Responsible Machines:

The Opportunities and Challenges of Artificial Intelligence

“No matter how virtual the subject becomes, there is always a body attached” – Stone 1990

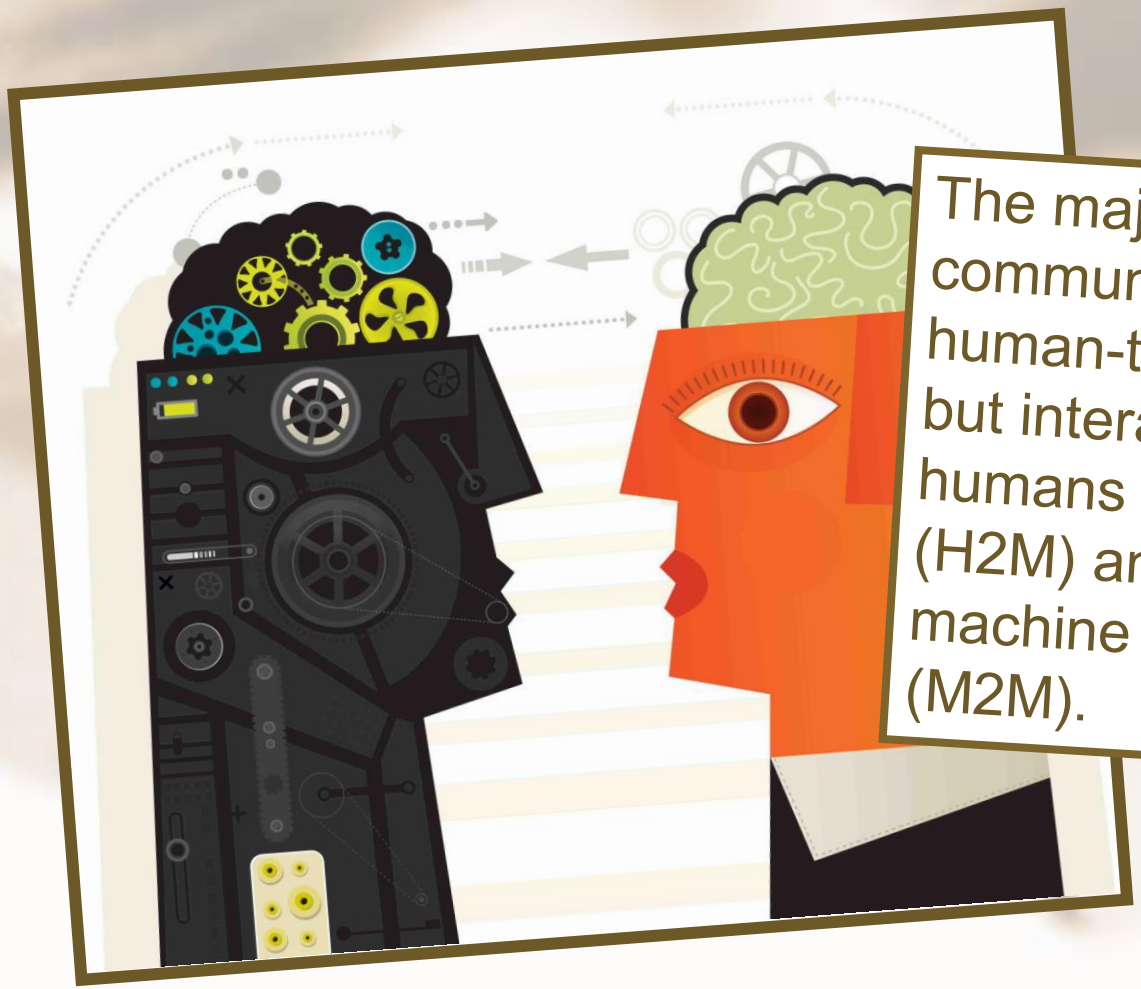
First Conference on Cyberspace

University of Texas - Austin, TX 4-5 May 1990

The Other may appear in the guise of different virtual characters, screen names, profiles, or avatars, but there is always some*body* behind it all.

Internet Folk Wisdom

Internet Folk Wisdom



The majority of online communication is not human-to-human (H2H) but interactions between humans and machines (H2M) and machines-to-machine exchanges (M2M).

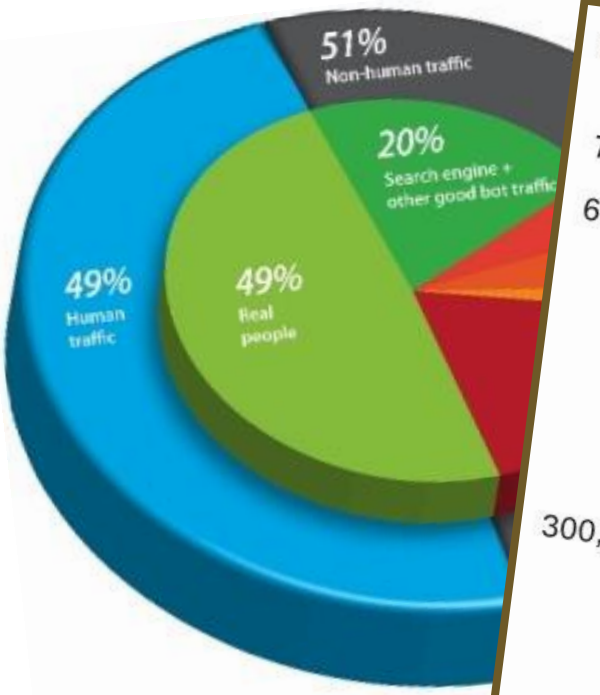
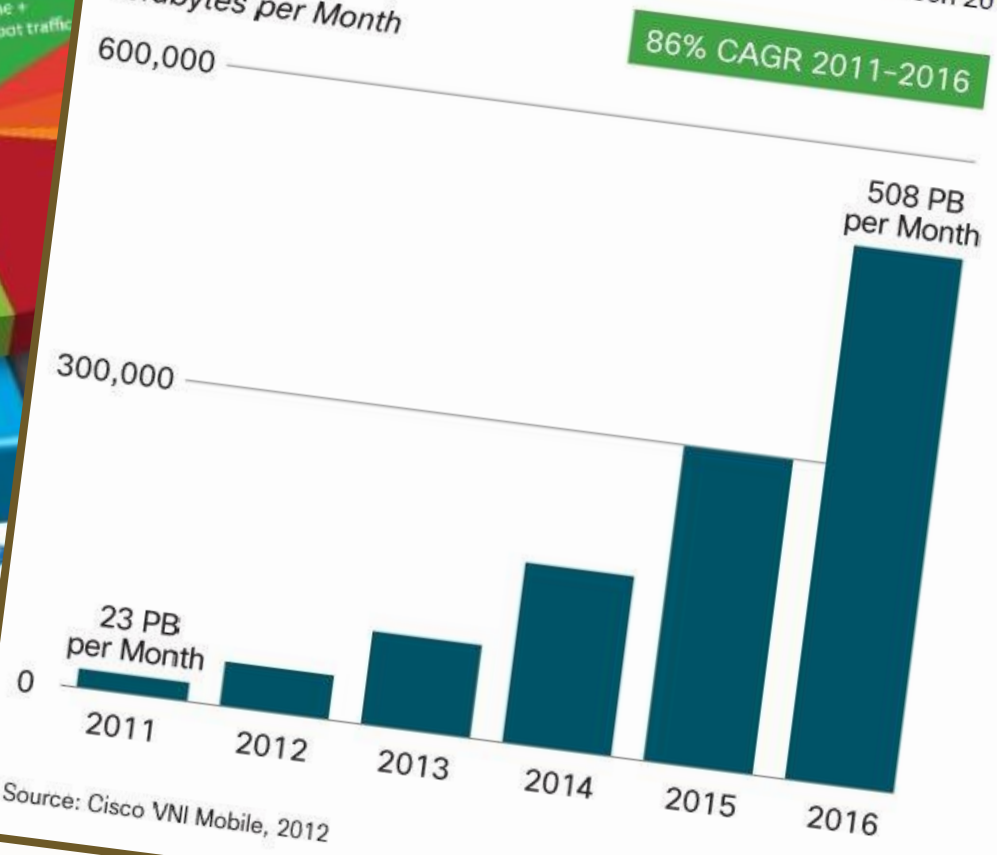
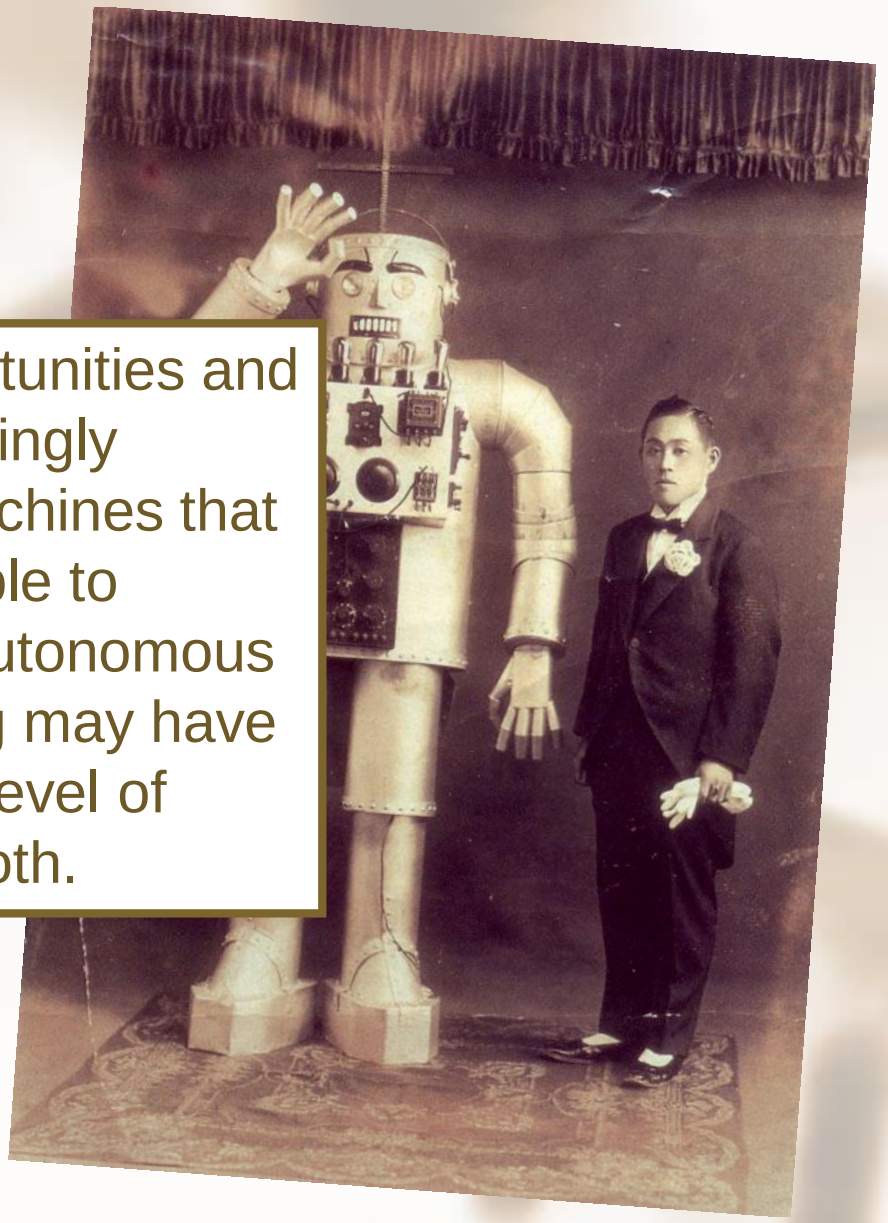


Figure 14. Machine-to-Machine Traffic to Increase 22-Fold Between 2011 and 2016



Source: Cisco VNI Mobile, 2012

Task: Investigate the opportunities and challenges of these increasingly *responsible machines*—machines that are designed for and are able to respond to us as another autonomous social entity and in so doing may have a legitimate claim to some level of rights, responsibilities, or both.



