

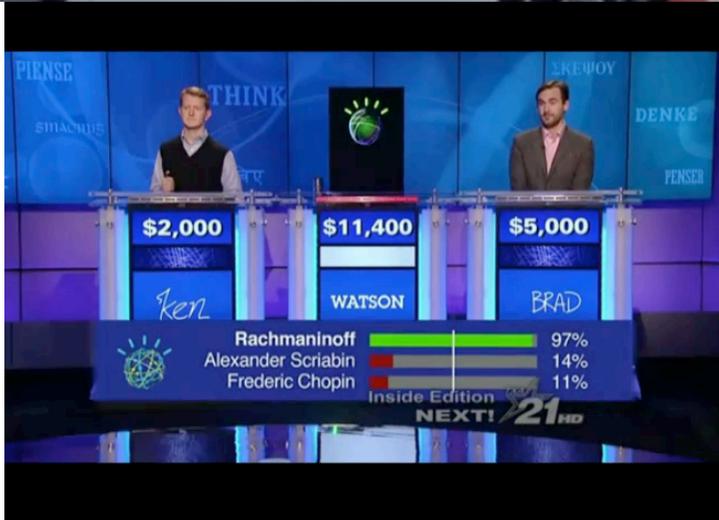
We're Already Succeeding & It's Already Scary / Wonderful

Joanna J. Bryson

Bath, Computer Science (IS)

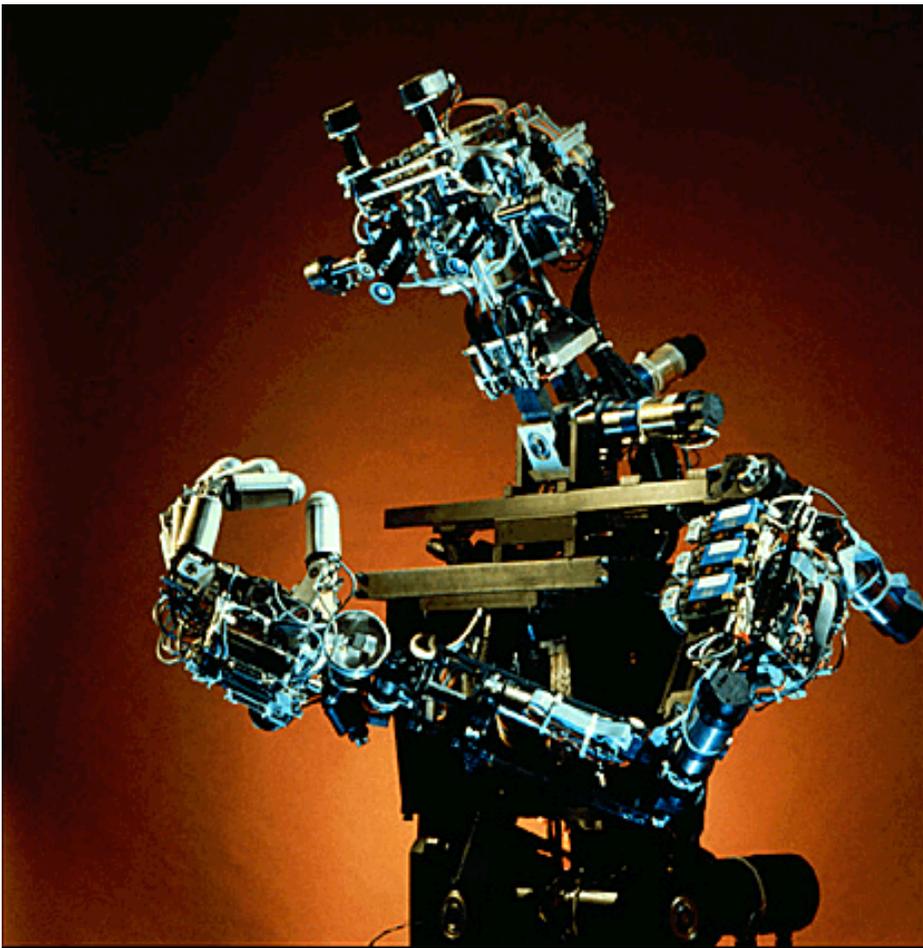
Mannheim, Social Science (MZES)

@j2bryson



This Is AI

Why isn't that obvious?



