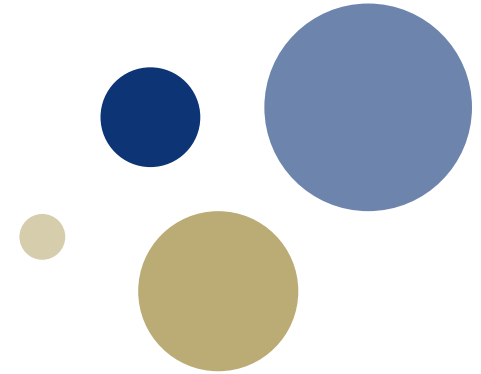




Norwegian University of
Science and Technology



Ethics in the age of Informatics, Big Data and AI

Professor dr. May Thorseth, Dept. of Philosophy and
Religious Studies, NTNU

Director of [Programme for Applied Ethics, NTNU](#)

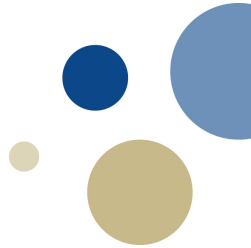
Email: may.thorseth@ntnu.no

Informatics Europe 2018
Chalmers 9 September 2018

Applied ethics: Janus face of ethical challenges



Janus face of security and protection



- Big data, surveillance for security:
 - Improvement of security and protection, against e.g. cyber attacks towards authorities → may safeguard authorities
- Big data, surveillance against protection:
 - Threat to privacy and misuse of data → may compromise protection of individual autonomy
- Must we choose between primacy of security or protection?

Basic values and assumptions about a good society



- Democratic society
- Autonomous citizens
- Protection of systems and people
- *Security and protection* of systems and people against attacks, e.g. **hacking**, cyber attacks, spreading of rumours, indoctrination, false news...
- Question: to what degree are protection and security concurrent?

Hacking – (il)legal or (il)legitimate?



○ White hat

- **Ethical hacker, e.g. identify places to repair.** The white-hat hacker uses their knowledge of computer security systems to compromise the organization's systems, just as a black hat hacker would. However, instead of using their access to steal from the organization or vandalize its systems, the white-hat hacker reports back to the organization and informs them of how they gained access, allowing the organization to improve their defenses.

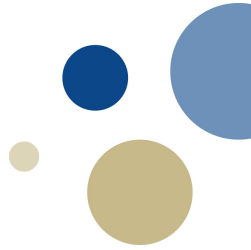
• Grey hat

- **E.g. notify an administrator about a defect system,** then offer to correct the defect for a fee; may publish the facts to the world. A gray hat doesn't work for their own personal gain or to cause carnage, but they may technically commit crimes and do arguably unethical things.

• Black hat

- **Violate computer security for personal gain,** such as stealing credit card numbers or harvesting personal data for sale to identity thieves or for pure maliciousness (such as creating a botnet and using that botnet to perform DDOS attacks against websites they don't like).

Janus face of hacking and surveillance



Same technology, different purposes
→ legitimacy depends on what?

Ethical dilemmas in context

- Choose between *either* system security *or* protection of privacy → what kind of info should be considered private/ not to be shared by all?



- Problem of fixed definitions of hacking because:
 - Divergent aims of security
 - Hacking may be illegitimate although not illegal?
 - Hacking may be illegal while still legitimate?

Scenarios of hacking

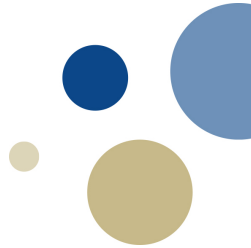
- Contextual definition of white/ grey/ black
- Blurred lines between different acts of breaking into security systems:
 - Repairing a system for the good of a company and sharing the knowledge with others → white hat, grey hat?
 - Authorities' surveillance of individuals, e.g. police hacking into personal computers, mobiles and tablets (cf. Aftenposten 25th of May, 2016) → white hat, grey hat, black hat?
 - Breaking into computers of suspected criminals?

Contexts of hacking

- Civil disobedience → due to lack of (rightful) information, cf. democracy, sharing of knowledge commons (e.g. citizens' right to know about spread of serious disease)?
- Industrial espionage → disseminating industrial secrets?
- Snowden-like cases?
- Authorities' hacking of individuals (social life)/ companies (industrial secrets)?



Hacktivism – hacking as civil disobedience



Because no legal status for data exists, organizations collect, process, exchange and sell data without gathering consent, compensating consumers or adequately protecting consumer data, in most cases.

Yet, hackers' accessing the same data is deemed criminal behavior, regardless of **motivation** or injury (pre GDPR, implemented May 25, 2018)



Should hacking be deemed criminal independent of motivation, intention or injury?

Ethical reflection on hacking – targeting minds of people: enlightenment or manipulation?

- **Question:** Possible to distinguish between good and malicious intentions of manipulations?
- **Question:** How to measure whether e.g (soft) cyber attacks – like rumours intended to induce hate/ fear/ hope – are legitimate? Relevant to ask who/ for what purpose?