Agenda

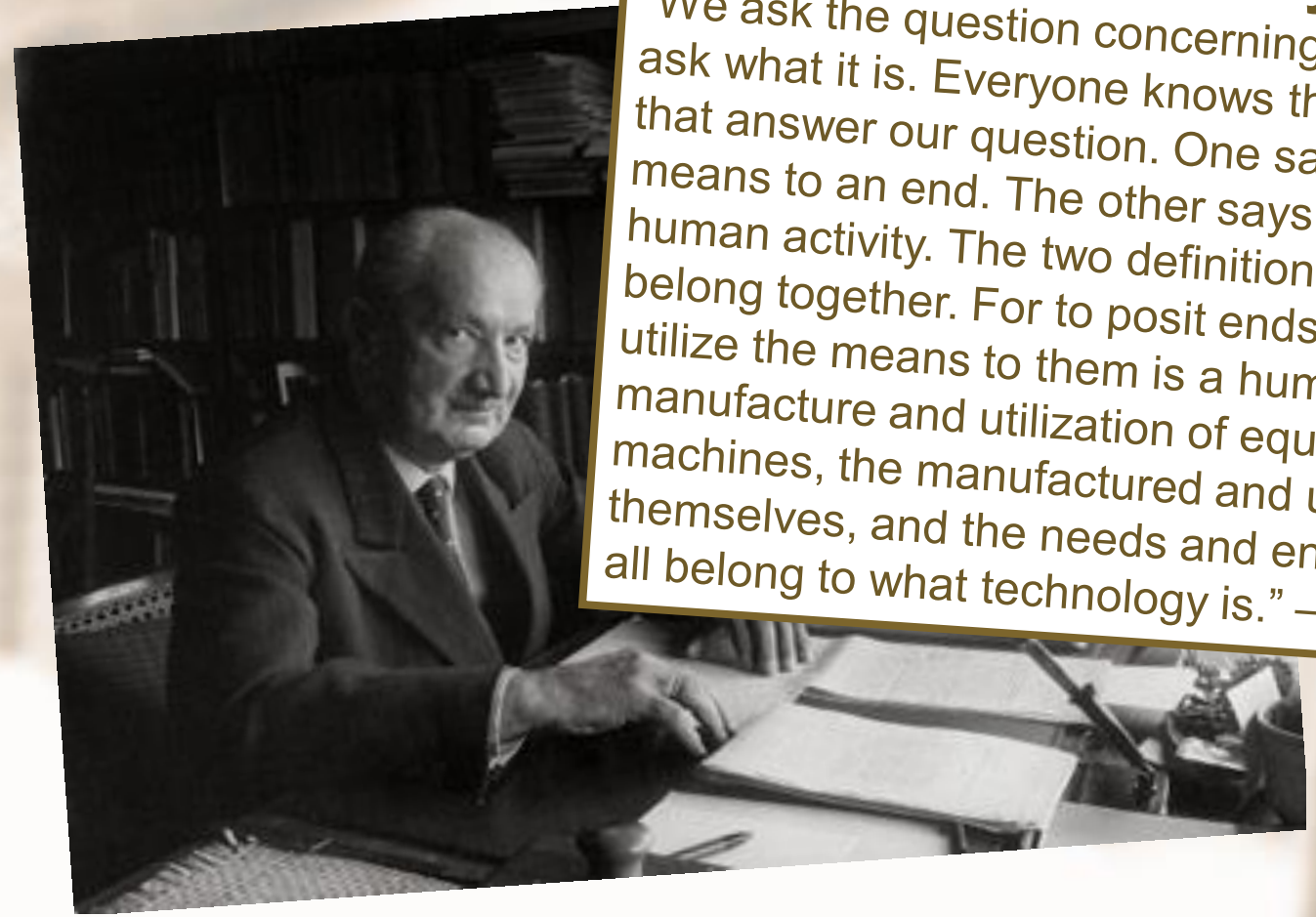
1) Default Setting
The Instrumental Theory of Technology

2) The New Normal
Recent Challenges to the Default Setting

3) Rise of the Machines
Consequences of this Machine Incursion

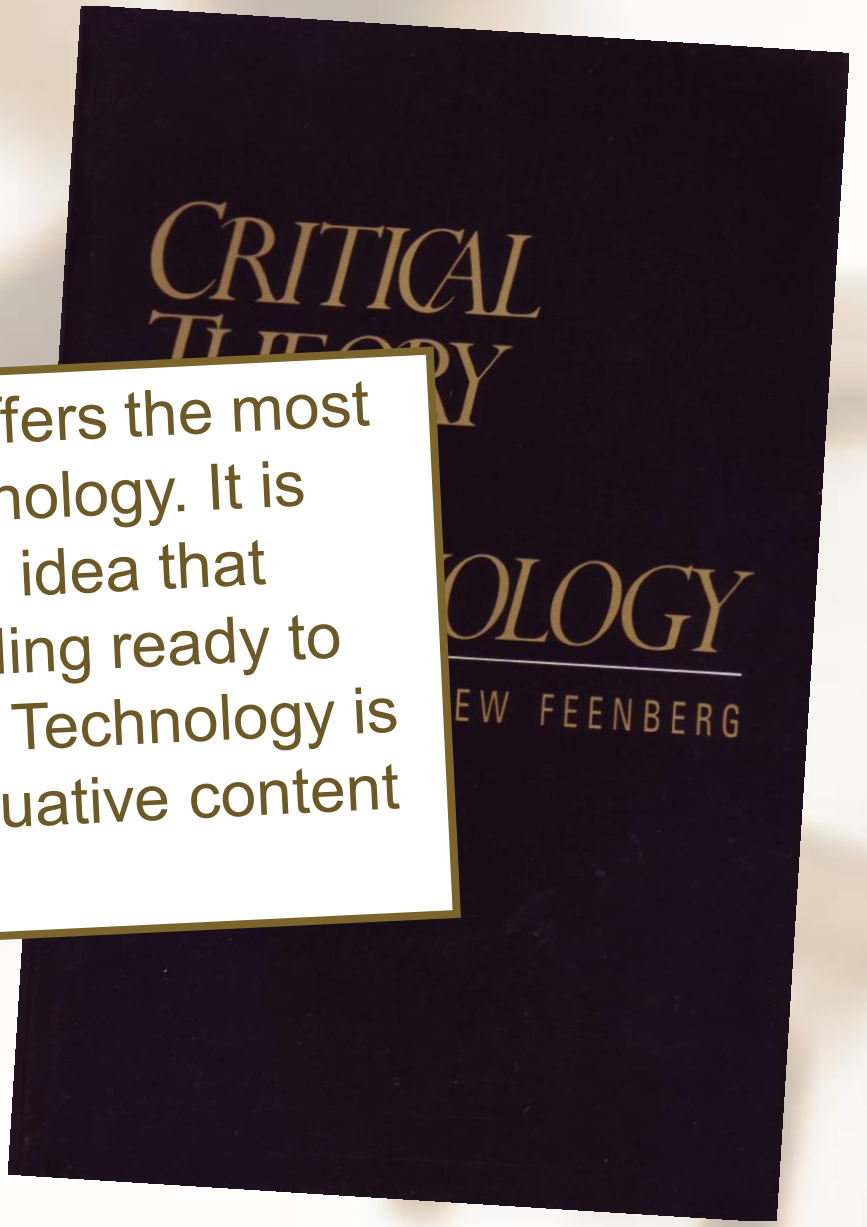
①
Default Setting


37"JOSE MARIA
12



Instrumental Theory

"We ask the question concerning technology when we ask what it is. Everyone knows the two statements that answer our question. One says: Technology is a means to an end. The other says: Technology is a human activity. The two definitions of technology belong together. For to posit ends and procure and utilize the means to them is a human activity. The manufacture and utilization of equipment, tools, and machines, the manufactured and used things themselves, and the needs and ends that they serve, all belong to what technology is." – Heidegger 1954



"The instrumentalist theory offers the most widely accepted view of technology. It is based on the common sense idea that technologies are 'tools' standing ready to serve the purposes of users. Technology is deemed 'neutral,' without valuative content of its own." – Feenberg 1991

Technology = Tool

"Morality rests on human shoulders, and if machines changed the ease with which things were done, they did not change responsibility for doing them. People have always been the only 'moral agents.'" – J. Storrs Hall 2001

Computer Mediated Communication

Basic Communication Model

1. SENDER has a thought
2. SENDER ENCODES thought into a MESSAGE.

4. RECEIVER DECODES message
5. RECEIVER INTERNALIZES message

Computer Mediated Communication

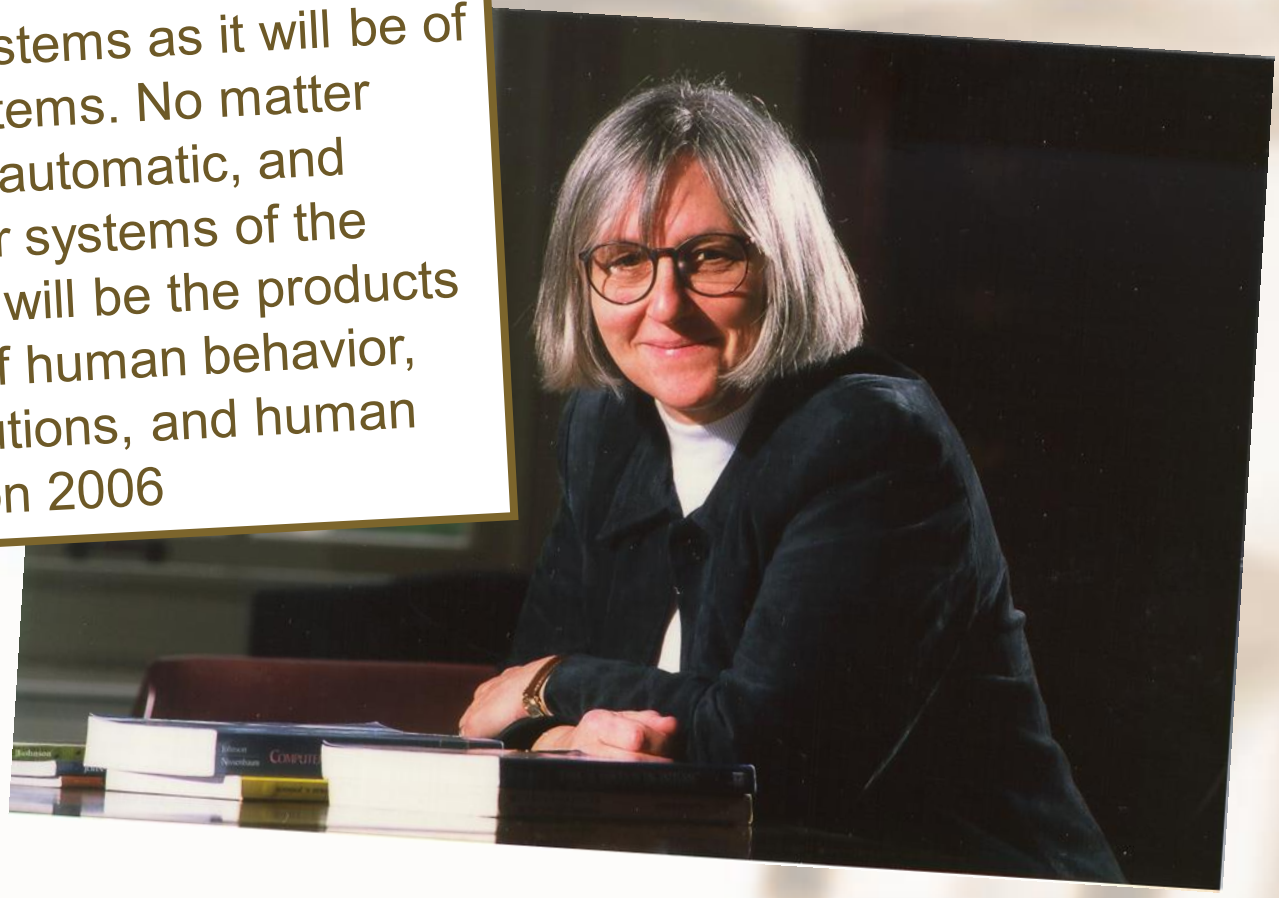
Communication Technology

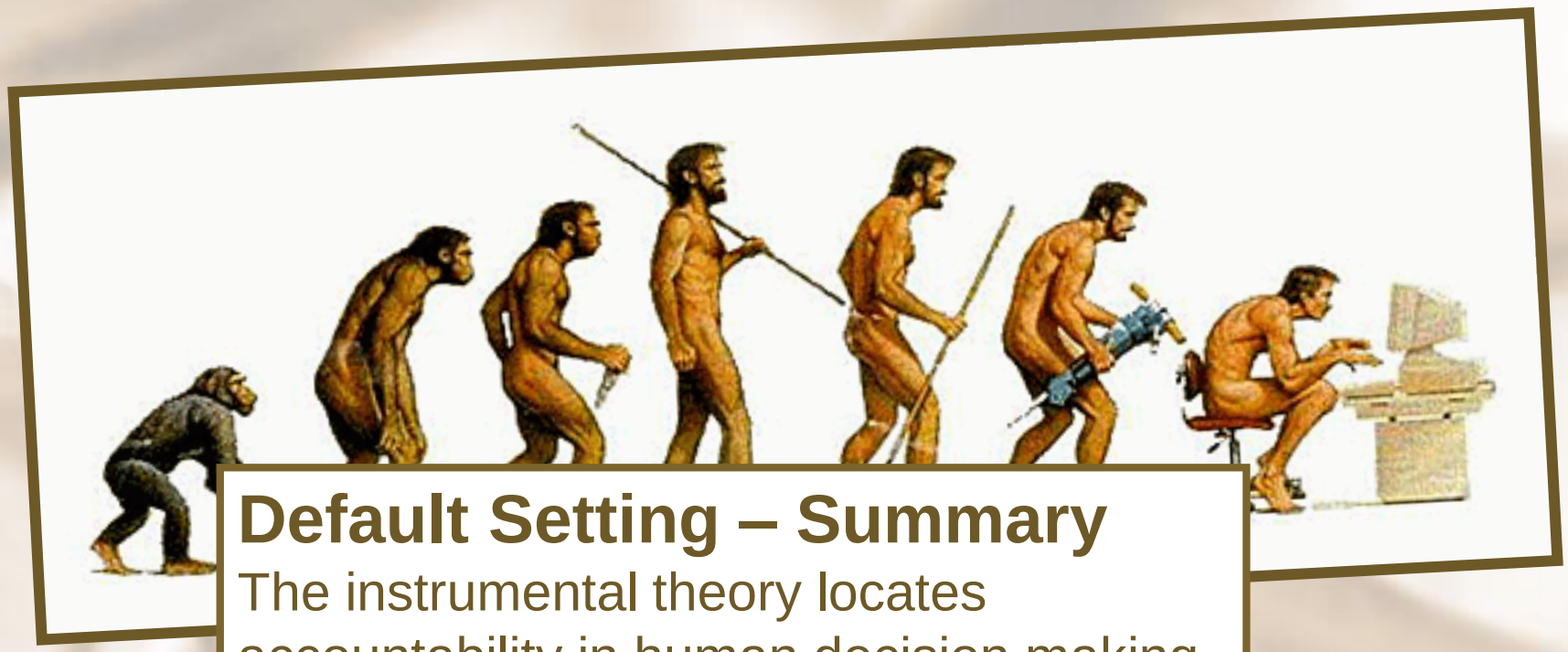
Something through which human messages pass

