

## TOWARDS A HUMAN ARTIFICIAL INTELLIGENCE FOR HUMAN DEVELOPMENT

Emmanuel Letouzé<sup>1</sup>, Alex ‘Sandy’ Pentland<sup>2</sup>

<sup>1</sup>Data-Pop Alliance and MIT Media Lab, <sup>2</sup>MIT and Data-Pop Alliance

---

**Abstract** – *This paper discusses the possibility of applying the key principles and features of current Artificial Intelligence (AI) to design future human systems in ways that could make them more efficient, fair, responsive, and inclusive.*

**Keywords** – *Artificial Intelligence, Big Data, Development, Fourth Industrial Revolution*

### 1. MOTIVATION

The rise of “Big Data” over the past decade and the more recent emergence of Artificial Intelligence (AI) in the mainstream discourse and our daily lives have stirred many hopes and, increasingly, fears about the fate of humankind in the “Fourth Industrial Revolution”. Are we heading towards brighter or darker times? Do Big Data and AI pose existential threats to democracy?<sup>i</sup> Will we (or our children) all lose our jobs and rights when robots and algorithms powering them take over? Or may we instead grow old in a world where decisions will be more rational, policies more efficient, processes fairer, politicians more accountable? In short: what may be the impact of data, algorithms, and AI on our increasingly data-infused digital societies?

As with past techno-political revolutions, the trajectories experienced by different groups will primarily depend on decisions made by humans—the most powerful or whom, sadly, do not typically care very much about how others are affected. To avoid a dystopian future shaped by and for elites and machines, a growing number of organizations, governments, and citizens feel a sense of urgency to act, but are unsure about how.

A first obstacle is lack of a clear understanding—of what is really happening and looming with Big Data, AI, and so forth. Another is lack of a long-term vision—of how humans and machines may cooperate in the future, and what the corresponding processes and ‘building blocks’ to put in place ought to

be. Yet another is lack of a clear roadmap for mobilizing and coordinating scarce resources including human and technological, towards that end. A last one is the dominance of personal agendas favoring naïve embrace or systematic fearmongering of all things AI.

In this paper, we aim to sketch an ambitious and optimistic vision and offer some reflections about how current and future human societies could shape this future, in particular “leverage” AI by not just using it but by applying some of its key principles to build a ‘Human AI’ reflecting and serving the objectives and drivers of human development (Pentland, 2017).

### 2. THE GIST AND “GOOD MAGIC” OF CURRENT AI

A first step is to understand the gist of what is happening. By and large, the “Data Revolution”<sup>ii</sup>, Big Data<sup>iii</sup>, and current AIs run on personal data emitted by people using digital devices and services for their daily actions and interactions; yielding digital signatures or “data bread crumbs”<sup>iv</sup> in the forms of cell phone records, bank transactions, web and social media content, geolocation data, pictures, videos, etc. The resulting large datasets can then be analyzed by algorithms to unveil patterns and correlations and make estimations, projections, predictions, prescriptions, etc.

Most of us already choose or are incentivized to rely on these tools to decide which roads to drive, articles to read, clothes to buy, content to like, flights to book, people to connect with,

etc. Our doctors will soon use the same types of tools to diagnose cancer and suggest a treatment plan.<sup>v</sup> There are and will be actual “robots” running our factories, doing home chores, and entertaining our kids, but, by and large, current and future AIs are and will be what was called Big Data a few years ago: computational analytics models fed and trained on large quantities of data crunched by machines (computers) to reach an objective: predict this, optimize that, suggest this, recognize that, and, in some cases, powering sophisticated machines (‘robots’) to implement decisions in a more or less autonomous manner.

When a driverless car is on the road, the computer that steers it looks at its surroundings, asking itself whether what it is in front is pavement or people and act accordingly, in ways it was trained and taught itself to do through millions of previous simulations. Before getting good at its job, the machine often ‘got it’ wrong, and was told so: it (virtually) crashed into trees, crushed people, fled when seeing the police—all things considered bad. Through these trials, errors, and feedbacks, it started being able to drive autonomously. Another machine looked at the picture of a cat and, when prompted, concluded it was a dog. It was told “wrong!” and asked to try again with a different photograph, many, many, times over.

