COMMENTARY   **Open Access**

Can an AI learn political theory?

Stephen J. DeCanio[1,2]

Abstract

Alan Turing's 1950 paper, "Computing Machinery and Intelligence," contains much more than its proposal of the "Turing Test." Turing imagined the development of what we today call AI by a process akin to the education of a child. Thus, while Turing anticipated "machine learning," his prescience brings to the foreground the yet unsolved problem of how humans might teach or shape AIs to behave in ways that align with moral standards. Part of the teaching process is likely to entail AIs' absorbing lessons from human writings. Natural language processing tools are one of the ways computer systems extract knowledge from texts. An example is given of how one such technique, Latent Dirichlet Allocation, can draw out the most prominent themes from works of classical political theory.

Keywords: Artificial intelligence, Machine learning, Latent Dirichlet allocation, Alan Turing, Human/AI interactions

Introduction – AI/human interactions

Speculation about future interactions between artificially intelligent computer systems (AIs hereafter) and humans elicits both optimism and concern. AIs are moving beyond being mere tools that increase human capacities; in some areas like classic board games – Checkers, Chess, Shogi, and Go – they already surpass humans [44].[1] An AI system won one debate against an expert human debater (but also lost one), with a human audience as judges [41]. Furthermore, AI systems are beginning to exhibit genuine *agency*---they are able to act on their own without human intervention. Self-driving vehicles are perhaps the most prominent example of this, but autonomous weapons systems are under development by all the major military powers [2, 13] and algorithms account for considerably more than half the trading volume in the U.S. stock market.[2]

How will future AIs relate to humans? It has been known, ever since Turing's (1950) paper "Computing Machinery and Intelligence" [48], that even purely deterministic systems can behave in intrinsically unpredictable ways. Turing's experience with the simple digital computer available to him at the time led to this insight:

> [S]uppose we could be sure of finding such laws [of the behavior of computers] if they existed. Then given a discrete-state machine it should certainly be possible to discover by observation sufficient about it to predict its future behaviour, and this within a reasonable time, say a thousand years. But this does not seem to be the case. I have set up on the Manchester computer a small programme using only 1000 units of storage, whereby the machine supplied with one sixteen figure number replies with another within two seconds. I would defy anyone to

---

[1]The machine learning of Alphabet/Google's AlphaZero takes place according to programmed rules, but produces a neural network whose internal functioning is not at all transparent. Learning systems like AlphaZero are different from the more conventional brute force, search-intensive models that solved Checkers in the 2000s and defeated the human world champion in Chess during the 1990s (for Checkers, see [39, 40]; for Chess, see [34]).

Correspondence: decanio@econ.ucsb.edu
[1]Economics, Emeritus, University of California, Santa Barbara, USA
[2]Tulsa, Oklahoma, USA

[2]Estimates range from 65 to 70% [18] to as high as 80% [1] or even 90% [11].

learn from these replies sufficient about the programme to be able to predict any replies to untried values ([48], pp. 452–453).

This realization that the results of some machine computations could not be predicted in advance lay fallow for over three decades. Most of the effort to harness the power of computers focused on particular applications, and these required that outputs be generated reliably and accurately,[3] so little attention was given to the possibility of *intrinsic* unpredictability.

Turing's insight subsequently has been developed further by Wolfram [54–56], who formulated it in terms of "computational irreducibility." According to Wolfram:

[N]ormally it has been assumed that if one can only find the underlying rules for the components of a system then in a sense these tell one everything important about the system.

But what we have seen over and over again...is that this is not even close to correct, and that in fact there can be vastly more to the behavior of a system than one could ever foresee just by looking at its underlying rules. And fundamentally this is a consequence of the phenomenon of computational irreducibility ([56], p. 751).

Computational irreducibility means that it requires as much computational effort to know the outcome of the system being examined as it does just to let the system run to completion. In other words, no compressed or reductionist model can embody all the essentials of an irreducible system. Wolfram shows that computational irreducibility arises in even the simplest systems (e.g., cellular automata); it follows that it cannot be ruled out entirely in some larger systems.[4]

In addition to this mathematical property, ordinary computer systems exhibiting some degree of agency will be unpredictable for more mundane reasons. For example, self-driving vehicles and other open systems absorb information from the environment that arrives in ways that cannot be predicted, and hence their behavior will be undetermined to some degree (see the papers in [14] for extended discussion of "interactive computing"). Self-motivated AIs, if such ever come to be, could interact with humans in strategic ways with outcomes that may or may not be beneficial to the humans [8].