Logical—Attribute agency to an inanimate object

Moral—Deflect responsibility to a mere instrument or tool

“Computer systems are produced, distributed, and used by people engaged in social practices and meaningful pursuits. This is as true of current computer systems as it will be of future computer systems. No matter how independently, automatic, and interactive computer systems of the future behave, they will be the products (direct or indirect) of human behavior, human social institutions, and human decision.” – Johnson 2006



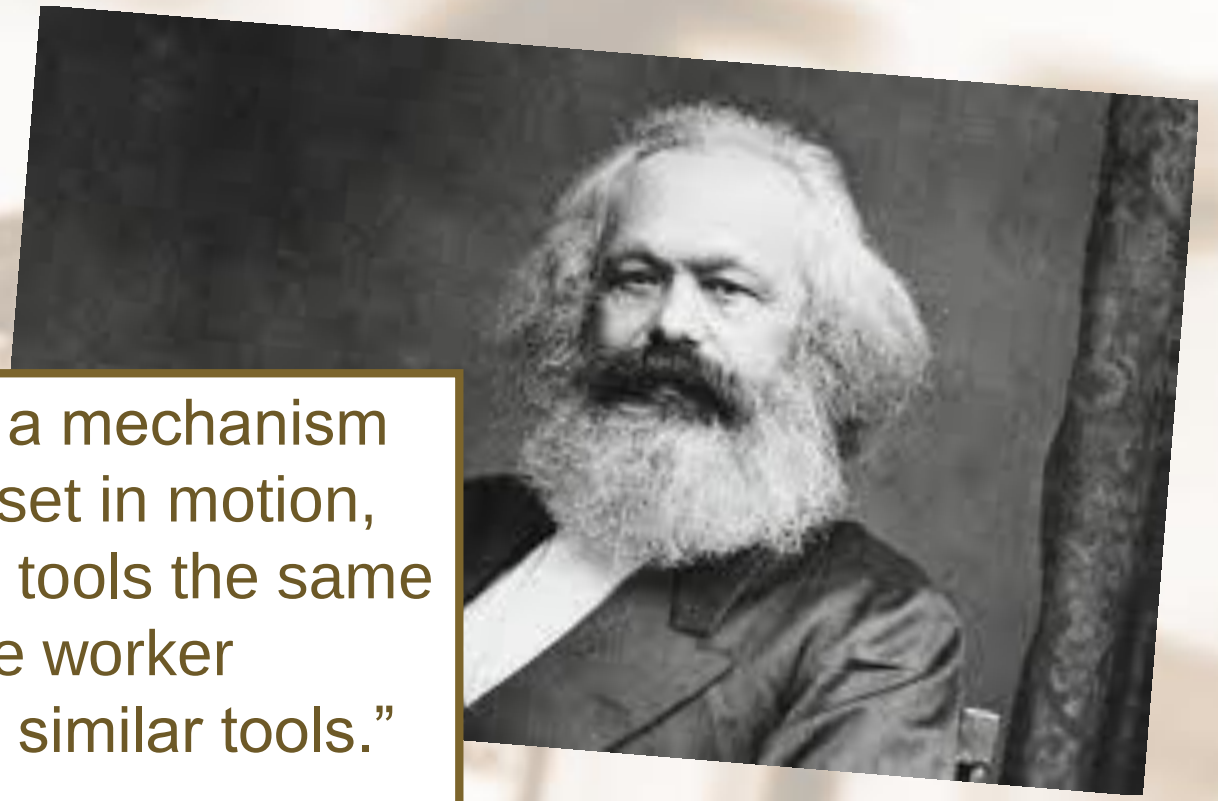


Default Setting – Summary

The instrumental theory locates accountability in human decision making and action, and it resists any and all efforts to defer responsibility to some inanimate object by blaming or scape-goating what are mere tools.

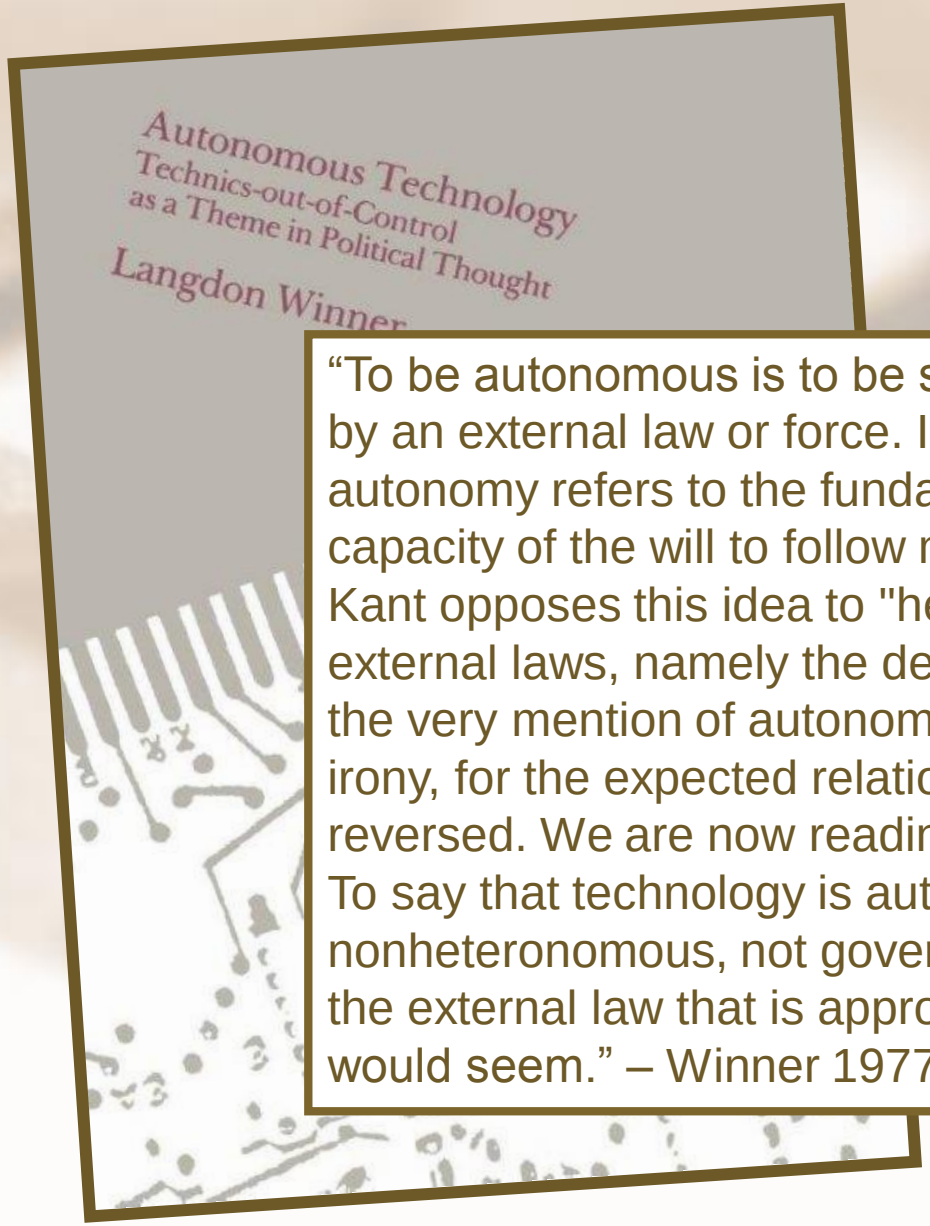
② The New Normal

Machine != Tool



"The machine is a mechanism that, after being set in motion, performs with its tools the same operations as the worker formerly did with similar tools."
– Marx 1867

The machine occupies the position not of the tool but of the human worker—the active agent who had used the tool.



"To be autonomous is to be self-governing, independent, not ruled by an external law or force. In the metaphysics of Immanuel Kant, autonomy refers to the fundamental condition of free will—the capacity of the will to follow moral laws which it gives to itself. Kant opposes this idea to "heteronomy," the rule of the will by external laws, namely the deterministic laws of nature. In this light the very mention of autonomous technology raises an unsettling irony, for the expected relationship of subject and object is exactly reversed. We are now reading all of the propositions backwards. To say that technology is autonomous is to say that it is nonheteronomous, not governed by an external law. And what is the external law that is appropriate to technology? Human will, it would seem." – Winner 1977

