

# INFO5993 / 4990 RESEARCH METHODS

Week 5

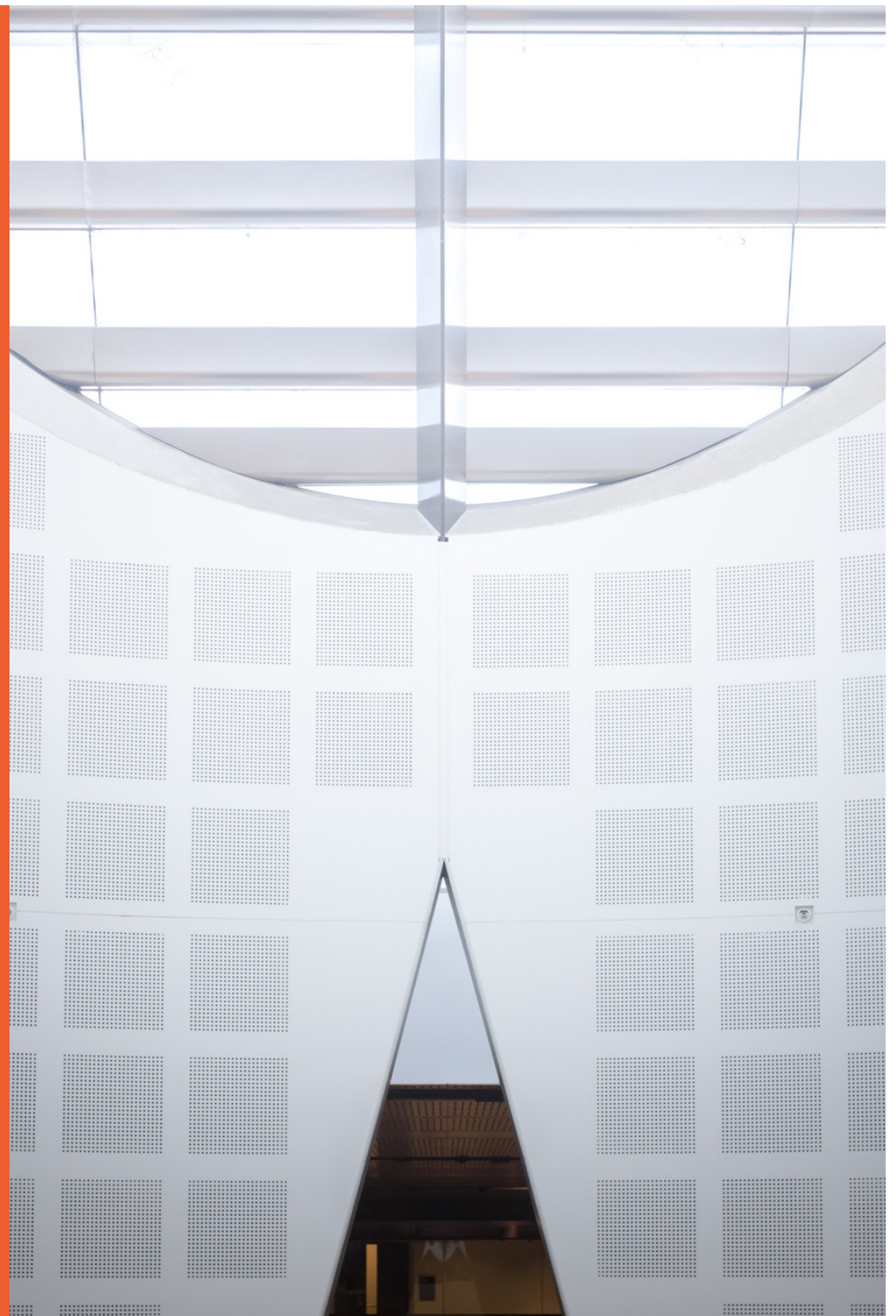
Quality Research  
Research Ethics

Kalina Yacef

School of Computer Science



THE UNIVERSITY OF  
SYDNEY

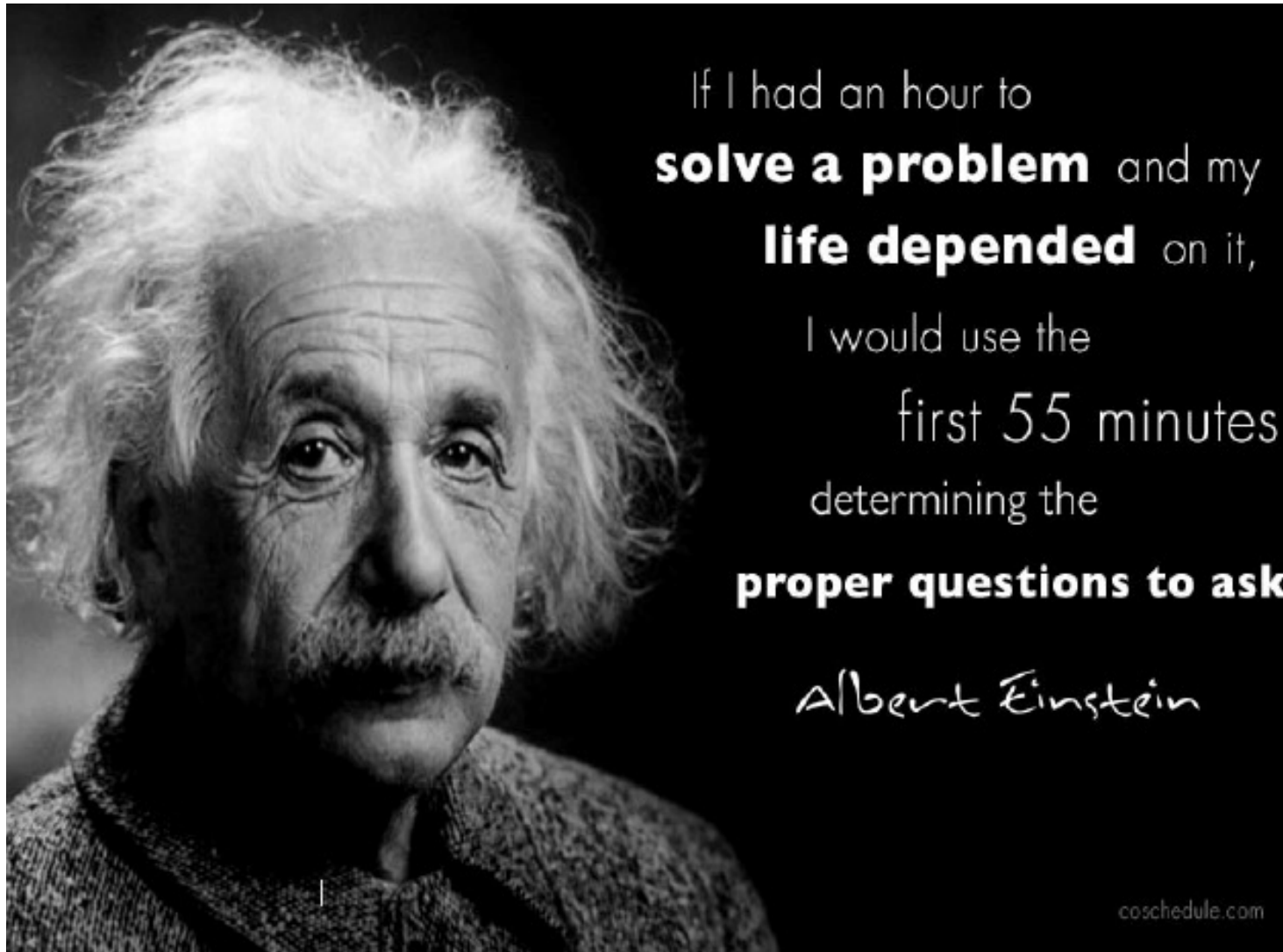


# Acknowledgement of Country

*We acknowledge the tradition of custodianship and law of the Country on which the University of Sydney campuses stand, the land of the Gadigal people of the Eora Nation.*