Whether or not future AIs ever achieve genuine *intentionality*, it will matter how they act towards and respond to humans. Such interactions, as they increasingly go beyond the simple relationship of user and tool, will have to be grounded on some kind of *knowledge base* possessed by the AI. How are future computer systems to acquire the knowledge required for moral behavior? Furthermore, what might give them the "motivation" to behave benignly?

Teaching AIs to be ethical

Turing addressed these questions indirectly in "Computing Machinery and Intelligence." A significant portion (approximately the final quarter) of his paper was devoted to a discussion of "Learning Machines." Turing speculated:

In the process of trying to imitate an adult human mind we are bound to think a good deal about the process which has brought it to the state that it is in. We may notice three components,

(a) The initial state of the mind, say at birth,
(b) The education to which it has been subjected,
(c) Other experience, not to be described as education, to which it has been subjected.

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain....The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child ([48], pp. 455–456).

Of course, "machine learning" is now at the cutting edge of AI research. It is not easy, however, to estimate how difficult it might be to follow Turing's suggestion across the whole range of adult capabilities. The purely economic investment in raising human children, socializing them, and teaching them moral behavior is enormous. Childrearing entails more than imparting basic literacy and numeracy---it involves teaching children to recognize and respect other persons, to adhere to interpersonal

---

[3]The best efforts of software developers have not been able to guarantee this, however. The literature on the rate of bugs per line of code is sobering [45]. In this paper Soergel estimates that "even the most careful software engineering practices in industry rarely achieve an error rate better than 1 per 1000 lines."

[4]Nor does it imply that computationally irreducible systems are entirely random, however. Such systems can exhibit statistical regularities [7], as well as predictability at coarse levels of granularity [19].

boundaries, to train their imaginations to empathize with others, etc.[5]

Setting aside the question of cost, the process of "bringing up baby-AI" would certainly entail some combination of building in or teaching rules of behavior, allowing for decision-making in situations where alternative actions are possible, and incorporating a knowledge base that could be drawn upon by the AI. Programming specific rules of behavior, in the spirit of a deontological ethical system, might work in many of the situations in which the AI might find itself, but could not be fully adequate to ensure moral behavior. Special cases can always be devised to confound any rigid set of rules. A consequentialist approach, by which the good and evil are judged by an action's outcome(s), is also problematic because the complexity of human behavior and social interactions makes it impossible to predict the ultimate outcomes of actions or policies.[6] Additionally, the end result of the educational process is itself unknowable. Human children grow up to behave in ways that cannot entirely be foreseen. Children may absorb some of the traits and values of their parents, but most assuredly have minds of their own. Child-AIs can be expected to show similar independence as they mature.

In any case, part of the education of an AI will almost certainly consist of the AI's *reading textual information* on the Internet. Such a vast store of material is available, however, that some human guidance is likely to be necessary (or at least efficient) in setting the AI's reading list. There is no simple way to measure the amount of information stored in all media because of difficulties of definition, coverage, and the rapid rate of technological progress, but a relatively recent attempt put the total amount of information stored as of 2007 at 295 optimally compressed exabytes[7] [17]. The authors noted that this amount of information stored on ordinary CD-ROMs would make a stack stretching from the Earth to the Moon and a quarter of that distance beyond (ibid., p. 62). These authors also estimated that the compound annual rate of growth of stored information had been 23% over the previous two decades; a 23% rate of growth extrapolated to 2019 would increase the 2007 figure by a factor of nearly 16. This rate of growth can probably be expected to increase over time.

Faced with so much information, the curriculum of a child-AI would have to be shaped and guided by human teachers, at least at first. Given a set of "recommended readings," a variety of methods for textual analysis that allow "M-assimilation" and "M-analysis" by computer systems without further direct human intervention.[8] The literature on "natural language processing" has been surveyed elsewhere [10, 15, 28, 29, 43]; the techniques include Latent Semantic Analysis (LSA) [22], Latent Dirichlet Allocation and other probabilistic topic models [3, 47], neural networks [5], and a variety of other statistical methods that fall under the general category of "stylometry."