Customer Service

Basic Communication Model

1. SENDER has a thought
2. SENDER ENCODES thought into a MESSAGE.

4. RECEIVER DECODES message
5. RECEIVER INTERNALIZES message

Customer Service

Sender

Receiver

Customer Service

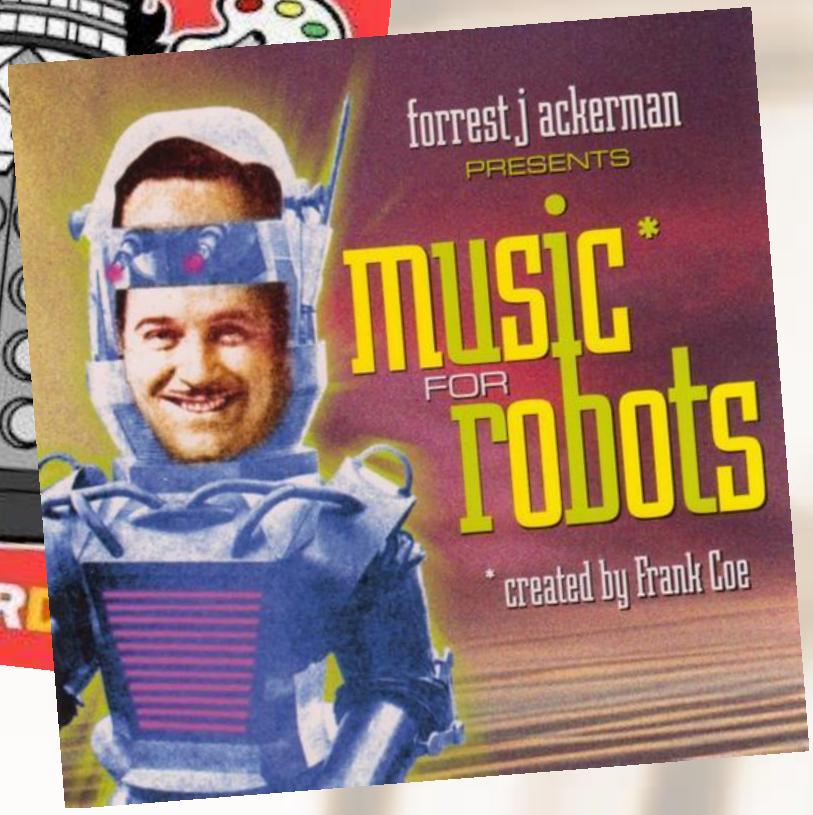
Sender

Medium

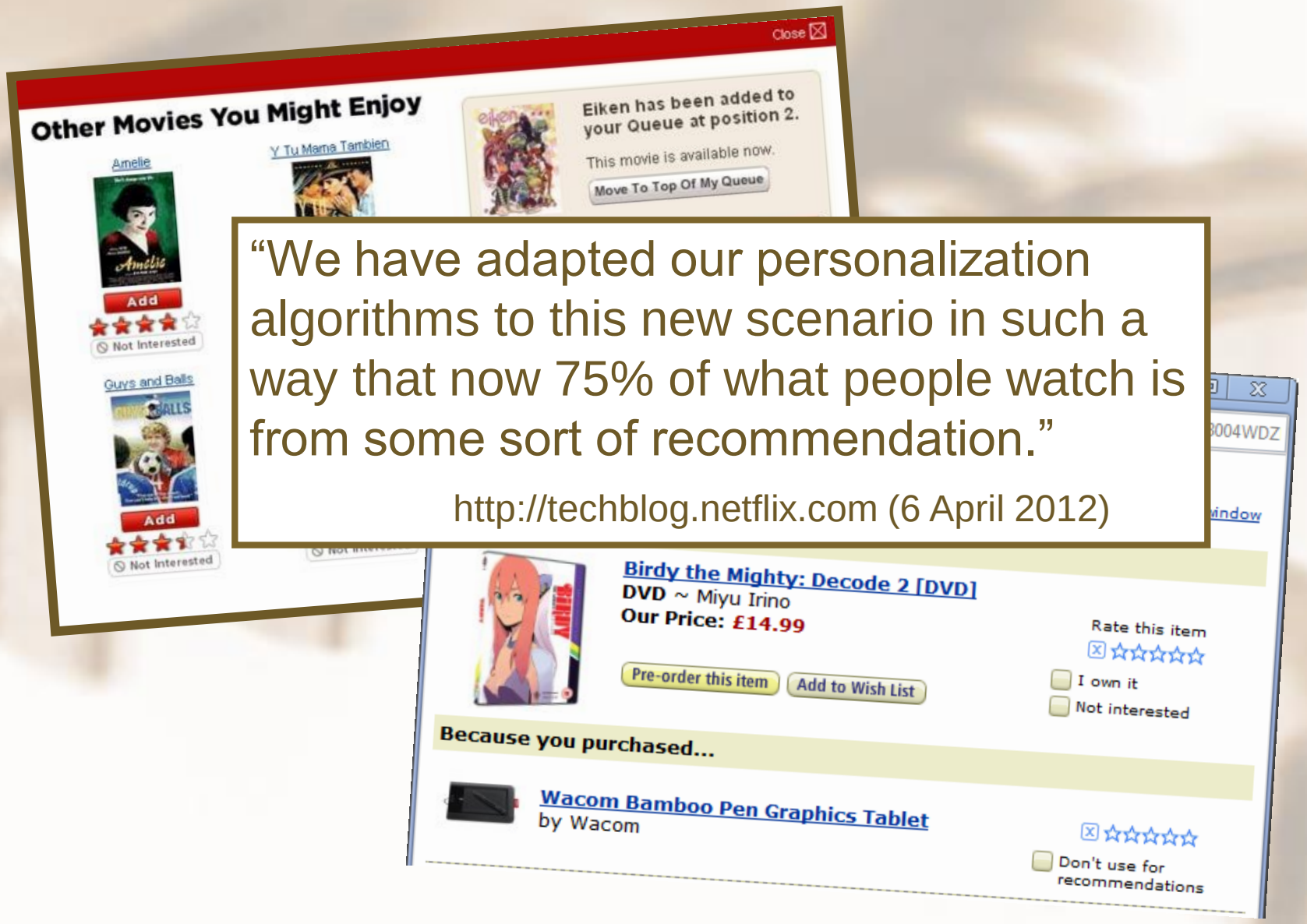
Receiver

Autonomous Artificial Agent

Algorithmic Culture



Algorithmic Culture



Algorithmic Culture

The rise of machine-written journalism - Mozilla Firefox

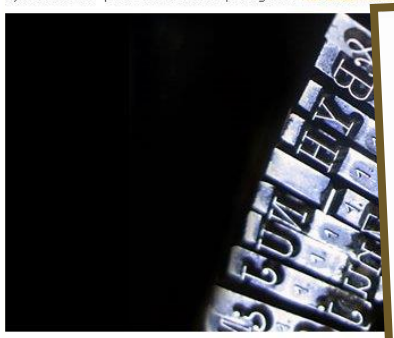
File Edit View History Bookmarks Tools Help

http://www.wired.co.uk/news/archive/2009-12/16/the-rise-of-

The rise of machine-written journali...

The rise of machine-written journalism

By Peter Kirwan | 16 December 2009 | Categories: The Great Transition, Technology



Technological change arrives in waves. Progress is always
persist amid the incoming tide.

The self-employed stocking-makers who lived and worked
early 19th century inhabited one such island. Theirs was a
harder to automate than cotton production.

Despite this, the stockingers were forced to adapt to the rh
that surrounded them. Their craft traditions died a lingering
violence. By 1811, Byron described them as "meagre with f
despair".

Famine isn't a worry for most journalists in the developed w
the news business do resemble the stocking-makers of Not

Transferring data from www.wired.co.uk...

future tense ASU | NEW AMERICA | SLATE

FUTURE TENSE THE CITIZEN'S GUIDE TO THE FUTURE MARCH 17 2014 5:30

The First News Rep
the L.A. Earthquake
Written by a Robot

By Will Oremus

The *LAT*'s Quakebot finished its story in seconds flat.

iurii / Shutterstock.com

Ken Schwencke, a journalist and programmer for the *Los Angeles Times*, was jolted awake at 6:25 a.m. on Monday by an earthquake. He rolled out of bed and went straight to his computer, where he found a brief story about the quake already written and

ENTER THE ROBOT JOURNALIST
Users' perceptions of automated content

Christer Clerwall

The advent of new technologies has always spurred questions about changes in journalism—its content, its means of production, and its consumption. A quite recent development in the realm of digital journalism is software-generated content, i.e. automatically produced content. This paper seeks to investigate how readers perceive software-generated content in relation to similar content written by a journalist. The study utilizes an experimental methodology where respondents were subjected to different news articles that were written either by a journalist or were software-generated. The respondents were then asked to answer questions about how they perceived the article—its overall quality, credibility, objectivity, etc. The paper presents the results from an initial small-scale study with findings suggesting that while the software-generated content is perceived as descriptive and boring, it is also considered to be objective although not necessarily discernible from content written by journalists. The paper discusses the results of the study and its implication for journalism practice.

KEYWORDS automated content; experimental study; online journalism; robot journalism

Introduction

Our technology humanizes big data sets by spotting patterns, trends and key insights and describing those findings in plain English that is indistinguishable from that produced by a human writer. (Automated Insights 2012)

Algorithmic Culture

How Quill Works

Identifies What Matters

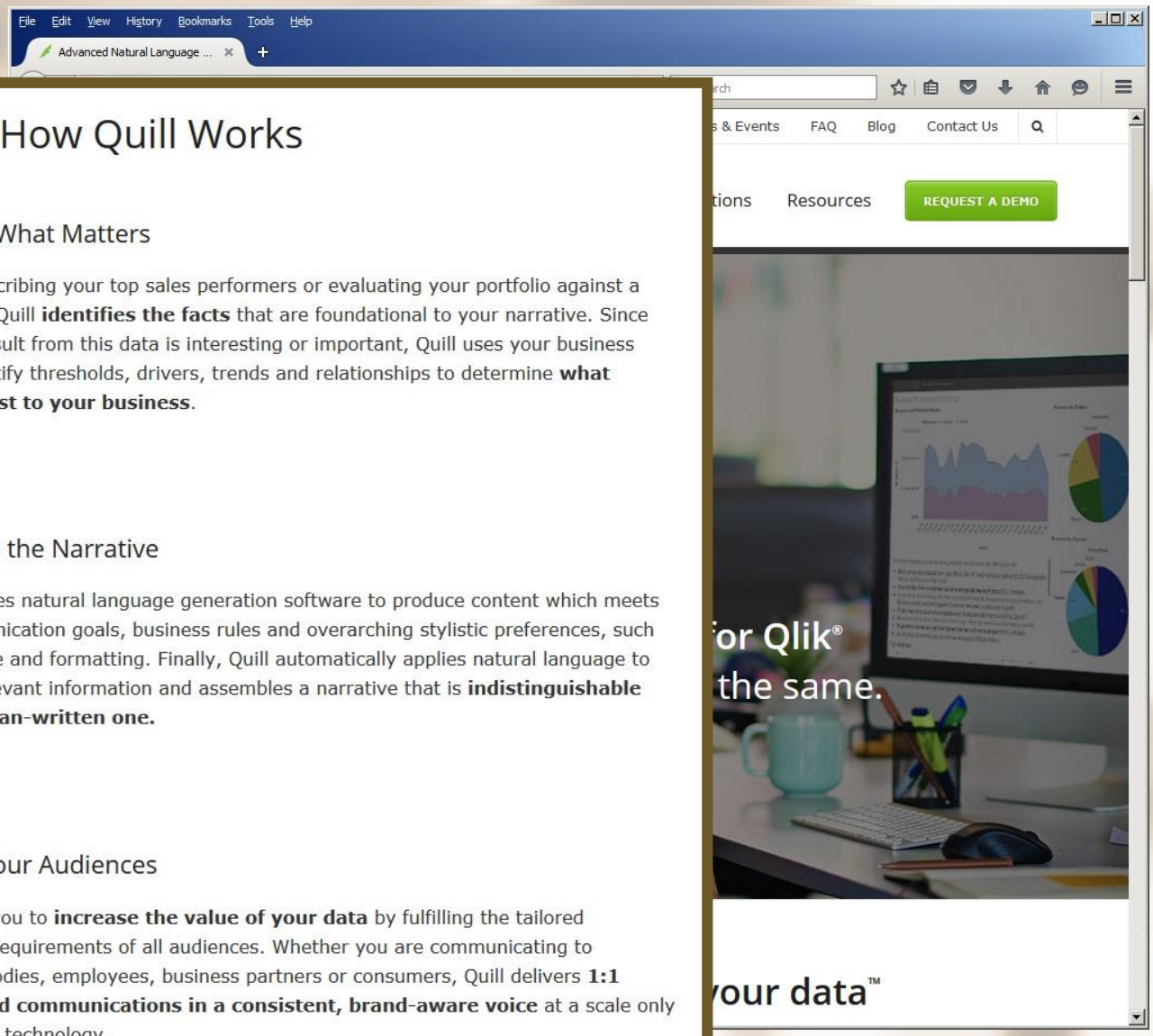
Whether describing your top sales performers or evaluating your portfolio against a benchmark, Quill **identifies the facts** that are foundational to your narrative. Since not every result from this data is interesting or important, Quill uses your business rules to identify thresholds, drivers, trends and relationships to determine **what matters most to your business**.

Generates the Narrative

Quill leverages natural language generation software to produce content which meets your communication goals, business rules and overarching stylistic preferences, such as tone, style and formatting. Finally, Quill automatically applies natural language to the most relevant information and assembles a narrative that is **indistinguishable from a human-written one.**

Informs Your Audiences

Quill allows you to **increase the value of your data** by fulfilling the tailored information requirements of all audiences. Whether you are communicating to regulatory bodies, employees, business partners or consumers, Quill delivers **1:1 personalized communications in a consistent, brand-aware voice** at a scale only possible with technology.