Case: Manipulative information spread under false identities

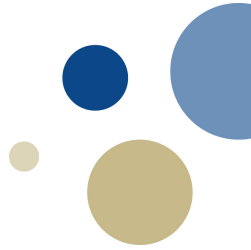


Intention (i): Create chaos, panic, disorder (cyber war, political elections)

Intention (ii): Induce critical reflection, appeal to judging from the viewpoint of divergent perspectives (cf. Kant: reflective judgment)


Question: Possible to keep intention (i) and (ii) apart (by appeal to ethical theories, e.g. deontological or consequentialist arguments)?

Moral concern: question of legitimacy



Ethical theories – lines of justification:

- Deontological theories, e.g. Kant, discourse ethics (Habermas) → appeal to intention → do right
- Consequentialist, e.g. Utilitarian → appeal to consequences → produce good consequences

- 
- **Deontology:** rightness/wrongness of actions themselves → act out of duty towards some law, e.g. Kant's categorical imperative; right action = in accordance with the moral law, possible to universalize, e.g. tell the truth ≠ lie
 - **Consequentialism:** rightness/wrongness of consequences of actions → right = produce good, e.g. utilitarianism

➤ **Deontological justification → good intention**

➤ **Utilitarian justification → good consequences**

➤ **Consequences:**

- Contribution to public debate?
- Silencing of political debate?



- Consequences impossible to foresee → cannot judge by appeal to consequences?



Principles

➤ Precautionary principle

– the link between precaution and innovation



➤ Principle of double effect

– foreseen but unintended vs. intended consequence of actions

Cyber attacks for the good or the bad?



Related to Russian/ Ukrainian case:

- ✓ Targeting minds of people – manipulate data, knowledge, opinion
- ✓ Inflict damage to data or services
- ✓ DoS (denial of service)
- ✓ Info leaks by hacktivist groups
- ✓ Espionage malware

Question: Relevant difference between the parties with respect to legitimacy?

Cyber warfare, main target: minds of people

- Proliferation of narratives → manipulate society's perceptions in order to cause disruptive behaviour
- Cyber Berkut (pro-Russian) vs. Cyber Hundred (pro-Ukrainian)
- Both cyber attacks:
 - Narratives and spreading of rumours to justify and promote activities
 - Cyber weapons: hacking electronic advertising billboards prior to election 24 October 2014; attacking and defacing websites
 - Desired effects: induce chaos, panic, mass disorders

Cyber warfare, surveillance and Big data – threats to democracy



- **Military context:**
 - info as weapon, e.g. malicious hacking to create chaos, panic, disorder and distrust ... *or* (soft) cyber attacks – e.g. rumours intended to induce hope
- **Civil context:**
 - targeting minds of people – manipulate data, knowledge, opinion ...*or* accomodate individual preferences
 - mainstreaming information → filtering and tailoring of viewpoints

Blurred division line between military and civil contexts in the age of informatics

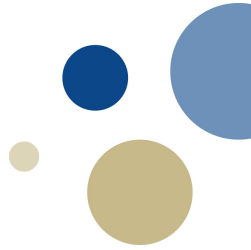


- Targeting minds of people for military reasons or political elections coincide – who are the victims of the battle?
- Dilemmas and ethical concerns partly embedded in the technology itself – who is to blame?
- Robotics and the problem of responsibility

Informatics in civil contexts – recalling basic values for a good society

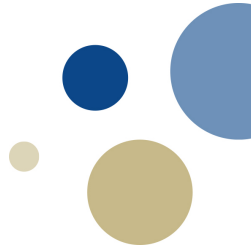
- Democratic society
- Autonomous citizens
- Protection of systems and people
- Security and protection of systems and people against attacks, e.g. **hacking**, cyber attacks, spreading of rumours, indoctrination, **false news...**

Autonomy



- Importance of autonomy → genuine communication and challenge of viewpoints
- Threats to autonomy: influence of political parties (e.g. Stephan Arkadyevich), cf. search engines → infringing upon autonomous choices

Democracy




- Importance of genuine communication and challenging of viewpoints
- Jeopardizing democracy and trust: misuse of information, spread of false news.
- Threats to democracy, i.a. tailoring of viewpoints due to filtering → confirming vs. opposing viewpoints → *Daily Me*

Filtering

- Internet does in principle give access to all kinds of information, but is in practice constrained by filtering
- Filtering, i.e. possibility of *Daily Me* (ref. C. Sunstein)
 - The Internet is substantially different from previous information media in that it does not provide sufficient public forums necessary for democracy.
 - Lack of public forums, characterised by access to a heterogenous group of listeners; unwanted encounters; being exposed to heterogenous viewpoints/speakers.
 - Risk of compromising sharing and exchange of viewpoints, which is basic to democracy

New situation due to information technology



- Weakening of shared public spaces (and traditional political bodies)
- Choices made on different arenas outside democratic governance and control
- Possibility of targeting peoples minds for the good and the bad → legitimate hactivism for enlightenment vs. manipulation

Informatics for a good society



- Emerging technology and the problem of ethics coming after → risk of?
- Big data, surveillance for security and protection, but also depriving people of privacy → main threat...?
- JANUS face of informatics as an emerging technology → no technology is a neutral tool!
- Importance of ethical awareness and reflection in the shaping of new technologies → challenge for robotics in particular: who is to blame when something goes wrong...?