to do so.
  - E.g. in task 2, most people disputed the specific percentages that Ackoff claimed, but some did not take a stance on the general question of whether there can be more information than data, etc.

TOPICS FOR WEEK 5

- Topic 1: Overview of ethics
- Topic 2: Ethics for IT workers and employers
- Topic 3: Security and cybersecurity
- Topic 4: Privacy
- Topic 5: Freedom of expression and IP

JAMES COOK UNIVERSITY AUSTRALIA

ETHICAL QUESTIONS

- Did *you* install the CovidSafe app?
- A New York Times technology writing was recently working on a story about a blog called *Slate Star Codex*. The author, a psychiatrist, goes by a pseudonym, Scott Alexander. NYT policy is to identify public figures by their legal names. Alexander claims this would be a security risk for him, and a risk to his patients. Is it ethical for the NYT to “dox” him?

TOPIC 1: OVERVIEW OF ETHICS

- **Morality** refers to social conventions about right and wrong that are so widely shared that they become the basis for an established consensus.
- **Ethics** is a set of beliefs about right and wrong behavior within a society.
- **Virtues** are habits of acceptable behavior.
- A person who acts with integrity acts in accordance with a personal **code of principles.**
- **Morals** are one's personal beliefs about right and wrong.
- **Law** is a system of rules that tells us what we can and cannot do.
- **Section 406 of the Sarbanes-Oxley Act** requires public companies to disclose whether they have codes of ethics and disclose any waiver to their code of ethics for certain members of senior management.
- The goal of the Sarbanes–Oxley Act was to **renew investor's trust in the content and preparation of disclosure documents by public companies.**
- **Code of ethics** highlights an organization's key ethical issues and identifies the overarching values and principles that are important to the organization and its decision-making process.