Through these iterations, these machines learned what features and combination of features of what they were seeing were most systematically associated with the right result. The algorithm, the series of steps classifying, organizing, ranking information and tasked with concluding “cat!” or “dog!” figured out that the longer the nose, the more likely the “thing” was to be “dog”, whereas considering at whether it had long or short hair was not a very valuable use of its time and neurons. It was learning how to “connect the dots” to get to the right results. The machine was learning. The gist of Big Data and current AI(s) is machine(s) learning.

Of course, there are many more caveats and complexities than this, but for most intents

and purposes it suffices to understand that current ‘narrow’ AI (as opposed to a ‘general’ AI that fuels the most vivid unsubstantiated fears about robots taking over the world, which doesn’t seem like a realistic outcome in the foreseeable future) is about this: get *lots* of data as inputs and learn how to connect them to output data (desirable or observed outcomes considered as the “right results”) through training testing, and learning based on past cases.

Applications and implications of this are already far-reaching. Is this person going to like this book because someone just like him or her (including him or her last month) did? Is this teenager on the verge of dropping out of school? Is this person Kieran McKay or Abigail Adeyemi? Should he or she get a loan, if so at what rate? Should the driverless car kill a pregnant woman or 5 elderly people if it has no choice but run over either? Several tough related questions come to mind and fuel ongoing debates: Are algorithms racist? Is it because their developers embed their biases or rather because predictions repeat past biases? What happens when the machine (here, the computer or the algorithm) encounters cases they have not seen before (e.g. a dog with a flat face? Someone with darker skin than in the dataset it was trained on)? Above all, how are and should those inferences, estimations, projections, predictions, suggestions, prescriptions, etc. be used, by whom, when, if at all?

These concerns and risks are real, and they need to be known and addressed to limit the worst typical side effects of technological change, at least in the short run, including widening inequities. But Big Data and AIs are neither ‘black magic’; nor are the algorithms running them completely ‘black boxes’. Given their ubiquity and power, it is important to understand how they do what they do, and what insights we could glean from them to promote positive social change. Critically, it is not (just) about *using AI* to optimize supply chains (and more), which will continue to have major impacts on societies and economies, but about *being inspired and helped by AI* to improve human systems.

What is the ‘good magic’ of current AIs? The good magic, is, in short and put simply, the “credit assignment (or reward) function”. It is the ability to assign credit for what “works”—i.e. what allows an algorithm to get the right (intended) result. In the example above, the computer tasked with telling a dog from a cat will extract millions of features from the image it sees, then assemble them in millions of ways, take guesses, and over time, learn which combinations of paths allow it to get the right answer (assuming everyone “calls a cat a cat”—as the French say<sup>vi</sup>) almost all the time. The reward function and learning through iterations lead to reinforcement of the combination of features to look for and use. In contrast, those that lead to the wrong result will be weakened. The machine will grow an incentive to not use them.

As it turns out, or so we think, applying the core principles and requirements of AI to entire human systems in a consistent, careful manner to design and deploy “human-machine (eco)systems” could be quite transformative—for the better.

### **3. APPLYING THE PRINCIPLES OF AI TO HUMAN SYSTEMS: TOWARDS A HUMAN AI**

We call such a system a Human Artificial Intelligence—a Human AI.<sup>vii</sup> What would this be and do? What would it not be and not do?

The basic principle is that as with current ‘simple’ (or narrow) AIs what “works” to “get it right”—think here policies, programs, behaviors, actions—would get rewarded and reinforced. Those that “don’t work” penalized and weakened. This too would be enabled by data-fed feedback loops. Over time, you would have human systems (societies, governments, organizations) with a pretty good sense of what “works”, i.e. what sets of policies, programs, behaviors, actions, yield good results. In addition to providing the core analogy, AIs would be part a central of this system, generating and crunching data and taking over tasks and helping decision-making under general human oversight.

A key to this is learning and agreeing through feedback what yields good vs. bad results, and acting accordingly the next time(s) around. Such processes already happen of course. Attempt to have a barbecue in a crowded subway car, and people will most probably tell you not to. Why? Because it will seem like a dangerous thing to do to most riders. How do they know? Through past experience or (more likely) through “common sense” based on past observations and inference. Sometimes we learn by insinuations or through intuitions: talk nonstop loudly at parties and, at some point, you will stop receiving invitations. Most of us will soon connect the dots. We also have instincts and reflexes nurtured through thousands of years of collective learning: we close or cover our eyes if a projectile gets near them, because that yields better results than keeping them wide open to take a closer look.