Algorithmic Culture

Shimon - Georgia Tech

David Cope’s EMI

Algorithmic Culture

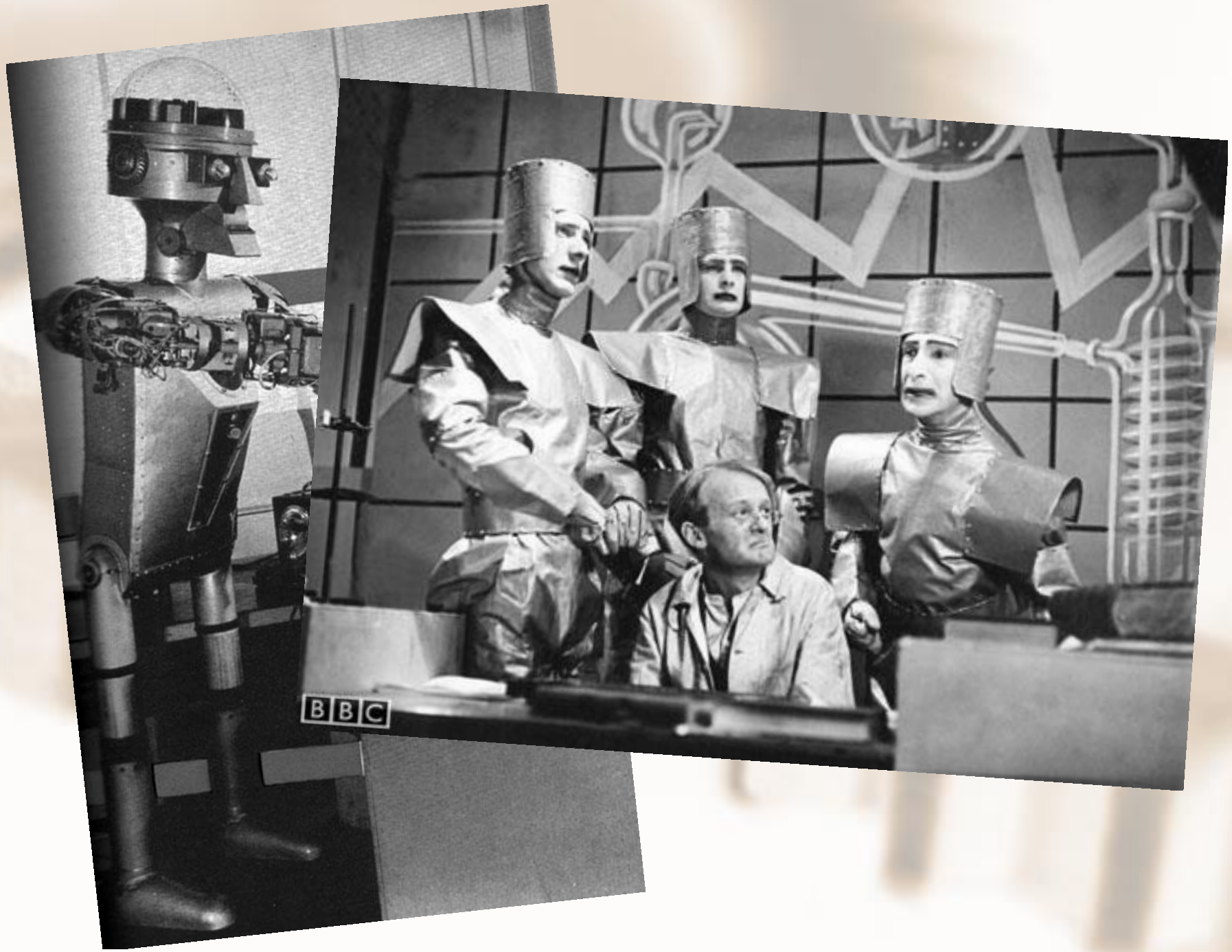
Human-Robot Jazz Improvisation

Guy Hoffman
Gil Weinberg

Social Robots

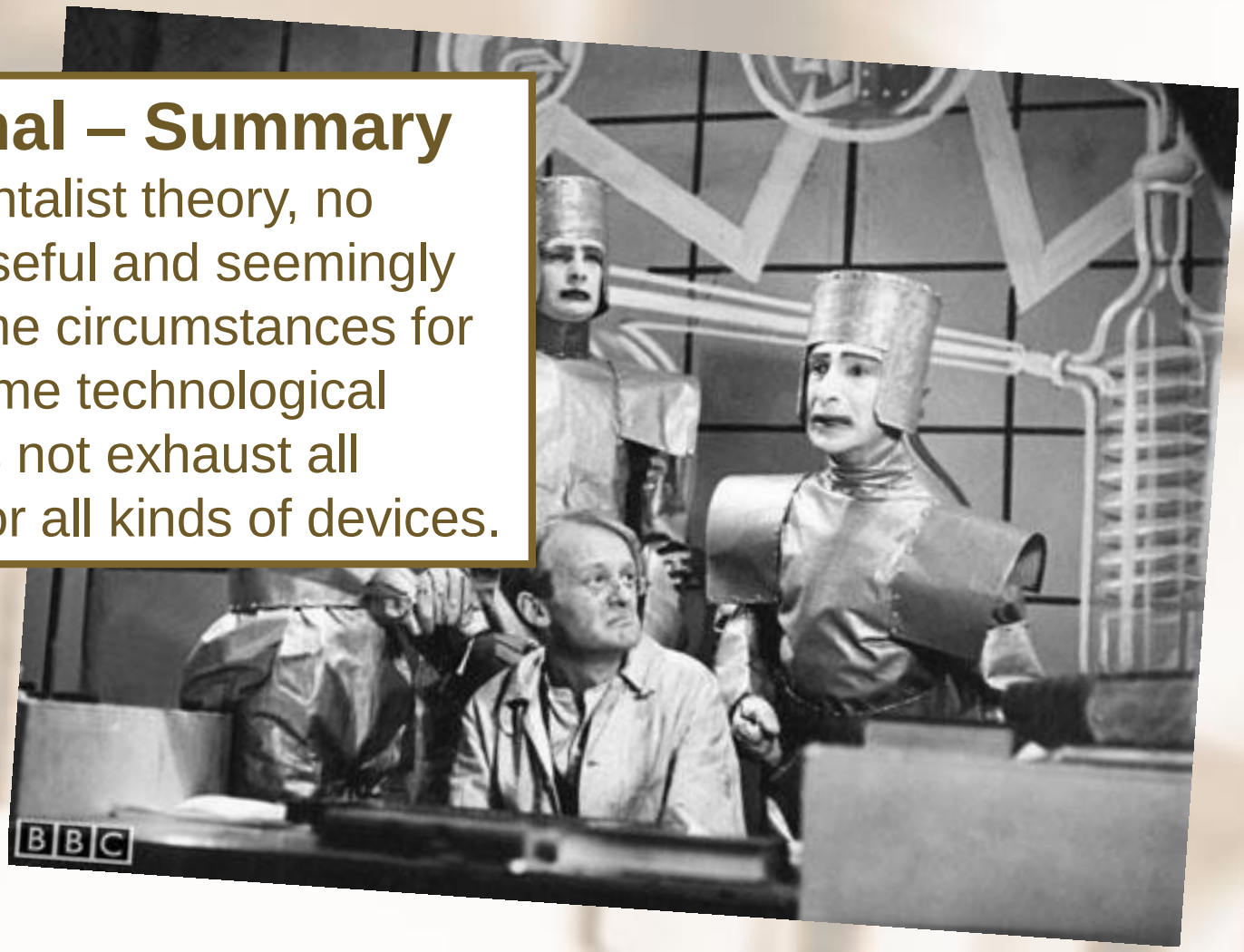
Social Robots are socially intelligent robot partners that interact with humans to promote social and intellectual benefits, work alongside with humans as peers, learn from people as apprentices, and foster more engaging interaction between people.

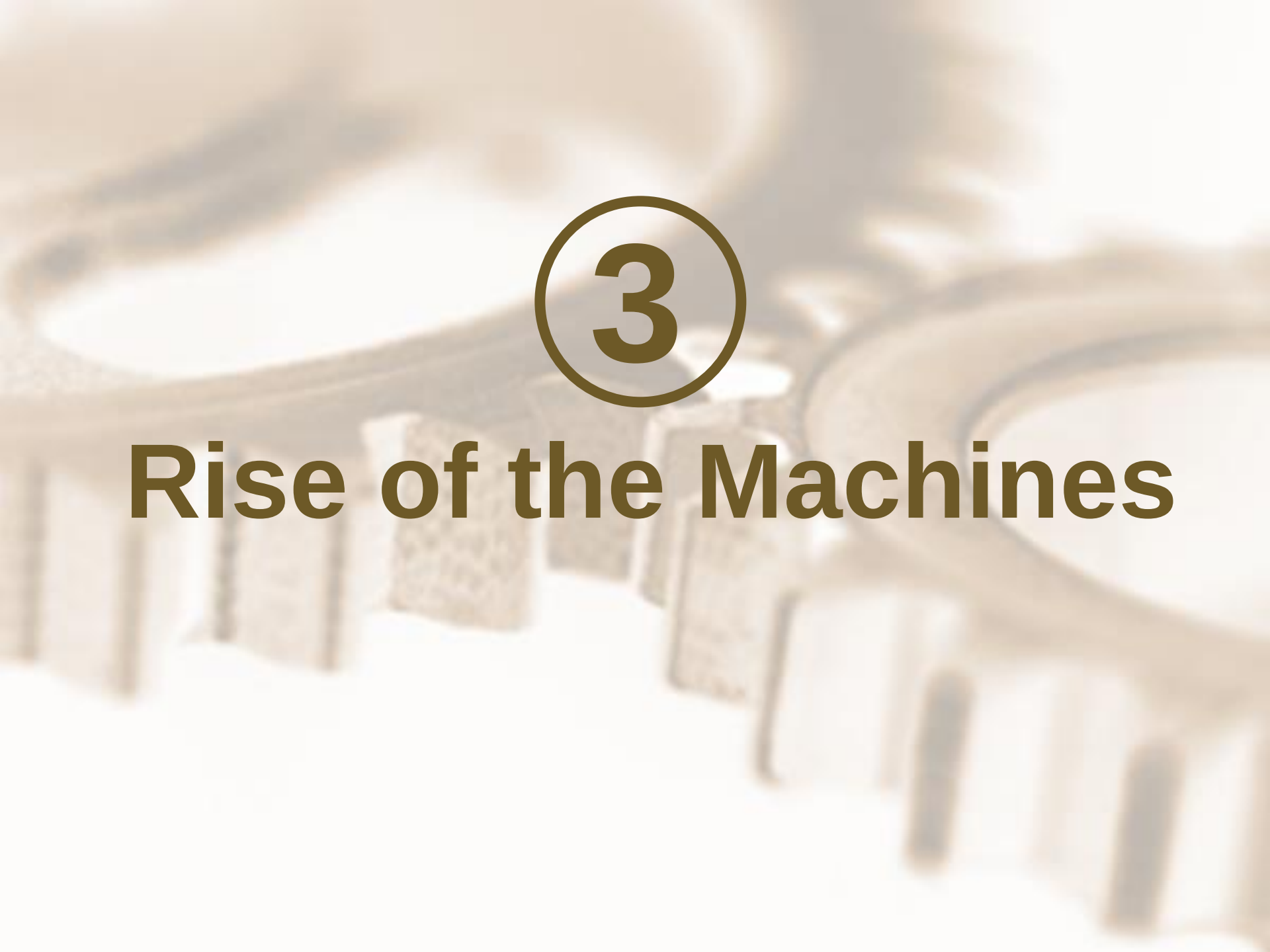
Social Robots

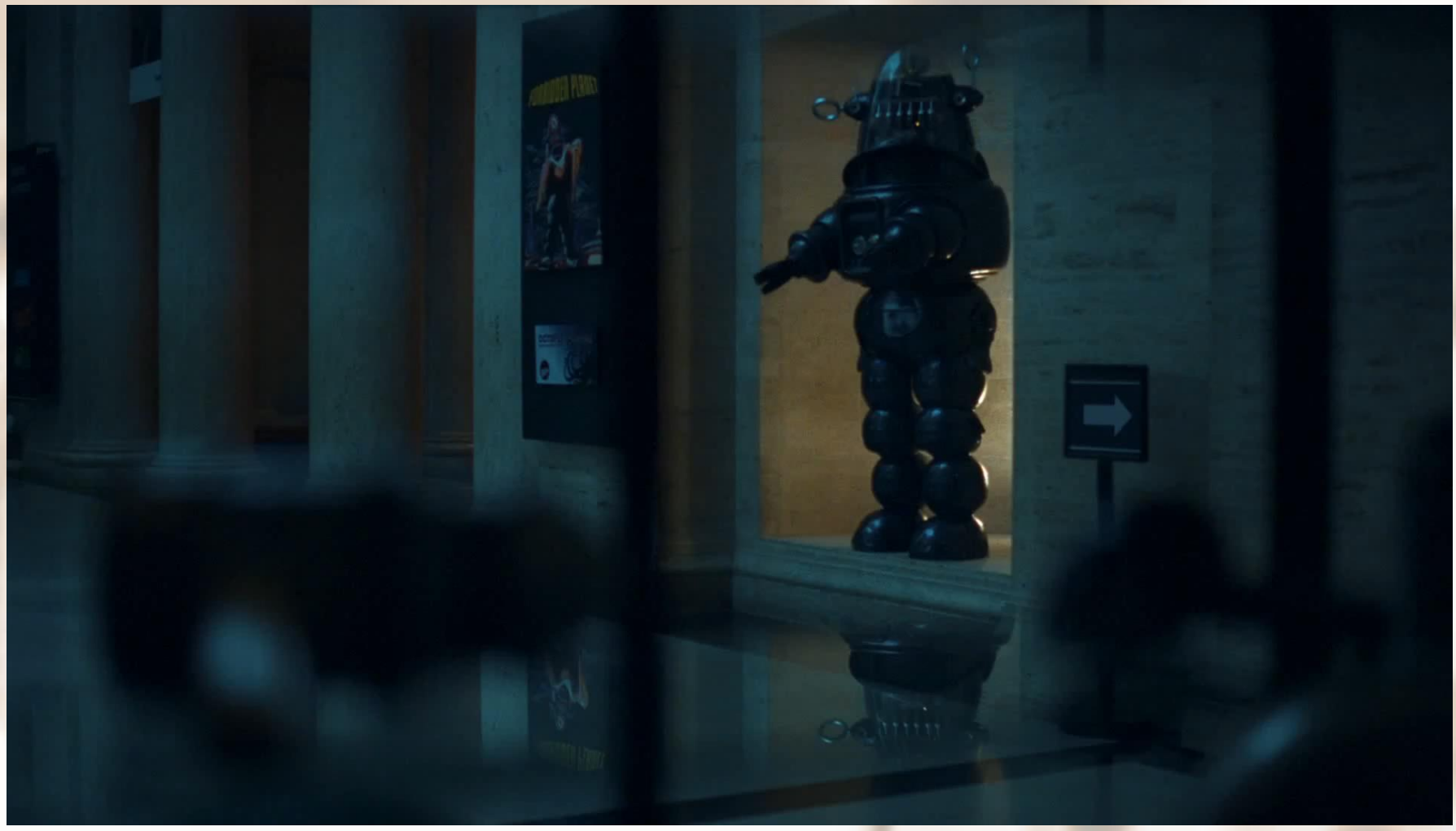

BBC

New Normal – Summary

The instrumentalist theory, no matter how useful and seemingly correct in some circumstances for explaining some technological devices, does not exhaust all possibilities for all kinds of devices.

