




HIM Lecture 7

Global, Local, and National Regulation: What Are Countries For?

Joanna J. Bryson

University of Bath, United Kingdom

@j2bryson

The background of the slide features two side-by-side views of Earth from space. The left view shows the Earth at night, with the continents of North and South America visible, their landmasses covered in a dense network of golden-yellow city lights. The right view shows the Earth during the day, with the blue oceans and white clouds of the Western Hemisphere visible. A vibrant green aurora borealis is seen swirling around the North Pole in this view.

Is AI Changing the World? Yes

thank you NASA

AI Is Changing Us

- Blurring distinction between customer and employee – citizens of corporations.
- Reducing (not eliminating) costs and advantages of geographic location, increasing inequality and transnational interdependence.
- Altering governance – makes stabilisation of policy through obscuring difficult or impossible.
- Allowing prediction of people from their data, and therefore their manipulation.

Science (a testable description)

MORE AI \Rightarrow MORE FASTER COALITIONING

Communication increases the chances of
finding new ways to cooperate.

Roughgarden, Oishi, & Akçay (2006)

Policy (normative agreement for achieving mutual goals)

Definition. Romanes in 1882 wrote *Animal Intelligence* (a monograph). Also the definition I learned as a social sciences major at Chicago in 1983, and as an MSc in AI at Edinburgh in 1993.

◆ **intelligence** *n.*

1. A person's faculty of understanding; perceiving and comprehending meaning; mental quickness; active intellect (Oxford English Dictionary 1972, entries from 1390; Michaelis 1963).
2. An animal's capacity to adjust its behavior in accordance with changing conditions (Romanes 1882 in McFarland 1985, 505).
3. An individual animal's associating stimuli (Thorndike 1911, 20–23).
4. A person's ability to adapt to new situations and to learn from experience (Michaelis 1963).
5. A person's inherent ability to seize the essential factors of a complex matter (Michaelis 1963).
6. An animal's learning ability (Wilson 1975, 473).
7. In more derived primates: an individual's ability to show reasoning or insight learning (Wilson 1975, 381).

cf. awareness and associated terms, learning

Policy

- How do you determine an appropriate course of action? **Normative Ethics**
- What do people actually do? **Descriptive Ethics**
- How can we achieve moral outcomes? **Applied Ethics**
- Can ethics even make sense? **Meta Ethics**

What you've actually decided to do.

Much like in Reinforcement Learning (there “policy” is the model you’ve learned for translating perception to action = intelligence).
For Science to alter behaviour, and better life, it has to inform policy.

Government

- Government does three things: (Landau 2016, cf Bryson 2019)
 1. Pick a problem to solve.
 2. Allocate resources to that problem = **redistribute**.
 3. Stabilise that solution so it can take effect, be tested.
- Corporations and countries both have government.
 - They also **both** work to police corruption in each other.
 - Good government by each makes this interaction easier.

Do We Still Need Nations?

Yes

- **Space matters.** Many problems are local:
 - House fires, vaccination, air pollution, food and water supplies, general education (including of your neighbours and their children), physical security, physical intimacy.
- United Nations agreement is that nations are responsible for the wellbeing (human rights) of all humans **within their borders**, citizens or not.

Geography ∈ Space

- **Moral agents** are considered responsible for their actions **by a society**.
- **Moral patients** are considered the responsibility **of a society's** agents.



Arguably, **ethics is determined by and determines a society** – a constantly renegotiated set of equilibria.

These equilibria depend on tradeoffs, such as security vs privacy.

The ideals for these tradeoffs will vary due to a nation's resources, neighbours, and its people's cultures.

The above is my
(educated) opinion.
Not everyone agrees with
me. (Google ATEAC)

Dealing with Reality – Facilitating and Leading Leaders

Joanna J. Bryson

University of Bath, United Kingdom

Invited talk, ICML 2018
(3 weeks after ATEAC)

@j2bryson

Why Engage?

