


# Ethics and Artificial Intelligence



From our earliest beginnings, mankind has contemplated a future in which our creations can replace human labour.

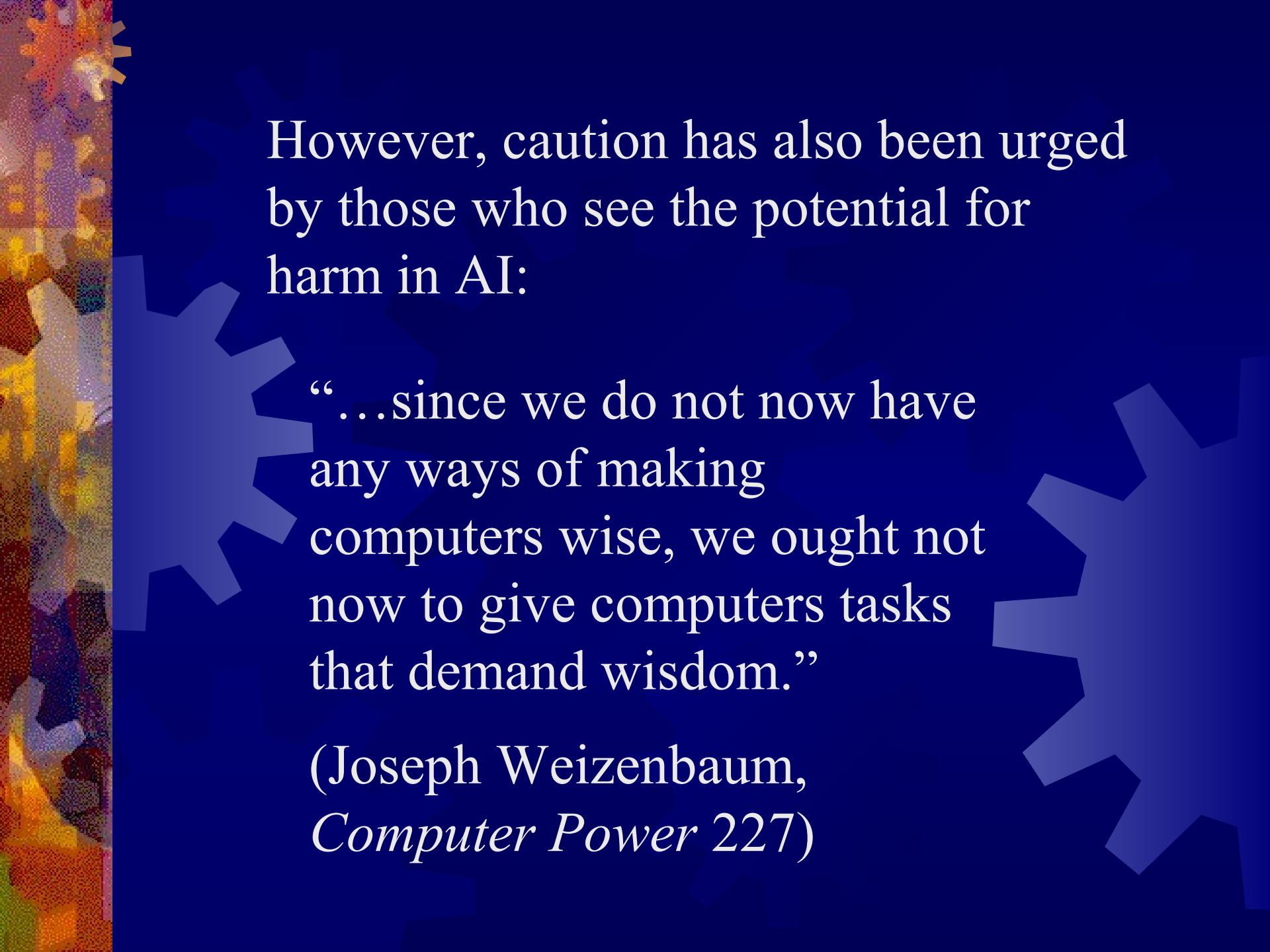




“If every tool, when ordered, or even of its own accord, could do the work that befits it, just as the creations of Daedalus moved of themselves . . . If the weavers' shuttles were to weave of themselves, then there would be no need either of apprentices for the master workers or of slaves for the lords.”