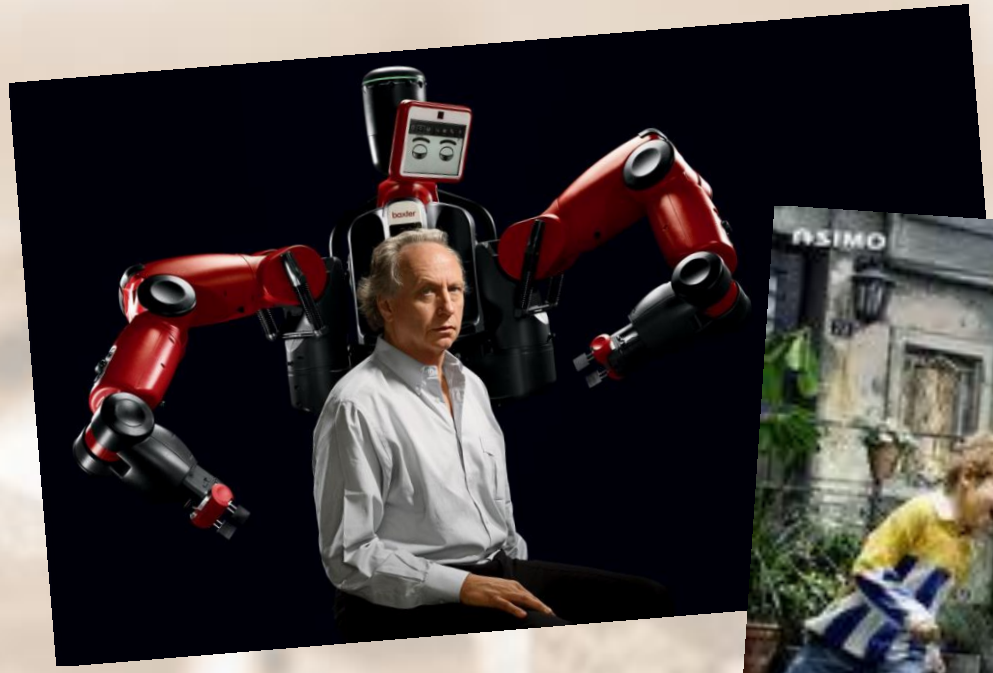



3
Rise of the Machines

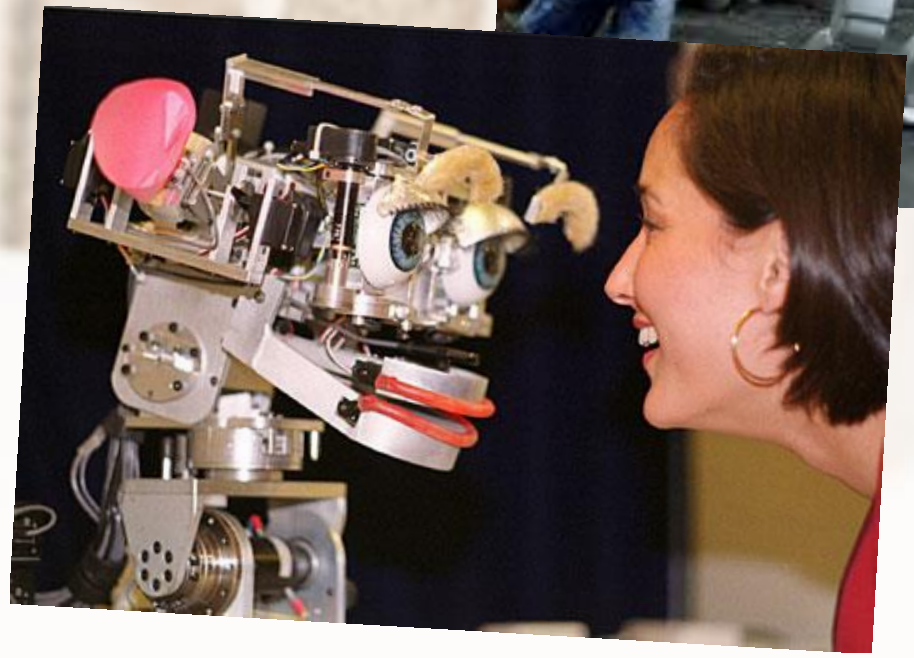


“Robots on the Move,” General Electric, November 2012

Robot Invasion!