*We pay our respects to those who have cared and continue to care for Country.*

# Great scholars do not **SOLVE** problems: they **CREATE** them



# What defines a Good Research Problem

- Core quality of a good researcher!
- The solution of such problems helps us in seeing the world in a new way.
- Finding a new problem or help re-define or clarify an old one is a great contribution to your research community.
- Not all problems are worth solving ("so what"?). Should be of interest to the community.
- Not always easy to formulate at the beginning!
  - However, thinking about the problem early in your degree will save you a lot of time later on. This also puts in you in the right frame of mind for advanced research
- From the research problem, narrow down to specific research questions

## Identifying research question

- Especially when you are learning to do research, it may be already chosen for you by supervisor
  - or supervisor may suggest an area, and leave you to find the question
- A question may arise from some previous research
  - Further work, issues not addressed, holes in the evidence collected
- A question may come from the combination of previous research
  - Bring two areas together, use a technique from one area in another
- A question may arise due to new technology
  - new hardware or technique may require new models, influence use or performance or feasibility
- A question may come from simple curiosity

# Suitable Research Questions

- **Specificity and answerability** – can the questions be answered through research?
- **Scale and scope** in relation to needs and available resources
- **Resource adequacy** in relation to available time
- Often start with **broad topic space**/ bigger question, then narrow down to a **specific** question

## Tips for Finding Research Questions (1)

- Read the papers your supervisor gave you
  - follow the references, check the web pages of the authors
  - read carefully the “Future research” sections
  - write down your ideas!!
- Follow the top conferences and research groups in your field
  - scan the call for papers and associated workshops for hot topics
  - scan the conference proceedings to identify important topics, key people and research groups. Check their web pages.
- Find review (survey) articles
- Talk to your supervisor, mentors, researchers, other students, etc.
  - don't be limited by your supervisors' suggestions – they are only a start, always strive to continue the thinking and find more !

## Tips for Finding Research Questions (2)

- Pick a rough idea that interest you and your supervisor
- Do a literature review, find 5 problems
- Write them up in short paragraphs with references
- Take those to your supervisor, pick one
- Repeat until the topic is sufficiently narrow

Callahan, 2001

→ Assignment 1 helps you with this

Try the research topic generator ☺- Disclaimer: this is a joke!  
<http://www.cs.purdue.edu/homes/dec/essay.topic.generator.html>



## Know your Research Community

- A community has conferences and journals of high prestige which they read and publish in
  - They meet often, and each knows (more or less) what others are doing
  - You must place your work in the context of a community
- Divided geographically
  - Europe vs America vs Asia

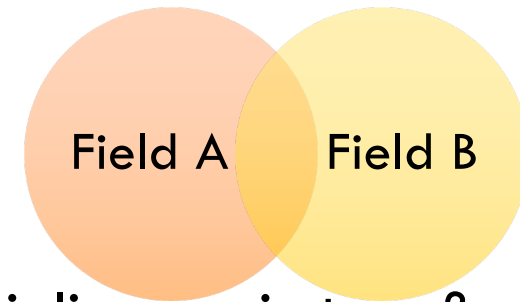
## Keep up with your research community

- Read lots of relevant papers
  - With a careful and critical eye
- Review papers from colleagues in your field
  - Ask your supervisor to practice reviewing
- Know who are the key players, and follow their work

# Your research community

- Belong to (at least) one community:
  - Critical to understand who your community is to know where to publish, who the top experts are, what the research practices are and so on.  
(Assignment 1)
- Contribute to the community:
  - Write, share your research, review papers, give critical & constructive feedback, mentor, organise events etc..
  - Ask your supervisor to help and mentor you in these
- Sometimes research is interdisciplinary... What is your community ?