These are, in many ways, core features and outcomes of evolutionary processes. White rabbits tend to have higher survival rates in snowy plains than Brown ones. After a while, there are only White rabbits left running around in snowy plains. It is also the gist of culture: societies learn and teach what “works” for them, and turn them into codes and norms. Most societies have learned that not providing basic education to their children is not great both for the children and for the society at large; that widespread corruption is harmful to the majority; that only providing candies and beer in corporate cafeterias wouldn’t be a good idea. All of those things tend to yield bad results. We could learn that hitting a child for educational purposes does not “work”—that it is more likely to yield an unstable, unhappy, and violent adult.<sup>viii</sup> We could learn that there is no conclusive evidence that the death penalty works as a deterrent to major crimes.<sup>ix</sup> We could learn that human activity over the past two centuries has caused many animal species to become extinct, while temperatures rose, and oceans became more acidic.<sup>x</sup>

The vision we sketch here is one where data would fuel those human systems by applying the ‘good magic’ of current AI systems—the credit assignment function—by identifying,

rewarding and reinforcing what yields good results. The core principle is learning through feedback; the system's fuel is data. In short, let's figure out what "works" best, possibly for the majority, reward it, and strive to only or mostly do what contributes to these ends. Over time, what helps yield good results will take over what does not, become most prevalent; ideally turned into norms that need less enforcement. Human systems would be better off—say safer, fairer, more civil, more sustainable, because the opposite results do not "work" for most people.

Let's give some simple examples. If a judge (or entire justice system) systematically pronounces harsher sentences for similar offenses against people of color, it should be fired (or reformed). If the way kids are taught impedes their learning abilities and lifelong prospects, they should be taught differently. If a government does a lousy job, steals money, or kills its citizens, it should be changed. This may simply feel like common sense or liberal democracy at work, but bad policies, bad actions, and bad results are pervasive in the most "advanced" liberal democracies—for many reasons.

Critically, this is not a call for better data in the hands of benevolent "Bismarckian" policymakers who would (finally!) be able to make the good decisions they would otherwise always have taken all along. We do not believe that some of the greatest threats and challenges of our time, and those to come, are primarily due to poor information available to the ruling classes. One reason is that a fair share of politicians and people in position of power are either uninterested in the goods of their fellow citizens or incompetent, or both. Instead, a Human AI would be a system where it would be difficult for an elected representative to claim credit or assign blame out of hot air, because citizens would say "*Really? show me the data!*". Of course, it would allow well-meaning politicians to do a much better job; but it would be a system where citizens could fight bad politicians armed with better data.

It is neither an "Orwellian" vision where citizens would be digitally monitored and rated in real time all the time. This point would and will merit to be given more attention than is possible in a few pages, but a few points can be noted. First, all societies have lots of systems in place to influence individual behaviors in ways they deem desirable—from taxes to credit scores. Second, the focus on Human AI is on collective action(s) rather than individual actions. It is about instilling a culture and setting up the necessary systems and standards to improve collective decisions.

For this to work, there first needs to be a general agreement that decisions and outcomes ought to be evaluated on the basis of data—here for now used as synonymous with facts. It may not be easy to agree on the features and factors of "good results", but at least we should agree to assess them on the basis of facts. We need not having a preconceived agreement on what level of income inequality is desirable, but we should start by agreeing that and how inequality should be measured. From there we can understand what contributes to different levels of inequality, and to what results they contribute in turn. Human AI requires a general agreement that facts should *matter*—in a "Northien" perspective.<sup>xi</sup>, because otherwise systems cannot learn; and if they cannot learn, they cannot improve.

In summary, for the time being, let us describe if not define a Human AI as human social system that would apply and leverage the power of data and the "good magic" of AI—the ability to assign credit and learn from feedback with data as key inputs and outputs—to reward and reinforce decisions and actions that contribute to good results, through and feeding fact-based discussions between its members.

#### 4. CHALLENGES AND IMPEDIMENTS TO DESIGNING A HUMAN AI

This sketch of the vision of a Human AI has left out many challenges and questions, the biggest of which we can only briefly discuss.

First, some of the examples mentioned above are voluntarily contentious—because there is no consensus about or on them. For example, many people around the globe still think that spanking a child is good for him or her; many people support the death penalty and torture; many people are skeptical or in denial about climate change. Some people still insist the earth is flat. More people believe tax cuts for billionaires create jobs. This raises the general and fundamental question of what, whether and how we do, could, or should learn, individually and collectively, and how to come to a consensus and reach a compromise on major societal issues.