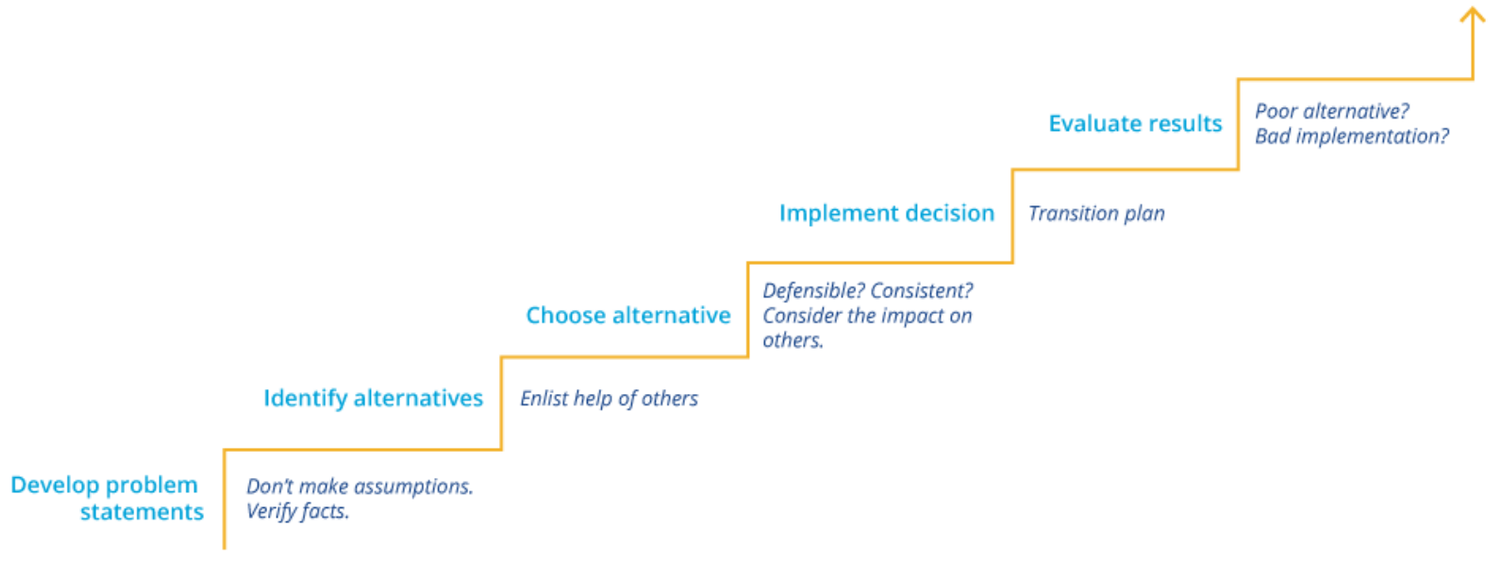
TOPIC 1: OVERVIEW OF ETHICS

- **Bathsheba syndrome -** The moral corruption of people in power
- **Common good approach** - Based on a vision of society as a community whose members work together to achieve a common set of values and goals
- **Fairness approach** - Focuses on how fairly actions and policies distribute benefits and burdens among people affected by the decision
- **Integrity** - Acting in accordance with a personal code of principles
- **Utilitarian approach** - Focusses on the ethical choice that produces the greatest excess of benefits over harm

JAMES COOK UNIVERSITY AUSTRALIA

ETHICAL DILEMMA 1

- **Give a salary increase?**
- You are currently being considered for a major promotion within your company – to vice president of marketing. In your current position, as manager of advertising, you supervise 15 managers and 10 hourly workers.
- As part of the annual salary review process, you have been given the flexibility to grant your employees an average three per cent annual salary increase; however, you are strongly considering a lower amount.
- This would ensure that your department's expenses stay under budget and would send the message that you are able to control costs.
- How do you proceed?
- Propose your course of action, giving a rationale for your decision.
- You may wish to consider the five-step decision-making process.