# Interdisciplinary research



- Where is the research contribution? In all disciplines, or just one?
- What about these:
  - you created a new algorithm and collaborate with a statistician to improve the analysis of your results ?
  - You collaborate with EIE engineers to build some sensor (which already exists) for your research experiment ?
  - You collaborate with EIE engineers to build a new type of sensor for your research study ?
  - You collaborate with medical researchers to explore novel ways to address an unsolved medical problem, and come up with new algorithms. This process requires you to define research objectives together. ?
- Publish in A, keeping B's material at a generally accessible level
- If it exists, publish at the intersection but also in your primary field

# Computing Sciences Accreditation Board, Inc.

- [https://en.wikipedia.org/wiki/CSAB\\_\(professional\\_organization\)](https://en.wikipedia.org/wiki/CSAB_(professional_organization))
- Computer Science is the body of knowledge concerned with computers and computation. It has theoretical, experimental, and design components and includes:
  - **Theories** for understanding computing devices, programs and systems.
  - **Experimentation** for the development and testing of concepts.
  - **Design methodology, algorithms and tools** for practical realization.
  - **Methods of analysis** for verifying that these realizations meet requirements.

# Assignment 1

- Identifying your research community and its top literature venues
  - Finding good examples of papers
  - Identifying Research Problems/Questions
  - Creating an annotated bibliography
- 
- Due Friday 25 March through Canvas/Turnitin and to your supervisor.

# Formulating research objectives

- Must be **original, significant and achievable**
- Can be of various forms!
  - Question
    - what is the best method for detecting micro-calcifications in mammograms?
  - Hypothesis
    - Fuzzy logic-based heuristics can improve automatic detection of micro-calcifications in mammograms
  - Task
    - Compare machine learning classifiers to detect micro-calcifications in mammograms
  - Investigation
    - Find the defining characteristics of micro-calcifications in mammograms pictures
- BE CLEAR about your research objectives, keep the focus

# How is research accepted as contribution?

- Must be written down!
  - Writing essential part of the research process
  - Log book + practice writing every day
- Must be situated in the context of previous knowledge (*Assignment 2*)
- Evaluate and argue why it is valid and new
  - Methods will be explored from next week
- Peer review: objective critique, feedback, validation from the community experts
- Share it, so that others can use it, connect with it, build from it

# Describe and Disseminate Your Research

- Have several clear, concise and succinct statements of your research problem
  - e.g. one minute (elevator) pitch
  - ten minutes introduction to full seminar
- Issues you must deal with:
  - Can it be understood by others without too much background?
  - Does it demonstrate a good understanding of the research community?
- Give lots of talks
- Writing up papers
  - Great practice for thesis write-up
  - Feedback from your community
  - Become famous... !



# Share your research verbally

## 1-2 min Elevator pitch about you and YOUR RESEARCH

- Goal : create a positive impression and start a discussion
- Write it down, edit it to make it **concise**, and **practice** it
- Begin by **posing a question** or **stating a problem**
- Describe the **unique** approach or insight that motivates your study
- The wording and delivery should be **conversational** and at a **high-level overview**, as opposed to dwelling on details
  - **Avoid any jargon**

15 sec introduction pitch, to present YOU in the context of a meeting/etc.