Aristotle

*(from Atheniensium Republica)*



However, caution has also been urged  
by those who see the potential for  
harm in AI:

“...since we do not now have  
any ways of making  
computers wise, we ought not  
now to give computers tasks  
that demand wisdom.”

(Joseph Weizenbaum,  
*Computer Power* 227)

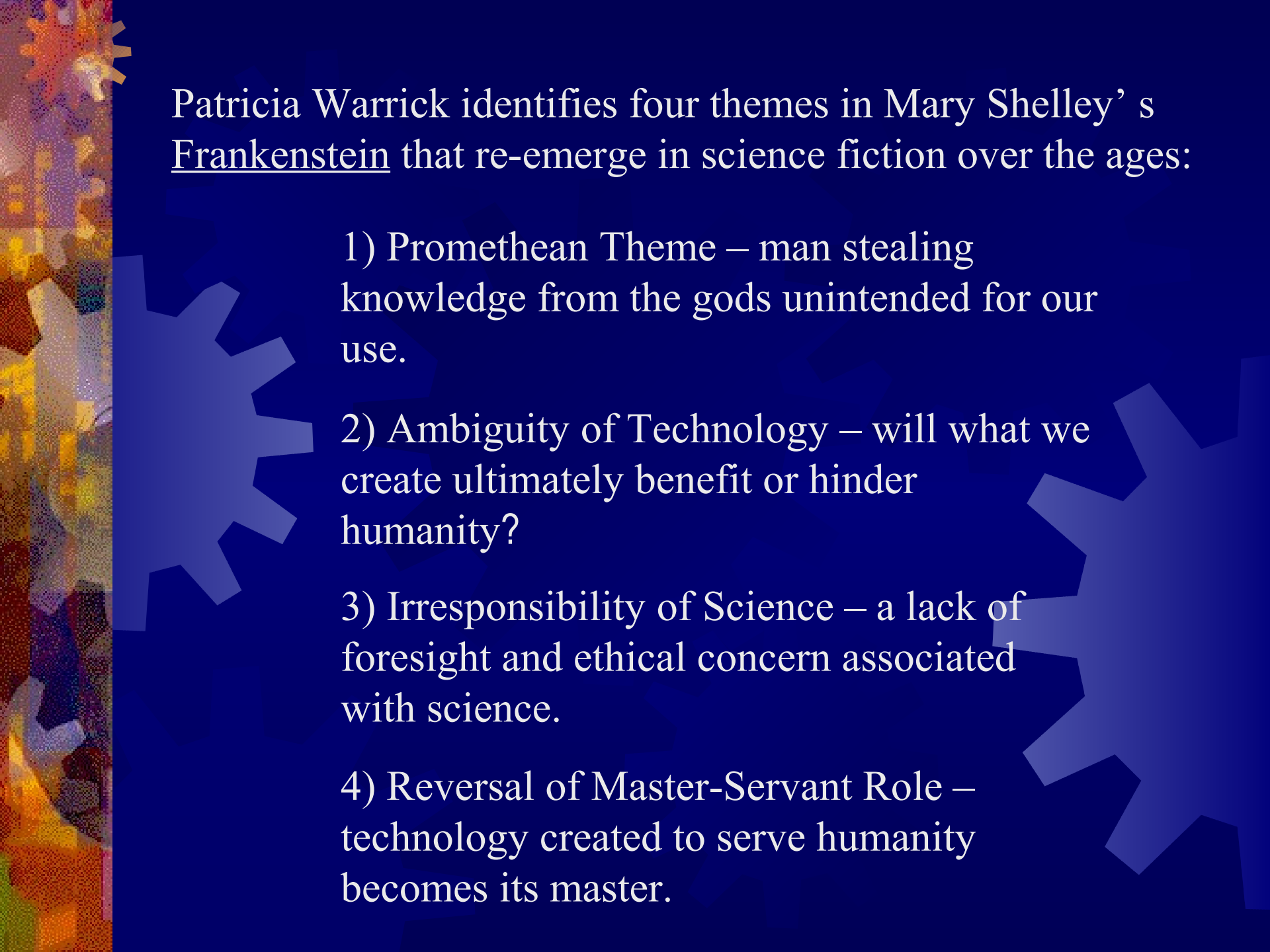
In this presentation, I will focus on the ethical dilemmas associated with AI as they αναφερθείτε to human beings...