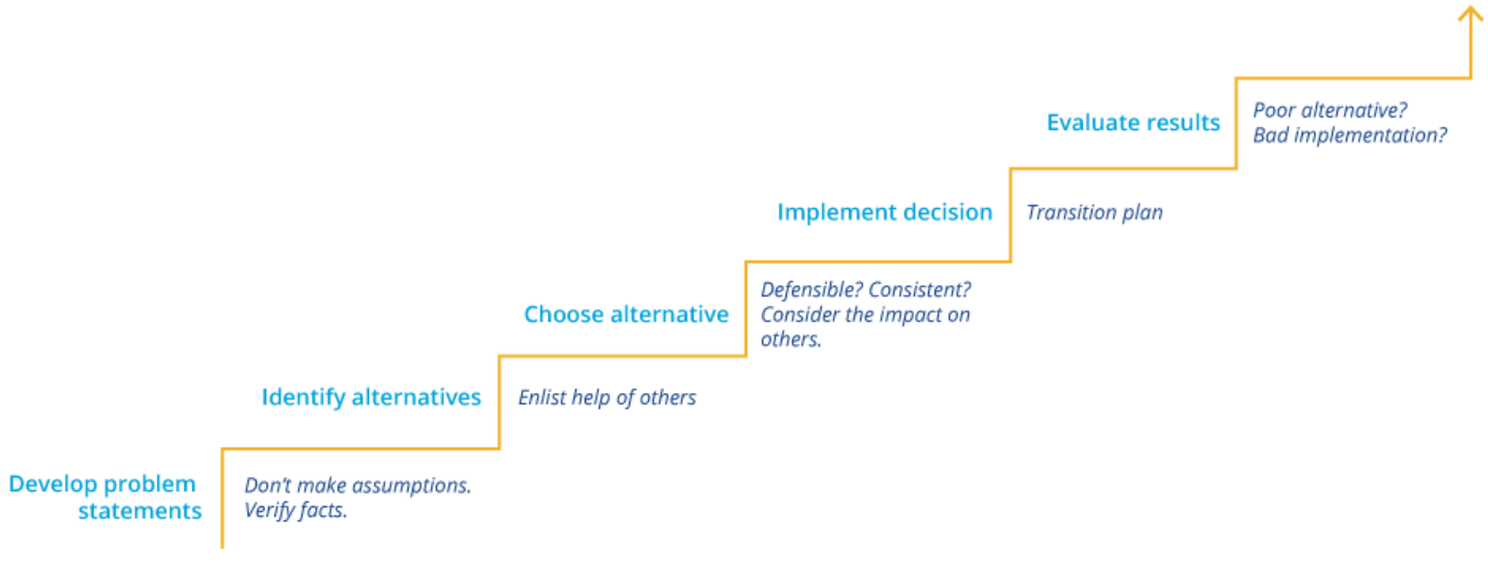
TOPIC 2: ETHICS FOR IT WORKERS AND EMPLOYERS

- A **professional** is someone who:
  - requires advanced training and experience
  - must exercise discretion and judgment in the course of his or her work
  - does work that cannot be standardized
- The mission of the **Business Software Alliance** is to stop the unauthorized copying of software produced by its members
- **Whistle-blowing** is an effort by an employee to attract attention to a negligent, illegal, unethical, abusive, or dangerous act by a company that threatens the public interest.
- **Fraud** is the crime of obtaining goods, services, or property through deception or trickery.
- **Compliance** means to be in accordance with established policies, guidelines, specifica- tions, or legislation.
- Society expects professionals to act in a way that:
  - causes no harm to society
  - provides significant benefits
  - establishes and maintains professional standards that protect the public
- **Certification** is a process that one undertakes voluntarily to prove competency in a set of skills.

JAMES COOK UNIVERSITY AUSTRALIA

ETHICAL DILEMMA 2

- **Exaggerated claims?**
- Your old roommate from college was recently let go from his firm during a wave of employee terminations to reduce costs. You two have kept in touch over the six years since school, and he has asked you to help him get a position in the IT organisation where you work.
- You offered to review his résumé, make sure that it gets to the “right person,” and even put in a good word for him. However, as you read the résumé, it is obvious that your friend has greatly exaggerated his accomplishments at his former place of work and even added some IT-related certifications you are sure he never earned.
- What would you do?
- Propose your course of action, giving a rationale for your decision.
- You may wish to consider the five-step decision-making process.



TOPIC 3: SECURITY AND CYBERSECURITY

- An **attack** on an information system that takes advantage of a vulnerability is called exploit.
- **Virtualization** software operates in a software layer that runs on top of the operating system and enables multiple virtual machines each with their own operating system to run on a single computer.
- A **zero-day attack** takes places before the security community or software developer knows about the vulnerability or has been able to repair it.
- Software that generates and grades tests that humans can pass but that all but the most sophisticated computer programs cannot is called **CAPTCHA**.
- The written statement that defines an organization's security requirements as well as the controls and sanctions used to meet those requirements is known as a **security policy.**