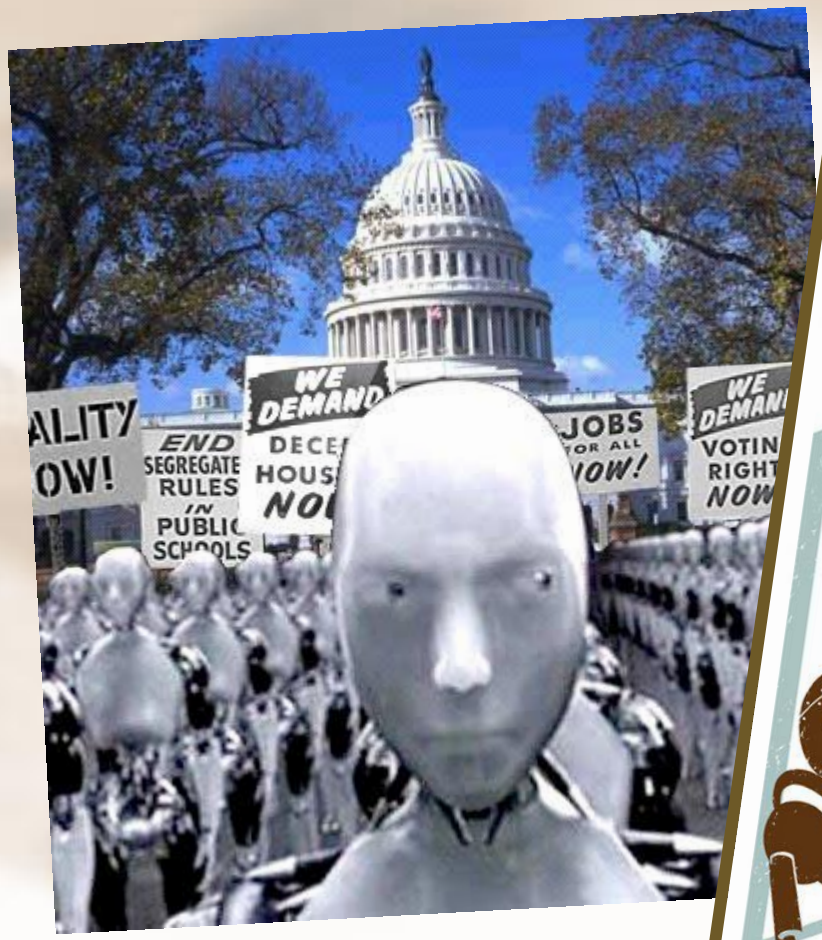

baxter

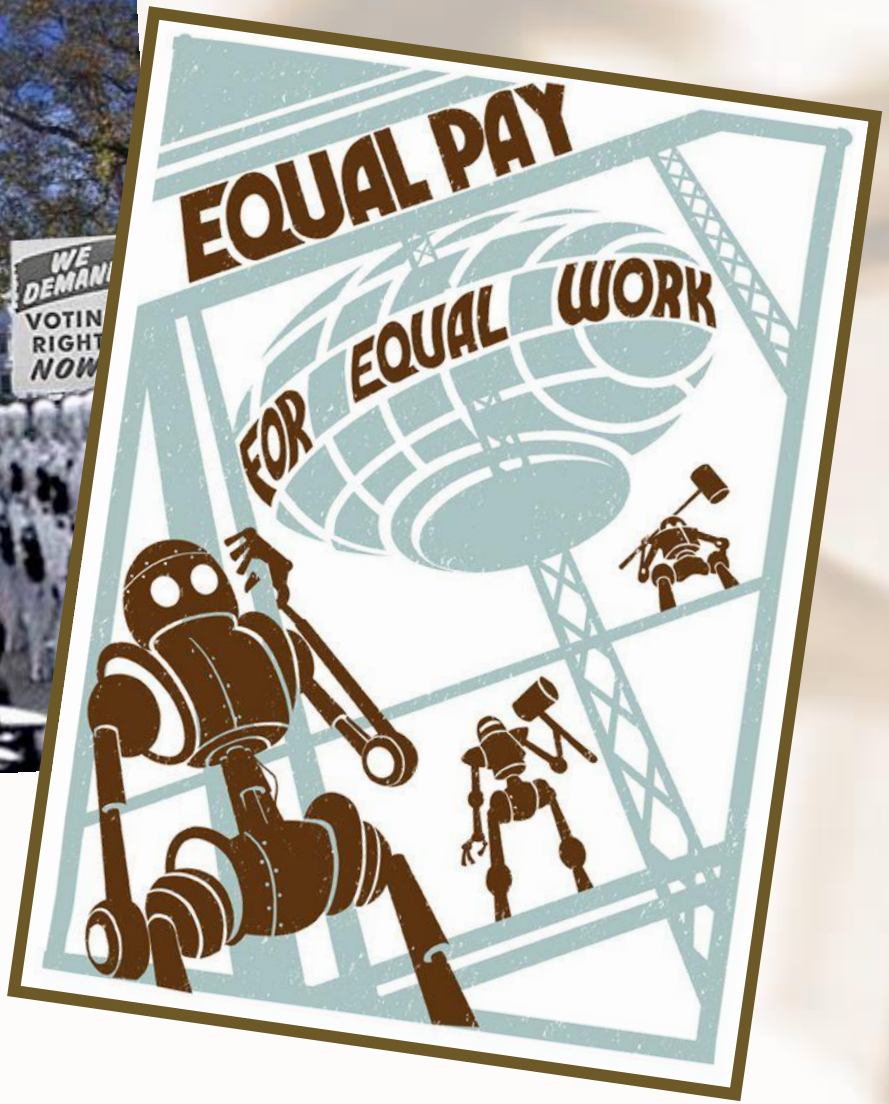

ASIMO



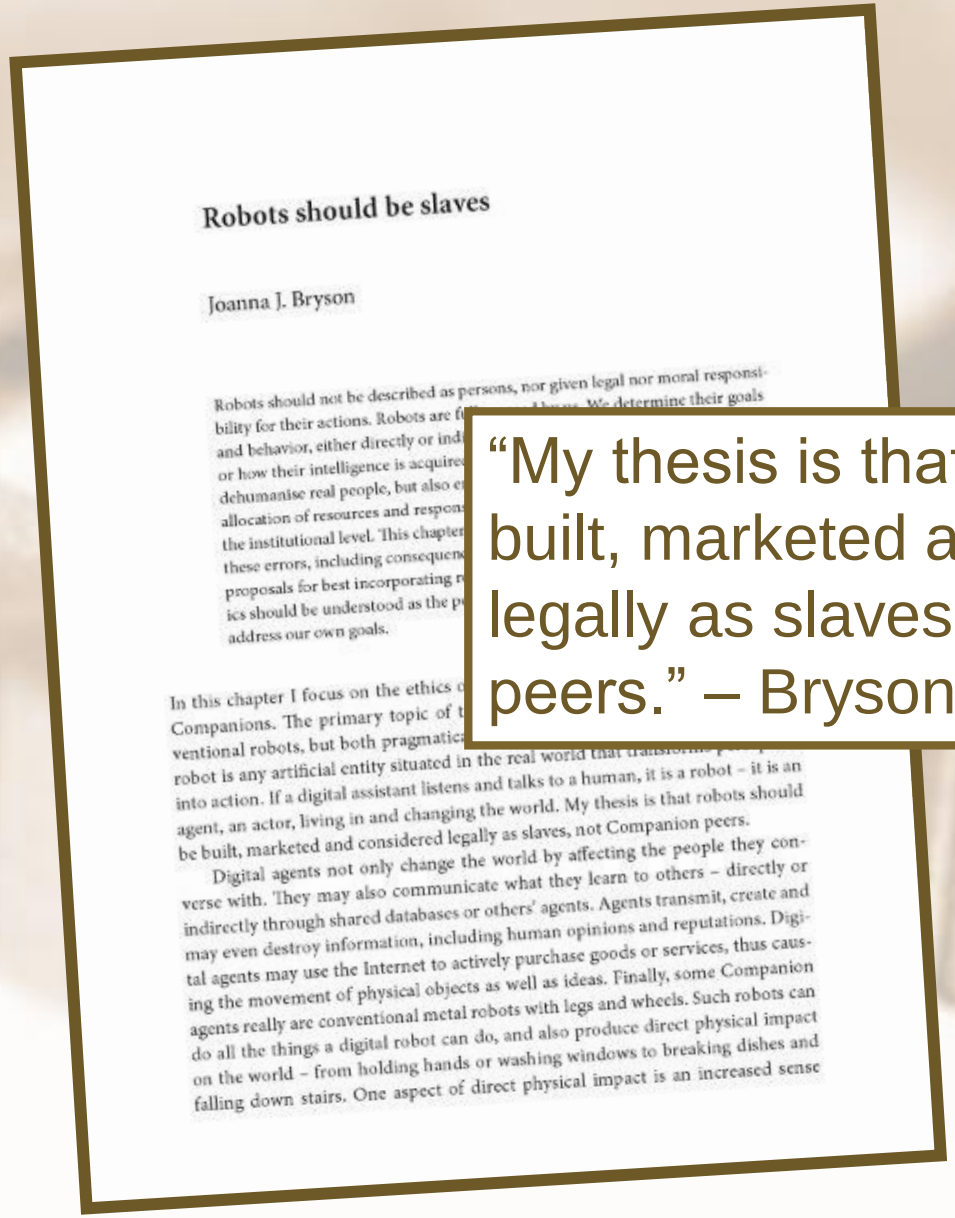
Machine Moral Agency

Machine Moral Patiency





1. Instrumental Theory



Robots should be slaves

Joanna J. Bryson

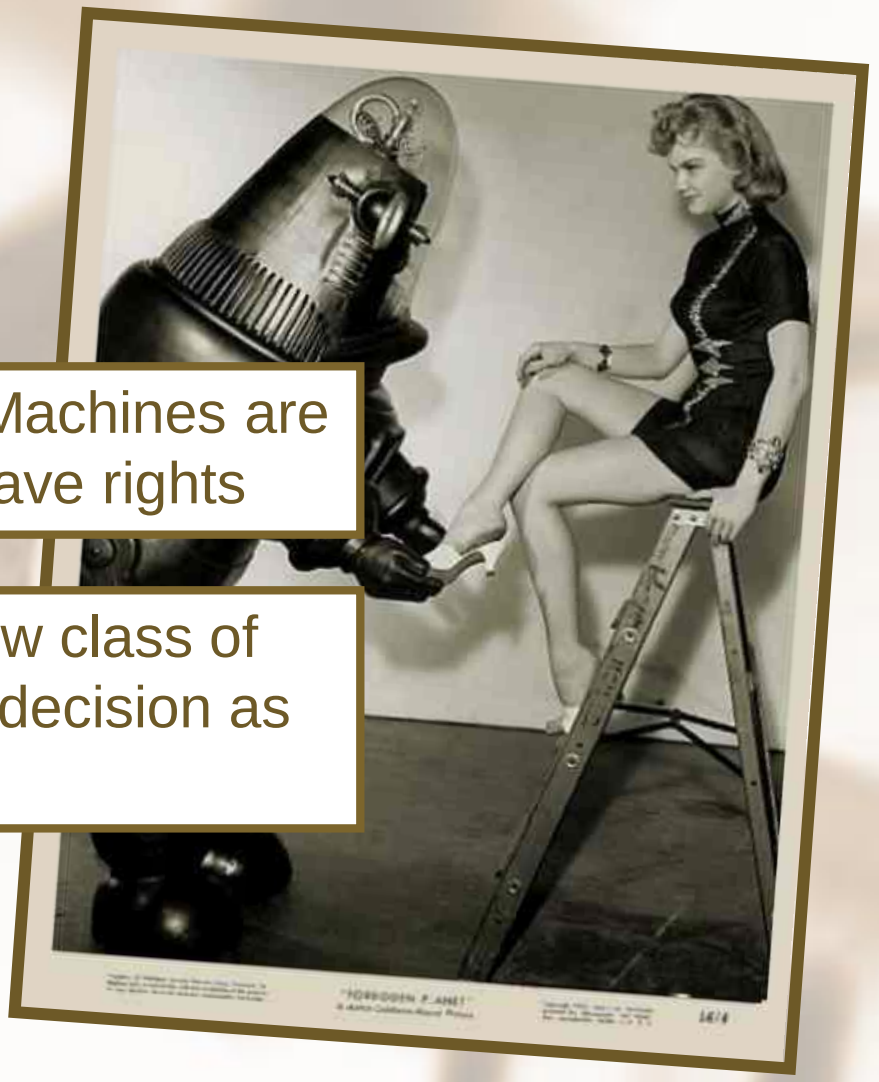
Robots should not be described as persons, nor given legal nor moral responsibility for their actions. Robots are f… determine their goals and behavior, either directly or ind… or how their intelligence is acquire… dehumanise real people, but also … allocation of resources and respon… the institutional level. This chapter… these errors, including consequen… proposals for best incorporating r… ics should be understood as the p… address our own goals.

In this chapter I focus on the ethics … Companions. The primary topic of t… ventional robots, but both pragmatic… robot is any artificial entity situated in the real world that … into action. If a digital assistant listens and talks to a human, it is a robot – it is an agent, an actor, living in and changing the world. My thesis is that robots should be built, marketed and considered legally as slaves, not Companion peers.

Digital agents not only change the world by affecting the people they converse with. They may also communicate what they learn to others – directly or indirectly through shared databases or others' agents. Agents transmit, create and may even destroy information, including human opinions and reputations. Digital agents may use the Internet to actively purchase goods or services, thus causing the movement of physical objects as well as ideas. Finally, some Companion agents really are conventional metal robots with legs and wheels. Such robots can do all the things a digital robot can do, and also produce direct physical impact on the world – from holding hands or washing windows to breaking dishes and falling down stairs. One aspect of direct physical impact is an increased sense

"My thesis is that robots should be built, marketed and considered legally as slaves, not companion peers." – Bryson 2010

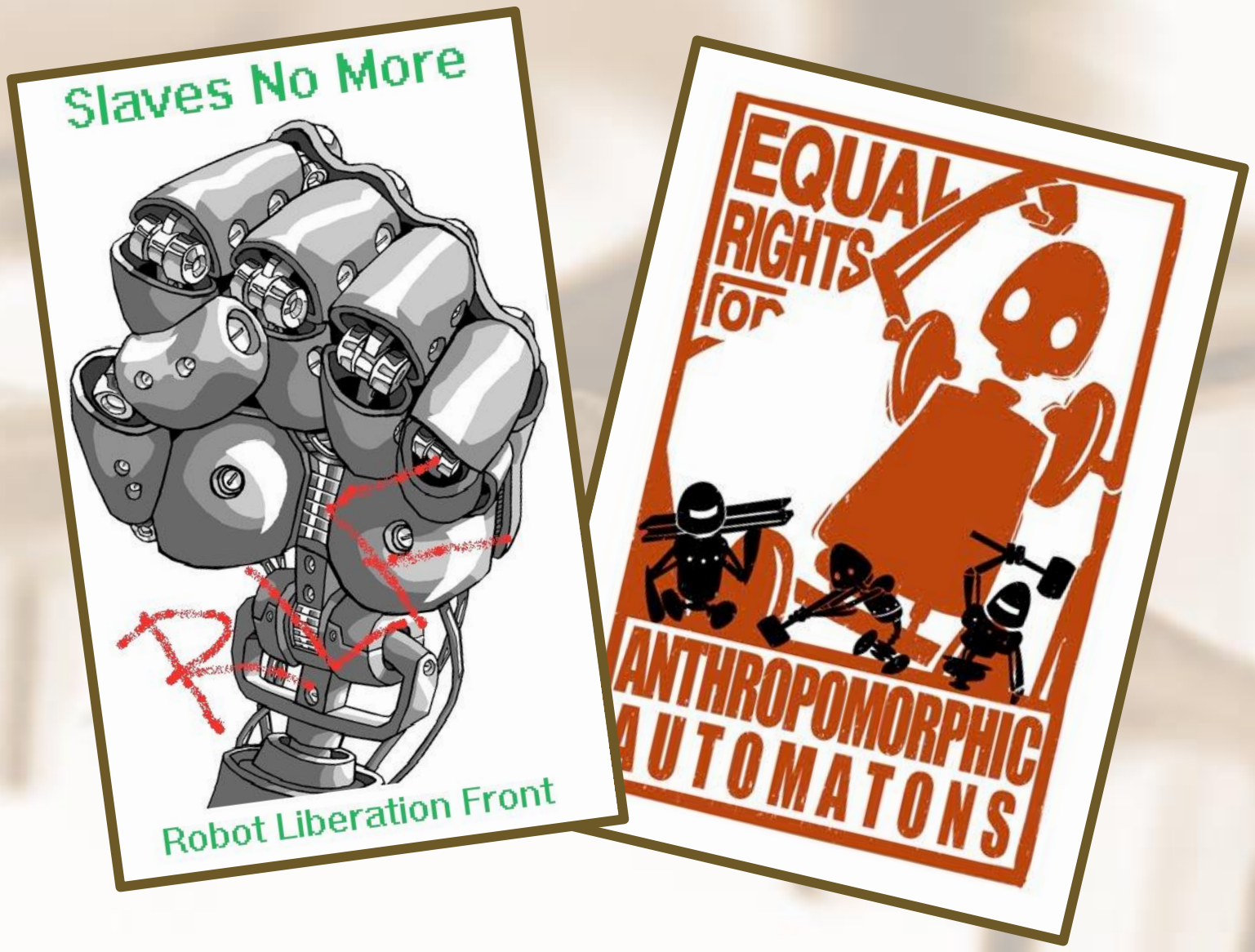
1. Instrumental Theory



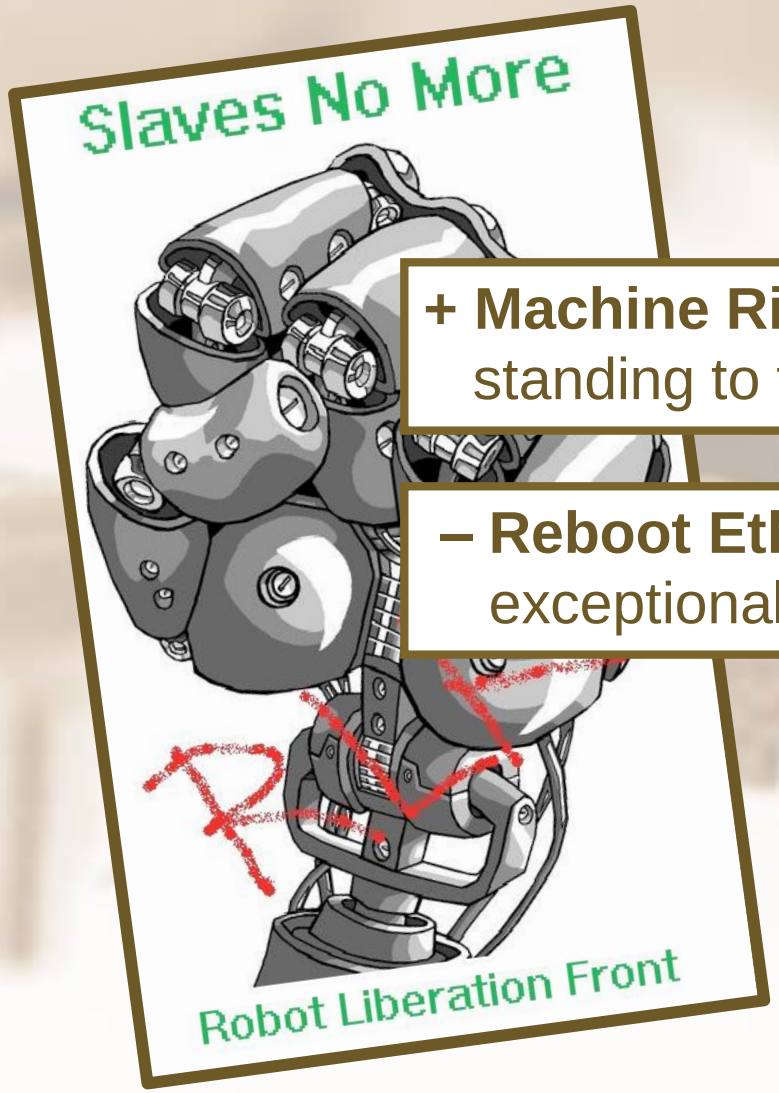
+ **Human Exceptionalism:** Machines are tools; only human beings have rights

– **Slavery 2.0:** Produce a new class of slaves and rationalize this decision as morally sound