A "simple" example – political theory

As an illustration, I will use one of the natural language processing approaches to see whether an AI system might begin to M-understand *political theory* by analyzing classical texts. Political theory is a good test case because it is so difficult; it illustrates how hard it is to think and act ethically in matters involving collective action. It hardly needs to be said that well-intentioned policies can fail, or may even produce terrible results, if the

---

[5]Childrearing is perhaps the largest single investment made by society. A rough calculation of the magnitude of this effort shows its approximate magnitude. The direct cost bringing up a child in the U.S. (housing, food, etc., exclusive of college education) has been estimated to be on the order of $233,610 per child in 2015 [50], not counting the indirect cost (the time and energy spent in parenting). Approximately 16% of this direct cost estimate consists of household expenditures for "child care and education" but the estimate does not include public expenditures for education [24]. Converting the resulting direct cost number to an annual average over the entire childhood period and multiplying by the proportion of children in the population yields direct expenses to be approximately 4.7% of U.S. GDP [4, 49, 53]. The OECD reports U.S. *public* expenditures on education (primary through tertiary) to be 4.1% of GDP [33] and *private* expenditures on education to be 2.0% [32]. (All OECD data reported here are as of 2016 or latest available and are rounded to the first decimal place.) To avoid double counting, the 16% from the USDA direct cost estimate (or approximately 0.9% of GDP) should be subtracted in computing the total resources devoted to childrearing. Alternatively, the OECD reports U.S. private spending on primary and secondary education to be 0.3% of GDP. The USDA-based estimate of this component is an overestimate because it includes "child care" costs. So even excluding the very large indirect parenting costs, the United States devotes about 10% of its national output to bringing children to full adulthood. But of course each human child needs to be brought up separately, even though there are some economies of scale in schooling, while AI "training" is likely to be much more easily transferable.

[6]These difficulties in "making machines moral" are discussed at length in [52] and the essays in [23].

[7]One exabyte equals approximately one billion gigabytes.

[8]The convention of preceding words like "analysis," "reading," or "comprehension" with an "M" denotes that the cognitive operations are being performed by machines, and sidesteps the question of what constitutes "thinking" that was dismissed by Turing as "too meaningless to deserve discussion" ([48], p. 442). The use of this "M-" terminology emphasizes that what AIs do (at least currently) is quite distinct from human understanding.

[9]Machiavelli understood this 500 years ago, although his insight is rarely acknowledged, because political decision-makers' interests generally lie in obscuring the disconnect between rhetoric and results. This has led to "Machiavellianism" becoming shorthand for amoral or immoral politics, whereas in fact Machiavelli's understanding that good intentions are not enough to produce good outcomes should be the starting point for the development of truly moral political practice.

designers of the policies do not understand how the socio-political system (including economics) works.[9] Sheer complexity frustrates purely technocratic efforts to solve social problems [12]. Ethical behavior in families or one-on-one interactions may be hampered by imperfect information and externalities, but these obstacles to benign outcomes are rampant in modern mass societies. In particular, the ordinary strictures against interpersonal violence must often be bypassed by the State as it carries out its necessary functions of national defense, provision of public goods through taxation, enforcement of contracts, and criminal justice.

The remainder of this paper is devoted to using Latent Dirichlet Allocation (LDA) to M-analyze six classical works of political theory: *The Prince* and *Discourses on the First Ten Books of Titus Livius* by Niccolò Machiavelli, Aristotle's *Politics*, John Locke's *Second Treatise of Government*, *Meditations* of Marcus Aurelius, and *The Federalist Papers* (written variously by John Jay, James Madison, and Alexander Hamilton).[10] LDA will be used to extract the main topic(s) from these works and to contrast the emphases of the different authors. For imagining M-learning by processing texts, it is informative to see whether or not an AI system can extract the most salient topics from these works without any direct human guidance or intervention.

LDA is based on a Bayesian model that works backwards to extract (unobserved) topics from (observed) texts under the assumption that the texts were (or could have been) generated by random sampling from particular (unobserved) probability distributions. The underlying model is a bag-of-words theory that asserts that most of the content of a text is derived from the words in association with each other, independent of structure and syntax. While the bag-of-words approach is by no means uncontroversial, "estimates of the relative amount that word order and word choice contribute to overall meaning of a sentence or paragraph suggest that the latter carries the lion's share, on the order of 80%-90%" ([20], p. 10, citing [21] "and later").

Detailed descriptions of LDA are readily available (e.g., [3, 47]), as well as step-by-step outlines of the Monte Carlo technique (Gibbs sampling) that enables LDA to be implemented in practice [6, 16, 38]. The model is based on two underlying probability distributions: the distribution of words across a "topic," and the distribution of topics across each document. Each topic i has its own distribution of words designated by ϕ^i, and each document has a distribution of topics designated by θ^j. It is imagined that document j is generated by first picking a topic i at random from the distribution θ^j and then a word at random from the distribution ϕ^i. Given a corpus of N documents, LDA estimates the (unobserved) distributions ϕ^i and θ^j by Bayesian inference. Loosely speaking, the posterior distributions of ϕ^i and θ^j conditional on the evidence can be recovered, given prior distributions for ϕ^i and θ^j and the evidence provided by the words in the actual documents in the corpus.