JAMES COOK UNIVERSITY AUSTRALIA

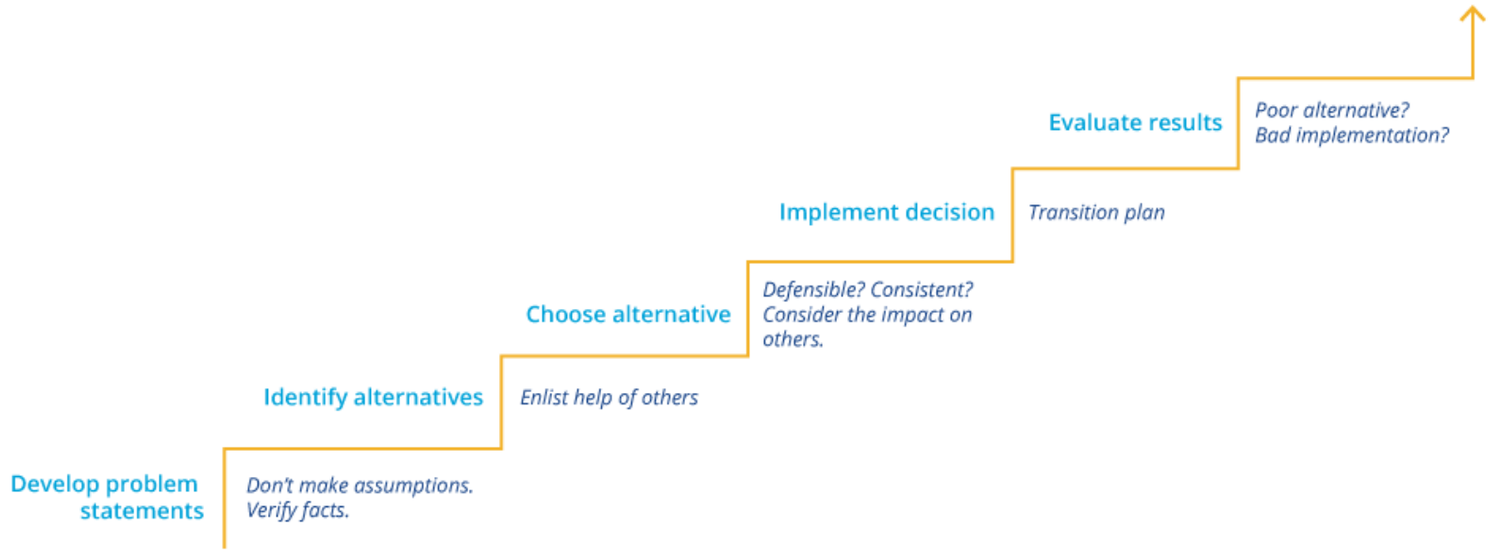
TOPIC 3: SECURITY AND CYBERSECURITY

- **Ransomware** is a form of malware that, if a user unknowingly downloads it to his or her smartphone, takes control of the device and its data until the owner agrees to pay a ran- som to the attacker.
- A **Distributed denial-of-service attack** is one in which a malicious hacker takes over computers via the Internet and causes them to flood a target site with demands for data and other small tasks.
- A **Trojan horse** is malicious code hidden inside a seemingly harmless program.
- A **botnet** is a large group of computers controlled from one or more remote locations by hackers, without the knowledge or consent of their owners.
- **Trustworthy computing** is a method of computing that delivers secure, private, and reliable com- puting experiences.
- The process of assessing security-related risks from both internal and external threats to an organization's computers and networks is called a **risk assessment.**

JAMES COOK UNIVERSITY AUSTRALIA

ETHICAL DILEMMA 3

- **Join a cyberterrorism unit?**
- You are one of the top students in your university's computer science program of 200 students. You are surprised when two representatives from a federal intelligence agency meet you after class. Over dinner, they talk to you about the increasing threat of cyberterrorist attacks launched on the United States by foreign countries and the need to counter those attacks.
- They offer you a position on the agency's super-secret cyberterrorism unit, at a starting salary 50 per cent higher than you know other computer science graduates are being offered. Your role would be to both develop and defend against new zero-day exploits that could be used to plant malware in the software used by the government and military computers. Would such a role be of interest to you?
- What questions might you ask to determine if you would accept their offer of employment?
- You may wish to consider the five-step decision-making process.



TOPIC 4: PRIVACY

- **Data perturbation**: adds 'noise' to databases allowing individual record confidentiality.
- **Data encryption**: translates **data** into another form, or code, so that only people with access to a secret key (formally called a decryption key) or password can read it
- **Data anonymisation**: encrypting or removing personally identifiable information from **data** sets