Some people won't figure out why they are no longer invited to parties, or won't ever be able to adjust and stop talking all the time. We are not fully rational beings either, and lots of other considerations get in the way. We (should) all know that getting that third piece of apple pie is bad for us, and yet many people do—because humans tend to value the present more, especially after a glass of good wine. Others chose to disregard or selectively pick data (here still synonymous with facts): they may believe that climate change is a hoax because they are told so by sources they grew up hearing; this is especially convenient if their income depends on fossil fuel. Others who have experienced or imposed violent behaviors may prefer their actions being comforted than confronted. Members of the world's "intellectual elite" pundit about climate change in conferences that require them to fly hundreds of thousands of miles—convincing themselves that on balance the world is better off that way.

There is also ample evidence that facts alone don't change people's minds<sup>xii</sup> and that recognizing facts may be getting harder in our data-rich era. Tensions between data, information and facts are not new to the data era; but there is a sense that in a world awash in data it becomes even harder for facts to be recognized, agreed upon, etc.<sup>xiii</sup> While so far we treated data and facts as synonyms for simplicity, there are obviously differences between the data we swim (or drown) in, and facts and truths. Ours is also the world of fake

news, alternative facts, where advances in digital imagery may soon mean we should not believe what we see.<sup>xiv</sup> Ill-intentioned powerful individuals and institutions have a rationale incentive for this to be perpetuated, as it works to their benefits. In many cases if citizens-customers were fully informed and aware of their motives and actions, they would not give them their money or their vote.

So the vision here is not one where people are bombarded by facts, surrounded by data dashboards, and told they are stupid or evil if they do not believe in them. Why not? Because this does not work. For such a system to work, there ought to be something more, a "connective tissue" that allows learning to happen, information to flow, facts to matter. Key ingredients for this seem to include greater trust, empathy, or "rational compassion", as discussed below, shared experiences and mingling, among and between individuals and groups. Somewhat, tautologically, Human AI would require and foster the kinds of societal characteristics and civic processes (especially with respect to social interactions) that would work best for itself. It is not entirely clear what those are (community discussions? elections?), but it seems reasonable to think that the answer are not Facebook battles and Twitter storms.

Another basic challenge is knowing what actually "works" and how / when / where it "works". The best economists in the world including Nobel Prize winners (or to be factually accurate, recipients of "The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel") still disagree about what fiscal, budgetary and monetary policies foster inclusive economic growth. There are broad areas of agreement, but no consensus on the right sets of policies. This applies to almost all domains of social life—because assigning causality or credit is difficult in complex systems where when so many variables (and values) interplay both as inputs and outputs. Most politicians claim that their actions should be given full credit for higher GDP growth or lower unemployment and conversely blame the business cycle for bad outcomes; when in reality "assigning credit" or

blame is hard in all cases, especially when few people try to do so with few data points.

Another major core challenge is agreeing on what the “good” end result ought to be. In most AI systems, the end result is a given (as in the “cat vs. dog” example); not in Human AI systems. Should societies aim for perfect income equality? Should economic policy aim to raise GDP, with all its limitations? Some will say, on balance, yes. Others will say, on balance, no. Should prolonging life be the end goal of any treatment? Soon, ‘values’ come into play. Opponents of hitting children or of the death penalty, or torture, will also argue along ‘moral’ lines, irrespective of outcomes and efficiency. And it would seem opinions often trump facts; that culture cards make the best of hands.

Yet there is probably still room and a need for rational outcome-based arguments in many of these debates. No study has ever concluded unambiguously that the death penalty deterred crime; some have concluded it may, at times. What is known for a hard fact is that some percentage of innocents will necessarily be killed in the process. Torture has been showed to “work” in few cases, and to lead to bad information in many, while a society that uses torture will probably not “work” for the majority over time. Female genital mutilation will have no place in a Human AI, because it leads to terrible results. Perspectives on social justice put forth by Rawls, Sen, and Nussbaum come into play<sup>xvixvi</sup>—in ways that would take much more space and time to give justice to. But fundamentally, as suggested above, a Human AI system is also one where what good results are may be discussed and determined on the basis of facts, to allow for gradual adjustments and improvements.