A “social” pitch: to present what you do to a non specialist audience, to friends, family, social occasions

## Elevator pitch

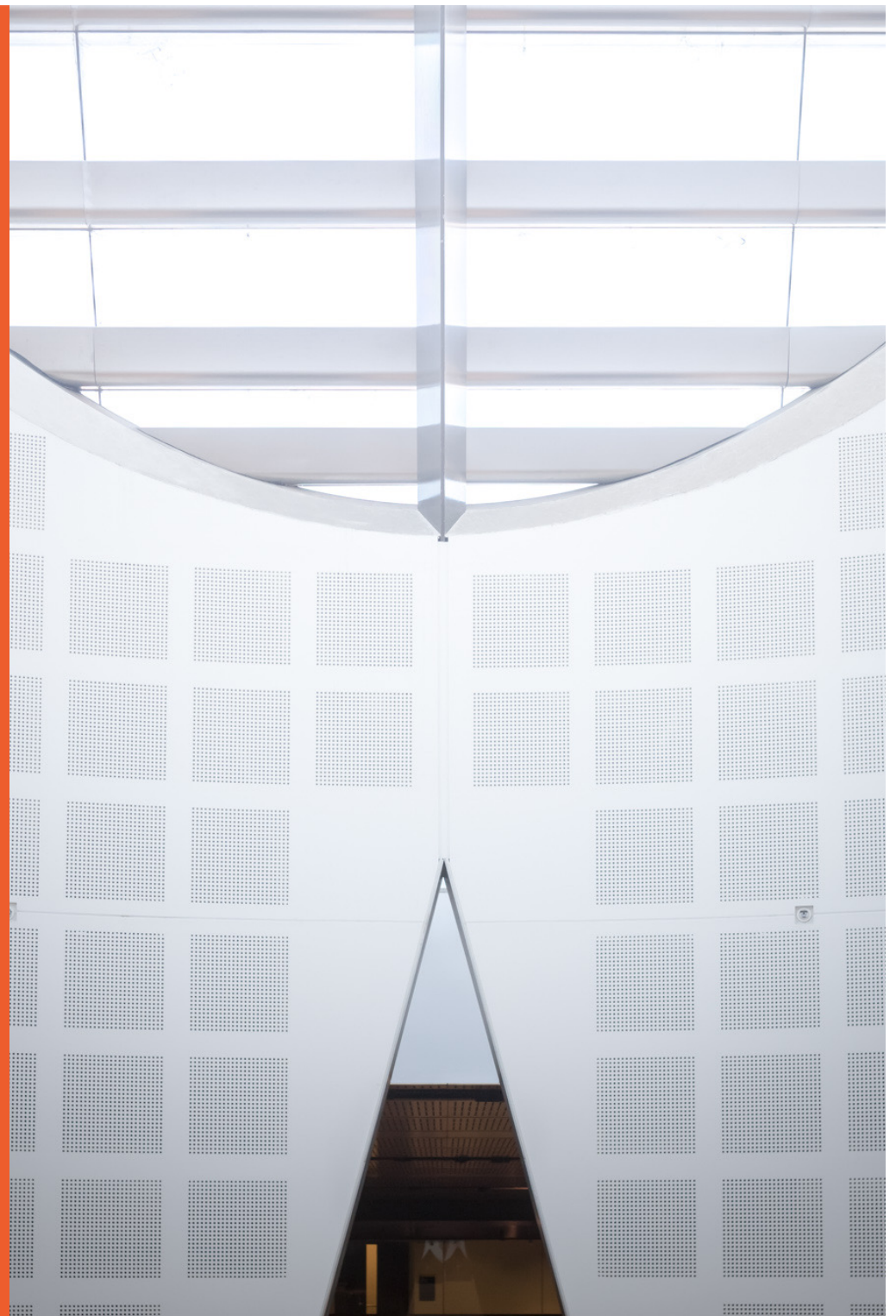
- What is the topic of your research?
- What is the problem, issue, or question that you are asking and addressing in your research?
- Why is that problem interesting and important? (i.e. So what?)
- How does your work connect with a broader disciplinary conversation about this topic/problem in your field, and what does it add to that conversation?