...not our creations.

What are our fears?

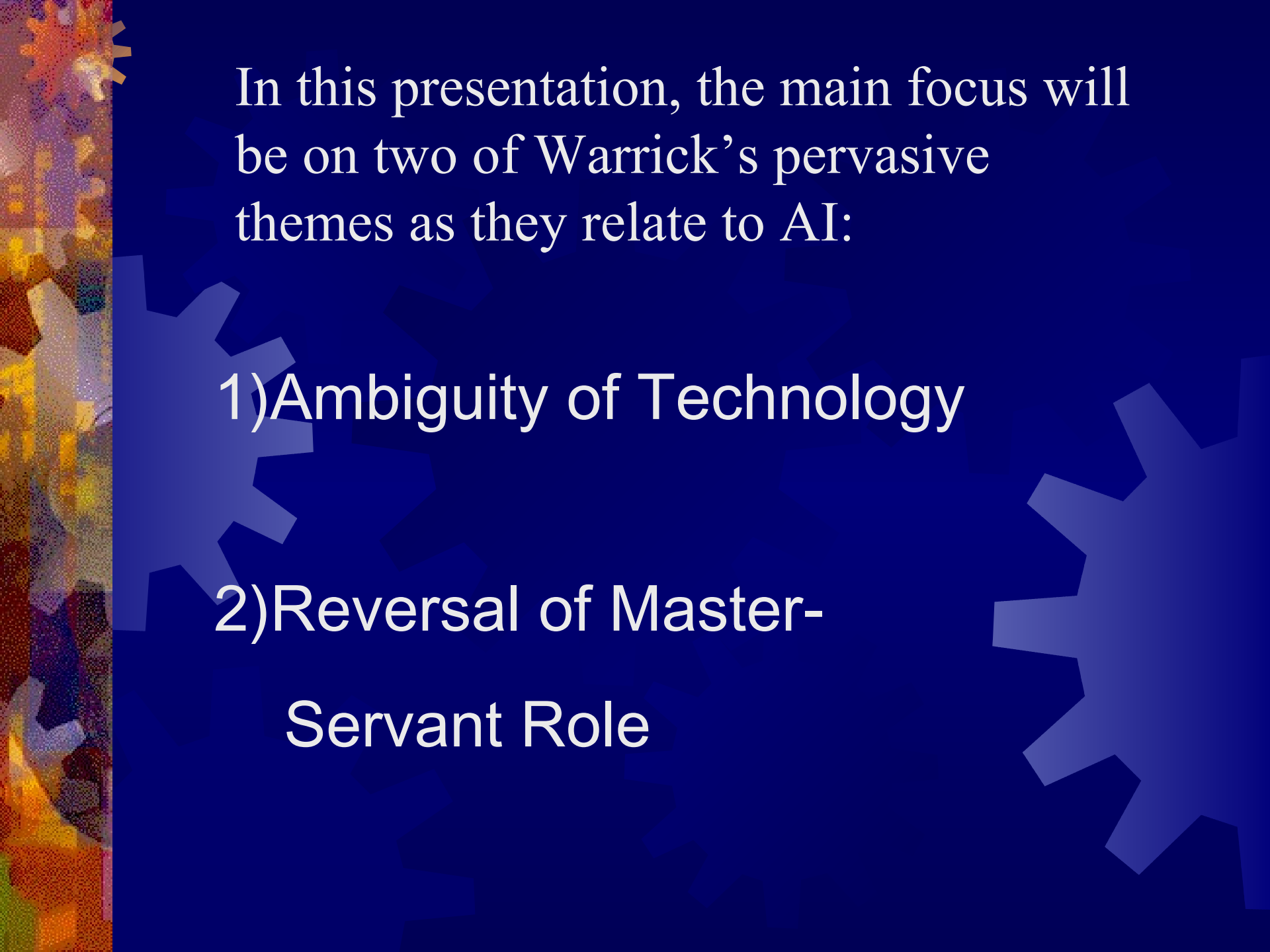




Patricia Warrick identifies four themes in Mary Shelley's Frankenstein that re-emerge in science fiction over the ages:

- 1) Promethean Theme – man stealing knowledge from the gods unintended for our use.
- 2) Ambiguity of Technology – will what we create ultimately benefit or hinder humanity?
- 3) Irresponsibility of Science – a lack of foresight and ethical concern associated with science.
- 4) Reversal of Master-Servant Role – technology created to serve humanity becomes its master.





In this presentation, the main focus will be on two of Warrick's pervasive themes as they relate to AI:

1) Ambiguity of Technology

2) Reversal of Master-Servant Role



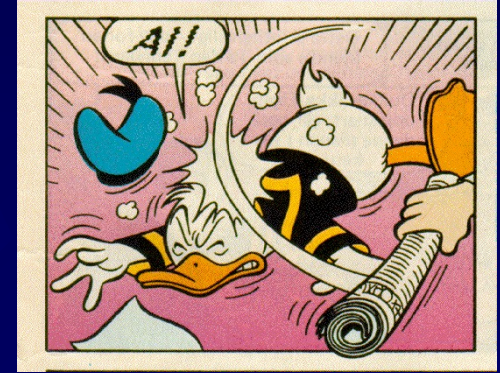
In her article addressing the ambiguity we feel towards the role of technology in our lives, Margaret Boden raises some interesting points.



# On the positive side...



- ★ AI could expand our capacity for knowledge, creativity, and problem solving
- ★ In the classroom, computer-assisted AI wouldn't undermine student progress and self-esteem
- ★ AI may “rehumanize” society as we have more time to devote to humanistic endeavours



# On the negative side...

- ★ Without political or social controls, AI may endanger humanity (ie. military defense) or impede our autonomy and privacy (ie. government databases)
- ★ What would be the social costs of an automated work force where the work of a few supported the many?
- ★ How would such changes affect who we are and our attitudes towards what we value?

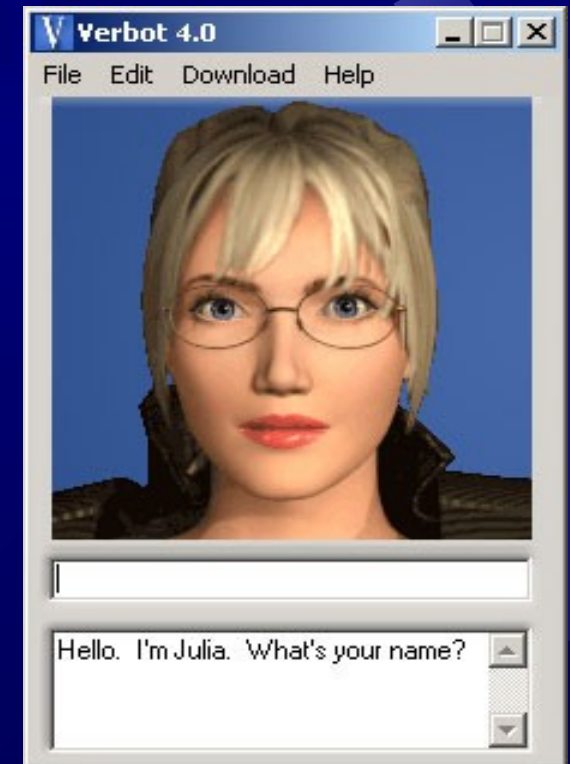
Nick Bostrom suggests other frightening insights related to AI:

*“Artificial minds can be easily copied.”*

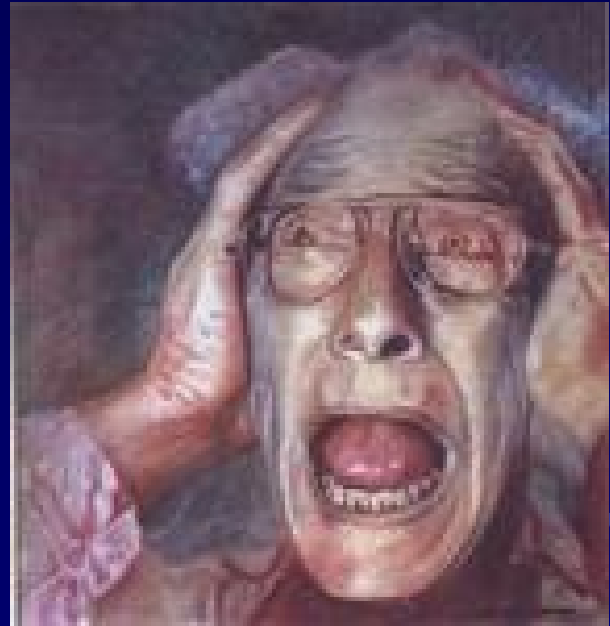
*“Human-level artificial intelligence leads quickly to greater-than-human-level artificial intelligence.”*

*“Technological progress in other fields will be accelerated by the arrival of artificial intelligence.”*

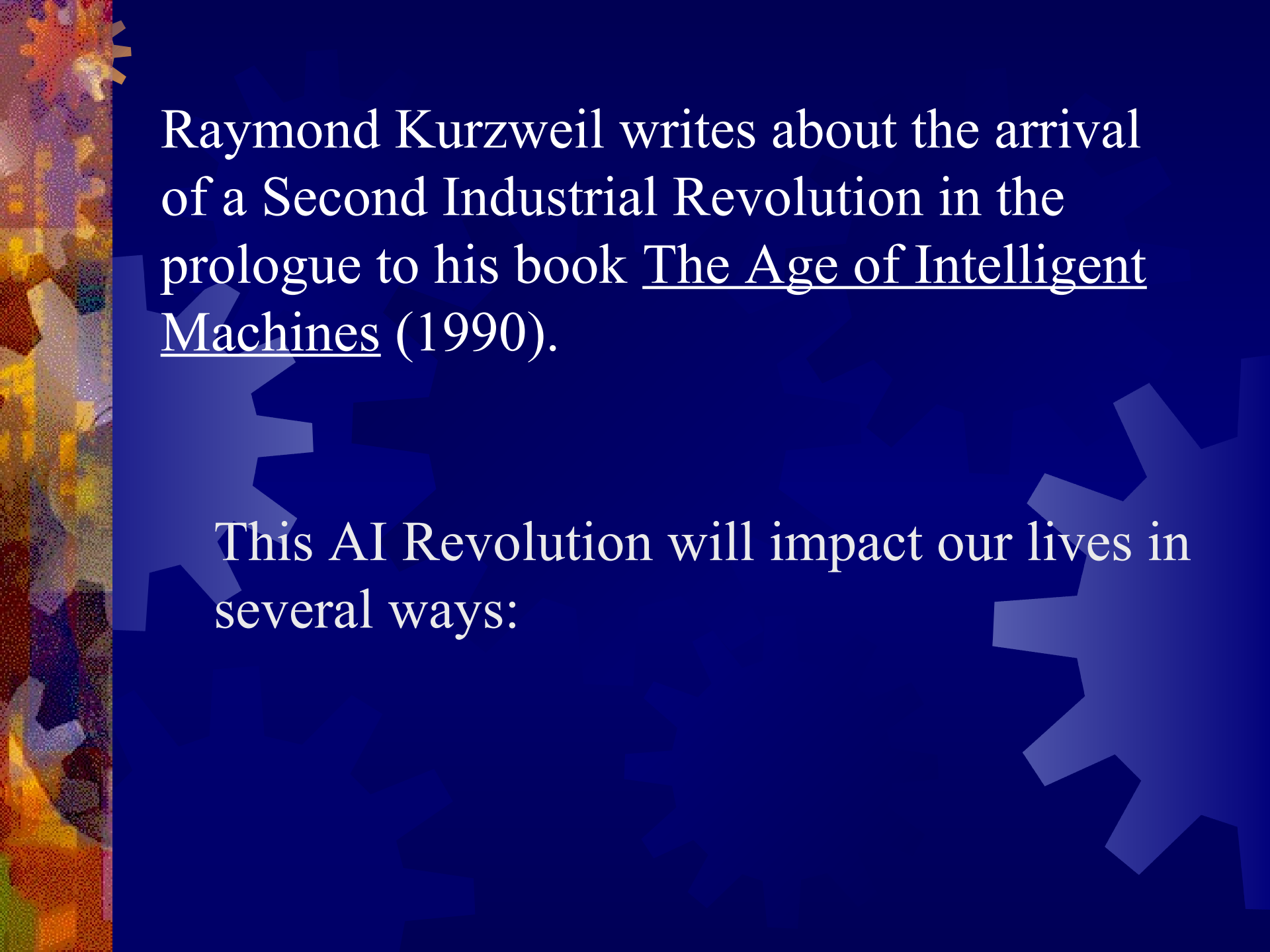
*“Unlike other technologies, artificial intelligences are not merely tools. They are potentially independent agents.”*