Another key challenge is access to data—to the kinds of data that would be necessary for a Human AI to start functioning—the data that power current narrow AIs. These sensitive data hold most keys to figuring out and advocating convincingly for what works, showing what doesn’t, and pointing towards means of improvement. For instance, assessing whether a new transportation

system may result or has resulted in increased economic opportunities and lower criminality would be significantly improved by having access to fine grained mobility data from cell-phones. The vast majority of “AI data” are collected and stored by private companies that legally act as “data controllers”. There have been many examples of and discussions about data sharing projects and agreements, but to date there are no systematic standards and norms for accessing those “AI data” ethically, and safely at scale in ways that could power the kind of Human AI described here.

Last indeed, there is the privacy imperative—as a fundamental human right. The vision of a Human AI is not an Orwellian one. It is not about looking into individual records; it is not about targeting specific individuals or groups. First, because this would not work. Societal reactions and legal trends observed in recent months suggest that while people’s attitudes towards privacy may be changing, we are not seeing the destruction of privacy as a marker and driver of human development. Second, there is no need to encroach on privacy for such a system to work: aggregated anonymized (strictly speaking, “pseudonymized”) indicators suffice.

Let us be very clear here: the Human AI is an aspirational analogy. It is a call for building human systems where facts matter; where the efficiency and relevance of policies and programs, and a multitude of socio-political processes and outcomes, can be assessed, discussed, and improved on the basis of data.

## 5. REQUIREMENTS AND PRIORITIES FOR DESIGNING A HUMAN AI

What is required for Human AI? It will take several key ingredients. It will take nurturing a strong, healthy data culture, including widespread data literacy, with more trust and interest in evidence-informed debates among the public. It will also take building better public governance for the systems that provide the data that can power a Human AI—including private sector data systems—allowing key data to be tapped into safely, ethically.

These are among others the key objectives of the Open Algorithms (OPAL) project. OPAL aims to allow accredited users to query private sector data through open algorithms running on the servers of partner private companies, behind their firewalls, to extract key aggregated indicators of interest, from cell-phone activity, bank transactions, possibly hospital records, police data, and more. With OPAL, no sensitive data ever leave the servers of the data partner organizations. All queries are logged, auditable; all algorithms are open, subject to scrutiny and redress.

OPAL also aims to develop governance standards and processes that will allow data subjects to weigh in on the kinds of analyses done using data about themselves; including through local oversight bodies referred to as Councils for the Orientation of Development and Ethics, or CODEs. Sensitive use cases are presented to the local CODEs, which may determine that some specific indicator, for example population density estimates, should not be provided beyond a certain level of temporal and geographic granularity for security reasons.

Currently piloted in Colombia and Senegal with two leading telecom operators and their National Statistical Offices, OPAL is the first ever real-world attempt at setting up technological systems and governance standards for building a Human AI. If successful, it will be expanded to other countries and industries. OPAL and other cases point to the fundamental discursive function of data, and to the importance of processes, for instilling positive social change. Setting up a project such as OPAL requires aligning incentives of large organizations around a common objective; that process alone has many virtues.

Another example of the value of processes that can be facilitated by data and algorithms is that of a controversy around changes to school bus routes in Boston, as recounted by Joi Ito, Director of the MIT Media Lab.<sup>xvii</sup> Protests over changes to school bus routes in Boston 'decided' by an algorithm designed by MIT researchers led to the (human) decision not to

use the algorithm. What appeared in hindsight was that the protesters were predominately wealthier families who has 'lost out' as a result of the changes suggested by the new algorithms. In the words of Joi Ito: "While I'm not sure privileged families would give up their good start times to help poor families voluntarily, I think that if people had understood what the algorithm was optimizing for—sleep health of high school kids, getting elementary school kids home before dark, supporting kids with special needs, lowering costs, and increasing equity overall—they would agree that the new schedule was, on the whole, better than the previous one."

Here, what lacked was transparency, openness, basic human communication. This looks like a missed opportunity to leverage the power of AI for the common good as a result of human flaws. Discussing complex social issues through the lens of data and algorithms, by explaining and discussing the objectives, features, potential pitfalls, of various algorithms, could change people's perceptions and attitudes towards social problems and fellow citizens.

Whether or not *empathy* is what ultimately needs strengthening as the core human ingredient for building better human systems is a matter of debate. It would seem like more empathy within and between groups could curb behaviors that hurt others, and lead to better overall social outcomes. This sounds consistent with and conducive to a Human AI that would, as described above, "reward and reinforce decisions and actions of its members that contribute to the common good, through and feeding candid fact-based discussions between its members."