# INFO5993 / 4990 RESEARCH METHODS

## Research Ethics

Kalina Yacef  
School of Computer Science

Based on material by Judy Kay, Alan  
Fekete, Peter Eades



# Agenda

- Background principles
- Honesty and integrity
- Credit & authorship
- Respect to participants
- Some scenarios
- Protect your rights

# Ethics: background principles

- Professional ethics
- Australian Government Research ethics
- University ethics

# Professional ethics

Emphasise obligations to clients, community, profession as a whole

- Association for Computing Machinery (ACM)

<http://www.acm.org/about/code-of-ethics>

*Relatively general*

- Australian Computer Society (ACS)

<http://www.acs.org.au/documents/codes/CodeofProfConductPractice.pdf>

*Oriented toward the IT profession*

# Australian Government Research bodies

- Australian Code for the Responsible Conduct of Research 2007

<https://www.nhmrc.gov.au/guidelines-publications/r39>

provides a framework for managing breaches of the Code and allegations of research misconduct, managing research data and materials, publishing and disseminating research findings, including proper attribution of authorship, conducting effective peer review and managing conflicts of interest. It also explains the responsibilities and rights of researchers if they witness research misconduct.

Research oriented

Common to all Australian Universities

NHMRC, ARC and others funding agencies

# University

- Research Integrity resource portal

<https://intranet.sydney.edu.au/research-support/ethics-integrity/research-integrity.html>

- Get familiar with:

- Code of Conduct for Students
- Responsible Research Practice training module
- Research Code of Conduct 2013

- If your research require human participants, Human Ethics (as well as Animal and Environment)

<https://intranet.sydney.edu.au/research-support/ethics-integrity/human-ethics.html>

- If needed, Research integrity advisor (Engineering Faculty : Dr Yixiang Gan)



## Honesty and integrity

- The rest of the community must be able to rely on your claims about your contribution
  - This allows the field to advance
- Be explicit about the evidence for your claims
- Be open and clear about limits or weaknesses of your work

# Tell the truth!

- Experiments must have been done as described
- If you say you built a system, it must exist
  - Complete and running (integrated)
  - Not just separate pieces for each aspect
  - it is OK to describe design where some features are not implemented if that is stated in the paper
- If you state performance values, they must be real and fair
  - Describe the environment in enough detail (eg was there other load, were the caches pre-filled, what hardware/software was used)

# Tell the truth!

- Data must be representative
  - Don't pick the most positive examples from several attempts
  - Don't reason circularly (ie don't use data to derive a hypothesis, or to determine key parameter values, and later claim the same data as evidence)
    - Instead, have separate training set and evaluation/test set
- Preserve relevant data for 5 years
  - include information to allow readers to obtain the data
  - inform supervisors of location of data
  - Use the University's [Research Data Store \(RDS\)](#) (access requested via [DashR](#)).
  - Consult [The University's Research Data Management](#) page for more details on how to manage your data

# Honorable research

- Do not use unlicensed software
- Obtain Ethics approval for human experimentation
  - eg user interface trials
  - Informed consent
  - Confidentiality [make anonymous]
- Respect confidential material
  - eg non-disclosure agreements
- Cross-reference all other papers from same research
  - each should have a distinct contribution, clearly explained
    - eg one paper describes system architecture, another presents network optimisation, another shows innovative UI
- Respect people/organisations that provided information, equipment, funding etc.
  - Licence conditions
  - Non-disclosure agreements
  - Possible approval process

## Credit others

- Cite sources you build on
  - Enough detail to find their work!
  - Use their words only in quotations (and then, be accurate)
- Acknowledge everyone who played a role
  - Eg tech support who provided patches to compiler
  - Eg friend who proofread for language idioms
  - Eg assistant who entered the data
  - Eg someone who gave an idea that turned out to be wrong but led to solution
- Authorship for those who provided important ideas
  - “conceiving, executing or interpreting” part of the research
  - Clearly distinguish ideas from other people
    - This includes thesis supervisor