- Because things are getting ropey and we all need to pitch in.
- **Fredrick Brooks: The Mythical Man Month** (1975, so old its title is sexist).
 - Programmers hate to be lead.
 - But hub & spokes is efficient.
 - Leading is just a team role like any other – about efficient coordination.
 - Leading: Don't waste more person seconds optimising solutions than the delta between solutions gives you back!
 - Following: Feed advice back to the leader when they seem to be ignorant, but also follow – coordinating is their job not yours.

Who Should Engage?

- Individual prominence is exactly like going viral:
 - Need to be reasonably good at what you do.
 - (Happen to?) Have acquired the skills / traits that the moment needs.
 - Some random / chance event gives you momentum others around your skill level don't happen to have (at least right now).
- If it happens to you: time to do your best (requires self-maintenance too).
- If it happens to someone else: help them be successful if you can (guard their time though.)

How Do You Split Up Your Time

- Foraging theory:
 - spend your effort on activities in proportion to their estimated likelihood to be efficacious.
- Explore / exploit:
 - If it looks like your efficacy model is out of date, you may want to invest time in updating it.
 - If you get offered an opportunity you've never had before (or recently) increases the probability you might get a valuable update.
- No one can do everything, but there may be somethings all professionals should do.

ICT Systems Are Designed, and Have Architecture

- Architects learn laws, policy, and how to work with governments & legislatures at university...
- because society decided collapsing buildings were unacceptable, and city alterations affect everyone.
- ICT systems are now falling on people and affecting everyone. Software needs to mature, as architecture did.
- Rate of **successful, sustainable** innovation is what matters, not just speed to market.

How do we regulate AI?

Transparency and Accountability

- In the **worst** case AI is as inscrutable as humans.
 - We audit **accounts**, not accountant's synapses.
- “But we can put can accountants on the witness stand and determine due diligence.”
 - **Really:** We **guess** diligence based on empathy.
- AI facilitates mandating **transparently-honest accounts**.
 - Fully document the **software engineering process**, **data** and **training**; **log the system's performance**.
 - These support regulation.

Key Local Regulatory Concepts

- **Do not** reward corporations by capping liabilities when they fully automate business processes – Legal lacuna of synthetic persons (Bryson, Diamantis & Grant 2017.)
- **Do not** motivate obfuscation of systems by reducing liabilities for poor revision control, badly-tested or poorly-monitored machine learning. **No special status** for systems with ill-defined properties, such as ‘conscious’, ‘sentient’, or ‘sufficiently complex’.
- **Clear code is safer** and can be more easily maintained, but **messy code is cheaper** to produce (in the short run.)
- **Regulation should motivate clarity** (transparency) by requiring proof of due diligence (accountability).

Law and Professional Societies

- **Governments** are good for enforcing law, redistribution.
- **Designed** to be slow – should provide **a stable environment** for developing businesses, families, etc.
- **Professional societies** (e.g. BCS, IEEE, IET, ACM) are composed of **experts**, should be able to keep up with contemporary issues.
- **Combination** – Professional societies maintain standards, **governments enforce these standards** (Bryson & Winfield, 2017).

AI Requires Security;

Security Is an Arms Race

- Google got hacked by the NSA (cf. Snowden). [The US Federal Government gets hacked](#). Political parties, [banks](#), cheap apps, [LinkedIn...](#)
- IoT devices generate less revenue than the cost of a security upgrade – lightbulbs & baby monitors **stay** compromised.
- There is no cybersecurity/autonomy “tradeoff” – you are encrypted or you aren’t. Backdoors get too many keys made.

What Counts as AI?

- The exact definition of intelligence (& therefore AI) shouldn't matter.
 - AI isn't so much the problem as digital – just regulate software.
- Economically, Boeing, Gazprom, and GPS present similar problems with respect to both security and inequality as Google, Facebook.
- Transnational problems require transnational solutions.
- Example: A 'Kantian solution' – Europe's proposed international ~~digital~~ **transnational megacompany** tax – every country could implement that law (EU-size market coordination helps with enforcement.)

Intelligence is computation—a transformation of information. Not math.

Computation is a physical process, taking time, energy, & space.

Finding the right thing to do at the right time requires search.

Cost of search = # of options^{# of acts} (serial computing).

Examples:

- Any 2 of 100 possible actions = $100^2 = 10,000$ possible plans.
- # of 35-move games of chess > # of atoms in the universe.