JAMES COOK UNIVERSITY AUSTRALIA

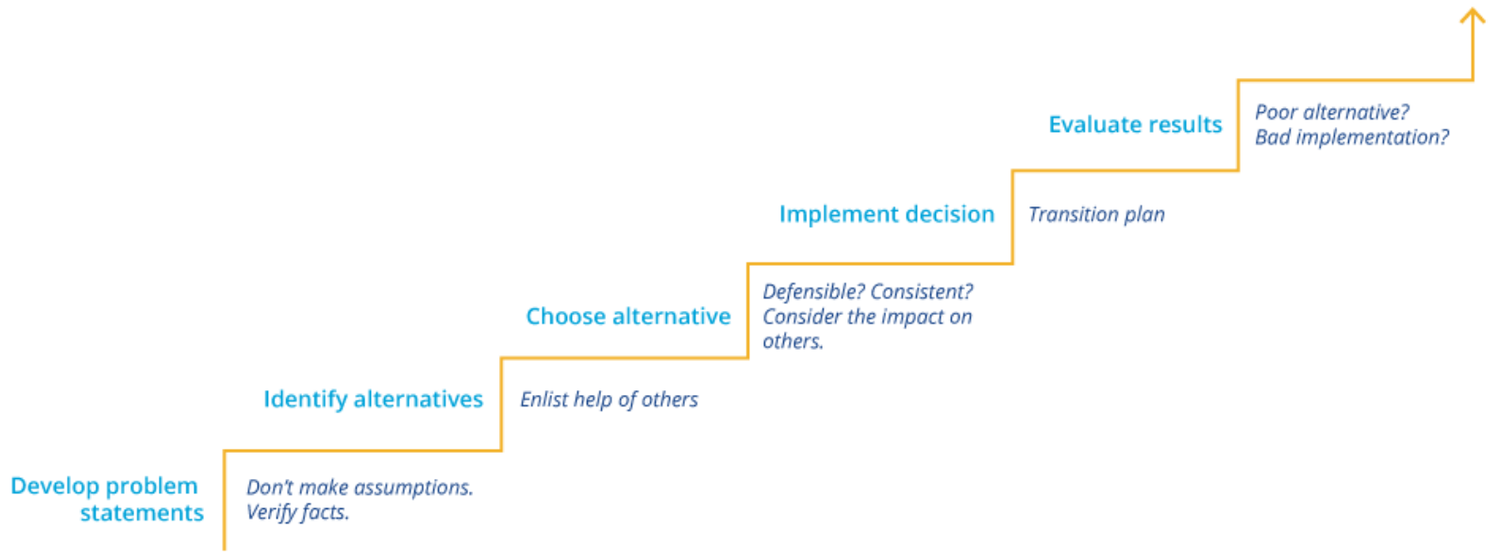
TOPIC 4: PRIVACY

- **Electronic discovery**: The collection, preparation, review and production of electronically stored information for use in criminal and civil actions and proceedings.
- **Data breach**: The intentional or unintentional release of private or confidential secure information to an untrusted environment.
- **Consumer profiling**: A way of describing a consumer categorically so that they can be grouped for marketing and advertising purposes.
- **Electronically stored information**: Any form of digital information, including emails, drawings, graphs, web pages, photographs, word-processing files, sound recordings and databases stored on any form of electronic storage device.
- **Fair information practices**: A term for a set of guidelines that govern the collection and use of personal data.
- **Information privacy**: The combination of communications privacy and data privacy.
- **Trap and trace**: A device that records and identifies the originating number of incoming calls for a particular phone number.
- **Right of privacy**: It is the right to be left alone, the most comprehensive of rights, and the right most valued by a free people.

JAMES COOK UNIVERSITY AUSTRALIA

ETHICAL DILEMMA 4

- **Accept the discount offer?**
- Your car insurance company has offered you a 15 per cent discount (roughly $200 per year), if you agree to let them install a sophisticated vehicle event data recorder (EDR) in your car. You have read over the terms of the agreement and discover that if you are involved in an accident, you must agree to let the data from the device be collected and analysed by a third-party accident investigation firm. You must also agree to let findings from this analysis be used in a court of law.
- What questions would you want answered and what advice might you seek before deciding whether to accept this discount offer?
- You may wish to consider the five-step decision-making process.



TOPIC 5: FREEDOM OF EXPRESSION AND IP

- **Defamation**: An oral or a written statement of alleged fact that is false and that harms another person.
- **Doxing**: The examination of Internet records in an attempt to reveal the identity of an anonymous poster.
- **Hate speech**: Persistent or malicious harassment aimed at a specific person, which can be prosecuted under the law.
- **Internet censorship**: The control or suppression of the publishing or accessing of information on the Internet.
- **Sexting**: Sending sexual messages, nude or semi-nude photos, or sexually explicit videos over a mobile device.
- **Slander**: An oral defamatory statement.
- **Copyright**: Grants the creators of original works the exclusive right to distribute, display, perform or reproduce the work in copies, or to prepare derivative works.