Off-the-shelf software can carry out this computationally-intensive process (e. g., [25, 37, 46]), and the present paper uses publically downloadable Mathematica code [27]. The "documents" of the analysis were individual chapters of the six works, suitably preprocessed to eliminate capitalization, punctuation, footnotes, and the word "Chapter" at the beginning of each document. Full Porter stemming [35] was not carried out, because Porter stemming might merge some words with the same root that have potentially different meanings. For example, "romans" refers to a people, while "rome" is a political entity.

Dirichlet priors are used in LDA, because the Dirichlet is a conjugate distribution to the Multinomial, and therefore guarantees that the posterior distributions conditioned on the data will also be Dirichlet. Use of conjugate distributions in a Bayesian framework simplifies the computation of the posterior distributions. The parameters of the Dirichlet priors were set as $\alpha = 0.01$ and $\beta = 0.01$; these low values of α and β make the distributions concentrate at the "corners" and hence will tend to sharply distinguish the topics. The number of topics was set at 40.

There were a total of 387 documents: 26 chapters of *The Prince*, 143 chapters of the *Discourses*, 103 chapters of *Politics*, 18 chapters of the *Second Treatise of Government*, 12 chapters of *Meditations*, and 85 chapters of *The Federalist Papers*. There is no automatic way to know whether the Gibbs sampling procedure implemented by the Mathematica code has converged to stable posterior distributions.[11] In the present analysis, 15,000 iterations of the sampler were employed, and comparisons of different runs suggested that this number of iterations was

---

[10]These books were downloaded in machine-readable form from Project Gutenberg.

[11]Raftery and Lewis suggest that in cases "reasonable accuracy may often be achieved with 5000 iterations or less" ([36], p. 1), but there is no hard-and-fast rule about this number.
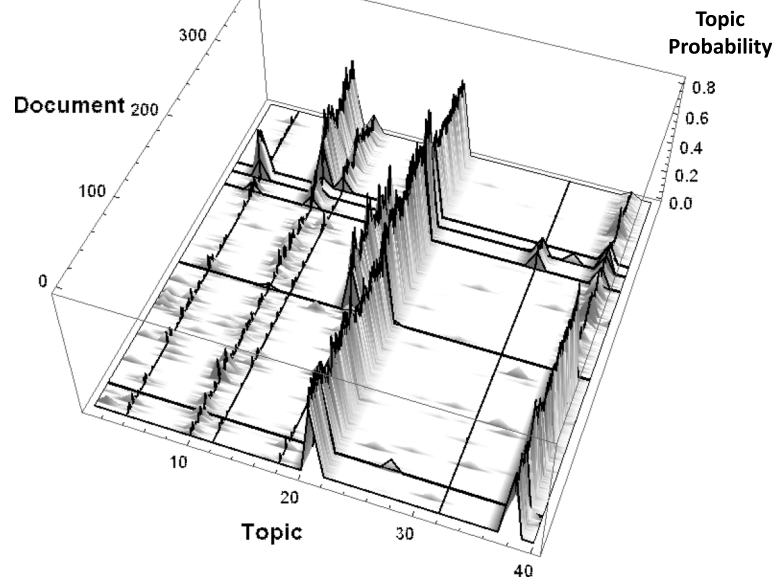
Fig. 1 topic probabilities for each of the documents

sufficient to extract topics reliably from the documents.

The results suggest that M-comprehension of the concepts of political theory is not inconceivable. Figure 1 shows a plot of topic probabilities for the documents from each of the six works.[12] In this figure, the lower axis signifies the topics (from 1 to 40), and the upper left of the box indicates the document number (from 1 to 387). The vertical axis shows the estimated posterior topic probability for each document.

Figure 1 illustrates in compact form the most important (i.e., highest probability) topics for each document and author. West-to-east grid lines separate the different authors and works according to their positions in the list of 387 chapters: 1–26 for *The Prince*, 27–169 for the *Discourses*, 170–272 for the *Politics*, 273–290 for the *Second Treatise of Government*, 291–302 for *Meditations*, and 303–387 for *The Federalist Papers*. South-to-north grid lines trace out the most important topics according to their probabilities. These lines follow the summits of the ridges of high probability.

All the documents show high probabilities for topic 21. The 15 most frequent words in topic 21

are: *people, great, time, power, shall, state, man, place, make, men, reason, government, citizens, public,* and *nature*. Clearly, all the works have to do with ***politics*** – the LDA algorithm is able to identify the clearest common theme. However, the works can be differentiated according to the other topics that are most prominent in them. Thus, second highest-probability topic in Machiavelli is number 38, and its 15 most frequent words are: *rome, prince, roman, city, romans, army, came, enemy, princes, arms, way, example, men, soldiers,* and *led*. Of course we know that Machiavelli drew heavily on examples from ancient Rome in his political writings, and that *The Prince* and the *Discourses* concentrate on war and conflict.