Concurrency can save real time, but not energy, and requires more space.

Quantum saves on space (sometimes) but not energy(?)

Omniscience (“AGI”) is not a real threat. No one algorithm can solve all of AI.

Viv Kendon, Durham





Google converts old paper mills, decommissioned coal plants into data centres.

Google uses only its own fiberoptic network (laid globally), chips designed and built in-house (unlike the EU), because of cybersecurity – even other fiberoptic cables in a bundle might spy on traffic.

AI is much more than algorithms or data.

Tech giants are significant transnational assets.

Only Humans Can Be Accountable

- Law and Justice are more about **dissuasion** than **recompense**.
- Safe, secure, accountable software systems are modular – **suffering*** in such is incoherent. ***e.g. systemic dysphoria of isolation, loss of status or wealth.**
- No penalty of law (or treaty) enacted directly against an artefact (including a shell company) can have efficacy.






[Artificial Intelligence and Law](#)

..... September 2017, Volume 25, [Issue 3](#), pp 273–291 | [Cite as](#)

Of, for, and by the people: the legal lacuna of synthetic persons

Authors

[Authors and affiliations](#)

Joanna J. Bryson , Mihailis E. Diamantis , Thomas D. Grant 

Bryson, Diamantis & Grant
(*AI & Law*, September 2017)

Fortunately, Humans Build Artefacts

- **Due diligence** in the systems engineering of software requires:
 - Architect a system (design components, process for development, use, and maintenance.)
 - **Secure the system.** Including **provenance** of all **software libraries**.
 - Document (**log**) with revision control **every change to the code base** (who made **the change**, **when**, and **why**). Need to be able to “roll back” changes that don’t work.
- For learning, need same for **data libraries**, and **model parameters**.
- For AI, need to **log inputs and decisions** of operational systems.
- **Already standard** e.g. in automotive industry, incl. for driverless.

Nations can lead, but...

UK Principles of Robotics (2011)

Asimov's Laws revised for Manufacturer Responsibility

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**. [devops]
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]

Owner /
Operator
Respon-
sibility

Boden et al 2011; cf. Bryson *AISB* 2000; Bryson; Prescott, *Connection Science*, 2017; Floridi 2018.

OECD Principles of AI 2019 (endorsed by 42 goats + G20)

1. AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.
2. AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, and they should include appropriate safeguards – for example, enabling human intervention where necessary – to ensure a fair and just society.
3. There should be transparency and responsible disclosure around AI systems to ensure that people understand when they are engaging with them [the AI systems] and can challenge outcomes.
4. AI systems must function in a robust, secure and safe way throughout their lifetimes, and potential risks should be continually assessed and managed.
5. Organisations and individuals developing, deploying or operating AI systems should be held accountable for their proper functioning in line with the above principles.



Numbers Will Always Matter

- Enforcement of ideals requires sufficient economic carrots and regulatory sticks.
- The GDPR demonstrates sovereign nations can coordinate to defend against and even correct regulatory failures of other regions.
- GDPR also acknowledges our persons include our data, which can be used to alter our behaviour, and must be defended, like airspace.



Questions for Governments

- **Governing:** select a problem, address it with resources, stabilise, & move on to next problem (Landau, 2016).

1. Was the law—or a treaty—deliberately violated?

- Special case: Are we under a persisting attack?

2. Were citizens or visitors **harmed** (includes: **security, human rights, economic flourishing**) due to negligence?

3. Has something fundamental changed in the world that requires new resources to help citizens flourish?

Some Answers for Government

1. Was the law (or a treaty) deliberately violated?
 - Manufacturers and operators should be able to prove it hasn't been, or be held accountable.
2. Were citizens or visitors harmed (including security, human rights, economic flourishing) due to negligence? (Same solution as above.)
3. Has something fundamental changed in the world that requires new resource to help citizens flourish?
 - Yes. Expertise is needed to help inspect software “accounts”. Transnational cooperation is needed for transnational redistribution, redundant versions of natural global monopolies.

If we legislate and adjudicate
for accountability,
transparency will follow.
Responsibility is already ours.