JAMES COOK UNIVERSITY AUSTRALIA

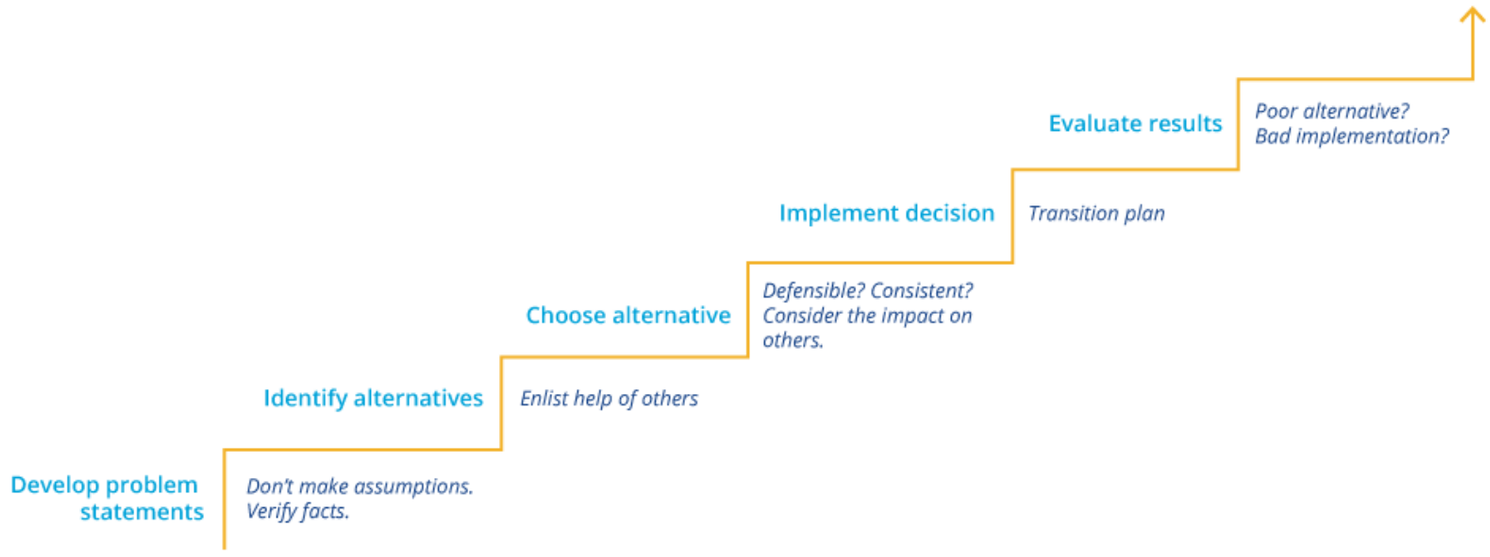
TOPIC 5: FREEDOM OF EXPRESSION AND IP

- **Cybersquatter**: A person or company that registers domain names for famous trademarks or company names to which they have no connection, with the hope that the trademark's owner will buy the domain name for a large sum of money.
- **Industrial espionage**: The gathering of information, not available to the public, through illegal means.
- **Intellectual property**: A term used to describe works of the mind, such as art, books, films, formulas, inventions, music and processes, that are distinct and 'owned' or created by a single person or a group
- **Patent**: A legally enforceable right that is granted for any device, substance, method or process that is new, inventive and useful.
- **Open source code**: Any program whose source code is made available for use or modification as users or other developers see fit.
- **Reverse engineering**: The process of taking something apart in order to understand it, build a copy of it or improve it.
- **Trademark**: Anything that enables a consumer to differentiate one company's products from another's.

JAMES COOK UNIVERSITY AUSTRALIA

ETHICAL DILEMMA 5

- **Send a warning?**
- You are a young, recently graduated attorney working part-time as part of the re-election campaign team for your midsized city's mayor. Several citizens have taken to writing strongly worded letters to the local newspaper voicing their displeasure with your candidate's actions in his initial term as mayor.
- The campaign manager has suggested that you file at least three defamation lawsuits against the most vocal complainers as a warning to others of what they can expect if they are too vocal in their disagreement with the mayor.
- The goal is to intimidate others who might be inclined to write negative letters to the newspaper.
- How would you respond to this suggestion?
- Propose your course of action, giving a rationale for your decision.
- You may wish to consider the five-step decision-making process.



ETHICAL DILEMMA 6

- **Reducing loss of trade secrets**
- You are interviewing for the role of human resources manager for a large software developer. Over the last year, the firm has lost a number of high-level executives who left the firm to go to work for competitors. During the course of your interview, you are asked what measures you would put in place to reduce the potential loss of trade secrets from executives leaving the firm.
- How would you respond?