By way of contrast, Aristotle's *Politics* (documents 170–272) in addition to "politics" (topic 21) shows high probability for topic 18: *evident, different, persons, manner, city, live, democracy, proper, children, oligarchy, governments, person, business, determine, established*. Aristotle's political theory addressed the advantages and disadvantages of different forms of government and offered guidelines about how humans should behave in order to prosper in society.

In addition to the "politics" topic, the *Federalist Papers* show the strong presence of topic 10: *constitution, authority, political, union, governments, members, federal, danger, legislative, legislature, interests, society, rights, principles, principle*. This topic

---

[12]All three contributors to *The Federalist Papers* were treated as a single author.

Table 1 First and Second Most Probable Topics, Six Classical Works

Topic	The Prince		The Discourses		Aristotle		Locke		Marcus Aurelius		The Federalist	
	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2
1	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	–	–	–	–	–
3	–	–	–	–	–	–	–	–	–	–	–	–
4	–	–	–	–	–	–	–	–	–	11	–	–
5	–	–	–	–	–	–	–	–	–	–	–	–
6	–	–	–	–	–	–	–	–	–	–	–	–
7	–	–	–	–	–	–	–	–	–	–	–	–
8	–	–	–	–	–	–	–	–	–	–	–	–
9	–	–	–	–	–	–	–	–	–	–	–	–
10	–	–	–	–	–	–	–	6	–	–	6	77
11	–	–	–	–	–	–	–	–	–	–	–	–
12	–	–	–	–	–	–	–	–	–	–	–	2
13	–	–	–	–	–	–	–	–	–	–	–	–
14	–	–	–	–	–	–	–	–	–	–	–	–
15	–	–	–	–	–	–	–	–	–	–	–	–
16	–	–	–	–	–	–	–	–	–	–	–	–
17	–	–	–	–	–	–	–	–	–	–	–	–
18	–	–	–	–	6	96	–	–	–	–	–	–
19	–	–	–	–	–	–	–	–	–	–	–	–
20	–	–	–	–	–	–	–	–	–	–	–	–
21	21	5	67	76	97	6	18	–	12	–	79	6
22	–	–	–	–	–	–	–	–	–	–	–	–
23	–	–	–	–	–	–	–	–	–	–	–	–
24	–	–	–	–	–	–	–	–	–	–	–	–
25	–	–	–	–	–	–	–	–	–	–	–	–
26	–	–	–	–	–	–	–	–	–	–	–	–
27	–	–	–	–	–	–	–	–	–	–	–	–
28	–	–	–	–	–	–	–	–	–	–	–	–
29	–	–	–	–	–	–	–	–	–	–	–	–
30	–	–	–	–	–	–	–	–	–	–	–	–
31	–	–	–	–	–	–	–	–	–	–	–	–
32	–	–	–	–	–	–	–	8	–	–	–	–
33	–	–	–	–	–	–	–	–	–	–	–	–
34	–	–	–	–	–	–	–	–	–	–	–	–
35	–	–	–	–	–	–	–	–	–	–	–	–
36	–	–	–	–	–	–	–	–	–	–	–	–
37	–	–	–	–	–	–	–	–	–	–	–	–
38	5	21	76	67	–	1	–	4	–	1	–	–
39	–	–	–	–	–	–	–	–	–	–	–	–
40	–	–	–	–	–	–	–	–	–	–	–	–

embodies the *Federalist*'s arguments in favor of adopting the U.S. Constitution.

The number of documents from Locke's *Second Treatise of Government* and Marcus Aurelius' *Meditations* are smaller than the number of documents from the other authors, so the signal that can be extracted from those two authors is perhaps less strong than for Machiavelli, Aristotle, and the *Federalist*. Nevertheless, Locke and Marcus Aurelius can be identified by the small "mountains" that appear where their documents are located (positions 273 to 290 for Locke, positions 201–302 for Marcus Aurelius). The *Meditations*, in addition to topic 21, shows some weight on topic 4: *universe, universal, needs, simplicity, rational, gods, soul, substance, plants, flesh, social, cast, beings, fate, fruit*. The *Meditations* is as much a work of philosophy as of politics, and this is reflected by the words in topic 4. Locke's *Second Treatise* shows several other relatively low-probability topics, but it shares an emphasis on topic 10 with the authors of *The Federalist Papers*. This is not surprising given the influence of Locke's political thought on the founders of the American political system.