People want to **make** AI they owe obligations to, can fall in love with, etc. – “equals” over which we have complete dominion.

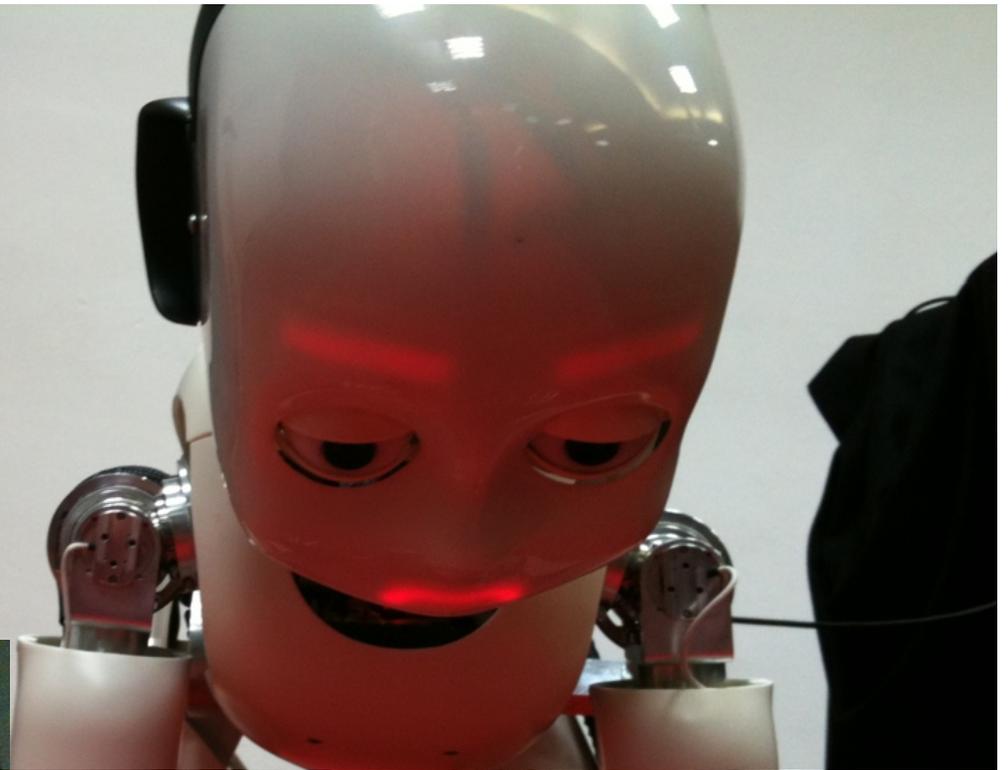
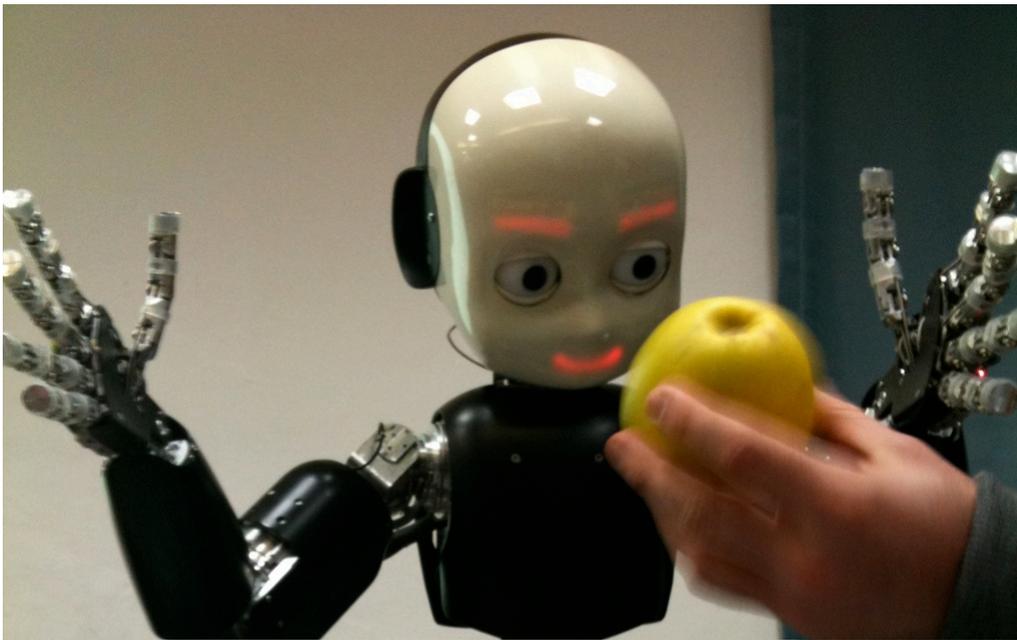
Joanna J. Bryson and Philip P. Kime, “Just an Artifact: Why Machines are Perceived as Moral Agents”, *The Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI)*, Barcelona, Spain, pp. 1641–1646, Morgan Kaufmann, 2011.
Joanna J. Bryson “Patience Is Not a Virtue: Suggestions for Co-Constructing an Ethical Framework Including Intelligent Artefacts”, *The Machine Question: AI, Ethics and Moral Responsibility*, D. Gunkel, S. Torrance and J. J. Bryson (eds.), AISB, Birmingham, 2012.