## Authorship (i)

- Authorship is substantial participation, including:
  - Conception and design, or analysis of data; and
  - Drafting/revising the paper; and
  - Final approval of the version to be published.
- An author's role must be sufficient for that person to take public responsibility for at least part of the paper.
- Before writing a paper, discuss with supervisor
  - scope of paper, where to publish
  - who are the co-authors, and in what order

# Authorship

- All people qualified to be authors should be authors
  - Everyone who is allowed to be an author should be an author, as long as he/she agrees.
- Often, supervisor will be co-author
  - contributed plan, advice, interpretation
- Often other members of the group will also be co-authors
  - Paper may combine results from several students' projects
  - This requires care and clarity when these results are discussed in an individual student's thesis
- Recommendation: Authors should sign an authorship statement

# Authorship is NOT

- Helping to get funding for the project
- The collection of data
- General supervision of the research group without any involvement in the work described in the paper
- “Honorary authorship” is not acceptable.
- Non-authors who have contributed (e.g., funding) should be acknowledged



## Note on Plagiarism/recirculation

- Recirculation, or self-plagiarism
- Publication of multiple papers based on the same data is not acceptable except where there is full cross-referencing within the papers
  - In any doubt, disclose this to both editors/publishers
- Always cite previous papers that you have written with a similar theme/content
- As a rule of thumb, don't copy-and-paste anything except some parts of the introduction
- All theses go through Turnitin

## Respect to participants

- Humans, animals, environment etc.
- Any research involving human participants requires an Ethics approval from the University
  - Questionnaires, surveys, observations, use of historical data eg logs etc..
- Asked to provide
  - Recruitment method, Participant consent form
  - Population targeted
  - Research protocol
  - Data anonymisation, storage, analysis, publication and so on
- Done through IRMA
- Ethics committee meet ~twice a month but are very busy – plan early!

# Some scenarios

## Scenario: Millicent and Mutter, with Dingle

1. An honours student, Millicent, writes a brilliant thesis on simplifying agent-oriented concept design (AOCD).
2. His supervisor, Professor Mutter, sees that it is brilliant and turns it into a joint paper, which is accepted to the rank A conference AOS2021.
3. Millicent gets first class honours and goes to work in SAAB in Sweden.
4. Professor Mutter presents the joint paper at AOS2021 in Tokyo.
5. Professor Dingle from Ohio State University sends an email to Professor Mutter pointing out that the brilliant simplification of AOCD was all in a paper that Dingle published in 2018. She accuses Mutter of plagiarism.

6. Mutter checks Millicent's thesis against Dingle's paper and finds that large sections have been copied, word for word; Mutter apologizes to Dingle.
7. Mutter writes to his dean and asks that Millicent's honours degree be rescinded.
8. The dean accuses Mutter of plagiarism.
9. The case works its way up the University disciplinary system.
10. The university offers Mutter a choice: accept a demotion to Associate Professor, or resign.
11. Mutter resigns.

## Scenario: Robbie and the Rapid Router

1. Robbie has a new routing algorithm RR that he thinks is faster than previous algorithms.
2. There are some published benchmark data sets for this kind of routing.
3. RR is a randomised algorithm (eg, genetic algorithm) that gives a different result every time you run it.
4. Robbie runs RR on the benchmark 1 000 times, and finds that the average runtime is 231.1ms; while the maximum runtime is 1 451.7ms and the minimum is 62.1 ms.
5. The best previous result on this benchmark used 81.3ms.
6. Robbie submits a paper reporting that his new algorithm is better than previous algorithms because it ran on the benchmark in 62.1 ms.
7. ....?

## Scenario: Bertie, Bogie and his wife

1. Bertie, the departmental director of research, does not like Associate Professor Bogie.
2. Bertie notes that Bogie has written a joint paper with his wife, who is a student at a different University.
3. Bertie begins to look through Bogie's many papers and finds three papers which are almost the same. They are published at three different conferences.
4. In the meantime, Bogie accepts a job as Professor at a different University.
5. Bertie writes to the director of research at Bertie's University, pointing out that at least three papers of Bogie are virtually the same.
6. ... ?