Another way of showing which topics are prominent in the different works is given in Table 1. This table indicates the number of times each topic is either the highest or the second-highest probability topic in the different authors' works. The same patterns that are displayed graphically in Fig. 1 appear also in Table 1. Topic 38, the one emphasizing Rome and the Romans, appears more frequently as the highest-probability topic in Machiavelli's *Discourses* than in *The Prince*, as would be expected given that examples from the history of ancient Rome are the starting point for many of the arguments of the *Discourses*. The sparsity of Table 1 is attributable to the choice of Dirichlet prior parameters that makes for sharp distinctions among the topics.

Conclusions

Turing's prescient thought experiment about educating a child-AI highlights just how far we are from being able to know whether actual AIs, if and when they emerge, will be benign. We have no idea "what it is like to be an AI."[13] Human beings are embodied creatures, and this is relevant both to learning [31, 51][14] and to moral development.[15] If we imagine the "motivations" of an emergent AI to be formed through a learning process, it is plausible that part of that learning will be the assimilation of human-generated works about ethics, morality, and politics. While the extraction of topics from classical texts in political theory is remote from an AI's actually being able to formulate plans and principles for action (political or otherwise), some kind of M-understanding of such works would surely be pertinent. At the very least, the long human history of the struggle to understand politics and formulate an appropriate moral foundation for collective action reflects experiences and insights that would be relevant to an AI in dealing with the same issues.

It should be emphasized that the results presented here do not (and cannot) demonstrate human-level understanding of texts by an unsupervised AI. All "bag-of-words" techniques ignore the syntactical structure of the texts. The LDA "topics" are not summaries of the contents of the works. Texts with differing or even opposing political implications might display highly similar LDA topic signatures. All of the authors whose writings are analyzed here exhibit a strong orientation to "politics," but their philosophies are quite different. Identification of topics in texts is not sufficient to define or develop political ethics.

LDA is a well-defined but circumscribed technique that enables classification and systematization of certain kinds of information contained in a set of texts. Given its limits, LDA can identify major themes---the "topics" that are its output---without human intervention. The results developed in this paper show that even corpora made up of only a few authors and works can successfully be processed by LDA. These results also show that natural language processing techniques are applicable to texts that are distant in time from the modern day. Just as an unsupervised AI can distinguish between fiction and non-fiction books [9], an AI employing LDA can extract the main topics from a set of classical texts dealing with a particular field (in this case, political theory). Deployment of more computing power would enable

---

[13]Paraphrasing the title of Nagel's famous paper [30]. The question "[w]hat is it like to be a ____" is connected to the unsolved "hard problem" of consciousness. A compilation of perspectives on this problem is provided by the papers in [42]. As the editor puts it at the close of his Introduction, "This sort of creative diversity is of course what should be expected as we wrestle with what has come to be recognized as a serious challenge for standard materialism, namely the existence of consciousness itself" (p. 6).

[14]O'Loughlin notes that "[e]ducational theorists have often appeared to be rather uncomfortable with the brute fact of corporeality. Their discussions of cognition, social phenomena, and the development of intellectual skills or moral reasoning have been frequently carried out as if bodies were something of an embarrassment" ([31], p. 16).
[15]Of course, human morality has been a central concern of religious thinkers and philosophers for thousands of years.
[16]All the calculations reported here were performed on a desktop PC running Mathematica 12 [26].

results like these to encompass a wider range of texts and a greater number of topics in finer detail.[16]