A Proposed Robot Ethics

**Robots are servants
we own.**

⇒ **Slaves**

Joanna J. Bryson, “Robots Should Be Slaves” in *Close Engagements with Artificial Companions: Key social, psychological, ethical and design issues*, Yorick Wilks (ed.), pp. 63–74, John Benjamins, Amsterdam, 2010.





Because we build robots and determine their goals. Our **authorship** is fundamentally different from our relationship to humans or other evolved systems.

We Get to Decide...

- Whether robot minds are unique or have autosave and offsite backup.
- Whether robots **suffer** (permanently degrade their behaviour) when neglected, insulted or otherwise **subordinate**.
- **We as authors are ethically obliged to ensure we have no ethical obligations to AI.**

UK EPSRC's Principles of Robotics (2011)

1. **Robots are multi-use tools.** Robots should not be designed solely or primarily to kill or harm humans, except in the interests of national security.
2. **Humans, not robots, are responsible agents.** Robots should be designed & operated as far as is practicable to comply with existing laws & fundamental rights & freedoms, **including privacy.**
3. **Robots are products.** They should be designed using processes which assure their safety and security.
(of 5...)

UK EPSRC's Principles of Robotics (2011)

4. **Robots are manufactured artefacts.** They should not be designed in a deceptive way to exploit vulnerable users; instead their machine nature should be **transparent**.
5. **The person with legal responsibility for a robot should be attributed.** [like automobile titles]

Joanna J. Bryson, Kerstin Dautenhahn and Geoff Pegman, "Man and the machine", letter published online, *The Economist*, 16 June 2012.

Joanna J. Bryson "The Making of the EPSRC Principles of Robotics", *The AISB Quarterly*, (133) Spring 2012.



Humans Are Just Chimps With History

- A brain 3x the size of a chimp's is not a big deal.



- We only became **really** different in the last 10,000 years, with writing, urbanisation, doctrinal religion (spreading knowledge to out-group) & agriculture.

AI Already Owns Our Advantage



Utopia: Solve hard problems like sustainability; reliably supporting everyone's efforts to self actualise.

Dystopia: Losing autonomy / ability to freely express; catastrophic disruption of the global ecosystem.