**Yikes!!**



Might this lead to a reversal of  
Master-Servant roles?



Raymond Kurzweil writes about the arrival of a Second Industrial Revolution in the prologue to his book The Age of Intelligent Machines (1990).

This AI Revolution will impact our lives in several ways:



# Economic Impact



- ★ Scarcity of jobs (especially mentally unstimulating ones)
- ★ Fewer blue collar and white collar jobs
- ★ Challenge to our Protestant work ethic
- ★ Countries that discourage independent thought in their citizens will be at a disadvantage when AI revolutionizes menial work
- ★ Countries with access to technology and resources will be key players in the future



# Educational Impact

- ★ Machines in the classroom would be non-judgemental teachers incapable of errors
- ★ Machines may not motivate or inspire students
- ★ A change in educational focus to “learning how to learn”
- ★ A need for “knowledge workers” and others to continue learning throughout their lifetimes



# Medical Impact



- ✦ Devices designed to give advanced warning before our bodily systems are endangered
- ✦ Doctors and nurses relegated to marginalized roles
- ✦ People take more responsibility for their own health using machines to diagnose their problems and propose treatment
- ✦ Silicon chips to replace damaged nerve cells
- ✦ Robotic surgeons to increase precision

# Military Impact


- ★ Laser and particle beam weapons
- ★ Smart weapons
- ★ Planes without pilots



# Social / Re-humanizing Impact

- ★ Less need for an urbanized workforce as people move away from cities
- ★ More devices available to assist handicapped people
- ★ Increase in the value of interpersonal relations as machines are impersonal
- ★ Direct relationship between increased leisure time and increased social contact
- ★ Impact of less competition in the workforce





I believe that in the short term, AI will continue to bring us many useful devices that will enhance our quality of life:

Medical advances

Leisure opportunities

Communication devices

Entertainment options

Labour-saving devices



Human nature being what it is, however, at some point in the near future we will be faced with the question:

Just because we CAN do something,  
does that mean we SHOULD?



Who will make these decisions?

On whose authority?



Once we've reached this point,  
will it be too late to choose a  
different path?



The decisions we make today are all steps towards a future we may not be prepared for...



...will we recognize ourselves by the end of this century?

## References

Boden, M. "The Age of Intelligent Machines." (1990)

<http://www.kurzweilai.net/meme/frame.html?main=/articles/art0162.html>

Bostrom, N. "When Machines Outsmart Humans." Futures. Vol. 35:7, pp.759-764, 2000.

<http://www.nickbostrom.com/2050/outsmart.html>

Davies, O., ed. The Omni Book of Computers and Robots. New York: Kensington Publishing Corp., 1982.

Kurzweil, R. "The Age of Intelligent Machines" 1990.

<http://www.kurzweilai.net/meme/frame.html?main=/meme/>

Warrick, P. The Cybernetic Imagination in Science Fiction. Cambridge, Mass.: MIT Press, 1980.

Weizenbaum, J. "Computer Power and Human Reason." San Francisco, CA: W.H. Freeman, 1976.

<http://www.gslis.utexas.edu/~palmquis/courses/reviews/amy.htm>