UNIVERSITÄT MANNHEIM

AI and Data Science in Fiction and Society

Seminar Kick Off

Heiko Paulheim

Goals

- Why should we read fictional literature at university at all?
  - given that we are not humanities students

- Fiction is a mirror of society
- Technology is developed in research labs
  - ...but its impact on society is seldom discussed therein

- In this seminar, we will discuss those impacts

Organization

- Requirements
  - Read a novel and present it in the seminar
  - Write a seminar paper
  - Review others' seminar papers
    - it is a good idea to also read the books for the seminar papers you review

- First step
  - Pick a novel
  - If not done yet, send a ranked list to Ms. Bianca Lermer
    - Until the end of Sunday, Sept. 18th
  - You are invited to propose books **not** on the list on the Web page

Organization

- Getting your book
  - we have copies in the library (of books listed on the Web page)
  - you may need them for some time

- After assignment is finished, you can borrow “your” copy
  - address the counter in the library in the castle west wing

Organization

- We will use a process called “peer review”
  - widely used (and discussed) in science
  - you will review your fellow students’ seminar papers

- Timeline
  - Prepare a draft until October 30th
  - You will get two seminar papers to review
  - Submit your reviews until November 6th
- Seminar (i.e. ,presentations, discussions)
  - November 17th, 24th, December 1st, 8th
  - Participate *actively*
- Final seminar paper submission: January 15th, 2022

Preparing Your Seminar Paper

- Synopsis
  - Brief summary of the novel's contents
  - Main themes, characters, plot points
- Characters
  - Main characters? What do they represent?
- Technology
  - Which technology is described?
  - What does already exist, what is still to be invented?
- Interaction of technology and society
  - Which of the characters appreciate the technology?
  - Who are rather sceptics?
- Anything else you want to discuss
- Use further literature!

Hints for Further Literature

- Book reviews
  - e.g., in newspapers
  - which aspects do they appraise? which do they criticize?
- Information on the author
  - e.g., interviews
  - did s/he comment on the book?
- Information on technology
  - e.g., research articles
- Information on ethical aspects
  - e.g., resarch articles
  - e.g., high level documents (EU HLEG, to come later)

Preparing a Review

- 1st rule: be constructive!
- What you should point at
  - can you follow easily? is there information missing at any point?
  - are all claims well supported?
  - do you have any questions not answered?
  - aspects underrepresented
- What you should not do
  - provide general criticism (“don’t like the paper”)
  - correct every spelling mistake
  - rewrite the seminar paper

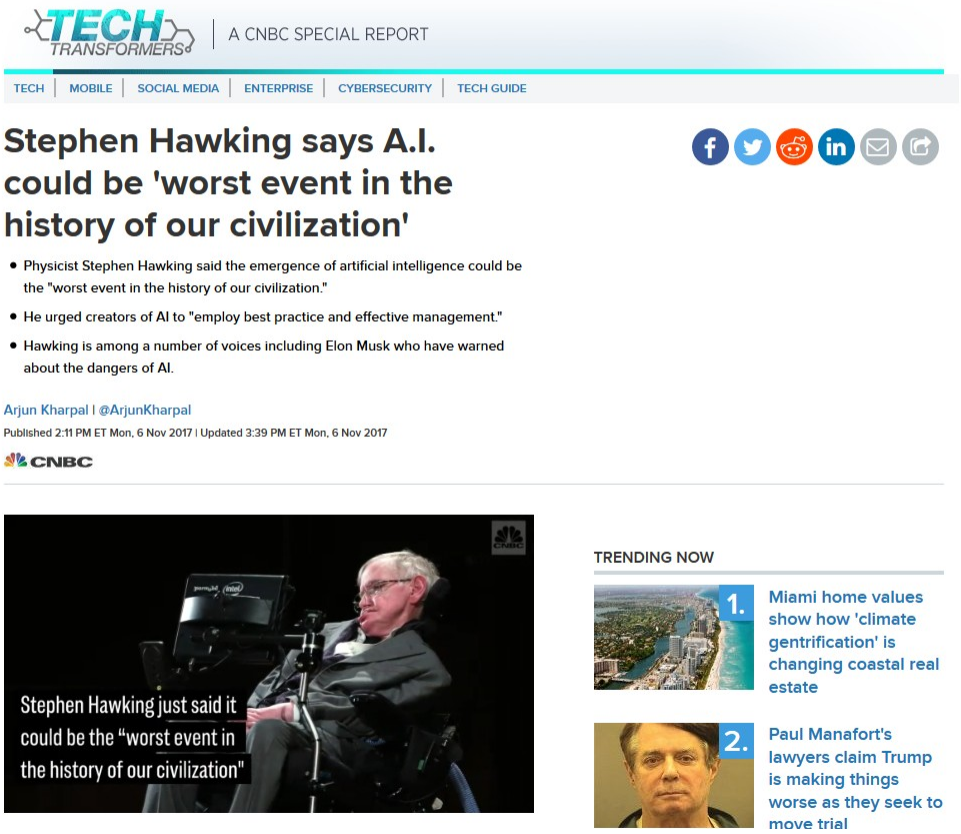
Preparing the Presentation

- Remember
  - we are here to discuss
- Therefore
  - don’t just report facts
  - raise questions
  - be provocative

Questions?

Straw Poll (will be repeated at the end)

- Do you consider AI rather beneficial or harmful?



Isaac Asimov's 3(+1) Laws of Robotics (1942)

- 0th law:
  - A robot may not harm humanity,
    or, by inaction, allow humanity to come to harm.
- 1st law:
  - A robot may not injure a human being
    or, through inaction, allow a human being
    to come to harm.
- 2nd law:
  - A robot must obey the orders given it
    by human beings, except where such orders
    would conflict with the 1st law.
- 3rd law:
  - A robot must protect its own existence as long as such protection does not conflict with the 1st or 2nd law.

EC Ethics Guidelines

- AIs should be...
  - “lawful, complying with all applicable laws and regulations”
  - “ethical, ensuring adherence to ethical principles and values”
  - “robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm”

Today's Challenges (among others)

- Transparency
  - Does AI need to be open source?
- Algorithmic bias
  - How to measure & prove? Is there useful bias?
- Liability
  - If an AI does harm, who is legally in charge?
- Human dignity
  - Are there areas which should not be controlled by AI?

Joseph Weizenbaum's AI-free Areas (1976)

- AIs should not be used for…
  - A customer service representative
  - A therapist
  - A nursemaid for the elderly
  - A soldier
  - A judge
  - A police officer
- Where are we almost 50 years later?

Thought Experiment: Universal Paperclips