But empathy tends to appeal to people's emotions, which can be manipulated and exploited. Paul Bloom, in his book "Against empathy: the case for rational compassion" argued that "[i]t is because of empathy that citizens of a country can be transfixed by a girl stuck in a well and largely indifferent to climate change", adding, provocatively: "We should aspire to a world in which a politician appealing to someone's empathy would be seen in the

same way as one appealing to people's racist bias".<sup>xviii</sup> And so perhaps a key ingredient of a Human AI is "rational compassion", which may be defined as the ability to consider different perspectives on the basis of facts—which we feel doesn't rule out reasonable interpersonal empathy.

A Human AI also requires developing incentives and means for civil society organizations, researchers, regulators, etc., to demand that public policies and programs be evaluated systematically using the best available data and methodologies, to adjust future iterations and contribute to a body of evidence on what yields which results. Data for transparency and rational compassion are a recipe for dealing with fake news and demagoguery.

This Human AI approach to improving society is not a techno-utopia; it is an "aspirational analogy" that places good data sources and rational discussion frameworks at the core of

a new social contract between humans as well as between humans and machines in 21<sup>st</sup> century societies. It is a vision of a AI that works for and with humans that is very different from one where machines are taking over jobs or the world, for better or worse. It is a vision where humans and machines (computers and the robots they will increasingly power) work together, each leveraging its comparative advantages.

It is also not a vision that should be assessed in the abstract or absolute; it is one that aims to improve the state of a world with many ills, a lot of which reflect and fuel bad information, bad faith, bad decisions, bad behaviors, and abuses of power that are seldom caught or addressed. The vision of a Human AI is letting the good magic of AI and the power of data challenge and improve old decision-making systems and power structures to improve human systems and the human experience, with humans in the drivers' seat.

---

<sup>i</sup> Pentland, Alex (2017) [https://www.edge.org/conversation/alex\\_sandy\\_pentland-the-human-strategy](https://www.edge.org/conversation/alex_sandy_pentland-the-human-strategy)

<sup>ii</sup> <https://weaponsofmathdestructionbook.com/> and <https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/>, notably <https://www.economist.com/prospero/2013/05/14/the-data-revolution>

<sup>iii</sup> <http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGGlobalPulseMay2012.pdf>

<sup>iv</sup> [https://www.edge.org/conversation/alex\\_sandy\\_pentland-reinventing-society-in-the-wake-of-big-data](https://www.edge.org/conversation/alex_sandy_pentland-reinventing-society-in-the-wake-of-big-data)

<sup>v</sup> <http://news.mit.edu/2018/artificial-intelligence-model-learns-patient-data-cancer-treatment-less-toxic-0810>

<sup>vi</sup> "Appeler un chat un chat", which means "Calling a spider a spider".

<sup>vii</sup> <http://thehumanstrategy.mit.edu/>

<sup>viii</sup> <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3768154/>

<sup>ix</sup> [https://www.washingtonpost.com/news/wonk/wp/2014/04/30/theres-still-no-evidence-that-executions-deter-criminals/?utm\\_term=.0f012c4991f7](https://www.washingtonpost.com/news/wonk/wp/2014/04/30/theres-still-no-evidence-that-executions-deter-criminals/?utm_term=.0f012c4991f7)

<sup>x</sup> <https://www.sciencedirect.com/science/article/pii/S0012821X1830356X>

<sup>xi</sup> Douglass C. North, *Structure and Change in Economic History*, New York, W. W. Norton, 1980

<sup>xii</sup> <https://www.newyorker.com/magazine/2017/02/27/why-facts-dont-change-our-minds>

<sup>xiii</sup> <https://www.media.mit.edu/videos/ml-talks-2018-04-23/>

<sup>xiv</sup> <https://www.newyorker.com/magazine/2018/11/12/in-the-age-of-ai-is-seeing-still-believing>

<sup>xv</sup> Rawls, John, 1921-2002. *A Theory of Justice*. Cambridge, Mass.: Belknap Press of Harvard University Press, 1971.

<sup>xvi</sup> Sen, Amartya. 2009. *The idea of justice*. Cambridge, Mass: Belknap Press of Harvard University Press. Nussbaum, Martha Craven. 1999. *Sex & social justice*.

<sup>xvii</sup> <https://www.wired.com/story/joi-ito-ai-and-bus-routes/>

<sup>xviii</sup> <https://www.economist.com/books-and-arts/2017/01/26/the-case-for-compassion-not-empathy>

+ Social Physics Sandy