## Scenario: Ellen and senior professor Miles

1. A PhD student, Ellen, goes to a conference, and gives a talk.
2. After the talk, a senior and respectable professor (called Miles) asks Ellen lots of questions, and asks her about her future directions.
3. Ellen tells him everything; she is very happy that Professor Miles is interested in her work.
4. Miles is ambitious, but he has only published two papers in the last three years.
5. Six months later, Miles publishes a paper which has all the stuff that that he and Ellen discussed.
6. The paper has no acknowledgement to Ellen.
7. ...?



## Scenario: Kathleen and Mabel, with Malmsbury error detection

1. A PhD student, Kathleen, is writing a paper for a conference (in Colorado) and discusses it a lot.
2. Mabel has a great idea that would fit right in Kathleen's paper.
3. Kathleen and Mabel chat and agree to include Mabel's idea; also to include Mabel as an author.
4. The paper is accepted, and presented at the conference by Mabel (Kathleen is in Norway at another conference).
5. At the conference, Professor Malmsbury sees a critical error in the paper.
6. Mabel says "It's really not my paper, Kathleen wrote it, it's her error".
7. ... ?

## Scenario: Banbury, Brightwhistle and the X-Rays

1. A PhD student, Banbury, invents a wobbly algorithm and applies it to 2018 chest X-ray data from Wentworth.
2. He publishes the paper in WOBBLY2021, claiming that it is better than the 2016 wobbly algorithm of Brightwhistle and Scott.
3. Brightwhistle gets annoyed, because she thinks her algorithm is the best.
4. Brightwhistle wants to test her 2016 algorithm on the 2018 chest X-ray data from Wentworth, and asks Banbury for the data.
5. Banbury replies that he spent half his grant extracting the data from the database, and if Brightwhistle wants the data, then she can get it herself.
6. ... ?

# Protect yourself

- Conflict with supervisor
  - Can occur during thesis (eg over funding, co-authorship, commercialisation, etc)
    - Best if issues discussed before work starts (eg agreement over ownership of IP)
    - If necessary, approach School director or your program
      - If PhD/Mphil, School Research Education director (Irena Koprinska), or Faculty's Associate Dean (Research Education), Kalina Yacef
    - Key is clear communication, and written records of agreements, presentation of ideas, etc.
- Planning for the future
  - Before you leave the university, talk to supervisor about future plans/possibilities
  - Will you continue the work?
    - What directions/extensions will supervisor give to other students (vs those you will do yourself, or give to your students)
  - Will you submit more of your work for publication?
    - Maybe supervisor will polish and submit, with both as co-authors
    - Maybe another student will continue, leading to combined publication

## Protect yourself (ii)

- Issues with referees, acquaintances, web surfers
  - People who see your work before publication might, if they are not ethical, publish it themselves, or improve on it
  - This can restrict your recognition
    - May even destroy your chance to get a degree!
  - Best protection is time-stamped publication (eg School Tech Report, or Arxiv)
    - Before submitting or sending out for comment

# Intellectual Property

- Four pillars of IP
  - Copyright in programs
  - Patents on techniques
  - (Trade secrets and Trademarks)
- Law is highly complicated
  - work done for a degree belongs to the student, not the university
  - work done when employed belongs to university
- Read University's regulations and policies on IP
- University IP policy

# Commercialisation

- If you think your project has commercial prospects, discuss things with supervisor and the University's Research support - Industry and Commercialisation
- Open publication of thesis (probably prevent patent claim)
  - can request thesis not to be openly published or to be embargoed for up to 18 months (requires Dean's approval), or have chapters embargoed or restricted.

## Next week

Deep Dive into CS Research methods:

1. Statistical Evaluation & Research evaluation
2. AI/Deep Learning Research methods