1. Instrumental Theory



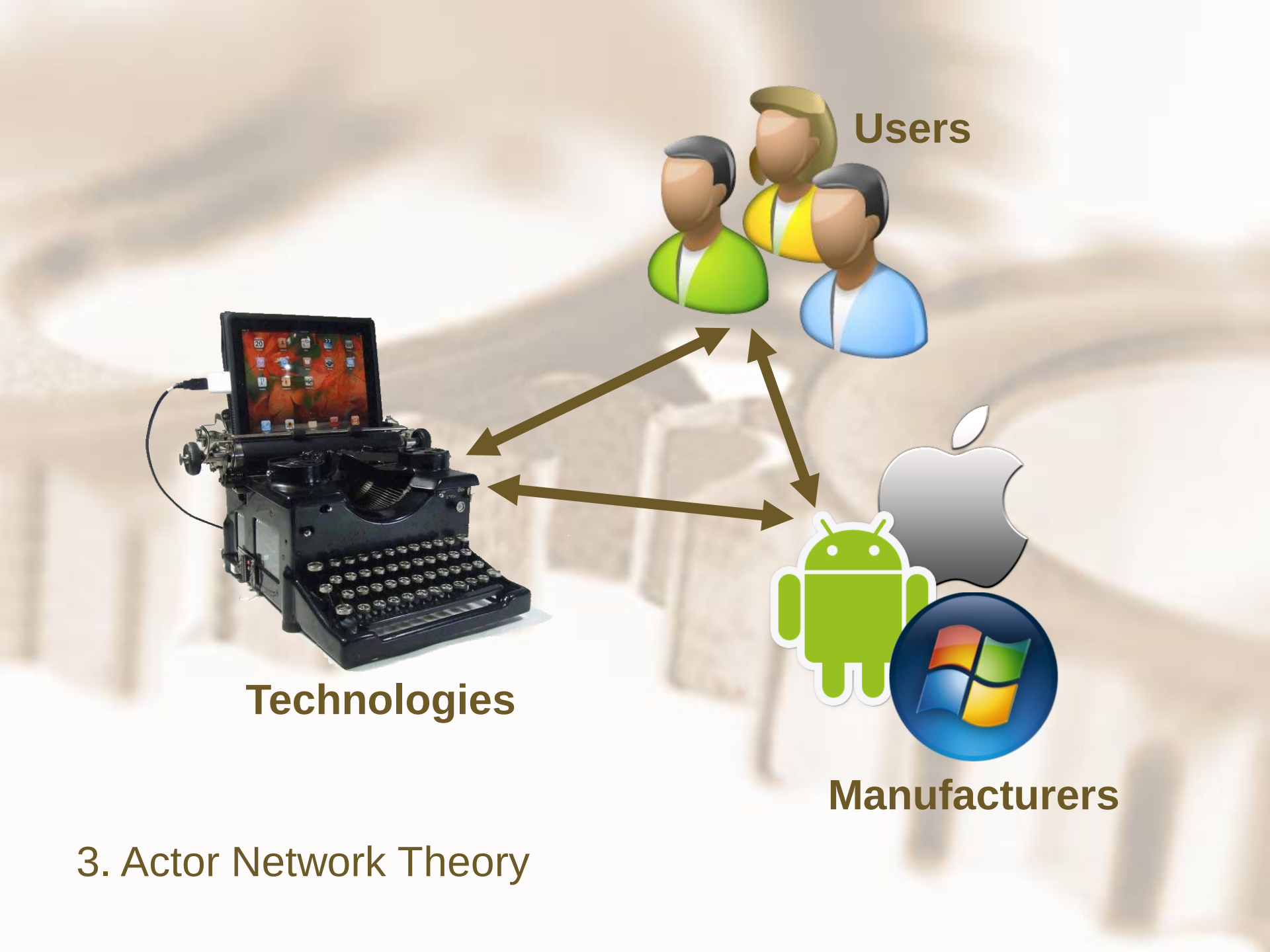
2. Machine Morality



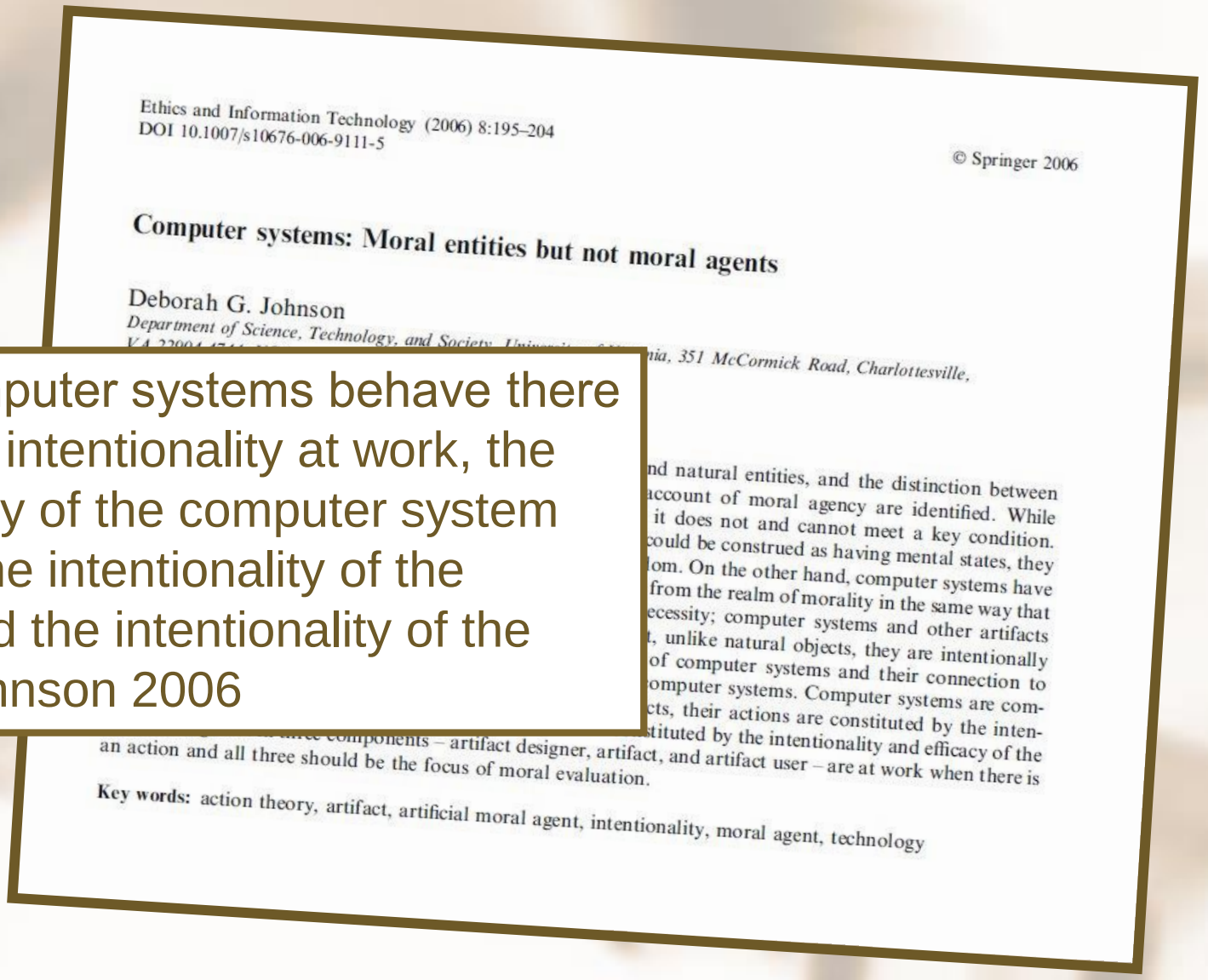
+ Machine Rights: Extend moral standing to these social aware entities

– Reboot Ethics: Think beyond human exceptionalism and instrumentalism

2. Machine Morality



3. Actor Network Theory



Ethics and Information Technology (2006) 8:195–204
DOI 10.1007/s10676-006-9111-5

© Springer 2006

Computer systems: Moral entities but not moral agents

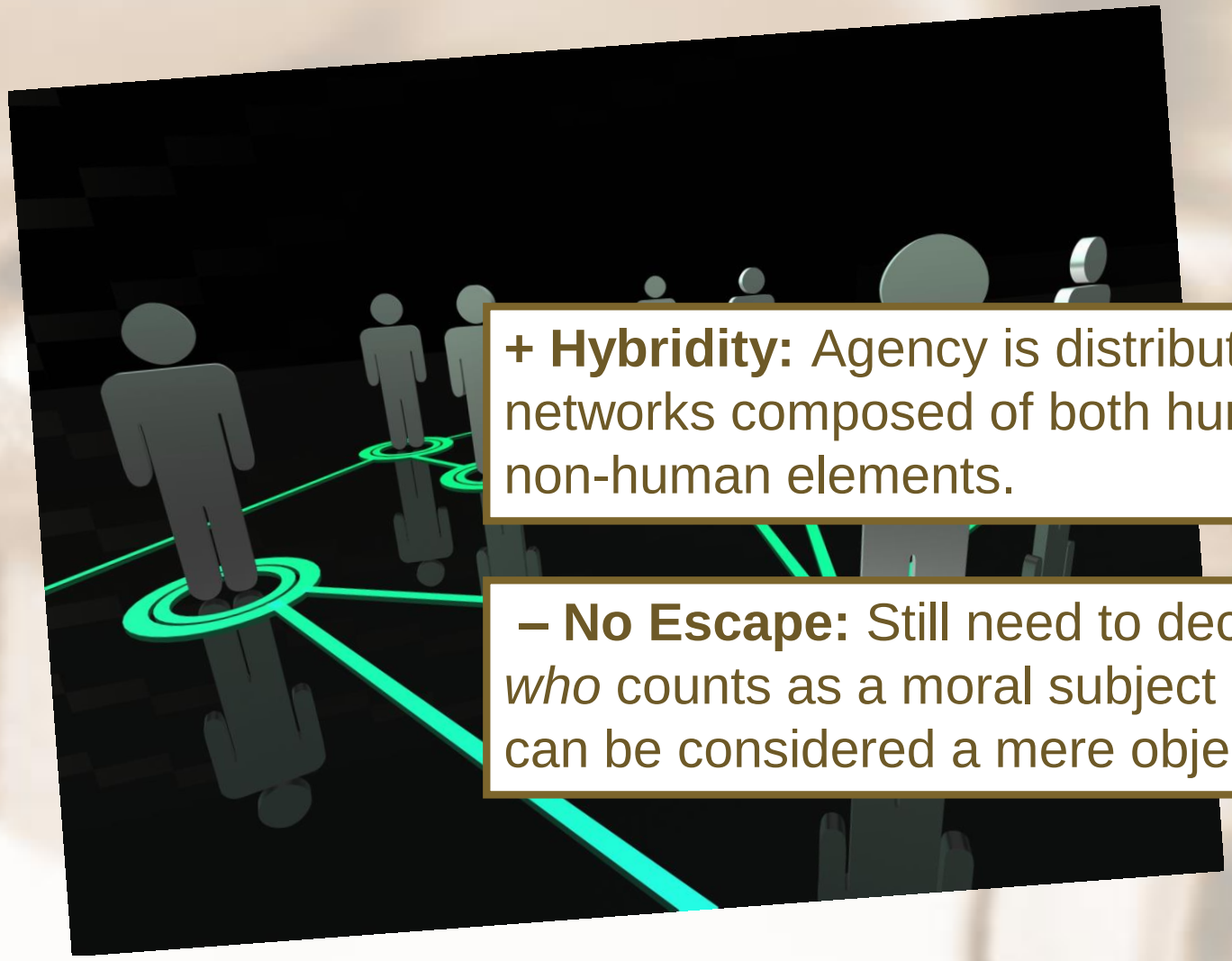
Deborah G. Johnson
Department of Science, Technology, and Society, University of Virginia, 351 McCormick Road, Charlottesville,

nd natural entities, and the distinction between
account of moral agency are identified. While
it does not and cannot meet a key condition.
could be construed as having mental states, they
lom. On the other hand, computer systems have
from the realm of morality in the same way that
ecessity; computer systems and other artifacts
t, unlike natural objects, they are intentionally
of computer systems and their connection to
omputer systems. Computer systems are com-
cts, their actions are constituted by the inten-
stituted by the intentionality and efficacy of the
components – artifact designer, artifact, and artifact user – are at work when there is
an action and all three should be the focus of moral evaluation.

Key words: action theory, artifact, artificial moral agent, intentionality, moral agent, technology

"When computer systems behave there is a triad of intentionality at work, the intentionality of the computer system designer, the intentionality of the system, and the intentionality of the user." – Johnson 2006

3. Actor Network Theory



+ Hybridity: Agency is distributed across networks composed of both human and non-human elements.

– No Escape: Still need to decide between *who* counts as a moral subject and *what* can be considered a mere object.

3. Actor Network Theory