References
1. Amaro S (2018) Sell-offs could be down to machines that control 80% of the US stock market, fund manager says. CNBC. https://www.cnbc.com/2018/12/05/sell-offs-could-be-down-to-machines-that-control-80percent-of-us-stocks-fund-manager-says.html. Accessed 8 June 2018.
2. Atherton KD. Are killer robots the future of war? Parsing the facts on autonomous weapons. New York: The New York times magazine; 2018.
3. Blei DM. Probabilistic topic models. Commun ACM. 2012;55(4):77–84.
4. Child Trends. (2019) Number of Children. Bethesda: Retrieved from https://www.childtrends.org/indicators/number-of-children. Accessed 1 Jul 2020.
5. Collobert R, Weston J. A unified architecture for natural language processing: deep neural networks with multitask learning. In: Proceedings of the 25th international conference on machine learning. Helsinki: ACM Digital Library; 2008.
6. Darling WM. A theoretical and practical implementation tutorial on topic modeling and Gibbs sampling: School of Computer Science, University of Guelph; 2011. http://u.cs.biu.ac.il/~89-680/darling-lda.pdf. Accessed 9 June 2019.
7. DeCanio SJ. Limits of economic and social knowledge. Houndmills, Basingstoke, Hampshire: Palgrave; 2014.
8. DeCanio SJ. Games between humans and AIs. AI Soc. 2018a;33:557–64 https://link.springer.com/article/10.1007/s00146-017-0732-5. Accessed 10 Sept 2019.
9. DeCanio SJ. AI recognition of differences among book-length texts. AI Soc. 2018b; https://link.springer.com/article/10.1007%2Fs00146-018-0851-7. Accessed 10 Sept 2019.
10. Foltz PW. Quantitative approaches to semantic knowledge representations. Discourse Process. 1998;25(2–3):127–30.
11. Foster A (2018) How much trading in the stock market is algorithmic trading and how much is non-algorithmic? Quora. https://www.quora.com/How-much-trading-in-the-stock-market-is-algorithmic-trading-and-how-much-is-non-algorithmic. Accessed 8 June 2018.
12. Friedman J. Power without knowledge: a critique of technocracy. New York: Oxford University Press; 2020.
13. Fryer-Biggs Z. Coming soon to the battlefield: robots that can kill: The Center for Public Integrity; 2019. https://publicintegrity.org/national-security/future-of-warfare/scary-fast/ai-warfare/. Accessed 8 Sept 2019.
14. Goldin D, Smolka SA, Wegner P, editors. Interactive computation: the new paradigm. Berlin: Springer-Verlag; 2006.
15. Gomaa WH, Fahmy AA. A survey of text similarity approaches. Int J Comput Appl (0975–8887). 2013;68(13):13–8.
16. Heinrich G (2009) Parameter estimation for text analysis. Technical report, Fraunhofer IGD, Darmstadt, Germany. http://www.arbylon.net/publications/text-est2.pdf. Accessed 14 June 2019.

17. Hilbert M, López P. The World's technological capacity to store, communicate, and compute information. Science. 2011;332:60–5.
18. Imburgia V. How much trading in the stock market is algorithmic trading and how much is non-algorithmic? Quora; 2018. https://www.quora.com/How-much-trading-in-the-stock-market-is-algorithmic-trading-and-how-much-is-non-algorithmic. Accessed 8 June 2018.
19. Israeli N, Goldenfeld N. On computational complexity and the predictability of complex physical systems: Department of Physics, University of Illinois at Urbana-Champaign; 2003. https://arxiv.org/pdf/nlin/0309047.pdf. Accessed 10 Sept 2019.
20. Landauer TK. LSA as a Theory of Meaning. In: Landauer TK, DS MN, Dennis S, Kintsch W, editors. Handbook of latent semantic analysis. New York: Routledge; 2007.
21. Landauer TK. On the computational basis of cognition: Arguments from LSA. In: Ross BH, editor. The psychology of learning and motivation. New York: Academic Press; 2002.
22. Landauer TK, McNamara DS, Dennis S, Kintsch W, editors. Handbook of latent semantic analysis. New York: Routledge; 2011.
23. Lin P, Abney K, Bekey GA, editors. Robot ethics: the ethical and social implications of robotics. Cambridge: The MIT Press; 2012.
24. Lino M, Kuczynski K, Rodriguez N, Schap T. Expenditures on Children by Families, 2015. In: Miscellaneous publication no. 1528-2015. Alexandria: U.S. Department of Agriculture, Center for Nutrition Policy and Promotion; 2017.
25. McCallum AK (2002) MALLET: a machine learning for language toolkit. http://mallet.cs.umass.edu. Accessed 11 June 2019.
26. Mathematica (Version 12) (2019) Wolfram Mathematica: The world's definitive system for modern technical computing. http://www.wolfram.com/mathematica/.
27. Mathematica Stack Exchange (2019) How to perform document classification (i.e., extracting topics from text)? https://mathematica.stackexchange.com/questions/66987/how-to-perform-document-classification-i-e-extracting-topics-from-text. Accessed 10 Sept 2019.
28. Mikolov TL. Chen K, Corrado G, Dean J (2013a) Efficient estimation of word representations in vector space. https://arxiv.org/abs/1301.3781. Accessed 7 Sept 2019.
29. Mikolov T. Sutskever I, Chen K, Corrado G, Dean J (2013b) Distributed representations of words and phrases and their compositionality. https://arxiv.org/pdf/1310.4546.pdf. Accessed 7 Sept 2019.
30. Nagel T. What Is It Like to Be a Bat? Philosophical Rev. 1974;83(4):435–50 Reprinted In: Nagel T (1979) Mortal Questions. Cambridge University Press, Cambridge.
31. O'Loughlin M. Embodiment in education: exploring creatural existence. Dordrecht: Springer; 2006.
32. OECD (Organization for Economic Co-operation and Development (2020a) Private spending on education (indicator). doi: https://doi.org/10.1787/6e70bede-en. https://data.oecd.org/eduresource/private-spending-on-education.htm Accessed 1 Jul 2020.
33. OECD (Organization for Economic Co-operation and Development) (2020b) Public spending on education (indicator). doi: https://doi.org/10.1787/f99b45d0-en. https://data.oecd.org/eduresource/public-spending-on-education.htm. Accessed 1 Jul 2020.
34. Pandolfini B. Kasparov and deep blue: the historic chess match between man and machine. New York: Fireside; 1997.
35. Porter M. An algorithm for suffix stripping. Program. 1980;14(3):130–7. https://doi.org/10.1108/eb046814 Accessed 10 Sept 2019.
36. Raftery AE, Lewis S (1991) How many iterations in the Gibbs sampler? University of Washington. http://people.ee.duke.edu/~lcarin/raftery92how.pdf. Accessed 14 June 2019.
37. Ramage D, Rosen E. Stanford topic modeling toolbox: The Stanford natural language processing group; 2009. https://nlp.stanford.edu/software/tmt/tmt-0.4/. Accessed 11 June 2019.
38. Resnik P, Hardesty E. Gibbs sampling for the uninitiated: Semantic scholar; 2010. https://pdfs.semanticscholar.org/6fb3/1dd73faa5aec66cfe414e235de63247e8b68.pdf?_ga=2.85528542.61205280.1560120102-1924383825.1559933214 Accessed 9 June 2019.
39. Schaeffer J. One jump ahead: challenging human supremacy in checkers. New York: Springer-Verlag; 1997.
40. Schaeffer J, Burch N, Björnsson Y, Kishimoto A, Müller M, Lake R, Lu P, Sutphen S. Checkers is solved. Science. 2007;317:1518–22.
41. Shankland S. An IBM computer debates humans, and wins, in a new, nuanced competition: c|net; 2018. https://www.cnet.com/news/an-ibm-

