# Topic Overview

## Key Terms

* What field are we in?
  + Informatics
    - Edinburgh started using this term for the department specifically because we needed something general to cover all of the subject areas that are essentially studying “information processing systems)
    - So computer science, cognitive science, software engineering, AI
    - People outside of those fields use these fairly interchangeably, and for our purposes the distinction is usually not going to be important. Most points made about “software engineers” and likely to be just as relevant to AI-ers (we don’t have a good word for that one, eh?)
    - Even inside of Informatics, the boundaries are pretty blurred. The only one of those I would not likely identify as is cognitive scientist, and I’ve still studied cognition and taught on cognitive science courses.
    - I usually use “computer scientist” as the most general term
* What is this course studying?
  + “Professional Issues” is, I think, intentionally vague (I didn’t name it)
  + In the “Professional Practice” part of SEPP, we covered things that might be considered “Law”, “Politics”, “Management”, “Philosophy”, “Sociology”, etc
  + So while most of your courses are teaching the practical and technical aspects of being a computer scientist, courses like the SE part of SEPP and the SDP teach you how your individual work fits into the wider context of a whole project or system, where you’ll likely be working alongside other people and taking into account users or a client.
  + The context of professionalism, then, can be seen as going one step further. We are interested in how your individual work, or any system you might be contributing to, fit into the wider context of the societies we live in.
  + How might what you do in your work as professional computer scientists influence, or be influenced by, that wider context? In what ways might that work cause harm? What might we do to try and mitigate that harm?
  + Unlike the mostly legal, rule-following topics in Professional Practice, here we are talking about things that often won’t have a right or wrong answer. What we are developing is your ability to engage with these issues and to discuss them with other people.
* AI Ethics
  + A lot of what we will cover here can loosely be called “Ethics”, and certainly a lot of similar courses at other institutions would be labelled as such.
  + Ethics is the study of what constitutes right or wrong, good or bad behaviour
  + Much of the reading you are likely to do around this topic will be talking about Ethical AI. Not because these issues are usually exclusive to AI, but because AI as a field is very much in the public eye at the moment, and so that is where issues of ethical behaviour often come to light.
  + Confusingly the terms “AI Ethics” or “Ethical AI” can get used to refer to a variety of things, so bear this in mind to avoid getting caught out:
    - Ethical AI where AI refers to the field, in terms of what gets researched or developed
    - Ethical AI where AI refers to the people in the field, in terms of how companies or employees of those companies behave
    - Ethical AI where AI refers to the actual artifact, the “artificial intelligence”, created by the process and how it behaves
    - And finally now there are “Ethical AIs” where people are specifically trying to design AIs to be able understand and reason about ethics

## Why are we studying Professional Issues?

* Practically, it was introduced because it’s a requirement of professional accreditation
* Because it’s a required part of most of your degrees
* Basically all respectable CS degrees are including some ethics in their programmes now
* Tech companies have come under a lot of ethical scrutiny and want to show that they are doing something about it, which will include favouring hiring engineers with ethical qualifications
* Big conferences are similarly starting to require ethics statements alongside submissions (if research is more your thing)

## Why should we be studying Professional Issues?

* The suite of things achievable by modern computer science is incredible and growing rapidly
  + I am not particularly old (I hope), but my childhood was pre-mobile phones, pre-internet
  + The suite of things people *claim* modern computer science can do is much larger.
  + Both of these, both what can and can’t be done, can have big impacts on people’s lives.
  + Technology is practically inescapable and so the people making it should be aware of
* There are lots of ways in which your work can influence other people, not just intended outputs of systems you build, but
  + Unintended outputs
  + Unexpected usage
  + Impact on society
  + Disproportionate output
  + Other people you work with
* We as a school believe it is important for all of our students to develop the capacity to use their technical skills responsibly.

## Key Takeaways

* Not to believe you can do it yourself
* Also not to believe that it’s someone else’s problem
* No easy answers. This is not a learn and parrot course.

## Previous Week’s Reading

* Impressed by the quality of comments on Hypothes.is, and appreciated that people were providing each other with definitions to help comprehension. Exactly how it should be used!
* This paper was talking about the state of ethics in computer science generally, and linking that in to how people are taught.
* Discourages “techno-solutionism”, where every problem is treated as something solvable by technology and/or technologists
* Or, at the other end of the spectrum, dismissing any responsibility
* Encourages participatory design: including social scientists and stakeholders in the process. It also encourages a shared understanding between those