computer-debates-humans-and-wins-in-a-new-nuanced-competition/. Accessed 7 Sept 2019.

42. Shear J, editor. Explaining Consciousness – The 'Hard Problem.'. Cambridge: The MIT Press; 1998.

43. Shiffrin R, Börner K. Mapping knowledge domains. Proc Natl Acad Sci. 2004; 101(suppl. 1):5183–5.

44. Silver D, Hubert T, Schrittweiser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T, Simonyan K, Hassabis D. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. Science. 2018;362:1140–4.

45. Soergel DAW. Rampant software errors may undermine scientific results [version 2; referees: 2 approved]. F1000Research. 2015;3:303 https://f1 000research.com/articles/3-303. Accessed 10 Sept 2019.

46. Steyvers M (2011) Matlab topic modeling toolbox 1.4. http://psiexp.ss.uci. edu/research/programs_data/toolbox.htm#Authors. Accessed 11 June 2019.

47. Steyvers M, Griffiths T. Probabilistic topic models. In: Landauer TK, DS MN, Dennis S, Kintsch W, editors. (2011) Handbook of Latent Semantic Analysis. New York: Routledge; 2011. p. 427–448.48.

48. Turing AM. Computing machinery and intelligence. Mind. 1950;59(236):433–60.

49. U.S. Bureau of Economic Analysis, Gross Domestic Product [GDP] (2020). Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred. stlouisfed.org/series/GDP. Accessed 1 Jul 2020.

50. United States Department of Agriculture (USDA) (2017) The Cost of Raising a Child. Posted by Mark Lino. Https://www.usda.gov/media/blog/2017/ 01/13/cost-raising-child. Accessed 5 Sept 2017.

51. Vlieghe J. The body in education. In: Smyers P, editor. International handbook of philosophy of education. Switzerland: Springer International Publishing AG, Springer International Handbooks of Education; 2018.

52. Wallach W, Allen C. Moral machines: teaching robots right from wrong. Oxford: Oxford University Press; 2009.

53. World Bank, Population, Total for United States [POPTOTUSA647NWDB] (2020). Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred. stlouisfed.org/series/POPTOTUSA647NWDB. Accessed 1 Jul 2020.

54. Wolfram S. Universality and complexity in cellular automata. Physica D. 1984a;10:1–35.

55. Wolfram S. Computation theory of cellular automata. Commun Math Phys. 1984b;96:15–57.

56. Wolfram S. A new kind of science. Champaign: Wolfram Media, Inc.; 2002.